A VARIATIONAL PRINCIPLE FOR A NON-INTEGRABLE MODEL

GEORG MENZ AND MARTIN TASSY

Abstract. We develop a new robust technique to deduce variational principles for non-integrable discrete systems. To illustrate this technique, we show the existence of a variational principle for graph homomorphisms from $\mathbb{Z}^m$ to a $d$-regular tree. This seems to be the first non-trivial example of a variational principle in a non-integrable model. Instead of relying on integrability, the technique is based on a discrete Kirszbraun theorem and a concentration inequality obtained through the dynamic of the model. Using those two results, we also obtain the existence of a continuum of translation-invariant, ergodic, gradient Gibbs measures for graph homomorphisms from $\mathbb{Z}^m$ to a regular tree.

Contents

1. Introduction 1
2. The variational principle for graph homomorphisms 9
3. The local surface tension 18
4. A Kirszbraun theorem and a concentration inequality 31
5. Existence of a continuum of shift-invariant ergodic gradient Gibbs measures 56
6. Proof of the variational principle 59
References 70

1. Introduction

The appearance of limit shapes as a limiting behavior of discrete systems is a well-known and studied phenomenon in statistical physics.

Date: February 6 2017.
2010 Mathematics Subject Classification. Primary: 82B20, 82B30, 82B41, Secondary: 60J10.

Key words and phrases. Variational principles, non-integrable models, limit shapes, domino tilings, Glauber dynamic, Azuma-Hoeffding, gradient-Gibbs measures, entropy, concentration inequalities, ergodic measures.
and combinatorics (e.g. [Geo88]). Among others, models that exhibit limits shapes are domino tilings and dimer models (e.g. [Kas63, CEP96, CKP01]), polymer models, lozenge tilings (e.g. [Des98, LRS01, Wil04]), Gibbs models (e.g. [She05]), the Ising model (e.g. [DKS92, Cer06]), asymmetric exclusion processes (e.g. [FS06]), sandpile models (e.g. [LP08]), the six vertex model (e.g. [BCG16, CS16, NR16]), Young tableaux (e.g. [LS77, VK77, PR07]).

Limit shapes appear whenever fixed boundary conditions force a certain response of the system. The main tool to explain those shapes is a variational principle. The variational principle asymptotically characterizes the number of microscopic states, i.e. the microscopic entropy $\text{Ent}_n$, via a variational problem. This means that for large system sizes $n$, the entropy of the system is given by maximizing a macroscopic entropy $\text{Ent}(f)$ over all admissible limiting profile $f \in \mathcal{A}$. The boundary conditions are incorporated in the admissibility condition. In formulas, the variational principle can be expressed as (see for example Theorem 2.9 below)

$$\text{Ent}_n \approx \inf_{f \in \mathcal{A}} \text{Ent}(f),$$

where the macroscopic entropy

$$E(f) = \int \text{ent}(\nabla f(x)) dx$$

can be calculated via a local quantity $\text{ent}(\nabla f(x))$. This local quantity is called local surface tension in this article.

Often, a consequence of a variational principle is that the uniform measure on the microscopic configurations, concentrates around configurations that are close to the minimizer of the variational problem (see comments before Theorem 2.11 below). This is related to the appearance of limit shapes on large scales.

In analogy to classical probability theory, one can understand the variational principle as an elaborated version of the law of large numbers. On large scales, the behavior of the system is determined by a deterministic quantity, namely the minimizer $f$ of the macroscopic entropy. Hence, deriving a variational principle is often the first step in analyzing discrete models, before one attempts to study other questions like the fluctuations of the model.
A lot of inspiration for this article comes from the the variational principle of domino tilings [CKP01] (see Figure 1a). It is one of the fundamental results for studying domino tilings and the other integrable discrete models. A detailed analysis of the limit shapes for domino tilings was given in [KOS06]. Recently a new constructive approach was developed in [CS16] for the determination of the Arctic curve (the frozen boundary of the limit shape). The approach is discussed mainly in the framework of the six vertex model, which is integrable. However, the method seems to be very robust. If it also can be used to determine the Arctic curve in a non-integrable model is an interesting open question.

So far, all the tools that were developed to study variational principles of discrete models rely on the integrability of the model. Up to the knowledge of the authors, there is no non-trivial example of a variational principle for which the underlying model is not integrable. However, simulations (see Figure 1b, Figure 1c and Figure 1d) show that limit shapes also appear for a large class of non-integrable models. Limit shapes appear to be a universal phenomenon. The purpose of this article is to go beyond integrability and to find out what properties of a discrete system lead to variational principles and limit shapes.

We are interested in graph-homomorphisms because they provide a natural framework to study systems with hard constraints (see [BW00]). In this article, we consider the non-integrable model of graph homomorphisms form $\mathbb{Z}^m$ to a $d$-regular tree. We want to point out the fact that in our variational principle the underlying lattice can have arbitrary dimension $m \geq 2$.

We identified two properties that a model of discrete maps needs to have in order to have a variational principle. The first one is a stability property. Perturbing the boundary condition on a microscopic scale does not change the macroscopic properties of the model. This allows the classification of configurations depending on the speed at which the configuration travel in the arrival space. Therefore, it implies that the macroscopic entropy $\text{Ent}(f)$ of the system is determined by a local surface tension $\text{ent}(\nabla f(x))$. The second one is a concentration property. Else, one cannot hope that the model satisfies a variational principle which is a type of law of large numbers.

In the case of discrete integrable models, such as domino tilings or the antiferromagnetic Potts model, both properties can be deduced
(a) An Aztec diamond for domino tilings. The combinatorics of the model is similar to Lipschitz functions from $\mathbb{Z}^2$ to $\mathbb{Z}$. (see [CKP01])

(b) An Aztec diamond for ribbon tilings. The combinatorics of the model is similar to Lipschitz functions from $\mathbb{Z}^2$ to $\mathbb{Z}^2$ (see [She02]).
(c) An Aztec diamond tiling by $3 \times 1$ bars. The combinatorics of the model is similar to Lipschitz functions from $\mathbb{Z}^2$ to $\mathbb{Z}_3 \ast \mathbb{Z}_3$ (see [KK92]).

(d) An Aztec diamond for Graph homomorphisms in a 3-regular tree. Each color represents one of the $\alpha_i$'s introduced in Section 3.
naturally. For deducing the first property, one uses that the space of configurations is a lattice. Then it is possible to quickly attach two configurations together provided that the boundary conditions are similar. This is done by using the minimum of two well-chosen extensions of those configurations. The second property, namely the concentration, is tackled by using a loop reversal argument or an analog version of this argument for other systems (see e.g. [CEP96, She05]).

We want to emphasize again that those arguments are based on the integrability of the underlying model and are not available for general graph homomorphisms. One of the main contributions of this article is that we provide alternative methods. These methods are not based on integrability but on weaker properties of graphs and of the underlying dynamic of the model. The authors believe that the principles behind those new arguments are robust. They should provide a possible line of attack to study variational principles and limiting behavior of a large class of non-integrable models.

Now, let us discuss how the two necessary properties, namely stability and concentration, are deduced without relying on integrability. The first property is obtained by using the discrete version of a well-known theorem for continuous metric spaces: the Kirszbraun theorem. Up to knowledge of the authors, the first version of a discrete Kirszbraun theorem was developed in the setting of tilings in [Tas14] and [PST16]. Using a new version of the Kirszbraun theorem for graph homomorphisms allows us to show that microscopic variations of the boundary conditions can be neglected on the macroscopic scale and thus do not influence the entropy of a system.

In order to deduce the second property, the authors combine a classical concentration inequality, namely the Azuma-Hoeffding inequality, with a coupling technique relying on dynamic properties of the model. Inspiration for this type of argument comes from [CEP96], where the Azuma-Hoeffding inequality was used to show concentration for domino tilings. However, it is very difficult to apply directly the Azuma-Hoeffding inequality for more complicated models. The reason is that this needs detailed information about the structure of the underlying space of configurations. Our dynamic approach circumvents this obstacle. In our approach, one only has to understand the response of the system to changing the value of one point.
There is a natural candidate for our dynamic approach. It is the Glauber dynamic (see [Cha16] for details). This dynamic would be sufficient for the simpler model of graph homomorphisms to $\mathbb{Z}$. However, using the Glauber dynamic does not work for the more complicated model of graph homomorphisms to a tree $\mathcal{T}$. Heuristically, this can be understood from the observation that the simple random walk on a tree is not commutative and tends to diverge. We overcome this technical obstacle by modifying the dynamic. More precisely, we add to the original Glauber dynamic an extra non-local resampling step. Even after this modification, the Glauber dynamic does not conserve the distance between two states. We circumvent this technical obstacle by introducing a suitable quantity called depth. It turns out the modified Glauber dynamic conserves this quantity. This allows us to apply the Azuma-Hoeffding inequality and deduce concentration in this quantity. We then show that concentration in depth is sufficient for deducing our variational principle.

The local surface tension $\text{ent}(s_1, \ldots, s_m)$ is defined as the limit of suitable chosen microscopic entropy i.e.

$$\text{ent}(s_1, \ldots, s_m) = \lim_{n \to \infty} \text{ent}_n(s_1, \ldots, n).$$

The existence of this limit is deduced by a combination of the Kirszbraun theorem and the concentration inequality. We also show that the local surface tension $\text{ent}(s_1, \ldots, s_m)$ is convex. We do not know if the local surface tension is strictly convex. From convexity it follows that variational problem given by our variational principle has a minimizer (cf. Theorem 2.9 below). However, we do not prove that this minimizing limiting profile is unique. The uniqueness of the minimizer would follow if the local surface tension is strictly convex, which we conjecture. Additionally, we conjecture that the convexity at a given slope increases when the degree of the tree increases.

In this article, we also show another consequence of the Kirszbraun theorem and the concentration inequality. It is the existence of a continuum of shift-invariant ergodic gradient Gibbs measures on tree-valued graph homomorphisms on $\mathbb{Z}^m$ as well as on any graph whose universal cover is a regular tree.

Compared to $\mathbb{Z}$, there are infinitely many ways to travel to infinity in a tree $\mathcal{T}$. Those pathways to infinity are described by geodesics. Limit shapes are sensitive to the choice of geodesics on which the graph homomorphism travels on the boundary (see Figure 1 for an illustration).
This adds another technical difficulty when deducing the variational principle for graph homomorphisms to a tree. It is the problem of defining the scaling limit of a graph homomorphism. For domino tilings the height function is an integer valued height function which allows a natural notion of a scaling limit. However, when the space of geodesics is more complex, as it is the case for trees, the notion of a scaling limit is less obvious. In order to define the scaling limit of a graph homomorphism to a tree, one has to additionally keep track of the information on which geodesic the graph homomorphism is traveling on. This leads to a more subtle definition of the limiting profile which involves several compatibility conditions (see Definition 2.2). Another consequence is that the variational principle for graph homomorphisms to a tree becomes more subtle. More precisely, the set $A$ of admissible limiting profiles $h$, over which the continuous entropy $\text{Ent}(h)$ is minimized, has an elaborated structure involving additional constraints.

**Overview over the article.** In Section 2, we describe the main result of this article, namely the variational principle for graph homomorphisms to a tree. In Section 3.1, we show the existence and convexity of the local surface tension $\text{ent}(s_1, \ldots, s_n)$. In Section 4, we provide the main technical tools needed in this article, namely the Kirszbraun theorem and the concentration inequality. Section 5 is independent of the variational principle. There, we use the Kirszbraun theorem and the concentration inequality to derive the existence of a continuum of
shift-invariant ergodic gradient Gibbs measures. And finally, in Section 6 we give the proof of the variational principle.

Notation

• $C$ and $c$ denote generic positive bounded universal constants.
• $|A|$ denotes the cardinality of the set $A$.
• $x, y, z$ denote elements $x, y, z \in \mathbb{Z}^m$.
• $S_n := \{0, \ldots, n-1\}^m$.
• $\vec{i}$ denotes the $i$-th Euclidean basis vector.
• $x \sim y$ indicates that the points $x$ and $y$ are neighbors.
• $e_{xy}$ is the oriented edge from $x \in \mathbb{Z}^m$ to $y \in \mathbb{Z}^m$.
• $d_G$ distance in the graph $G$.
• $T$ denotes a $d$-regular tree.
• $w, v$ denote elements $w, v \in T$.
• $r \in T$ denotes the root of the tree $T$.
• $g \subset T$ denotes a geodesic of the tree $T$.
• $\Pi_g : T \to g$ is the projection on the geodesic $g$.
• $\partial T := \{\infty_g : g$ is a geodesic of $T\}$.
• $\theta(\varepsilon)$ denotes a generic smooth function with $\lim_{\varepsilon \to 0} \theta(\varepsilon) = 0$.
• $\alpha_i$ denotes colors of edges.
• $(\alpha_1, \ldots, \alpha_d)$ denotes the generating set of a $d$-regular graph.
• $h : \mathbb{Z}^m \to T$ is a graph homomorphism.
• $\mathcal{H}_n^g(s)$ is the set of all graph-homomorphisms $h : \mathbb{Z}^m \to T$ that are $n$-invariant with slope $s$ and supported on the geodesic $g$.
• $\text{ent}_n(s) = -\frac{1}{m^2} \ln |\mathcal{H}_n^g(s)|$.
• $\text{ent}(s) = \lim_{n \to \infty} \text{ent}_n(s)$.

2. The variational principle for graph homomorphisms

Let us start with clarifying the underlying model. For $n \in \mathbb{N}$, we consider a finite subset $R_n \subset \mathbb{Z}^m$ of the $m$-dimensional lattice $\mathbb{Z}^m$. We assume that for $n \to \infty$ the scaled sublattice $\frac{1}{n} R_n$ converges in the Gromov-Hausdorff sense to a compact and simply connected region $R \subset \mathbb{R}^m$ with Lipschitz boundary $\partial R$. The basic objective is to study graph homomorphisms $h : R_n \to T$, where $T$ denotes a $d$-regular tree.

Definition 2.1. (Graph homomorphism, height function) Let $\mathcal{T}$ denote the $d$-regular rooted tree and let $\Lambda \subset \mathbb{Z}^m$ be a finite set. We denote with $d_G$ the natural graph distance on a graph $G$. A function $h : \Lambda \to \mathcal{T}$
is called graph-homomorphism, if
\[ d_T(h(k), h(l)) = 1 \]
for all \( k, l \in \Lambda \) with \( d_{\mathbb{Z}^m}(k, l) = 1 \). In analogy to [CKP01], we may also call \( h \) a \( T \)-valued height function. Let \( \partial \Lambda \) denote the inner boundary of \( \Lambda \subset \mathbb{Z}^m \) i.e.
\[ \partial \Lambda = \{ x \in \Lambda \mid \exists y \notin \Lambda : \text{dist}_{\mathbb{Z}^m}(x, y) = 1 \} . \]

We call a homomorphism \( h : \partial \Lambda \to T \) boundary graph homomorphism or boundary height function.

We want to study the question of how many \( T \)-valued height functions exist that extend a fixed prescribed boundary height function \( h_{\partial R_n} : \partial R_n \to T \). Hence, let us consider the set \( M(R_n, h_{\partial R_n}) \) that is defined as
\[ M(R_n, h_{\partial R_n}) = \{ h : R_n \to T \mid h \text{ is a height function and } h(\sigma) = h_{\partial R_n}(\sigma) \quad \forall \sigma \in \partial R_n \} . \]

The goal of the article is to derive an asymptotic formula as \( n \to \infty \) of the microscopic entropy
\[ \text{Ent} (R_n, h_{\partial R_n}) := -\frac{1}{|R_n|^2} \log M(R_n, h_{\partial R_n}) . \]

For this purpose, let us introduce the notion of an asymptotic height profile and the notion of an asymptotic boundary height profile. Those two objects will serve as the possible limits of sequences of graph homomorphisms \( h_{R_n} : R_n \to T \) and boundary graph homomorphisms \( h_{\partial R_n} : \partial R_n \to T \).

**Definition 2.2 (Asymptotic height profile).** Let \( k \in \mathbb{N} \), let \( h_R : R \to \mathbb{R}^+ \times \{1, \ldots, k\} \) be a function and let \( (a_{ij})_{k \times k} \) be a set of non-negative real numbers satisfying the following compatibility conditions
\[ a_{i,j} = a_{j,i} \quad \text{and} \quad a_{i,i} = 0 \]
and (cf. Figure 2)
\[ a_{i,j} < a_{i,k} \Rightarrow a_{j,k} = a_{i,k} \]
for all \( i, j \in \{1, \ldots, k\} \). We say that \( (h_R, (a_{ij})_{k \times k}) \) is an asymptotic height profile if:
- The first coordinate of the map \( h_R \) is 1-Lipschitz with respect to the \( l_1 \)-norm, i.e. for all \( x, y \in R \)
  \[ |h^1_R(x) - h^1_R(y)| \leq |x - y|_1 . \]
The map $h_R$ is $(a_{ij})_{k \times k}$-admissible in the sense that for all $i \neq j$:

$$h_R^{-1}(\mathbb{R}^+, i) \cap h_R^{-1}(\mathbb{R}^+, j) \subset h_R^{-1}([0, a_{ij}], \{1, \ldots, k\}).$$

The Definition 2.2 has the following interpretation. Firstly, we note that compared to a classical asymptotic height function $h_R : R \to \mathbb{R}$ (see for example [CKP01]) our notion of an asymptotic height profile $(h_R, (a_{ij})_{k \times k})$ has two coordinates $h_1^R$ and $h_2^R$. The reason for having two coordinates is that, in contrast to $\mathbb{R}$, there are infinitely many ways to travel from zero to infinity on a tree $\mathcal{T}$. Those pathways to infinity are described by directed geodesics $g$ starting in the root $r$. We assume that the asymptotic boundary height profile will travel only on finitely many geodesics $g_i$ that are indexed by $1, \ldots, k$.

The second coordinate $h_2^R(x) = i$ indicates on which geodesic $g_i$, the point $x \in R$ is mapped to (see Definition 2.6 from below). More precisely, the point $x \in R$ will be mapped onto a point on the geodesic $g_{h_2^R(x)}$. The first coordinate $h_1^R(x)$ of the asymptotic boundary profile $h_R$ specifies the exact location on the geodesic $g_{h_2^R(x)}$. This means that $x \in R$ will be mapped onto a point on the geodesic $g_{h_2^R(x)}$ that has distance $h_1^R(x)$ from the root $r$. Working with directed geodesics allows to assume that $h_1^R(x) \in \mathbb{R}^+$ is non negative. The Lipschitz condition (5) on $h_1^R$ is very natural and follows from the fact that graph homomorphisms are $1-$Lipschitz. This is very similar to the setting of
Let us now describe the meaning of the numbers $a_{ij}$, the compatibility condition (4) and the condition (6). The numbers $a_{ij}$ have their origin in the following observation. Any two geodesics $g_1 \subset T$ and $g_2 \subset T$ starting in $r \in T$ have a nonzero intersection $g_1 \cap g_2 \neq \emptyset$. However, they must split up at some vertex $v_{12} \in T$ (see also discussion below and Figure 4). If seen from the root $r$, the vertex $v_{12} \in T$ can be interpreted as the splitting point of the geodesics $g_1$ and $g_2$. If seen from infinity, the vertex $v_{12} \in T$ can be interpreted as the meeting point of the geodesics $g_1$ and $g_2$. The number $a_{12}$ denotes the asymptotic height of this meeting point (see also (10) in Definition 2.8). When traveling on a geodesic from infinity, it is only possible to change to the other geodesic by passing through the meeting point $v_{12}$. The admissibility condition (6) enforces that the asymptotic height profile has a similar property. Using this interpretation of the number $a_{ij}$ it also becomes clear why the compatibility condition (4) is needed (see also Figure 2). This interpretation is made precise in the following lemma.

**Lemma 2.3.** We consider a map $h_R : R \to \mathbb{R}^+ \times \{1, \ldots, k\}$. We assume that the first coordinate $h^1$ is continuous and that the numbers $a_{ij}$ satisfy the conditions (3) and (4). Then it is equivalent:

- The map $h$ satisfies the condition (6).
- For any two points $x, y \in R$ and any path $p \subset R$ that connects $x$ and $y$ there is a point $z \in p$ such that

$$h^1_R(z) \leq a_{h^2(x), h^2(y)}.$$  \hspace{1cm} (7)

We state the proof of Lemma 2.3 in Section 6. Let us consider an example and assume that there are points $x, y \in R$ such that the second coordinate $h^2_R(x) = 1$ and $h^2_R(y) = 2$. This indicates that the asymptotic height function $h^1_R$ travels at $x$ on the geodesic $g_1$ and at $y$ on the geodesic $g_2$. Now, let us consider a path $p \subset R$ from $x$ to $y$. Then the asymptotic height function has to change geodesics on that path. The admissibility condition (6) enforces that changing from the geodesic $g_1$ to the geodesic $g_2$ can only take place below the meeting point, which is characterized by the height $a_{12}$.

