UNIVERSITY OF CALIFORNIA

Los Angeles

Homogenization in Discrete Random Models

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Mathematics

by

Andrew Krieger

2021

ABSTRACT OF THE DISSERTATION

Homogenization in Discrete Random Models

by

Andrew Krieger

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2021

Professor Georg Menz, Chair

This dissertation studies the asymptotic behavior of two probabilistic models. It consists of two parts, one from the specific area of random surface models within the wider branch of statistical physics, and the other from the area of random graphs, and more specifically on random graphs built upon a (non-trivial) underlying geometry. In both cases, the random structure homogenizes as the system size tends to infinity.

The first model under study is $\mathbb{Z}$-valued graph homomorphisms from the lattice $\mathbb{Z}^d$. It is known that this model exhibits limit shapes: a graph homomorphism chosen uniformly at random subject to fixed boundary values will, with high probability, lie uniformly close to a certain limiting profile over the bulk of the lattice subset. We extend the limit shape result, as quantified via a variational principle, large deviations principle, and concentration inequality, to a new version of the model. In the new version, the uniform distribution over graph homomorphism is perturbed by a random potential. This illustrates the robustness of the results and the methods used to prove them.

The second model is long-range percolation on the lattice graph $\mathbb{Z}^d$. This is a random graph that includes all nearest-neighbor edges in $\mathbb{Z}^d$ plus a random selection of longer edges.

Longer edges are included or excluded at random and independently, where the probability that the edge with endpoints x and y is included is asymptotic to $\beta|x-y|^{-s}$ for some $s \in (d, 2d)$ and some $\beta > 0$. We sharpen the best known asymptotics for the graph distance under this choice of edge inclusion probability. The proof is inspired by the recent work [BL19], which introduced and studied a continuum analogue of the model, and the conclusion is similar to the corresponding result contained therein.

The dissertation of Andrew Krieger is approved.

Marek Biskup

Guido Montúfar

Tim Austin

Georg Menz, Committee Chair

University of California, Los Angeles

2021

To my wife Di,

who has supported me throughout my Ph.D. program

and who inspires me to reach higher in all my endeavors.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I am immensely grateful to my advisor, Georg Menz. He was generous with his time and patient even when my progress sometimes slowed down. He would always encourage me and suggest next steps so that I could complete my PhD, and apparently his efforts were fruitful. I must also credit Georg for introducing me to the research problem that makes up the bulk of this dissertation, at least by page count: the study of limit shapes for discrete random surface models. The problem captured my attention while I was looking for an area to specialize in, it provided a good test against which I sharpened my technical skills, and it presents plenty of interesting open problems and connections to other fields. Georg has an excellent understanding of the literature and of useful tricks and techniques, so his mathematical advice was a huge benefit. I could not have gotten to this point without Georg's advice and support, and I am grateful for it.

I am also thankful towards my other committee members. In particular Marek Biskup sometimes stepped in to serve as nearly a second advisor, especially to provide guidance relating to my ongoing progress as a student and me future career options. My only regret in working with him is that I did not take up more of his time in our recent collaboration regarding the other research topic in this dissertation, the geometry of the long range percolation model. In working with him, it's impossible not to notice that he is a brilliant researcher who catches on to problems and arrives at solutions very quickly. Marek was also instrumental in convincing me to start writing this dissertation and to finish it in a timely manner, or at least an on-time manner. On that note, I must thank Guido Montúfar and Tim Austin as well for their commitment in hearing my defense and reviewing my dissertation quickly, despite the unfair workload I created for them by completing it so close to the deadline.

I am thankful to my past collaborators for the opportunity to work with them, and for their generosity in allowing me to include versions of some of our joint work in this dissertation. In particular, some of the material in Chapter 2 of this dissertation is a version of [KMT20b] and of [Kri21]. The former is a collaboration with Georg and with Martin Tassy. Together they provided a starting point for the project with their earlier work [MT20a] and contributed the high level structure of [KMT20b]. They were also instrumental in solving some of the technical challenges in the project. Also, some of the material in Chapter 3 is a version of [BK21]. That article is collaboration with Marek. He provided invaluable advice and editing for the portion of [BK21] which is reflected in here. He also solely produced a second major result that was not included here. Additionally, my work was supported in part by NSF grant DMS-1712632.

Finally, I must offer my heartfelt thanks to my family. It is impossible to overstate how much they've helped. My parents taught me well and pushed me to excel from a young age, and they've always been there to cheer me on. My wife Di picked up where my parents left off in terms of motivating me. She leads by example, pushing herself hard and encouraging me to keep up. Now that I've caught up on the PhD, I can keep growing professionally and personally alongside her. Di and her parents deserve special credit for helping so much to take care of our daughter while I was occupied with writing this dissertation. I couldn't have done this without the support of my family, and I'm blessed to have had them in my life.

VITA

2020–2021 Research Assistant, Mathematics Department, UCLA.

2014–2020 Teaching Assistant, Mathematics Department, UCLA.

2014 B.S. (Mathematics and Computer and Information Science), Ohio State University.

PUBLICATIONS

M. Biskup and A. Krieger, *Arithmetic oscillations for the chemical distance in long-range percolation on* $\mathbb{Z}^d$. In preparation.

A. Krieger, *The concentration inequality for a discrete height function model perturbed by random potential*. In preparation, preprint available at `arXiv:2111.05907`.

A. Krieger, G. Menz, M. Tassy, *Homogenization of the variational principle for discrete random maps*. Under revision, preprint available at `arXiv:1710.11330`.

A. Krieger, G. Menz, M. Tassy, *Deducing a variational principle with minimal a priori assumptions*. Electron. J. Comb. 27(4), 2020. Available at `doi:10.37236/9121`.

M. Springer, A. Krieger, S. Manilov, H. Masuhara, *dart2java: Running Dart in Java-based Environments*. Proc. ICOOOLPS, 2017, 1–6. Available at `doi:10.1145/3098572.3098575`.

CHAPTER 1

Introduction

Many classical and modern results in probability theory are fundamentally asymptotic results, in that they describe the limiting behavior of large random systems quantitatively or qualitatively. Classical examples include essential results such as the law of large numbers and the central limit theorem, which form the basis for modern statistics. A common principle is that, under certain mild assumptions, the underlying randomness homogenizes or is averaged out, leading to a deterministic limit. This dissertation addresses two specialized models for which there are interesting asymptotic results illustrating the principle of homogenization.

The first model studied is the $\mathbb{Z}$-homomorphism model, where the fundamental objects are graph homomorphisms from a lattice graph to $\mathbb{Z}$, called height functions in the text below. For the purposes of this dissertation, the codomain of the height functions will always be a connected, finite subgraph $R_n \subset \mathbb{Z}^d$, with the both $\mathbb{Z}$ and $\mathbb{Z}^d$ endowed with the nearest-neighbor graph structure and with R_n taken to be the subgraph induced by its vertices. In other words, if $x, y \in R_n$ where x, y are nearest-neighbors in $\mathbb{Z}^d$, then the edge between x and y is included in R_n as well. As discussed below, we study the asymptotic characteristics of a typical such height functions, subject to boundary value constraints. Here the asymptotic parameter is the domain size parameter n; we take as hypotheses that the domains R_n converge under a scaling limit, i.e. $\frac{1}{n} R_n \to R \subset \mathbb{R}^d$ in the Hausdorff distance, and that the boundary conditions converge under a similar scaling limit (see Definition 4 in Chapter 2). When the notion of typicality is defined by the uniform probability measure the set of height functions (subject to the boundary conditions), then several asymptotic results are known in the existing literature. We focus in particular on the following: First, the variational principle, which characterizes the scaling limit of typical height functions $h_n : R_n \to \mathbb{Z}$ as the minimizer of a certain integral over R (or in some cases as a family of distinct minimizers, if uniqueness is not established). Second, the large deviations principle, which implies that the probability of sampling a height function that is not uniformly close to the (or a) minimizing limiting height profile tends to zero. Third, the concentration inequality, from which one can deduce quantitative bounds on the concentration of the probability measure on height functions on R_n around its mean. Our work is novel in that we do not study the uniform probability measure, but a random perturbation of it. The results below therefore show the robustness of these results and of the methods of proof used to derive them.

The second model under study is the long-range percolation model. This is a geometric random graph, constructed by taking an underlying deterministic graph or metric space and adding edges between arbitrary pairs of vertices independently, with the probability that an edge is included given as a function of the distance between the endpoints (generally decaying as the distance becomes large). As suggested by the name, this model was introduced as a percolation model, and percolation does occur under certain conditions on the edge inclusion probabilities. However, here we work in a regime where percolation is guaranteed to occur, and percolation is not the subject of our work. Instead we are interested in the asymptotic growth rate of the graph distance in the resulting random graph, as a function of the edge inclusion probabilities. We provide sharper asymptotic results than were previously known.

1.1 The $\mathbb{Z}$-homomorphism model

The asymptotic behavior of the $\mathbb{Z}$-homomorphism model is a problem of statistical physics, and more specifically this model is within the class of random surface models. Statistical physics originates in the natural sciences as a broadly successful attempt to explain humanscale or macroscopic physical phenomena as the bulk sum of simple physical interactions between particles. As a very partial list, we mention macroscopic properties like pressure, temperature, volume, work, and energy of gasses in confined spaces, which can be modelled as the result of elastic collisions and other interactions between individual gas particles, and we also mention the phenomena of ferromagnetic and spontaneous magnetization, which can be modelled as arising from the magnetic spin of metal particles or magnetic domains.

Random surface models address a different set of physical or mathematical phenomena, namely the formation and characteristics of some surface that is governed by an underlying microscopic process. Physical examples of such include the interface between liquid and air on the surface of a body of liquid or the surface of a bubble; mathematical examples include the height function model discussed in the sequel as well height functions corresponding to other combinatorial models, such as domino tilings [Kas63, CEP96, CKP01], Young tableaux [LS77, VK77, PR07], and the six-vertex model [BCG16, CS16, RS17, Sri16]. These models study a discrete system that can be expressed as or identified with a height function or state $h : R \to E$ where R is a "nice" lattice graph (say, $R \subset \mathbb{Z}^d$ a simply connected subgraph), where the codomain E is a "nice" space such as $\mathbb{Z}$ or $\mathbb{R}$, and where the height function must satisfy certain model-specific constraints.

As mentioned above, in the $\mathbb{Z}$-homomorphism model that we study, we take $R \subset \mathbb{Z}^d$ a subgraph of the integer lattice endowed with nearest-neighbor edges and we take $E = \mathbb{Z}$, again with nearest-neighbor edges. In other words, whenever $x, y \in R$ satisfy $|x - y|_1 = 1$, then $|h(x) - h(y)| = 1$. This model is described in detail in Section 2.2.1 below. It also goes by other names in the literature, such as the $\mathbb{Z}^d$-indexed random walk in [Kah01] or the body-centered solid-on-solid model in [Bei77]. In two dimensions this model is equivalent to the six-vertex model; this connection is discussed further in Section 2.2.2 below.

Many other random surface models have been studied in the mathematical literature. Let us give a few examples, and refer the reader to [She05] and references therein for more. First, there is the Ising model (e.g. [DKS92, Cer06]), where the height values are in $E = \{-1, +1\}$.

Next, many tiling problems can be studied via random surface models. Examples include domino tilings in the square lattice $\mathbb{Z}^2$ (e.g. [Kas63, CEP96, CKP01]) , ribbon tilings in the square lattice (e.g. [She01, CL90]), and lozenge tilings in the triangular lattice (e.g. [Des98, LRS01, Wil04]). The height functions for domino and lozenge tilings take values in $\mathbb{Z}$, and those for ribbon tilings take values in a more general group.

The general question that we study is what a typical such height function "looks like" as the size of the domain tends to infinity. More specifically, one might hope to characterize the global or local asymptotic behavior of a typical height function, in terms of e.g. its overall scaling limit (as is studied in this dissertation) or the local statistics of short-range patterns in its values (as has been studied in famous examples like the arctic circle phenomenon for domino tilings of the Aztec diamond [CEP96]). The work recorded in this dissertation does not address local statistics for the $\mathbb{Z}$-homomorphism model, but below we discuss other results regarding local statistics from the literature, to better explain the limit shape phenomenon.

A major contribution to the study of random surfaces is the work by Sheffield [She05], which introduced and studied the class of simply attractive potentials. These are nearest-neighbor potentials where the energy term corresponding to a pair $x, y \in R$ is given by a convex function of the height difference, i.e. $V_{x,y}(h(x) - h(y))$ for some convex function $V_{x,y} : \mathbb{R} \to [0, \infty]$. The work also introduces two sub-classes: isotropic simply attractive potentials, where the functions $V_{x,y}$ are even (i.e. $V(-\eta) = V(\eta)$ for $\eta \in \mathbb{R}$) and are independent of the pair x, y, and Lipschitz simply attractive potentials, for which there exists a compact interval $K \subset \mathbb{R}$ such that $\eta \in \mathbb{R} \setminus K$ implies that $V_{x,y}(\eta) = \infty$ for every x, y. Note that the naïve attempt to represent the $\mathbb{Z}$-homomorphism model via such a potential fails to be simply attractive, because the nearest-neighbor potential terms are not convex; rather, $V_{x,y}(\eta) = 0$ for $\eta \in \{\pm 1\}$ and $V(\eta') = \infty$ for $\eta' \in \{0, \pm 2, \pm 3, \dots\}$, corresponding to the hard constraint $h(x) - h(y) \in \{\pm 1\}$. Indeed an alternate representation of the $\mathbb{Z}$-homomorphism model as a Lipschitz simply attractive potential is possible, e.g. using the bijection to the six-vertex model discussed in Section 2.2.2 below, but we shall not pursue that connection further at this point.

There are several families of results that one might reasonably hope to establish for different random surface models. For example, there are the aesthetically interesting and (hence) motivating examples of limit shape results, where one concludes that the overwhelming majority of states of the model are asymptotically close to a limit shape. An example is the arctic circle phenomenon exhibited by domino tilings of the Aztec diamond, as in [CEP96, JPS98]. The Aztec diamond is a lattice region comprising points $x \in \mathbb{Z}^2$ such that $|x|_1 \leq n$, with the extremal points $\{(\pm n, 0), (0, \pm n)\}$ truncated. One considers tilings of the faces in the dual graph by dominoes, i.e. rectangles of size 1×2. A typical such tiling is shown in Figure 1.1. Note the "frozen" regions at top, bottom, left, and right, and the non-frozen region in the center. The shape of the boundary heavily biases the uniform distribution on tilings in favor of those where the topmost domino to lies horizontally. Indeed, if the topmost two squares are not covered by a horizontal domino, then they must be included in two vertical dominoes instead. Then both of the two top sides of the square are forced to contain only vertical dominoes. This significantly reduces the number of possible tilings. However as one proceeds inward away from the corners, eventually the entropy of allowable domino orientations increases. It is shown in [CEP96] that as the size of the Aztec diamond tends to infinity, the border between the frozen and liquid regions tends toward the inscribed circle on the (rotated) square. The existence of this arctic circle is a stunning example of a limit shape.

Let us now turn to the $\mathbb{Z}$-homomorphism model. Even without a rigorous proof, occurrence of the arctic circle phenomenon is easy to observe empirically by sampling and visualizing states of the model. Figure 1.2 shows two states of the $\mathbb{Z}$-homomorphism that we study below. In both cases the height functions are constrained by boundary conditions similar to the Aztec diamond, and in each an arctic circle is visible. The height functions have extremal slope 1 near the corners, which is evident from the solid color in the left and right corners in each of the schematic pictures on the left side, and from the checkerboard

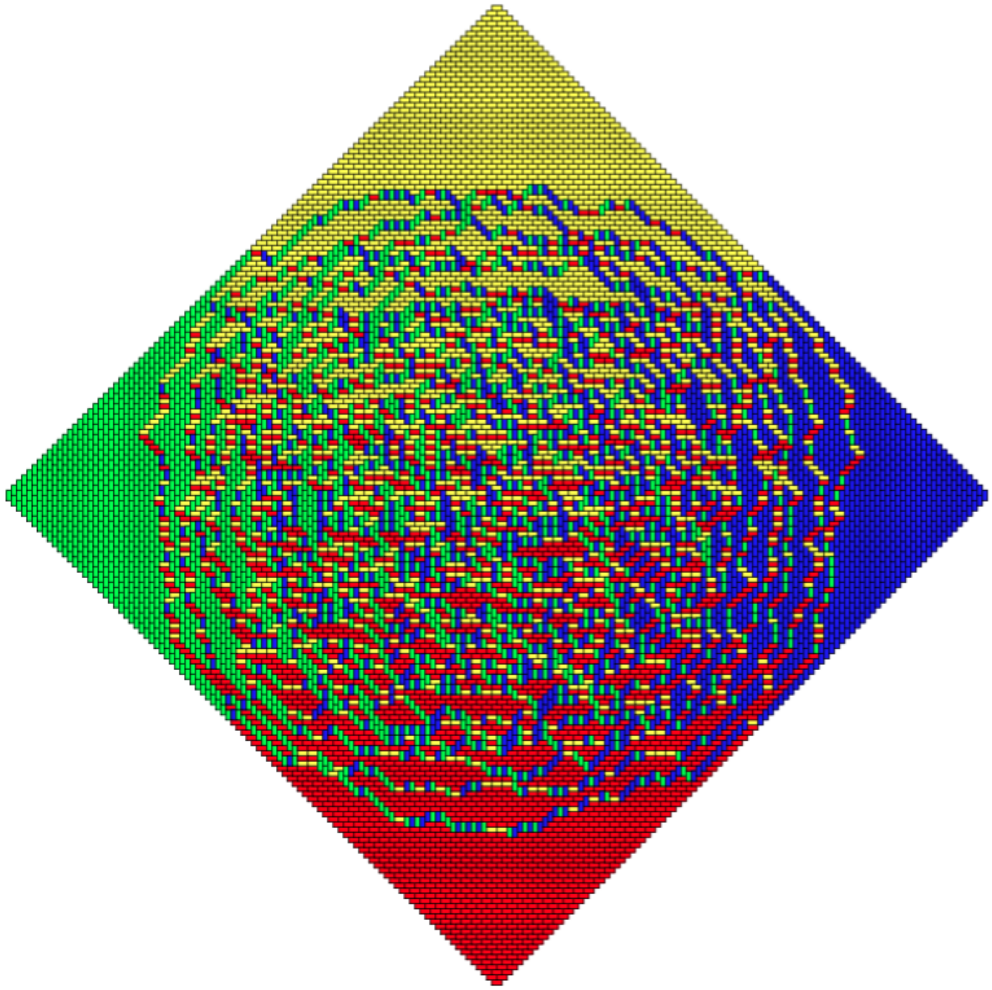
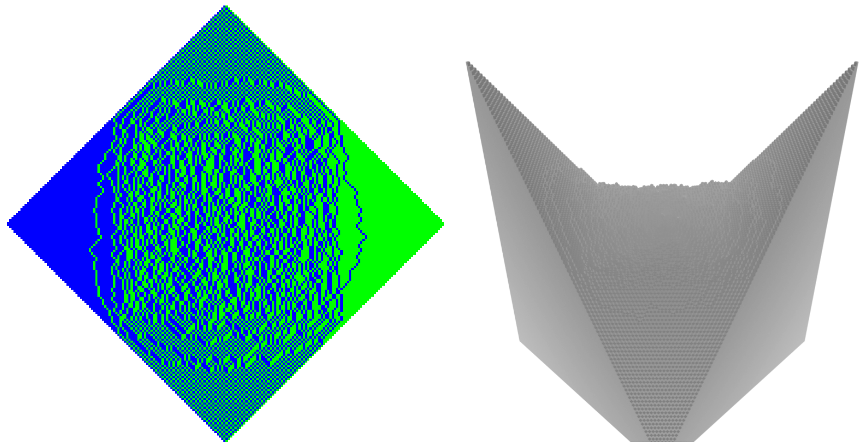


Figure 1.1: A typical domino tiling of the Aztec diamond. Dominoes are colored according to both their orientation (horizontal or vertical) and their parity (whether the left or top square of the domino lies on a "black" or "white" square, assuming an underlying checkerboard coloring of the lattice). The arctic circle is clearly visible as the boundary between frozen regions in the four corners (where every domino is of the same color) and the non-frozen central region (where the colors are randomly mixed). The arctic circle is a limit phenomenon: if the analogous picture is drawn for a much larger Aztec diamond, scaled to the same dimensions on page, then the random law of the color at a given point in the picture either is close to a delta law (if the point is outside the arctic circle) or has significantly higher entropy (inside).

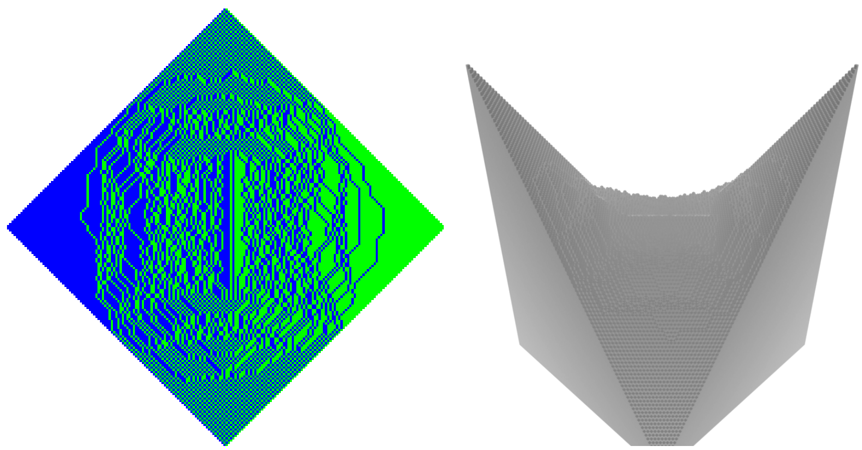
pattern exhibited on the top and bottom corners (assuming that the picture is displayed with a sufficient level of detail to make the pattern visible). One does not see countervailing slopes until well away from the corners, where colors mix more freely. Note that the local distribution of height changes across an edge in the central region are not uniform, since the boundary still has an effect even at long range. The bias towards the "preferred" configuration in each of the four corners is more substantial when the potential is larger. The emergence of the solid-colored region in the center of the left picture in Figure 1.2b is intriguing, and it suggests that the limit shape might have interesting macroscopic features. The limit shape has not been investigated in detail as of the time of writing.

Another example of a limit shape is a limiting height profile. Given instances of the discrete model on domains R_n that converge under scaling in some suitable sense to a limit domain R, e.g. $R_n \subset \mathbb{Z}^d$ and $R \subset \mathbb{R}^d$ such that $\frac{1}{n}R_n \to R$ in the Hausdorff metric, a limiting height profile $h : R \to \mathbb{R}$ is a continuous function such that a typical discrete height function $h_n : R_n \to \mathbb{Z}$ converges in a suitable sense, i.e. $\frac{1}{n}h_n \to h$ uniformly on R_n. The existence of limiting height profiles for the $\mathbb{Z}$-homomorphism model is also evident from Figure 1.2. The work in Chapter 2 below, combined with a result from [LT20] cited below in the current chapter, rigorously establishes the existence of a limiting height profile.

Several technical results serve to establish the existence of a limit shape, to characterize it, and to quantify or characterize the rate of convergence to the limit shape and the vanishing probability that a rescaled height function differs substantially from the limit shape. First is the concentration inequality (see Theorem 10 below), which establishes a sort of "finite limit shape," or more precisely: for large n, most height functions, either with respect to the uniform measure or with respect to the perturbed measures introduced below, assign to each point in the domain a height value close to the expected value, under the same measure respectively. For the rest of this paragraph we assume the results are stated in terms of the uniform measure, i.e. without random (or non-random) perturbations; the results are analogous when a perturbed measure is used. The concentration inequality is interesting in



(a) Smaller random potential. (ω_e) are i.i.d. with $\mathbb{P}(\omega_e = \frac{1}{4}) = \mathbb{P}(\omega_e = -\frac{1}{4}) = \frac{1}{2}$.



(b) Larger random potential. (ω_e) are i.i.d. with $\mathbb{P}(\omega_e = 8) = \mathbb{P}(\omega_e = -8) = \frac{1}{2}$.

Figure 1.2: Two randomly sampled height functions with the same boundary values, each pictured in two representations. The height functions are sampled from distributions subject to the same boundary constraints but perturbed by respectively a smaller (less influential) and larger (more influential) random potential. On the left are schematic representations, with blue points representing edges that go down as one moves right or up across edges in the 2d domain R_n, and green points representing edges that go up. On the right is a 3d representation, with heights indicated via the visual height of a square column above the face $(x, y) \in R_n$. In other words, the pictures on the right are plots of the 3d surfaces $\{(x, y, h(x, y)) \colon (x, y) \in R_n\}$.

its own right, but on its own it has a significant weakness: the putative limit shapes, namely the expected height values, need not converge as the system size $n \to \infty$. The next two results remedy this, by dealing directly with the limiting process. The profile theorem gives an estimate for the number of height functions that are uniformly close after rescaling to an arbitrary limiting height profile. The variational principle builds on this to identify the total number of height functions satisfying a boundary constraint. This number is characterized by means of a variational problem, namely optimizing a particular integral function over all limiting height profiles. The number thus calculated coincides with the number of height functions predicted by the profile theorem to lie close to an optimizer, which suffices to show that a positive fraction of height functions sit close to an optimizer, but no more. Finally the large deviations principle gives bounds on the number of height functions that are not close to an optimizer. This finally provides a weak limit shape result; the weakness is that *a priori* there could be many optimizers, in which case there is not a unique limit shape, but rather a sort of "limiting family." The question of uniqueness of the optimal limiting height profile is answered in the affirmative by the work of Lammers and Tassy in [LT20], at least in the absence of random perturbations. There the surface tension function used to define the variational principle objective function (cf. Definition 10 and Theorem 4 below) is shown to be strictly convex. As a matter of ordinary convex analysis, and as discussed there and in Chapter 2 below, strict convexity ensures uniqueness of the optimal height function.

The variational principle, large deviations principle, large deviations principle, and concentration inequality for the $\mathbb{Z}$-homomorphism model with uniform measure are all known in the existing literature; see for example [KMT20a] and references cited therein. Our contribution in this area consists in extending these results to the randomly perturbed case, and we present that in Chapter 2. Our purpose in doing so is to demonstrate the robustness of these technical tools and of the known methods for proving them. That the statements and proofs carry over into the perturbed setting without very much difficulty is evidence of robustness and universality; the source and presence of some difficulties is also of interest to understand the limitations of the statements and proofs.

1.2 Long-Range Percolation

The second model studied in this dissertation is the long-range percolation model. This is a random graph model, with vertex set $\mathbb{Z}^d$ and a random set of undirected edges. The probability distribution of the edge set depends upon the geometry of the underlying lattice graph: any two distinct lattice points are potentially connected by an edge, but the probability that a particular edge is present in the edge set vanishes as the distance between the points (measured via a norm $|\cdot|$ on $\mathbb{Z}^d$, e.g. the ℓ^1 norm) tends to infinity. Edges between distinct unordered pairs of endpoints are independently present or absent.

Five regimes of typical asymptotic behavior have been identified in the case that the edge inclusion probabilities are translation invariant with power-law decay. To be precise: for $x, y \in \mathbb{Z}^d$, let $\mathfrak{p}(x-y) \in [0,1]$ denote the probability that the undirected edge between x and y is included in the graph. Assume that for some $s \in (d, 2d)$, the limit $\beta := \lim_{|x|\to\infty} \mathfrak{p}(x)|x|^s$ exists with $\beta \in (0, \infty)$. Then, depending on how the exponent s relates to the dimension d, the model may be in any of 5 distinct regimes, under which the distance exhibits different asymptotic growth rates:

- The case $s < d$: In this regime, the long-range percolation has a.s. finite diameter. For example, in Benjamini and Berger [BB01] the following observation is recorded: consider the long-range percolation process on the (one-dimensional) cycle graph $\mathbb{Z}/n\mathbb{Z}$, i.e. the graph with vertex set $\{0, \ldots, n-1\}$ endowed with the n undirected edges $\{(i, i+1) \colon i = 0, \ldots, n-2\} \cup \{(n-1, 0)\}$. This long-range percolation graph stochastically dominates the Erdős–Rényi graph $G(n,p)$ with edge probability $p = \beta n^{-s}$. In this regime the Erdős–Rényi graph has

$$\lim_{n\to\infty} P\big(\operatorname{diam} G(n,p) < C\big) = 1 \tag{1.1}$$

for some constant $C > 0$ depending on d, s, and β but not on n; see e.g. [Bar16, New18] and references therein. As adding edges can only decrease the diameter of a graph, the same long-range percolation graph with $d = 1$ and $s \in (0, 1)$ also has bounded diameter with high probability.

Moreover, in the $s < d$ regime the diameter of the long-range percolation graph on the infinite lattice $\mathbb{Z}^d$ with parameter $s < d$ is almost surely finite. Indeed, Benjamini, Kesten, Peres, and Schramm [BKP04] showed using the notion of stochastic dimension that the diameter of the graph is almost surely $\lceil d/(d - s) \rceil$.

- The case $s = d$: Distances grow sub-logarithmically, but just barely so. Specifically, Coppersmith, Gamarnik, and Sviridenko [CGS02] showed there exist constants $c_1, c_2 > 0$ such that

$$\lim_{n\to\infty} P\left(\frac{c_1 \log n}{\log \log n} \le D_n \le \frac{c_2 \log n}{\log \log n} \right) = 1, \tag{1.2}$$

where D_n denotes the diameter of the long-range percolation graph on the cube of side length n.

- The case $d < s < 2d$: Distances in the graph grow polylogarithmically, i.e. the distance $D(0, x)$ is asymptotic to $(\log |x|)^\Delta$ for an explicitly known exponent $\Delta > 0$. Because this is the regime that the work below addresses, the relevant prior results are summarized in greater detail after a brief discussion of the remaining two regimes.

- The case $s = 2d$: The distance grows sublinearly in the underlying norm. Indeed Coppersmith, Gamarnik, and Sviridenko [CGS02] established the upper bound

$$\lim_{n\to\infty} P\Big(D(n) \le n^\theta \Big) = 1, \tag{1.3}$$

where $D(n)$ is the diameter of the long-range percolation graph on the cube (in any dimension $d \ge 1$) and where the exponent $\theta \in (0, 1)$ depends on d and β. The same work established a similar lower bound

$$\lim_{n\to\infty} P\Big(D(n) \ge n^{\theta'} \Big) = 1, \tag{1.4}$$

but only in the case $d = 1$ and $\beta < 1$, and with exponent $\theta' \leq \theta$. Ding and Sly [DS15] extended the lower bound, proving that for $d = 1$ and for any $\beta > 0$,

$$\lim_{n\to\infty} P\Big(c_1 n^{\theta''} \leq D(n) \leq c_2 n^{\theta''}\Big) = 1, \tag{1.5}$$

for some $c_1, c_2 > 0$ and some $\theta'' \in (0,1)$. Note in particular that the exponent θ'' is the same in both the lower and upper bounds. Since the existence of the limiting exponent θ'' is proved via subadditivity, no explicit expression for θ'' is known.

- The case $s > 2d$: The distance scales linearly with the underlying norm. Indeed this was known for the case $d = 1$ from Benjamini and Berger [BB01], where proof was given that

  $$\lim_{n\to\infty} P\Big(D(n) \geq cn\Big) = 1, \tag{1.6}$$

  for some $c > 0$ depending on s and β. The linear lower bound was extended to dimensions $d \geq 1$ by Berger [Ber04]. In the case we consider, where nearest-neighbor edges are deterministically present, the corresponding upper bound, i.e. $D(n) \leq Cn$ for some $C > 0$ depending on the norm, is trivial.

  Without assuming that all nearest-neighbor edges are present a.s., the upper bound depends on the percolation structure of the graph. For example, in dimension 1 percolation always fails if $s > 2$ and percolation depends on β for $s \in (1,2]$. When percolation failed, and hence when every graph component is a.s. finite, distance asymptotics are not a meaningful object to study. In higher dimensions, if $\min_{x\colon |x|_1=1} \mathfrak{p}(x)$ is large enough, then percolation occurs a.s. In that case a linear upper bound for the distance (restricted to points in the same infinite cluster) follows by comparison to nearest-neighbor bond percolation; the corresponding bound for the nearest-neighbor model is due to Antal and Pisztora [AP96]. Furthermore it is conjectured in [Ber04] that a linear upper bound holds for the distance $D(0,x)$, restricted to points x lying in the same cluster as 0, whenever the model achieves percolation (even without comparison to the nearest-neighbor percolation). However this conjectured bound appears to still be open.

As mentioned above, this work addresses the regime with $d < s < 2d$. Let us review some developments relating to the graph distance in this regime. The following bounds were provided in the one-dimensional case $d = 1$ by Benjamini and Berger [BB01] and by a similar proof for $d \geq 1$ by Coppersmith, Gamarnik, and Sviridenko [CGS02]: there exist constants $c, C > 0$ and $\delta > 1$ depending on s, d, and β such that

$$\lim_{n\to\infty} P\Big(c \log n \leq D(n) \leq C(\log n)^{\delta}\Big) = 1. \tag{1.7}$$

These bounds for the distance were improved to poly-logarithmic and the optimal exponent was derived explicitly by Biskup [Bis04]. There it was proved that for any $\varepsilon > 0$,

$$\lim_{|x|\to\infty} P\Big((\log|x|)^{\Delta-\varepsilon} \leq D(0,x) \leq (\log|x|)^{\Delta+\varepsilon}\Big) = 1, \tag{1.8}$$

where $\Delta = \Delta(s,d) := 1/\log_2(2d/s) \in (1,\infty)$. The corresponding improvement for the diameter is due to Biskup [Bis11a], where it was proved that, for any $\varepsilon > 0$ and with Δ as before,

$$\lim_{n\to\infty} P\Big((\log n)^{\Delta-\varepsilon} \leq D(n) \leq (\log n)^{\Delta+\varepsilon}\Big) = 1. \tag{1.9}$$

Next, the gap in the exponents was removed by Biskup and Lin [BL19], with the conclusion that there exists constants $c, C > 0$ such that

$$\lim_{|x|\to\infty} P\Big(c(\log|x|)^{\Delta} \leq D(0,x) \leq C(\log|x|)^{\Delta}\Big) = 1. \tag{1.10}$$

Chapter 3 of this dissertation tightens the result further. The main result of that chapter, Theorem 11, states the following: Recall that $\beta = \lim_{x\to\infty} \mathfrak{p}(x)|x|^s \in (0,\infty)$. For almost every $\beta \in (0,\infty)$, there is a function $\phi_\beta : (0,\infty) \to (0,\infty)$ such that, for any $\varepsilon > 0$,

$$\frac{1}{r^d}\,\#\left\{x \in B(0,r)\colon \left|\frac{D(0,x)}{\phi_\beta(r)(\log r)^{\Delta}} - 1\right| > \varepsilon\right\} \underset{r\to\infty}{\overset{P}{\longrightarrow}} 0, \tag{1.11}$$

where $B(0,r) := \{x \in \mathbb{Z}^d \colon |x| < r\}$, where $\#$ denotes the set cardinality, and where $\xrightarrow{P}$ denotes convergence in probability. Moreover, the function ϕ_β is continuous, bounded away from 0 and from ∞, and log-log-periodic, in the sense that $\phi_\beta(r^\gamma) = \phi_\beta(r)$, where $\gamma = s/2d \in (\frac{1}{2}, 1)$.

Let us make a few observations to make the conclusion (1.11) more clear. First note that both $r \mapsto \phi_\beta(r)$ and $r \mapsto \log r$ are slowly varying, in the sense that for any $a > 0$, $\lim_{r\to\infty} \phi_\beta(ar)/\phi_\beta(r) = \lim_{r\to\infty}(\log ar)/(\log r) = 1$. Indeed, to deduce the limit for $\phi_\beta(r)$, use log-log periodicity: fix $n \in \mathbb{Z}$ such that $r_0 := r^{\gamma^n} \in [\mathrm{e}^\gamma, \mathrm{e})$. Then $\phi_\beta(r) = \phi_\beta(r_0)$ and $\phi_\beta(ar) = \phi_\beta(a^{\gamma^n} r_0)$. Since $n \to \infty$ as $r \to \infty$ and $a^{\gamma^n} \to 1$ as $n \to \infty$, the desired limit follows by continuity of $r \mapsto \phi_\beta(r)$. The corresponding result for $\log(r)$ follows from an easy calculation. Because these functions are slowly varying, one can ignore constant multiples in their argument when taking limits. Next, recall that the cardinality of $B(0,r)$ is bounded above and below by a constant times r^d (e.g. because the cardinality of the ℓ^2-ball is likewise bounded and all finite-dimensional norms are equivalent). Moreover, for small $\delta > 0$, the fraction of points $x \in B(0,r)$ with $|x| \geq \delta r$ tends to 1 as $r \to \infty$. So, the result can be paraphrased to say that as $r \to \infty$, the fraction of points x with $|x| < r$ for which the ratio $D(0,x)/[\phi_\beta(|x|)(\log|x|)^\Delta]$ differs from one by more than ε tends to 0, in probability.

It is an ongoing project [BK21] to further characterize the function $\phi_\beta(r)$. In [BL19] a continuum version of the long-range percolation model is studied. It is shown that in the continuum version, for almost every $x \in \mathbb{R}^d$ one has $D(0,rx)/[\phi_\beta(r)(\log r)^\Delta] \to 1$ in probability as $r \to \infty$. There it was conjectured that the function $\phi_\beta(r)$ for the continuum model was constant as a function of r. However the results in the preprint [BK21] imply that $\phi_\beta(r)$ is not generally constant. There the discrete long-range percolation model that was described above is considered, i.e. with parameters $s \in (d, 2d)$ and $\beta > 0$ and with edge inclusion probability $\mathfrak{p}(x)$ asymptotic to $\beta|x|^{-s}$ for all $|x|_1 > 1$. It is shown that as $\beta \to \infty$, the limiting function $\phi_\beta(r)$ tends to an explicit limit $\phi_\infty(r)$ locally uniformly in r, and the limit $\phi_\infty(r)$ is not constant in r.

CHAPTER 2

The $\mathbb{Z}$-homomorphism model

As mentioned in Chapter 1 above, the first model studied in this dissertation is the $\mathbb{Z}$-homomorphism model, a discrete random surface model. The underlying combinatorial mode is quite simple, but we consider an additional random perturbation in the form of the random potential ω introduced below. Our purpose is to study the limiting behavior of this model, and in particular the limit shape phenomenon that occurs with respect to the asymptotic height profile under the proper scaling limit. We shall give an informal description of the model and of the main results in the next few paragraphs. After that, and after introducing some preliminary notations and definitions, we state the main results formally. Finally we give rigorous proofs of the claimed results.

As was described in Chapter 1, the underlying combinatorial model is graph homomorphisms from domains $R_n \subset \mathbb{Z}^m$ to $\mathbb{Z}$, where both $\mathbb{Z}^m$ and $\mathbb{Z}$ are endowed with the nearest-neighbor edge structure. The subset R_n is assumed to be a subgraph induced by its vertices, in the sense that whenever $x, y \in R_n$ and x and y are adjacent in $\mathbb{Z}^m$, then the edge between x and y is included in the subgraph R_n. We also impose a parity condition, requiring that the graph homomorphisms under study must preserve parity. Indeed, both $\mathbb{Z}^m$ and $\mathbb{Z}$ are bipartite, with the even elements (those whose ℓ_1 norm is even) and the odd elements (the rest) forming the two halves of the graph, both in $\mathbb{Z}^m$ generally and in $\mathbb{Z}$ in particular. A graph homomorphism must therefore either map all even points in $R_n \subset \mathbb{Z}^m$ to even points in $\mathbb{Z}$ and likewise map odd points to odd points, or else it must map even points to odd and odd to even. There are obvious bijections between the two classes of homomorphisms (e.g. the map taking $h : R_n \to \mathbb{Z}$ to the homomorphism $R_n \ni z \mapsto h(z) + 1$), and our work is technically easier if we treat only the parity-preserving case. The main reason that this restriction helps is as follows. Consider subdomains $R' \subset R$, with R perhaps much larger. We start with a graph homomorphism $h' : R' \to \mathbb{Z}$ on R' and boundary height values $h_{\partial R} : \partial R \to \mathbb{Z}$ on ∂R. If both h' and $h_{\partial R}$ preserve parity, or if both invert parity, then there may exist a compatible extension to the entirety of R, but if h' preserves parity and $h_{\partial R}$ inverts (or vice versa), then automatically no extension is possible. (See Theorem 6, the Kirszbraun theorem for $\mathbb{Z}^m$, for the full conditions on extensibility of graph homomorphisms.) To avoid qualifying most formal statements with conditions like "assuming that either both h' and $h_{\partial R}$ preserve parity or that both invert parity," we simply restrict our attention to the parity-preserving case. We call parity-preserving graph homomorphisms height functions.

The main results of this chapter partially describe the asymptotic structure of the set of height functions on R_n subject to prescribed boundary constraints. The focus of the first main result, the profile theorem, is to measure how many height functions are close to an admissible asymptotic height function (after rescaling). In other words, we want to estimate the probability mass of the ball

$$B(R_n, h_R, \delta) := \Big\{ h_{R_n} : R_n \to \mathbb{Z} \,\Big|\, \max_{z \in R_n} \big| \tfrac{1}{n} h_{R_n}(z) - h_R(\tfrac{1}{n} z) \big| < \delta \Big\}, \tag{2.1}$$

measured in terms of a random probability measure μ_ω that incorporates the random potential ω. See Section 2.2.5 for further details. For the sake of simplicity in this introductory text, let us ignore the contribution of the random potential ω. Then the probability measure on sets of height functions is uniform, and the profile theorem gives estimates on the number of height functions in the ball $B(R_n, h_R, \delta)$. In particular, the profile theorem states that

$$\operatorname{Ent}_{R_n}\big(B(R_n, h_R, \delta)\big) \approx \operatorname{Ent}_R(h_R), \tag{2.2}$$

where $\operatorname{Ent}_{R_n} := -\frac{1}{|R_n|} \log |B(R_n, h_R, \delta)|$ is the microscopic entropy of the ball (see Definition 8), where $\operatorname{Ent}_R(h_R) := \frac{1}{|R|} \int_R \operatorname{ent}(\nabla h_R(x))\, dx$ is the macroscopic entropy of the asymptotic height function (see Definition 11), and where ent : $[-1,1]^m \to \mathbb{R}$ is the local surface tension (see Definition 10).

Next, Theorem 4, the variational principle, estimates the microscopic entropy of the entire set of height functions satisfying a boundary condition. The result is given by optimizing the macroscopic entropy integral over admissible asymptotic height functions:

$$\mathrm{Ent}_{R_n}\big(\big\{h_{R_n} : R_n \to \mathbb{Z} \,\big|\, h_{R_n}|_{\partial R_n} = h_{\partial R_n}\big\}\big) = \inf\big\{\mathrm{Ent}_R(h_R) \,\big|\, h_R|_{\partial R} = h_{\partial R}\big\}, \tag{2.3}$$

where the microscopic domains R_n and their boundary values $h_{\partial R_n}$ are assumed to converge (in a suitable scaling limit) to the limiting domain $R \subset \mathbb{R}^m$ and boundary values $h_{\partial R} : \partial R \to \mathbb{R}$ respectively.

The large deviations principle gives probability estimates for sets of height functions that are geometrically distant from to the optimizer of the variational principle. The conclusion is that the probability measures μ_ω over height functions on R_n satisfy a large deviations principle, with rate function essentially given by the macroscopic entropy integral. See Theorem 5 for a formal statement.

A consequence of the large deviations principle is that the probability of the collection of all height functions that differ from the entropy-optimizing function in (2.3) above vanishes exponentially as the system size goes to infinity. Note that the infimum in (2.3) admits a unique minimizer, at least in the absence of random potential, since the local surface tension ent is strictly convex (convexity is the content of Lemma 11 below; strict convexity for the uniform case is due to [LT20]). In the case with the random potential, there still exists at least one minimizer by convexity alone. We conjecture that strict convexity and hence unique of the minimizer can be obtained by methods similar to those in [LT20].

The last main result of this chapter is a concentration inequality for the measures μ_ω. In many ways this is a simpler result than the three mentioned above, and in fact it could be used in proving them. However it is not necessary to use the concentration inequality in those other proofs, and so we do not make reference to it until after completing proof of the profile theorem, the variational principle, and the large deviations principle.

The concentration inequality states that the measures μ_ω concentrate around their mean, uniformly over the space R_n. Apart from being somewhat simpler to prove, it is also easier to extract quantitative results on the rate of convergence from the concentration inequality, although applications of such bounds are beyond the scope of the work presented.

Outline of the rest of the chapter

Having summarized the main results, we shall proceed in the rest of the chapter as follows:

- In Section 2.1, we introduce notation and conventions used throughout the rest of the chapter.
- In Section 2.2, we introduce formal definitions of the model, the random measures μ_ω, the quantities like $\mathrm{Ent}_{R_n}(\cdot)$ referenced above, and other necessities.
- In Section 2.3, we formally state the first batch of results: the profile theorem, the variational principle, and the large deviations principle.
- In Section 2.4, we derive intermediate results related to the entropy and local surface tension.
- In Section 2.5 we prove the profile theorem.
- In Section 2.6 we prove the variational principle.
- In Section 2.7 we prove the large deviations principle.
- In Section 2.8 we discuss the strategy for proving the concentration inequality and make some necessary definitions.
- In Section 2.9 we give the proof of the concentration inequality.

2.1 Notation and conventions

Throughout the rest of this chapter, we use the following notations and conventions. The first several items in the listing are relatively standard, but we list them in hopes of reducing ambiguity.

- Usually $m \in \mathbb{N} \cup \{0\}$ denotes the dimension of an ambient space, e.g. $\mathbb{Z}^m$ or $\mathbb{R}^m$.
- For $p \in [1, \infty]$, $|\cdot|_p$ denotes the usual p-norm on $\mathbb{R}^m$ or $\mathbb{Z}^m$.
- For vectors $x, y \in \mathbb{R}^m$, $x \cdot y$ denotes the usual dot product, i.e. the inner product associated to the norm $|\cdot|_2$.
- Usually $s \in [-1, 1]^m$ denotes a "slope," i.e. a vector with $|s|_\infty \leq 1$. Note that the ℓ^∞ condition should not be surprising, since the Lipschitz condition for height functions is stated with respect to the ℓ^1 norm.
- $|A|$ denotes either the cardinality or the Lebesgue measure of the set A, depending on whether the set is finite (or rarely, countably infinite) or is a subset of some space $\mathbb{R}^m$.
- Given a function $f : A \to B$ and a subset $A' \subseteq A$ of the domain, we write $f|_{A'} : A' \to B$ for the restriction of f to A'.
- Given vertices u, v in a graph, particularly the graph $\mathbb{Z}^m$, we write $u \sim v$ to mean that u and v are adjacent.

The following notations are less standard, but they will prove useful in the upcoming exposition.

- $S_n := \{-n, -(n-1), \ldots, n-1, n\}^m \subset \mathbb{Z}^m$ denotes a hypercube in the lattice, centered at the origin.
- For $S \subset \mathbb{Z}^m$, $\partial S := \{z \in S \mid \exists \tilde{z} \in \mathbb{Z}^m \setminus S,\, \tilde{z} \sim z\}$ is the interior boundary of S.

- $e_{zz'}$ is the unoriented edge between neighbors $z \sim z'$ in $\mathbb{Z}^m$.

- For $h : \mathbb{Z}^m \to \mathbb{Z}$ and $e = e_{zz'} \in E(\mathbb{Z}^m)$, we abuse notation and write $h(e)$ for the edge $e_{h(z),h(z')} \in E(\mathbb{Z})$.

- τ_w denotes the shift by $w \in \mathbb{Z}^m$ on edges of the graph $\mathbb{Z}^m$. That is, $\tau_w e_{zz'} = e_{z+w,z'+w}$.

- $\theta(\varepsilon) \geq 0$ denotes a smooth function with $\lim_{\varepsilon\to 0} \theta(\varepsilon) = 0$. We explain the $\theta(\varepsilon)$ asymptotic notation further in Section 2.1.1 below.

2.1.1 Asymptotic notation

In this section we introduce a notation for asymptotic error. In comparison to the familiar Landau big-O notation, our θ-notation further abstracts away the rate of convergence of the error, but it still makes explicit the dependence on parameters. For this purpose we write $\theta_\alpha(\delta)$ for a family of unspecified functions, parameterized by a symbol α, such that $\theta_\alpha(\delta) \to 0$ at a rate depending on the value of the parameter α. That is, for any $\varepsilon > 0$ and any admissible parameter value α, there exists $\delta_0 = \delta_0(\alpha) > 0$ such that $0 < \delta < \delta_0$ implies $\theta_\alpha(\delta) < \varepsilon$.

Extending the above notation, we frequently replace the single parameter α by a list of parameters $\alpha, \beta, \gamma, \ldots$. For example, we might write an identity like

$$\begin{aligned}\min_{h_R \in M(R,h_{\partial R})} \operatorname{Ent}_R(h_R) = \operatorname{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}, \delta)\big) \\ + \theta_{m,R,h_{\partial R},R_n,h_{\partial R_n}}(\delta) \\ + \theta_{m,R,h_{\partial R},R_n,h_{\partial R_n},\delta}(\tfrac{1}{n}).\end{aligned} \tag{2.4}$$

The identity states that the two entropy terms on the first line differ by a small amount; the difference vanishes as δ and $\frac{1}{n}$ go to zero, and the rate of convergence depends on several parameters. The "$\theta(\delta)$" term depends on the parameters from the setting, namely the ambient dimension m, the region R, the height function h_R of interest, and the corresponding discrete objects R_n and h_{R_n}. The "$\theta(\frac{1}{n})$" term depends on these parameters along with the value of δ. We find that listing out the setting parameters m, R, h_R, R_n, and h_{R_n} makes the expression harder to read. So for the rest of the chapter we suppress these parameters from the subscripts of θ terms. Under this convention (2.4) becomes:

$$\min_{h_R \in M(R, h_{\partial R})} \mathrm{Ent}_R(h_R) = \mathrm{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}, \delta)\big) + \theta(\delta) + \theta_\delta(\tfrac{1}{n}). \tag{2.5}$$

As mentioned above, the advantage of our θ notation is that it abstracts away the exact rates of convergence, but leaves explicit the dependencies between parameters. For example, suppose we want to make the error in approximation in (2.4) to be less than ε. We should first choose δ so that (say) $\theta(\delta) < \frac{1}{2}\varepsilon$, then choose n depending on δ (and on the suppressed parameters m, R, etc.) so that $\theta_\delta(\frac{1}{n}) < \frac{1}{2}\varepsilon$.

2.2 Formal definition of model

In this section we formally describe the model under study and introduce related notation. The setting, notation, and main results are similar to those of [KMT20a] and [KMT20b].

2.2.1 Basic definitions

Throughout the sequel, we fix a dimension $m \in \mathbb{N}$, a macroscopic domain $R \subset \mathbb{R}^m$, and a sequence of microscopic domains $R_n \subset \mathbb{Z}^m$ satisfying these assumptions:

Assumption 1 (Assumptions on domain R and R_n)**.** *We assume that $R \subset \mathbb{R}^m$ is compact and connected, that R is the closure of its interior, and that the boundary of R has zero Lebesgue measure. We assume that $R_n \subset \mathbb{Z}^m$ is contained in R after rescaling, i.e. that $\frac{1}{n}R_n \subset R$, although this is just a simplifying assumption. Moreover, we assume that $\frac{1}{n}R_n \to R$ in the Hausdorff metric, i.e. the metric on $\{A \subset \mathbb{R}^m\}$ defined by*

$$d_H(A, B) := \left(\sup_{x\in A}\inf_{y\in B}|x-y|_1\right) \vee \left(\sup_{y\in B}\inf_{x\in A}|x-y|_1\right). \tag{2.6}$$

Now, we define precisely the height functions in our model.

Definition 1 (Height function). A *height function* on R_n is a parity-preserving graph homomorphism $h_{R_n} : R_n \to \mathbb{Z}$. In other words, if $z, w \in R_n$ and $z \sim w$, then $|h_{R_n}(z) - h_{R_n}(w)| = 1$, and for any $z = (z_1, \ldots, z_m) \in R_n$,

$$h_{R_n}(z) \equiv z \pmod 2\,, \quad \text{i.e. } h_{R_n}(z) \equiv \sum_{i=1}^{m} z_i \pmod 2\,. \tag{2.7}$$

The condition (2.7) states that a height function preserves the parity of the lattice $\mathbb{Z}^m$. Indeed, every graph homomorphism either preserves parity at all points or inverts parity at all points, since the source space $\mathbb{Z}^m$ and the target space $\mathbb{Z}$ are both bipartite. Our main results are also valid without the parity-preserving condition, but for the same reasons as outlined in [KMT20a, Section 2.1] we import the parity-preserving condition for simplicity.

We introduce the following symbols to refer to sets of height functions:

Definition 2 (Sets of height functions). Let R_n be a microscopic domain as above, let $h_{R_n} : R_n \to \mathbb{Z}$ be a boundary height function, and let $\delta > 0$. We define:

$$M(R_n) := \left\{ h_{R_n} : R_n \to \mathbb{Z} \,\middle|\, h_{R_n} \text{ is a height function} \right\}, \tag{2.8}$$

$$M(R_n, h_{\partial R_n}) := \left\{ h_{R_n} \in M(R_n) \,\middle|\, h_{R_n}|_{\partial R_n} = h_{\partial R_n} \right\}, \tag{2.9}$$

$$M(R_n, h_{\partial R_n}, \delta) := \left\{ h_{R_n} \in M(R_n) \,\middle|\, \sup_{z \in \partial R_n} |h_{R_n}(z) - h_{\partial R_n}(z)| < \delta n \right\}, \text{ and} \tag{2.10}$$

$$B(R_n, h_R, \delta) := \left\{ h_{R_n} \in M(R_n) \,\middle|\, \sup_{z \in R_n} |h_R(\tfrac{1}{n} z) - \tfrac{1}{n} h_{R_n}(z)| < \delta \right\}. \tag{2.11}$$

In the last definition, the expression "$h_R(\frac{1}{n}z)$" makes sense because of the assumption that $\frac{1}{n} R_n \subset R$ in Assumption 1.

The limiting object for convergent sequences of height functions is:

Definition 3 (Asymptotic height function). We call a function $h_R : R \to \mathbb{R}$ an *asymptotic height function* if h_R is Lipschitz with Lipschitz constant at most 1, with respect to the ℓ^1-norm on $\mathbb{R}^m$; that is, if

$$\mathrm{Lip}(h_R) := \sup_{x \neq y \in R} \frac{|h_R(x) - h_R(y)|}{|x - y|_1} \leq 1\,. \tag{2.12}$$

Likewise, if $h_{\partial R} : \partial R \to \mathbb{R}$ is 1-Lipschitz (with respect to the ℓ^1-norm), we call $h_{\partial R}$ an *asymptotic boundary height function*.

The limit of height functions is defined as follows.

Definition 4 (Convergence of height functions). Given a sequence of height functions $h_{R_n} : R_n \to \mathbb{Z}$ and an asymptotic height function $h_R : R \to \mathbb{R}$, we say that h_{R_n} converges in the scaling limit to h_R if

$$\lim_{n\to\infty} \sup_{z\in R_n} \sup_{\substack{x\in R \\ |x-\frac{1}{n}z|_1 \le d_n}} \Big| \frac{1}{n} h_{\partial R_n}(z) - h_{\partial R}(x) \Big| = 0 \,, \tag{2.13}$$

where $d_n := d_H(\frac{1}{n}R_n, R)$.

Finally, we define the following sets of asymptotic height functions:

Definition 5 (Sets of asymptotic height functions). Let $R \subset \mathbb{R}^m$ be a domain satisfying Assumption 1, let $h_{\partial R} : \partial R \to \mathbb{R}$ be an asymptotic boundary height function, and let $\delta > 0$. We define:

$$M(R) := \big\{ h_R : R \to \mathbb{R} \,\big|\, h_R \text{ is an asymptotic height function} \big\} \,, \tag{2.14}$$

$$M(R, h_{\partial R}) := \big\{ h_R : R \to \mathbb{R} \,\big|\, h_R|_{\partial R} = h_{\partial R} \big\} \,, \tag{2.15}$$

$$M(R, h_{\partial R}, \delta) := \big\{ h_R : R \to \mathbb{R} \,\big|\, \forall x \in \partial R \,,\, |h_R(x) - h_{\partial R}(x)| \le \delta \big\} \,, \text{ and} \tag{2.16}$$

$$B(R, \tilde{h}_R, \delta) := \big\{ h_R : R \to \mathbb{R} \,\big|\, \forall x \in R \,,\, |h_R(x) - \tilde{h}_R(x)| < \delta \big\} \,. \tag{2.17}$$

2.2.2 Connection to the six-vertex model

It is interesting to observe that in two dimensions, the $\mathbb{Z}$-homomorphism model without random potential is equivalent to the six-vertex model with uniform weights. Recall that a configuration of the six-vertex is an assignment of one of six admissible states to each vertex in a specific subset of the lattice $\mathbb{Z}^2$, subject to certain local compatibility conditions. Weights $w_1, \ldots, w_6$ are associated to the six possible vertex states, and at least in the case of finite volume, configurations are sampled in proportion to the product of the weights of the vertices. When the weights are uniform (i.e. $w_1 = \cdots = w_6 = 1$) then the induced measure is uniform over admissible configurations, and the related partition function counts the number of configurations. If one follows the conventions of [RS18], then each configuration of the six-vertex model has a unique (up to additive constant) associated height function, defined on the faces of the lattice, such that the heights of two adjacent faces differ by ± 1. This height function is a $\mathbb{Z}$-homomorphism defined on the dual lattice. (Note that there is another common convention used to define height functions for the six-vertex model, used in [BCG16] among others; a review of the six-vertex model is beyond the scope of this dissertation.) A configuration of the six-vertex model is shown in Figure 2.1, along with its associated height function. Note that the six-vertex configuration pictured there satisfies the well-studied domain wall boundary conditions, and therefore the height function has extremal slope along all four edges of the boundary. As mentioned before, the partition function of the six-vertex model counts the number of configurations (because we take uniform weights). Since the configurations are in bijection with their height functions, the six-vertex partition function is closely related to the microscopic entropy defined above. As such, the results in this chapter can be translated to corresponding results for the six-vertex model, with uniform weights and appropriately translated boundary conditions.

2.2.3 Affine height functions

Affine height functions play an important role in defining and studying the entropy of our model. For an asymptotic height function $h_R : R \to \mathbb{R}$, we mean by "affine" the usual property: there exist $s \in [-1, 1]^m$ and $b \in \mathbb{R}$ such that $h_R(x) = s \cdot x + b$. The bounds on s ensure that h_R satisfies the Lipschitz property (2.12), so all such functions are indeed asymptotic height functions as per Definition 3.

On microscopic domains R_n, we consider best-possible approximations to affine functions. Fix $s \in [-1, 1]^m$ and $b \in \mathbb{R}$. At a lattice point $z \in \mathbb{Z}^m$, we define $h_{R_n}^{s \cdot x+b}(z)$ to be $s \cdot z + b$, rounded to the nearest integer of correct parity (see Figure 2.2). In the rest of this subsection,

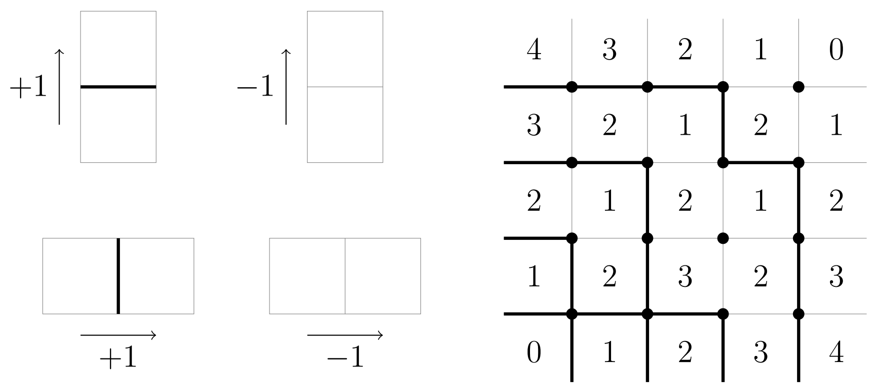


Figure 2.1: The height functions rules for the six-vertex model (cf. [RS18, Figure 4]), and an example of a configuration of the six-vertex model and the associated height function. The results in this chapter imply the variational principle and large deviations principle for the six-vertex model with uniform weights and any boundary data. The example above shows the domain wall boundary data, which corresponds to a boundary height function with extremal slope along all four edges of the square. These are the same boundary conditions as were imposed for the samples from the $\mathbb{Z}$-homomorphism model pictured in Figure 1.2.

we formalize this definition, verify that it actually does define a height function, and check that it is consistent.

Let us introduce an auxiliary notation that is used only in this subsection. Given a point $z = (z_1, \dots, z_m) \in \mathbb{Z}^m$, we say z has even or odd parity as $(\sum_{i=1}^m z_i) \in \mathbb{Z}$ has even or odd parity respectively, and we write $z \bmod 2$ for the parity of z.

Given $z \in \mathbb{Z}^m$ and $y \in \mathbb{R}$, we write $[y]_{z \bmod 2}$ for the closest integer to y that has parity $z \bmod 2$. In case of a tie, i.e. if y is an integer that has opposite parity to z, we arbitrarily choose to "round up" and set $[y]_{z \bmod 2} = y + 1 \in \mathbb{Z}$.

For example, let $z = (1, 2, 3) \in \mathbb{Z}^3$ and $z' = (4, -6, 7)$. Then z is an even point and z' is an odd point. So:

$$[5.4]_{z \bmod 2} = 6, \qquad [-3]_{z \bmod 2} = -2, \tag{2.18}$$

$$[5.4]_{z' \bmod 2} = 5, \qquad [-3]_{z' \bmod 2} = -3. \tag{2.19}$$

$$\tag{2.20}$$

Now, given $s \in [-1, 1]^m$ and $b \in \mathbb{R}$, we define the affine height functions $h_{R_n}^{s \cdot x + b}$ by

$$h_{R_n}^{s \cdot x + b}(z) := [s \cdot z + b]_{z \bmod 2}\,. \tag{2.21}$$

Note that the expression in the superscript of $h_{R_n}^{s \cdot x + b}$ is merely formal; "$s \cdot x + b$" should be read as "the function mapping x to $s \cdot x + b$". Moreover, the choice of domain R_n in the subscript does not affect the values of $h_{R_n}^{s \cdot x + b}$ at any point; for any sets $A_n, B_n \subseteq \mathbb{Z}^m$ and any point $z \in A_n \cap B_n$, one has $h_{A_n}^{s \cdot x + b}(z) = h_{B_n}^{s \cdot x + b}(z)$. An example of a function $h_{R_n}^{s \cdot x + b}$ is provided in Figure 2.2.

From the definition above, it is not clear that $h_{R_n}^{s \cdot x + b}$ are height functions. This is the content of Lemma 1.

Lemma 1. *Let $s \in [-1, 1]^m$ and $b \in \mathbb{R}$. For any adjacent points $z \sim z' \in \mathbb{Z}^m$, the values $h_{R_n}^{s \cdot x + b}(z)$ and $h_{R_n}^{s \cdot x + b}(z')$ differ by exactly 1.*

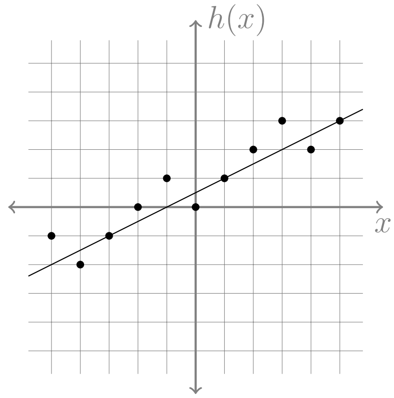


Figure 2.2: An affine height function $h_{R_n}^{s\cdot x+b}$ and the corresponding continuous affine function $x \mapsto s \cdot x + b$. Here $s = \frac{1}{2}$ and $b = \frac{1}{2}$.

Proof. From the definition of $h_{R_n}^{s\cdot x+b}$, we note two inequalities:

$$|h_{R_n}^{s\cdot x}(z) - (s \cdot z + b)| \leq 1, \tag{2.22}$$

and

$$|h_{R_n}^{s\cdot x+b}(z') - (s \cdot z' + b)| \leq 1. \tag{2.23}$$

Additionally, since $s \in [-1, 1]^m$, we have

$$|(s \cdot z + b) - (s \cdot z' + b)| \leq 1. \tag{2.24}$$

By the triangle inequality, $|h_{R_n}^{s\cdot x+b}(z) - h_{R_n}^{s\cdot x+b}(z')| \leq 3$. We shall show that equality cannot hold. Since the difference $h_{R_n}^{s\cdot x+b}(z) - h_{R_n}^{s\cdot x+b}(z')$ is obviously an odd integer, it will follow that the difference is ± 1.

Suppose towards a contradiction that

$$\left|h_{R_n}^{s\cdot x+b}(z) - h_{R_n}^{s\cdot x+b}(z')\right| = 3. \tag{2.25}$$

Then (2.22) and (2.23) must be equalities. From the definition of $[\cdot]_{z \bmod 2}$, necessarily then $s \cdot z + b$ is an integer with parity opposite that of z, and so

$$h_{R_n}^{s\cdot x+b}(z) = (s \cdot z + b) + 1. \tag{2.26}$$

Likewise

$$h_{R_n}^{s\cdot x+b}(z') = (s \cdot z' + b) + 1. \tag{2.27}$$

But then

$$\left|h_{R_n}^{s\cdot x+b}(z) - h_{R_n}^{s\cdot x+b}(z')\right| = \left|(s \cdot z + b + 1) - (s \cdot z' + b + 1)\right| \leq 1. \tag{2.28}$$

□

We end this section with the following lemma. The conclusion (2.30) is exactly what is needed later to apply the Kirszbraun theorem (Theorem 6):

Lemma 2. *Let $s, s' \in [-1,1]^m$, $b, b' \in \mathbb{R}$, and $z, z' \in \mathbb{Z}^m$. If*

$$\left|(s \cdot z + b) - (s' \cdot z' + b')\right| \leq |z - z'|_1, \tag{2.29}$$

then

$$\left|h_{\{z\}}^{s\cdot x+b}(z) - h_{\{z'\}}^{s'\cdot x+b'}(z')\right| \leq |z - z'|_1. \tag{2.30}$$

Proof. The proof is similar to that of Lemma 1. By the triangle inequality and (2.29),

$$\begin{aligned} &\left|h_{\{z\}}^{s\cdot x+b}(z) - h_{\{z'\}}^{s'\cdot x+b'}(z')\right| \\ &\quad \leq \left|h_{\{z\}}^{s\cdot x+b}(z) - (s \cdot z + b)\right| + \left|(s \cdot z + b) - (s' \cdot z' + b')\right| \\ &\qquad + \left|(s' \cdot z' + b') - h_{\{z'\}}^{s'\cdot x+b'}(z')\right| \\ &\quad \leq |z - z'|_1 + 2. \end{aligned} \tag{2.31}$$

Since $h_{\{z\}}^{s\cdot x+b}(z)$, $h_{\{z'\}}^{s'\cdot x+b'}(z')$ and $|z - z'|_1$ are all integers,

$$\left|h_{\{z\}}^{s\cdot x+b}(z) - h_{\{z'\}}^{s'\cdot x+b'}(z')\right| - |z - z'|_1 \in \{\ldots, -2, -1, 0, 1, 2\}. \tag{2.32}$$

We want to prove that the left-hand side of (2.32) is ≤ 0. By parity considerations it must be even, and we need only prove it is $\neq 2$. Assume for a contradiction that the left-hand side of (2.32) equals 2. Then equality holds in (2.31), and in particular

$$\left|h_{\{z\}}^{s\cdot x+b}(z) - (s \cdot z + b)\right| = 1 \quad \text{and} \quad \left|h_{\{z'\}}^{s'\cdot x+b'}(z') - (s' \cdot z' + b')\right| = 1. \tag{2.33}$$

As in the proof of Lemma 1, this implies that

$$h^{s\cdot x+b}_{\{z\}}(z) = (s \cdot z + b) + 1 \quad \text{and} \quad h^{s'\cdot x+b'}_{\{z'\}}(z') = (s' \cdot z' + b') + 1. \tag{2.34}$$

Therefore

$$\left|h^{s\cdot x+b}_{\{z\}}(z) - h^{s'\cdot x+b'}_{\{z'\}}(z')\right| = \left|(s \cdot z + b) - (s' \cdot z' + b')\right| \leq |z - z'|_1. \tag{2.35}$$

This is the desired contradiction, which completes the proof. □

2.2.4 Random Potential

Now let us formally describe the random potential ω. We will impose the following assumptions on ω throughout the rest of this chapter of the dissertation:

Assumption 2 (Random potential ω)**.** *We consider a real-valued random potential*

$$\omega = (\omega_e)_{e\in E(\mathbb{Z})} \in \mathbb{R}^{E(\mathbb{Z})} \tag{2.36}$$

defined on the set of edges $E(\mathbb{Z})$ of $\mathbb{Z}$. We write $\mathbb{P}$ for the law of ω and $\mathbb{E}$ for the expectation with respect to that law. We assume that $\mathbb{P}$ satisfies the following assumptions:

- *The elements ω_e of random potential are almost surely finite, and moreover the random variable C_ω defined by*

  $$C_\omega := 1 \vee \sup_{e\in E(\mathbb{Z})} |\omega_e| \tag{2.37}$$

  is in L^1, i.e. $\mathbb{E}[C_\omega] < \infty$.

- *The random potential ω is shift invariant. This means that for any finite number of edges $e_1, \dots e_k \in E(\mathbb{Z})$, any integer $z \in \mathbb{Z}$, and any bounded and measurable function $\xi : \mathbb{R}^k \to \mathbb{R}$,*

  $$\mathbb{E}\big[\xi(\omega_{e_1}, \dots, \omega_{e_k})\big] = \mathbb{E}\big[\xi(\omega_{\tau_z(e_1)}, \dots, \omega_{\tau_z(e_k)})\big]\,, \tag{2.38}$$

  where $\tau_z : E(\mathbb{Z}) \to E(\mathbb{Z})$ is the shift by z (as per the Notation and Conventions above).

- *Moreover, the random potential ω is ergodic with respect to the set of shifts $\{\tau_z \,|\, z \in \mathbb{Z},\, z \equiv 0 \pmod 2\}$. This means that if $E \subset \Omega$ is a shift invariant event, i.e. if $E = \tau_2^{-1}(E)$, then $\mathbb{P}(E) \in \{0, 1\}$.*

- *As a matter of normalization, we assume w.l.o.g. that*

$$\mathbb{E}[\omega_{e_{0,1}}] = 0\,, \tag{2.39}$$

where $e_{0,1}$ is the edge from 0 to 1 in $\mathbb{Z}$.

Example 1. A simple non-trivial example of a random potential ω that satisfies Assumption 2 is the i.i.d. potential. Let X denote a bounded (real) random variable with mean 0, and let $(\omega_e)_{e \in E(\mathbb{Z})}$ denote a family of i.i.d. copies of X.

Remark 1. The assumptions of shift invariance and ergodicity are standard in homogenization literature; see for example the "usual conditions" for the random conductance model from [Bis11b, Definition 3.1]. However we point out one difference: the random potential ω is ergodic with respect to the even shifts $\{\tau_z \,|\, z \equiv 0 \pmod 2\}$. This is a stronger condition than being ergodic with respect to the full set of shifts $\{\tau_z \,|\, z \in \mathbb{Z}\}$. This requirement is due to the earlier assumption made in Definition 1 that height functions preserve parity. As such, we cannot simply shift a height function up or down by 1 in the height space; if $h_{S_n}(z) = k \in \mathbb{Z}$, then there is no (parity-preserving) height function "$\tau_1 h_{S_n}$" such that $\tau_1 h_{S_n}(z) = k + 1$. More concretely, the family of measure-preserving translations used in the proof of Lemma 9 below includes all of the shifts $\{\tau_z \,|\, z \equiv 0 \pmod 2\}$ and none of the shifts $\{\tau_z \,|\, z \equiv 1 \pmod 2\}$, hence the stronger ergodicity assumption is technically required.

2.2.5 Entropy and surface tension

Recall that our goal in this chapter of the dissertation is to study limiting height profiles under the random potential defined by ω. In homogenization generally, one considers two different situations: In the quenched case, one considers the measure μ_ω for fixed ω. In the annealed case, one takes the expectation with respect to ω. Our goal is to show that the variational principle holds with high probability. With that context in mind, we define the quenched Hamiltonian $H_{R_n}(\cdot) = H_{R_n}(\cdot, \omega)$ and the quenched measure μ_ω as follows:

Definition 6 (The quenched Hamiltonian). For finite subsets $R_n \subset \mathbb{Z}^m$, we define the Hamiltonian H_{R_n} as follows: for a fixed boundary height function $h_{\partial R_n} : \partial R_n \to \mathbb{Z}$, for any height function $h_{R_n} \in M(R_n, h_{\partial R_n})$, and for any realization ω of the random potential,

$$H_{R_n}(h_{R_n}, \omega) = \sum_{e \in E(R_n)} \omega_{h_{R_n}(e)}, \tag{2.40}$$

where $E(R_n) = \{e_{x,y} \,|\, x, y \in R_n\}$ is the edge set of the subgraph of $\mathbb{Z}^m$ induced by R_n.

Definition 7 (Quenched Gibbs measure). Given a realization ω of the random potential and a set $A \subset M(R_n)$ of height functions, the partition function $Z_\omega(A)$ is given by

$$Z_\omega(A) = \sum_{h_{R_n} \in A} \exp\big(H_{R_n}(h_{R_n}, \omega)\big) \,. \tag{2.41}$$

For a fixed boundary data function $h_{\partial R_n} \in M(\partial R_n)$, the quenched Gibbs measure μ_ω on $M(R_n, h_{\partial R_n})$ is defined by

$$\mu_\omega(h_{R_n}) = \frac{1}{Z_\omega\big(M(R_n, h_{\partial R_n})\big)} \exp\big(H_{R_n}(h_{R_n}, \omega)\big) \,. \tag{2.42}$$

Remark 2. If one chooses the constant potential $\omega = \mathbf{0} = (0)_{e \in E(\mathbb{Z})}$, then the associated quenched Gibbs measure $\mu_{\mathbf{0}}$ is the uniform measure on $M(R_n, h_{\partial R_n})$. In this case one recovers the variational principle of [KMT20a].