For an illustration of an asymptotic height profile $(h_R, a_{1,2})$ we refer to Figure 3. On the blue region $R_{\text{blue}} \subset R$, the height profile travels on the geodesic $g_1$. On the red region $R_{\text{red}} \subset R$, the height profile travels on the geodesic $g_2$. Mathematically, this means that the second
The coordinate of $h_{R}$ satisfies
\[ h_{R}^{2}(x) = \begin{cases} 1, & \text{if } x \in R_{\text{blue}}, \\ 2, & \text{if } x \in R_{\text{red}}. \end{cases} \]

The yellow line $L_{\text{yellow}}$ separates the blue region $R_{\text{blue}}$ and the red region $R_{\text{red}}$. The admissibility condition (6) means that one can only cross from $R_{\text{blue}}$ to $R_{\text{red}}$ below the meeting point of $g_{1}$ and $g_{2}$. Hence, the first coordinate of $h_{R}$ satisfies for all $x \in L_{\text{yellow}}$
\[ h_{R}^{1}(x) \leq a_{1,2}. \]

In a variational principle only the boundary condition is prescribed. For that reason, we now adapt Definition 2.2 and define the notion of an asymptotic boundary height profile.

**Definition 2.4 (Asymptotic boundary height profile).** Let $k \in \mathbb{N}$, let $h_{\partial R} : \partial R \to \mathbb{R}^{+} \times \{1, \ldots, k\}$ be a function and let $(a_{ij})_{k \times k}$ be a set of non-negative real numbers satisfying the condition (3) and (4). We say that $(h_{\partial R}, (a_{ij})_{k \times k})$ is an asymptotic boundary height profile if it satisfies the conditions (5) and (6) from above and the following condition: For all $x, y \in \partial R$ it holds
\[ |h_{\partial R}^{1}(x) - a_{h_{\partial R}^{2}(x), h_{\partial R}^{2}(y)}l_{1}| + |a_{h_{\partial R}^{2}(x), h_{\partial R}^{2}(y)} - h_{\partial R}^{1}(y)|l_{1} \leq |x - y|l_{1}. \] (8)

Compared to the Definition 2.2 of an asymptotic height profile, the condition (8) is new. It is needed to guarantee that every asymptotic boundary height profile $h_{\partial R}$ can be extended to an asymptotic boundary height function.

**Lemma 2.5.** Let $(h_{\partial R}, (a_{ij})_{k \times k})$ be an asymptotic boundary height function in the sense of Definition 2.4. Then it can be extended a asymptotic height profile $(h_{\partial R}, (a_{ij})_{k \times k})$ on the full region $R$. 
The proof of Lemma 2.5 is stated in Section 6. Lemma 2.5 is important because otherwise the statement of the variational principle, formulated in Theorem 2.9 below, could be empty.

The next step toward the variational principle is to define in which sense a sequence of (boundary) graph homomorphisms \( h_{\partial R_n} : \partial R_n \to T \) convergences to an asymptotic height profile \( (h_{\partial R}, (a_{ij})_{k \times k}) \). For this purpose, let us introduce some necessary definitions.

**Definition 2.6. (Geodesic on the tree \( T \))** Let \( T \) denote the \( d \)-regular tree with root \( r \). A graph homomorphism \( g : \mathbb{N} \to T \) is called geodesic if the map \( g \) is one-to-one.

Let \( \infty_g \) denote a boundary point associated to a directed geodesic \( g \subset T \) starting at the root \( r \in g \). We denote with \( \partial T \) the set of all such boundary points i.e. (see for example [Klo08])

\[
\partial T := \{ \infty_g : g \text{ is a one-sided geodesic of } T \text{ starting in } r \in g \}.
\]

It follows from the definition that for two boundary points \( \infty_{g_1}, \infty_{g_2} \in \partial T \) there is a unique element \( v_{12} \in T \) such that (cf. Figure 4)

\[
\max_{v \in g_1 \cap g_2} d_T(r, v) = d_T(r, v_{12}).
\]

We will write

\[
|\infty_{g_1} \cap \infty_{g_2}| := \max_{v \in g_1 \cap g_2} d_T(r, v) = d_T(r, v_{12}),
\]

and call \( |\infty_{g_1} \cap \infty_{g_2}| \) the height of the meeting point of the two geodesics \( g_1 \) and \( g_2 \). We also need the following observation.

**Lemma 2.7.** Let \( g \subset T \) a geodesic on the graph \( T \) containing the root \( r \in T \). Then the geodesic \( g \) can be identified with a map \( g : \mathbb{Z} \to T \) such that:

- the map \( g \) is a graph homomorphism;
- the map \( g \) is one-to-one;
- \( g(0) = r \).

We are now ready to define the convergence of a sequence of boundary graph homomorphisms to an asymptotic boundary height profile.

**Definition 2.8.** Let \( h_{\partial R_n} : \partial R_n \to T \) be a sequence of boundary height functions and let \( (h_{\partial R}, (a_{ij})_{k \times k}) \) be an asymptotic boundary height profile in the sense of Definition 2.2. We say that the sequence \( h_{\partial R_n} \) converges to \( (h_{\partial R}, (a_{ij})_{k \times k}) \) (i.e. \( \lim_{n \to \infty} h_{\partial R_n} = (h_{\partial R}, (a_{ij})_{k \times k}) \)), if the following two conditions are satisfied:
There exist \( k \) sequences of boundary points \( \{ \infty_{g_{i,n}}, \ldots, \infty_{g_{k,n}} \} \) such that for all \( 1 \leq i, j \leq k \):

\[
\lim_{n \to \infty} \frac{1}{n} |\infty_{g_{i,n}} \cap \infty_{g_{j,n}}| = a_{ij}
\]

(10)

where \( |\infty_{g_{i,n}} \cap \infty_{g_{j,n}}| \) is the height of the meeting point of the two geodesics (see (9)).

For \( z \in \partial R_n \) we define the set

\[
S(z) := \partial R \cap \left\{ x \in \mathbb{R}^m : \|x - \frac{z}{n}\|_{\infty} \leq \frac{1}{2n} \right\}.
\]

Then it holds that

\[
\lim_{n \to \infty} \sup \sup_{\{z \in \partial R_n : S(z) \neq \emptyset\}} \frac{1}{n} \mathrm{dist}_T \left( h_{\partial R_n}(z), g_{h_{\partial R}(x,n)} \left( \lfloor nh_{\partial R}(x) \rfloor \right) \right) = 0.
\]

(11)

where \( h_{1_{\partial R}} \) and \( h_{2_{\partial R}} \) are the two components of the map \( h_{\partial R} \).

Definition 2.8 is illustrated in Figure 4. The condition (10) ensures that the quantity \( a_{ij} \) characterizes the asymptotic meeting point of the geodesics \( g_i \) and \( g_j \). One can observe that the compatibility condition (4) on \( a_{ij} \) is actually a consequence of the condition (10). The condition (11) asymptotically characterizes the values of graph homomorphism \( h_{\partial R_n} \) via the the asymptotic height profile.

Let us now formulate the main result of this article, namely the variational principle for graph homomorphisms to a regular tree. As we outlined in the introduction, a variational principle contains two statements. The first statement, namely Theorem 2.9, gives a variational characterization of the entropy (cf. (2))

\[
\mathrm{Ent} \left( \mathbb{R}_n, h_{\partial R_n} \right) = -\frac{1}{n^2} \ln |M(R_n, h_{\partial R_n})|.
\]

Hence it asymptotically characterizes the number of possible graph homomorphisms \( h_n \in M(R_n, h_{\partial R_n}) \) with boundary data \( h_{\partial R_n} \).

**Theorem 2.9** (Variational principle). We assume that \( R \subset \mathbb{R}^m \) is a compact, simply connected region with Lipschitz boundary \( \partial R \). We consider a lattice discretization \( R_n \subset \mathbb{Z}^m \) of \( R \) such that the rescaled sublattice \( \frac{1}{n} R_n \) converges to \( R \) in the Gromov-Hausdorff sense.

We assume that the boundary height functions \( h_{\partial R_n} \) converge to an asymptotic boundary height profile \( (h_{\partial R}, (a_{ij})_{k \times k}) \) in the sense of Definition 2.8.

Let \( \mathrm{AHP}(h_{\partial R}, (a_{ij})_{k \times k}) \) denote the set of asymptotic height profiles that extend \( (h_{\partial R}, (a_{ij})_{k \times k}) \) from \( \partial R \) to \( R \).
Given an element $h_R \in AHP(h_{\partial R}, (a_{ij})_{k \times k})$, we define the macroscopic entropy via

$$\text{Ent} (R, h_R) = \int_R \text{ent} (\nabla h_R^1(x)) \, dx,$$

where the local surface tension $\text{ent}(s_1, \ldots, s_m)$ is given by Theorem 3.1 from below. Then it holds that

$$\lim_{n \to \infty} \text{Ent}(\Lambda_n, h_{\partial R_n}) = \min_{h_R \in AHP(h_{\partial R}, (a_{ij})_{k \times k})} \text{Ent} (R, h_R).$$

The local surface tension will be defined in Section 3.1 as a limit of carefully chosen entropies. Contrary to the case of domino tilings, we do not have an explicit formula for the local surface tension $\text{ent}(s_1, \ldots, s_m)$. The convexity of the local surface tension $\text{ent}(s_1, \ldots, s_m)$ is deduced in Section 3.1. In analogy to domino tilings, the authors believe that the local surface tension is strictly convex, but they are missing a proof. As a consequence, we do not know if the minimizer of the continuous entropy is unique.

Let us now turn to the second part of the variational principle, namely the profile theorem (see Theorem 2.11 from below). The profile theorem contains information about the profile of a graph homomorphisms $h_n$ that is chosen uniformly random from $M(R_n, h_{\partial R_n})$.

In a non-rigorous way, the statement of Theorem 2.11 is the following. Let us consider an asymptotic boundary height profile $h_R \in AHP(h_{\partial R}, (a_{ij})_{k \times k})$. Then the continuous entropy $\text{Ent}(h_R)$ is given
by the number of graph homomorphisms $h_n \in M(R_n, h_{\partial R_n})$ that are close to $h_R$. Applying this statement to the minimizer $h_{\text{min}}$ of the continuous entropy $\text{Ent}(h)$ has the following consequence. The uniform measure on the set of graph homomorphisms $M(R_n, h_{\partial R_n})$ concentrates on graph homomorphisms $h_n$ that have a profile that is close to $h_{\text{min}}$. As a consequence, a uniform sample of $M(R_n, h_{\partial R_n})$ will have a profile that is close to the minimizing profile $h_{\text{min}}$ for large $n$.

Let us now make this discussion precise. For that purpose, we have to specify when the profile of a graph homomorphism $h_n$ is close to an asymptotic height profile $h$.

**Definition 2.10.** For fixed $\varepsilon > 0$, let us consider the grid $R_{\text{grid}, \varepsilon}$ with $\varepsilon$-spacing contained in $R$. More precisely, $R_{\text{grid}, \varepsilon}$ is given by (see Figure 5)

$$R_{\text{grid}, \varepsilon} := \{ x = (z_1, \ldots, z_m) \in R \mid \exists 1 \leq k \leq m : |x_k| \in \varepsilon \mathbb{N} \}.$$ 

For a given asymptotic height profile $h$, we define the ball $HP_n(h, \delta, \varepsilon)$ of size $\delta > 0$ on the scale $\varepsilon > 0$ by the formula

$$HP_n(h, \delta, \varepsilon) = \left\{ h_n \in M(R_n, h_{\partial R_n}) \mid \sup_{x \in R_n, z \in R_{\text{grid}, \varepsilon}} \left| \frac{1}{n} d_T(h_n(x), r) - h^1 \left( \frac{x}{n} \right) \right| \leq \delta \right\},$$

where the set $M(R_n, h_{\partial R_n})$ of graph homomorphisms is given by (1).

Now, let us formulate the profile theorem.
Theorem 2.11. (Profile theorem) Let \((h_R, (a_{ij})_{k \times k})\) be an extension of the asymptotic boundary height profile \((h_{\partial R}, (a_{ij})_{k \times k})\). Then

\[
\text{Ent}(R, h_R) = -\frac{1}{|R_n|} \ln |HP_n(h_R, \delta, \varepsilon)| + \theta(\varepsilon) + \theta(\delta) + \theta\left(\frac{1}{\varepsilon n}\right)
\]

(13)

where \(\theta\) denotes a generic smooth function with \(\lim_{x \to 0} \theta(x) = 0\).

Remark 2.12. We want to point out that the second coordinate \(h^2\) does not play a role in the definition (12) of \(HP_n(h, \delta, \varepsilon)\). This means that we neglect the information which geodesic a graph homomorphism \(h_n \in M(R_n, h_{\partial R_n})\) follows. This can be done because the entropic effect of choosing the geodesics is of lower order. Rigorously, this fact is deduced in Lemma 6.1 below. Let us now give a heuristic argument. The variational principle lives on the scale \(|R|\). Approximating the set \(R\) by blocks of side length \(\varepsilon n\) it follows that \(|R| \approx l \varepsilon^n m^n\), where \(l\) is the number of blocks. Having a close look at the definition (12) of \(HP_n(h, \delta, \varepsilon)\) shows that only the grid \(R_{\text{grid}, \varepsilon}\) is important. Hence, the entropic effect of choosing different geodesics lives at most on the scale of the length of the grid \(R_{\text{grid}, \varepsilon}\). The length of the grid \(R_{\text{grid}, \varepsilon}\) is of the order \(l \varepsilon^{m-1} n^{m-1}\) and therefore negligible on the scale of the variational principle.

It follows from Remark 2.12 that choosing different geodesics \(g_1, \ldots, g_k\) and meeting points \((a_{ij})_{k \times k}\) has no effect on Theorem 2.11. However, it still has an effect on the variational principle formulated in Theorem 2.9. Choosing different geodesics and meeting points changes the set \(AHP(h_{\partial R}, (a_{ij})_{k \times k})\) of asymptotic height functions \(h_R\) over which the continuous entropy \(\text{Ent}(R, h_R)\) is minimized. This is another main aspect how the variational principle of Theorem 2.9 is distinct from the variational principle of domino tilings [CKP01].

The proof of Theorem 2.9 and of Theorem 2.11 is stated in Section 6. The argument needs a lot of preparation. For example, we first have to define the local surface tension \(\text{ent}(s_1, \ldots, s_m)\). This is done in Section 3. There, we also show that the local surface tension is convex. The main technical tools for the proof of the variational principle, namely the Kirszbraun theorem for graphs and the concentration inequality, are provided in Section 4.

3. The local surface tension

The purpose of this section is to show the existence of the local surface tension \(\text{ent}(s), s \in \mathbb{R}^m\) (see Theorem 3.1 from below). Additionally, we will also show in this section that the local surface tension \(\text{ent}(s)\) is convex (see Theorem 3.2 from below). We use a similar approach
A VARIATIONAL PRINCIPLE FOR A NON-INTEGRABLE MODEL

as in [CKP01] in the sense that the local surface tension \( \text{ent}(s) \) will be defined as limit of a microscopic surface tension \( \text{ent}_n(s) \), i.e.

\[
\text{ent}(s) := \lim_{n \to \infty} \frac{1}{n} \text{ent}_n(s).
\]

Even if the strategy is clear there are a lot of challenges. The first one is to find the right definition of the microscopic surface tension \( \text{ent}_n(s) \). For this, we have to generalize the notion of periodicity from \( \mathbb{Z} \)-valued height functions to general graph homomorphisms. The precise definition of \( \text{ent}_n(s) \) is given in Section 3.1 below.

The bigger challenge is to show that the limit of the microscopic surface tensions exists.

**Theorem 3.1.** Let \( s \in \mathbb{R}^m \) such that \( |s|_\infty < 1 \) and let \( \text{ent}_n(s) \) be given by Definition 3.7. Then the limit

\[
\text{ent}(s) := \lim_{n \to \infty} \text{ent}_n(s)
\]

exists and defines the local surface tension \( \text{ent}(s) \).

The proof of Theorem 3.1 is stated in Section 3.2. It is complex and needs auxiliary technical preparations. Usually, in the field of discrete maps the existence of the local surface tension is shown by exact calculation using the integrability of the underlying model (cf. for example [CKP01]). We cannot use a similar procedure due to the absence of integrability in our model. In our new method, we substitute the integrability by using two ingredients:

- The first ingredient is a Kirszbraun theorem for graphs (see Theorem 4.1 from below). It states under which conditions one can attach together two different graph homomorphisms.
- The second ingredient is a concentration inequality (see Theorem 4.6 from below). It states that, for canonical boundary data, a graph homomorphism cannot deviate too much from a linear height profile.

Those two ingredients are not only fundamental for the proof of Theorem 3.1 but also for deducing the variational principle (i.e. Theorem 2.11). Let us explain this remark in more detail. One of the main ingredients in the proof of the variational principle is that the entropy of a large box with fixed boundary condition is asymptotically close to the entropy with a well-chosen free boundary condition (see Theorem 3.9 from below). From the definition it is clear that the entropy
of the box with free boundary conditions is larger than the entropy of
the box with a fixed boundary condition. Hence, it is only left to show
that one can control the entropy with free boundary condition from
above by the entropy with fixed boundary condition. In order to do so,
one first applies the concentration inequality to show that the entropy
with free boundary conditions is controlled by the entropy on a slightly
smaller box with a boundary condition that allows fluctuations. In the
second step one applies the Kirszbraun theorem to show that such a
graph homomorphism on the smaller box can be extended to a graph
homomorphism on the original box satisfying the fixed boundary con-
dition. For details we refer to the proof of Theorem 3.9.

Once the existence of the local surface tension $\text{ent}(s)$ is established it
is natural to ask if the local surface tension is convex. This is the case
in our model.

**Theorem 3.2.** The local surface tension $\text{ent}(s)$ given by Definition 3.7
is convex in every coordinate. In particular, this implies that $\text{ent}(s)$ is
convex.

The proof of Theorem 3.2 is stated in Section 3.3. As in the proof of
the existence of the local surface tension $\text{ent}(s)$, the main tools of
the argument are the Kirszbraun theorem (see Theorem 4.1 from below)
and the concentration inequality (see Theorem 4.6 from below). Given
that the local surface tension $\text{ent}(s)$ is convex, it is natural to ask
if $\text{ent}(s)$ is also strict convex. We believe that this is the case but we
are missing a proof.

### 3.1. Definition of the microscopic surface tension.

In order to
define the microscopic surface tension $\text{ent}_n(s)$ we need to study the
translation invariant measures of our model. For this reason, we start
with generalizing of the notion of periodicity of height functions to
graph homomorphisms. For this purpose, we will identify the $d$-regular
rooted tree $T$ with the group

$$G = \langle \alpha_1, \ldots, \alpha_d | \alpha_1^2 = \ldots = \alpha_d^2 = e \rangle.$$ 

This is done through the natural bijection induced by the Cayley graph
of $G$ generated by the $\alpha_i$’s. We use the convention that the root of $T$
is represented by the identity of $G$. The reason for this identification
is that the group structure provides an easy way to define gradient
measures. Using the previous bijection we can choose a canonical way
to associate a unique $\alpha_i$ to each edge of $T$. As a consequence, there is
a natural way to associate to a graph homomorphism $h$ a dual function
$\tilde{h}$ acting on edges of $\mathbb{Z}^m$:
Definition 3.3 (Dual of a graph homomorphism). Let $h : \mathbb{Z}^m \to \mathcal{T}$ be a graph homomorphism. We define its dual map

$$\tilde{h} : \{e_{x,y} | x, y \in \mathbb{Z}^m : |x - y| = 1\} \to \{\alpha_1, \ldots, \alpha_d\}$$

in the following way. Note that for any $x \sim y \in \mathbb{Z}^m$ there is a unique $a_i$ such that $h(y) = a_i h(x)$. Then, the value of dual function $\tilde{h}$ on the edge $e_{xy}$ is given by $\tilde{h}(e_{xy}) = a_i$.

The dual map $\tilde{h}$ determines the graph-homomorphism up to translations in the graph $\mathcal{T}$. If there is no source of confusion, we will denote the dual map and the graph homomorphism with the same symbol $h$.

The dual map $\tilde{h}$ maps each path $p = \{x_0, \ldots, x_n\}$ in $\mathbb{Z}^m$ onto a word in the alphabet $\{\alpha_1, \ldots, \alpha_d\}$. We will use the notation $\tilde{h}(p)$ for this word.

We are now ready to define the analog of periodicity for graph homomorphisms.

Definition 3.4 (Translation invariant graph homomorphism). We denote by $\vec{i}_k$ the $k$-th vector of the standard basis of $\mathbb{Z}^m$. Let $h : \mathbb{Z}^m \to \mathcal{T}$ be a graph homomorphism. We say that $h$ is $n$-translational invariant if for all $x \sim y \in \mathbb{Z}^m$ and $k \in \{1, \ldots, m\}$

$$\tilde{h}(e_{xy}) = \tilde{h}(e_{(x+n\vec{i}_k)(y+n\vec{i}_k)})$$

Figure 6. A translation invariant configuration on a 3-regular tree. Each color of an edge represents one of the $\alpha_i$'s

In order to define the microscopic surface tension we need to associate to every $n$-translation invariant homomorphism $h : \mathbb{Z}^m \to \mathcal{T}$ a slope, which indicates the speed at which the homomorphism travels on the graph in every direction of the plane.
Definition 3.5 (Slope of a translation invariant graph homomorphism). Let $h : \mathbb{Z}^m \to \mathcal{T}$ be a $n$-translation invariant homomorphism. The slope $s = (s_1, \ldots, s_m)$ of $h$ is defined by

$$s_k = \frac{1}{n} \min_{x \in \mathbb{Z}^m} d_{\mathcal{T}}(h(x), h(x + n\vec{i}_k)), \quad \text{for } 1 \leq k \leq m.$$ 

An essential property of $n$-invariant homomorphisms is that they must stay within finite distance of a unique geodesic of $\mathcal{T}$ if the slope is nonzero or stay within finite distance of a single point if the slope is zero. This statement is made more precise in the next lemma.