Now let us introduce the microscopic entropy of our model. Again there are two situations: first, the quenched case, defined for a fixed realization ω and the annealed case.

Definition 8 (Quenched and annealed microscopic entropy). Given a domain $R_n \subset \mathbb{Z}^m$ and a finite non-empty subset $A \subset M(R_n)$, the quenched microscopic entropy $\mathrm{Ent}_{R_n}(A, \omega)$ is given by

$$\mathrm{Ent}_{R_n}(A, \omega) := -\frac{1}{|R_n|} \log Z_\omega(A) \tag{2.43}$$

$$\left(= -\frac{1}{|R_n|} \log \sum_{h_{R_n} \in A} \exp\big(H_{R_n}(h_{R_n}, \omega)\big) \right) . \tag{2.44}$$

The annealed microscopic entropy $\text{Ent}(R_n, h_{\partial R_n})$ is given by

$$\text{Ent}_{R_n,\mathtt{an}}(A) := \mathbb{E}\big[\text{Ent}_{R_n}(A,\omega)\big]. \tag{2.45}$$

Remark 3. As in Remark 2, if one chooses the constant potential $\omega = \mathbf{0}$, then the quenched microscopic entropy $\text{Ent}_{R_n}(M(R_n, h_{\partial R_n}), \mathbf{0})$ is the same as the microscopic entropy that was studied in [KMT20a].

Next, we define the local surface tension. As with the microscopic entropy, the local surface tension admits both a quenched and an annealed version, at least *a priori*.

Definition 9 (Quenched microscopic and local surface tension). The quenched local surface tension is the a.s.-limit

$$\text{ent}(s,\omega) := \lim_{n\to\infty} \text{ent}_n(s,\omega)\,, \tag{2.46}$$

where $\text{ent}_n(s,\omega)$ is the quenched microscopic surface tension, defined by

$$\text{ent}_n(s,\omega) := \text{Ent}_{S_n}\left(M(S_n, h^s_{\partial S_n}),\omega\right). \tag{2.47}$$

Recall from Notation and Conventions above that $S_n = \{-n,\ldots,n\}^m$, and note that the existence of the limit in (2.46) is the content of Lemma 9.

Definition 10 (Annealed microscopic and local surface tension). The annealed microscopic surface tension $\text{ent}_{n,\mathtt{an}}(s)$ is given by

$$\text{ent}_{n,\mathtt{an}}(s) := \mathbb{E}\left[\text{ent}_n(s,\omega)\right], \tag{2.48}$$

and the annealed local surface tension $\text{ent}(s)$ is given by

$$\text{ent}_{\mathtt{an}}(s) := \mathbb{E}\left[\text{ent}(s,\omega)\right]. \tag{2.49}$$

Remark 4. Similarly to Remark 2 and Remark 3, we obtain back the local surface tension for the uniform measure if we consider a constant random potential $\omega = \mathbf{0}$. In the case of random potential, it follows from Assumption 2 and Lemma 3 that $\text{ent}_n(s,\omega)$ is uniformly integrable and therefore that $\text{ent}_{n,\mathtt{an}}$ and $\text{ent}_{\mathtt{an}}$ are well-defined.

Remark 5. It is not hard to see that the annealed local surface tension is also the limit of the annealed microscopic surface tension. Indeed, from Assumption 2 the quenched microscopic surface tension $\mathrm{ent}_n(s,\omega)$ is dominated by an L^1 function (see Lemma 3). Therefore, the dominated convergence theorem implies that

$$\lim_{n\to\infty} \mathrm{ent}_{n,\mathtt{an}}(s) = \lim_{n\to\infty} \mathbb{E}\left[\mathrm{ent}_n(s,\omega)\right] = \mathbb{E}\left[\lim_{n\to\infty} \mathrm{ent}_n(s,\omega)\right] = \mathrm{ent}_{\mathtt{an}}(s)\,. \tag{2.50}$$

Moreover, it is the content of Lemma 10 that the quenched local surface tension is a.s. equal to the annealed local surface tension.

The annealed macroscopic entropy is defined by:

Definition 11 (Annealed macroscopic entropy). Given an asymptotic height function $h_R \in M(R, h_{\partial R})$, the annealed macroscopic entropy $\mathrm{Ent}_{R,\mathtt{an}}(h_R)$ is defined by

$$\mathrm{Ent}_{R,\mathtt{an}}(h_R) := \int_R \mathrm{ent}_{\mathtt{an}}(\nabla h(x))\, dx\,. \tag{2.51}$$

2.3 Main results

The first main result of this chapter is the profile theorem. Its proof is the content of Section 2.5.

Theorem 3 (Profile theorem). *Recall that $C_\omega := 1 \vee \sup_{e\in E(\mathbb{Z})}|\omega_e|$ is by Assumption 2 an L^1 random variable. For any $h_R \in M(R, h_{\partial R})$ and any $\eta > 0$, there exist functions $\theta_{h_R}(\delta)$ and $\theta_{h_R,\delta}(\frac{1}{n})$ with $\theta_{h_R}(\delta) \to 0$ as $\delta \to 0$ and $\theta_{h_R,\delta}(\frac{1}{n}) \to 0$ as $n \to \infty$ such that*

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\Big(&\Big|\mathrm{Ent}_{R_n}\big(B(R_n, h_R, \delta), \omega\big) - \mathrm{Ent}_{\mathtt{an}}(R, h_R)\Big| \\ &\geq \eta + C_\omega \theta_{h_R}(\delta) + C_\omega \theta_{h_R,\delta}\big(\tfrac{1}{n}\big)\Big) = 0.\end{aligned} \tag{2.52}$$

The second main result is the variational principle. Its proof is the content of Section 2.6.

Theorem 4 (Variational principle)**.** *The random variables*

$$\mathrm{Ent}_{R_n}(M(R_n, h_{\partial R_n}, \delta), \omega) \tag{2.53}$$

converge in probability to the infimum of $\mathrm{Ent}_{\mathtt{an}}(R, h_R)$ *over asymptotic height functions* $h_R \in M(R, h_{\partial R})$, *i.e. for every* $\eta > 0$,

$$\limsup_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}\bigg(\Big| \mathrm{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}), \omega\big) - \inf_{h_R \in M(R, h_{\partial R})} \mathrm{Ent}_{R,\mathtt{an}}(h_R) \Big| \geq \eta \bigg) = 0. \tag{2.54}$$

The third main result is the large deviations principle, which we state using the standard notation from large deviations theory. Its proof is the content of Section 2.7.

Theorem 5 (Large deviations principle)**.** *Consider the space* $M(R)$ *of asymptotic height functions on* R, *endowed with the topology of uniform convergence. For* $\delta > 0$ *and* $n \in \mathbb{N}$, *define a random probability measure* $\mu_{\delta,n}(\cdot, \omega)$ *on* $M(R)$ *by*

$$\mu_{\delta,n}(A, \omega) := \frac{Z_\omega\big(\{h_{R_n} \in M(R_n, h_{\partial R_n}, \delta) \,\big|\, \tilde{h}_{R_n} \in A\}\big)}{Z_\omega\big(M(R_n, h_{\partial R_n}, \delta)\big)}, \tag{2.55}$$

where $\tilde{h}_{R_n} \in M(R)$ *denotes the asymptotic height function given by rescaling and interpolating* $h_{R_n} \in M(R_n)$, *i.e.* $\tilde{h}_{R_n}(\frac{1}{n}z) = \frac{1}{n}h_{R_n}(z)$ *for* $z \in R_n$.

Then the measures $\mu_{\delta,n}$ *satisfy a large deviations principle in probability with rate functional* I *given by*

$$I(h_R) := \begin{cases} \mathrm{Ent}_{R,\mathtt{an}}(h_R) - E & \text{if } h_R \in M(R, h_{\partial R}), \\ +\infty & \text{otherwise,} \end{cases} \tag{2.56}$$

where $E := \inf_{h_R \in M(R, h_{\partial R})} \mathrm{Ent}_{R,\mathtt{an}}(h_R)$. *Specifically, this means that for any Borel set* $A \subset M(R)$ *and any* $\eta > 0$,

$$\limsup_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}\bigg(\frac{1}{|R_n|} \log \mu_{\delta,n}(A, \omega) \;\leq\; -\inf_{h_R \in A^\circ} I(h_R) - \eta \bigg) = 0 \tag{2.57}$$

and

$$\limsup_{\delta\to 0}\limsup_{n\to\infty}\mathbb{P}\left(\frac{1}{|R_n|}\log\mu_{\delta,n}(A,\omega)\ \geq\ -\inf_{h_R\in\overline{A}}I(h_R)+\eta\right)=0, \tag{2.58}$$

where A° denotes the interior of A and $\overline{A}$ denotes the closure.

2.4 The quenched and annealed local surface tension

The purpose of this section is to establish several fundamental properties of the quenched entropy and local surface tension of our model. We proceed as follows:

- In Section 2.4.1 we state the Kirszbraun theorem, used heavily in the rest of this section and beyond.
- In Section 2.4.2 we derive robustness of the entropy and local surface tension under boundary value changes.
- In Section 2.4.3 we prove the existence of the quenched local surface tension and the equivalence between the quenched and annealed local surface tension.
- In Section 2.4.4 we study the local surface tension as a function $s \mapsto \mathrm{ent}_{\mathtt{an}}(s)$, and we show that this function is convex and continuous.

2.4.1 Kirszbraun theorem

The Kirszbraun theorem quoted below is a discrete analogue of the classical Kirszbraun theorem from [Kir34]. The classical theorem gives a condition under which a Lipschitz continuous function can be extended from a subset of a domain to the entirety of that domain. Likewise, the Kirszbraun theorem for graph homomorphisms quoted below gives a condition under which a $\mathbb{Z}$-valued graph homomorphism may be extended from a subset of a domain to the entire domain. Note that the property of being a $\mathbb{Z}$-valued graph homomorphism is stronger than the Lipschitz property with constant 1, since if $z \sim \tilde{z}$ are two adjacent points in the domain of a graph homomorphism $h : R_n \to \mathbb{Z}$, then $h(z) \neq h(\tilde{z})$.

Theorem 6. *Let R_n be a connected region of $\mathbb{Z}^m$, let R'_n be a subset of R_n, and let $\bar{h} : R'_n \to \mathbb{Z}$ be a graph homomorphism that preserves parity. There exists a graph homomorphism $h : R_n \to \mathbb{Z}$ such that $h = \bar{h}$ on R'_n if and only if for all $x, y \in R'_n$,*

$$d_{\mathbb{Z}}(\bar{h}(x), \bar{h}(y)) \leq d_{R_n}(x, y), \tag{2.59}$$

where $d_{\mathbb{Z}}$ and d_{R_n} denote respectively the graph distance on $\mathbb{Z}$ and on $R_n \subset \mathbb{Z}^m$.

Remark 6. The parity condition is necessary in general; consider for example the function $\overline{h}$ defined on $\{0, 2\} \subset \mathbb{Z}$ by $\overline{h}(0) = 0$, $\overline{h}(2) = 1$. The parity condition in Theorem 6 is the reason for the parity condition in Definition 1.

There is a proof of a more general version of this theorem in [MT20b, Theorem 4.1]. The proof is restated below for the reader's convenience, except that it simplified by only addressing the model from this dissertation, where the height functions take values in $\mathbb{Z}$ rather than in an infinite regular tree.

Proof of Theorem 6. Obviously if an extension h of $\overline{h}$ exists, then $\overline{h}$ satisfies (2.59). So, suppose instead that (2.59) holds, and let us prove that an extension h exists. For $y \in R_n$, set

$$h(y) := \max\left\{\overline{h}(x) - |x - y|_1 \,\middle|\, x \in R'_n\right\}. \tag{2.60}$$

We must check two things: first that $h(y) = \overline{h}(y)$ when $y \in R'_n$, and second that $|h(y) - h(\tilde{y})| = 1$ when $y \sim \tilde{y}$ are adjacent points in R_n.

To prove that $h|_{R'_n} = \overline{h}$, let $y \in R'_n$ and consider any point $x \in R'_n$. By the Lipschitz property of $\overline{h}$,

$$\overline{h}(x) - \overline{h}(y) \leq \left|\overline{h}(x) - \overline{h}(y)\right| \leq |x - y|_1, \tag{2.61}$$

so $\overline{h}(x) - |x - y|_1 \leq \overline{h}(y)$. Therefore the maximum in (2.60) is attained when $x = y$, so $h(y) = \overline{h}(y) + |y - y|_1 = \overline{h}(y)$.

To prove that h is a graph homomorphism, let $y \sim \tilde{y}$ be adjacent points in R_n, and let $x, \tilde{x}$ be points in R'_n that attain the maximum in (2.60) for $y, \tilde{y}$ respectively, i.e. $h(y) = \overline{h}(x) - |x-y|_1$ and $h(\tilde{y}) = \overline{h}(\tilde{x}) - |\tilde{x}-\tilde{y}|_1$. Then

$$\begin{aligned} h(y) &= \max\left\{\overline{h}(z) + |z-y|_1 \,\middle|\, z \in R'_n\right\} \\ &\geq \overline{h}(\tilde{x}) - |\tilde{x}-y| \\ &\geq \overline{h}(\tilde{x}) - |\tilde{x}-\tilde{y}| - 1 \\ &= h(\tilde{y}) - 1, \end{aligned} \tag{2.62}$$

and likewise $h(\tilde{y}) \geq h(y) - 1$.

For every $x \in R'_n$, the map $y \mapsto \overline{h}(x) + |x-y|_1$ preserves parity (recall the assumption that $\overline{h}$ preserves parity), and therefore so does h. So h is a parity-preserving map such that $|h(y) - h(\tilde{y})| \leq 1$ whenever y and $\tilde{y}$ are neighbors. This proves that h is a graph homomorphism. □

As an illustration of the usefulness of the Kirszbraun theorem, we prove the following lemma, which justifies the choice of the normalizing factor $\frac{1}{|R_n|}$ in Definition 8:

Lemma 3. *Almost surely (in terms of the distribution $\mathbb{P}$ of the random potential ω),*

$$-\log(2) - 2mC_\omega \leq \mathrm{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}), \omega\big) \leq 2mC_\omega\,. \tag{2.63}$$

Proof. As a corollary of the Kirszbraun theorem (Theorem 6), there is always at least one height function $h_0 \in M(R_n, h_{\partial R_n})$. So,

$$\mathrm{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}), \omega\big) \leq -\frac{1}{|R_n|} \log \sum_{h \in \{h_0\}} \exp\left(\sum_{e \in E(R_n)} \omega_{e_{h(x),h(y)}}\right) \tag{2.64}$$

$$\leq \frac{|E(R_n)|}{|R_n|} C_\omega \tag{2.65}$$

$$\leq 2mC_\omega\,. \tag{2.66}$$

On the other hand, we overestimate the cardinality of $M(R_n, h_{\partial R_n})$ as follows: enumerate the points of the interior of R_n, in such a way that each point x_i is adjacent to the previous point x_{i-1} (and the first point x_1 is adjacent to $x_0 \in \partial R_n$). For each point x_i in the enumeration, we require that $h(x_i) = h(x_{i-1}) \pm 1$, so there are at most 2 choices for $h(x_i)$. All together, $|M(R_n, h_{\partial R_n})| \leq 2^{|R_n|}$. It follows that

$$\begin{aligned} \operatorname{Ent}(R_n, h_{\partial R_n}, \omega) &\geq -\frac{1}{|R_n|} \log\Big(|M(R_n, h_{\partial R_n})| \exp\big(C_\omega |E(R_n)|\big)\Big) && (2.67) \\ &\geq -\frac{1}{|R_n|} \log 2^{|R_n|} - \frac{|E(R_n)|}{|R_n|} C_\omega && (2.68) \\ &\geq -\log(2) - mC_\omega. && (2.69) \end{aligned}$$

□

In the sequel, we will usually use the Kirszbraun theorem in the following setting. Given two domains $R_{n_1} \subset R_{n_2} \subset \mathbb{Z}^m$, a height function $h_{R_{n_1}} \in M(R_{n_1})$, and a boundary height function $h_{\partial R_{n_2}} \in M(\partial R_{n_2})$, there exists an extension $\tilde{h}_{R_{n_2}} \in M(R_{n_2})$ with $\tilde{h}_{R_{n_2}}|_{R_{n_1}} = h_{R_{n_1}}$ and $\tilde{h}_{R_{n_2}}|_{\partial R_{n_2}} = h_{\partial R_{n_2}}$ if and only if

$$|h_{R_{n_1}}(z_1) - h_{\partial R_{n_2}}(z_2)| \leq |z_1 - z_2|_1 \quad \text{for all } z_1 \in \partial R_{n_1}, z_2 \in \partial R_{n_2}\,. \tag{2.70}$$

2.4.2 Robustness of the quenched entropy

The quenched microscopic entropy and local surface tensions are robust, in the sense that small changes in boundary values cause small changes in the numeric value of the entropy. There are two steps in proving these robustness results: First, just as for the unperturbed model of [KMT20a], compare the two sets of height functions associated with the two boundary value functions, perhaps by exhibiting an injection from one set into the second or by estimating cardinalities directly. Second, show that individual height functions from each of the two sets contribute comparable amounts to the entropy after applying the random potential, e.g. by showing that every height function in one set admits a "similar" height function in the second set, whose Hamiltonian value is not much different; this step is sometimes straightforward and other times quite subtle.

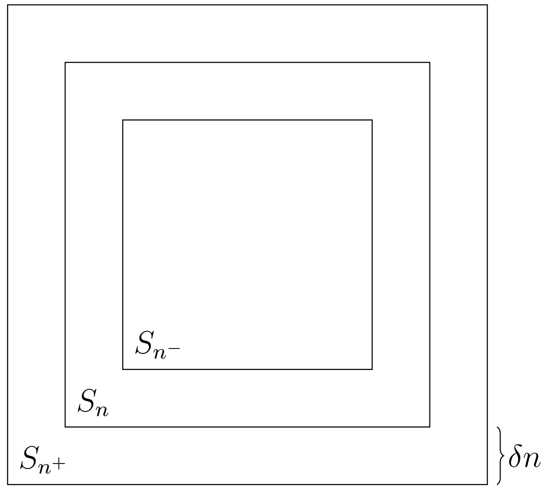


Figure 2.3: Nested domains from Lemma 4.

Lemma 4. *Let* $\alpha > 0$*, let* $s \in \mathbb{R}^m$ *with* $|s|_\infty \leq 1 - \alpha$*, let* $\varepsilon \in (0, \frac{\alpha}{2})$*, let* $n \in \mathbb{N}$ *with* $n \geq (1 - \frac{2\varepsilon}{\alpha})^{-1})$*, and let* $h_{\partial S_n} \in M(\partial S_n, s, \varepsilon)$*. Write*

$$n^+ := \left\lceil (1 + \tfrac{2\varepsilon}{\alpha})n \right\rceil \qquad \textit{and} \qquad n^- := \left\lfloor (1 - \tfrac{2\varepsilon}{\alpha})n \right\rfloor . \tag{2.71}$$

(We remark that $1 \leq n^- < n < n^+$*.) Then,*

$$\begin{aligned} \mathrm{ent}_{n^+}(s,\omega) - C_\omega\, \theta\big(\tfrac{\varepsilon}{\alpha}\big) &\leq \mathrm{Ent}_{S_n}\big(M(S_n, h_{\partial S_n}), \omega\big) \\ &\leq \mathrm{ent}_{n^-}(s,\omega) + C_\omega\, \theta_m\big(\tfrac{\varepsilon}{\alpha}\big) . \end{aligned} \tag{2.72}$$

Proof of Lemma 4. We prove the inequality

$$\mathrm{ent}_{n^+}(s,\omega) - C_\omega \theta\big(\tfrac{\varepsilon}{\alpha}\big) \leq \mathrm{Ent}_{S_n}\big(M(S_n, h_{\partial S_n}), \omega\big). \tag{2.73}$$

The proof of the reverse inequality is similar.

Note that the smaller square $S_n = \{-n, -(n-1), \ldots, n-1, n\}^m$ is contained inside the larger square S_{n^+}, and that

$$|x - y|_1 \geq \tfrac{2\varepsilon}{\alpha} n \qquad \text{whenever } x \in \partial S_n \text{ and } y \in \partial S_{n^+} . \tag{2.74}$$

We construct an injection from $M(S_n, h_{\partial S_n})$ into $M(S_{n^+}, h^s_{\partial S_{n^+}})$ using the Kirszbraun theorem, Theorem 6. Let $h_{S_n} \in M(S_n, h_{\partial S_n})$, let $x \in \partial S_n$, and let $y \in \partial S_{n^+}$. By the definitions of $M(S_n, h_{\partial S_n})$ and of $h^s_{\partial S_n}$,

$$\left|h_{S_n}(x) - h^s_{S_{n^+}}(y)\right| \tag{2.75}$$

$$\leq \left|h_{S_n}(x) - s \cdot x\right| + \left|s \cdot (x-y)\right| + \left|h^s_{S_{n^+}}(y) - s \cdot y\right| \tag{2.76}$$

$$\leq \varepsilon n + |s|_\infty |x-y|_1 + 1. \tag{2.77}$$

By hypothesis $|s|_\infty \leq 1-\alpha$ and by (2.74), $\varepsilon n \leq \frac{\alpha}{2}|x-y|_1$. Therefore for $n \geq \frac{2}{\alpha}$,

$$|h_{S_n}(x) - h^s_{S_{n^+}}(y)| \leq |x-y|_1\,, \tag{2.78}$$

so h_{S_n} admits an extension $h_{S_{n^+}} \in M(S_{n^+}, h^s_{\partial S_{n^+}})$. The map $h_{S_n} \mapsto h_{S_{n^+}}$ is an injection from $M(S_n, h_{\partial S_n})$ into $M(S_{n^+}, h^s_{\partial S_{n^+}})$. The existence of such an injection implies immediately that

$$\begin{aligned}\mathrm{Ent}_{S_n}\big(M(S_n, h_{\partial S_n}), \omega\big) &\geq \frac{|S_n|}{|S_{n^+}|}\,\mathrm{Ent}_{S_{n^+}}\big(M(S_{n^+}, h^s_{\partial S_{n^+}}), \omega\big) \\ &\qquad - \frac{2mC_\omega(|S_{n^+}| - |S_n|)}{|S_n|} \\ &= \mathrm{Ent}_{S_{n^+}}\big(M(S_{n^+}, h^s_{\partial S_{n^+}}), \omega\big) - C_\omega \theta_m\big(\tfrac{\varepsilon}{\alpha}\big)\,.\end{aligned} \tag{2.79}$$

This proves the first inequality of (2.72). As mentioned at the beginning of the proof, the other inequality is similar. Since $n^- < n$, one extends height functions from $M(S_{n^-}, h_{\partial S^s_{n^-}})$ to $M(S_n, h_{\partial S_n})$. We omit the details. □

Lemma 4 does not extend to the case where $|s|_\infty = 1$. As $|s|_\infty \to 1$ the ratio of the box sizes $\frac{|S_{n+}|}{|S_n|} \approx 1 + \frac{\varepsilon}{\alpha}$ and the error bound $\theta(\frac{\varepsilon}{\alpha})$ both diverge. Fundamentally these difficulties come from the Kirszbraun theorem. When $|s|_\infty$ is close to 1, the "margin" $S_{n^+} \setminus S_n$ must be large in order to connect $h_{\partial S_n}$ to $h_{\partial S_{n^+}}$ and when $|s|_\infty = 1$, such an extension is not generally possible. Therefore we take a different approach for $|s|_\infty \approx 1$, using elementary combinatorics to count the number of height functions. The two following calculations are intermediate results used to prove the robustness lemma, Lemma 7.

Lemma 5 (Counting height functions near $|s|_\infty = 1$)**.** *Let $\varepsilon > 0$. Let $s \in \mathbb{R}^m$ with $1-\varepsilon < |s|_\infty \le 1$, and let $h_{\partial S_n} \in M(\partial S_n, h^s_{\partial S_n}, \varepsilon)$. Then,*

$$\frac{1}{|S_n|} \log\big|M(S_n, h_{\partial S_n})\big| = \theta(\varepsilon)\,. \tag{2.80}$$

Proof of Lemma 5. Fix a coordinate index $1 \le i \le m$ such that $|s_i| > 1-\varepsilon$, and assume without loss of generality that $s_i > 1-\varepsilon$. Decompose S_n into $(2n+1)^{m-1}$ lines in the i^{th} coordinate direction. Along each such line h_{S_n} must increase by at least $2(1-2\varepsilon)n$. Therefore, the $2n$ edges in the line split into two subsets: at least $2(1-2\varepsilon)n$ "increasing" edges, and at most $4\varepsilon n$ "decreasing" edges. Counting each line independently, we conclude that

$$\big|M(S_n, h_{\partial S_n})\big| \le \binom{2n}{\lceil 4\varepsilon n \rceil}^{(2n+1)^{m-1}}. \tag{2.81}$$

The conclusion (2.80) follows immediately. For a more verbose version of this proof, see [KMT20a, Lemma 21]. □

Lemma 6 (Height functions at slope $|s|_\infty = 1$)**.** *Let $s' \in \mathbb{R}^m$ with $|s'|_\infty = 1$. Then $|M(S_n, h^{s'}_{\partial S_n})| = 1$, and the sole element of $M(S_n, h^{s'}_{\partial S_n})$ is the canonical height function $h^{s'}_{S_n}$.*

Proof of Lemma 6. As in the proof of Lemma 5, fix a coordinate index $1 \le i \le m$ such that $|s_i| = 1$. Decompose S_n into lines in the i^{th} coordinate direction. Along each line, any height function $h_{S_n} \in M(S_n, h^{s'}_{\partial S_n})$ must increase by exactly $2n$. Since h_{S_n} is a graph homomorphism, that is only possible if h_{S_n} increases along every edge, i.e. $h_{S_n}(x+1, y) - h_{S_n}(x, y) = 1$ for $x = -n, \ldots, n-1$. It follows that $|M(S_n h^{s'}_{\partial S_n})| \le 1$. To complete the proof, observe that $h^{s'}_{S_n} \in M(S_n, h^{s'}_{\partial S_n})$. □

Having recorded Lemma 5 and 6, we return to establishing robustness results. Our goal is to compare the microscopic surface tension $\text{ent}_n(s, \omega) := \text{Ent}_{S_n}(M(S_n, h^s_{\partial S_n}), \omega)$ and the entropy $\text{Ent}_{S_n}(M(S_n, h_{\partial S_n}), \omega)$ associated to an "approximately affine" boundary height function $h_{\partial S_n} \in M(\partial S_n, h^s_{\partial S_n}, \varepsilon)$. The difference is that Lemma 4 took $|s|_\infty \leq 1-\alpha$ and the lemma below takes $|s|_\infty > 1-\alpha$.

Lemma 7. *Let $\varepsilon > 0$. Let $s, s' \in \mathbb{R}^m$ with $|s|_\infty \leq 1$, $|s'|_\infty = 1$, and $|s - s'|_\infty < \varepsilon$. Let $n \in \mathbb{N}$ be sufficiently large (specifically, $n \geq \frac{1}{\varepsilon}$) and let $h_{\partial S_n} \in M(\partial S_n, h^s_{\partial S_n}, \varepsilon)$. Then:*

$$\left|\text{Ent}(S_n, h_{\partial S_n}, \omega) - \text{ent}_n(s', \omega)\right| \leq C_\omega \, \theta(\varepsilon)\,. \tag{2.82}$$

Because of the $\theta(\varepsilon)$ error term, Lemma 7 will not be useful for slopes s with $|s|_\infty$ far from 1.

Remark 7. There are two ingredients to the proof. The first is counting results of Lemma 5 and Lemma 6, and the second is a comparison between the Hamiltonian $H_{S_n}(h_{S_n}, \omega)$ of a generic height function $h_{S_n} \in M(S_n, h_{\partial S_n})$ and the Hamiltonian $H_{S_n}(h^{s'}_{S_n}, \omega)$ of the unique element $h^{s'}_{S_n} \in M(S_n, h^{s'}_{\partial S_n})$. Since proofs were already given for the two lemmas, most of the argument below is spent on the comparison of Hamiltonians.

The comparison of Hamiltonians is also fundamentally a combinatorial argument that relies on the rigidity caused by the slopes s and s' being close to (or on) the boundary of the slope space $[-1, 1]^m$. It is surprising that such a subtle argument is (apparently) needed in the case of homogenization, since the two counting lemmas are sufficient in the uniform case, and these lemmas are not very complicated to prove.

The subtlety is similar to that of the proof of Lemma 8 below. In both cases, difficulties arise when comparing Hamiltonians for two height functions defined on the same domain S_n. In comparison, the proof of Lemma 4 (which has a similar statement to the current Lemma 7)) is based on extending height functions from one domain to another larger domain via the Kirszbraun theorem. Comparing the Hamiltonian of a height function on the larger domain to the Hamiltonian of the same function on a sub-domain is simple, since the difference in Hamiltonians can be controlled by the difference in cardinality of the domains.

Proof of Lemma 7. As mentioned above, we will compare the Hamiltonians $H_{S_n}(h_{S_n}, \omega)$ and $H_{S_n}(h^{s'}_{S_n}, \omega)$, where $h_{S_n} \in M(S_n, h_{\partial S_n})$ and $h^{s'}_{S_n} \in M(S_n, h^{s'}_{\partial S_n})$. More precisely, we will later deduce the inequality

$$\left|H_{S_n}(h_{S_n}, \omega) - H_{S_n}(h^{s'}_{S_n}, \omega)\right| \leq 210 m^2 (2n+1)^m C_\omega \varepsilon \,. \tag{2.83}$$

Given that (2.83) holds, the proof is straight-forward: For one inequality, we calculate

$$\begin{aligned}
&\mathrm{Ent}_{S_n}\big(M(S_n, h_{\partial S_n}), \omega\big) && (2.84)\\
&\quad = -\frac{1}{|S_n|} \log \sum_{h_{S_n} \in M(S_n, h_{\partial S_n})} \exp\big(H_{S_n}(h_{S_n}, \omega)\big) && (2.85)\\
&\quad \overset{(2.83)}{\leq} -\frac{1}{|S_n|} \log \sum_{h_{S_n} \in M(S_n, h_{\partial S_n})} \exp\bigg(H_{S_n}(h^{s'}_{S_n}, \omega) && (2.86)\\
&\qquad\qquad\qquad\qquad - 210 m^2 (2n+1)^m C_\omega \varepsilon\bigg) && (2.87)\\
&\quad \overset{Lemma\ 5}{\leq} \quad -\frac{1}{|S_n|} H_{S_n}\big(h^{s'}_{S_n}, \omega\big) + \theta(\varepsilon) + 210 m^2 C_\omega \varepsilon && (2.88)\\
&\quad = \mathrm{Ent}_{S_n}\big(M(S_n, h^{s'}_{S_n}), \omega\big) + \theta(\varepsilon)\,. && (2.89)
\end{aligned}$$

The opposite inequality is derived in the same way, which concludes the proof of Lemma 7 up to the verification of (2.83).

For convenience, let us use for the remaining argument the following convention: When denoting the Hamiltonian of $H_{S_n}(h_{S_n}, \omega)$ we just write $H(h_{S_n})$, omitting the dependency on the random potential ω.

Informal verification of (2.83): Heuristically, the estimate (2.83) makes sense. Because the slopes s and s' are ε-close to each other, and s' has slope 1, every height function $h_{S_n} \in$

$M(S_n, h_{\partial S_n})$ has to behave similar to the canonical height function $h^{s'}_{S_n}$ of slope s'. Therefore, the difference in the associated energies, as measured by the Hamiltonian $H_{S_n}(h_{S_n})$ and $H_{S_n}(h^{s'}_{S_n})$, should vanish as $\varepsilon \to 0$.

To make this argument rigorous one needs to precisely estimate the number of heights that each height function h_{S_n} visits, i.e. the set $\{h_{S_n}(e) \,|\, e \in E(S_n)\}$ with multiplicities, and compare to the corresponding set for $h^{s'}_{S_n}$. This is relatively straight-forward on a one-dimensional lattice but unfortunately becomes much more subtle on a higher-dimensional lattice. To see why, consider the decomposition of the box S_n into lines. This leads a decomposition of the edges in $E(S_n)$ into *parallel edges* within a line, and *cross edges* connecting two lines. Without cross edges the one-dimensional argument would easily carry over, but controlling the cross edges is necessary as well. This control is accomplished by the sets G_y below.

To begin the rigorous verification of (2.83), pick an arbitrary height function $h_{S_n} \in M(S_n, h_{\partial S_n})$. As mentioned above, we decompose S_n into lines parallel to one of the coordinate axes. Assume by symmetry that $s = (s_1, s_2, \ldots, s_m)$ and $s' = (s'_1, \ldots, s'_m)$ satisfy $s'_1 = 1$ and (therefore) $s_1 > 1 - \varepsilon$. For $y \in \{-n, \ldots, n\}^{m-1}$ let ℓ_y denote the line in the first coordinate direction through $(0, y)$ in S_n, i.e.

$$\ell_y := \big\{(-n, y),\, (-n+1, y),\, \ldots,\, (n-1, y),\, (n, y)\big\}\,. \tag{2.90}$$

Observe that S_n is the disjoint union of the $(2n+1)^{m-1}$ lines ℓ_y. In particular, the Hamiltonian $H_{S_n}(h_{S_n})$ decomposes with respect to the lines ℓ_y as

$$\begin{aligned} H_{S_n}(h_{S_n}) &:= \sum_{e \in E(S_n)} \omega_{h_{S_n}(e)} \\ &= \sum_y \left(\sum_{e \in E(\ell_y)} \omega_{h_{S_n}(e)} + \frac{1}{2} \sum_{y' \sim y} \sum_{e \in \tilde{E}_{y,y'}} \omega_{h_{S_n}(e)} \right) \\ &= \sum_y \tilde{H}_y(h_{S_n})\,, \end{aligned} \tag{2.91}$$

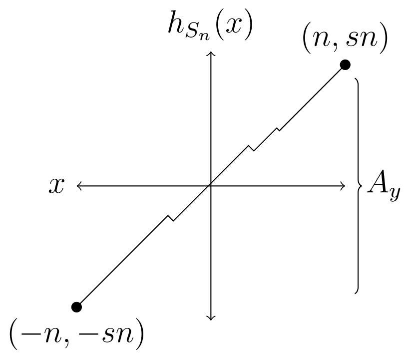
where $\tilde{E}_{y,y'}$ is the set of edges in $E(S_n)$ with one endpoint in ℓ_y and the other in $\ell_{y'}$ (we call these *cross edges*), and where $\tilde{H}_y$ is defined to be the parenthesized quantity from the line above. Note that the factor $\frac{1}{2}$ is necessary because each cross edge in $\tilde{E}_{y,y'}$ also contributes to $\tilde{H}_{y'}(h_{S_n})$, so without the factor $\frac{1}{2}$ the contributions from the cross edges would be double-counted.

We define two families of sets $A_y \subset E(\mathbb{Z})$ and $G_y \subset A_y$, indexed by points $y \in \{-n, \ldots, n\}^{m-1}$. In terms of the heuristic argument above, these sets roughly correspond to the heights visited by h_{S_n} and $h^{s'}_{S_n}$, although in fact both A_y and G_y are subsets of $\{h_{S_n}(e) \,|\, e \in E(S_n)\}$.

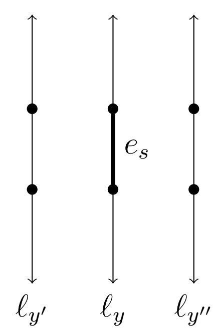
Let A_y denote the edges $e \in E(\mathbb{Z})$ that lie inside the interval from $(s \cdot (-n, y) + 2\varepsilon n)$ to $(s \cdot (+n, y) - 2\varepsilon n)$. Based on the boundary conditions and homomorphism property of h_{S_n} and $h^{s'}_{S_n}$, every edge $e \in A_y$ occurs both in the image $\{h_{S_n}(\tilde{e}) \,|\, \tilde{e} \in E(\ell_y)\}$ and in the image $\{h^{s'}_{S_n}(\tilde{e}) \,|\, \tilde{e} \in E(\ell_y)\}$. (The factors of 2 in the definition of A_y are necessary since the boundary height function $h_{\partial S_n}$ may differ from $h^s_{\partial S_n}$ by up to εn, in addition to s_1 differing from 1 by up to ε.) The situation in dimension $m = 1$ is illustrated in Figure 2.4a.

We define $G_y \subset A_y$ in the following way: These are the edges $e \in A_y \subset E(\mathbb{Z})$ satisfying these three constraints with respect to h_{S_n} (illustrated in Figure 2.4b):

- e occurs with multiplicity 1 in the multi-set $\{h_{S_n}(\tilde{e}) \,|\, \tilde{e} \in E(\ell_y)\}$. (By choice of A_y, e occurs with multiplicity ≥ 1.) Write e_s for the unique edge $e_s \in E(\ell_y)$ such that $h_{S_n}(e_s) = e$.
- Both endpoints of e occur with multiplicity 1 in the multi-set $\{h_{S_n}(z) \,|\, z \in \ell_y\}$.
- For each endpoint z of e_s and each neighboring vertex $z' \sim z$ that lies in $S_n \setminus \ell_y$, $h_{S_n}(z')$ occurs with multiplicity 1 in the multi-set $\{h_{S_n}(\tilde{z}) \,|\, \tilde{z} \in \ell_{y'}\}$ for the line $\ell_{y'}$ that



(a) Here h_{S_n} is a one-dimensional height function with slope $s \geq 1 - \varepsilon$. The set A_y comprises the $1 - 4\varepsilon$ fraction of the $2n$ edges in ℓ_y, centered around 0. (The central height in higher dimensions is instead $s \cdot (0, y)$.) Both h_{S_n} and $h^{s'}_{S_n}$ must contain all of these edges in their image. They might contain additional edges.



(b) The three lines are ℓ_y in the center and two of its neighbors, $\ell_{y'}$ and $\ell_{y''}$. The highlighted edge is the edge $e_s \in E(\ell_y)$ for $e \in G_y$, i.e. the unique edge in ℓ_y with $h_{S_n}(e_s) = e$. There is also an edge $e_{s'}$ (not shown), satisfying the corresponding uniqueness property for $h^{s'}_{S_n}$. Finally, all six highlighted vertices are good, i.e. each vertex has a unique height within its line.

Figure 2.4: Figures relating to the proof of Lemma 7.

contains it.

Further on in the argument, we will call elements of G_y "good" edges. We will call a vertex $z \in \ell_y$ "good" if its height $h_{S_n}(z)$ occurs in with multiplicity 1 in $\{h_{S_n}(\tilde{z}) \,|\, \tilde{z} \in \ell_y\}$, and likewise for $z' \in \ell_{y'}$.

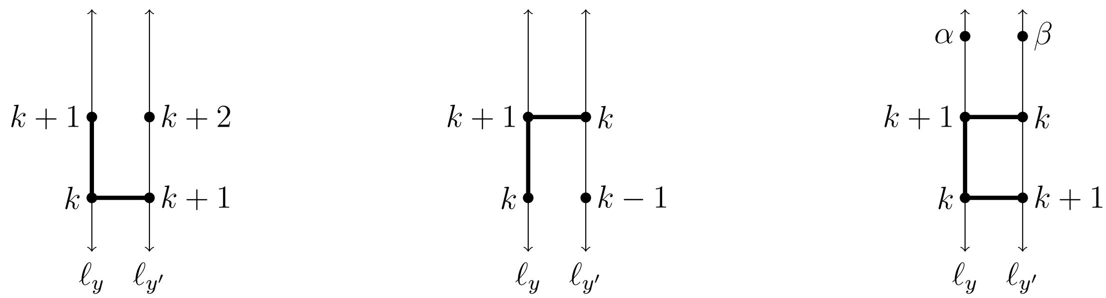
Later on, we will need that for an arbitrary "good" edge $e \in G_y$ it holds:

$$\sum_{y' \sim y} \sum_{\substack{\tilde{e} \in \tilde{E}_{y,y'} \\ h_{S_n}(\tilde{e}) = e}} \omega_{h_{S_n}(\tilde{e})} = \big|\{y' \sim y\}\big|\, \omega_e. \tag{2.92}$$

Note that $|\{y' \sim y\}| \leq 2m$ for all y, with equality unless y is a boundary point (implicitly we assume that $y' \in \{-n, \dots, n\}^{m-1}$). Argument for (2.92): We observe that for each $y' \sim y$, by using the second and third constraints and considering cases, there is a unique cross edge $e_{s,y'}$ between ℓ_y and $\ell_{y'}$ such that $h_{S_n}(e_{s,y'}) = e$. For a proof of this simple fact we refer to Figure 2.5. The identity (2.92) follows then immediately.

We will also need count $|G_y|$. Heuristically, since the slope s is close to 1, G_y must be a large subset of $E(\ell_y)$. To be precise, recall that $|A_y| \geq 2n - 4\lceil \varepsilon n \rceil$ by construction, and that G_y is the subset of edges $e \in A_y$ that satisfy the three constraints above. The second constraint actually implies the first, so to count G_y we simply count how many edges in A_y satisfy the last two constraints. Actually we count the complement, i.e. how many edges do not satisfy these two constraints. Indeed, each "bad" vertex in ℓ_y (in the sense described after the constraints) causes at most two edges in $E(\ell_y)$ to violate the second constraint. Likewise, each "bad" vertex in an adjacent line ℓ_y' causes at most two edges in $E(\ell_y)$ to violate the second constraint. All other edges in A_y are "good," i.e. are included in G_y.

It remains to count the "bad" vertices in any line ℓ_y. Since $s_1 > 1 - \varepsilon$ and since h_{S_n} approximates the slope-s height function $h^s_{S_n}$ on ∂S_n, the height values $h_{S_n}(-n, y)$ and $h_{S_n}(+n, y)$ on the endpoints of ℓ_y differ by at least $2n - 4\varepsilon n$. Since h_{S_n} is a graph homomorphism, it



(a) Case 1 (both adjacent height values larger): Clearly there is one edge between ℓ_y and $\ell_{y'}$ that is mapped to $e = e_{k,k+1}$. Suppose that another cross edge has heights k and $k+1$. Then its left endpoint would have either height k or height $k+1$, which contradicts the fact that the two labelled vertices in ℓ_y are "good," i.e. that their heights occur only once in ℓ_y.

(b) Case 2 (both adjacent height values smaller): Again there is one edge between ℓ_y and $\ell_{y'}$ that is mapped to $e = e_{k,k+1}$, and again no other vertices in ℓ_y can have either height k or height $k+1$.

(c) Case 3 (cannot occur because $e \in G_y$): Here there would be two edges between the lines that both map to $e_{k,k+1}$. But since the vertex at height k in ℓ_y is "good", the vertex labelled α must have height $k+2$. Likewise since the vertex at height $k+1$ in $\ell_{y'}$ is "good", vertex β must have height $k-1$. Since $\alpha \sim \beta$, this violates the graph homomorphism property.

Figure 2.5: Consideration of cases for part of the proof of Lemma 7. The claim to be shown is: given $e \in G_y$ (say $e = e_{k,k+1}$), there is a unique cross edge $e_{s,y'} \in \tilde{E}_{y,y'}$ which is mapped to e by the height function h_{S_n}. In the figure, the vertices are labelled by their heights, i.e. by the values of h_{S_n}. The bolded edge in ℓ_y is $e_s \in E(\ell_y)$, i.e. the unique edge in ℓ_y with $h_{S_n}(e_s) = e$. In Figure 2.5a and Figure 2.5b, the bolded edge between the lines is the unique edge between the lines with height $e_{k,k+1}$. Figure 2.5c shows two such edges, but in fact this case cannot occur. By the homomorphism property, these three cases exhaust the possibilities for heights on the two vertices in $\ell_{y'}$ that are adjacent to the endpoints of e_s.

maps the $2n+1$ vertices in ℓ_y surjectively onto the set of $\geq 2n - 4\lceil \varepsilon n \rceil + 1$ integers between the heights of the endpoints. By the pigeonhole principle, at most $8\lceil \varepsilon n \rceil$ of these integers occur with multiplicity ≥ 2, i.e. at most $8\lceil \varepsilon n \rceil$ vertices are "bad." Thus

$$\begin{aligned} |G_y| &\geq |A_y| - 2\left|\left\{\text{“bad” vertices in } \ell_y \text{ or } \ell_{y'} \text{ (for } y' \sim y)\right\}\right| && (2.93)\\ &\geq \underbrace{2n - 4\lceil \varepsilon n \rceil}_{|A_y|} - 2 \cdot \underbrace{(2m+1)}_{\#\text{ lines}} \cdot \underbrace{8\lceil \varepsilon n \rceil}_{\text{“bad” vertices per line}} && (2.94)\\ &= 2n - (32m+20)\lceil \varepsilon n \rceil && (2.95)\\ &\geq 2n - 52m\lceil \varepsilon n \rceil\,. && (2.96) \end{aligned}$$

Now we work towards the Hamiltonian estimate (2.83). Let $e \in G_y$, and recall that e_s is the unique edge in $E(\ell_y)$ such that $h_{S_n}(e_s) = e$, and that $e_{s,y'}$ is the unique cross edge between ℓ_y and $\ell_{y'}$ such that $h_{S_n}(e_{s,y'}) = e$. As a result (recall the definitions of $\tilde{H}_{\ell_y}$ and $\tilde{E}_{y,y'}$ from (2.91) above):

$$\begin{aligned} \tilde{H}_{\ell_y}(h_{S_n}) &= \Big(\sum_{\tilde{e} \in E(\ell_y)} \omega_{h_{S_n}(\tilde{e})}\Big) + \frac{1}{2}\Big(\sum_{y' \sim y}\sum_{\tilde{e} \in \tilde{E}_y} \omega_{h_{S_n}(\tilde{e})}\Big)\\ &\overset{(2.92)}{=} \Big(\sum_{e \in G_y} \omega_e + \sum_{\substack{\tilde{e} \in E(\ell_y)\\ h_{S_n}(\tilde{e}) \notin G_y}} \omega_{h_{S_n}(\tilde{e})}\Big)\\ &\qquad + \frac{1}{2}\Big(\left|\{y' \sim y\}\right| \sum_{e \in G_y} \omega_e + \sum_{y' \sim y} \sum_{\substack{\tilde{e} \in \tilde{E}_{y,y'}\\ h_{S_n}(\tilde{e}) \notin G_y}} \omega_{h_{S_n}(\tilde{e})}\Big)\,, \end{aligned} \tag{2.97}$$

so

$$\Big|\tilde{H}_{\ell_y}(h_{S_n}) - \big(\tfrac{1}{2}\big|\{y' \sim y\}\big| + 1\big) \sum_{e \in G_y} \omega_e\Big| \tag{2.98}$$

$$\leq C_\omega\big(|E(\ell_y)| - |G_y|\big) + \frac{1}{2}\sum_{y' \sim y} C_\omega\big(|\tilde{E}_{y,y'}| - |G_y|\big) \tag{2.99}$$

$$\overset{(2.96)}{\leq} 52mC_\omega\lceil\varepsilon n\rceil + \frac{1}{2}\sum_{y' \sim y} C_\omega\big(52m\lceil\varepsilon n\rceil + 1\big) \tag{2.100}$$

$$\leq 52mC_\omega\lceil\varepsilon n\rceil(1+m) + mC_\omega \tag{2.101}$$

$$\leq 104m^2C_\omega\lceil\varepsilon n\rceil + mC_\omega \tag{2.102}$$

$$\leq 105m^2(2n+1)C_\omega\varepsilon\,. \tag{2.103}$$

(In the last line, we assume that $n \geq \frac{1}{\varepsilon}$, so that $(2n+1)\varepsilon \geq \lceil\varepsilon n\rceil \geq 1$.)

Because $s'_1 = 1$, $h^{s'}_{S_n}|_{\ell_y}$ is an injection, the three bullet points above are also satisfied with $h^{s'}_{S_n}$ in place of h_{S_n}. Therefore the calculation above also applies with $h'^{s}_{S_n}$ in place of h_{S_n}, so

$$\Big|\tilde{H}_{\ell_y}(h^{s'}_{S_n}) - \big(\tfrac{1}{2}\big|\{y' \sim y\}\big| + 1\big) \sum_{e \in G_y} \omega_e\Big| \leq 105m^2(2n+1)C_\omega\varepsilon\,. \tag{2.104}$$

By the triangle inequality,

$$\big|\tilde{H}_{\ell_y}(h_{S_n}) - \tilde{H}_{\ell_y}(h^{s'}_{S_n})\big| \leq 210m^2(2n+1)C_\omega\varepsilon\,. \tag{2.105}$$

By summing over $y \in \{-n, \cdots, n\}^{m-1}$, we get the desired inequality (2.83), i.e.

$$\big|H_{S_n}(h_{S_n}) - H_{S_n}(h^{s'}_{S_n})\big| \leq 210m^2(2n+1)^mC_\omega\varepsilon\,. \tag{2.106}$$

□

Both Lemma 4 and Lemma 7 imply that the microscopic entropy is robust to changes in boundary data, but they apply in different regimes. The former result applies when the boundary data has slope s with norm $|s|_\infty$ bounded away from 1, and the latter when the slope s has norm close to 1. For convenience later on, we combine the two results into a single theorem.

Theorem 7. *For any $\varepsilon \in (0, \frac{1}{9})$ and any slope $s \in [-1,1]^m$, there exist $A = A(s,\varepsilon) > 0$, $B = B(s,\varepsilon) > 0$, and $n_0 = \lceil \frac{1}{\varepsilon} \rceil \in \mathbb{N}$ such that, for any $n \geq n_0$ and any boundary height function $h_{\partial S_n} \in M(\partial S_n, h^s_{\partial S_n}, \varepsilon)$,*

$$\begin{aligned} \mathrm{ent}_{An}(s,\omega) - C_\omega\, \theta(\varepsilon) \leq \mathrm{Ent}\big(M(S_n, h_{\partial S_n}), \omega\big) \\ \leq \mathrm{ent}_{Bn}(s,\omega) + C_\omega\, \theta(\varepsilon). \end{aligned} \tag{2.107}$$

Moreover, the functions $A(s,\varepsilon)$ and $B(s,\varepsilon)$ are bounded away from 0 and ∞ uniformly in s and ε. More precisely,

$$1 \leq A(s,\varepsilon) \leq \big(1 + 2\varepsilon^{1/2} + \tfrac{1}{n}\big) < \infty \tag{2.108}$$

and

$$0 < \big(1 - 2\varepsilon^{1/2} - \tfrac{1}{n}\big) < B(s,\varepsilon) \leq 1. \tag{2.109}$$

Proof of Theorem 7. Take $\alpha = \varepsilon^{1/2}$ and proceed according to two cases. For slopes s with $|s|_\infty \leq 1 - \alpha$, use Lemma 4 to choose $A = n^+/n \approx (1 + 2\varepsilon^{1/2})$ and $B = n^-/n \approx (1 - 2\varepsilon^{1/2})$. Note that $\varepsilon < \frac{1}{9}$ implies that $\varepsilon < \frac{\alpha}{2}$ and $n \geq \frac{1}{\varepsilon} \geq (1 - 2\varepsilon^{1/2})^{-1}$, as required by the lemma. Moreover $1 - 2\varepsilon^{1/2} - \frac{1}{n} > \frac{2}{9}$, so B is indeed bounded away from 0. The error terms $\theta(\frac{\varepsilon}{\alpha})$ from the lemma are equivalent to $\theta(\varepsilon^{1/2}) = \theta(\varepsilon)$.

For slopes with $|s|_\infty > 1 - \alpha$, take $A = B = 1$ and apply Lemma 7 twice, using $\alpha = \varepsilon^{1/2}$ in place of ε: once for the boundary height function $h_{\partial S_n}$ given in the statement of the theorem, and once for the canonical boundary height function $h^s_{\partial S_n}$. The estimate on

$$\big|\mathrm{Ent}(M(S_n, h_{\partial S_n}), \omega) - \mathrm{ent}_n(s,\omega)\big| \tag{2.110}$$

follows from the triangle inequality. $\square$

The robustness results above focused on boundary height functions that differed at macroscopic scale, i.e. $|h_{\partial S_n} - \tilde{h}_{\partial S_n}|_u \leq \varepsilon n$. For boundary height functions with sub-linear differences, we will derive stronger robustness results. Lemma 8 addresses the case where the two boundary height functions differ at only a single point on ∂S_n, and Corollary 1 extends to the sub-linear case (actually, only to $|h_{\partial S_n} - \tilde{h}_{\partial S_n}|_u = o(\frac{\log n}{n})$, but that is sufficient for our purposes.)

Lemma 8 (Robustness for minimally different boundary height functions)**.** *Fix $n \in \mathbb{N}$, and let $h^+_{\partial S_n}$ and $\tilde{h}^-_{\partial S_n}$ be two boundary height functions on the hypercube S_n which differ at exactly one point $z_0 \in \partial S_n$, i.e. $h^+_{\partial S_n}|_{S_n \setminus \{z_0\}} = h^-_{\partial S_n}|_{S_n \setminus \{z_0\}}$ and $h^+_{\partial S_n}(z_0) = h^-_{\partial S_n}(z_0) + 2$.*

Then,

$$\Big|\mathrm{Ent}_{S_n}\big(M(S_n, h^+_{\partial S_n}), \omega\big) - \mathrm{Ent}_{S_n}\big(M(S_n, h^-_{\partial S_n}), \omega\big)\Big| \leq \frac{4mC_\omega + \log(2n)}{|S_n|}\,. \tag{2.111}$$

Remark 8. The $\log(2n)$ term is necessary at least in some extreme cases. For example, suppose that $\omega \equiv 0$, $m = 1$, $z_0 = -n$, $h^+_{\partial S_n}(-n) = 2$, $h^-_{\partial S_n}(-n) = 0$, and $h^\pm_{\partial S_n}(n) = 2n$. Then $\mathrm{Ent}_{S_n}(M(S_n, h^+_{\partial S_n}), \mathbf{0}) = -\frac{1}{n}\log(2n)$ and $\mathrm{Ent}_{S_n}(M(S_n, h^-_{\partial S_n}), \mathbf{0}) = 0$; cf. Lemma 5 and Lemma 6 for calculations.

Proof of Lemma 8. For concreteness and w.l.o.g., assume that the boundary values at z_0 are $h^-_{\partial S_n}(z_0) = 0$ and $h^+_{\partial S_n}(z_0) = 2$. (Technically this assumption is only valid if z_0 has even parity because we require that height functions preserve parity, and one should instead assume e.g. that $h^\pm_{\partial S_n}(z_0) \in \{1, 3\}$ in the other case. For simplicity we ignore this detail in the rest of the proof.)

Consider the line $z_0, z_1, \ldots, z_{2n}$ of points in S_n starting from z_0 and going into S_n, perpendicular to the boundary. Classify each height function $h^+_{S_n} \in M(S_n, h^+_{\partial S_n})$ based on the number of initial "up" steps, i.e.

$$k_{\mathtt{up}}(h^+_{S_n}) := \max\big\{\tilde{k} \geq 0 \,|\, h^+_{S_n}(z_k) = h^+_{S_n}(z_{k-1}) + 1 \,\text{for}\; 1 \leq k \leq \tilde{k}\big\}\,. \tag{2.112}$$

Note that from our initial assumption, $h^+_{S_n}(z_k) = k+2$ for $0 \le k \le k_{\mathtt{up}}$. Necessarily $k_{\mathtt{up}}(h^+_{S_n}) < 2n$, since if $h^+_{S_n}$ went up along all $2n$ edges, then the values $h^-_{S_n}(z_{2n}) = h^+_{S_n}(z_{2n}) = 2n+2$ and $h^-_{S_n}(z_0) = 0$ would violate the Kirszbraun theorem.

On the line segment $\{z_0, \dots, z_{n_{\mathtt{up}}}\} \subset S_n$, $h^+_{S_n}$ is "too high," in the sense that no height function in $M(S_n, h^-_{\partial S_n})$ can match it. But by the Kirszbraun theorem, there exists $h^-_{S_n} \in M(S_n, h^-_{\partial S_n})$ such that $h^-_{S_n}(z_{n_{\mathtt{up}}+1}) = h^+_{S_n}(z_{n_{\mathtt{up}}+1})$. In fact, we may define $h^-_{S_n}$ by

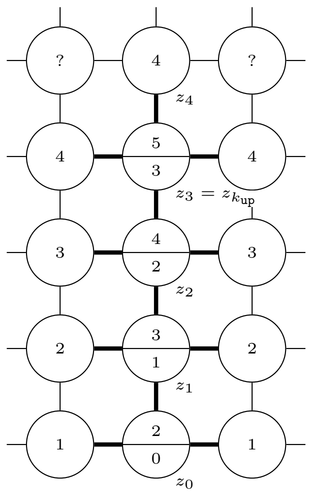
$$h^-_{S_n}(z) = \begin{cases} k = h^+_{S_n}(z) - 2, & \text{if } z = z_k \text{ for } 0 \le k \le k_{\mathtt{up}}, \text{ and} \\ h^+_{S_n}(z), & \text{otherwise}\,. \end{cases} \tag{2.113}$$

It follows that $h^+_{S_n}$ and $h^-_{S_n}$ have the same Hamiltonian, except for the contribution from the edges incident to a vertex z_k ($0 \le k \le k_{\mathtt{up}}$). There are $(2m-1)(k_{\mathtt{up}}+1)$ such edges, which leads to the naïve estimate $|H_{S_n}(h^+_{S_n}, \omega) - H_{S_n}(h^-_{S_n}, \omega)| \le (2m-1)(k_{\mathtt{up}}+1)C_\omega$. This estimate is not useful because $k_{\mathtt{up}}$ on the right-hand side leads to an error of order n in the worst case. However, as shown in Figure 2.6, a more careful estimate is possible. Indeed, both $h^+_{S_n}$ and $h^-_{S_n}$ map the edges e in question to the same collection of edges $\{e_{k,k+1} \,|\, 0 \le k \le k_{\mathtt{up}}\} \subset E(\mathbb{Z})$, with each $e_{k,k+1}$ repeated about $2m-1$ times. We omit the details, but a careful count of the edge heights yields the inequality

$$\big|H_{S_n}(h^+_{S_n}, \omega) - H_{S_n}(h^-_{S_n}, \omega)\big| \le 4mC_\omega\,. \tag{2.114}$$

Now we turn to the entropy inequality. For $0 \le k < 2n$, let

$$M_k := \big\{h^+_{S_n} \in M(S_n, h^+_{\partial S_n}) \,\big|\, k_{\mathtt{up}}(h^+_{S_n}) = k\big\}. \tag{2.115}$$



(a) The values of the height functions $h^+_{S_n}$ and $h^-_{S_n}$ from the proof of Lemma 8. On the vertices $z_0, \dots, z_{k_{\mathtt{up}}}$ where the two height functions differ, the larger value is the height that $h^+_{S_n}$ takes and the smaller value is $h^-_{S_n}$. Here $k_{\mathtt{up}} = 3$, since $h^+_{S_n}$ increases across the first three edges in the center line. The $(2m-1)(k_{\mathtt{up}}+1)$ shaded edges are exactly the set up edges incident to any of $z_0, \dots, z_{k_{\mathtt{up}}}$, and these are the only edges on which $h^{\pm}_{S_n}$ differ.

$e \in E(\mathbb{Z})$	$h^+_{S_n}$	$h^-_{S_n}$
$e_{0,1}$	0	$2m-1$
$e_{1,2}$	$2m-2$	$2m-1$
$e_{2,3}$	$2m-1$	$2m-1$
$e_{3,4}$	$2m-1$	$2m-1$
$\vdots$	$\vdots$	$\vdots$
$e_{k_{\mathtt{up}},k_{\mathtt{up}}+1}$	$2m-1$	$2m-1$
$e_{k_{\mathtt{up}}+1,k_{\mathtt{up}}+2}$	$2m$	0

(b) Number of shaded edges on which $h^+_{S_n}$, $h^-_{S_n}$ attain certain heights. For example, from the last row of the table: $h^+_{S_n}(e) = e_{k_{\mathtt{up}}+1,k_{\mathtt{up}}+2}$ for all $2m$ edges incident on $z_{k_{\mathtt{up}}}$. In the difference $H_{S_n}(h^+_{S_n}) - H_{S_n}(h^-_{S_n})$, the bulk of the height values in the table cancel, leaving only boundary terms. That is why the bound in (2.114) does not depend on $k_{\mathtt{up}}$.

Figure 2.6: Explanation of inequality (2.114) from the proof of Lemma 8.

Then the sets M_k ($0 \le k < 2n$) partition $M(S_n, h^+_{\partial S_n})$, so

$$\begin{aligned}
&\mathrm{Ent}_{S_n}\big(M(S_n, h^+_{\partial S_n}), \omega\big) && (2.116)\\
&\quad = -\frac{1}{|S_n|} \log \sum_{k=0}^{2n-1} \sum_{h^+_{S_n} \in M_k} \exp\big(H_{S_n}(h^+_{S_n}, \omega)\big) && (2.117)\\
&\quad \overset{(2.114)}{\ge} -\frac{1}{|S_n|} \log \sum_{k=0}^{2n-1} \sum_{h^+_{S_n} \in M_k} \exp\big(H_{S_n}(h^-_{S_n}, \omega) + 4mC_\omega\big) && (2.118)\\
&\quad \ge -\frac{1}{|S_n|} \log \sum_{k=0}^{2n-1} \sum_{h^-_{S_n} \in M(S_n, h^-_{\partial S_n})} \exp\big(H_{S_n}(h^-_{S_n}, \omega) + 4mC_\omega\big) && (2.119)\\
&\quad = \mathrm{Ent}_{S_n}\big(M(S_n, h^-_{\partial S_n}), \omega\big) - \frac{4mC_\omega + \log(2n)}{|S_n|}\,. && (2.120)
\end{aligned}$$

The reverse inequality is derived by exchanging the roles of $h^\pm_{\partial S_n}$, considering the number $k_{\mathtt{down}}$ of initial downward steps of $h^-_{\partial S_n}$ on the line $\{z_0, \ldots, z_{2n}\}$, and proceeding as before with the necessary changes. □

Lemma 8 applies only when the two boundary height functions $h^+_{S_n}$ and $h^-_{S_n}$ differ minimally. However by applying Lemma 8 repeatedly, we can compare two height functions with more differences. That idea is captured in the following corollary.

Corollary 1 (Robustness with respect to sub-linear height differences)**.** *Let $h_{\partial S_n}$ and $\tilde{h}_{\partial S_n}$ be boundary height functions on S_n, and let $M = \|h_{\partial S_n} - \tilde{h}_{\partial S_n}\|_\infty$. Then*

$$\begin{aligned}
&\Big|\mathrm{Ent}_{S_n}\big(M(S_n, h_{\partial S_n}), \omega\big) - \mathrm{Ent}_{S_n}\big(M(S_n, \tilde{h}_{\partial S_n}), \omega\big)\Big| \\
&\qquad \le \frac{M}{2}\big(4mC_\omega + \log(2n)\big)\frac{|\partial S_n|}{|S_n|}\,.
\end{aligned} \tag{2.121}$$

Remark 9. The main idea of the proof is to interpolate the boundary height function from $h_{\partial S_n}$ to $\tilde{h}_{\partial S_n}$, where each step in the interpolation changes the value of the boundary height function at exactly one boundary point. Note that each interpolation step changes the height by 2 at that distinguished boundary point, which is the reason for the factor $\frac{M}{2}$ rather than simply M. Given such an interpolation, all that remains is to apply Lemma 8 and the triangle inequality.

Proof of Corollary 1. We claim that there exists a finite sequence $h^{(1)}_{\partial S_n}, \ldots, h^{(k)}_{\partial S_n}$ such that each pair $h^{(j)}_{\partial S_n}$ and $h^{(j+1)}_{\partial S_n}$ differ at exactly one point, such that $h^{(1)}_{\partial S_n} = h_{\partial S_n}$ and $h^{(k)}_{\partial S_n} = \tilde{h}_{\partial S_n}$, and such that $k \leq \frac{M}{2}|\partial S_n|$. Each element of the sequence is constructed from the previous element by a "flip" operation: Given a boundary height function $h^{(j)}_{\partial S_n}$ and a vertex $z_j \in \partial S_n$ where all the neighboring vertices $z' \in \partial S_n$, $z' \sim z_j$ have the same height $h_{\partial S_n}(z') = a \in \mathbb{Z}$, the height function $h^{(j+1)}_{\partial S_n}$ is identical to $h^{(j)}_{\partial S_n}$ on $\partial S_n \setminus \{z_j\}$ and takes the other valid value on z_j. Specifically, if $h^{(j)}_{\partial S_n}(z_j) = a+1$, then $h^{(j+1)}_{\partial S_n}(z_j) = a-1$; otherwise $h^{(j+1)}_{\partial S_n}(z_j) = a+1$.

It remains to show that the vertices $z_1, \ldots, z_{k-1}$ can be chosen so that $h^{(k)}_{\partial S_n} = \tilde{h}_{\partial S_n}$ and so that $k \leq \frac{M}{2}|\partial S_n|$. To prove both these points, consider the metric $d : M(\partial S_n) \times M(\partial S_n) \to \mathbb{Z}$ defined by

$$d(h'_{\partial S_n}, h''_{\partial S_n}) := \sum_{z \in \partial S_n} \left| h'_{\partial S_n}(z) - h''_{\partial S_n}(z) \right|. \tag{2.122}$$

As long as $d(h^{(j)}_{\partial S_n}, \tilde{h}_{\partial S_n}) > 0$, we will find a vertex z_j for which the flip operation both is valid and decreases the distance d. Towards this end, let $E_j := \{z \in \partial S_n \,|\, h^{(j)}_{\partial S_n}(z) > \tilde{h}_{\partial S_n}(z)\}$. If $E_j \neq \varnothing$, choose $z_j := \operatorname{argmax}_{z \in E_j} h^{(j)}_{\partial S_n}$.

We claim that flipping at z_j is valid, and more specifically that for all neighbors $z' \sim z_j$ in ∂S_n, $h^{(j)}_{\partial S_n}(z') = h^{(j)}_{\partial S_n}(z_j) - 1$. Indeed, there are two cases. If $h^{(j)}_{\partial S_n}(z') = \tilde{h}_{\partial S_n}(z')$ for any $z' \sim z_j$, then necessarily $\tilde{h}_{\partial S_n}(z_j) = h^{(j)}_{\partial S_n}(z_j) - 2$ and $\tilde{h}_{\partial S_n}(z') = h^{(j)}_{\partial S_n}(z') = h^{(j)}_{\partial S_n}(z_j) - 1$ for all $z' \sim z$. Otherwise all $z' \sim z$ are also in E_j, so the claim follows since z_j maximizes $h^{(j)}_{\partial S_n}$ over E_j. So as claimed, it is valid to flip the height function $h^{(j)}_{\partial S_n}$ at z_j, and this flip decreases the difference $|h^{(j+1)}_{\partial S_n}(z_j) - \tilde{h}_{\partial S_n}(z_j)|$ by two, and therefore decreases the distance $d(h^{(j+1)}_{\partial S_n}, \tilde{h}_{\partial S_n})$ by two.

If E_j is empty, use instead the set $F_j := \{z \in \partial S_n \,|\, h^{(j)}_{\partial S_n}(z) < \tilde{h}_{\partial S_n}(z)\}$, pick $z_j := \operatorname{argmin}_{z \in F_j} h^{(j)}_{\partial S_n}$, and repeat the argument, changing inequalities and signs accordingly. If F_j is also empty, then $h^{(j)}_{\partial S_n} = \tilde{h}_{\partial S_n}$ and the process is complete.

At most $\frac{1}{2} d(h_{\partial S_n}, \tilde{h}_{\partial S_n}) \leq \frac{M}{2}|\partial S_n|$ steps are needed in total, since each step decreases the distance by 2.

To complete the proof of the corollary, apply Lemma 8 to each pair $\{h^{(j)}_{\partial S_n}, h^{(j+1)}_{\partial S_n}\}$ and use the triangle inequality. □

2.4.3 Existence and equivalence

Recall from Definition 9 that the quenched local surface tension is defined as the limit of the quenched microscopic surface tension. Because of the random potential ω, the existence of this limit is not obvious. We prove the existence of the limit using an ergodic theorem for almost superadditive random families.

First, we introduce the notation needed for stating the ergodic theorem. Let $\mathcal{B}$ denote the set of all (non-empty) boxes in $\mathbb{Z}^m$, i.e.

$$\mathcal{B} = \left\{ \big([a_1, b_1) \times \cdots \times [a_m, b_m) \big) \cap \mathbb{Z}^m \;\middle|\; a_1 < b_1, \, \ldots, \, a_m < b_m \in \mathbb{Z}^m \right\}.$$

Note that the sets $S_n := [-n, n]^m \cap \mathbb{Z}^m$ are included in $\mathcal{B}$. We say that a family of L^1 random variables $F = (F_B)_{B \in \mathcal{B}}$ is *almost superadditive* if, for any finitely many disjoint boxes $B_1, \ldots, B_n \in \mathcal{B}$ whose union $B = B_1 \cup \cdots \cup B_n$ also lies in $\mathcal{B}$,

$$F_B \geq \sum_{i=1}^{n} F_{B_i} - A \sum_{i=1}^{n} |\partial B_i| \quad \text{a.s.}\,, \tag{2.123}$$

where $A = A(\omega) : \Omega \to [0, \infty)$ is an L^1 random variable, and where $\partial B_i = \{x \in B_i \,|\, \exists y \in \mathbb{Z}^m \setminus B_i, \, x \sim y\}$ is the inner boundary of B_i.

Theorem 8 (Ergodic theorem for almost superadditive random families)**.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, let* $\tau = (\tau_u)_{u \in \mathbb{Z}^m}$ *be a family of measure-preserving transformations on* Ω*, and let* $F = (F_B)_{B \in \mathcal{B}}$ *be a family of* L^1 *random variables satisfying the following three conditions:*

- F *is almost superadditive, i.e.* F *satisfies* (2.123),

- *For all $u \in \mathbb{Z}^m$,*

$$\lim_{n\to\infty} \sup_{u\in\mathbb{Z}^m} \frac{1}{|S_n|}\Big|F_{u+S_n} - F_{S_n} \circ \tau_u\Big| = 0\,, \tag{2.124}$$

where $u + B = \{u + x \,|\, x \in B\}$ is the translation of B by u.

- *The quantity $\tilde{\gamma}(F) = \limsup_{n\to\infty} \frac{1}{|S_n|}\,\mathbb{E}[F_{S_n}]$ is finite.*

Then the limit $\lim_{n\to\infty} \frac{1}{|S_n|}\,F_{S_n}$ exists almost surely and in L^1. If moreover $\{\tau_u\}_{u\in\mathbb{Z}^m}$ is ergodic, then the limit is

$$\lim_{n\to\infty} \frac{1}{|S_n|}\,F_{S_n} = \tilde{\gamma}(F). \tag{2.125}$$

This theorem is based on [AK81, Theorem 2.4], which is a multidimensional extension of the subadditive ergodic theorem proven in [Kin68, Lig85, KMT20b] among other sources. The version stated here is adapted to notion of almost superadditivity that the quenched microscopic entropy satisfies. Now let us turn to the application of this ergodic theorem:

Lemma 9 (Existence of the quenched local surface tension)**.** *For almost every realization ω of the random potential, the limit* (2.46) *exists.*

The proof is a straightforward application of the ergodic theorem.

Proof of Lemma 9. Fix $s \in [-1,1]^m$. Let the family of measure-preserving transformations $\tau = (\tau_u)_{u\in\mathbb{Z}^m}$ be given by

$$\big(\tau_u\omega\big)_e := \omega_{e-[s\cdot u]_u \bmod 2} \quad \text{for } e \in E(\mathbb{Z}) \text{ and } u \in \mathbb{Z}^m. \tag{2.126}$$

Define the random process $F = (F_B)_{B\in\mathcal{B}}$ by

$$F_B := -|B|\ \mathrm{Ent}\big(M(B, h^s_{\partial B}), \omega\big) = \log Z_\omega\big(M(B, h^s_{\partial B})\big)\,. \tag{2.127}$$

Now we verify the hypotheses of the ergodic theorem (Theorem 8). First, the fact that $|\omega_e| \le C_\omega$ for all edges $e \in E(\mathbb{Z})$ implies that each variable F_B ($B \in \mathcal{B}$) is in L^1.

Next, the almost superadditivity property (2.123) follows from distributivity:

$$\begin{aligned}\sum_{i=1}^{n} F_{B_i} &= \log \prod_{i=1}^{n} \sum_{h_{B_i}\in M(B_i,h^s_{\partial B_i})} \exp\big(H_{B_i}(h_{B_i},\omega)\big) \\ &= \log \sum_{\substack{h_{B_1}\in M(B_1,h^s_{\partial B_1})\\ \cdots \\ h_{B_n}\in M(B_n,h^s_{\partial B_n})}} \exp\left(\sum_{i=1}^{n} H_{B_i}(h_{B_i},\omega)\right).\end{aligned} \tag{2.128}$$

The final sum is indexed by n-tuples of height functions, i.e. it is the sum over the Cartesian product of the sets $M(B_i,h^s_{\partial B_i})$. This Cartesian product is a subset of $M(B,h_B)$, so

$$\sum_{i=1}^{n} F_{B_i} \le \log \sum_{h_B\in M(B,h^s_{\partial B})} \exp\left(\sum_{i=1}^{n} H_{B_i}(h_B|_{B_i},\omega)\right). \tag{2.129}$$

The quantity on the right-hand side of (2.129) differs from F_B by at most $mC_\omega \sum_{i=1}^{n}|\partial B_i|$, since the Hamiltonian terms in (2.129) do not include edges that cross from one box B_i to another box B_j. This error term satisfies (2.123).