Lemma 3.6. Let $h : \mathbb{Z}^m \to \mathcal{T}$ be a $n$-invariant homomorphism with slope $s \neq 0$. Then there exist a unique geodesic $g \subset \mathcal{T}$ such that for all $x \in \mathbb{Z}^m$

$$\text{dist}_{\mathcal{T}}(h(x), g) \leq \frac{n}{2}. \quad (14)$$

In this case we say that $h : \mathbb{Z}^m \to \mathcal{T}$ is supported on the geodesic $g$.

If $h$ has slope $s = 0$ then $h$ has finite range, and for all $x \in \mathbb{Z}^m$

$$\text{dist}_{\mathcal{T}}(h(x), h(0)) \leq \frac{mn}{2}. \quad (15)$$

Proof of Lemma 3.6. We start with considering the case where the slope of $h$ is $(0, \ldots, 0)$. In this case the $n$-invariance yields that for $(l_1, \ldots, l_m) \in \mathbb{Z}^m$:

$$h(x_1, \ldots, x_m) = h(x_1 + l_1n, \ldots, x_m + l_mn).$$

Now, the estimate (15) follows directly from the observation that any point $x \in \mathbb{Z}$ is within graph distance $\frac{mn}{2}$ of the set

$$\left\{ k_1\vec{i}_1 + \ldots + k_m\vec{i}_m \in \mathbb{Z}^m \mid k_i \in \mathbb{Z} \right\}.$$

Consider now the case where the slope of $h$ is not zero. We start by noticing that any geodesic $g$ that satisfies (14) must be unique. Indeed, since $G$ is hyperbolic, two geodesics cannot stay within finite distance. Therefore $h(\mathbb{Z}^m)$ can only stay within finite distance of at most one geodesic $g$.

Let us now deduce the estimate (14). Let $s_i$ be a non-zero coefficient of the slope $s = (s_1, \ldots, s_m)$. Without loss of generality we assume that $s_1 > 0$. This means that along the first coordinate the map $h$ travels on a geodesic $g$ with speed $s_1$. This implies that there is an integer $x_1 \in \mathbb{Z}$ such that

$$\{h(ln + x_1, 0, \ldots, 0), l \in \mathbb{Z} \} \subset g$$

With a simple indirect argument contradicting the graph homomorphism property of $h$ it follows that the image of all lines parallel to $\vec{i}_1$
under \( h \) must also travel on the geodesic \( g \) with speed \( s_1 \). This means that for all integers \( x_2, \ldots, x_m \in \mathbb{Z} \) there is an integer \( x_1 \) such that
\[
\{h(ln + x_1, x_2, \ldots, x_m), l \in \mathbb{Z}\} \subset g.
\]

Now, let \( z_1 \in \mathbb{Z} \) be arbitrary. Then we can write \( z_1 = ln + x_1 + r \) for some numbers \( l \in \mathbb{Z} \) and \(-\frac{n}{2} \leq r < \frac{n}{2}\). By additionally using that \( h \) is a \( n \)-translational invariant graph homomorphisms it follows that
\[
\begin{align*}
dist_T(h(z_1, x_2, \ldots, x_m), g) \\
\leq dist_T(h(z_1, x_2, \ldots, x_m), h(ln + x_1, x_2, \ldots, x_m)) \\
+ dist_T(h(ln + x_1, x_2, \ldots, x_m), g) \\
\leq \frac{n}{2},
\end{align*}
\]
which is the desired estimate (14). \( \square \)

Now, we have everything that is needed to define the microscopic surface tension \( \text{ent}_n(s) \).

**Definition 3.7** (Microscopic surface tension \( \text{ent}_n(s) \)). Let \( g \in T \) be an arbitrary geodesic. For \( s = (s_1, \ldots, s_m) \in [-1, 1]^m \) we denote by \( \mathcal{H}_n^g(s) \) the set
\[
\mathcal{H}_n^g(s) := \{ h : \mathbb{Z}^m \to T : h \text{ is } n \text{-invariant} \}
\]
with slope \( \left( \frac{|s_1|}{n}, \ldots, \frac{|s_m|}{n} \right) \) supported on \( g \)
and \( \Pi_g(h(0)) = g(0) \),
where \( \Pi_g : T \to g \subset T \) denotes the projection onto the geodesic \( g \).
For \( (s_1, \ldots, s_m) = 0 \) we define
\[
\mathcal{H}_n^g(0, \ldots, 0) := \{ h : \mathbb{Z}^m \to T : h \text{ is } n \text{-invariant} \}
\]
with slope \( (0, \ldots, 0) \) and \( h(0) = r \).

The microscopic surface tension \( \text{ent}_n(s) \) is defined as
\[
\text{ent}_n(s) := -\frac{1}{n^m} \ln |\mathcal{H}_n^g(s)|.
\]

Because all geodesics are equivalent, the definition of \( \text{ent}_n(s) \) is independent from the particular choice of \( g \). In particular by re-orientating the geodesic \( g \) we can assume wlog. that \( s_1 \geq 0 \). We denote by \( \mathbb{P}_n^g \) the uniform probability measure on \( \mathcal{H}_n^g(s) \).

We want to note that the Definition 3.7 of \( \text{ent}_n(s) \) is well posed, i.e. the set \( \mathcal{H}_n^g(s) \) is not empty. Indeed, one can easily construct elements of \( \mathcal{H}_n^g(s) \) by using the Kirszbraun theorem for graphs (see Theorem 4.1 from below).
3.2. Existence of the local surface tension. The purpose of this section is to state the proof of Theorem 3.1, i.e. showing the existence of the local surface tension

$$\operatorname{ent}(s) = \lim_{n \to \infty} \operatorname{ent}_n(s).$$

The first step towards the proof of Theorem 3.8 is the following statement, which shows that the microscopic surface tension is not oscillating wildly.

Lemma 3.8. Let $n_1 \leq n_2$. Then

$$|\operatorname{ent}_{n_1}(s) - \operatorname{ent}_{n_2}(s)| \leq C \left( 1 - \left( \frac{n_1}{n_2} \right)^m + \frac{n_2 - n_1}{n_2} + \frac{1}{n_2^m} \right).$$

Proof of Lemma 3.8. Before starting the argument, let us recall the definition of $\operatorname{ent}_n(s)$. It is defined via

$$\operatorname{ent}_n(s) := -\frac{1}{n^m} \ln |\mathcal{H}_n^g(s)|.$$

In the first step of the argument, we show that, if we denote by $\cdot$ the usual inner product of $\mathbb{R}^m$, the size of the set $\mathcal{H}_n^g(s)$ is comparable to the set

$$M_n := \left\{ h \in \mathcal{H}_n^g(s) \mid \max_{x \in S_n} d_T(g(|s \cdot x|), h(x)) \leq n^{0.6} \right\}.$$

Indeed, the concentration inequality (41) of Theorem 4.6 yields

$$\mathbb{P}_n \left( \max_{x \in S_n} d_T(g(|s \cdot x|), h(x)) \geq n^{0.6} \right) \leq Ce^{-\frac{n^{0.2}}{122}}$$
This implies that for $n$ large enough
\[ \mathbb{P}_n^s (h \in M_n) \geq 1/2. \]
Because $\mathbb{P}_n^s$ is the uniform measure on $H^s_n(s)$, this yields that for all sufficiently large $n \in \mathbb{N}$
\[ |M_n| \leq |H^s_n(s)| \leq 2|M_n|. \tag{17} \]
For the second step of the argument, let us assume that the box $S_{n_1}$ is centered within the box $S_{n_2}$ (see Figure 7). We will show that
\[ |M_{n_1}| \leq |M_{n_2}| \leq |M_{n_1}| d^{(n_2-n_1)n_2^{(m-1)}}. \tag{18} \]
Indeed, the first inequality $|M_{n_1}| \leq |M_{n_2}|$ follows from the fact that any element $h \in M_{n_1}$ can be extended to a function $\bar{h} \in M_{n_2}$. The inequality $|M_{n_2}| \leq |M_{n_1}| d^{(n_2-n_1)n_2^{(m-1)}}$ follows from the following argument. Let $\bar{h} \in M_{n_2}$. Then its restriction $\bar{h}|_{S_{n_1}}$ to the box $S_{n_1}$ is in $M_{n_1}$. Additionally, we observe that there are less than $d^{(n_2-n_1)n_2^{(m-1)}}$ many ways to extend a configuration on $S_{n_1}$ to the box $S_{n_2}$. This yields the desired estimate
\[ |M_{n_2}| \leq |M_{n_1}| d^{(n_2-n_1)n_2^{(m-1)}}, \]
and verifies the estimate 18.

Let us now turn to the third and last step of the proof. First we observe that due to (17) it holds
\[ \left| -\frac{1}{n_2^m} \ln |M_{n_2}| - \text{ent}_{n_2}(s) \right| \leq \frac{1}{n_2^m} \ln 2. \tag{19} \]
Additionally, we observe by taking the log of (18) that
\[
\left( \frac{n_1}{n_2} \right)^m \text{ent}_{n_1}(s) \geq -\frac{1}{n_2^m} \ln |M_{n_1}| - \frac{1}{n_2^m} \ln 2
\geq -\frac{1}{n_2^m} \ln |M_{n_2}| - \frac{1}{n_2^m} \ln 2
\geq -\frac{1}{n_2^m} \ln |M_{n_1}| - \frac{(n_2-n_1)n_2^{(m-1)}}{n_2^m} \ln d - \frac{1}{n_2^m} \ln 2
= \left( \frac{n_1}{n_2} \right)^m \text{ent}_{n_1}(s) - \frac{n_2-n_1}{n_2} \ln d - \frac{1}{n_2^m} \ln 2.
\]
By observing that $-\ln d \leq \text{ent}_n(s) \leq 0$, the last estimate yields that
\[ \left| -\frac{1}{n_2^m} \ln |M_{n_2}| - \text{ent}_{n_1}(s) \right| \leq C \left( 1 - \left( \frac{n_1}{n_2} \right)^m + \frac{n_2-n_1}{n_2} + \frac{1}{n_2^m} \right). \tag{20} \]
A combination of (19) and (20) yields the desired inequality (16). \( \square \)
We want to point out that $\text{ent}_n(s)$ corresponds to a microscopic entropy with free boundary conditions. Another ingredient for the proof of Theorem 3.1 is that the microscopic entropy with free boundary condition and fixed boundary condition are equivalent:

**Theorem 3.9.** Let $\delta > 0$, $S_n$ be a $n \times n$ square, $g$ be a fixed geodesic in $\mathcal{T}$ and $s \in \mathbb{R}^M$ such that $|s|_\infty \leq 1$. Let $h_{\partial S_n} : \partial S_n \to \mathcal{T}$ such that for all $x \in \partial S_n$ it holds

$$d_\mathcal{T}(h_{\partial S_n}(x), g([s \cdot x])) \leq \delta n,$$

where $g([s \cdot x])$ is given by Lemma 2.7. Then it holds that

$$\text{Ent}(S_n, h_{\partial S_n}) = \text{ent}_n(s) + \theta \left( \frac{1}{n} \right) + \theta(\delta),$$

where $\text{Ent}(S_n, h_{\partial S_n})$ is given by (2) and $\text{ent}_n(s)$ is the microscopic surface tension given by Definition 3.7.

**Proof.** In order to deduce the estimate it suffices to show that

$$\text{ent}_n(s) \leq \text{Ent}(S_n, h_{\partial S_n}) + \theta(\delta) + c \frac{\delta}{n^{m-1}}$$

and

$$\text{ent}_n(s) \geq \text{Ent}(S_n, h_{\partial S_n}) + \theta(\delta) + c \frac{\delta}{n^{m-1}}.$$  

(23)

We start with deducing the estimate (22). By the concentration estimate of Theorem 4.6 from below, it follows that

$$\mathbb{P}_{(1-2\delta)n}^s \left( \max_{x \in \mathbb{Z}^m} d_\mathcal{T}(h(x), g([s \cdot x])) \geq \delta n \right) \leq C e^{-c\delta^2 n},$$

(24)

where $\mathbb{P}_{(1-2\delta)n}^s$ denotes the the uniform measure on $\mathcal{H}_{(1-2\delta)n}^g(s)$. Let us define the set $M_{(1-2\delta)n, \delta} \subset \mathcal{H}_{(1-2\delta)n}^g(s)$ according to

$$M_{(1-2\delta)n, \delta} := \left\{ h \in \mathcal{H}_{(1-2\delta)n}^g(s) : \max_{x \in \mathbb{Z}^m} d_\mathcal{T}(h(x), g([s \cdot x])) \leq \delta n \right\}.$$

Then it follows from (24) that

$$\text{ent}_{(1-2\delta)n}(s) \geq -\frac{1}{(1 - 2\delta)m^{n}} \ln |M_{(1-2\delta)n, \delta}| - \frac{C}{n^{m}}.$$  

(25)

We observe that due to (21) and the Kirszbraun theorem (cf. Theorem 4.1) any element $h_{(1-2\delta)n} \in M_{(1-2\delta)n, \varepsilon}$ can be extended to a graph homomorphism $h : S_n \to \mathcal{T}$ such that $h = h_{\partial S_n}$ on $\partial S_n$ and $h = h_{(1-2\delta)n}$
on $S_{(1-2\delta)n}$ (cf. the argument in the proof of Theorem 3.1). This implies that for large enough $n$

$$\text{Ent}(S_n, h_{\partial S_n}) \leq -\frac{1}{nm} \ln |M_{(1-2\delta)n,\delta}|$$

$$\leq (1 - 2\delta)^m \text{ent}_{(1-2\delta)n}(s) + c \frac{\delta}{n^{m-1}},$$

where we have used the estimate (25) from above and the identity

$$\text{ent}_{(1-2\delta)n}(s) = \text{ent}_n(s) + \theta(\delta),$$

which follows from Lemma 3.8 for large enough $n$. This verifies the estimate (22).

The estimate (23) can be verified by a similar argument as was used for (22). Instead of restricting $S_n$ to a smaller box and compare $S_n$ with $H_{g_{k(1-2\delta)n}}$ and $M_{(1-2\delta)n,\delta}$, one has to extend the box $S_n$ and comparing it with $H_{g_{kl(1+2\delta)n}}$ and $M_{(1-2\delta)n,\delta}$. We omit the details. □

The merit of Theorem 3.9 is the following: In order to show that the limit $\lim_{n \to \infty} \text{ent}_n(s)$ exists it suffices to show that the limit

$$\lim_{n \to \infty} \text{Ent}(S_n, h_{\partial S_n})$$

exists. The advantage of considering $\text{Ent}(S_n, h_{\partial S_n})$ is that the boundary data $h_{\partial S_n}$ is fixed and can be chosen such that (21) is satisfied. Therefore, let us fix from now on one particular sequence of boundary data $h_{\partial S_n}$ that additionally to (21) also satisfies the following condition.

**Definition 3.10.** (Periodic boundary data) Let $h_{\partial S_n}$ denote a boundary graph homomorphism on the boundary $\partial S_n$ of the box $S_n$. Let $\tilde{h}_{\partial S_n}$ denote the dual boundary graph homomorphism on the edge set $\tilde{E}_{\partial S_n}$ of $\partial S_n$ (see Definition 3.3). We say that the boundary graph homomorphism $h_{\partial S_n}$ has well-periodic boundary data if the following two conditions are satisfied:

- If $e_{x,y} \in \tilde{E}_{\partial S_n}$ and $e_{x+ni,y+ni} \in \tilde{E}_{\partial S_n}$ for some $i \in \{1, \ldots, n\}$, then $\tilde{h}_{\partial S_n}(e_{x,y}) = \tilde{h}_{\partial S_n}(e_{x+ni,y+ni}).$

- If $e_{x,y} \in \tilde{E}_{\partial S_n}$ and $e_{x-ni,y+ni} \in \tilde{E}_{\partial S_n}$ for some $i \in \{1, \ldots, n\}$, then $\tilde{h}_{\partial S_n}(e_{x,y}) = \tilde{h}_{\partial S_n}(e_{x+ni,y+ni}).$

The Definition 3.10 has the following simple interpretation. The dual boundary graph homomorphism $\tilde{h}_{\partial S_n}$ can be understood as a coloring
of the edges of the set $\partial S_n$. Then the boundary data $h_{\partial S_n}$ is periodic, if the coloring of one face of $\partial S_n$ matches the coloring of the opposite face.

The advantage of using periodic boundary data is that one gets monotonicity of a subsequence of $\text{Ent}(S_n, h_{\partial S_n})$ for free.

**Lemma 3.11.** Under the same assumptions as in Theorem 3.9, let us consider the entropy $\text{Ent}(S_n, h_{\partial S_n})$. We additionally assume that the boundary graph homomorphism $h_{\partial S_n}$ is periodic in the sense of Definition 3.10. On the box $S_{kn}$ we consider the boundary condition $h_{\partial S_{kn}}$ that arises from attaching $k$ copies of $h_{\partial S_n}$ to each other. Then it holds that for all integers $k \in \mathbb{N}$

$$\text{Ent}(S_n, h_{\partial S_n}) \geq \text{Ent}(S_{nk}, h_{\partial S_{nk}}).$$

The proof of Lemma 3.11 follows from a simple underestimation of the configurations in $\text{Ent}(S_{nk}, h_{\partial S_{nk}})$. Because the boundary data $h_{\partial S_n}$ is periodic one can just take a configuration $h_{S_n}$ on $S_n$ and extend it to the box $S_n$ by attaching $k$ copies of $h_{S_n}$ to each other. By construction, the resulting configuration on $S_{nk}$ will have the correct boundary data $h_{\partial S_{nk}}$ and therefore the estimate (26) follows automatically. We omit the details of this proof.

Now, we have everything that is needed for the proof of Theorem 3.1.

**Proof of Theorem 3.1.** The main idea is to consider a sequence of periodic boundary data $h_{\partial S_n}$ (see Definition 3.10) that satisfies (21) and show that the limit

$$\lim_{n \to \infty} \text{Ent}(S_n, h_{\partial S_n}) := E(s)$$

exists. Then it easily follows from statement of Theorem 3.9 that limit

$$\lim_{n \to \infty} \text{ent}_n(s) = E(s)$$

also exists. This would verify the statement of Theorem 3.1.

We begin with observing that

$$0 \geq e_n := \text{Ent}(S_n, h_{\partial S_n}) \geq -\ln d.$$ 

The reason is that edges take at most $d$-values. Therefore, it suffices to show that the sequence $e_n$ cannot have two distinct accumulations points $x_1$ and $x_2$. We argue by contradiction and assume that $x_1$ and $x_2$ are two accumulation points of the sequence $e_n$ satisfying the relation

$$-\ln d \leq x_1 < x_2 \leq 0.$$
Then there exists a number \( l \) such that
\[
e_l \leq \frac{x_1 + x_2}{2}.
\]
By Lemma 3.11, the subsequence \( k \to e_{kl} \) is decreasing, hence for all \( k \in \mathbb{N} \)
\[
e_{kl} \leq \frac{x_1 + x_2}{2}.
\]
We will now show that this implies for large enough \( l \) and \( k \) that also for all \( n \geq kl \)
\[
e_n \leq \frac{x_1 + x_2}{2} + \varepsilon,
\]
for some small constant \( \varepsilon > 0 \). This would be a contradiction to the assumption that \( x_1 \) and \( x_2 \) are accumulation points of the sequence \( e_n \) and therefore would verify (27).

Hence, it is left to deduce the estimate (27). Let \( n \geq lk \). Then we know that we can write
\[
n = \tilde{k}l + v,
\]
where \( \tilde{k} \geq k \) and \( 0 \leq v \leq l \). By using a combination of Theorem 3.9 and Lemma 3.8, it follows that
\[
|e_n - e_{\tilde{k}l}| \leq |\text{ent}_{\tilde{k}l+v}(s) - \text{ent}_{\tilde{k}l}| + \theta \left( \frac{1}{k\tilde{l}} \right) + \theta(\delta)
\]
\[
\leq \theta \left( \frac{1}{k} \right) + \theta(\delta).
\]
Hence, we see that if choosing \( k \) large enough and \( \delta \) small enough that
\[
e_n \leq e_{\tilde{k}l} + \varepsilon \leq \frac{x_1 + x_2}{2} + \varepsilon,
\]
which verifies (27) and closes the argument.