Now let us show that F satisfies the translation invariance estimate (2.124). For $h_{\partial(u+B)} \in M(\partial(u+B))$, consider the shifted boundary height function $\Psi_u h_{\partial(u+B)} \in M(\partial B)$ defined by

$$(\Psi_u h_{\partial(u+B)})(z) := h_{\partial(u+B)}(u+z) - \lfloor s\cdot u\rfloor \qquad \text{for } z\in\partial B\,. \tag{2.130}$$

Since both $h^s_{\partial B}$ and $h^s_{\partial(u+B)}$ are rounded to the nearest integer (of appropriate parity), the shifted boundary height function $\Psi_u h^s_{\partial(u+B)}$ may not agree exactly with $h^s_{\partial B}$. However it holds that

$$\big|\Psi_u h_{\partial(u+B)}(z) - h_{\partial B}(z)\big| \le 4 \qquad \text{for all } z\in\partial B\,. \tag{2.131}$$

Therefore by Corollary 1,

$$\big|F_{u+B} - F_B\circ\tau_u\big| \le |B|\,\theta\big(\tfrac{1}{n}\big)\,. \tag{2.132}$$

The last condition to check is $\tilde\gamma(F) = \limsup_{n\to\infty}\frac{1}{|S_n|}\mathbb{E}[F_{S_n}] < \infty$, which follows from boundedness of the quenched entropy. Indeed by Lemma 3, the inequality $F_B \le m|B|C_\omega$

holds almost surely, so $\tilde{\gamma}(F) \le \mathbb{E}(C_\omega) < \infty$.

At this point we have checked all the hypotheses of the ergodic theorem (Theorem 8). From the ergodic theorem we conclude that the pointwise limit

$$\mathrm{ent}(s,\omega) = \lim_{n\to\infty} \mathrm{ent}_n(s,\omega) = \lim_{n\to\infty} \frac{1}{|S_n|} F_{S_n}(\omega) \tag{2.133}$$

exists almost surely. In addition, when $s \neq 0$, the family of measure-preserving transformations $(\tau_u)_{u\in\mathbb{Z}^m}$ is ergodic with respect to $\mathbb{P}$, since the family includes every shift $\omega \mapsto (\omega_{k+e})_{e\in E(\mathbb{Z})}$ for $k \in \mathbb{Z}$. Therefore whenever $s \neq 0$, the limit $\mathrm{ent}(s,\omega)$ is almost surely equal to its expectation, $\mathbb{E}[\mathrm{ent}(s,\omega)] = \mathrm{ent}_{\mathtt{an}}(s)$. □

The failure of ergodicity in the case $s = 0$ is evident from the definition of $(\tau_u)_{u\in\mathbb{Z}^m}$ in (2.126): there we have $(\tau_u\omega)_e := \omega_{e-[s\cdot u]_{u \bmod 2}}$ for each $e \in E(\mathbb{Z})$. When $s = 0$ the quantity $s \cdot u$ is zero even as $u \to \infty$, so the entire family of transformations $(\tau_u)_{u\in\mathbb{Z}^m}$ is actually finite rather than ergodic. As such, a different argument is needed for $s = 0$. We credit Marek Biskup for suggesting the following argument.

Lemma 10 (Equivalence of quenched and annealed local surface tension)**.** *For almost every ω, it holds that*

$$\mathrm{ent}(s,\omega) = \mathrm{ent}_{\mathtt{an}}(s). \tag{2.134}$$

Moreover, the quenched microscopic surface tension $\mathrm{ent}_n(s,\omega)$ *converges in* L^1 *to* $\mathrm{ent}_{\mathtt{an}}(s)$.

Proof of Lemma 10. For $s \neq 0$, the desired identity (2.134) follows from the ergodic theorem, as mentioned at the end of the proof of Lemma 9.

For $s = 0$, we will establish translation invariance of $\mathrm{ent}(s,\omega)$ directly. First we replace the environmental shift τ_2 by a shift in heights, i.e.

$$\mathrm{Ent}_{S_n}\left(M(S_n, h^0_{\partial S_n})\right) \circ \tau_2 = \mathrm{Ent}_{S_n}\left(M(S_n, h^{0\cdot x+2}_{\partial S_n})\right). \tag{2.135}$$

This identity is justified simply by expanding definitions; both sides are equal to

$$-\frac{1}{|S_n|}\log\sum_{h_{S_n}}\exp\Bigg(\sum_{e}\omega_{h_{S_n}(e)+2}\Bigg), \tag{2.136}$$

where the first sum runs over $h_{S_n} \in M(S_n, h^0_{\partial S_n})$ and the second runs over $e \in E(S_n)$.

Now, the square S_n sits inside of S_{n+2}. The boundary values $h^{0\cdot x+2}_{\partial S_n}$ and $h^0_{\partial S_{n+2}}$ satisfy the Kirszbraun criterion (2.59); in fact, each $h \in M(S_n, h^{0\cdot x+2}_{\partial S_n})$ admits a unique extension $\tilde{h}$ in $M(S_{n+2}, h^0_{\partial S_{n+2}})$. Since $\tilde{h}$ is an extension of h to a domain with $O(n^{m-1})$ more points and $O(n^{m-1})$ more edges, the Hamiltonians satisfy

$$\big|H_{S_n}(h,\omega) - H_{S_{n+2}}(\tilde{h},\omega)\big| \le cn^{m-1}C_\omega \tag{2.137}$$

for some $c > 0$. Therefore

$$\begin{aligned}
&\mathrm{Ent}_{S_n}\big(M(S_n, h^{0\cdot x+2}_{\partial S_n}),\omega\big)\\
&\qquad\ge -\frac{1}{|S_n|}\log\sum_{h\in M(S_n,h^{0\cdot x+2}_{\partial S_n})}\exp\big(H_{S_{n+2}}(\tilde{h},\omega)\big) - \frac{cC_\omega}{n}\\
&\qquad\ge \mathrm{Ent}_{S_{n+2}}\big(M(S_{n+2}, h^0_{\partial S_{n+2}}),\omega\big) - \frac{cC_\omega}{n}\,.
\end{aligned} \tag{2.138}$$

Now we combine (2.135) and (2.138) and send $n \to \infty$, which yields

$$\mathrm{ent}(0,\omega)\circ\tau_2 \ge \mathrm{ent}(0,\omega)\,. \tag{2.139}$$

By a similar argument with τ_2 replaced by τ_{-2}, we conclude that $\mathrm{ent}(0,\tau_2\omega) = \mathrm{ent}(0,\omega)$, i.e. $\mathrm{ent}(0,\omega)$ is invariant under τ_2. Since the distribution $\mathbb{P}$ of ω is ergodic with respect to τ_2 (cf. Assumption 2), this implies that $\mathrm{ent}(0,\omega) = \mathbb{E}[\mathrm{ent}(0,\omega)] = \mathrm{ent}_{\mathtt{an}}(0)$ almost surely. □

2.4.4 Convexity and continuity

The last results that we need about the annealed local surface tension $\mathrm{ent}_{\mathtt{an}}(s)$ are that is is convex and continuous as a function of the slope s.

Convexity allows us to apply standard analytic techniques to conclude that the macroscopic entropy functional $\mathrm{Ent}_{R,\mathsf{an}}(\cdot)$ is lower semi-continuous (see, for example, [CKP01, Section 2]). By semi-continuity, there exists a (perhaps non-unique) minimizer of the entropy functional, so the minimum in the variational principle (Theorem 4) is achieved.

Lemma 11. *The function $s \mapsto \mathrm{ent}_{\mathsf{an}}(s)$ is convex for $s \in (-1,1)^m$.*

Remark 10. The proof follows a standard argument based on buckled height functions; see e.g. [KMT20a, She05] for the uniform case. The energetic effect of the random potential contributes only on the boundary scale, and so is negligible in the limit. The proof could be considered an exercise for the reader; we work out the details below.

Proof of Lemma 11. We shall prove that for any choice of fixed coordinates

$$(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots s_m) \in [-1,1]^{m-1}, \tag{2.140}$$

the single-variate functions $s_i \mapsto \mathrm{ent}_{\mathsf{an}}((s_1, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots, s_m))$ are convex. It follows from elementary analysis that $s \mapsto \mathrm{ent}_{\mathsf{an}}(s)$ is a convex function on the m-dimensional domain $[-1,1]^m$. To simplify notation, we state the proof in the case $m = 2$. The proof generalizes to higher dimensions.

So, choose $u_0, u_1, u_2, v \in [-1,1]$ such that such that

$$u_1 = \frac{1}{2}u_0 + \frac{1}{2}u_2. \tag{2.141}$$

Our goal is to prove that

$$\mathrm{ent}_{\mathsf{an}}((u_1, v)) \leq \frac{1}{2}\,\mathrm{ent}_{\mathsf{an}}((u_0, v)) + \frac{1}{2}\,\mathrm{ent}_{\mathsf{an}}((u_2, v)). \tag{2.142}$$

We proceed as follows, in four steps.

- First, consider a discrete hypercube S_{2n+1}, which we recall is the hypercube $\{-(2n+1), \ldots, (2n+1)\}^2$ of side length $2(2n+1)+1$ centered at the origin. We subdivide it into $2m = 4$ smaller boxes. We choose height functions with slope (u_0, v) or (u_2, v) on the smaller boxes, and we construct a bijection which maps from a choice of height functions on the four smaller boxes to a height function on the larger box.

- Second, we use the bijection to derive an inequality between the microscopic entropy on the four smaller boxes and an entropy-like quantity on the larger box.

- Third, we relate this entropy-like quantity to the annealed surface tension $\mathrm{ent}_{\mathtt{an}}((u_1, v))$.

- Fourth, we relate the entropy on the smaller boxes to the right-hand side of (2.142), which concludes our proof.

So, let us make precise how we decompose S_{2n+1}. We write

$$S_{2n+1} = S_n^1 \cup S_n^2 \cup S_n^3 \cup S_n^4 \cup S' \tag{2.143}$$

where

$$S_n^1 := \tau_{(-n,+n)} S_n, \tag{2.144}$$

$$S_n^2 := \tau_{(-n,-n)} S_n, \tag{2.145}$$

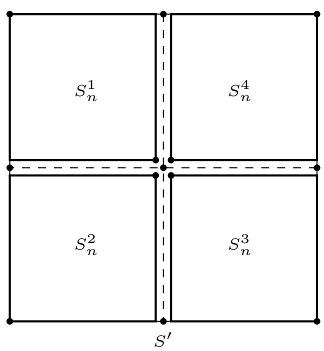
$$S_n^3 := \tau_{(+n,-n)} S_n, \tag{2.146}$$

$$S_n^4 := \tau_{(+n,+n)} S_n, \text{ and} \tag{2.147}$$

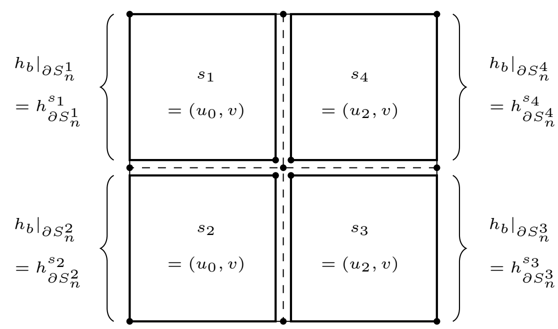
$$S' := \{(x_1, x_2) \in S_{2n+1} \,|\, x_1 = 0 \text{ or } x_2 = 0\}. \tag{2.148}$$

This decomposition is illustrated in Figure 2.7a.

As an aside, it would be simpler if we could decompose S_{2n+1} into just the four boxes S_n^k without needing the extra set S'. But, both S_{2n+1} and S_n^k are centered boxes, with an odd number of points along their edges ($4n+3$ and $2n+1$ points, respectively), so such a decomposition is arithmetically impossible. Centered boxes are a requirement of the ergodic



(a) The large box S_{2n+1} is divided into five subsets. The smaller boxes S_n^k, for $k = 1, 2, 3, 4$, are translated copies of the box S_n centered at the origin. The set S', indicated by dashed lines, is the intersection of S_{2n+1} and the x- and y-axes.



(b) To each of the four smaller boxes S_n^k, we associate a slope s_k. The two boxes on the left have $s_k = (u_0, v)$ and the two on the right have $s_k = (u_2, v)$. If h_b is a *buckled height function* (defined after (2.151)), then h_b satisfies the indicated boundary conditions on the four smaller boxes.

Figure 2.7: Decomposition of S_{2n+1} into subsets, as used in the proof of Lemma 11.

theorem that we used to prove Lemma 9 (the existence of the quenched local surface tension) and Lemma 10 (the equivalence of the quenched and annealed local surface tension). One could state the results without requiring odd-sized boxes centered exactly at the origin, but the statements become more complicated. We choose instead to keep the odd-sized boxes, and to keep the extra set S'. Because $|S'| = o(|S_n|)$, S' will be asymptotically negligible.

To continue with the current proof, we consider boundary height functions of slope s_k on the small boxes S_n^k, where

$$s_1 = s_2 = (u_0, v), \text{ and} \tag{2.149}$$

$$s_3 = s_4 = (u_2, v). \tag{2.150}$$

This assignment of slopes to the small boxes is illustrated in Figure 2.7b.

Fix a 4-tuple of height functions

$$(h_n^k)_{k=1,2,3,4} \in \prod_{k=1}^{4} M(S_n^k, h^{s_k}_{\partial S_n^k}). \tag{2.151}$$

We claim that there exists a height function $h_b : S_{2n+1} \to \mathbb{Z}$ such that for each of the four boxes S_n^k, $h_b|_{S_n^k} = h_n^k$. We call h_b a *buckled height function*, since if h_b stays close to the linear height functions $h^{s_k}_{S_n^k}$ over the entirety of the small boxes S_n^k, and if we view the graph of h_b in profile from along the y-axis, we see a buckled shape: slope (u_0, v) along the left half, which changes abruptly to slope (u_2, v) along the right half. Figure 2.7b illustrates the boundary conditions that are imposed on a buckled height function on the boundaries ∂S_n^k of the small boxes.

One can prove the existence of the height function h_b that extends the 4-tuple (h_n^k) to all of S_{2n+1} by using the Kirszbraun theorem. However it is also easy to construct a concrete extension using the canonical height functions. Briefly, on either side of a point on the x-axis, the slopes s_k are equal. For a point on the y-axis, the adjacent slopes differ only in the first coordinate. That is not a problem because for a point to be on the y-axis means that the value in its first coordinate is 0.

We write M_b for the set of all height functions $h_b : S_{2n+1} \to \mathbb{Z}$ that can be realized by the above extension process. Clearly, the set M_b is in bijection with the Cartesian product of the four sets $M(S_n^k, h_n^k)$. This bijection completes the first step of our proof.

In the second step of the proof, we derive the following approximation:

$$-\frac{1}{|S_{2n+1}|}\log Z_\omega(M_b) \leq \frac{1}{4}\sum_{k=1}^{4}\mathrm{Ent}_{S_n}\big(M(S_n^k, h^{s_k}_{\partial S_n^k}), \omega\big) + C_\omega\,\theta\big(\tfrac{1}{n}\big). \tag{2.152}$$

The key idea is that for a height function $h : S_{2n+1} \to \mathbb{Z}$, the Hamiltonian $H_{S_{2n+1}}(h,\omega)$ splits as

$$H_{S_{2n+1}}(h,\omega) = \sum_{k=1}^{4}\sum_{e\in E(S_n^k)}\omega_{h(e)} + \sum_{e\in E(S')}\omega_{h(e)} + \sum_{\tilde{e}\in\tilde{E}}\omega_{h(e)} \tag{2.153}$$

$$\geq \sum_{k=1}^{4}\sum_{e\in E(S_n^k)}\omega_{h(e)} - C_\omega\, O(n^{m-1}), \tag{2.154}$$

$$\tag{2.155}$$

where $\tilde{E}$ is the set of edges from $E(S_{2n+1})$ that cross between two distinct parts of the decomposition $S_{2n+1} = S_n^1 \cup S_n^2 \cup S_n^3 \cup S_n^4 \cup S'$. It follows that

$$-\frac{1}{|S_{2n+1}|}\log Z_\omega(M_b) = -\frac{1}{|S_{2n+1}|}\log \sum_{h_b\in M_b} \exp H_{S_{2n+1}}(h_b,\omega) \tag{2.156}$$

$$\leq -\frac{1}{|S_{2n+1}|}\log \left[\prod_{k=1}^{4}\left(\sum_{h_n\in M\left(S_n^k, h^{s_k}_{\partial S_n^k}\right)} \exp H_{S_n}(h_n,\omega)\right)\right. \tag{2.157}$$

$$\left.\exp\bigl(-C_\omega O(n^{m-1})\bigr)\right] \tag{2.158}$$

$$= -\frac{1}{4}\sum_{k=1}^{4}\frac{1}{|S_n|}\log \sum_{h_n\in M\left(S_n^k, h^{s_k}_{\partial S_n^k}\right)} \exp H_{S_n}(h_n,\omega) \tag{2.159}$$

$$+\frac{1}{|S_{2n+1}|} C_\omega\, O(n^{m-1}) \tag{2.160}$$

$$= \frac{1}{4}\sum_{k=1}^{4}\mathrm{Ent}_{S_n}\bigl(M(S_n^k, h^{s_k}_{S_n^k}),\omega\bigr) + C_\omega\,\theta\bigl(\tfrac{1}{n}\bigr). \tag{2.161}$$

This proves (2.152) and completes the second step of the proof.

Two steps remain. The third step is to relate the expression $-\frac{1}{|S_{2n+1}|}\log Z(M_b)$ (which we described as "entropy-like" earlier when describing the steps of this proof) to the annealed surface tension $\mathrm{ent}_{\mathtt{an}}((u_1,v))$. The fourth and final step is to verify that the microscopic entropy $\mathrm{Ent}_{S_n}(M(S_n^k, h^{s_k}_{\partial S_n^k}),\omega)$ converges to the annealed surface tension $\mathrm{ent}_{\mathtt{an}}(s_k)$ for $k = 1, 2, 3, 4$. This will suffice to prove the convexity inequality 2.142.

To relate $-\frac{1}{|S_{2n+1}|}\log Z(M_b)$ and $\mathrm{ent}_{\mathtt{an}}((u_1,v))$, we first pass to the microscopic entropy $\mathrm{Ent}_{S_{2n+1}}(M(S_{2n+1}, h^b_{\partial S_{2n+1}}),\omega)$. The boundary height function $h^b_{\partial S_{2n+1}}$ is given by $h^b_{\partial S_{2n+1}} = h_b|_{\partial S_{2n+1}}$ for any $h_b \in M_b$. All the buckled height function h_b have the same boundary data because of the boundary conditions on ∂S_n^k, plus the consistent (albeit arbitrary) choice of extension to S'. Obviously $M_b \subseteq M(S_{2n+1}, h^b_{\partial S_{2n+1}})$, and therefore by monotonicity,

$$-\frac{1}{|S_{2n+1}|} \log Z_\omega[M_b] \geq \mathrm{Ent}_{S_{2n+1}}\big(M(S_{2n+1}, h^b_{\partial S_{2n+1}}), \omega\big). \tag{2.162}$$

To estimate $\mathrm{Ent}_{S_{2n+1}}(M(S_{2n+1}, h^b_{\partial S_{2n+1}}), \omega)$, let us consider any boundary point $x = (x_1, x_2) \in \partial S_{2n+1}$. If $x_1 \leq 0$, then the boundary height function $h^b_{\partial S_{2n+1}}(x)$ is equal to $u_0 \cdot x$ up to a rounding error of at most 1, so

$$|h^b_{\partial S_{2n+1}}(x) - (u_1, v) \cdot x| \leq |u_0 - u_1||x_1| \leq \varepsilon n, \tag{2.163}$$

where $\varepsilon = |u_0 - u_1| = |u_0 - u_2|$. If instead $x_1 \geq 0$, then $h^b_{\partial S_{2n+1}}(x) = u.x$ up to rounding error, so still (2.163) holds. Therefore, the boundary data h_b is approximately linear, i.e. $h_b \in M(\partial S_{2n+1}, h^u_{\partial S_{2n+1}}, \varepsilon)$. By Theorem 7, there exists $A = A((u_1, v), \varepsilon) > 0$ such that

$$\mathrm{Ent}_{S_{2n+1}}\big(M(S_{2n+1}, h^b_{\partial S_{2n+1}}), \omega\big) \geq \mathrm{ent}_{A(2n+1)}((u_1, v), \omega). \tag{2.164}$$

We have proved the following inequality, which concludes the third step:

$$\mathrm{ent}_{A(2n+1)}((u_1, v), \omega) \leq -\frac{1}{|S_{2n+1}|} \log Z_\omega(M_b). \tag{2.165}$$

In the last step, we consider the quenched microscopic entropy on the four sub-boxes, i.e. $\mathrm{Ent}_{S^k_n}(M(S^k_n, h^k_n), \omega)$. Recall that each box S^k_n is a translation $\tau_{(\pm n, \pm n)} S_n$ of the box S_n centered at the origin. We transfer the translation over to the height function and environment. Let $\tau_1 : E(\mathbb{Z}) \to E(\mathbb{Z})$ denote the shift by $s_k \cdot (-n, +n)$, so that $\tau_1 \circ h^1_n = h^{s_1}_{S_n}$ and $\mathrm{Ent}_{S^1_n}(M(S^1_n, h^{s_1}_{\partial S^1_n}), \omega) = \mathrm{ent}_n(s_1, \tau_1 \omega)$. Likewise, define $\tau_2, \tau_3, \tau_4 : E(\mathbb{Z}) \to E(\mathbb{Z})$ so that for each $k = 1, 2, 3, 4$, it holds that

$$\mathrm{Ent}_{S^k_n}\big(M(S^k_n, h^{s_k}_{\partial S^k_n}), \omega\big) = \mathrm{ent}_n(s_k, \tau_1 \omega). \tag{2.166}$$

Combining this identity with (2.152) and (2.165), we deduce a quenched microscopic inequality

$$\mathrm{ent}_{A(2n+1)}((u_1, v), \omega) \leq \frac{1}{4} \sum_{k=1}^{4} \mathrm{ent}_n(s_k, \tau_k \omega) + C_\omega\, \theta\big(\tfrac{1}{n}\big). \tag{2.167}$$

Taking expectations, the annealed microscopic inequality is

$$\mathrm{ent}_{A(2n+1),\mathtt{an}}((u_1,v)) \le \frac{1}{4}\sum_{k=1}^{4} \mathrm{ent}_{n,\mathtt{an}}(s_k) + \mathbb{E}(C_\omega)\,\theta\big(\tfrac{1}{n}\big). \tag{2.168}$$

Inequality (2.142), which states that the annealed local surface tension is convex, follows immediately by sending $n \to \infty$. □

2.5 Proof of the profile theorem

Before proving the profile theorem, Theorem 3, in its full generality, it is useful to prove a special case of the theorem with the extra assumptions that the asymptotic height function is piecewise affine on a domain which is of a collection of simplices. In this special case it is not difficult to relate the microscopic entropy $\mathrm{Ent}_{R_n}(B(R_n, h_R, \delta), \omega)$ to the quenched microscopic surface tension $\mathrm{ent}_n(s, \omega)$, and then to derive the desired conclusion (2.52). The special case is stated in Lemma 12 below, after some necessary notation is introduced in Definitions 12 and 13.

Definition 12 (Simplices of scale ℓ; cf. [KMT20a, Definition 27] and [She05, Section 5.2.1]). Let $\mathrm{Sym}(m)$ denote the group of permutations on $\{1, \dots, m\}$, and for $w = (w_1, \dots, w_m) \in \mathbb{R}^m$, let $\lfloor w \rfloor$ denote the integer point $\lfloor w \rfloor := (\lfloor w_1 \rfloor, \dots, \lfloor w_m \rfloor)$. Let $v \in \mathbb{Z}^m$, let $\sigma \in \mathrm{Sym}(m)$, and let $\ell > 0$. Define $C(v, \sigma)$ to be the closure of the set

$$\left\{ w \in \mathbb{R}^m \,\middle|\, \lfloor w \rfloor = v \text{ and } w_{\sigma(1)} - \lfloor w_{\sigma(1)} \rfloor > \cdots > w_{\sigma(m)} - \lfloor w_{\sigma(m)} \rfloor \right\}, \tag{2.169}$$

and define the simplex of scale ℓ to the scaled set

$$\ell C(v, \sigma) := \{\ell w \,|\, w \in C(v, \sigma)\}. \tag{2.170}$$

See also Figure 2.8 and Figure 2.9.

Definition 13 (Piecewise affine asymptotic height functions). Let $\Delta_1, \dots, \Delta_k$ be simplices of scale ℓ and let $K = \Delta_1 \cup \cdots \cup \Delta_k$ be their union. We say that an asymptotic height function

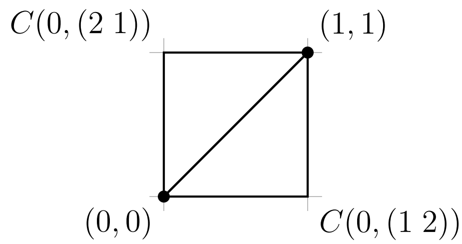


Figure 2.8: The two simplices in dimension 2 that tile the unit square. The simplex $C(0,(1\ 2))$ is the closure of the set of points $(x, y) \in [0, 1]^2$ such that $x > y$, and $C(0,(2\ 1))$ is the closure of the points with $y > x$. The other simplices $\{C(v, \sigma) \mid v \in \mathbb{Z}^m, \sigma \in S_2\}$ are translates of these two simplices.

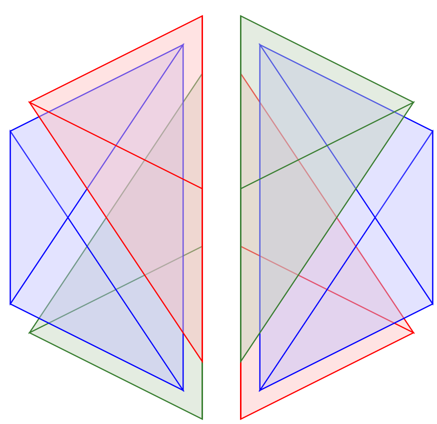


Figure 2.9: Decomposition of a unit cube into $\{C(0, \sigma) \mid \sigma \in S_3\}$. The simplices have been separated for a more clear figure.

$h_K \in M(K)$ is piecewise affine if each restriction $h_K|_{\Delta_i}$ is an affine function, i.e. if there exist $s_i \in [-1,1]^m$ and $b_i \in \mathbb{R}$ such that $h_K|_{\Delta_i}(x) = s_i \cdot x + b_i$ for all $x \in \Delta_i$. We write

$$\begin{aligned} M_{\mathtt{aff}}(K) &= \big\{ h_K \in M(K) \;\big|\; h_K \text{ is piecewise affine} \big\} \\ M_{\mathtt{aff}}(K, h_{\partial K}) &= M_{\mathtt{aff}}(K) \cap M(K, h_{\partial K}) \,. \end{aligned} \tag{2.171}$$

Lemma 12 (Profile theorem, simplicial case)**.** *Let $\Delta_1, \dots, \Delta_k$ be simplices of scale ℓ and let $K = \Delta_1 \cup \dots \cup \Delta_k$ be their union.*

For any $h_K \in M_{\mathtt{aff}}(K, h_{\partial K})$ and any $\eta > 0$, there exists $\varepsilon = \varepsilon_0(h_K, \eta)$ such that, for any $\varepsilon \in (0, \varepsilon_0]$ and any $p_{\mathtt{max}} \in (0,1)$, there exists $n_0 = n_0(h_K, \eta, \varepsilon, p_{\mathtt{max}})$ such that for all $n \geq n_0$,

$$\begin{aligned} \mathbb{P}\Big(\Big| &\mathrm{Ent}_{K_n}\big(B(K_n, h_K, \varepsilon\ell), \omega\big) - \mathrm{Ent}_{K,\mathtt{an}}(h_K) \Big| \\ &> \eta + C_\omega \theta_{h_K}(\varepsilon) + C_\omega \theta_{h_K,\varepsilon}\big(\tfrac{1}{n}\big) \Big) < p_{\mathtt{max}} \,. \end{aligned} \tag{2.172}$$

Proof. We will prove two bounds on $\mathrm{Ent}_{K_n}(B(K_n, h_{K_n}, \varepsilon\ell), \omega)$: an upper bound

$$\begin{aligned} \mathbb{P}\Big(&\mathrm{Ent}_{K_n}\big(B(K_n, h_K, \varepsilon\ell), \omega\big) > \mathrm{Ent}_{K,\mathtt{an}}(h_K) \\ &+ \eta + C_\omega \theta_{h_K}(\varepsilon) + C_\omega \theta_{h_K,\varepsilon}\big(\tfrac{1}{n}\big) \Big) \leq \theta_{h_K,\eta,\varepsilon}\big(\tfrac{1}{n}\big) \end{aligned} \tag{2.173}$$

and a lower bound

$$\begin{aligned} \mathbb{P}\Big(&\mathrm{Ent}_{K_n}\big(B(K_n, h_K, \varepsilon\ell), \omega\big) < \mathrm{Ent}_{K,\mathtt{an}}(h_K) \\ &- \eta - C_\omega \theta_{h_K}(\varepsilon) - C_\omega \theta_{h_K,\varepsilon}\big(\tfrac{1}{n}\big) \Big) \leq \theta_{h_K,\eta,\varepsilon}\big(\tfrac{1}{n}\big) \,. \end{aligned} \tag{2.174}$$

Assuming that both (2.173) and (2.174) hold, the conclusion (2.172) follows immediately by taking n_0 large enough based on the two $\theta_{h_k,\eta,\varepsilon}(\frac{1}{n})$ terms and applying the union bound on probabilities. So first let us verify the upper bound (2.173), and later we will verify the lower bound (2.174). For (2.173) we undercount the set of height functions $B(K_n, h_{K_n}, \varepsilon\ell)$. We choose a fine mesh of hypercubes $Q_{i,n}$ that approximate K_n and consider only those height functions that agree with the canonical boundary height functions $h^{s_i \cdot x + b_i}_{\partial Q_{i,n}}$ on $\partial Q_{i,n}$, where $s_i \in [-1,1]^m$ and $b_i \in \mathbb{R}$ are chosen such that $s_i \cdot x + b_i = h_K|_{Q_i}$. The mesh size is small enough that every such height function is in $B(K_n, h_{K_n}, \varepsilon\ell)$.

To be precise, let $q = \frac{1}{4}\varepsilon\ell$ be the mesh size. Let $Q_1, \dots, Q_k \subset \mathbb{R}^m$ enumerate the set of hypercubes in $\mathbb{R}^m$ that have side length q, have vertices in $q\mathbb{Z}^m$, and lie entirely in one of the simplices Δ_j. That last property ensures that there exist $s_i \in [-1,1]^m$ and $b_i \in \mathbb{R}$ such that

$$h_K(x) = s_i \cdot x + b_i \qquad \text{for all } x \in Q_i\,. \tag{2.175}$$

For $n \in \mathbb{N}$, let $Q_{i,n} := \{z \in \mathbb{Z}^m \mid \frac{1}{n}z \in Q_i\}$. Then as desired, for any choice of height functions

$$\big(h_{Q_{i,n}}\big)_{i=1}^{k} \in \prod_{i=1}^{k} M\big(Q_{i,n}, h^{s_i\cdot x+b_i}_{\partial Q_{i,n}}\big)\,, \tag{2.176}$$

there exists at least one extension $h_{K_n} \in M(K_n)$ to the whole of K_n (i.e. $h_{K_n}|_{Q_{i,n}} = h_{Q_{i,n}}$ for each $i = 1, \dots, k$), and any such extension lies in $B(K_n, h_K, \varepsilon\ell)$ by choice of q. Therefore,

$$\begin{aligned}\mathrm{Ent}_{K_n}\big(B(K_n, h_K, \varepsilon\ell), \omega\big) \le \frac{1}{k}\sum_{i=1}^{k} \mathrm{Ent}_{Q_{i,n}}\big(M(Q_{i,n}, h^{s_i\cdot x+b_i}_{\partial Q_{i,n}}), \omega\big)& \\ + C_\omega\theta_m(\varepsilon) + C_\omega\theta_{m,\varepsilon,\ell}\big(\tfrac{1}{n}\big)&\,,\end{aligned} \tag{2.177}$$

where the θ error terms come from the contribution of the set $K_n \setminus \bigcup_{i=1}^{k} Q_{i,n}$. For each $i = 1, \dots, k$, let us abuse notation and write "qn" to denote the side length of the hypercube $Q_{i,n}$. (In fact, the actual product $q \cdot n$ is generally not an integer, but the quantity we call qn satisfies $|qn - q \cdot n| < 1$.) Consider $Q_{i,n}$ as a translate $Q_{i,n} = v_i + S_{qn}$ for $v_i \in \mathbb{Z}^m$. Then the boundary values $h^{s_i\cdot x+b_i}_{\partial Q_{i,n}}$ are close to the translated values of $h^{s_i}_{\partial S_{qn}}$; in particular, for $z \in \partial S_{qn}$,

$$\Big| h^{s_i\cdot x+b_i}_{\partial Q_{i,n}}(v_i + z) - \Big(h^{s_i}_{\partial S_{qn}}(z) + \lfloor s_i \cdot v_i + nb_i \rfloor\Big)\Big| \le 4\,. \tag{2.178}$$

(A non-zero error occurs when s_i is irrational, or more generally when qns_i is not integral or has the wrong parity.) By Corollary 1 it follows that

$$\mathrm{Ent}_{Q_{i,n}}\big(M(Q_{i,n}, h^{s_i\cdot x+b_i}_{\partial Q_{i,n}}), \omega\big) = \mathrm{ent}_{qn}(s_i, \tau_{\lfloor s_i\cdot v_i+nb_i\rfloor}\omega) + C_\omega\theta_m\big(\tfrac{1}{n}\big)\,. \tag{2.179}$$

Combining (2.177) and (2.179) and abbreviating $\tau_{i,n} := \tau_{\lfloor s_i\cdot v_i+nb_i\rfloor}$ yields

$$\begin{aligned}&\mathrm{Ent}_{K_n}\big(B(K_n, h_K, \varepsilon\ell), \omega\big) \\ &\qquad \le \frac{1}{k}\sum_{i=1}^{k} \mathrm{ent}_{qn}(s_i, \tau_{i,n}\omega) + C_\omega\theta_m(\varepsilon) + C_\omega\theta_{m,\varepsilon,\ell}\big(\tfrac{1}{n}\big)\,.\end{aligned} \tag{2.180}$$

We note that the sequences $\{\mathrm{ent}_{qn}(s_i, \tau_{i,n}\omega)\}_{n\in\mathbb{N}}$ may not necessarily converge to $\mathrm{ent}_{\mathtt{an}}(s)$ as $n \to \infty$, despite the almost-sure convergence result of Lemma 10, due to the potential shifts $\tau_{i,n}$. However, since each $\mathrm{ent}_{qn}(s_i, \cdot) \to \mathrm{ent}_{\mathtt{an}}(s_i)$ in L^1, we can apply the Markov bound:

$$\begin{aligned} \mathbb{P}\bigg(\bigg| \frac{1}{k}\sum_{i=1}^{k} \mathrm{ent}_{qn}(s_i, \tau_{i,n}\omega) - \frac{1}{k}\sum_{i=1}^{k} \mathrm{ent}_{\mathtt{an}}(s_i)\bigg| > \eta\bigg) \\ \le \frac{1}{k}\sum_{i=1}^{k}\frac{1}{\eta}\left\|\mathrm{ent}_{qn}(s_i,\cdot) - \mathrm{ent}_{\mathtt{an}}(s_i)\right\|_{L^1} \\ = \theta_{h_K,\eta,\varepsilon,\ell}\left(\tfrac{1}{n}\right). \end{aligned} \tag{2.181}$$

The last step in verifying (2.173) is to compare $\mathrm{Ent}_{K,\mathtt{an}}(h_K)$ to a sum involving $\mathrm{ent}_{\mathtt{an}}(s_i)$. This is straightforward: because h_K is affine on each hypercube Q_i, the integrand $x \mapsto \mathrm{ent}_{\mathtt{an}}(\nabla h_K(x))$ in the macroscopic entropy is constant on each Q_i, so

$$\begin{aligned} \mathrm{Ent}_{K,\mathtt{an}}(h_K) &\overset{\text{def.}}{=} \frac{1}{|K|}\int_K \mathrm{ent}_{\mathtt{an}}(\nabla h_K(x))\,dx \\ &= \frac{1}{k}\sum_{i=1}^{k}\frac{1}{|Q_i|}\int_{Q_i}\mathrm{ent}_{\mathtt{an}}\big(\nabla h_K|_{Q_i}\big) + \theta_K(\varepsilon) \\ &= \frac{1}{k}\sum_{i=1}^{k}\mathrm{ent}_{\mathtt{an}}(s_i) + \theta_K(\varepsilon). \end{aligned} \tag{2.182}$$

The only error is from the contribution of the region $K \setminus \bigcup_{i=1}^k Q_i$. Combining inequalities (2.180), (2.181), and (2.182) proves the desired upper bound (2.173), i.e.

$$\begin{aligned} \mathbb{P}\Big(\mathrm{Ent}_{K_n}\big(B(K_n, h_K, \varepsilon\ell), \omega\big) > \mathrm{Ent}_{K,\mathtt{an}}(h_K) \\ + \eta + \theta_{h_K}(\varepsilon) + \theta_{h_K,\varepsilon}\left(\tfrac{1}{n}\right)\Big) \le \theta_{h_K,\eta,\varepsilon}\left(\tfrac{1}{n}\right). \end{aligned} \tag{2.183}$$

Now we turn to the lower bound (2.174). Similar to before, let $q = \varepsilon^{1/2}\ell$ and let $Q_1, \ldots, Q_k$ enumerate the hypercubes that have side length q, have vertices in $q\mathbb{Z}^m$, and lie entirely inside of one of the simplices Δ_j. Note that the side length q is different now compared to above when we were justifying the upper bound (2.173), and hence $Q_1, \ldots, Q_k$ denotes a different set of hypercubes.

To prove (2.174) we overcount height functions, using the same idea as in the article [KMT20a]. In summary, define a subset of "exceptional" points $E_n \subset K_n$ as follows: let

$$G_n = \bigcup_{i=1}^{k} \partial Q_{i,n}\,, \quad U_n = K_n \setminus \bigcup_{i=1}^{k} Q_{i,n}\,, \quad \text{and} \quad E_n = G_n \cup U_n\,. \tag{2.184}$$

Informally, G_n is the "grid" formed by the boundaries of the hypercubes and U_n is the "uncovered" region, i.e. the part of K_n that is not covered by the hypercubes. We group height functions $h_{K_n} \in B(K_n, h_K, \varepsilon\ell)$ based on their values on the set E_n. For each fixed assignment of heights $h_{K_n}|_{E_n} \in M(E_n)$, the entropy of the set of extensions to the hypercubes $\bigcup_1^k Q_n \approx K_n \setminus E_n$ is asymptotically equal to the macroscopic entropy $\mathrm{Ent}_{K,\mathtt{an}}(h_K)$. The set E_n is not too large, so even after counting all admissible assignments $h_{K_n}|_{E_n}$, the resulting asymptotics match (2.174).

To make the above argument rigorous, let $\mathrm{Adm}(E_n)$ denote the set of admissible height functions on E_n, i.e. those height functions $h_{E_n} \in M(E_n)$ that admit an extension to a height function in $B(K_n, h_K, \varepsilon\ell)$. There is an obvious injection from $B(K_n, h_K, \varepsilon\ell)$ into

$$\biguplus_{h_{E_n} \in \mathrm{Adm}(E_n)} \prod_{i=1}^{k} M\big(Q_{i,n}, h_{E_n}|_{\partial Q_{i,n}}\big)\,, \tag{2.185}$$

where "$\biguplus$" denotes the disjoint union (so for distinct height functions h_{E_n} and h_{E_n} in $\mathrm{Adm}(E_n)$, the product sets $\prod_1^k M\big(Q_{i,n}, h_{E_n}|_{\partial Q_{i,n}}\big)$ and $\prod_1^k M\big(Q_{i,n}, h_{E_n}|_{\partial Q_{i,n}}\big)$ are considered disjoint inside the set from (2.185)). It follows that

$$\begin{aligned} &Z_\omega\big(B(K_n, h_K, \varepsilon\ell), \omega\big) \\ &\quad \leq \sum_{h_{E_n} \in \mathrm{Adm}(E_n)} Z_\omega\left(\prod_{i=1}^{k} M\big(Q_{i,n}, h_{E_n}|_{Q_{i,n}}\big)\right) \\ &\quad \leq \big|\mathrm{Adm}(E_n)\big| \max_{h_{E_n} \in \mathrm{Adm}(E_n)} Z_\omega\left(\prod_{i=1}^{k} M\big(Q_{i,n}, h_{E_n}|_{Q_{i,n}}\big)\right). \end{aligned} \tag{2.186}$$

Therefore

$$\begin{aligned}
&\operatorname{Ent}_{K_n}\big(B(K_n, h_K, \varepsilon\ell), \omega\big) \\
&\qquad \geq \min_{h_{E_n} \in \operatorname{Adm}(E_n)} \sum_{i=1}^{k} \frac{|Q_{i,n}|}{|K_n|} \operatorname{Ent}_{Q_{i,n}}\big(M(Q_{i,n}, h_{E_n}|_{Q_{i,n}}), \omega\big) \\
&\qquad\qquad - \frac{\log|\operatorname{Adm}(E_n)|}{|K_n|} \,.
\end{aligned} \tag{2.187}$$

Clearly $\frac{|Q_{i,n}|}{|K_n|} = \frac{1}{k} + \theta_m(\varepsilon) + \theta_{m,\varepsilon,\ell}\big(\frac{1}{n}\big)$.

To control $|\operatorname{Adm}(E_n)|$, we argue as follows. First, $\frac{|G_n|}{|K_n|} = \theta_m(q) = \theta_m(\varepsilon)$ and $\frac{|U_n|}{|K_n|} = \theta_m(\varepsilon)$. Second, for an arbitrary base point $z_0 \in E_n$, there are at most $2\varepsilon\ell n + 1$ admissible values for $h_{E_n}(z_0)$ if $h_{E_n} \in \operatorname{Adm}(E_n)$, since h_{E_n} must extend to a height function in the ball $B(K_n, h_K, \varepsilon\ell)$. Third, the set E_n is connected, so for each of the admissible values of $h_{E_n}(z_0)$, there are at most $2^{|E_n|}$ height functions in $\operatorname{Adm}(E_n)$ taking that value at z_0. Putting these observations together, we conclude that $\frac{1}{|K_n|}\log|\operatorname{Adm}(E_n)| = \theta_m(\varepsilon) + \theta_{m,\varepsilon,\ell}\big(\frac{1}{n}\big)$.

Applying these asymptotic results in (2.187) yields

$$\begin{aligned}
&\operatorname{Ent}_{K_n}\big(B(K_n, h_K, \varepsilon\ell), \omega\big) \\
&\qquad \geq \min_{h_{E_n} \in \operatorname{Adm}(E_n)} \frac{1}{k} \sum_{i=1}^{k} \operatorname{Ent}_{Q_{i,n}}\big(M(Q_{i,n}, h_{E_n}|_{Q_{i,n}}), \omega\big) \\
&\qquad\qquad - \theta_m(\varepsilon) - \theta_{m,\varepsilon,\ell}\big(\tfrac{1}{n}\big) \,.
\end{aligned} \tag{2.188}$$

Whenever $h_{E_n} \in \operatorname{Adm}(E_n)$,

$$\max_{z \in E_n} \big| h_K(\tfrac{1}{n} z) - \tfrac{1}{n} h_{E_n}(z) \big| < \varepsilon\ell \,, \tag{2.189}$$

so for each $i = 1, \ldots, k$, by analogy to (2.178),

$$\max_{z \in \partial S_{qn}} \Big| \big(h_{E_n}(v_i + z) - \lfloor s_i \cdot v_i + qnb_i \rfloor\big) - h^{s_i}_{\partial S_{qn}}(z) \Big| \leq \varepsilon\ell n \,. \tag{2.190}$$

We apply Theorem 7 to the height function

$$\big(z \mapsto h_{E_n}(v_i + z) - \lfloor s_i \cdot v_i + qnb_i \rfloor\big) \in M(S_{qn}) \tag{2.191}$$

to conclude that

$$\begin{aligned}\mathrm{Ent}_{Q_{i,n}}&\big(M(Q_{i,n}, h_{E_n}|_{\partial Q_{i,n}}), \omega\big) \\ &\geq \mathrm{ent}_{Aqn}\big(s_i, \tau_{\lfloor s_i \cdot v_i + qnb_i \rfloor}\omega\big) - C_\omega \theta(\varepsilon) \,.\end{aligned} \tag{2.192}$$

The two almost-sure inequalities (2.188) and (2.192), the probability estimate (2.181), and the macroscopic bound (2.182) together imply the desired lower bound (2.174), which completes the proof of Lemma 12. □

The remainder of the proof of the profile theorem (Theorem 3) for general asymptotic height functions follows closely the proof in Section 6 of the article [KMT20a]. Below we state an approximation result (Theorem 9), which concludes that any asymptotic height function h_R admits a “good” approximation h_K satisfying the hypotheses of Lemma 12 above. Following that result are three robustness lemmas (Lemmas 14, Lemma 15, and Lemma 16). With these tools it is straightforward to reduce the general case of Theorem 3 to the special case of Lemma 12.

The approximation result, Theorem 9, is unchanged from that the article [KMT20a] where the uniform $\mathbb{Z}$-homomorphism model was studied. This should not be surprising because the random potential in the current model does not affect the class of limit objects that our model admits, i.e. domains satisfying Assumption 1 and asymptotic height functions. This theorem is similar to [CKP01, Lemma 2.2] or [Sch14, Theorem 1].

Theorem 9 (Simplicial Rademacher theorem)**.** *Let $R \subseteq \mathbb{R}^m$ be a region satisfying Assumption 1, and let $h_R \in M(R, h_{\partial R})$ be an asymptotic height function on R. For any $\varepsilon > 0$ and any $\ell > 0$ sufficiently small (depending on ε), we may choose a simplex domain $K = \Delta_1 \cup \cdots \cup \Delta_k \subseteq R$ of scale ℓ (see Definition 12) and a piecewise affine asymptotic height function $h_K : K \to \mathbb{R}$ (that is, an asymptotic height function such that each restriction $h_K|_{\Delta_i} : \Delta_i \to \mathbb{R}$ is affine) that satisfy the following properties:*

1. *$|R \setminus K| < \varepsilon$ and $d_H(K,R) < \varepsilon$), where we recall that for subsets of $\mathbb{R}^m$, $|\cdot|$ denotes the Lebesgue measure and $d_H(\cdot,\cdot)$ denotes Hausdorff metric;*

2. *$\max_{x\in K}|h_K(x) - h_R(x)| < \frac{1}{2}\varepsilon\ell$; and*

3. *on at least a $(1-\varepsilon)$ fraction of the points in K (by Lebesgue measure), the gradients $\nabla h_K(x)$ and $\nabla h_R(x)$ agree to within ε, i.e. $\frac{1}{|K|}\left|\left\{x \in K \;\middle|\; |\nabla h_K(x) - \nabla h_R(x)|_2 \geq \varepsilon\right\}\right| < \varepsilon$.*

Remark 11. We recall that the Rademacher theorem states that a Lipschitz function h_R is differentiable almost everywhere. However ∇h_R may be poorly behaved. The Rademacher theorem gives no control over ∇h_R, and the Lipschitz property only implies boundedness of the derivative, not regularity. The simplicial Rademacher theorem provides an approximation both to h_R and to its derivative. Moreover the approximating function h_K has a very simple derivative, despite the potential wildness of ∇h_R. The cost is that h_K only approximates h_R well on a (large) portion of the domain rather than almost everywhere, but for our purposes this is a good trade-off.

In fact, it is not necessary that the function h_R be Lipschitz. Almost everywhere differentiability is sufficient.

A proof of this lemma is given in the article [KMT20a]. We include it below for completeness. Before giving the proof however, we state and prove the following lemma about the standard simplices from Definition 12.

Lemma 13. *Let Δ be any of the simplices $C(v,\sigma)$ for $v \in \mathbb{Z}^m$ and $\sigma \in S_m$. The $m+1$ vertices of Δ can be labelled $x^{(0)}, \dots, x^{(m)}$ in such a way that, for each $i = 1, \dots, m$,*

$$x^{(i)} - x^{(i-1)} = e^{(\sigma(i))}, \tag{2.193}$$

where for $1 \leq j \leq m$, $e^{(j)}$ denotes the j-th standard basis vector (i.e., all entries of $e^{(j)}$ are 0, except the j-th entry, which is 1).

Remark 12. We encourage the reader to keep Figure 2.9 in mind (or better, in sight) while reading this proof.

Proof. For simplicity, we assume without loss of generality that $v = 0$. We use the permutation σ to define a path between vertices of the simplex $C(0, \sigma)$ starting at $(0, \dots, 0)$ and ending at $(1, \dots, 1)$. To construct the path, first observe that

$$C(0, \sigma) = \big\{x = (x_1, \dots, x_m) \in [0, 1]^m \,\big|\, x_{\sigma(i)} \geq x_{\sigma(j)} \text{ for all } i < j\big\}. \tag{2.194}$$

In other words, the $\sigma(1)$-th component of x must be greater than the $\sigma(2)$-th, which is greater than or equal to the $\sigma(3)$-th, and so on. The path travels from $(0, \dots, 0)$ along the $\sigma(1)$-th axis to $e_{\sigma(1)}$, then parallel to the $\sigma(2)$-th axis to $e_{\sigma(1)} + e_{\sigma(2)}$, and so on up to $\sum_{i=1}^m e_i = (1, \dots, 1)$. Numbering the vertices of the path from $x^{(0)}$ to $x^{(m)}$ proves the lemma. □

Now we turn to the robustness lemmas, which will be used when applying Theorem 9 to approximate h_R by another asymptotic height function. The three lemmas below are almost direct analogues of Lemmas 35, 36, and 37 from [KMT20a] respectively.

Lemma 14 (Robustness of macroscopic entropy under approximations)**.** *Let $\varepsilon > 0$, and let $\tilde{R} \subseteq R \subset \mathbb{R}^m$ be sets meeting the assumptions from Assumption 1 with $|R \setminus \tilde{R}| < \varepsilon$. Let $h_{\tilde{R}} \in M(\tilde{R})$ and $h_R \in M(R)$ be such that*

$$\Big|\Big\{x \in \tilde{R} \,\Big|\, \big|\nabla h_{\tilde{R}}(x) - \nabla h_R(x)\big|_2 \geq \varepsilon\Big\}\Big| < \varepsilon\,. \tag{2.195}$$

Then,

$$\mathrm{Ent}_{R,\mathsf{an}}(h_R) = \mathrm{Ent}_{\tilde{R},\mathsf{an}}(h_{\tilde{R}}) + \theta_m(\varepsilon)\,. \tag{2.196}$$

Proof. Recall from Definition 11 that

$$\mathrm{Ent}_{R,\mathsf{an}}(h_R) := \frac{1}{|R|} \int_R \mathrm{ent}_{\mathsf{an}}\big(\nabla h_R(x)\big)\, dx\,, \tag{2.197}$$

and likewise for $\mathrm{Ent}_{\tilde{R},\mathtt{an}}(h_{\tilde{R}})$. The conclusion follows from three observations: first that the domains of integration are bounded sets with small symmetric difference, second that the function $s \mapsto \mathrm{ent}_{\mathtt{an}}(s)$ is continuous, and third that the functions ∇h_R and $\nabla h_{\tilde{R}}$ almost agree (as per (2.195)) on most of the intersection of their domains (by measure). □

Lemma 15 (Robustness of microscopic entropy under change in profile)**.** *Let* $\varepsilon > 0$ *and* $n \in \mathbb{N}$*. Let* $R \subset \mathbb{R}^m$ *satisfy Assumption 1, and let* $R_n \subset \mathbb{Z}^m$ *satisfy* $\frac{1}{n}R_n \subset R$*. Let* $h_R, \tilde{h}_R \in M(R)$ *be two asymptotic height functions such that* $\sup_{x\in R}|h_R(x) - \tilde{h}_R(x)| \leq \varepsilon$*. Then,*

$$\mathrm{Ent}_{R_n}\big(B(R_n, h_R, 2\varepsilon), \omega\big) \leq \mathrm{Ent}_{R_n}\big(B(R_n, \tilde{h}_R, \varepsilon), \omega\big)\,. \tag{2.198}$$

Proof. For any fixed ω, the functional $\mathrm{Ent}_{R_n}(\cdot, \omega) : M(R_n) \to \mathbb{R}$ is monotonic, and it follows from Definition 2 that

$$B(R_n, \tilde{h}_R, \varepsilon) \subseteq B(R_n, h_R, 2\varepsilon)\,. \tag{2.199}$$

□

Lemma 16 (Robustness of microscopic entropy under domain approximations)**.** *Let* $c \in (0, 1]$*,* $\varepsilon \in (0, 1]$*, and* $n \in \mathbb{N}$*. Let* $\tilde{R} \subset R \subset \mathbb{R}^m$ *and* $\tilde{R}_n \subset R_n \subset \mathbb{Z}^m$ *satisfy these assumptions:*

$$\tfrac{1}{n}R_n \subset R\,, \qquad\qquad \tfrac{1}{n}\tilde{R}_n \subset \tilde{R}\,, \tag{2.200}$$

$$d_H(\tfrac{1}{n}R_n, R) = \theta_R(\varepsilon)\,, \qquad\qquad d_H(\tfrac{1}{n}\tilde{R}_n, \tilde{R}) = \theta_R(\varepsilon)\,, \tag{2.201}$$

$$\frac{|R_n|}{n^m|R|} = 1 + \theta_R(\varepsilon) + \theta_{R,\varepsilon}\big(\tfrac{1}{n}\big)\,, \qquad\qquad \frac{|\tilde{R}_n|}{n^m|\tilde{R}|} = 1 + \theta_R(\varepsilon) + \theta_{R,\varepsilon}\big(\tfrac{1}{n}\big)\,, \tag{2.202}$$

$$\frac{|R|}{|\tilde{R}|} = 1 + \theta_R(\varepsilon)\,. \tag{2.203}$$

Let $h_R \in M(R)$ *be an asymptotic height function with* $\mathrm{Lip}(h_R) \leq 1 - c\varepsilon$*. Then,*

$$\begin{aligned}\mathrm{Ent}_{\tilde{R}_n}\big(B(\tilde{R}_n, h_R, \varepsilon), \omega\big) &- C_\omega\theta_R(\varepsilon) - C_\omega\theta_{R,\varepsilon}\big(\tfrac{1}{n}\big)\\ &\leq \mathrm{Ent}_{R_n}\big(B(R_n, h_R, \varepsilon), \omega\big)\\ &\leq \mathrm{Ent}_{\tilde{R}_n}\big(B(\tilde{R}_n, h_R, \tfrac{c}{3}\varepsilon^2), \omega\big) + C_\omega\theta_R(\varepsilon) + C_\omega\theta_{R,\varepsilon}\big(\tfrac{1}{n}\big)\,.\end{aligned} \tag{2.204}$$

Proof. We prove the two inequalities in (2.204) separately. For the first inequality, observe that the map

$$\begin{aligned} B(R_n, h_R, \varepsilon) &\to B(\tilde{R}_n, h_R, \varepsilon) \\ h_R &\mapsto h_R|_{\tilde{R}} \end{aligned} \tag{2.205}$$

is not generally an injection, but it is at most $(2^{|R_n \setminus \tilde{R}_n|})$-to-1 (by the graph homomorphism property and connectedness of R_n). For any $h_{R_n} \in B(R_n, h_R, \varepsilon)$,

$$H_{R_n,\omega}(h_{R_n}) \leq H_{\tilde{R}_n,\omega}(h_{R_n}|_{\tilde{R}_n}) + C_\omega |R_n \setminus \tilde{R}_n| \,, \tag{2.206}$$

so

$$Z_\omega(B(R_n, h_R, \varepsilon)) \leq 2^{|R_n \setminus \tilde{R}_n|} Z_\omega(B(\tilde{R}_n, h_R, \varepsilon)) \exp\big(C_\omega |R_n \setminus \tilde{R}_n|\big) \tag{2.207}$$

and

$$\begin{aligned} &\mathrm{Ent}_{R_n}\big(B(R_n, h_R, \varepsilon), \omega\big) \\ &\quad \geq \frac{|\tilde{R}_n|}{|R_n|} \mathrm{Ent}_{\tilde{R}_n}\big(B(\tilde{R}_n, h_R, \varepsilon), \omega\big) \\ &\qquad - \log(2) \frac{|R_n \setminus \tilde{R}_n|}{|R_n|} - C_\omega |R_n \setminus \tilde{R}_n| \\ &\quad = \mathrm{Ent}_{\tilde{R}_n}\big(B(\tilde{R}_n, h_R, \varepsilon), \omega\big) - C_\omega \theta_R(\varepsilon) - C_\omega \theta_{R,\varepsilon}\big(\tfrac{1}{n}\big) \,. \end{aligned} \tag{2.208}$$

To prove the second inequality in (2.204), we first note that there exists an injection from $B(\tilde{R}_n, h_R, \frac{c}{3}\varepsilon^2)$ into $B(R_n, h_R, \varepsilon)$. A height function $h_{\tilde{R}_n} \in B(\tilde{R}_n, h_R, \frac{c}{3}\varepsilon^2)$ is extended to $h_{R_n} \in B(R_n, h_R, \varepsilon)$ in such a way that $\big|h_{R_n}(z) - n h_R\big(\frac{1}{n} z\big)\big| \leq 1$ when z is in R_n and sufficiently far away from $\tilde{R}_n$; the parameter value $\frac{c}{3}\varepsilon^2$ is chosen so that such an extension is admissible by the Kirszbraun theorem. For details, see the proof of [KMT20a, Lemma 37]. For this injection $h_{\tilde{R}_n} \mapsto h_{R_n}$,

$$H_{\tilde{R}_n,\omega}(h_{\tilde{R}_n}) \leq H_{R_n,\omega}(h_{R_n}) + C_\omega |R_n \setminus \tilde{R}_n| \,, \tag{2.209}$$

so

$$Z_\omega\big(B(\tilde{R}_n, h_R, \tfrac{c}{3}\varepsilon^2)\big) \leq Z_\omega\big(B(R_n, h_R, \varepsilon)\big) \exp\big(C_\omega |R_n \setminus \tilde{R}_n|\big) \tag{2.210}$$

and

$$\begin{aligned}
&\mathrm{Ent}_{\tilde{R}_n}\big(B(\tilde{R}_n, h_R, \tfrac{c}{3}\varepsilon), \omega\big) \\
&\qquad \geq \frac{|R_n|}{|\tilde{R}_n|}\,\mathrm{Ent}_{R_n}\big(B(R_n, h_R, \varepsilon), \omega\big) \\
&\qquad\qquad - \log(2)\,\frac{|R_n \setminus \tilde{R}_n|}{|R_n|} - C_\omega |R_n \setminus \tilde{R}_n| \\
&\qquad = \mathrm{Ent}_{R_n}\big(B(R_n, h_R, \varepsilon), \omega\big) - \theta_R(\varepsilon) - \theta_{R,\varepsilon}\big(\tfrac{1}{n}\big)\,.
\end{aligned} \tag{2.211}$$

□

To prove the profile theorem, we reduce to the special case of Lemma 12, where the domain is a collection of simplices and the asymptotic height function is piecewise affine. Before that, in order to apply Lemma 16, we reduce to the case where h_R has Lipschitz constant strictly less than 1. Both reductions are simple applications of the robustness results above.

Proof of the profile theorem (Theorem 3). For the reader's convenience we recall the conclusion of the theorem that we are about to prove, namely:

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}\Big(\Big|&\mathrm{Ent}_{R_n}\big(B(R_n, h_R, \delta), \omega\big) - \mathrm{Ent}_{\mathtt{an}}(R, h_R)\Big| \\
&\geq \eta + C_\omega \theta_{h_R}(\delta) + C_\omega \theta_{h_R,\delta}\big(\tfrac{1}{n}\big)\Big) = 0\,.
\end{aligned} \tag{2.212}$$

For the first step of the proof, we reduce from the case of an arbitrary asymptotic height function $h_R \in M(R, h_{\partial R})$, i.e. a continuous function $h_R : R \to \mathbb{R}$ with Lipschitz constant at most 1 (with respect to the ℓ^1 norm on R), to an asymptotic height function with Lipschitz constant strictly less than 1. Indeed, let $c := (2\,\mathrm{diam}_1 R)^{-1} \wedge 1$, where $\mathrm{diam}_1 R$ denotes the diameter of R under the ℓ^1 norm. By translation invariance of the random potential ω, we assume that there exists $x_0 \in R$ with $h_R(x_0) = 0$. Define

$$\tilde{h}_R := (1 - c\delta) h_R\,. \tag{2.213}$$

We make the following observations. First,

$$\mathrm{Lip}(\tilde{h}_R) \;=\; (1 - c\delta)\,\mathrm{Lip}(h_R) \;\leq\; 1 - c\delta\,. \tag{2.214}$$

Second, for any $x \in R$,

$$|h_R(x) - \tilde{h}_R(x)| \;\le\; c\delta|h_R(x)| \;\le\; c\delta|x - x_0|_1 \;\le\; \tfrac{\delta}{2}\,. \tag{2.215}$$

Third, for any $x \in R$,

$$|\nabla h_R(x) - \nabla \tilde{h}_R(x)| \le c\delta\,. \tag{2.216}$$

Lemma 14, together with (2.216) and the choice of constant $c = c(R)$, yields

$$\mathrm{Ent}_{R,\mathtt{an}}(h_R) = \mathrm{Ent}_{R,\mathtt{an}}(\tilde{h}_R) + \theta_R(\delta)\,. \tag{2.217}$$

Similarly, Lemma 15 and (2.215) imply that almost surely,

$$\begin{aligned} &\mathrm{Ent}_{R_n}\big(B(R_n, \tilde{h}_R, 2\delta), \omega\big) \\ &\qquad \le \mathrm{Ent}_{R_n}\big(B(R_n, h_R, \delta), \omega\big) \\ &\qquad \le \mathrm{Ent}_{R_n}\big(B(R_n, \tilde{h}_R, \tfrac{1}{2}\delta), \omega\big)\,. \end{aligned} \tag{2.218}$$

Assume for the sake of the proof that (2.212) holds for $\tilde{h}_R$. Then almost surely,

$$\begin{aligned} \mathrm{Ent}_{R_n}\big(B(R_n, h_R, \delta), \omega\big) &\le \mathrm{Ent}_{R_n}\big(B(R_n, \tilde{h}_R, \tfrac{\delta}{2}), \omega\big) \\ &\le \mathrm{Ent}_{R,\mathtt{an}}(\tilde{h}_R) \;+\; \eta \;+\; \theta_{\tilde{h}_R}\big(\tfrac{\delta}{2}\big) \;+\; \theta_{\tilde{h}_R,\delta/2}\big(\tfrac{1}{n}\big) \\ &= \mathrm{Ent}_{R,\mathtt{an}}(h_R) \;+\; \eta \;+\; \theta_{h_R}(\delta) \;+\; \theta_{h_R,\delta}\big(\tfrac{1}{n}\big), \end{aligned} \tag{2.219}$$

where in the last line, we combine the $\theta_R(\delta)$ term from (2.217) together with the $C_\omega \theta_{\tilde{h}_R}(\frac{\delta}{2})$ term above; this is admissible since $C_\omega \ge 1$ by definition (recall that $C_\omega := 1 \vee \sup_{e \in E(\mathbb{Z})} |\omega_e|$) and since the various factors of $\frac{1}{2}$ do not affect the asymptotics. The reverse inequality is similar, and so we have reduced to the problem of proving (2.212) with the added assumption that $\mathrm{Lip}(h_R) \le 1 - c\delta$ for $c = c(R) \in (0, 1)$.

We reduce further to the special case from Lemma 12, i.e. a piecewise affine asymptotic height function defined on a collection of simplices. First, we choose parameter values $\varepsilon = \varepsilon(\delta)$ and $\ell = \ell(\varepsilon, \delta)$ satisfying three criteria:

1. $\varepsilon \to 0$ as $\delta \to 0$,

2. $\delta = \varepsilon\ell$,

3. ℓ is sufficiently small so that the simplicial Rademacher theorem (Theorem 9) applies.

The choices of ε and ℓ may be realized as follows, from [KMT20a]: Choose a sequence $\varepsilon_k \searrow 0$ arbitrarily, e.g. $\varepsilon_k = \frac{1}{k}$. Let ℓ_k be the largest admissible ℓ value based on ε_k, but not larger than 1. For any given δ choose the smallest ε_k such that $\varepsilon_k\ell_k > \delta$; this ensures the first criterion. Set $\varepsilon = \varepsilon_k$ and $\ell = \frac{\delta}{\varepsilon_k} \leq \ell_k$; this ensures the last two criteria.

For the remainder of the argument, fix $\delta > 0$. Let ε and ℓ satisfy the above criteria, and let $K \subseteq R \subset \mathbb{R}^m$ be a simplicial domain and $h_K \in M(K)$ an asymptotic height function satisfying the conclusions of the simplicial Rademacher theorem (Theorem 9). Since $\nabla h_K \approx \nabla h_R$ (cf. conclusion 3 of Theorem 9) and since the macroscopic entropy is robust (Lemma 14),

$$\Big|\mathrm{Ent}_{R,\mathtt{an}}(h_R) - \mathrm{Ent}_{K,\mathtt{an}}(h_K)\Big| \leq \theta_R(\varepsilon) = \theta_R(\delta)\,, \tag{2.220}$$

where we use the fact that $\varepsilon \to 0$ as $\delta \to 0$ in order to replace ε by δ in the θ error term.

Similarly, by conclusions 1 and 2 of Theorem 9 and the microscopic entropy robustness,

$$\begin{aligned}
&\mathrm{Ent}_{R_n}\big(B(R_n, h_R, \varepsilon\ell), \omega\big)\\
&\overset{(Lemma\ 16)}{\leq} \mathrm{Ent}_{K_n}\big(B(K_n, h_R|_K, \tfrac{c}{3}(\varepsilon\ell)^2), \omega\big) + C_\omega\theta(\varepsilon) + C_\omega\theta_\varepsilon\big(\tfrac{1}{n}\big)\\
&\overset{(Lemma\ 15)}{\leq} \mathrm{Ent}_{K_n}\big(B(K_n, h_K, \tfrac{c}{6}(\varepsilon\ell)^2), \omega\big) + C_\omega\theta(\varepsilon) + C_\omega\theta_\varepsilon\big(\tfrac{1}{n}\big)
\end{aligned} \tag{2.221}$$

and

$$\begin{aligned}
&\mathrm{Ent}_{R_n}\big(B(R_n, h_R, \varepsilon\ell), \omega\big)\\
&\overset{(Lemma\ 16)}{\geq} \mathrm{Ent}_{K_n}\big(B(K_n, h_R|_K, \varepsilon\ell), \omega\big) - C_\omega\theta(\varepsilon) - C_\omega\theta_\varepsilon\big(\tfrac{1}{n}\big)\\
&\overset{(Lemma\ 15)}{\geq} \mathrm{Ent}_{K_n}\big(B(K_n, h_K, \tfrac{1}{2}\varepsilon\ell), \omega\big) - C_\omega\theta(\varepsilon) - C_\omega\theta_\varepsilon\big(\tfrac{1}{n}\big)\,.
\end{aligned} \tag{2.222}$$

Combining (2.220), (2.221), (2.222), and the special case of the profile theorem proved in Lemma 12 completes the proof. □

2.6 Proof of the variational principle

In this section we prove the variational principle (Theorem 4). The proof follows the steps of the corresponding proof for the uniform case in [KMT20a]. The main difference and the step that needs attention is that the deterministic convergence needs to be lifted to a convergence in probability. The two main inequalities in the proof follow from first comparing the set of height functions $M(R_n, h_{\partial R_n}, \delta)$ to the subset $B(R_n, h_R^*, \delta)$ for a well-chosen asymptotic height function h_R^*, and second from comparing to a superset $\bigcup_{i=1}^k B(R_n, h_R^{(i)}, \delta_i)$ for a collection of asymptotic height functions $h_R^{(1)}, \ldots, h_R^{(k)}$. Especially in the second part of the argument, some care is needed in regards to the asymptotic parameters. In particular:

- The choice (and number) of height functions $h_R^{(i)}$ depends on δ,
- the radii δ_i of the balls around these height functions depends on η,
- the probability that the profile theorem fails (i.e. the probability that $\mathrm{Ent}_{R,\mathtt{an}}(h_R^{(i)})$ and $\mathrm{Ent}_{R_n}(B(R_n, h_R^{(i)}, \delta_i), \omega)$ differ by a large amount due to the exact configuration ω of the random potential) depends not just on the error tolerance η but also on the number of height functions $h_R^{(i)}$.

Proof of Theorem 4. Let $\eta > 0$ and $p_{\mathtt{max}} > 0$. First we will establish that

$$\limsup_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}\Big(\mathrm{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}, \delta), \omega\big) > \inf_{h_R \in M(R, h_{\partial R})} \mathrm{Ent}_{R,\mathtt{an}}(h_R) + \eta \Big) \leq p_{\mathtt{max}} . \tag{2.223}$$

Choose $h^* \in M(R, h_{\partial R})$ such that

$$\mathrm{Ent}_{R,\mathtt{an}}(h_R^*) \leq \inf_{h_R \in M(R, h_{\partial R})} \mathrm{Ent}_{R,\mathtt{an}}(h_R) + \tfrac{\eta}{4} . \tag{2.224}$$

For any $\delta > 0$ and $n \in \mathbb{N}$, $B(R_n, h^*_R, \delta) \subseteq M(R_n, h_{\partial R_n}, \delta)$. Hence almost surely,

$$\mathrm{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}, \delta), \omega\big) \leq \mathrm{Ent}_{R_n}\big(B(R_n, h^*_R, \delta), \omega\big) . \tag{2.225}$$

By the profile theorem (applied to h^*_R),

$$\begin{aligned}\mathbb{P}\Big(\big|\mathrm{Ent}_{R_n}\,\big(B(R_n, h^*_R, \delta), \omega\big) - \mathrm{Ent}_{R,\mathtt{an}}(h^*_R)\big| \\ > \tfrac{\eta}{4} + C_\omega \theta_{h^*_R}(\delta) + C_\omega \theta_{h^*_R,\delta}\big(\tfrac{1}{n}\big)\Big) \underset{n\to\infty}{\to} 0 .\end{aligned} \tag{2.226}$$

Let us spend a part of the available probability $p_{\mathtt{max}}$ to establish a bound on C_ω. Specifically, since $C_\omega \in L^1$, Markov's inequality implies that

$$\mathbb{P}\big(C_\omega > \tfrac{2\|C_\omega\|_1}{p_{\mathtt{max}}}\big) \leq \tfrac{1}{2} p_{\mathtt{max}} . \tag{2.227}$$

Therefore as long as δ is small enough so that the $\theta_{h^*_R}(\delta)$ term is less than $\frac{\eta}{4} \cdot \frac{p_{\mathtt{max}}}{2\|C_\omega\|_1}$, and as long as n is large enough that the $\theta_{h^*_R,\delta}(\frac{1}{n})$ term is less than $\frac{\eta}{4} \cdot \frac{p_{\mathtt{max}}}{2\|C_\omega\|_1}$ and the probability in (2.226) is less than $\frac{1}{2} p_{\mathtt{max}}$, we have

$$\mathbb{P}\Big(\mathrm{Ent}_{R_n}\big(B(R_n, h^*_R, \delta), \omega\big) > \mathrm{Ent}_{R,\mathtt{an}}(h^*_R) + \tfrac{3\eta}{4}\Big) < p_{\mathtt{max}} . \tag{2.228}$$

The first desired inequality (2.223) follows immediately from (2.225), (2.228), and (2.224).

Now we turn to the second half of the variational principle, namely:

$$\begin{aligned}\limsup_{\delta\to 0} \limsup_{n\to\infty} \mathbb{P}\Big(\mathrm{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}, \delta), \omega\big) \\ < \inf_{h_R \in M(R, h_{\partial R})} \mathrm{Ent}_{R,\mathtt{an}}(h_R) - \eta\Big) \leq p_{\mathtt{max}} .\end{aligned} \tag{2.229}$$

In order to establish (2.229), we overcount the set $M(R_n, h_{\partial R}, \delta)$ using compactness of the space of asymptotic height functions $M(R, h_{\partial R}, \delta)$ (with respect to the topology of uniform convergence). Indeed, choose asymptotic height functions $h^{(1)}_R, \ldots, h^{(k)}_R$ such that

$$M(R, h_{\partial R}, \delta) \subset \bigcup_{i=1}^{k} B(R, h^{(i)}_R, \delta_i) , \tag{2.230}$$

where the values $\delta_i > 0$ are such that the $\theta_{h_R^{(i)}}(\delta_i)$ terms from the profile theorem (Theorem 3) are each less than $\frac{\eta}{4} \cdot \frac{p_{\mathtt{max}}}{2\|C_\omega\|_1}$.

As in the first part of the proof, we restrict to the event

$$\Omega' := \left\{ C_\omega < \frac{2\|C_\omega\|_1}{p_{\mathtt{max}}} \right\}, \tag{2.231}$$

which has $\mathbb{P}(\Omega') \geq 1 - \frac{p_{\mathtt{max}}}{2}$. Furthermore, we assume implicitly that n is large enough that:

- each of the $\theta_{h_R^{(i)},\delta_i}(\frac{1}{n})$ terms from the profile theorem is less than $\frac{\eta}{4} \cdot \frac{p_{\mathtt{max}}}{2\|C_\omega\|_1}$, and
- the exceptional events

$$E_{i,n} := \Omega' \cap \left\{ \left| \operatorname{Ent}_{R_n}\big(B(R_n, h_R^{(i)}, \delta_i), \omega\big) - \operatorname{Ent}_{R,\mathtt{an}}(h_R^{(i)}) \right| > \tfrac{3\eta}{4} \right\} \tag{2.232}$$

  satisfy $\mathbb{P}(E_{i,n}) < \frac{p_{\mathtt{max}}}{2k}$ for $i = 1, \ldots, k$.