3.3. Convexity of the local surface tension. The purpose of this section is to prove Theorem 3.2 which states that the local surface tension \( \text{ent}(s) \) is convex.

Proof of Theorem 3.2. We will show that the local surface tension \( \text{ent}(s) \) is convex in every coordinate which yields that \( \text{ent}(s) \) is convex. By symmetry it suffices to show that \( \text{ent}(s) \) is convex in the first coordinate. For convenience, we only give the argument for the case \( m = 2 \). The argument for the general case is similar. For that reason let \( s_2 \) be fixed. We argue by contradiction. Hence, let us suppose that \( \text{ent}(s) \)
is not convex in the first coordinate. Then there are numbers \( s_{1,0} < s_{1,1} < s_{1,2} \) such that

\[
\frac{1}{2} s_{1,0} + \frac{1}{2} s_{1,2} = s_{1,1}
\]

and

\[
\text{ent}(s_{1,1}, s_2) > \frac{1}{2} \text{ent}(s_{1,0}, s_2) + \frac{1}{2} \text{ent}(s_{1,2}, s_2).
\]

(28)

For an integer \( n \) we consider the microscopic entropy \( \text{ent}_n(s_{1,1}, s_2) \) given by Definition 3.7 i.e.

\[
\text{ent}_n(s_{1,1}, s_2) := -\frac{1}{n^2} \ln |\mathcal{H}_n^g(s_{1,1}, s_2)|,
\]

where \( g \) denotes a geodesic in \( \mathcal{T} \). We want to recall that elements \( h \in \mathcal{H}_n^g(s_{1,1}) \) are graph homomorphisms \( h : S_n \to \mathcal{T} \), where \( S_n \subset \mathbb{Z}^2 \) denotes the \( n \times n \) box

\[
S_n := \{0, \ldots, n-1\}^2.
\]

We assume that without loss of generality that \( n \) is odd. The idea is to split up the box \( S_n \) into four boxes of side length \( n/2 \) i.e.

\[
S_n = B_1 \cup B_2 \cup B_3 \cup B_4.
\]

Down below, we will compare the number of graph homomorphisms in \( \mathcal{H}_n^g(s_{1,1}, s_2) \) to the number the number of graph homomorphisms on each sub-box \( B_i \) with a fixed buckled boundary (see Figure 8 and Figure 9). For that purpose let \( h_b \in \mathcal{H}_n^g(s_{1,1}, s_2) \) be a graph homomorphism such that

- for all \( x \in \partial S_n \) with \( 0 \leq x_1 \leq n-1/2 \) and \( x_2 \in \{0, n-1/2, n-1\} \) it holds
  \[
d_{\mathcal{T}} \left( h_b(x), g \left( \lfloor s_{1,0}x_1 + s_2x_2 \rfloor \right) \right) \leq \delta \frac{n}{2};
  \]
- for all \( x \in \partial S_n \) with \( n/2 \leq x_1 \leq n-1 \) and \( x_2 \in \{0, n-1/2, n-1\} \) it holds
  \[
d_{\mathcal{T}} \left( h_b(x), g \left( \lfloor s_{1,0}n - \frac{1}{2} + s_{1,2} \left( x_1 - \frac{n-1}{2} \right) + s_2x_2 \rfloor \right) \right) \leq \delta \frac{n}{2};
  \]
- for all \( x \in \partial S_n \) with \( x_1 = 0 \) and \( x_2 \in \{0, n-1\} \) it holds
  \[
d_{\mathcal{T}} \left( h_b(x), g \left( \lfloor s_2x_2 \rfloor \right) \right) \leq \delta \frac{n}{2};
  \]
- and
- for all \( x \in \partial S_n \) with \( x_1 = n-1 \) and \( x_2 \in \{0, n-1\} \) it holds
  \[
d_{\mathcal{T}} \left( h_b(x), g \left( \lfloor s_{1,1}n + s_2x_2 \rfloor \right) \right) \leq \delta \frac{n}{2}.
  \]
Figure 8. Schematic drawing of a typical graph homomorphism $h \in \mathcal{H}_g^S(s_1, s_2)$ on the block $S_n$. A blue line means that the graph homomorphism $h$ travels with speed $s_1$ on $g$ and a black line means that $h$ travels with speed $s_2$.

The role of the graph homomorphism $h_b$ is to fix the boundary condition on each box $B_i$, $i = 1, \ldots, 4$. Theorem 3.9 from above states that asymptotically the microscopic entropy of free and fixed boundary conditions are the same. Using Theorem 3.9 and underestimating the possible number of graph homomorphisms yields that

$$\operatorname{ent}_n(s_{1,1}, s_2) \leq 2 \left( \frac{n}{2} \right)^2 \frac{1}{n^2} \operatorname{ent}_2(s_{1,0}, s_2) + 2 \left( \frac{n}{2} \right)^2 \frac{1}{n^2} \operatorname{ent}_2(s_{1,2}, s_2) + o(\delta) + o \left( \frac{1}{n} \right).$$

The last estimate in combination with the fact that (cf. Theorem 3.1)

$$\operatorname{ent}_n(s_1, s_2) = \operatorname{ent}(s_1, s_2) + o \left( \frac{1}{n} \right)$$

implies that

$$\operatorname{ent}(s_{1,1}, s_2) \leq \frac{1}{2} \operatorname{ent}(s_{1,0}, s_2) + \frac{1}{2} \operatorname{ent}(s_{1,2}, s_2) + o(\delta) + o \left( \frac{1}{n} \right),$$

which contradicts (28) by choosing $\delta > 0$ small enough and $n$ large enough and therefore closes the argument. \(\square\)

4. A Kirszbraun theorem and a concentration inequality

In this section, we provide the technical tools that are needed in our proof of the variational principle to overcome the difficulty that our
model is not integrable. In Section 4.1 we deduce the Kirszbraun theorem in regular trees. In Section 4.2 we deduce the concentration inequality for graph homomorphisms. In Section 3 those tools were used to show the existence and convexity of the local surface tension \( \text{ent}(s) \) (see Theorem 3.1 and Theorem 3.2) and the equivalence of fixed and free boundary conditions (see Theorem 3.9). Those tools are also the technical foundation to derive the existence of a continuum of shift-invariant ergodic gradient Gibbs measures in Section 5.4.

4.1. A Kirszbraun theorem for graph homomorphisms. For continuous metrics, Kirszbraun theorems state that under the right conditions a \( k \)-Lipschitz function defined on a subset of a metric space can be extended to the whole space (cf. [Kir34, Val43, Sch69]). The goal of this section is to show that such theorems also exist for various spaces of discrete functions. We only consider the special case of graph homomorphisms from \( \mathbb{Z}^m \) to a \( d \)-regular tree for the convenience of the reader. The concepts of this section are quite universal and certainly could be applied to more general situations.

Since \( \mathbb{Z}^m \) and \( \mathcal{T} \) are both bipartite let us fix a 2-coloring of the two graphs. The color of a vertex is called parity. The next statement is the Kirszbraun theorem for graphs.

**Theorem 4.1** (Kirszbraun theorem for graphs). Let \( \Lambda \) be a connected region of \( \mathbb{Z}^m \), \( S \) be a subset of \( \Lambda \) and \( h : S \to \mathcal{T} \) be a graph homomorphism which conserves the parity. There exists a graph homomorphism...
\( h : \Lambda \to T \) such that \( h = \bar{h} \) on \( S \) if and only if for all \( x, y \) in \( S \)

\[
d_T(\bar{h}(x), \bar{h}(y)) \leq |x - y|_{l_1}.
\]

(29)

where \( |x - y|_{l_1} \) is the \( l_1 \)-norm in the graph \( \Lambda \).

**Remark 4.2.** If \( \Lambda \) is not convex then the \( l_1 \)-norm in \( \Lambda \) is not the \( l_1 \)-norm of \( \mathbb{Z}^m \).

**Remark 4.3.** The parity condition in Theorem 4.1 is necessary. Indeed, let us consider a situation where

\[ |x - y|_{l_1} = 2 \]

and

\[ d_T(\bar{h}(x), \bar{h}(y)) = 1. \]

Then it follows that the condition (29) is satisfied but there cannot be an extension \( h \) of \( \bar{h} \). However, the graph homomorphism \( \bar{h} \) in this example is violating the parity condition.

For the proof of Theorem 4.1 we need the following observation which states that once the image of a single point is fixed it is always possible to build a graph homomorphism that goes as fast as possible in one direction of the tree (here towards \( w_0 \)).

**Lemma 4.4.** Let \( \Lambda \) be a connected region of \( \mathbb{Z}^m \), \( x \) be a point in \( \Lambda \), \( w \) be a vertex of \( T \) and \( p = \{ w = v_0, \ldots, w_0 = v_k \} \) be the geodesic path in \( T \) going from \( w \) to \( w_0 \). The map \( h^w_x : \Lambda \to T \) given by \( h^w_x(y) = v_d(x, y) \) is a graph homomorphism.

**Proof.** The function \( h^w_x \) is 1-Lipschitz since the graph distance is 1-Lipschitz. Moreover two neighbors cannot have the same image because for bipartite graphs, the parity of the graph distance to a single point depends on the parity of the vertex. Therefore \( h^w_x \) is a graph homomorphism. \( \square \)

For the proof of Theorem 4.1 let us introduce the natural analogue of the norm of \( \cdot \) on a tree, which we call depth.

**Definition 4.5** (Depth on a tree). Let \( g \) be a geodesic of \( T \) and let \( \pm \infty_g \) denote the boundary point of the geodesic \( g \). The depth associated to the geodesic \( g \) is given by the unique function \( | \cdot | : T \to \mathbb{Z} \) such that \( |r| = 0 \) and for nearest neighbors \( v \sim w \in T \)

\[
|v| = |w| + \begin{cases} 
-1, & \text{if } v \text{ is closer to } \infty_g \text{ than } w, \\
1, & \text{if } v \text{ is closer to } \infty_g \text{ than } w.
\end{cases}
\]

We want to note that the depth can be negative. On the set \( \mathcal{H}_g(s) \) we always consider the depth associated with the geodesic \( g \).
Let us now turn to the proof of Theorem 4.1.

**Proof of Theorem 4.1.** The condition (29) is clearly necessary since a graph homomorphism is 1-Lipschitz. Suppose now that (29) holds. In order to prove Theorem 4.1, we only need to construct a graph homomorphism $h : \Lambda \to T$ such that $h = \bar{h}$ on $S$. For that purpose, let us consider

$$h(y) = \arg\max\{|h^\bar{h}(x)(y)|, x \in S\},$$

(30)

In other words, $h(y)$ is defined in the following way: Given $y \in \Lambda$ one has to find a vertex $\tilde{x} \in S$ such that

$$|h^\bar{h}(\tilde{x})(y)| = \max_{x \in S} |h^\bar{h}(x)(y)|.$$

Then one sets

$$h(y) = h^\bar{h}(\tilde{x})(y).$$

It is left to show that $h$ is well defined, that $h = \bar{h}$ on $S$ and that $h$ is a graph homomorphism. We start with deducing that $h$ is well defined. The fact that $h$ is well defined will follow from the following observation: If there are $y \in \Lambda$ and $x_1, x_2 \in S$ such that $|h^\bar{h}(x_1)(y)| = |h^\bar{h}(x_2)(y)|$ then

$$h^\bar{h}(x_1)(y) = h^\bar{h}(x_2)(y).$$

(31)

Therefore, let us deduce now the statement (31). We can assume without loss of generality that $x_1 \neq x_2$. By definition of $h^\bar{h}(x_i)$ it holds

$$d_{Z^m}(x_i, y) = d_T(\bar{h}(x_i), h^\bar{h}(x_i)(y)).$$

Using this fact, the subadditivity of the graph distance and (29) yields that

$$d_T(\bar{h}(x_1), \bar{h}(x_2)) \leq d_{Z^m}(x_1, x_2)$$

$$\leq d_{Z^m}(x_1, y) + d_{Z^m}(y, x_2)$$

$$= d_T(\bar{h}(x_1), h^\bar{h}(x_1)(y)) + d_T(\bar{h}(x_2), h^\bar{h}(x_2)(y)).$$

(32)

Now, let $v \in T$ be the unique vertex on the geodesic path between $\bar{h}(x_1)$ and $w_0$ such that

$$d_T(\bar{h}(x_1), \bar{h}(x_2)) = d_T(\bar{h}(x_1), v) + d_T(v, \bar{h}(x_2)).$$

(33)

Combining (32) and (33) gives

$$d_T(\bar{h}(x_1), v) + d_T(v, \bar{h}(x_2)) \leq d_T(\bar{h}(x_1), h^\bar{h}(x_1)(y)) + d_T(\bar{h}(x_2), h^\bar{h}(x_2)(y)).$$

Thus either

$$d_T(\bar{h}(x_1), v) \leq d_T(\bar{h}(x_1), h^\bar{h}(x_1)(y))$$

(34)
or
\[ d_T(\bar{h}(x_2), v) \leq d_T(\bar{h}(x_2), h_{x_2}(y)). \tag{35} \]

Due to the tree structure, all the vertices \( w \) on one geodesic path between \( \{\bar{h}(x_i)\}_{i=1,2} \) and \( w_0 \) and such that \( d_T(\bar{h}(x_i), v) \leq d_T(\bar{h}(x_i), w) \) must also be on the other geodesic path. Hence due to (34) and (35) at least one of the \( h_{x_2}(y) \) is on both geodesic paths to \( w_0 \). Since the depth is one-to-one on those paths, the only way that \( |h_{x_1}(y)| = |h_{x_2}(y)| \) is if \( h_{x_1}(y) = h_{x_2}(y) \) which deduces (31).

We prove now that \( h = \bar{h} \) on \( S \). For any pair of points \( x_1, x_2 \in S \), we know from the definition of the map \( h_{x_2} \) that
\[ d_T(\bar{h}(x_2), h_{x_2}(x_1)) = d_{\Xi^n}(x_2, x_1). \]

Thus it follows from the definition of \( |w| = \text{dist}_T(w, w_0) \) that
\[
|h_{x_2}(x_1)| = |\bar{h}(x_2)| - d_T(\bar{h}(x_2), h_{x_2}(x_1))
\leq |\bar{h}(x_2)| - (|\bar{h}(x_2)| - |\bar{h}(x_1)|)
\leq |\bar{h}(x_1)|.
\]

This means that, at the point \( x_1 \), the maximum argument in (30) must be reached for \( x_1 \) and thus \( h(x_1) = \bar{h}(x_1) \).

Now, we will show that \( h \) is a graph homomorphism. For this purpose let \( y \sim z \in \Lambda \) be nearest neighbors. We have to show that this implies \( \text{dist}_T(h(y), h(x)) = 1 \). We distinguish two cases. In the first case we assume that there is \( x \in S \) such that
\[ h(y) = h_{x}(y) \quad \text{and} \quad h(z) = h_{x}(z). \]

In this case, the fact that \( \text{dist}_T(h(y), h(z)) = 1 \) directly follows from Lemma 4.4 which states that all the \( h_{x} \) are graph homomorphisms themselves.

Let us now consider the second case where we assume that there exist \( x_1, x_2 \in S, x_1 \neq x_2 \) such that
\[ h(y) = h_{x_1}(y) \quad \text{and} \quad h(z) = h_{x_2}(z). \]

Then, we have to show that
\[ \text{dist}_T(h_{x_1}(y), h_{x_2}(z)) = 1. \]
In this situation we will show below that either
\[ h_{x_1}^\tilde{h}(y) = h_{x_2}^\tilde{h}(y) \quad \text{or} \quad h_{x_1}^\tilde{h}(z) = h_{x_2}^\tilde{h}(z). \]
Indeed, if this statement is true we have reduced the second case to the first case.
Let us now deduce the statement (40). By the definition of the map \( \tilde{h} \) it must hold that
\[ \left| |h_{x_1}^\tilde{h}(z)| - |h_{x_2}^\tilde{h}(y)| \right| \leq 2, \]
else one could easily construct a contradiction. The case
\[ |h_{x_1}^\tilde{h}(z)| = |h_{x_2}^\tilde{h}(y)| \]
cannot happen. Else one would get a contradiction to the fact that \( \tilde{h} \) conserves the parity. Hence it follows that
\[ |h_{x_1}^\tilde{h}(z)| = |h_{x_2}^\tilde{h}(y)| \pm 1. \]
Hence we have that either
\[ |h_{x_1}^\tilde{h}(z)| = |h_{x_2}^\tilde{h}(y)| \]
or
\[ |h_{x_1}^\tilde{h}(y)| = |h_{x_2}^\tilde{h}(y)|. \]
Now, we are in the same situation as in the second step and can conclude from from (31) that (40), which concludes the argument. \( \square \)

4.2. A concentration inequality. The main purpose of this section is to deduce a concentration inequality for the the uniform measure on the set \( H_n^g(s) \), which is defined in Definition 3.7. The concentration inequality is:

**Theorem 4.6** (Concentration inequality). There exists a universal constant \( C \) such that, under the uniform measure \( \mathbb{P}_n^s \) on \( H_n^g(s) \), for all \( x = (x_1, \ldots, x_m) \) in \( \mathbb{Z}^m \), and for all \( \varepsilon \geq n^{-0.45} \) we have

\[ \mathbb{P}_n^s \left( \max_{x \in \mathbb{Z}^m} d_T(h(x), g(\lfloor s \cdot x \rfloor)) \geq \varepsilon n \right) \leq Ce^{-\frac{\varepsilon^2 n}{212}}, \]

where the function \( g(n) \) is given by Lemma 2.7.

We will deduce Theorem 4.6 from an auxiliary concentration inequality.

**Lemma 4.7** (Auxiliary concentration inequality). For all \( \varepsilon > 0 \) and \( x \in \mathbb{Z}^m \) it holds

\[ \mathbb{P}_n^s (||h(x)| - \mathbb{E}[|h(x)||] \geq \varepsilon n) \leq Ce^{-\frac{\varepsilon^2 n}{32}}, \]

where \( |h(x)| \) is given by Definition 4.5.
The proof of Lemma 4.7 is in several steps. In [CEP96] a concentration inequality was deduced for domino tilings of an Aztec diamond. Using this argument as an inspiration, our argument is also based on the Azuma-Hoeffding inequality (see Lemma 4.8 from below). In the setting of graph homomorphisms to a tree, verifying the assumptions of the Azuma-Hoeffding inequality becomes very challenging. For this purpose we developed a completely new argument based on coupling. We state the proof of Lemma 4.7 with all the details in Section 4.2.1 and now continue to state the proof of Theorem 4.6. The only additional ingredient that is needed in the proof of Theorem 4.6 is the fact that every element $h \in \mathcal{H}_n^g(s)$ has to stay close to the geodesic $g$ (see Lemma 3.6).

**Proof of Theorem 4.6.** The goal is to derive the estimate (41), which is verified in several steps. Let $x = (x_1, \ldots, x_m) \in \mathbb{Z}^m$. The auxiliary concentration inequality of Lemma 4.7 states that

$$
\mathbb{P} (||h(x) - \mathbb{E}[|h(x)|]| \geq \epsilon n) \leq C e^{-\frac{2n^2}{3\epsilon^2}},
$$

(42)

where we used the simplified notation $\mathbb{P} = \mathbb{P}_n$. We argue that a combination of the estimate (42) and the fact that the map $h \in \mathcal{H}_n^g(s)$ is $n-$translational invariant yields that

$$
\mathbb{P}(\max_{x \in \mathbb{Z}^m} ||h(x) - \mathbb{E}[|h(x)|]| \geq \epsilon n) \leq C e^{-\frac{2n^2}{3\epsilon^2}}.
$$

(43)

Indeed, due to the fact that $h$ is $n-$translational invariant it holds

$$
\mathbb{P}(\max_{x \in \mathbb{Z}^m} ||h(x) - \mathbb{E}[|h(x)|]| \geq \epsilon n)
\leq \sum_{x \in \{0, n-1\}^m} \mathbb{P} (||h(x) - \mathbb{E}[|h(x)|]| \geq \epsilon n)
\leq C n^m e^{-\frac{2n^2}{3\epsilon^2}}
\leq C e^{-\frac{2n^2}{3\epsilon^2}},
$$

which verifies the estimate (43).

In the next step, we show that the estimate (43) yields

$$
\mathbb{P}(d_T(h(x), \Pi_g(h(x))) \geq \epsilon n) \leq C e^{-\frac{2n^2}{3\epsilon^2}},
$$

(44)

where $\Pi_g : T \rightarrow g$ denotes the projection onto the nearest point in $g$. We start with arguing that the function $k \rightarrow \mathbb{E}[|h(k, 0, \ldots, 0)|]$ is non-decreasing. Since wlog. the slope is non-negative, i.e. $s_1 \geq 0$ it follows
that
\[
\lim_{k \to \infty} |h(k, 0, \ldots, 0)| - |h(k, 0, \ldots, 0)| = \infty \quad \text{a.s.}
\]
By the Fatou's Lemma this yields that
\[
\lim_{k \to \infty} \mathbb{E}[|h(k, 0, \ldots, 0)|] - \mathbb{E}[|h(k, 0, \ldots, 0)|] = \infty
\]
Therefore, there is a number \(k \in \mathbb{N}\) such that
\[
\mathbb{E}[|h(k, 0, \ldots, 0)|] - \mathbb{E}[|h(k, 0, \ldots, 0)|] \geq 0.
\]
Using a telescope sum it follows that
\[
\mathbb{E}[|h(k, 0, \ldots, 0)|] - \mathbb{E}[|h(k, 0, \ldots, 0)|] = \sum_{0 \leq i < k} \mathbb{E}[|h(i + 1, 0, \ldots, 0)|] - \mathbb{E}[|h(i, 0, \ldots, 0)|].
\]
We can deduce from the translation invariance that for all \(i, j \in \mathbb{N}\)
\[
\mathbb{E}[|h(i + 1, 0, \ldots, 0)|] - \mathbb{E}[|h(i, 0, \ldots, 0)|] = \mathbb{E}[|h(j + 1, 0, \ldots, 0)|] - \mathbb{E}[|h(j, 0, \ldots, 0)|].
\]
It follows that
\[
\mathbb{E}[|h(i + 1, 0, \ldots, 0)|] - \mathbb{E}[|h(i, 0, \ldots, 0)|] \geq 0.
\]
Therefore, the function \(k \to \mathbb{E}[|h(k, 0, \ldots, 0)|]\) must be non-decreasing, which yields that for \(l \geq k:\)
\[
\mathbb{E}[|h(l, 0, \ldots, 0)|] - \mathbb{E}[|h(k, 0, \ldots, 0)|] \leq 0. \quad (45)
\]
Now, let \(k\) be such that \(d_T(h(k, 0, \ldots, 0), \Pi_g(h(k, 0, \ldots, 0))) \geq c\). Since all configurations are supported on the geodesic \(g\), we know there exist \(l > k\) such that: \(|h(l, 0, \ldots, 0)| \leq |h(k, 0, \ldots, 0)| - c\). Hence, by using the estimate (45) we get that
\[
c \leq |h(k, 0, \ldots, 0)| - |h(l, 0, \ldots, 0)|
\]
\[
= |h(k, 0, \ldots, 0)| - \mathbb{E}[|h(k, 0, \ldots, 0)|]
\]
\[
+ \mathbb{E}[|h(k, 0, \ldots, 0)|] - \mathbb{E}[|h(l, 0, \ldots, 0)|]
\]
\[
+ \mathbb{E}[|h(l, 0, \ldots, 0)|] - |h(l, \ldots, 0)|
\]
\[
\leq ||h(k, 0, \ldots, 0)| - \mathbb{E}[|h(k, 0, \ldots, 0)|]|
\]
\[
+ ||h(l, 0, \ldots, 0)| - \mathbb{E}[|h(l, 0, \ldots, 0)|]|.
\]
It follows that either
\[
||h(l, 0, \ldots, 0)| - \mathbb{E}[|h(l, 0, \ldots, 0)|]| \geq c/2
\]
or
\[
||h(k, 0, \ldots, 0)| - \mathbb{E}[|h(k, 0, \ldots, 0)|]| \geq c/2.
\]
Combining this with the estimate (43) yields the desired estimate (44).
We observe that the concentration inequality (44) together with the universal bound (see Lemma 3.6)
\[ d_T (h(x), \Pi_g(h(x))) \leq m \frac{n}{2} \]
yields that
\[ \mathbb{E} [d_T (h(x), \Pi_g(h(x)))] \leq C n^{0.51}. \tag{46} \]

In this step, we will deduce the concentration inequality
\[ \mathbb{P} (d_T(h(x), g(s \cdot x)) \geq \varepsilon n) \leq C e^{-\frac{\varepsilon^2 n}{m}}, \tag{47} \]
where the constant $C$ is universal i.e. independent of the choice of $x \in \mathbb{Z}^m$. Indeed, the translation invariance yields that
\[ \mathbb{E} [||\Pi_g(h(x))||] = s \cdot x. \]
Without loss of generality we assume that $0 \leq s \cdot x \in \mathbb{Z}$. Then $g(s \cdot x)$ denotes the unique point $u \in g$ on the geodesic $g \subset T$ such that $|u| = s \cdot x$. It follows that
\[
\begin{align*}
d_T (\Pi_g(h(x)), g(s \cdot x)) &= ||\Pi_g(h(x))| - \mathbb{E} [||\Pi_g(h(x))||]| \\
&\leq |d_T (h(x), \Pi_g(h(x))) + ||\Pi_g(h(x))||| \\
&\quad - \mathbb{E} [d_T (h(x), \Pi_g(h(x)))] - \mathbb{E} [||\Pi_g(h(x))||] \\
&\quad + |d_T (h(x), \Pi_g(h(x))) - \mathbb{E} [d_T (h(x), \Pi_g(h(x)))]| \\
&= ||h(x)| - \mathbb{E} [||h(x)||] \\
&\quad + |d_T (h(x), \Pi_g(h(x))) - \mathbb{E} [d_T (h(x), \Pi_g(h(x)))]|,
\end{align*}
\]
where we used the fact that
\[ |h(x)| = d_T (h(x), \Pi_g(h(x))) + ||\Pi_g(h(x))|| \]
and also that
\[ \mathbb{E} [||h(x)||] = \mathbb{E} [d_T (h(x), \Pi_g(h(x)))] + \mathbb{E} [||\Pi_g(h(x))||]. \]
Overall, using now the bound (46) yields that
\[
\begin{align*}
d_T (\Pi_g(h(x)), g(s \cdot x)) &\leq ||h(x)| - \mathbb{E} [||h(x)||] + d_T (h(x), \Pi_g(h(x))) + C n^{0.51}.
\end{align*}
\]
We observe that due to the fact that $\varepsilon \geq n^{-0.45}$ it holds that for large enough $n$:

$$\frac{\varepsilon}{2} n - Cn^{0.51} \geq \frac{n^{0.55}}{2} - cn^{0.51} \geq 0.$$ 

Hence, if we assume that

$$d_T(\Pi_\gamma(h(x)), \gamma(s \cdot x)) \geq \varepsilon n,$$

it follows that

$$\frac{\varepsilon}{2} n \leq ||h(x)| - \mathbb{E}[|h(x)|] + d_T(h(x), \Pi_\gamma(h(x))).$$

The last estimate yields that either

$$||h(x)| - \mathbb{E}[|h(x)|] \geq \frac{\varepsilon}{4} n$$

or

$$d_T(h(x), \Pi_\gamma(h(x))) \geq \frac{\varepsilon}{4} n.$$ 

Hence, we get from the concentration inequality (42) and (44) that

$$\mathbb{P}(d_T(h(x), \gamma(s \cdot x)) \geq \varepsilon n)$$

$$\leq \mathbb{P}(d_T(h(x), \Pi_\gamma(h(x))) \geq \frac{\varepsilon}{4} n)$$

$$+ \mathbb{P}(||h(x)| - \mathbb{E}[|h(x)|] \geq \frac{\varepsilon}{4} n)$$

$$\leq 2Ce^{-\frac{\varepsilon^2 n}{2}}.$$ 

which is the desired estimate (47).

Now, the estimate (41) follows easily from (47) and the observation that by the translational invariance of $h$ i.e.

$$\mathbb{P}\left(\max_{x \in \mathbb{Z}^m} d_T(h(x), \gamma(s \cdot x)) \geq \varepsilon n\right)$$

$$= \mathbb{P}\left(\max_{x \in S_n} d_T(h(x), \gamma(s \cdot x)) \geq \varepsilon n\right)$$

$$\leq \sum_{x \in S_n} \mathbb{P}(d_T(h(x), \gamma(s \cdot x)) \geq \varepsilon n)$$

$$\leq n^m 2Ce^{-\frac{\varepsilon^2 n}{2}}$$

$$\leq Ce^{-\frac{\varepsilon^2 n}{2}}.$$
4.2.1. **Proof of Lemma 4.7**. We will derive the statement of Lemma 4.7 from the well-known Azuma-Hoeffding inequality.

**Lemma 4.8** (Azuma-Hoeffding [Azu67, Hoe63]). Suppose that \((X_k)_{k \in \mathbb{N}}\) is a martingale and
\[
|X_k - X_{k-1}| < c_k \quad \text{almost surely.}
\]
Then for all \(N \in \mathbb{N}\) and all \(\varepsilon > 0\)
\[
\mathbb{P}(|X_N - X_0| \geq t) \leq 2 \exp \left( - \frac{t^2}{2 \sum_{k=1}^{N} c_k^2} \right).
\]
In order to apply Azuma-Hoeffding we have to specify which martingale \(X_k\) we are considering. For this let us first introduce a filtration of sigma algebras \(\mathcal{F}_k\) we are using. We consider a path \(p : \mathbb{N} \cup \{0\} \to \mathbb{Z}^m\) from \(0 \in \mathbb{Z}^m\) to \(x \in S_n\) i.e. \(p(0) = 0\) and \(p(|x|_H) = x\). We define the functions \(g_k : \mathcal{H}_n^g \to \mathbb{R}\) for \(k = 0\) via
\[
g_0(h) = |h(p(0))|
\]
and for \(k \geq 1\) via
\[
g_k(h) = |h(p(k))| - |h(p(k - 1))|.
\]
Then, the sigma algebras \(\mathcal{F}_k\) are defined via
\[
\mathcal{F}_k = \sigma(g_l, 0 \leq l \leq k).
\]
Now, we define the martingale \(X_k, 0 \leq k \leq |x|_H\) in the usual way using conditional expectations i.e.
\[
X_k = \mathbb{E}[|h(x)||\mathcal{F}_k],
\]
where \(\mathbb{E}\) denotes the expectation under the uniform probability measure on \(\mathcal{H}_n^g(s)\). We note that for \(k = |x|_H\) it holds
\[
X_{|x|_H} = |h(x)|.
\]
As a consequence, the statement of Lemma 4.7 follows directly from Azuma-Hoeffding by choosing \(N = |x|_H\) (see Lemma 4.8), if we can show that almost surely
\[
|X_k - X_{k-1}| \leq 4.
\]
This is exactly the statement of the following lemma.

**Lemma 4.9.** Using the definitions from above, it holds for \(k = 1 \ldots, |x|_H\) that almost surely
\[
|X_k - X_{k-1}| = |\mathbb{E}[|h(x)||\mathcal{F}_k] - \mathbb{E}[|h(x)||\mathcal{F}_{k-1}]| \leq 4. \quad (48)
\]
Hence, we see that in order to deduce Lemma 4.7 it is only left to verify Lemma 4.9. For deducing the estimate (48), we need to show that changing the depth of a single point $y$ does not influence too much the expected depth of the point $x$. The proof of Lemma 4.9 is quite subtle. The reason being that the structure of the uniform probability measure $P$ on the set $H_n(s)$ is extremely hard to break down. Unfortunately, without classical tools related to the integrability of the model (see for example [CEP96]) it seems that there is no direct way to compare $E_s^*\left[|F_{k+1}|\right]$ and $E_s^*\left[|F_k|\right]$.

Instead, we use a dynamic approach. We construct a coupled Markov chain $(X_n, Y_n)$ on $H_n(s)^2$ such that the law of $X_n$ converges to $E_s^*\left[|F_{k+1}|\right]$ and the law of $Y_n$ converges to $E_s^*\left[|F_k|\right]$. The crucial property will be that the Markov chain keeps the depth deviation of $X_n$ and $Y_n$ invariant. This means that if

$$||X_n(x)| - |Y_n(x)|| \leq 4$$

then also

$$||X_{n+1}(x)| - |Y_{n+1}(x)|| \leq 4.$$  \hspace{1cm} (49)

Because this property is verified at each step and the law of the Markov chain converges to $E_s^*\left[|F_{k+1}|\right]$ and $E_s^*\left[|F_k|\right]$, it also holds that

$$|E_s^*\left[|h(x)||F_{k+1}|\right] - E_s^*\left[|h(x)||F_k|\right]| \leq 4,$$

which is the desired statement of Lemma 4.9.

There is a natural choice for the Markov chain sampling the uniform measure on $H_n(s)$, which is the Glauber dynamic (see Definition 4.13). Unfortunately, even when coupled, the Glauber dynamic does not conserve the depth deviation. More precisely, the Glauber dynamic does not have the desired property (49). This can be seen by constructing counter examples involving fake minima (cf. Definition 4.10 and the proof of Lemma 4.9 from below). If one would consider graph homomorphisms to $\mathbb{Z}$ then this technical problem would not appear and one could use the classical Glauber dynamic.

We overcome this technical difficulty in the following way: We carefully analyze the situations in which the depth deviation can increase under the Glauber dynamic. Then we add a resampling step before every Glauber step to prevent those situations. When adding the resampling step one has to be very careful not to introduce a bias. We will
show that this is not the case. More precisely, we show that the modified Glauber dynamic still converges to the uniform measure on $\mathcal{H}_n(s)$ given $\mathcal{F}_k$ and $\mathcal{F}_{k-1}$.

We begin with introducing some necessary definitions.

**Definition 4.10** (Minimum and extremum of a graph-homomorphism). Let $h : \mathbb{Z}^m \to \mathcal{T}$ be a graph-homomorphism. We say that a point $x \in \mathbb{Z}^m$ is minimum for $h$ if

$$|h(x)| < |h(z)| \text{ for all } z \sim x.$$ 

If additionally

$$h(z) = h(\tilde{z}) \text{ for all } z, \tilde{z} \sim x$$

we say that $x$ is a true minimum, otherwise we say that $x$ is a fake minimum for $h$. Finally, if

$$h(z) = h(\tilde{z}) \text{ for all } z, \tilde{z} \sim x$$

holds without the first condition we say that $x$ is an extremum of $h$.

For an illustration of Definition 4.10 we refer to Figure 10.

**Remark 4.11.** We want to observe that there is no fake local maximum for a graph homomorphism $h$. The reason is that for a given point $h(x) \in \mathcal{T}$ there is only one neighbor $h(x) \sim v \in \mathcal{T}$ with lower depth.

**Definition 4.12.** (Pivoting) If $x$ is an extrema of $h$, we call pivoting $x$ the following operation:

- For $z \sim x$ choose $w \in \mathcal{T}$ randomly and with equal probability among the neighbors of $h(z) \in \mathcal{T}$.
- Set $h(x) = w$. 

![Figure 10. Extremum of a graph homomorphism](image)
In Figure 11 we illustrate what pivoting means for a graph homomorphism to a tree. With the notion of pivoting, we can define the natural Glauber dynamic on $\mathcal{H}_n^R(s)$.

**Definition 4.13. (Glauber dynamic)** The Glauber dynamic on $\mathcal{H}_n^R(s)$ is a Markov chain given by the following procedure: Let $X_n \in \mathcal{H}_n^R(s)$ then then $X_{n+1}$ is attained via:

1. Choose a vertex $x \in S_n \subset \mathbb{Z}^m$ uniformly at random.
2. If $X_n(x)$ is an extremum in the sense of Definition 4.10 then pivot $X_n$ around $x$. Else do nothing.

As we mentioned above, it is possible that the original Glauber dynamics increases the depth deviation. This can only happen if one configuration has a fake minimum and the other configuration does not. The purpose of the adapted Glauber dynamics is to eliminate this situation. In order to define the adapted Glauber dynamic, let us first introduce the concepts of excursion and resampling an excursion.
Definition 4.14 (Excursion). Let $h : \mathbb{Z}^m \to \mathcal{T}$ be a translation invariant homomorphism and $\mathcal{T}_0$ be a complete branch of $\mathcal{T}$ such that the vertex of $\mathcal{T}_0$ with lowest depth has degree one (see Figure 12). This vertex is called the root of the branch $\mathcal{T}_0$. We say that an excursion of $h$ is a connected component $\mathcal{C}$ of $h^{-1}(\mathcal{T}_0)$ on which the depth is bounded from above. If $x \in \mathbb{Z}^m$ is mapped by $h(x)$ onto the root of $\mathcal{T}_0$ then we say that the excursion starts at $x$.

We call $\mathcal{C}$ excursion because for any element $x \in \mathcal{C}$ and path $p \subset \mathbb{Z}^m$ from $x$ to infinity, there must be an element $z \in p$ such that $h(z)$ is mapped onto the root of the branch $\mathcal{T}_0$ (see Figure 13).

Definition 4.15 (Excursion resampling). Let $h : \mathbb{Z}^m \to \mathcal{T}$ be a graph homomorphism and $\mathcal{C}$ be an excursion of $h$ starting in $x$. We call the following operation resampling the excursion $\mathcal{C}$:

1. Observe that there is an index $j \in \{1, \ldots, d\}$ such that $|h(x) + \alpha_i| < |h(x)|$.
2. Choose an index $i \in \{1, \ldots, d\} \setminus \{j\}$ randomly with equal probability.
3. We define a new dual graph homomorphism $\tilde{g}$ by:
   $$\tilde{g}(e_{xy}) := \begin{cases} \alpha_i & \text{if } e_{xy} \text{ on the outer boundary of } \mathcal{C}, \\ \hat{h}(e_{xy}) & \text{else}. \end{cases}$$
4. We set the new graph homomorphism $h$ to be the graph homomorphism that is naturally associated to $\tilde{g}$.

For an illustration of resampling an excursion we refer to Figure 14 and Figure 15. One could ask why one does not choose the index $i$ uniformly at random out of the set $\{1, \ldots, d\}$. The reason for choosing the index $i$ out of the set $\{1, \ldots, d\} \setminus \{j\}$ is that by this procedure one guarantees that resampling an excursion does not change the depth profile of the configuration $h$.

For our adapted Glauber dynamic it is important to decide if an edge $e_{xy}$ of a fake minimum at $x$ is in an excursion $\mathcal{C}$ or not. The next lemma helps a lot in that task.

Lemma 4.16. Let $x \in \Lambda$ be a fake minimum for the homomorphism $h$ and let $e_{xy}$ be an edge starting from $x$. The edge $e_{xy}$ is not in an excursion $\mathcal{C}$ that starts at $x$ if and only if there exist an infinite path $p = \{x_0 = x, x_1 = y, \ldots\}$ and such that:

1. For all $i \geq 1 : |h(x_i)| > |h(x)|$
2. $\sup_{i \geq 1} |h(x_i)| = \infty$
Moreover, two edges $e_{xy}$ and $e_{xy}'$ are in a common excursion of $h$ starting at $x$ if there exist a path $\mathbf{p} = \{x_0 = x, \ldots, x_n = x\}$ whose first and last edges are $e_{xy}$ and $e_{y'x}$ and such that for all $1 < i < n : |h(x_i)| > |h(x)|$.

**Proof of Lemma 4.16.** Suppose that $e_{xy}$ is not in an excursion, and consider the connected component $C$ of $h^{-1}(T_0)$ which contains $y$ where $T_0$ is the complete tree with root $h(y)$. By definition $e_{xy}$ is not in an excursion if and only the depth is not bounded from above on $C$. Consider a sequence $\{x_1, \ldots, x_n, \ldots\}$ in $C^\mathbb{N}$ such that for all $i \geq \mathbb{N} : |h(x_i)| = |h(x)| + 1$, and build a path $\mathbf{p}$ that goes through all the $x_i$’s while staying in $C$ (this is always possible since $C$ is connected), then $\mathbf{p}$ verifies the conditions of the lemma. Reciprocally, if $e_{xy}$ is in an excursion for $h$ then
the depth is bounded from above on $C$ and any path $p$ on which the depth is not bounded from above must leave $C$ at some point. □

The resampling-step will be used in our construction of the adapted Glauber dynamic. It is important to show that resampling maps an element from $\mathcal{H}_n^g(s)$ onto an element in $\mathcal{H}_n^g(s)$. This is a direct consequence of the following statement.

**Lemma 4.17.** Assume that $h \in \mathcal{H}_n^g(s)$. Then it holds that an edge $e_{xy}$ is in an excursion if and only if for all $1 \leq k \leq m$ the edge $e_{x+n\vec{n}_k y+n\vec{n}_k}$ is also in an excursion. Additionally, it holds that $e_{xy}$ and $e_{x'y'}$ are on the border of a common excursion if and only if for all $1 \leq k \leq m$ the edges $e_{(x+n\vec{n}_k)(y+n\vec{n}_k)}$ and $e_{(x'+n\vec{n}_k)(y'+n\vec{n}_k)}$ are also on the border of a common excursion.

The proof of Lemma 4.17 is based on Lemma 4.16.

**Proof of Lemma 4.17.** Both claims follow from the fact that translating by $n\vec{n}_k$ leaves the depth difference invariant since the geodesic $g$ which support the configuration is left unchanged by translation. Hence, if there exist an infinite path $p = \{x_0 = x, x_1 = y, ...\}$ such that for all $i \geq 1 : |h(x_i)| > |h(x)|$ and $\sup_{i \geq 1} |h(x_i)| = \infty$. Then the path $p + n\vec{n}_k = \{x_0 = x + n\vec{n}_k, x_1 = y + n\vec{n}_k, ...\}$ also verifies for all $i \geq 1 : |h(x_i)| > |h(x)|$ and $\sup_{i \geq 1} |h(x_i)| = \infty$.