Then for sufficiently small δ and sufficiently large n, the "good" event

$$\Omega_{\delta,n} := \Omega' \cap E_{1,n}^c \cap \cdots \cap E_{k,n}^c \tag{2.233}$$

satisfies $\mathbb{P}(\Omega_{\delta,n}) \geq 1 - p_{\mathtt{max}}$ and, for $\omega \in \Omega_{\delta,n}$,

$$\left| \operatorname{Ent}_{R_n}\big(B(R_n, h_R^{(i)}, \delta_i), \omega\big) - \operatorname{Ent}_{R,\mathtt{an}}(h_R^{(i)}) \right| \leq \tfrac{3\eta}{4} \,. \tag{2.234}$$

Assume in the sequel that $\omega \in \Omega_{\delta,n}$. By the set inclusion (2.230),

$$\operatorname{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}, \delta), \omega\big) \geq -\frac{1}{|R_n|} \log\left(\sum_{i=1}^{k} Z_\omega\big(B(R_n, h_R^{(i)}, \delta_i)\big) \right). \tag{2.235}$$

To handle the sum inside the logarithm, we compare each summand $Z_\omega(B(R_n, h_R^{(i)}, \delta_i))$ against $\inf_{h_R} \operatorname{Ent}_{R,\mathtt{an}}(h_R)$. Indeed,

$$\begin{aligned} \operatorname{Ent}_{R_n}\big(B(R_n, h_R^{(i)}, \delta), \omega\big) &\overset{(2.234)}{\geq} \operatorname{Ent}_{R,\mathtt{an}}\big(h_R^{(i)}\big) - \frac{3\eta}{4} \\ &\geq \inf_{h_R \in M(R, h_{\partial R})} \operatorname{Ent}_{R,\mathtt{an}}(h_R) - \frac{3\eta}{4} \,, \end{aligned} \tag{2.236}$$

and so

$$Z_\omega\big(B(R_n, h_R^{(i)}, \delta_i)\big) \le \exp\Big[|R_n|\Big(-\inf_{h_R\in M(R,h_{\partial R})} \mathrm{Ent}_{R,\mathtt{an}}(h_R) + \tfrac{3\eta}{4}\Big)\Big] \tag{2.237}$$

and

$$\sum_{i=1}^{k} Z_\omega\big(B(R_n, h_R^{(i)}, \delta_i)\big) \le k\exp\Big[|R_n|\Big(-\inf_{h_R\in M(R,h_{\partial R})} \mathrm{Ent}_{R,\mathtt{an}}(h_R) + \tfrac{3\eta}{4}\Big)\Big]. \tag{2.238}$$

Returning to (2.235), this yields

$$\begin{aligned}&\mathrm{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}, \delta), \omega\big)\\ &\qquad \ge \inf_{h_R\in M(R,h_{\partial R})} \mathrm{Ent}_{R,\mathtt{an}}(h_R) - \frac{\log k}{|R_n|} - \frac{3\eta}{4}.\end{aligned} \tag{2.239}$$

As long as n is large enough (depending on k, which in turn depends on δ), we have $\frac{\log k}{|R_n|} < \frac{\eta}{4}$, and so

$$\mathrm{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}, \delta), \omega\big) \ge \inf_{h_R\in M(R,h_{\partial R})} \mathrm{Ent}_{R,\mathtt{an}}(h_R) - \eta\,, \tag{2.240}$$

for any $\omega \in \Omega_{\delta,n}$. This establishes (2.229) and thereby proves the variational principle (Theorem 4). □

2.7 Proof of the large deviations principle

In this section we prove Theorem 5, the large deviations principle. Much like the proof variational principle in the section above, this proof for the homogenized model follows the same strategy as the corresponding proof for the uniform model. Before beginning the actual proof, we recall the following definitions from the statement of the theorem in Section 2.3

for the reader's convenience. For $\delta > 0$, $n \in \mathbb{N}$, and $h_R \in M(R)$:

$$\mu_{\delta,n}(A,\omega) := \frac{Z_\omega\big(\{h_{R_n} \in M(R_n, h_{\partial R_n}, \delta) \,|\, \tilde{h}_{R_n} \in A\}\big)}{Z_\omega\big(M(R_n, h_{\partial R_n}, \delta)\big)},$$

$$I(h_R) := \begin{cases} \mathrm{Ent}_{R,\mathtt{an}}(h_R) - E & \text{if } h_R|_{\partial R} \in M(R, h_{\partial R}), \\ \infty & \text{otherwise,} \end{cases}, \quad \text{and} \tag{2.241}$$

$$E := \inf_{h_R \in M(R,h_{\partial R})} \mathrm{Ent}_{R,\mathtt{an}}(h_R).$$

where $\tilde{h}_{R_n}$ is the piecewise-affine interpolation of the function $\frac{1}{n}z \mapsto \frac{1}{n}h_{R_n}(z)$ on the simplex domain with vertices $\{\frac{1}{n}K\}_{n \in R_n}$. Now we begin the proof.

Proof of Theorem 5. First, we prove the LDP lower bound (2.57), which we repeat for convenience. Given $\eta > 0$ and Borel $A \subset M(R)$, we must prove:

$$\overline{\lim_{\delta \to 0}}\, \overline{\lim_{n \to \infty}}\, \mathbb{P}\left(\frac{1}{|R_n|} \log \mu_{\delta,n}(A,\omega) \;\le\; -\inf_{h_R \in A^\circ} I(h_R) - \eta\right) = 0. \tag{2.242}$$

Without loss of generality we may assume that A is open. We may assume also that $\inf_{h_R \in A} I(h_R) < \infty$, or else (2.242) is trivial. By using these assumptions and replacing the symbols $\mu_{\delta,n}$ and $I(h_R)$ by their definitions, (2.242) simplifies to

$$\overline{\lim_{\delta \to 0}}\, \overline{\lim_{n \to \infty}}\, \mathbb{P}\bigg(\inf_{h_R \in A} \mathrm{Ent}_{R,\mathtt{an}}(h_R) - E + \eta \;\le\; \mathrm{Ent}_{R_n}\Big(\big\{h_R \in M(R_n, h_{\partial R_n}, \delta) \,\big|\, \tilde{h}_{R_n} \in A\big\}, \omega\Big) - \mathrm{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}, \delta), \omega\big)\bigg) = 0. \tag{2.243}$$

By the variational principle (Theorem 4),

$$\overline{\lim_{\delta \to 0}}\, \overline{\lim_{n \to \infty}}\, \mathbb{P}\Big(\Big|\mathrm{Ent}_{R_n}\big(M(R_n, h_{\partial R_n}, \delta), \omega\big) - E\Big| \ge \tfrac{\eta}{2}\Big) = 0. \tag{2.244}$$

So, under the limit superior in probability, we may cancel the corresponding terms in (2.243). Hence it suffices to show that

$$\overline{\lim_{\delta \to 0}}\, \overline{\lim_{n \to \infty}}\, \mathbb{P}\Big(\inf_{h_R \in A} \mathrm{Ent}_{R,\mathtt{an}}(h_R) + \tfrac{\eta}{2} \le \mathrm{Ent}_{R_n}\Big(\big\{h_{R_n} \in M(R_n, h_{\partial R_n}, \delta) \,\big|\, \tilde{h}_{R_n} \in A\big\}, \omega\Big)\Big) = 0. \tag{2.245}$$

Note the analogy between (2.245) and inequality (2.223) from the proof of the variational principle. We will prove (2.245) in a similar manner to (2.223). Choose an asymptotic height functions $h_R^\eta \in A$ that satisfies

$$\mathrm{Ent}_{R,\mathtt{an}}\big(h_R^\eta\big) \le \inf_{h_R \in A} \mathrm{Ent}_{R,\mathtt{an}}(h_R) + \frac{\eta}{4}. \tag{2.246}$$

Since $A \subset M(R)$ is open (with respect to the uniform norm) and since $h_R^\eta \in A$, it follows that for all $\delta > 0$ less than some $\delta_0 = \delta_0(\eta)$

$$B\big(R_n, h_R^\eta, \delta\big) \subset \big\{h_{R_n} \in M(R_n, h_{\partial R_n}, \delta) \,\big|\, \tilde{h}_{R_n} \in A\big\}. \tag{2.247}$$

Using this inclusion together with the profile theorem for h_R^η, we deduce

$$\begin{aligned} &\overline{\lim_{\delta\to 0}}\, \overline{\lim_{n\to\infty}}\, \mathbb{P}\Big(\inf_{h_R \in A} \mathrm{Ent}_{R,\mathtt{an}}(h_R) + \tfrac{\eta}{2} \le \mathrm{Ent}_{R_n}\Big(\big\{h_{R_n} \in M(R_n, h_{\partial R_n}, \delta) \,\big|\, \tilde{h}_{R_n} \in A\big\}, \omega\Big)\Big) \\ &\qquad \le \overline{\lim_{\delta\to 0}}\, \overline{\lim_{n\to\infty}}\, \mathbb{P}\Big(\mathrm{Ent}_{R,\mathtt{an}}\big(h_R^\eta\big) + \tfrac{\eta}{4} \le \mathrm{Ent}_{R_n}\big(B(R_n, h_R^\eta, \delta), \omega\big)\Big) = 0. \end{aligned} \tag{2.248}$$

Thus we have proven (2.245).

Now, we turn to the LDP upper bound (2.58), which for convenience we reproduce here:

$$\overline{\lim_{\delta\to 0}}\, \overline{\lim_{n\to\infty}}\, \mathbb{P}\left(\frac{1}{|R_n|} \log \mu_{\delta,n}(A,\omega) \ge -\inf_{h_R \in \overline{A}} I(h_R) + \eta\right) = 0. \tag{2.249}$$

We observe that $(\mu_{\delta,n})_{\delta,n}$ is exponentially tight, i.e. that for every $b \in (0,\infty)$, there exists $K_b \subset M(R)$ such that

$$\overline{\lim_{\delta\to 0}}\, \overline{\lim_{n\to\infty}}\, \frac{1}{|R_n|} \log \mu_{\delta,n}(K_b^c, \omega) \le -b, \quad \mathbb{P}(d\omega)\text{-a.s.} \tag{2.250}$$

Indeed, we may take K_b to be the closure of $M(h_R, h_{\partial R}, 1)$, independent of b. For $\delta < \frac{1}{3}$ and n large enough that

$$\max_{z \in \partial R_n} \big|\tfrac{1}{n} h_{\partial R_n}(z) - h_{\partial R}\big(\tfrac{1}{n} z\big)\big| \le \tfrac{1}{3}, \tag{2.251}$$

any $h_{R_n} \in M(R_n, h_{\partial R}, \delta)$ satisfies $\tilde{h}_{R_n} \in M(h_R, h_{\partial R}, 1)$ by the triangle inequality, and so we have $\mu_{\delta,n}(K_b^c, \omega) = 0$. By the general theory of large deviations, exponential tightness implies that it is sufficient prove the upper bound (2.249) for compact sets $A \subset M(R)$.

If $\inf_{h_R \in A} I(h_R) = \infty$, then every height function in A differs from $h_{\partial R}$ at some point on the boundary. In fact by compactness, there exists δ_0 such that for every $h_R \in A$, $\sup_{x \in \partial R} |h_{\partial R}(x) - h_R(x)| \geq \delta_0$. Clearly, as in the proof of exponential tightness above, this implies that $\{h_{R_n} \in M(R_n, h_{\partial R_n}, \delta) \,|\, \tilde{h}_{R_n} \in A\}$ is empty once δ is small enough and n large enough. For all such δ, n we have $\mu_{\delta,n}(A) = 0$ and (2.249) follows.

It remains to prove the upper bound (2.249) when $\inf_{h_R \in A} I(h_R) < \infty$ and A is compact. Just like for the lower bound before, we reduce to proving:

$$\overline{\lim_{\delta \to 0}}\, \overline{\lim_{n \to \infty}}\, \mathbb{P}\Big(\inf_{h_R \in A} \mathrm{Ent}_R(h_R) - \tfrac{\eta}{2} \geq \mathrm{Ent}_{R_n}\Big(\big\{h_{R_n} \in M(R_n, h_{\partial R_n}, \delta),\, \big|\, \tilde{h}_{R_n} \in A\big\}, \omega\Big)\Big) = 0. \tag{2.252}$$

We will follow the proof of (2.229) from the proof of the variational principle above. Similar to (2.230), use compactness to choose $h_R^{(1)}, h_R^{(2)}, \ldots, h_R^{(k)} \in A$ such that

$$A \subset \bigcup_{i=1}^{k} B\big(R, h_R^{(i)}, \delta_i\big), \tag{2.253}$$

where $\eta_1, \ldots, \eta_k$ are chosen so that for each i, the error term from the profile theorem for $h_R^{(i)}$ is smaller than $\frac{\eta}{4}$. Exactly as in the proof of Theorem 4 (see in particular (2.235)),

$$\mathrm{Ent}_{R_n}\Big(\big\{h_{R_n} \in M(R_n, h_{\partial R_n}, \delta) \,\big|\, \tilde{h}_{R_n} \in A\big\}, \omega\Big) \geq -\frac{1}{|R_n|} \log\Big(\sum_{i=1}^{k} Z_\omega\big(B\big(R_n, h_R^{(i)}, \eta_i\big)\big)\Big). \tag{2.254}$$

From this we deduce the analogue of (2.239), namely:

$$\begin{aligned} &\mathrm{Ent}_{R_n}\Big(\big\{h_{R_n} \in M(R_n, h_{\partial R_n}, \delta) \,\big|\, \tilde{h}_{R_n} \in A\big\}, \omega\Big) \\ &\quad \geq \inf_{h_R \in A} \mathrm{Ent}_{R,\mathtt{an}}(h_R) - \frac{\log k}{|R_n|} - \frac{3\eta}{4}. \end{aligned} \tag{2.255}$$

As $n \to \infty$ the right-hand side becomes greater than $\inf_{h_R \in A} \mathrm{Ent}_{R,\mathtt{an}}(h_R) - \eta$. Then, relabelling from η to $\frac{\eta}{2}$, completes the proof of (2.252) and of the large deviations principle. □

2.8 Preliminaries for proof of the concentration inequality

Now, let us describe our strategy for proving the homogenized concentration inequality. For the simpler case of random surface models without random perturbations, there are two approaches in the literature to proving the concentration inequality directly: an approach that leverages monotonicity and stochastic dominance (e.g. [CEP96, LT20]) and a dynamic approach using a natural Markov chain on height functions (e.g. [MT20b]). The second approach does not easily carry over to the case of random perturbations. It relies on monotonicity on the quenched level (i.e. with fixed random potential ω), but monotonicity only occurs in the annealed setting (i.e. after averaging over ω). This complicates the dynamic structure of the second approach. Therefore we use the first approach. See the statement and proof of Lemma 17 for details.

In this section and in Section 2.9, we write the random measure μ on sets of height functions, perturbed by the random potential ω, as follows: given a non-empty set $M \subset M(R)$ of height functions,

$$\mu_M(\cdot, \omega) := Z_\omega(\cdot)/Z_\omega(M) \tag{2.256}$$

denotes the measure conditioned on M. This notation is more useful here than the notation $\mu_\omega(\cdot)$ from (2.42) since we have more reason to consider several different sets of height functions in the context of this and the next section.

We note that a hypothesis sometimes imposed on intermediate results is that a certain set of heigh functions must not be empty, which is often equivalent to asserting that a chosen boundary height function must admit extensions. Also frequently throughout the proof, we frequently consider cases where either a subdomain admits height functions extending to the full domain or does not. This seems inevitable; the existence of a height function extension depends on the boundary values. The best characterization of suitable boundary data seems to be the Kirszbraun theorem, quoted as Theorem 6 above. We copy the statement of the theorem here for convenience, and note that a proof is given following the earlier statement.

Theorem. *Let R_n be a connected region of $\mathbb{Z}^m$, let R'_n be a subset of R_n, and let $\bar{h} : R'_n \to \mathbb{Z}$ be a graph homomorphism that preserves parity. There exists a graph homomorphism $h : R_n \to \mathbb{Z}$ such that $h = \bar{h}$ on R'_n if and only if for all $x, y \in R'_n$,*

$$d_{\mathbb{Z}}(\bar{h}(x), \bar{h}(y)) \leq d_{R_n}(x, y), \tag{2.257}$$

where $d_{\mathbb{Z}}$ and d_{R_n} denote respectively the graph distance on $\mathbb{Z}$ and on $R_n \subset \mathbb{Z}^m$.

Remark 13. One might hope to derive sufficient conditions for extensibility that are easier to verify, such as the following: suppose that R' is a line segment, i.e. $R' = \{v, v+e, v+2e, \ldots, v+(\ell-1)e\}$ for some $\ell \in \mathbb{N}$ and $v, e \in \mathbb{Z}^d$ with $|e|_1 = 1$. Then R' is an isometric subgraph of R, meaning that for any points $x, y \in R'$, $d_{R'}(x, y) = d_R(x, y)$. Indeed, clearly $d_{R'}(x, y) \geq d_R(x, y)$, and for any i, j, by observation $d_{R'}(v + ie, v + je) = |i - j| = d_{\mathbb{Z}^d}(v + ie, v + je) \leq d_R(x, y)$. Since R' is an isometric subgraph of every $R \supseteq R'$, it follows that for any $h_{R'} \in M(R')$ and any $x, y \in R'$, we have $|h_{R'}(x) - h_{R'}(y)| \leq d_{R'}(x, y) = d_R(x, y)$. Thus by the Kirszbraun theorem $h_{R'}$ admits at least one extension to R. This idea can be pushed a little bit farther. For example, the exact same argument works whenever R' is a geodesic, i.e. a shortest path between two points in R; a similar argument via the isometric subgraph property applies when ever R' is a box, i.e. when $R' = \{(z_1, \ldots, z_i) \in \mathbb{Z}^d \colon a_i \leq z_i \leq b_i\}$ for some $a_1 \leq b_1, \ldots, a_d \leq b_d$. However, this approach ultimately is not fruitful for our current purposes. The motivating example we use is the case where R is a box and R' is its boundary, and one can check easily that this subgraph R' is not isometric. For example, in two dimensions, if $R = \{0, 1, 2\} \times \{0, 1, 2\}$, then the opposite midpoints $(0, 1)$ and $(2, 1)$ have $d_{R'}((0, 1), (2, 1)) = 4 > 2 = d_R((0, 1), (2, 1))$. Indeed, one can check that there is a unique height function on R' with $h_{R'}((0, 1)) = 1$ and $h_{R'}((2, 1)) = 5$, which violates the Kirszbraun hypothesis (2.257). As such, in the sequel we must account for the possibility that $M(R; h_{R'})$ may be empty for some or even for all $h_{R'} \in M(R')$.

Finally, in order to state and prove the concentration inequality, we must introduce the annealed measure on height functions. This combines the randomness in the random potential with the randomness used to select a height function.

Definition 14 (Annealed measure). Below we prove that, for any $R' \subset R$ and any $h_{R'} \in M(R')$, the function $(\omega, A) \mapsto \mu_{M(R;h_{R'})}(A, \omega)$ from $\Omega \times \mathcal{P}(M(R; h_{R'}))$ to $[0, 1]$ is a probability kernel. In other words, for all $\omega \in \Omega$, $A \mapsto \mu_{M(R;h_{R'})}(A, \omega)$ is a probability measure, and for all $A \subset M(R; h_{R'})$, $\omega \mapsto \mu_{M(R;h_{R'})}(A, \omega)$ is measurable. Thus, the formula

$$(\mu_{M(R;h_{R'})} \circ \mathbb{P})(A) := \mathbb{E}[\mu_{M(R;h_{R'})}(A, \omega)], \quad A \subset M(R; h_{R'})$$

defines a probability measure $\mu_{M(R;h_{R'})} \circ \mathbb{P}$ on $M(R; h_{R'})$. Moreover, for any function $f : M(R; R') \to \mathbb{R}$, if h_R is a random variable with law $\mu_{M(R;h_{R'})} \circ \mathbb{P}$ then

$$E[f(X)] = \mathbb{E}\big[E_{\mu_{M(R;h_{R'})}(\cdot,\omega)}(f)\big]. \tag{2.258}$$

Proof (of claims in Definition 14). By definition $\mu_{M(R;h_{R'})}(\cdot, \omega)$ is a probability measure for fixed ω. The second property follows from how $\mu_{M(R;h_{R'})}(h_R, \omega)$ is defined: the numerator is a sum of finitely many random potential values ω_e inside of the (continuous, hence measurable) exponential function, and the denominator is a finite sum over copies of the numerator, only using different height functions to select the random potential values ω_e. The equation (2.258) is a standard identity for regular conditional distributions and its proof is a straightforward exercise; see e.g. [Dur10, Exercise 5.1.14]. Note that since $M(R; R')$ is finite, each function $f : M(R; R') \to \mathbb{R}$ is measurable. □

2.9 Proof of the concentration inequality

Having completed the preliminaries above, we are prepared to state the concentration inequality then move on to proofs.

Theorem 10. *Let $R_n \subset \mathbb{Z}^m$ be a sequence of finite, connected subgraphs with* $\operatorname{diam}(R_n) := \max_{x,y \in R_n} |x - y|_1 \leq An$ *for some $A > 0$. Let $\varepsilon > 0$ and $h_{\partial R_n} \in M(\partial R_n)$ be given, and let $\mu_n = \mu_{M(R;h_{\partial R_n})}(\cdot, \omega)$ denote the (perturbed) distribution on $M(R; h_{\partial R_n})$. Then for any $c > 0$*

and any $n \in \mathbb{N}$,

$$\mu_n \circ \mathbb{P}\left(\max_{v \in R_n} \left| h_{R_n}(v) - E_{\mu_n}(h_{R_n}(v)) \right| \geq c\sqrt{n}\right) \leq 2\,|R_n|\, e^{-nc^2/A}. \tag{2.259}$$

Remark 14. Theorem 10 gives quantitative bounds for the probability that a height function h_R differs from expected value on the scale of $\sqrt{n}$. The probability bounds are exponential in a constant times n, which comes about because of the one-dimensional nature of the Azuma–Hoeffding inequality. In comparison, the large deviations principle stated earlier achieves a volume-order term in the exponential. The cost of using the large deviations principle is that control over the size of fluctuations is not quantitative. For example, ignoring the random potential ω and using the uniform measure on $M(R; h_{\partial R})$ instead, the large deviations principle implies that $\mu_n(\max_v |h_{R_n}(v) - E_{\mu_n}(h_{R_n}(v))| \geq \varepsilon) \leq \exp(-n^m I(\varepsilon))$. Here $I(\varepsilon) > 0$ is (related to) the rate function of the large deviations principle, and it does not admit an obvious closed form expression in terms of ε. It would be interesting if we could obtain a quantitative concentration result with volume-order term in the exponential of the probability bound.

We will build up to the proof of the concentration inequality via a few intermediate results. Lemma 17 below establishes the monotonicity property, which is the main ingredient of the proof. We derive from it Corollary 2, which is used in the proof of an auxiliary concentration inequality in Lemma 18. The difference between Lemma 18 and the main theorem is that the former addresses only a single point $v \in R$, whereas the latter concerns the maximum deviation from the mean over the entire domain. The statements and proofs of these results are based on the method presented in [CEP96]; we cite the analogous steps where appropriate below. Differences arise starting with Corollary 2 below, where the shift-invariant and ergodic properties of the law of ω must be used to account for the fact that height functions with different base heights "see" different random potential values ω_e. However, the essential steps of the proof still goes through, since even under the influence of ω the relevant measures are Gibbs measures, based upon a finite-range potential (indeed, a nearest-neighbor potential,

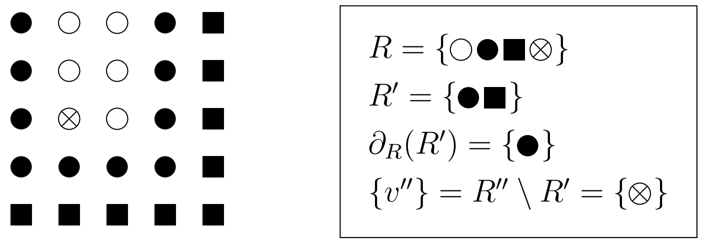


Figure 2.10: An example of the sets relevant to the proof of Lemma 17. On the left is a 5×5 subset of $\mathbb{Z}^2$, with points decorated according to the key on the right.

which enforces the Lipschitz property). It seems like the proof should extend to other finite-range models and perhaps beyond, but for brevity we will not explore that idea further here.

Lemma 17 (cf. [CEP96, Lemma 18])**.** *Let $R' \subset R$ and let $h_{R'}, \widetilde{h}_{R'} \in M(R')$ be such that $h_{R'} \leq \widetilde{h}_{R'}$ and that $M(R; h_{R'})$ and $M(R; \widetilde{h}_{R'})$ are not empty. Then for any realization ω, $\mu_{M(R;h_{R'})}(\cdot, \omega)$ is stochastically dominated by $\mu_{M(R;\widetilde{h}_{R'})}(\cdot, \omega)$. More precisely, there exists a measurable function $\pi : M(R; h_{R'}) \times M(R; \widetilde{h}_{R'}) \times \Omega \to [0, 1]$ such that:*

- *for almost every ω, $\pi(\cdot, \cdot, \omega)$ is a coupling, i.e.*

$$\sum_{\widetilde{h}_R} \pi(h_R, \widetilde{h}_R, \omega) = \mu_{M(R;h_{R'})}(h_R, \omega) \qquad \text{and} \qquad \sum_{h_R} \pi(h_R, \widetilde{h}_R, \omega) = \mu_{M(R;\widetilde{h}_{R'})}(\widetilde{h}_R, \omega),$$

  and

- $\pi\big(\{h_R \leq \widetilde{h}_R\}, \omega\big) = 1.$

Proof. Consider the "relative boundary" $\partial_R(R') := \{v' \in R' \colon \exists v \in R \setminus R', v \sim v'\}$, i.e. the points in R' that are directly adjacent to R. As we shall see below, these are the only essentially relevant points of R'. Indeed, we split the proof into two cases, depending on the restrictions $h_{R'}|_{\partial_R(R')}$ and $\widetilde{h}_{R'}|_{\partial_R(R')}$ on $\partial_R(R')$. The first case is the easier of the two. In the first case, the two height functions agree on $\partial_R(R')$. In the second case, there is a strict inequality $h_{R'}(v) < \widetilde{h}_{R'}(v)$ for at least one point $v \in \partial_R(R')$.

Case 1: Assume first that $h_{R'}(v) = h_{R'}(v)$ for all $v \in \partial_R(R')$. Since $\partial_R(R')$ may be a proper subset of R', this does not imply that $h_{R'} = \widetilde{h}_{R'}$. (Although in the case where $\partial_R(R')$, then this case does indeed reduce to the trivial assertion that $\mu_{M(R;h_{R'})}$ is stochastically dominated by itself.) However, it does hold that $h_{R'}$ and $\widetilde{h}_{R'}$ have the same extensions to $R \setminus R'$; to be very precise,

$$\big\{h_R|_{R\setminus R'} \colon h_R \in M(R; h_{R'})\big\} = \big\{\widetilde{h}_R|_{R\setminus R'} \colon \widetilde{h}_R \in M(R; \widetilde{h}_{R'})\big\}. \tag{2.260}$$

For clarity, let us repeat the above paragraph in the context of Figure 2.10. Case 1 of the proof concerns data $h_{R'}$ and $\widetilde{h}_{R'}$ that agree on the solid black circle region (i.e. $\partial_R(R')$), though they may differ on the solid black square points (i.e. $R' \setminus \partial_R(R')$). By definition of $\partial_R(R')$, the solid black circle points surround the white circle region (i.e. $R \setminus R'$), at least relative to the domain R. (Often we will assume that $\partial R \subset R'$, but that assumption isn't necessary here, and it does not hold in the figure.) Since $h_{R'}$ and $\widetilde{h}_{R'}$ agree on $\partial_R(R')$, they have the same extensions to the white circle region, in the sense of (2.260).

An easy calculation shows that for $h_R \in M(R; h_{R'})$,

$$\mu_{M(R;h_{R'})}(h_R, \omega) = \frac{\exp\big(H^+_{R\setminus R'}(h_R, \omega)\big)}{\sum_{f_R \in M(R;h_{R'})} \exp\big(H^+_{R\setminus R'}(f_R, \omega)\big)}, \tag{2.261}$$

where we recall that $H^+_{R\setminus R'}$ denotes the Hamiltonian on domain $R \setminus R'$, including the edges that cross between $R \setminus R'$ and $\partial_R(R')$. Therefore in particular that the right-hand expression (2.261) depends only on the values of the extension h_R restricted $(R \setminus R') \cup \partial_R(R')$. The same is true of extensions $\widetilde{h}_R \in M(R; \widetilde{h}_{R'})$. As such, the obvious bijection between the two sets in (2.260) is measure-preserving in both directions. The existence of a coupling π satisfying the claims of the lemma follows immediately.

Case 2: Assume instead that $h_{R'}(v) < \widetilde{h}_{R'}(v)$ for some $v \in R'$ adjacent to a vertex $v'' \in R \setminus R'$. In Figure 2.10, v'' is the white circle marked with an "X," and v might be either of the adjacent solid black circles. Let $R'' = R' \cup \{v''\}$. We proceed by induction on the cardinality of $R \setminus R'$. The induction hypothesis states that given any height functions

$h_{R''}, \widetilde{h}_{R''} \in M(R'')$ such that $h_{R''} \le \widetilde{h}_{R''}$ and such that both $M(R; h_{R''})$ and $M(R; \widetilde{h}_{R''})$ are nonempty, the measure $\mu_{M(R;h_{R''})}$ is stochastically dominated by $\mu_{M(R;\widetilde{h}_{R''})}$. Note that the base case of the induction occurs when $R \setminus R' = \{v''\}$ has cardinality 1, and so $R'' = R$; the lemma is trivial in this case.

So let us extend induction hypothesis from R'' to R'. From the hypotheses of the lemma, each of $h_{R'}$ and $\widetilde{h}_{R'}$ admits at least one extension to R. Therefore each admits at least one extension to R'' that in turn admits an extension to R. On the other hand since $R'' \setminus R' = \{v''\}$ is a set of cardinality 1, each of $h_{R'}$ and $\widetilde{h}_{R'}$ admits at most two extensions to R'. Formally, let $h^+_{R''}$ and $h^-_{R''}$ denote the two possible extensions of $h_{R'}$ to R'', where $h^{\pm}_{R''}(v'') = h_{R'}(v) \pm 1$. Below we will address the possibility that one or the other of these putative extensions does not exist. Likewise, let $\widetilde{h}^{\pm}_{R''} \in M(R''; \widetilde{h}_{R'})$ denote the two extensions of $\widetilde{h}_{R'}$ to R'', subject to the possibility that one or the other of the two extensions may not exist.

By conditioning on the height value at v'', we see that

$$\begin{aligned} \mu_{M(R;h_{R'})} &= p^+ \mu_{M(R;h^+_{R''})} + p^- \mu_{M(R;h^-_{R''})} \quad \text{and} \\ \mu_{M(R;\widetilde{h}_{R'})} &= \widetilde{p}^+ \mu_{M(R;\widetilde{h}^+_{R''})} + \widetilde{p}^- \mu_{M(R;\widetilde{h}^-_{R''})}, \end{aligned} \tag{2.262}$$

where

$$\begin{aligned} p^{\pm} &:= \mu_{M(R;h_{R'})}(\{h_R(v'') = h_{R'}(v) \pm 1\}) \in [0,1] \quad \text{and} \\ \widetilde{p}^{\pm} &:= \mu_{M(R;\widetilde{h}_{R'})}(\{\widetilde{h}_R(v'') = \widetilde{h}_{R'}(v) \pm 1\}) \in [0,1]. \end{aligned} \tag{2.263}$$

This addresses the issue noted above, about the possibility that one (but not both) of $h^{\pm}_{R''}$ may not exist; if so, the corresponding $p^{\pm}$ term is 0, and the other $p^{\mp}$ term is 1. By parity considerations, it must hold (assuming that the various extensions exist), that

$$h^-_{R''}(v'') < h^+_{R''}(v'') = h_{R'}(v) + 1 \le \widetilde{h}_{R'}(v) - 1 = \widetilde{h}^-_{R''}(v'') < \widetilde{h}^+_{R''}(v'').$$

By (up to) four applications of the induction hypothesis, we conclude that each of the measures $\mu_{M(R;h^{\pm}_{R''})}$ is stochastically dominated by each of the measures $\mu_{M(R;\widetilde{h}^{\pm}_{R''})}$. Since all the measures are probability measures and since all four of $p^{\pm}, \widetilde{p}^{\pm}$ are nonnegative, the identities (2.262) implies that $\mu_{M(R;h_{R'})}$ is stochastically dominated by $\mu_{M(R;\widetilde{h}_{R'})}$. □

We will make use of stochastic dominance via expectations, as captured in the following corollary.

Corollary 2 (cf. [CEP96, Corollary 19]). *Let $R' \subset R$, let $v \in R \backslash R'$, and let $h_{R'}, \widetilde{h}_{R'} \in M(R')$ with $h_{R'} \leq \widetilde{h}_{R'} + 2$. Let $h_R(v)$ and $\widetilde{h}_R(v)$ denote the $\mathbb{Z}$-valued random variables obtained by sampling h_R from $\mu_{M(R;h_{R'})} \circ \mathbb{P}$ and $\widetilde{h}_R$ from $\mu_{M(R;\widetilde{h}_{R'})} \circ \mathbb{P}$ and evaluating the respective height functions at v. Then*

$$E_{\mu_{M(R;h_{R'})} \circ \mathbb{P}}[h_R(v)] \leq E_{\mu_{M(R;\widetilde{h}_{R'})} \circ \mathbb{P}}[\widetilde{h}_R(v)] + 2.$$

Remark 15. Notice that unlike the stochastic monotonicity result of Lemma 17, which is almost sure in ω, the corollary above requires an expectation over the law $\mathbb{P}$ of ω. This is a substantial difference from [CEP96] caused by the random potential. Indeed, the requirement arises from the fact that $h_{R'}$ and $\widetilde{h}_{R'} + 2$ "see" a different part of the random potential ω. Since μ is a shift-invariant Gibbs measure, we can average out this height shift by annealing over the random potential.

Proof. By Lemma 17, for each fixed ω we have $\mu_{M(R;h_{R'})}(\cdot, \omega) \overset{\text{law}}{\leq} \mu_{M(R;\widetilde{h}_{R'}+2)}(\cdot, \omega)$, so

$$E_{\mu_{M(R;h_{R'})}(\cdot,\omega)}[h_R(v)] \leq E_{\mu_{M(R;\widetilde{h}_{R'}+2)}(\cdot,\omega)}[\widetilde{h}_R(v)], \quad \text{for a.e. } \omega.$$

We will transfer the height shift from the "$\widetilde{h}_{R'} + 2$" into the random potential ω and into the height function inside the expectation. Indeed, from the definition of the Hamiltonian, we have for $R \subset \mathbb{Z}^d$ and $f_R \in M(R)$ that

$$H_R^\circ(f_R + 2, \omega) = H_R^\circ(f_R, \tau_2\omega), \tag{2.264}$$

where $\tau_2 : \Omega \to \Omega$ is defined by $(\tau_2(\omega))_{x,x+1} = \omega_{x+2,x+3}$ for all $x \in \mathbb{Z}$. A straightforward calculation establishes that, for any $h_R \in M(R)$,

$$\mu_{M(R;\widetilde{h}_{R'}+2)}(h_R, \omega) = \mu_{M(R;\widetilde{h}_{R'})}(h_R - 2, \tau_2\omega); \tag{2.265}$$

indeed,

$$\begin{aligned}
\mu_{M(R;\widetilde{h}_{R'}+2)}(h_R,\omega) &= \frac{\exp\bigl(H_R^\circ(h_R,\omega)\bigr)}{\sum_{f_R\in M(R;h_{R'}+2)}\exp\bigl(H_R^\circ(f_R,\omega)\bigr)}\\
&= \frac{\exp\bigl(H_R^\circ(h_R,\omega)\bigr)}{\sum_{f_R\in M(R;h_{R'})}\exp\bigl(H_R^\circ(f_R+2,\omega)\bigr)}\\
&= \frac{\exp\bigl(H_R^\circ(h_R,\omega)\bigr)}{\sum_{f_R\in M(R;h_{R'})}\exp\bigl(H_R^\circ(f_R,\tau_2\omega)\bigr)}\\
&= \frac{\exp\bigl(H_R^\circ(h_R-2,\tau_2\omega)\bigr)}{\sum_{f_R\in M(R;h_{R'})}\exp\bigl(H_R^\circ(f_R,\tau_2\omega)\bigr)}\\
&= \mu_{M(R;\widetilde{h}_{R'})}(h_R-2,\tau_2\omega).
\end{aligned}$$

By change of variables,

$$E_{\mu_{M(R;h_{R'})}(\cdot,\omega)}[h_R(v)] \le E_{\mu_{M(R;\widetilde{h}_{R'})}(\cdot,\tau_2\omega)}[\widetilde{h}_R(v)+2], \quad \text{for a.e. } \omega.$$

Take expectations with respect to $\mathbb{P}$. Under the expectation the shift τ_2 vanishes, by ergodicity. The result follows by construction of the measures $\mu_{M(R;h_{R'})}\circ\mathbb{P}$ and $\mu_{M(R;\widetilde{h}_{R'})}\circ\mathbb{P}$; cf. equation (2.258). $\square$

Now we are prepared to prove a limited version of the concentration inequality, where we are concerned with only a single point $v\in R$. The key to the proof is the monotonicity of Corollary 2. We translate this into an inductive bound on martingale differences: each time we take a "step" starting at the boundary ∂R and "walking" towards v, the two possible extensions at that step differ by at most 2. Then we use the Azuma–Hoeffding inequality to establish the probability bound. From this point on the proof is standard, following closely to the methods used in [CEP96] and other works.