Similarly two edges $e_{xy}$ and $e_{x'y'}$ are on the border of a common excursion if there exist a path $p = \{x_0 = x, x_1 = y, ..., x_{n-1} = y', x_n = x\}$ such that for all $1 < i < n : |h(x_i)| > |h(x)|$ which is true if and only if the path $p + n\vec{n}_k = \{x_0 = x + n\vec{n}_k, x_1 = y + n\vec{n}_k, ..., x_{n-1} = y' + n\vec{n}_k, x_n = x + n\vec{n}_k\}$ also verifies for all $1 < i < n : |h(x_i + n\vec{n}_k)| > |h(x + n\vec{n}_k)|$. □
Now, let us describe the adapted Glauber dynamic that is used in the proof of Lemma 4.19.

**Definition 4.18 (Adapted Glauber dynamic).** Let \((c_1, \ldots, c_k) \in \mathbb{Z}^k\) and denote by \(\mathcal{H}_n^g(s)[c_1, \ldots, c_k]\) to be the set of homomorphisms \(h \in \mathcal{H}_n^g(s)\) such that \(|h(x_i)| - |h(0)| = c_i\) for \(1 \leq i \leq k\). We consider the following dynamic on \(\mathcal{H}_n^g(s)[c_1, \ldots, c_k]\). Let \(X_n \in \mathcal{H}_n^g(s)[c_1, \ldots, c_k]\). Then the new configuration \(X_{n+1} \in \mathcal{H}_n^g(s)[c_1, \ldots, c_k]\) is obtained in the following way:

1. Choose a vertex \(x \in S_n\).
2. If \(|X_n(x)| - |X_n(0)|\) is not fixed then resample all the excursions \(C\) of \(X_n\) that start in \(x\). If \(x\) becomes an extremum, pivot \(x\) after that and resample again all the excursions \(C\) of \(X_n\) that start in \(x\).
3. If \(|X_n(x)| - |X_n(0)|\) is fixed (through \(\mathcal{F}_k\)) then only resample all the excursions \(C\) of \(X_n\) that start in \(x\) but do not pivot \(x\).

In the next lemma, we show that the adapted Glauber dynamic converges to the correct law.

**Lemma 4.19.** The Markov chain \(\{X_n\}_{n \in \mathbb{N}}\) given by Definition 4.18 is reversible and irreducible on the state space \(\mathcal{H}_n^g(s)[c_1, \ldots, c_k]\). As a direct consequence, the law of \(\{X_n\}_{n \in \mathbb{N}}\) converges to the uniform measure on \(\mathcal{H}_n^g(s)[c_1, \ldots, c_k]\).

**Proof of Lemma 4.19.** We start with observing that the adapted Glauber dynamic leaves \(\mathcal{H}_n^g(s)[c_1, \ldots, c_k]\) invariant. Resampling excursions around a vertex \(x\) does not change the depth of \(x\) and the dynamic only pivot vertices whose depth is not fixed.

We now show that the Markov chain is reversible wrt. the uniform probability measure on \(\mathcal{H}_n^g(s)[c_1, \ldots, c_k]\). Because we consider the uniform probability measure on \(\mathcal{H}_n^g(s)[c_1, \ldots, c_k]\) it suffices to show that

\[
P[X_n, X_{n+1}] = P[X_{n+1}, X_n].
\]

Let \(Z_n \in S_n\) denote the position that is chosen in the first step of the adapted Glauber dynamics. We notice that after conditioning on \(Z_n\) both the law given of resampling an excursion and pivoting are uniform over the spaces of reachable configurations. Thus we have that

\[
P[X_n, X_{n+1}] = \sum_{x \in S_n} P[X_n, X_{n+1} \mid Z_n = x] \frac{1}{n^2}
= \sum_{x \in S_n} P[X_{n+1}, X_n \mid Z_{n+1} = x] \frac{1}{n^2}
= P[X_{n+1}, X_n].
\]
This shows that the adapted Glauber dynamics is reversible on the state space $H^g_n(s)[c_1, ..., c_k]$.

Now, we prove that the chain is irreducible. The main idea is the following. In every step, there is a positive probability that the resampling excursions does not change the configuration $h \in H^g_n(s)[c_1, ..., c_k]$. Therefore, the adapted Glauber dynamics is irreducible if the original Glauber dynamic is irreducible. To show the irreducibility of the original Glauber dynamic, one shows in the first step that there is a positive probability that the original Glauber dynamics transforms $h \in H^g_n(s)[c_1, ..., c_k]$ after finitely many steps to a configuration $g \in H^g_n(s)[c_1, ..., c_k]$ that is only supported on the geodesic $g$. Now, one can only consider moves of the original Glauber dynamics that do not move $g$ away from $g$. This means that the question of irreducibility of the adapted Glauber dynamics on $H^g_n(s)[c_1, ..., c_k]$ has been reduced to the question if the original Glauber dynamics is irreducible for graph homomorphisms to $\mathbb{Z}$, which we show now.

We define the distance between two homomorphisms $h$ and $h'$ supported on $g$ by

$$d(h, h') = \sum_{x \in S_n} ||h(x)| - |h'(x)||.$$

It is clear that $d(h, h') = 0$ implies $h$ equal $h'$ and that $d(h, h')$ must be finite since $S_n$ is finite. Hence, if we show that there exists always a pivot move which decreases the distance between two configurations, we can use inductive argument to prove that the chain is irreducible. Let $C$ be the cluster on which $||h(x)| - |h'(x)||$ is maximal and let $x_0$ be a point in $C$. Suppose without restriction than $h$ is greater than $h'$ on $C$ and define $p = \{x_0, ..., x_q\}$ to be a maximal increasing path starting form $x_0$. That is a path of maximal length such that for all $0 \leq i < q : |h(x_{i+1})| > |h(x_i)|$. This path has to be finite otherwise there exist two points $x_i$ and $x_j$ along the path $p$ which are $n$-translates of each other and the slope verifies $s \cdot x = 1$. All points on the path have to be in $C$ since for all $0 \leq i < q$:

$$h'(x_{i+1}) - h'(x_i) \leq h(x_{i+1}) - h(x_i)$$

Moreover, the fact that $p$ is of maximal length impose that $h(x_q)$ is a local maximum for the depth. Hence it is possible to do a pivoting move at $x_q$ which decrease the depth of $x_q$ by 2 and decrease the distance between $h$ and $h'$.

Because the Markov chain has a finite state space and is reversible, it follows from standard theory of Markov processes that the law of the
chain converges to the unique invariant probability measure (see for example [Dur10]). □

Now, we are ready to state the proof of Lemma 4.9.

Proof of Lemma 4.9. We consider a Markov chain $X_l$ on the state space $\mathcal{H}^g_n(s)[c_1, \ldots, c_{k+1}]$ given by the adapted Glauber dynamic of Definition 4.18. Let $Y_l$ denote the Markov chain on the state space $\mathcal{H}^g_n(s)[c_1, \ldots, c_k]$ that is also given by the adapted Glauber dynamic of Definition 4.18. As outlined before, the strategy is to define a coupling $(X_l, Y_l)$ of those Markov chains such that

$$\max_{y \in \mathbb{Z}^m} |X_l(y)| - |Y_l(y)| \leq 2 \Rightarrow \max_{y \in \mathbb{Z}^m} |X_{l+1}(y)| - |Y_{l+1}(y)| \leq 2. \quad (50)$$

We postpone the verification of (50) and show how it is used to derive the statement of Lemma 4.9. We pick $h \in \mathcal{H}^g_n(s)[c_1, \ldots, c_{k+1}]$ arbitrary and set $X_0 = h$. Then by the Kirszbraun theorem (cf. Theorem 4.1) there exists an element $\tilde{h} \in \mathcal{H}^g_n(s)[c_1, \ldots, c_k]$ such that

$$\max_{y \in \mathbb{Z}^m} |h(y)| - |\tilde{h}(y)| \leq 2.$$ 

Hence, we set $Y_0 = \tilde{h}$ and we get from (50) that for any realization of the Markov chain and all $l \in \mathbb{N}$

$$\max_{y \in \mathbb{Z}^m} |X_l(y)| - |Y_l(y)| \leq 2.$$ 

The last estimate implies

$$|\mathbb{E} [|X_l(x)|] - \mathbb{E} [|Y_l(x)|]| \leq \mathbb{E} [||X_l(x)| - |Y_l(x)||] \leq 2.$$ 

Now, Lemma 4.8 yields that

$$\lim_{l \to \infty} \mathbb{E} [|X_l(x)|] = \mathbb{E}_n^s [|h(x)| |\mathcal{F}_{k+1}]$$

and

$$\lim_{n \to \infty} \mathbb{E} [|Y_l(x)|] = \mathbb{E}_n^s [|h(x)| |\mathcal{F}_k].$$

Hence, we overall get the desired estimate (48)

$$|\mathbb{E}_n^{(s_1, \ldots, s_m)} [|h(x)| |\mathcal{F}_{k+1}] - \mathbb{E}_n^{(s_1, \ldots, s_m)} [|h(x)| |\mathcal{F}_k]| \leq 2.$$ 

The only thing left to show is that such a coupling $(X_n, Y_n)$ indeed exists. Let us consider an element $y \in \mathbb{Z}^m$ such that

$$|X_n(y)| = |Y_n(y)|.$$
A VARIATIONAL PRINCIPLE FOR A NON-INTEGRABLE MODEL

Then every coupling of the chain $X_n$ and $Y_n$ works because the adapted Glauber dynamic can only increase the depth deviation in one time step by 2.

Therefore, let us now consider an element $y \in \mathbb{Z}^m$ such that

$$||X_n(y)| - |Y_n(y)|| = 2.$$ 

Without loss of generality we can assume that (else we just interchange the role of $X_n$ and $Y_n$ in the argument)

$$|X_n(y)| \geq |Y_n(y)| + 2.$$ 

In this situation, there is only one scenario in which the original Glauber dynamic would increases the depth deviation to 4, more precisely, such that

$$||X_{n+1}(y)| - |Y_{n+1}(y)|| = 4.$$ 

The scenario is when $y$ is a true local minima for $X_n$ and a fake local minimum for $Y_n$ (recall that fake local maxima cannot exist for tree-valued height functions, cf. Remark 4.11). The depth deviation could now increase if the Glauber dynamic selects the site $y$. Then the Glauber dynamic could pivot $X_n(y)$ but not $Y_n(y)$, which could possibly increase the depth deviation to 4. However, using the excursion resampling we can show that there is a coupling that prevents this scenario from happening.

Let us explain this strategy in more details. First of all, we want to mention that resampling an excursion does not change the depth of a configuration. Hence, the additional resampling steps of our adapted Glauber dynamic cannot create a violation of the desired conclusion (50).

Down below, we construct a coupling such that

$\{y$ is a true local minima of $L_yX_l\} \Rightarrow \{y$ is a true local minimum of $L_yY_l\}$,

where $L_yX_l \in \mathcal{H}_n^g(s)$ and $L_yY_l \in \mathcal{H}_n^g(s)$ denote the states obtained after resampling the excursions of $h$ and $\hat{h}$ starting at $y$. The next step of our Markov chain $(X_n, Y_n)$ is pivoting $L_yX_l$ and $L_yY_l$ around $y$. Now, this step can easily be coupled such that if $|L_yX_l(y)|$ increases or decreases, then also does $|L_yY_l(y)|$. This implies the desired conclusion (50).

We explain now how to construct a coupling that satisfies the statement (51). The auxiliary Lemma 4.20 from below states that

$$P(X_l, y) \leq P(Y_l, y),$$

(52)
where \( P(X_l, y) \) is the probability that \( y \) is a true minimum of \( X_l \) after resampling of the excursions starting at \( x \). The quantity \( P(Y_l, y) \) is defined analogously. It is a direct consequence of (52) is that there is a coupling of the resampling step \( L_y \) such that (51) is satisfied: One throws a random variable \( U \) that is uniformly distributed on \([0, 1]\). If \( U \leq P(X_l, y) \) one decides that both \( L_y X_l \) and \( L_y Y_l \) will have a true local minimum around \( y \), and \( L_y X_l \) and \( L_y Y_l \) are chosen uniformly among those states. If \( P(X_l, y) \leq U \leq P(Y_l, y) \) one decides that only \( L_y Y_l \) will have a true local minimum around \( y \), but not \( L_y X_l \). And finally if \( P(Y_l, y) \leq U \leq 1 \), one decides that both \( L_y X_l \) and \( L_y Y_l \) will not have a true local minimum around \( y \). This completes the argument.

□

In the proof of Lemma 4.9 we used the following auxiliary statement.

**Lemma 4.20.** Let \( h_u, h_d \in \mathcal{H}_n^g(s) \) such that

\[
\max_{y \in \mathbb{Z}^m} ||h_u(y)| - |h_d(y)|| \leq 2.
\]

Let \( x \in \mathbb{Z}^m \) be a local minimum for both \( h_u \) and \( h_d \) such that

\[
|h_u(x)| = |h_d(x)| + 2.
\]

Let \( L_x h_u \in \mathcal{H}_n^g(s) \) and \( L_x h_d \in \mathcal{H}_n^g(s) \) denote the states one obtains after the excursions of \( h_u \) and \( h_d \) starting at \( x \). Then it holds that

\[
P(h_u, x) \leq P(h_d, x),
\]

where \( P(h_u, x) \) and \( P(h_d, x) \) denote the probability that \( L_x h_u \) and \( L_x h_d \) respectively have a true local minimum at \( x \).

**Proof of Lemma 4.20.** We split up the proof of Lemma 4.20 in several steps. In the first step, we show that the following inequality which is an important consequence of Lemma 4.16. For every different excursion of a graph homomorphism \( h \in \mathcal{H}_n^g(s) \) starting at \( x \), let us choose one edge \( e_{xz} \) contained in that excursion. We denote by \( E(h, x) \) the set containing those edges. It follows that \( |E(h, x)| \) equals the the number of different excursions around \( x \) for the homomorphism \( h \) (see Picture 13 for an example). We want to show that

\[
|E(h_d, x)| \leq |E(h_u, x)|.
\]

In order to prove this, it is sufficient to show the following statements. The first statement is:

If an edge \( e_{xy} \) is not in an excursion of \( h_u \) starting at \( x \),

\[
\text{then } e_{xy} \text{ is also not in an excursion of } h_d \text{ starting at } x.
\]
The second statement is: Assume that the edges \( \{e_{xy_1}, \ldots, e_{xy_k}\} \) are the same excursion for \( h_u \) around \( x \). Then either all the edges
\[
\{e_{xy_1}, \ldots, e_{xy_k}\}
\]
are in a common excursion for \( h_d \) around \( x \) or
\[
\text{no edge in } \{e_{xy_1}, \ldots, e_{xy_k}\} \text{ is in an excursion for } h_d \text{ starting at } x.
\]
Indeed, by the contraposition of the statement (55), the edges \( e_{xz} \in E(h_d, x) \) are in an excursion of \( h_u \) starting at \( x \). By the contraposition of the statement (56), it follows that any two edges \( e_{xz}, e_{xz'} \in E(h_d, x) \), \( e_{xz} \neq e_{xz'} \), are in different excursions of \( h_u \). This implies the desired estimate (54).

Let us now show the statement (55). Let \( e_{xy} \) be not in an excursion for \( h_u \) that starts in \( x \). We know from Lemma 4.16 that there exists an infinite path \( p = \{x_0 = x, x_1 = y, \ldots\} \) such that \( |h_u(x_i)| > |h_u(x)| \) for all \( i \in \mathbb{N} \). Since \( |h_u(x)| = |h_d(x)| + 2 \) and
\[
\max_{x \in \mathbb{Z}^m} |h_u(x)| = |h_d(x)| = 2,
\]
this also imposes that for all \( |h_d(x_i)| > |h_d(x)| \) for all \( i \in \mathbb{N} \). This means that \( e_{xy} \) cannot be in an excursion for \( h_d \).

Now, let us show (56). We show the statement for only two edges \( e_{xy_1} \) and \( e_{xy_2} \). The generalization to arbitrary many edges \( \{e_{xy_1}, \ldots, e_{xy_k}\} \) is straightforward. It follows from Lemma 4.16 that if two edges \( e_{xy} \) and \( e_{xy'} \) are in the same excursion for \( h_u \) then there exists a path \( p = \{x_0 = x, x_1 = y, \ldots, x_{n-1} = y', x_n = x\} \) such that \( h_u(x_i) > h_u(x) \) for all \( 1 \leq i \leq n - 1 \). For the same reason as in the previous paragraph, this imposes that \( |h_d(x_i)| > |h_d(x)| \) for all \( 1 \leq i \leq n - 1 \). This yields that either \( e_{xy} \) and \( e_{xy'} \) are in the same excursion for \( h_d \) or both are not in an excursion for \( h_d \).

In the next step, let us show the following statement. Assume that \( x \) is a local minimum for a graph homomorphism \( h \in \mathcal{H}_{n}^g(s) \) and two edges \( e_{xy} \) and \( e_{xy'} \) are not in an excursion of \( h \) starting at \( x \). Then
\[
\tilde{h}(e_{xy}) = \tilde{h}(e_{xy'}),
\]
where \( \tilde{h} \) denotes the dual of the graph-homomorphism \( h \) (cf. Definition 3.3). Indeed, if an edge is not in an excursion then it is contained in a path to a boundary point of \( \mathcal{T} \). Because \( h \in \mathcal{H}_{n}^g(s) \) is supported on one geodesic there can only be one such path. We observe that \( x \) is a local minimum for \( h \). Hence, \( \tilde{h}(e_{xy}) \) has to increase the depth of \( |h(x)| \). This means that one has to move forward on the geodesic \( g \) and
therefore there is only one choice left for \( \tilde{h}(e_{x,y}) \). This verifies (57).

In the next step, we deduce the following formula for an arbitrary element \( h \in H^s_n(s) \):

\[
P(h, x) = \begin{cases} 
\frac{1}{(d-1)|E(h, x)|-1}, & \text{if all edges around } x \text{ are in an excursion of } h \text{ starting at } x, \\
\frac{1}{(d-1)|E(h, x)|}, & \text{if there is an edge around } x \text{ that is not in an excursion of } h \text{ starting at } x.
\end{cases}
\]

(58)

Indeed, let us first consider the case in which all edges around \( x \) are in an excursion of \( h \) starting at \( x \). The resampling step means that each excursion of \( h \) starting in \( x \) will be attached to an uniformly chosen direction that increases the depth (see Definition 4.15 and Figure 14 and Figure 15). Note that in a d-regular tree there are \( d-1 \)-many such directions. We recall that \( P(h, x) \) denotes the probability that \( h(x) \) becomes a true local minimum after resampling the excursions starting at \( x \) (see Definition 4.10). To get a local minimum all excursions must head into the same direction. Hence, for the first excursion one can choose any direction, but all other excursions must head into the same direction, which yields the desired formula

\[
P(h, x) = \frac{1}{(d-1)|E(h, x)|-1}.
\]

Now, let us consider the second case in which there is an edge around \( x \) that is not in an excursion of \( h \). It follows from the statement (57) that the graph homomorphism \( h \) heads in the same direction for all the edges around \( x \) that are not in an excursion. Hence, in order that \( h \) becomes a true local minimum after resampling the excursions around \( x \), all the excursions have to head into the same direction. This yields the desired formula

\[
P(h, x) = \frac{1}{(d-1)|E(h, x)|}.
\]

Finally, using the estimate (54) and the formula (58) we verify the desired estimate (53) by considering several cases:

In the first case, let us assume that there is an edge around \( x \) that is not in an excursion for \( h_u(x) \) starting at \( x \). It follows from the statement (55) that there is also an edge that is not in an excursion for \( h_d(x) \). Hence, it follows from a combination of (54) and (58) that

\[
P(h_u, x) = \frac{1}{(d-1)|E(h_u)|} \leq \frac{1}{(d-1)|E(h_d)|} = P(h_d, x).
\]
Let us consider the second case in which all edges around \( x \) for \( h_u(x) \) are in an excursion for \( h_u \). We make a further distinction and additionally assume that all the edges around \( x \) are also in an excursion for \( h_d(x) \). In this case, a combination of (54) and (58) yields that

\[
P(h_u, x) = \frac{1}{(d - 1)|E(h_u)| - 1} \leq \frac{1}{(d - 1)|E(h_d)| - 1} = P(h_d, x).
\]

Let us now consider the last case in which we assume that all edges around \( x \) are in an excursion for \( h_u(x) \) but there is an edge around \( x \) that is not in excursion for \( h_d(x) \). In this case it we will show that

\[
|E(h_d, x)| + 1 \leq |E(h_u, x)|.
\]

Postponing the verification of (59) we get by using (59), (54) and (58) that

\[
P(h_u, x) = \frac{1}{(d - 1)|E(h_u)| - 1} \leq \frac{1}{(d - 1)|E(h_d)|} \leq P(h_d, x),
\]

which closes the argument.

The only step remaining is to prove (59). For each excursion \( C^d_j \) of \( h_d \) starting at \( x \) let us choose one representative edge \( e_{xz_j} \in C^d_j \) that starts in \( x \). By using the statement (56), we know that the edges \( e_{xz_1}, \ldots, e_{xz_k} \) must be in different excursions \( C^u_i \) of \( h_u \). Hence, re-indexing allows us to assume that \( e_{xz_j} \in C^d_j \) for \( j \in \{1, \ldots, l\} \). By assumption, there must be an edge \( e_{xz_{l+1}} \) that is not in an excursion of \( h_d \). Because by assumption all edges near \( x \) are in an excursion of \( h_u \) it follows that there is an \( i \in \{1, \ldots, k\} \) such that \( e_{xz_{l+1}} \in C^u_i \). We show in a moment that for all \( j \in \{1, \ldots, l\} \)

\[
C^u_i \neq C^u_j.
\]

This means that the graph homomorphism \( h_u \) has at least \( l + 1 \) many different excursions that start at \( x \), which verifies (59).

Let us turn to the verification of (60). We use an indirect argument and assume that wlog.

\[
C^u_i = C^u_j.
\]

Hence, the edge \( e_{xz_1} \) and \( e_{xz_{l+1}} \) are in the same excursion \( C^u_i \). By the statement (56) this implies that either \( \{e_{xz_1}, e_{xz_{l+1}}\} \subset C^d_1 \) or both edges \( e_{xz_1} \) and \( e_{xz_{l+1}} \) are not in an excursion. Which is a contradiction to the fact that by construction

\[
e_{xz_1} \in C^d_1 \quad \text{and} \quad e_{xz_{l+1}} \notin C^d_1.
\]

□
5. Existence of a Continuum of Shift-Invariant Ergodic Gradient Gibbs Measures

This section is independent of the variational principle (cf. Theorem 2.9 and Theorem 2.11) and of its own interest. The Kirszbraun theorem and the concentration estimate obtained in Section 4 carry important information about the set \( exG(Z^m, T_d) \) of gradient Gibbs measures that are ergodic wrt. the translations of \( Z^m \). In this section, we show the existence of a continuum of translation-invariant, ergodic, gradient Gibbs measures. In order to make our statement precise, we begin with recalling some classical results about Gibbs measures for discrete systems. They can all be found in [Geo88].

**Definition 5.1** (Slope). Let \( \nu \) be a gradient Gibbs measure that is ergodic wrt. translations of \( Z^m \). Then for all \( 1 \leq i \leq m \) the limit

\[
    s_i(\nu) = \lim_{n \to \infty} \frac{1}{n} d_T(h(0), h(ne_i))
\]

exists \( \nu \)-almost surely and we call \((s_1(\nu), ..., s_m(\nu))\) the slope of \( \nu \).

It is a classical result (e.g. [Cha16]) that the subadditive ergodic theorem implies the existence of this limit. Moreover, every translation invariant gradient Gibbs measure can be decomposed into a mixture of ergodic gradient Gibbs measures. The latter allows to define the slope of a translation invariant Gibbs measure in the following way:

**Definition 5.2.** Let \( \mu \) be a translation invariant Gibbs measure. Then the slope \( s(\mu) = (s_1(\mu), ..., s_m(\mu)) \) of \( \mu \) is

\[
    s_i(\mu) = \int_{exG(Z^m, T_d)} s_i(\nu) w_\mu(\nu) \, d\nu,
\]

where \( w_\mu \) is the ergodic decomposition of \( \mu \). This means that for any test function \( f \) it holds that

\[
    \int f(x) \mu(dx) = \int_{exG(Z^m, T_d)} \int f(x) \nu(dx) w_\mu(\nu). \]

We will now show how the concentration results (41) obtained in Section 4 imply the existence of an ergodic gradient Gibbs measure for each slope \( s \) whose \( l_\infty \)-norm is less or equal to 1.

**Lemma 5.3.** Let \( s \in \mathbb{R}^m \) be such that \( |s|_\infty \leq 1 \). Then there exist a sequence of \( n \)-translation invariant homomorphisms \( \{h_n\}_{n \in \mathbb{N}} \) with slope \( \left( \frac{|s_1 n|}{n}, ..., \frac{|s_m n|}{n} \right) \) (cf. Definition 3.4).
Proof. The proof is a direct consequence of the Kirszbraun theorem. For all \((k, l) \in \mathbb{Z}^m\), set \(h(k_1 n, \ldots, k_m n) = g(k_1 \lfloor s_1 n \rfloor + \ldots + k_m \lfloor s_m n \rfloor)\). Since \(g\) is isomorphic to \(\mathbb{Z}\) we can apply the Kirszbraun theorem between \(\mathbb{Z}^m\) and \(g\). Hence, there exist an extension of \(h\) on the whole space \(\mathbb{Z}^m\) which is entirely supported on the geodesic \(g\). The slope of this extension must be \((\lfloor s_1 n \rfloor, \ldots, \lfloor s_m n \rfloor)\) which concludes our proof. \(\square\)

**Theorem 5.4.** For all \(s \in \mathbb{R}^m\) such that \(|s|_\infty \leq 1\), there exist an ergodic gradient Gibbs measure \(\mu\) with slope \(s\).

Proof. We already know from the previous lemma that for all \(n \in \mathbb{N}\) the set \(\mathcal{H}_n^g(\lfloor s_1 n \rfloor, \ldots, \lfloor s_m n \rfloor)\) is non empty. Hence, the uniform probability measure \(\mu_n(s)\) on \(\mathcal{H}_n^g(\lfloor s_1 n \rfloor, \ldots, \lfloor s_m n \rfloor)\) exists. By compactness of the space of gradient Gibbs measures in the topology of local convergence, we know that there exist a subsequence \(\mu_{n_k}(s)\) which converges to a gradient Gibbs measure \(\mu\) on \(\mathbb{Z}^m\) (see Lemma 8.2.7 of [She05] for more details). The shift invariance of the \(\mu\) follows from the shift invariance of the spaces \(\mathcal{H}_n^g(\lfloor s_1 n \rfloor, \ldots, \lfloor s_m n \rfloor)\). This means that we can write the measure \(\mu\) as a mixture of ergodic Gibbs measures (see Definition 5.2). Suppose, by contradiction that the slope is not almost surely equal to \(s\) on the average \(w_\mu\). Then there exist \(i \in \{1..m\}\), \(\delta\) and \(\epsilon\) such that \(w_\mu(|s_i(\nu) - s_i| \geq \epsilon) \geq \delta\). We know from the proof of Theorem 4.6 that for all \(\epsilon > 0\)

\[
\mathbb{P}_\mu \left( |d_T(h(0), h(n)) - \mathbb{E}[d_T(h(0), h(n))]| \geq \epsilon n \right) \leq C e^{-\frac{\epsilon^2 n}{2m}}.
\]

Dividing by \(n\) we can rewrite this inequality as

\[
\mathbb{P}_\mu \left( \left| \frac{1}{n} d_T(h(0), h(n)) - \mathbb{E} \left[ \frac{1}{n} d_T(h(0), h(n)) \right] \right| \geq \epsilon \right) \leq C e^{-\frac{\epsilon^2 n}{2m^2}}
\]

and thus it holds on the one hand that

\[
\lim_{n \to \infty} \mathbb{P}_\mu \left( \left| \frac{1}{n} d_T(h(0), h(n)) - \mathbb{E} \left[ \frac{1}{n} d_T(h(0), h(n)) \right] \right| \geq \epsilon \right) = 0.
\]
On the other hand, we have that
\[
\lim_{n \to \infty} \mathbb{P}_\mu \left( \left| \frac{1}{n} d_T(h(0), h(n)) - \mathbb{E} \left[ \frac{1}{n} d_T(h(0), h(n)) \right] \right| \geq \varepsilon \right) = \lim_{n \to \infty} \int \mathbb{P}_\nu \left( \left| \frac{1}{n} d_T(h(0), h(n)) - \mathbb{E} \left[ \frac{1}{n} d_T(h(0), h(n)) \right] \right| \geq \varepsilon \right) w_\mu(d\nu) 
\]
\[
\geq \lim_{n \to \infty} \int_{\{|s_\nu| - s_i| \geq \varepsilon\}} \mathbb{P}_\nu \left( \left| \frac{1}{n} d_T(h(0), h(n)) - \mathbb{E} \left[ \frac{1}{n} d_T(h(0), h(n)) \right] \right| \geq \varepsilon \right) w_\mu(d\nu) 
\]
\[
\geq \int_{\{|s_\nu| - s_i| \geq \varepsilon\}} \lim_{n \to \infty} \mathbb{P}_\nu \left( \left| \frac{1}{n} d_T(h(0), h(n)) - \mathbb{E} \left[ \frac{1}{n} d_T(h(0), h(n)) \right] \right| \geq \varepsilon \right) w_\mu(d\nu) 
\]
\[
\geq \delta, 
\]
which is a contradiction. Therefore, the slope is $w_\mu$-almost surely equal to $(s_1, \ldots, s_m)$ and there exists at least one ergodic gradient Gibbs measure with slope $(s_1, \ldots, s_m)$. □

It is also possible to define the slope in a similar way for any graph $G$ whose universal cover is a regular tree as described in [Cha16] and obtain the same result characterization of shift-invariant ergodic gradient Gibbs measure.

**Corollary 5.5.** Let $G$ be a finite $d$-regular graph with no four cycle and let $\text{Hom}(\mathbb{Z}^m, G)$ be the space of ergodic gradient Gibbs measures from $\mathbb{Z}^m$ to $G$ for the shift of $\mathbb{Z}^m$. For all $s \in \mathbb{R}^m$ such that $|s|_{\infty} \leq 1$ then there exist an ergodic gradient Gibbs measure with slope $s$.

**Proof.** It is known that the $d$-regular tree is the universal cover of all finite $d$-regular graphs with no four cycle (see [Cha16]). Let $G$ be such a graph, the push forward of the measure $\mu$ obtained in Theorem 5.4 defines a shift-invariant Gibbs measure on $\text{Hom}(\mathbb{Z}^m, G)$ which conserves the slope (one can choose the image of 0 uniformly among the vertices of $G$ to conserve the shift invariance). Since this push forward also contracts the distance between two vertices, the concentration inequality (41) still holds and one can use the exact same argument as in Theorem 5.4 to show the existence of an ergodic gradient Gibbs measure with slope $s$. □
The purpose of this section is to prove the results stated in Section 2. More precisely, we give the proof of Lemma 2.3, Lemma 2.5, Theorem 2.9 and Theorem 2.11.

We start with deducing Lemma 2.3.

**Proof of Lemma 2.3.** We start with showing that the path property (7) implies the condition (6). For arbitrary indexes $i, j \in \{1, \ldots, k\}$ we consider an element $x \in h_{R}^{-1}(\mathbb{R}^+, i) \cap h_{R}^{-1}(\mathbb{R}^+, j)$. We have to show that then $h_{R}^{1}(x) \leq a_{ij}$.

For that purpose, let $z_{l}^{i} \in h_{R}^{-1}(\mathbb{R}^+, i)$ and $z_{l}^{j} \in h_{R}^{-1}(\mathbb{R}^+, i)$ be two sequences of points converging to $x$ i.e. $\lim_{l \to \infty} z_{l}^{i} = x$ and $\lim_{l \to \infty} z_{l}^{j} = x$. For $l \in \mathbb{N}$ we consider consider a straight path $p_{l}$ connecting $z_{l}^{i}$ and $z_{l}^{j}$. By the path property (7) there exists a point $z_{l} \in p_{l}$ such that $h_{R}^{1}(z_{l}) \leq a_{ij}$. Additionally, we note that by construction $\lim_{l \to \infty} z_{l} = x$. Then, by continuity of the function $h_{R}^{1}$ it follows that $h_{R}^{1}(x) = \lim_{l \to \infty} h_{R}^{1}(z_{l}) \leq a_{ij}$, which verifies the property (6).

Now, let us show how the condition (6) together with the compatibility condition (4) yields the condition (7). We consider two points $x, y \in R$ and an arbitrary path $p \subset R$ that connects $x$ and $y$. Let $h_{R}^{2}(x) = i$ and $h_{R}^{2}(y) = j$. The path $p$ will cross several domains $h_{R}^{-1}(\mathbb{R}^+, i_{l})$, where $i_{l} \in \{1, \ldots, k\}$ and $1 \leq l \leq \ell$. We argue that by the compatibility condition (4) there must exist indexes $i_{l_{1}}$ and $i_{l_{2}}$ such that

$$a_{i_{l_{1}}, i_{l_{2}}} \leq a_{i_{l}, j}.$$  

Indeed, let us consider the sequence

$$i = i_{0}, i_{1}, \ldots, i_{l} = j.$$

and assume that for all $1 \leq \ell \leq l - 1$ it holds

$$a_{i_{0}, i_{\ell}} < a_{i_{\ell}, i_{\ell+1}}.$$
Then by using the condition \((4)\) on the numbers \(a_{i_0 i_j}\) and \(a_{i_0 i_l}\) it follows that
\[
a_{i_1, i_l} = a_{i_0, i_l}.
\]
Repeating this argument recursively yields that in the end
\[
a_{i_1 - 1, i_l} = a_{i_0, i_l},
\]
which is a contradiction to \((62)\).

Let us now be \(i_{t_1}\) and \(i_{t_2}\) be elements that satisfy \((61)\). By construction we now that
\[
p \cap h_R^{-1}(\mathbb{R}^+, i_{t_1}) \cap h_R^{-1}(\mathbb{R}^+, i_{t_2}) \neq \emptyset.
\]
Therefore, let us choose an element
\[
z \in p \cap h_R^{-1}(\mathbb{R}^+, i_{t_1}) \cap h_R^{-1}(\mathbb{R}^+, i_{t_2}).
\]

Then by using the admissibility condition \((6)\) it follows that
\[
h_R^1(z) \leq a_{i_{t_1}} a_{i_{t_2}} \leq a_{i, j},
\]
which verifies the path condition \((2.3)\). \qed

Let us now turn to the proof of Lemma \(2.5\). The argument uses a standard construction to extend Lipschitz functions from the boundary \(\partial R\) to the whole set \(R\).

**Proof of Lemma \(2.5\)** Let \((h, (a_{i j})_{k \times k})\) be an asymptotic boundary height profile and define \(g : R \to \mathbb{R}^+ \times \{1, \ldots, k\}\) by:
\[
\begin{align*}
g_1(y) &= \max\{0, \max_{x \in \partial R} \{h_1(x) - |x - y|_1\}\} \quad \text{(63)} \\
g_2(y) &= h_2(\arg \max_{x \in \partial R} \{h_1(x) - |x - y|_1\}).
\end{align*}
\]
We will show that under that \(g\) extends \(h\) to an asymptotic height profile on the whole region \(R\). In order to prove this we show the following three properties:
- \(g = h\) on \(\partial R\);
- \(g\) satisfies the condition \((5)\);
- \(g\) satisfies the condition \((6)\).

The first property is a simple consequence of the inequality \((5)\) that holds for the function \(h_1\). Indeed using \((5)\) yields that for all \(x, y \in \partial R\) we have \(\max_{x \in \partial R} \{h_1(x) - d(x, y)\} \leq h_1(y)\) and thus \(g_1(y) = h_1(y)\). For the second property, we first observe that by a combination of the triangle inequality and the fact that \(h_1\) satisfies \((5)\) on \(\partial R\), the function \(\tilde{g}_1 : R \to \mathbb{R}\) given by
\[
\tilde{g}_1(y) := \max_{x \in \partial R} \{h_1(x) - |x - y|_1\}
\]
satisfies on $R$ the condition (5). Then it is a simple consequence that also $g_1$ satisfies (5) on $R$.

Let’s turn to the third property. We show the third property by contradiction. Let us assume that the map $g$ does not satisfy the condition (6). Then there is a point $x \in g^{-1}(\mathbb{R}^+, i) \cap g^{-1}(\mathbb{R}^+, j)$ such that

$$g_1(x) > a_{ij}.$$ 

Then by continuity there are two points $x_1, x_2 \in R$ such that

$$g_2(x_1) = i, \quad g_2(x_2) = j, \quad g_1(x_1) > a_{ij} \geq 0, \quad \text{and} \quad g_1(x_2) > a_{ij} \geq 0.$$ 

By definition (63) of the map $g$ it follows that there is a point $z_1 \in \partial R$ and $z_2 \in \partial R$ such that

$$g_1(x_1) = h_1(z_1) - |x_1 - z_1|l_1 > a_{ij} \geq 0 \quad (64)$$

and

$$g_1(x_2) = h_1(z_2) - |x_2 - z_2|l_1 > a_{ij} \geq 0.$$ 

Because we can choose $|x_1 - x_2|l_1$ to be arbitrarily small the last estimate yields that

$$h_1(z_2) - |x_1 - z_2|l_1 \geq h_1(z_2) - |x_2 - z_2|l_1 - |x_1 - x_2|l_1 > a_{ij} \geq 0 \quad (65)$$

A combination of (64) and (65) yields that

$$|h_1(z_1) - a_{ij}| + |h_1(z_2) - a_{ij}| > |x_1 - z_1|l_1 + |x_1 - z_2|l_1,$$

$$\geq |z_1 - z_2|l_1.$$ 

The last estimate contradicts inequality (8) and completes the argument. □

Let’s turn to the verification of Theorem 2 and Theorem 2.11. Inspired by the argument of [CKP01] we first deduce Theorem 2.11 and then use Theorem 2.11 to verify Theorem 2.9 via a compactness argument.

The main idea of the proof of Theorem 2.11 is to establish the desired identity (13) in two steps: In the first step one underestimates the number of graph homomorphisms $h : R_n \to \mathcal{T}$ and in the second step one overestimates the number of graph homomorphisms. The main ingredient for underestimating the number of graph homomorphisms is the equivalence of the entropy of fixed and free boundary conditions on a square (see Theorem 3.9). The tool for overestimating the number of graph homomorphisms is provided now. As mentioned in Remark 2.12 of Section 2, the following lemma also justifies that the entropic effect,
that results from the additional freedom of choosing the geodesic \( g \), is of lower order.

**Lemma 6.1.** Let \( \epsilon > 0 \) and let \( S_n \subset \mathbb{Z}^m \) be the \( n \times n \) square given by

\[
S_n := \{ x \in \mathbb{Z}^m \mid \forall i: 0 \leq x_i \leq n - 1 \}.
\]

Let \( H \) be the set of homomorphisms \( \{ h : S_n \to T \} \) such that

- \( 0 \) is mapped to the root \( r \), i.e. \( h(0) = r \in T \) and
- for all \( x = (x_1, \ldots, x_m) \in \partial S_n \) we have
  \[
  |d_T(h(x), r) - \langle s \cdot x \rangle| \leq \epsilon n.
  \]

Then there is an integer \( n_0 \) such that for all \( n \geq n_0 - 1 \)

\[
-\frac{1}{n^m} \ln |H| = \text{ent}_n(s) + \theta(\epsilon),
\]

where \( \text{ent}_n(s) \) is the local surface tension given by Definition 3.7.

**Proof.** Let \( h \in H \). We start with showing that there exist a geodesic segment \( s \subset T \) such that for all \( x \in \partial S_n \)

\[
d_T(h(x), s([s \cdot x])) \leq 2\epsilon n,
\]

where \( s([s \cdot x]) = v \) denotes the unique element \( v \in s \) such that \( |v| := d_T(v, r) = \langle s \cdot x \rangle \).

Indeed, let us define \( s \) to be the segment of \( h \) between the root \( h(0) = r \) and a corner of \( S_n \) i.e. \( h(\pm n, \ldots, \pm n) \). Without loss of generality we choose the corner \( h(\pm n, \ldots, \pm n) \) and assume that \( (s_1, \ldots, s_m) \) are all non-negative (else one would have to choose another corner). We assume that (66) is not satisfied. Hence, we suppose that there exist \( x \in \partial S_n \) that

\[
d_T(h(x), s([s \cdot x])) > 2\epsilon n.
\]

There exists a path \( p \subset \mathbb{Z}^m \) between \( x \) and \( (n, \ldots, n) \) on the boundary of \( \partial S_n \) on which all the coordinates are non-decreasing. This means that for all \( y \in p \) we have

\[
s \cdot x \leq s \cdot y.
\]

Now, consider \( z \) to be the first point on \( p \) whose image is on \( s \), we must have \( |h(z)| < |h(x)| - 2\epsilon n \), where we used the short notation \( |h(x)| = d_T(h(x), h(0)) \). A direct calculation yields (cf. the proof of Theorem 4.6) that

\[
2\epsilon < |h(x)| - |h(z)|
\]

\[
= |h(x)| - s \cdot x + s \cdot x - s \cdot y + s \cdot y - |h(z)|
\]

\[
\leq ||h(x)| - s \cdot x| + |s \cdot y - |h(z)||
\]
Hence, it follows that either
\[ ||h(z)| - s \cdot z| > \epsilon n \]
or
\[ ||h(x)| - s \cdot x| > \epsilon n, \]
which is a contradiction with \( h \in H \) and verifies (66).

Now, we observe that due to (66) we can define for all \( h \in H \) a geodesic \( g \) such that \( s \subset g \) and
\[ d_T(h(x), g(\lfloor s \cdot x \rfloor)) \leq 2\epsilon n. \]

Now, let us estimate how many different geodesics \( g \) can exist. On \( S_n \) there are less than \( d^{m^{-1}} \) possible distinct boundary graph homomorphisms. Hence, the number of possible distinct geodesics is also bounded by \( d^{m^{-1}} \). All, in all, we obtain from Theorem 3.9 that
\[ \frac{1}{n^m} \ln |H| \geq -\frac{1}{n^m} \log d^{m^{-1}} + \text{ent}_n(s) + \theta(\epsilon) \]
\[ = -\frac{1}{n} \ln d + \text{ent}_n(s) + \theta(\epsilon) \]
\[ \geq \text{ent}_n(s) + \theta(\epsilon), \]
where we have chosen \( n \geq n_0 \) large enough in the last step. The reverse inequality easily follows from Theorem 3.9 by underestimating the elements in \( H \) i.e.
\[ -\frac{1}{n^m} \ln |H| \leq \text{ent}_n(s) + \theta(\epsilon), \]
this completes the argument. \qed
Proof of Theorem 2.11. The first step of the proof is to show that one can assume wlog. that the asymptotic height profile \((h, a_{ij})\) has a simpler structure. For that reason, let us consider the discretized region \(R_\varepsilon\) given by

\[
R_\varepsilon = \{ z \in R | z \in \varepsilon \mathbb{Z}^d \}.
\]

Given the asymptotic height profile \((h, a_{ij})\), we consider an associated asymptotic height profile \((h_\varepsilon, a_{ij})\) such that: The first coordinate \(h_1^\varepsilon\) is the linear interpolation of \(h_1(z)\) with respect to the lattice \(R_\varepsilon\) and the second coordinate \(h_2^\varepsilon = h^2\).

Later, we will need the following relation i.e.

\[
HP_n(h_\varepsilon, 0.5\delta, \varepsilon) \subset HP_n(h, \delta, \varepsilon) \subset HP_n(h_\varepsilon, 2\delta, \varepsilon), \quad (67)
\]

which follows easily from the definition \((12)\) of \(HP_n(h, \delta, \varepsilon)\) by choosing \(\varepsilon\) small enough and exploiting the fact that \(h^1\) is 1-Lipschitz.

The main step of the argument is to show that that \((13)\) is satisfied for \(h_\varepsilon\) i.e.

\[
\text{Ent}(h_\varepsilon) = -\frac{1}{|R_n|} \ln |HP_n(h_\varepsilon, \delta, \varepsilon)| + \theta(\varepsilon) + \theta(\delta) + \theta \left( \frac{1}{\varepsilon n} \right). \quad (68)
\]

We will deduce \((68)\) later and now argue that then \((13)\) is also satisfied for the original asymptotic height profile \((h, (a_{ij})_{k \times k})\).

We need the following fact. Due to the convexity of the local surface tension \(\text{ent}(s)\) and that \(h^1\) is 1-Lipschitz, it holds for a.e. \(x \in R\) that

\[
\text{ent}(\nabla h^1_\varepsilon(x)) = \text{ent}(\nabla h^1(x)) + \theta(\varepsilon)
\]
and therefore
\[
\text{Ent}(h_\varepsilon) = \int_R \text{ent}(\nabla h_\varepsilon^1(x)) dx
\]
(69)
\[
= \int_R \text{ent}(\nabla h^1(x)) dx + \theta(\varepsilon)
\]
\[
= \text{Ent}(h) + \theta(\varepsilon).
\]
As a consequence, we get by combining (67), (68) and (69) that
\[
-\frac{1}{|R_n|} \ln |HP_n(h, \delta, \varepsilon)| \leq -\frac{1}{|R_n|} \ln |HP_n(h_\varepsilon, 0.5\delta, \varepsilon)|
\]
\[
= \text{Ent}(h_\varepsilon) + \theta(\varepsilon) + \theta(0.5\delta) + \theta \left( \frac{1}{\varepsilon n} \right)
\]
\[
= \text{Ent}(h) + \theta(\varepsilon) + \theta(\delta) + \theta \left( \frac{1}{\varepsilon n} \right)
\]
as well as
\[
-\frac{1}{|R_n|} \ln |HP_n(h, \delta, \varepsilon)| \geq -\frac{1}{|R_n|} \ln |HP_n(h_\varepsilon, 2\delta, \varepsilon)|
\]
\[
= \text{Ent}(h_\varepsilon) + \theta(\varepsilon) + \theta(2\delta) + \theta \left( \frac{1}{\varepsilon n} \right)
\]
\[
= \text{Ent}(h) + \theta(\varepsilon) + \theta(\delta) + \theta \left( \frac{1}{\varepsilon n} \right).\]
The last two estimates yield the desired conclusion that (13) is also valid for the height profile \((h, (a_{ij})_{k \times k})\).

Hence, it is left to deduce the identity (69). This identity is deduced in two steps. In the first step, we show that
\[
-\frac{1}{|R_n|} \ln |HP_n(h_\varepsilon, \delta, \varepsilon)| \leq \int_R \text{ent}(\nabla h_\varepsilon^1(z)) dz + \theta(\varepsilon) + \theta(\delta) + \theta \left( \frac{1}{n} \right)
\]
(70)
In the second step, we show that
\[
-\frac{1}{|R_n|} \ln |HP_n(h_\varepsilon, \delta, \varepsilon)| \geq \int_R \text{ent}(\nabla h_\varepsilon^1(z)) dz + \theta(\varepsilon) + \theta(\delta) + \theta \left( \frac{1}{n} \right)
\]
(71)
Let us turn to the first step i.e. the verification of (70). Because \(R_n\) converges to \(R\) in the Gromov-Hausdorff sense, we may assume wlog.
$R \subset \mathbb{R}^m$ can be written as a finite union of $l_\infty$-blocks of side length $\varepsilon$ aligned to the grid $R_{\text{grid},\varepsilon}$ i.e. (see also Figure 16 and Figure 17)

$$R = \bigcup_{k=1}^l B_k. \quad (72)$$

This implies in particular that

$$|R_n| = l \left(1 + \theta \left(\frac{1}{n}\right)\right) \varepsilon^m n^m. \quad (73)$$

In order to deduce (70), we have to underestimate the number of graph homomorphism in $HP_n(h_\varepsilon, \delta, \varepsilon)$. For that purpose let us fix an element $g_n \in HP_n(h_\varepsilon, \delta, \varepsilon)$. The graph homomorphism $g_n$ gives us boundary values on the set $R_{\text{grid},\varepsilon}$ (cf. Figure 16). We define the set

$$UP_n(h_\varepsilon, \delta, \varepsilon) := \left\{h_n \in M(R_n, h_{\partial R_n}) : \forall x \in R_n : \frac{x}{n} \in R_{\text{grid},\varepsilon} h_n(x) = g_n(x)\right\}.$$

Then it follows from the definition that

$$|HP_n(h_\varepsilon, \delta, \varepsilon)| \geq |UP_n(h_\varepsilon, \delta, \varepsilon)|, \quad (74)$$

which leads to the estimate

$$-\frac{1}{|R_n|} \ln |HP_n(h_\varepsilon, \delta, \varepsilon)| \leq -\frac{1}{|R_n|} \ln |UP_n(h_\varepsilon, \delta, \varepsilon)|$$

$$= -\frac{1}{l(1 + (\frac{1}{n}))} \varepsilon^m n^m \ln |UP_n(h_\varepsilon, \delta, \varepsilon)|$$

where we used the identity (73) in the last step.

Now the idea is that every block $B_k$ of $R_{\text{grid},\varepsilon}$ becomes independent and can be estimated by itself. In order to estimate the entropy of a block we use Lemma 3.9, which can be applied due to the fact that the function $h_\varepsilon$ is linear on the boundary of a block $B_k$ (see also the
definition (74) of the set $UP_n(h, \delta)$. Hence, we get that

$$-\frac{1}{l} \left(1 + \theta \left(\frac{1}{l} \right)\right) \varepsilon^n m \ln |UP_n(h, \delta, \varepsilon)|$$

$$= -\frac{1}{l} \left(1 + \theta \left(\frac{1}{l} \right)\right) \sum_{k=1}^l \left(\text{ent}_{\varepsilon n} (\nabla h_1^1(x_k)) + \theta(\delta) + \theta \left(\frac{1}{n}\right)\right)$$

$$= -\frac{1}{l} \sum_{k=1}^l \text{ent}_{\varepsilon n} (\nabla h_1^1(x_k)) + \theta(\delta) + \theta \left(\frac{1}{n}\right)$$

$$= -\frac{1}{l \varepsilon^m} \sum_{k=1}^l \varepsilon^m \text{ent} (\nabla h_1^1(x_k)) + \theta(\delta) + \theta \left(\frac{1}{n}\right) + \theta \left(\frac{1}{\varepsilon n}\right),$$

where $x_k \in B_k \subset R$ chosen arbitrary for $1 \leq k \leq l$. By using that due to (72) it holds that $|R| = l \varepsilon^m$, we get that

$$\frac{1}{l \varepsilon^m} \sum_{k=1}^l \varepsilon^m \text{ent} (\nabla h_1^1(x_k)) + \theta(\delta) + \theta \left(\frac{1}{n}\right) + \theta \left(\frac{1}{\varepsilon n}\right)$$

$$= \frac{1}{|R|} \int_R \text{ent} (\nabla h_1^1(x)) dx + \theta(\delta) + \theta \left(\frac{1}{n}\right) + \theta \left(\frac{1}{\varepsilon n}\right),$$

which overall deduces the desired estimate (70).

Let us now turn to the verification of (71). The strategy is to overestimate the number of graph homomorphisms contained in $HP_n(h, \delta, \varepsilon)$. By having a close look at the definition of $HP_n(h, \delta, \varepsilon)$, we can overestimate the number of graph homomorphisms by looking at the values of a height function on each block $B_k$ independently (see also Figure 16). To do so, on every block $B_k$ we consider a boundary condition of the same type as in Lemma 6.1. Then, using Lemma 6.1 and the identity (73) yields that

$$-\frac{1}{|R|} \ln |HP_n(h, \delta, \varepsilon)|$$

$$\geq -\frac{1}{m} \left(1 + \theta \left(\frac{1}{m} \right)\right) \sum_{k=1}^m \left(\text{ent}_{\varepsilon n} (\nabla h_1^1(x_k)) + \theta(\delta) + \theta \left(\frac{1}{n}\right)\right)$$

$$= \frac{1}{m} \sum_{k=1}^m \text{ent}_{\varepsilon n} (\nabla h_1^1(x_k)) + \theta(\delta) + \theta \left(\frac{1}{n}\right)$$

$$= \frac{1}{m \varepsilon^m} \sum_{k=1}^m \varepsilon^m \text{ent} (\nabla h_1^1(x_k)) + \theta(\delta) + \theta \left(\frac{1}{n}\right) + \theta \left(\frac{1}{\varepsilon n}\right).$$
which is the desired estimate (71) and closes the argument. $\square$

We deduce Theorem 2.9 from Theorem 2.11 with a standard compactness argument similar to the one used in [CKP01].

**Proof of Theorem 2.9.** We consider a fixed asymptotic boundary height profile $(h_{\partial R}, (a_{ij})_{k \times k})$ and consider the set $AHP(h_{\partial R}, (a_{ij})_{k \times k})$ of all possible extensions to an asymptotic height profile $(h, (a_{ij})_{k \times k})$ (see Definition 2.2). We observe that

$$AHP(h_{\partial R}, (a_{ij})_{k \times k}) \neq \emptyset.$$  

It follows from the fact that the space $AHP(h_{\partial R}, (a_{ij})_{k \times k})$ is closed and the local surface tension $\text{ent}$ is convex (cf. Theorem 3.2) and uniformly bounded from below by $-d$ that

$$\inf_{h_R \in SA(h_{\partial R}, a_{ij})} \text{Ent} (R, h_R) = \min_{h_R \in SA(h_{\partial R}, a_{ij})} \text{Ent} (R, h_R) = \text{Ent} (R, h_{\text{min}}),$$

where $(h_{\text{min}}, a_{ij})$ is the minimizer of the macroscopic entropy $\text{Ent}(R, h_R)$.

Let us fix an $\dd > 0$. We have to show that there exists an integer $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ it holds

$$\text{Ent} (\Lambda_n, h_{\partial R_n}) \leq \text{Ent} (R, h_{\text{min}}) + \dd$$ \hfill (75)

and

$$\text{Ent} (\Lambda_n, h_{\partial R_n}) \geq \text{Ent} (R, h_{\text{min}}) - \dd.$$ \hfill (76)

We start with deducing the estimate (75). Underestimating the number of graph homomorphisms and using the identity (13) yields that

$$\text{Ent} (\Lambda_n, h_{\partial R_n}) = -\frac{1}{|R_n|} \ln |M(R_n, h_{\partial R_n})|$$

$$\leq -\frac{1}{|R_n|} \ln |H_{\text{min}} (h_{\text{min}}, \delta) |$$

$$= \text{Ent} (R, h_{\text{min}}) + \theta(\varepsilon) + \theta(\delta) + \theta \left( \frac{1}{\varepsilon n} \right).$$

Choosing now first $\varepsilon > 0$ and $\delta > 0$ small, and then $n$ large (depending on $\varepsilon$) yields the desired upper bound (75).
Let us now deduce the lower bound \((76)\). By compactness we know that for a fixed number \(\delta > 0\), there is an integer \(l\), depending on \(\delta\) but not on \(n\), and asymptotic height profiles \(h_1, \ldots, h_l\) such that

\[
M(R_n, h_{\partial R_n}) \subset \bigcup_{i=1}^l HP_n(h_i, \delta, \varepsilon).
\]  

(77)

Without loss of generality we may assume that the function \(h_1 = h_{\min}\). From the definition of \(h_{\min}\) and Theorem 3.9 it follows that (cf. 6)

\[
\frac{1}{|R_n|} \ln |HP_n(h_{\min}, \delta, \varepsilon)| = \text{Ent}(R, h_{\min}) + \theta(\delta)
\]

\[
\leq \text{Ent}(R, h_i) + \theta(\delta)
\]

\[
= -\frac{1}{|R_n|} \ln |HP_n(h_i, \delta, \varepsilon)| + \theta(\delta)
\]

for all \(i \in \{1, \ldots, l\}\). The last inequality yields that for all \(i \in \{1, \ldots, l\}\)

\[
|HP_n(h_i, \delta, \varepsilon)| \leq |HP_n(h_{\min}, \delta, \varepsilon)| \exp(|R_n|\theta(\delta)).
\]

This yields in combination with (13) and (77) and \(h_1 = h_{\min}\) that

\[
\text{Ent}(\Lambda_n, h_{\partial R_n}) = -\frac{1}{|R_n|} \ln |M(R_n, h_{\partial R_n})|
\]

\[
\geq -\frac{1}{|R_n|} \ln \sum_{i=1}^l |HP_n(h_i, \delta, \varepsilon)|
\]

\[
\geq -\frac{1}{|R_n|} \ln \big( l |HP_n(h_1, \delta, \varepsilon)| \big) - l\theta(\delta)
\]

\[
= \text{Ent}(R, h_{\min}) + \theta(\varepsilon) + (1 - l)\theta(\delta) + \theta \left( \frac{1}{\varepsilon n} \right) - \frac{1}{|R_n|} \ln l
\]

\[
\geq \text{Ent}(R, h_{\min}) + (1 - l)\theta(\varepsilon) + \theta(\delta) + \theta \left( \frac{\ln l}{\varepsilon n} \right).
\]

Choosing now first \(\varepsilon > 0\) and \(\delta > 0\) small and then \(n\) large enough (depending on \(\varepsilon\) and \(l\)) verifies the lower bound \((76)\) and finishes the argument.

\[\square\]

**Acknowledgement**

The authors want to thank Marek Biskup, Nishant Chandgotia, Filippo Colomo, Nicolas Destainville, Richard Kenyon, Michel Ledoux, Igor
Pak, Laurent Saloff-Coste, Scott Sheffield, and Tianyi Zheng for helpful discussions and comments.

REFERENCES

[Azu67] Kazuoki Azuma. Weighted sums of certain dependent random variables. *Tohoku Math. J. (2)*, 19:357–367, 1967.

[BCG16] Alexei Borodin, Ivan Corwin, and Vadim Gorin. Stochastic six-vertex model. *Duke Math. J.*, 165(3):563–624, 02 2016.

[BW00] Graham R. Brightwell and Peter Winkler. Gibbs measures and dismantlable graphs. *J. Combin. Theory Ser. B*, 78(1):141–166, 2000.

[CEP96] Henry Cohn, Noam Elkies, and James Propp. Local statistics for random domino tilings of the Aztec diamond. *Duke Math. J.*, 85(1):117–166, 1996.

[Cer06] R. Cerf. *The Wulff crystal in Ising and percolation models*, volume 1878 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, With a foreword by Jean Picard.

[Cha16] Nishant Chandgotia. Four-cycle free graphs, height functions, the pivot property and entropy minimality. *Ergodic Theory and Dynamical Systems*, FirstView:1–31, 3 2016.

[CKP01] Henry Cohn, Richard Kenyon, and James Propp. A variational principle for domino tilings. *J. Amer. Math. Soc.*, 14(2):297–346 (electronic), 2001.

[CS16] F. Colomo and A. Sportiello. Arctic curves of the six-vertex model on generic domains: The tangent method. *Journal of Statistical Physics*, 164(6):1488–1523, 2016.

[Des98] N. Destainville. Entropy and boundary conditions in random rhombus tilings. *J. Phys. A*, 31(29):6123–6139, 1998.

[DKS92] R. Dobrushin, R. Kotecký, and S. Shlosman. *Wulff construction*, volume 104 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1992. A global shape from local interaction, Translated from the Russian by the authors.

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[FS06] Patrik L. Ferrari and Herbert Spohn. Scaling limit for the space-time covariance of the stationary totally asymmetric simple exclusion process. *Comm. Math. Phys.*, 265(1):1–44, 2006.

[Geo88] Hans-Otto Georgii. *Gibbs measures and phase transitions*, volume 9 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1988.

[Hoe63] Wassily Hoeffding. Gibbs measures and phase transitions. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[Kas63] P. W. Kasteleyn. Dimer statistics and phase transitions. *J. Mathematical Phys.*, 4:287–293, 1963.

[Kir34] M. Kirszbraun. Über die zusammenziehende und Lipschitzsche Transformationen. *Fundamenta Mathematicae*, 22(1):77–108, 1934.

[KK92] Claire Kenyon and Richard Kenyon. Tiling a polygon with rectangles. In *Foundations of Computer Science, 1992. Proceedings., 33rd Annual Symposium on*, pages 610–619. IEEE, 1992.
A VARIATIONAL PRINCIPLE FOR A NON-INTEGRABLE MODEL

Benot Kloeckner. On Lipschitz compactifications of trees. *Comptes Rendus Mathematique*, 346(7):413 – 416, 2008.

Richard Kenyon, Andrei Okounkov, and Scott Sheffield. Dimers and amoebae. *Ann. of Math. (2)*, 163(3):1019–1056, 2006.

Lionel Levine and Yuval Peres. Strong spherical asymptotics for rotor-router aggregation and the divisible sandpile. *Potential Analysis*, 30(1):1, 2008.

Michael Luby, Dana Randall, and Alistair Sinclair. Markov chain algorithms for planar lattice structures. *SIAM J. Comput.*, 31(1):167–192, 2001.

B. F. Logan and L. A. Shepp. A variational problem for random Young tableaux. *Advances in Math.*, 26(2):206–222, 1977.

Ananth Sridhar Nicolai Reshetikhin. Limit shapes of the stochastic six vertex model. *arXiv:1609.01756*, 2016.

Boris Pittel and Dan Romik. Limit shapes for random square Young tableaux. *Adv. in Appl. Math.*, 38(2):164–209, 2007.

Igor Pak, Adam Sheffer, and Martin Tassy. Fast domino tileability. *Discrete Comput Geom*, 56(2):377–394, 2016.

J. T. Schwartz. *Nonlinear functional analysis*. Gordon and Breach Science Publishers, New York-London-Paris, 1969. Notes by H. Fattorini, R. Nirenberg and H. Porta, with an additional chapter by Hermann Karcher, Notes on Mathematics and its Applications.

Scott Sheffield. Ribbon tilings and multidimensional height functions. *Transactions of the American Mathematical Society*, 354(12):4789–4813, 2002.

Scott Sheffield. Random surfaces. *Astérisque*, (304):vi+175, 2005.

Martin Tassy. *Translation invariant Gibbs measure on tilings*. In preparation, Brown University, 2014.

F. A. Valentine. On the extension of a vector function so as to preserve a Lipschitz condition. *Bull. Amer. Math. Soc.*, 49:100–108, 1943.

A. M. Versik and S. V. Kerov. Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux. *Dokl. Akad. Nauk SSSR*, 233(6):1024–1027, 1977.

David Bruce Wilson. Mixing times of lozenge tiling and card shuffling Markov chains. *Ann. Appl. Probab.*, 14(1):274–325, 02 2004.

**Department of Mathematics, University of California, Los Angeles**

*E-mail address: gmenz@math.ucla.edu*

**Department of Mathematics, University of California, Los Angeles**

*E-mail address: mtassy@math.ucla.edu*