Lemma 18 (Auxiliary concentration inequality cf. [CEP96, Theorem 21])**.** *Let $h_{\partial R}\in M(\partial R)$ and let $v\in R$ be such that there is a path $x_0\in\partial R, x_1,\dots,x_{l-1}=v$ of length l with $x_i\sim x_{i-1}$ for $i=1,\dots,l-1$. Then for any $c>0$,*

$$\mu_{M(R;h_{\partial R})}\circ\mathbb{P}\Bigl(\Bigl\{h_R\in M(R;h_{\partial R})\colon \bigl|h_R(v)-E_{\mu_{M(R;h_{\partial R})}}[h_R(v)]\bigr|>lc\Bigr\}\Bigr)<2e^{-lc^2/2}.$$

Proof. For $k = 1, \ldots, l$, define σ-algebras $\mathcal{F}_k := \sigma(h_R \mapsto h_R(x_i), 0 \le i < k) \subset \mathcal{P}(M(R; h_{\partial R}))$ and define a martingale $M_k := E[h_R(v)|\mathcal{F}_k]$, where $E[\,\cdot\,]$ denotes the expectation with respect to the measure $\mu_{M(R;h_{R'})} \circ \mathbb{P}$. Note that $M_1 = E[h_R(v)]$ and that $M_l = h_R(v)$.

We claim that for each $k = 1, \ldots, l-1$, the martingale difference $|M_{k+1} - M_k|$ is less than or equal to 2 almost surely. To this end, fix k and condition on $h(x_i) = z_i \in \mathbb{Z}$ for $i = 0, \ldots, k-1$. To avoid events of probability zero, assume that $z_0, \ldots, z_{k-1}$ are such that there exists at least one extension in $M(h_R; h_{\partial R})$ with $h_R(x_i) = z_i$ for each i; by hypothesis $M(h_R; h_{\partial R})$ is nonempty, so at least one such assignment of heights z_i exists.

Having fixed these height values, there are at most two assignments of the height value $z_k := h_R(x_k)$ which admit further extensions in $M(R; h_{\partial R})$: namely, $z_k = z_{k-1} \pm 1$. Therefore the martingale $M_{k+1} = E[h_R(v)|\mathcal{F}_{k+1}]$ takes at most two distinct values conditioned on $\{h_R(x_i) = z_i,\, i = 0, \ldots, k-1\}$. Because the (at most) two possible values of $h_R(x_k)$ differ by at most 2, and because the height values at $x_0, \ldots, x_{k-1}$ have been fixed, Corollary 2 applied with $R' = \{x_0, \ldots, x_k\}$ implies that the (at most) two distinct values of M_{k+1} differ by at most 2. Since $M_k = E[M_{k+1}|\mathcal{F}_k]$ is the weighted average of these (at most) two values of M_{k+1}, it follows that $|M_{k+1} - M_k| \le 2$. The conclusion follows immediately from the Azuma–Hoeffding inequality. $\square$

Now we are prepared to prove the main result of this section, i.e. the concentration inequality.

Proof of Theorem 10. By the union bound and Lemma 18,

$$\mu_n \circ \mathbb{P}\Big(\max_{v \in R_n} \big|h_{R_n}(v) - E_{\mu_n}(h_{R_n}(v))\big| \ge c\sqrt{n}\Big) \le \sum_{v \in R_n} \mu_n \circ \mathbb{P}\Big(\big|h_{R_n}(v) - E_{\mu_n}(h_{R_n}(v))\big| \ge c\sqrt{n}\Big). \tag{2.266}$$

For each $v \in R_n$, apply Lemma 18 with the path length parameter l_v chosen as small as possible and with parameter c_v chosen such that $l_v c_v = c\sqrt{n}$. Recall that by hypothesis the diameter of R_n is at most An, so $l \leq An/2$. It follows that

$$l_v c_v^2 = \frac{(l_v c_v)^2}{l_v} = \frac{c^2 n}{l_v} \geq \frac{2c^2 n}{A}.$$

Using also the hypotheses that $|R_n| \leq Bn^m$, we have

$$\begin{aligned} &\mu_n \circ \mathbb{P}\Big(\max_{v\in R_n} \big|h_{R_n}(v) - E_\mu(h_{R_n}(v))\big| \geq c\sqrt{n}\Big) \\ &\quad \leq \sum_{v\in R_n} \mu_n \circ \mathbb{P}\Big(\big|h_{R_n}(v) - E_{\mu_n}\big(h_{R_n}(v)\big)\big| \geq l_v c_v\Big) \\ &\quad \leq \sum_{v\in R_n} 2e^{-l_v c_v^2/2} \\ &\quad \leq 2|R_n|e^{-c^2 n/A}. \end{aligned} \tag{2.267}$$

This is the desired bound from (2.259), and concludes the proof of Theorem 10, the concentration inequality. □

CHAPTER 3

Long-Range Percolation

The next problem addressed in this dissertation is the scaling of the graph distance (a.k.a. chemical distance) in the long-range percolation graph on $\mathbb{Z}^d$. This is a random graph with vertex set $\mathbb{Z}^d$ and a random set of undirected edges. In the version of long-range percolation studied here, the edge set $\mathscr{E} \subset E(\mathbb{Z}^d)$ always contains the nearest-neighbor edges, i.e. the (undirected) edges between pairs $x, y \in \mathbb{Z}^d$ with $|x-y|_1 = 1$, where $|\cdot|_1$ denotes the ℓ^1 norm. Furthermore the edge set contains longer edges selected at random. The presence or absence of these longer edges in $\mathscr{E}$ is governed by independent random variables, with the edge between vertices $x, y \in \mathbb{Z}^d$ present in $\mathscr{E}$ with probability $\mathfrak{p}(x-y) \in [0,1]$. In this chapter, our setting will be as follows: We start with collection of numbers $(\mathfrak{q}(x))_{x\in\mathbb{Z}^d} \subseteq [0,\infty)$ satisfying the symmetry condition $\mathfrak{q}(x) = \mathfrak{q}(-x)$ for all $x \in \mathbb{Z}^d$ and also the asymptotic condition $\mathfrak{q}(x) \sim |x|^{-s}$ for $s \in (d, 2d)$; in other words, $\lim_{|x|\to\infty} \mathfrak{q}(x)|x|^s = 1$. We also fix a parameter $\beta \in (0,\infty)$. Then, set

$$\mathfrak{p}_\beta(x) := \begin{cases} 1 - \exp(-\beta\mathfrak{q}(x-y)), & \text{if } |x-y|_1 > 1 \\ 1, & \text{if } |x-y|_1 = 1, \text{ and} \\ 0, & \text{if } |x-y|_1 = 0, \end{cases} \tag{3.1}$$

where $|\cdot|_1$ denotes the ℓ^1 distance. Observe that $\mathfrak{p}_\beta(x)/|x|^{-s} \sim \beta$.

From the references cited earlier in Chapter 1, particularly [BL19], it is known that

$$\lim_{|x|\to\infty} P\Big(c(\log|x|)^\Delta \le D(0,x) \le C(\log|x|)^\Delta\Big) = 1, \tag{3.2}$$

for some coefficients $C > c > 0$, where

$$\Delta := \frac{1}{\log_2(2d/s)}. \tag{3.3}$$

The proof there proceeded by introducing a continuum version of the long-range percolation model. Since the proof in Section 3.2 below is inspired by the proof for the continuum model there, let us discuss it in detail. To define the continuum model, start with a Poisson point process $\mathscr{I}'$ on $\mathbb{R}^d \times \mathbb{R}^d$ with intensity measure

$$\mu(\mathrm{d}x\,\mathrm{d}y) = 1_{|x|_2<|y|_2}\beta|x-y|^{-s}\,\mathrm{d}x\,\mathrm{d}y, \tag{3.4}$$

and let $\mathscr{I} = \mathscr{I}' \cup \{(y,x)\colon (x,y) \in \mathscr{I}\}$ denote its symmetrization. The pairs in $\mathscr{I}$ are interpreted as edges on $\mathbb{R}^d$. A path is defined by its start and end point, plus a finite (possibly empty) sequence of edges. The path length is the sum of the number of edges, plus sum of the distances between their vertices. Specifically, let x and y denote respectively the path start and end points. Let $(x_1, y_1), \ldots, (x_m, y_m) \in \mathscr{I} \subset \mathbb{R}^d \times \mathbb{R}^d$ denote the selected edges. The path then has length

$$m + \sum_{i=0}^{m} |y_i - x_{i+1}|, \tag{3.5}$$

where by convention $y_0 = x$ and $x_{m+1} = y$. The distance between points $x, y \in \mathbb{R}^d$ is defined as the infimum over all finite paths of this form with edges drawn from $\mathscr{I}$.

By changing the value of the parameter β, one may couple a copy of this continuum model and a copy of the original (discrete) long-range percolation model, in such a way that for every edge (x, y) in the continuum model there is a corresponding edge $(\lfloor x \rfloor, \lfloor y \rfloor)$ in the discrete model, where $\lfloor x \rfloor$ denotes the unique point in the lattice $\mathbb{Z}^d$ such that $x \in \lfloor x \rfloor + [0,1)^d$. Then any path in the continuum model induces a path in the discrete model, with edges in the continuum path mapped to by their corresponding discrete edges and with the linear spans in the continuum model, i.e. the spans between edges, which contribute the sum $\sum_{i=0}^{m} |y_i - x_{i+1}|$ in the formula (3.5), realized using a chain of nearest-neighbor edges in $\mathbb{Z}^d$. One then checks that the path thus created has distance bounded by a constant times the original path length, and hence one can use bounds on the continuum model (discussed below) to control distances in the discrete model. This approach was sufficient to derive the asymptotics stated in (3.2) for the original discrete model.

One might hope to derive tighter asymptotics for the discrete model by coupling the two processes more tightly. Indeed, if one considers only long edges in the two random processes, say of length $\geq K \gg 1$, then one can couple copies of the two processes with parameters β, β' chosen very close together. One hopes then to derive a limit result easily from this comparison. In fact, this coupling-based approach and the approach written out in Section 3.2 below both encounter the same issue, regarding continuity in the parameter β, as is discussed further near (3.11) below. The approach in Section 3.2 instead adapts the proof from [BL19]. A coupling argument is used twice in the process of adapting the proof, but that is not central to the method. In preparation for giving the adapted proof, let us review the statement and proof for the continuum version of the long-range percolation from [BL19].

The asymptotic result that was derived for the continuum model is: for $x \in \mathbb{R}^d \smallsetminus \{0\}$,

$$\frac{D(0, rx)}{\phi_\beta(r)(\log r)^\Delta} \xrightarrow[r\to\infty]{P} 1, \tag{3.6}$$

where the superscript "P" denotes convergence in probability. The function $\phi_\beta : (0,\infty) \to (0,\infty)$ is continuous in r, is bounded away from 0 and ∞, and satisfies the log-log-periodicity condition

$$\phi_\beta(r^\gamma) = \phi_\beta(r), \tag{3.7}$$

where $\gamma := s/(2d)$ is an exponent deeply related to the structure of optimal paths in the graph. Indeed, the consistent idea among the references [Bis04, Bis11a, BL19, BK21] that establish distance and diameter results in the regime $d < s < 2d$ is as follows. With high probability, the optimal path in the graph from 0 to a points x with $|x| \gg 1$ uses a single edge with length of order $|x|$. Such an edge can be found with high probability by searching for and edge $(X, Y) \in \mathscr{I}$ whose endpoints are respectively "close" to 0 and to x. Specifically, one expects to find $|X| \approx |x - Y| \approx |x|^\gamma$, and γ is a threshold for the appearance of such edges. Indeed, whenever one considers balls with radius $\gg |x|^\gamma$ there are usually many such edges, and if the radius is $\ll |x|^\gamma$ there are usually none (where "usually" means with probability tending to 1 as $|x| \to \infty$). Then one considers sub-paths from 0 to X and from Y to x, leading one to iterate through radii of the form $|x|, |x|^\gamma, |x|^{\gamma^2}$, and so on. Given that the structure of paths is tied to these double-exponential sequences in γ, the log-log-periodicity condition (3.7) should not be seen as mysterious.

To be more concrete about how the result in [BL19] is established, the key step is a subadditive estimate that builds off of the argument sketched in the paragraph above. A different criterion is used to choose the edge (X, Y): instead of minimizing the "graph distance" $D(0, x)$, one minimizes the function $f(X, Y) = |x|^{-\gamma}|X|^{2d} + |x|^{-\gamma}|Y|^{2d}$. Using the Poisson structure in the continuum process, the joint law of (X, Y') can be computed. After multiplying by $|X|^{-\gamma}$ to remove the scale dependence on $|x|$, one is left with a pair $Z = |x|^{-\gamma}X$ and $Z' = |x|^{-\gamma}(x - Y)$ which are independent. In other words,

$$D(0, x) \overset{\text{law}}{\leq} D(0, |x|^\gamma Z) + D(x, x + |x|^\gamma Z') + 1. \tag{3.8}$$

The actual result is [BL19, Proposition 2.7], which differs from (3.8) in several ways. First, and perhaps least importantly, the actual result includes an extra error term: with exponentially vanishing probability the edge (X, Y) selected above is not suitable (for technical reasons that we elide in this informal discussion), and one falls back to $D(0, x) \leq |x|$ in this low-probability event. More substantial is the fact that rather than dealing with the continuum distance $D(0, x)$, the actual statement in [BL19] has $\widetilde{D}(0, x)$, the restricted distance. In the definition of $\widetilde{D}(x, y)$ one minimizes path distance only over paths that stay within distance $2|y - x|$ of x. Hence the random variables $\widetilde{D}(x, y)$ and $\widetilde{D}(x', y')$ are independent if the pairs $\{x, y\}$ and $\{x', y'\}$ are separated by a large distance in the norm, as the pairs $\{0, |x|^\gamma Z\}$ and $\{x, x + |x|^\gamma |Z'|\}$ usually are in (3.8). Applying this independent in (3.8) one can replace $\widetilde{D}(x, x + |x|^\gamma Z')$ by an independent copy $\widetilde{D}'(x, x + |x|^\gamma Z')$, then use translation invariance to simplify to $\widetilde{D}(0,|x|^\gamma Z')$. The result is

$$\widetilde{D}(0,x) \overset{\text{law}}{\le} \widetilde{D}(0,|x|^\gamma Z) + \widetilde{D}'(0,0+|x|^\gamma Z') + 1 + \text{error}. \tag{3.9}$$

Now, using the law of the i.i.d. pair (Z,Z') that was computed above, one can show that the infinite product $W := (\prod_{k=1}^\infty |Z|^{\gamma^k})Z$ converges with $|W| \in (0,\infty)$ a.s. This variable W has the property that $|W|^\gamma Z \overset{\text{law}}{=} W$, so it is a fixed point of the subadditive iteration: plugging in $x \leftarrow rW$ in (3.9) and taking expectations yields

$$E\widetilde{D}(0,rW) \le 2E\widetilde{D}(0,r^\gamma |W|^\gamma Z) + \text{error} = 2E\widetilde{D}(0,r^\gamma W) + \text{error}. \tag{3.10}$$

Hence the limit $L(r) := \lim_{n\to\infty} 2^{-n} E\widetilde{D}(0,r^{\gamma^{-n}}W)$ exists. One then builds upon this first limit result, first passing from convergence in expectation to almost sure convergence. Then one replaces the limit along the double-exponential sequence $r^{\gamma^{-n}}$ by a limit with just $r\to\infty$; this requires weakening from almost sure convergence to convergence in probability. Then it remains to pass from the restricted distance $\widetilde{D}$ back to the full distance; this is accomplished by relaxing the restriction on how far paths can travel away from x; the limit of the hierarchy of relaxed restricted distances is the full distance, and [BL19] shows how to carry the distance asymptotics forward to it.

When we carry out the same strategy of proof sketched above for the original discrete long-range percolation process, a few points will necessarily differ from [BL19]. First, the proof above, when written out completely, relies on varying the parameter β at a certain point. In the continuum model, it is possible to exchange a multiplicative adjustment to β for a multiplicative scaling of the spatial coordinates; the exact result is [BL19, Lemma 2.2]: for $a \ge 1$,

$$D_\beta(0,ax) \overset{\text{law}}{\le} D_{a^{s-2d}\beta}(0,ax) \overset{\text{law}}{\le} aD_\beta(0,x). \tag{3.11}$$

This relies on the fact that the Poisson process from which the continuum model is realized can be scaled by the map $x \mapsto ax$. The result is a Poisson process with nearly the same intensity, except thinned by the factor of a^{s-2d} that appears above. Of course, one cannot generally apply the scaling map $x \mapsto ax$ when points are restricted to the integer lattice. Without this trick, it is more difficult to control the process when changing the parameter β. A more abstract approach is used below, relying on the monotonicity in β, which implies continuity for all but countably many values of β.

Interestingly, the scaling relation (3.11) is used for the opposite purpose as well: to absorb linear scaling in the spatial coordinate into the parameter β. Indeed, from the outermost expressions in (3.11) one obtains the useful relation $D(0, ax) \overset{\text{law}}{\leq} aD(0, x)$. The corresponding inequality in the discrete model is difficult to establish. Even when $|x| \gg 1$ so that ax is very close to $\lfloor ax \rfloor \in \mathbb{Z}^d$, one cannot easily control the distance $D(0, ax)$ using $D(0, x)$, since every point along the path shifts by some amount, and possibly by a large amount, when those points are far from the origin. In [BL19] the scaling inequality (3.11) was combined with a rotational inequality (which again cannot be easily translated to the discrete model, for essentially the same reason) to deduce [BL19, (4.22)]: for each $\varepsilon > 0$ there exists $\delta > 0$, such that for all $x, y \in \mathbb{R}^d$ and all $t > 0$,

$$|y - x| < \delta |x| \qquad \Longrightarrow \qquad P_\beta\big(D(0, x) \leq t\big) \leq P_{(1+\varepsilon)\beta}\big(D(0, y) \leq (1 + \varepsilon)t\big). \tag{3.12}$$

In other words, ignoring the ε factors, the distance to x is stochastically dominated by the distance to y for nearby points y, where the meaning of "nearby" is relative to $|x|$. Since the limits we take involve scaling to infinity, it is important that the error $|y-x|$ is allowed to scale with $|x|$. This result (3.12) is used when conditioning on $|W - x| < \delta$, to replace $D(0, rW)$ by $D(0, rx)$. The lack of a result corresponding to (3.12) for the discrete model is why we are limited to claiming the limit for most points $x \in B(0, r)$ in Theorem 11 below, whereas in [BL19, Theorem 1.2] the conclusion was for almost all x.

Outline of the rest of the chapter

Having discussed the strategy of proof, the rest of the chapter will proceed as follows:

- In Section 3.1, we formaly state the main result of this chapter, namely Theorem 11.
- In Section 3.2, we prove the main theorem.

3.1 Model and results

Our formal setting is as follows: Let $|\cdot|$ denote a norm on $\mathbb{Z}^d$ and let $s \in (d, 2d)$. Let $\mathfrak{q}$ denote a function $\mathfrak{q} : \mathbb{Z}^d \to [0, \infty)$, such that $\mathfrak{q}(x) = \mathfrak{q}(-x)$ for all $x \in \mathbb{Z}^d$, and such that

$$\mathfrak{q}(x) \sim \frac{1}{|x|^s}. \tag{3.13}$$

For $\beta \in (0, \infty)$ define

$$\mathfrak{p}_\beta(x) := \begin{cases} 1 - \exp(-\beta\mathfrak{q}(x-y)), & \text{if } |x-y|_1 > 1 \\ 1, & \text{if } |x-y|_1 = 1, \text{ and} \\ 0, & \text{if } |x-y|_1 = 0, \end{cases} \tag{3.14}$$

The long-range percolation with edge probability $\mathfrak{p}_\beta$ is the random graph with vertices $\mathbb{Z}^d$ and an undirected edge between x and y present with probability $\mathfrak{p}_\beta(x-y)$, independently of other edges. The chemical distance (i.e. graph distance) $D(x, y)$ between vertices $x, y \in \mathbb{Z}^d$ is then defined as the minimal number of edges in any path connecting x to y.

Having fixed notation and definitions, we are now prepared to state the main result of this chapter:

Theorem 11. *Let $d \geq 1$ and $s \in (d, 2d)$ and assume $\mathfrak{q}$ obeys (3.13). Let $\Delta = 1/\log_2(2d/s)$, as in (3.3). For each $\beta > 0$ there exists a continuous function $\phi_\beta \colon (1, \infty) \to (0, \infty)$ subject to the log-log-periodicity condition*

$$\forall r > 1\colon \quad \phi_\beta(r^\gamma) = \phi_\beta(r) \tag{3.15}$$

with $\gamma := \frac{s}{2d}$, and there is at most countable $\Sigma \subseteq (0, \infty)$ such that, for all $\beta \in (0, \infty) \smallsetminus \Sigma$,

$$\forall \varepsilon > 0\colon \quad \frac{1}{r^d}\#\left(\left\{x \in B(0, r)\colon \left|\frac{D(0, x)}{\phi_\beta(r)(\log r)^\Delta} - 1\right| > \varepsilon\right\}\right) \underset{r\to\infty}{\overset{P}{\longrightarrow}} 0. \tag{3.16}$$

The map $\beta \mapsto \phi_\beta(r)$ is non-increasing and left-continuous. It is continuous at all $\beta \notin \Sigma$.

3.2 Proof of asymptotic bounds on chemical distance

The proof of of Theorem 11 follows closely that of its continuum predecessor [BL19, Theorem 1.2]. Many steps of the proof can in fact be taken over nearly *verbatim*. The main novelty is the need for a coupling between the lattice and continuum edge processes and an argument by-passing discontinuity points of $\beta \mapsto \phi_\beta(r)$.

3.2.1 Subadditivity inequality

Given a sample of the percolation graph, let $\mathscr{E}$ denote for the set of all undirected edges included in the graph, including the nearest-neighbor edges. We then echo definition (2.1) of [BL19] and introduce $\widetilde{D}\colon \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{Z}$ via

$$\widetilde{D}(x,y) := \inf\left\{ n \geq 0\colon \begin{array}{l} \{(x_{k-1},x_k)\colon k=1,\dots,n\} \subseteq \mathscr{E},\ x_0 = x, \\ x_n = y,\ \forall k=1,\dots,n\colon\ |x_k - x| < 2|x-y|_1 \end{array} \right\}. \tag{3.17}$$

We will refer to $\widetilde{D}(x,y)$ as the *restricted distance* from x to y as it is non-negative, strictly positive for $x \neq y$ and arises by optimizing lengths of paths, although $\widetilde{D}$ is not a distance in proper sense as it is not symmetric in general. What matters in the sequel is

$$\forall x,y \in \mathbb{Z}^d\colon \quad D(x,y) \leq \widetilde{D}(x,y) \leq |x-y|_1 \tag{3.18}$$

and the fact that the law of $\widetilde{D}$ is translation invariant with

$$\forall x,y,x',y' \in \mathbb{Z}^d\colon\ |x-x'|_1 > 2|x-y|_1 + 2|x'-y'|_1 \ \Rightarrow\ \widetilde{D}(x,y) \perp\!\!\!\perp \widetilde{D}(x',y'). \tag{3.19}$$

Here and henceforth $|\cdot|_1$ denotes the ℓ^1-norm on $\mathbb{R}^d$.

For $x \in \mathbb{R}^d$, let $\lfloor x \rfloor$ denote the unique $z \in \mathbb{Z}^d$ such that $x - z \in [0,1)^d$. The independence property (3.19) enabled by the consideration of the restricted distance permits us to prove the following analogue of [BL19, Proposition 2.7] that drives the bulk of the subsequent derivations in this section.

Proposition 1 (Subadditivity inequality)**.** *Fix $\eta \in (0,1)$ and $\overline{\gamma} \in (\gamma, 1)$. Let Z, Z' be i.i.d. $\mathbb{R}^d$-valued random variables with common law given by*

$$P(Z \in B) = \sqrt{\eta\beta} \int_B e^{-\eta\beta c_0 |z|^{2d}} dz, \tag{3.20}$$

where

$$c_0 := \int 1_{\{|z|^{2d} + |\tilde{z}|^{2d} \le 1\}} \, dz \, d\tilde{z}. \tag{3.21}$$

Let $\widetilde{D}'$ be an independent copy of $\widetilde{D}$ with $\widetilde{D}$ and $\widetilde{D}'$ assumed independent of Z and Z'. For each $\gamma_1, \gamma_2 \in (0, \overline{\gamma})$ with $\gamma_1 + \gamma_2 = 2\gamma = s/d$, there are $c_1, c_2 \in (0, \infty)$ and, for each $x \in \mathbb{Z}^d$, there is an event $A(x) \in \sigma(Z, Z')$ such that

$$\widetilde{D}(0, x) \overset{law}{\le} \widetilde{D}\big(0, \lfloor |x|^{\gamma_1} Z \rfloor\big) + \widetilde{D}'\big(0, \lfloor |x|^{\gamma_2} Z' \rfloor\big) + 1 + |x|_1 1_{A(x)} \tag{3.22}$$

and

$$P\big(A(x)\big) \le c_1 e^{-c_2 |x|^{\vartheta}} \tag{3.23}$$

hold with $\vartheta := 2d[\overline{\gamma} - \max\{\gamma_1, \gamma_2\}]$.

Proof. Fix $\eta \in (0,1)$, $\overline{\gamma} \in (\gamma, 1)$ and $\gamma_1, \gamma_2 \in (0, \overline{\gamma})$ with $\gamma_1 + \gamma_2 = 2\gamma$. Let $x \in \mathbb{Z}^d$. Following the overall strategy of the proof in [BL19], consider Borel measures μ and μ' on $\mathbb{R}^d \times \mathbb{R}^d$ defined by

$$\mu(\mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y}) := \eta\beta 1_{\{|\tilde{x}| < |\tilde{y}|\}} 1_{\{|\tilde{x}| \vee |\tilde{y} - x| \le |x|^{\overline{\gamma}}\}} \frac{\mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y}}{|x|^s} \tag{3.24}$$

and

$$\mu'(\mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y}) := \eta\beta \frac{\mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y}}{|x|^s} - \mu'(\mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y}). \tag{3.25}$$

Next observe that, for $|x|$ larger than an η-dependent constant, for any $(\tilde{x}, \tilde{y}) \in \mathbb{R}^d \times \mathbb{R}^d$, the inequalities $|\tilde{x}|_2 \le |\tilde{y}|_2$, $|\tilde{x}| \le |x|^{\overline{\gamma}}$, and $|\tilde{y} - x| \le |x|^{\overline{\gamma}}$ imply

$$\left(\tfrac{1+\eta}{2}\right)^{1/s} |x| < |\tilde{x} - \tilde{y}| < \left(\tfrac{1+\eta}{2}\right)^{-1/s} |x| \tag{3.26}$$

and so, by the inequality on the right,

$$\mu\Big(\big(\lfloor\tilde{x}\rfloor+[0,1)^d\big)\times\big(\lfloor\tilde{y}\rfloor+[0,1)^d\big)\Big)\le\frac{\eta\beta}{|x|^s}\le\frac{\eta}{\big(\frac{1+\eta}{2}\big)}\,\frac{\beta}{|\tilde{x}-\tilde{y}|^s}. \tag{3.27}$$

By (3.14–3.13), the left-inequality in (3.26) and $\eta(\frac{1+\eta}{2})^{-1}<1$, this is less than $\mathfrak{p}_\beta(\lfloor\tilde{x}\rfloor-\lfloor\tilde{y}\rfloor)$ as soon as $|x|$ is sufficiently large. Hence for $|x|$ large we can couple a Poisson point process $\mathscr{I}$ with intensity measure μ' to the discrete edge set $\mathscr{E}$ so that

$$\forall(\tilde{x},\tilde{y})\in\mathscr{I}\colon\ \ (\lfloor\tilde{x}\rfloor,\lfloor\tilde{y}\rfloor)\in\mathscr{E} \tag{3.28}$$

holds pointwise and, by (3.19) and restriction built into the definition of $\widetilde{D}$, the families

$$\big\{\widetilde{D}(0,\lfloor\tilde{x}\rfloor)\colon|\tilde{x}|\le|x|^{\overline{\gamma}}\big\},\ \big\{\widetilde{D}(x,\lfloor\tilde{y}\rfloor)\colon|\tilde{y}-x|\le|x|^{\overline{\gamma}}\big\},\ \mathscr{I} \tag{3.29}$$

are independent.

Let $\mathscr{I}'$ be a Poisson point process with intensity measure μ'' independent of $\mathscr{I}$ and $\mathscr{E}$. Then $\mathscr{I}\cup\mathscr{I}'$ is a homogeneous Poisson process with intensity $\eta\beta|x|^{-s}\in(0,\infty)$ and, as is readily checked, there is almost surely a unique pair $(X,Y)\in\mathscr{I}\cup\mathscr{I}'$ that minimizes the function

$$f_x(\tilde{x},\tilde{y}):=\big(|x|^{-\gamma_1}|\tilde{x}|\big)^{2d}+\big(|x|^{-\gamma_2}|\tilde{y}-x|\big)^{2d}. \tag{3.30}$$

The joint law of X and Y can be computed explicitly as in [BL19], using the Poisson structure of the edge set $\mathscr{I}$:

$$P\big((X,Y)\in B\big)=\frac{\eta\beta}{|x|^s}\int_B\exp\Big\{-\frac{\eta\beta}{|x|^s}\int 1_{\{f_x(\tilde{x}',\tilde{y}')\le f_x(\tilde{x},\tilde{y})\}}\mathrm{d}\tilde{x}\mathrm{d}\tilde{y}'\Big\}\,\mathrm{d}\tilde{x}\mathrm{d}\tilde{y}. \tag{3.31}$$

The random variables

$$Z:=|x|^{-\gamma_1}X\quad\text{and}\quad Z':=|x|^{-\gamma_2}(Y-x) \tag{3.32}$$

then have the joint law

$$P\big((Z,Z')\in B\big)=\eta\beta\int_B\exp\Big\{-\eta\beta\int 1_{\{|\tilde{z}|^{2d}+|\tilde{z}'|^{2d}\le|z|^{2d}+|z'|^{2d}\}}\mathrm{d}\tilde{z}\mathrm{d}\tilde{z}'\Big\}\mathrm{d}z\mathrm{d}z' \tag{3.33}$$

Scaling $\tilde{z}$ and $\tilde{z}'$ by $(|z|^{2d}+|z'|^{2d})^{1/d}$, the integral in the exponent is shown to equal $c_0(|z|^{2d}+|z'|^{2d})$. (This is where using $2d$-powers in (3.30) is crucial.) Thus Z and Z' are indeed independent with above law (3.20).

Next define the event $A(x)$ as follows: When $|x|$ is large enough (with exact bounds including those mentioned above and those further given below), set

$$A(x) := \bigl\{|Z| > |x|^{\overline{\gamma}-\gamma_1}\bigr\} \cup \bigl\{|Z'| > |x|^{\overline{\gamma}-\gamma_2}\bigr\} \tag{3.34}$$

and let $A(x)$ be the entire probability space otherwise. On the event $A(x)^{\mathrm{c}}$ the edge (X,Y) lies in $\mathscr{I}$ since $|X| \vee |Y-x| \le |x|^{\overline{\gamma}}$ and so $(\lfloor X\rfloor, \lfloor Y\rfloor) \in \mathscr{E}$ by (3.28). Moreover, both X and Y are within distance $2|x|$ of the origin (as long as $|x|$ is large enough; this is part of the bounds on $|x|$). Recalling the notation $B(y,r) := \{z \in \mathbb{R}^d \colon |z-y| < r\}$, similar arithmetic as in [BL19, eq. (2.29) and (2.30)] establishes

$$B\bigl(0, 2|\lfloor X\rfloor|\bigr) \subseteq B\bigl(0, 2|x|\bigr) \quad \text{and} \quad B\bigl(x, 2|\lfloor Y\rfloor - x|\bigr) \subseteq B\bigl(0, 2|x|\bigr). \tag{3.35}$$

Picking a path achieving $\widetilde{D}(0,\lfloor X\rfloor)$, concatenating it with edge $(\lfloor X\rfloor, \lfloor Y\rfloor)$ and a path achieving $\widetilde{D}(x,\lfloor Y\rfloor)$ then produces a path in $B(0,2|x|)$ whose length dominates the restricted distance $\widetilde{D}(0,x)$.

Using (3.18) to bound $\widetilde{D}(0,x)$ by $|x|_1 1_{A(x)}$ when $A(x)$ occurs, this yields the pointwise inequality

$$\widetilde{D}(0,x) \le \widetilde{D}\bigl(0, \lfloor |x|^{\gamma_1} Z\rfloor\bigr) + \widetilde{D}\bigl(x, x + \lfloor |x|^{\gamma_2} Z'\rfloor\bigr) + 1 + |x|_1 1_{A(x)}. \tag{3.36}$$

In light of (3.29), the two instances of $\widetilde{D}$ on the right can be regarded as independent of each other and of the variables Z and Z'. Invoking translation invariance of the law of $\widetilde{D}$, the proof is reduced to (3.23). This follows readily from (3.34) and (3.20). □

3.2.2 Convergence for restricted distance

The next several steps hew closely to the original article. Indeed, taking expectation in (3.22) with $\gamma_1 = \gamma_2 = \gamma$ gives

$$E\widetilde{D}(0,x) \le 2E\widetilde{D}\big(0, \lfloor |x|^\gamma Z \rfloor\big) + 1 + |x|_1 P\big(A(x)\big) \tag{3.37}$$

In order to unite the arguments in the two expectations and get an expression that can be iterated, we replace x by the random variable

$$W := Z_0 \prod_{k=1}^{\infty} |Z_k|^{\gamma^k}, \tag{3.38}$$

where $Z_0, Z_1, \dots$ are i.i.d. copies of Z. As shown in [BL19, Lemma 3.1], the infinite product converges and $W \in (0,\infty)$ a.s., with W admitting a continuous, a.e.-non-vanishing probability density and finite moments of all orders. Noting that for W and Z independent we get $|W|^\gamma Z \overset{\text{law}}{=} W$, taking W independent of the $\widetilde{D}$'s then yields

$$E\widetilde{D}(0, rW) \le 2E\widetilde{D}(0, r^\gamma W) + c \tag{3.39}$$

for $c := 1 + \sup_{x\in\mathbb{R}^d} |x|_1 P(A(x))$. This implies the existence of the limit

$$L_\beta(r) := \lim_{n\to\infty} \frac{E\widetilde{D}\big(0, \lfloor r^{\gamma^{-n}} W \rfloor\big)}{2^n} \tag{3.40}$$

giving us

$$\forall r > 1\colon \quad \phi_\beta(r) := L_\beta(r)(\log r)^{-\Delta} \tag{3.41}$$

From (3.40) we get $L_\beta(r^\gamma) = 2L_\beta(r)$, which is responsible for the log-log-periodicity (3.15). The construction via a (essentially) decreasing limit then ensures that ϕ_β is bounded from above on $(1,\infty)$ while [BL19, Theorem 2.5] implies that ϕ_β is also uniformly positive.

While simple, the construction of L_β via (3.40) harbors several conceptual problems. First, it concerns the restricted distance. Second, it depends on W which itself depends on β and η. In [BL19, Section 3], these concerns are dispelled by subsequently proving that, for all $r \geq 1$ and Lebesgue a.e. $x \in \mathbb{R}^d$,

$$\frac{\widetilde{D}\big(0, \lfloor r^{\gamma^{-n}} x \rfloor\big)}{2^n} \underset{n\to\infty}{\longrightarrow} L_\beta(r), \quad P\text{-a.s.} \tag{3.42}$$

see [BL19, Proposition 3.3]. The proof of this is based on the subadditivity estimate (3.22) and, modulo rounding of the arguments of $\widetilde{D}$, it can be taken over *varbatim*.

Another concern is the regularity of $r \mapsto L_\beta(r)$. As in [BL19], this can again be handled using the subadditivity bound (3.22) which gives

$$\widetilde{D}\big(0, \lfloor r^{\gamma^{-n}} x \rfloor\big) \overset{\text{law}}{\leq} \widetilde{D}(0, \lfloor r^{\gamma_1 \gamma^{-n}} |x|^{\gamma_1} Z \rfloor\big) + \widetilde{D}'\big(0, \lfloor r^{\gamma_2 \gamma^{-n}} |x|^{\gamma_2} Z' \rfloor\big) + O(1), \tag{3.43}$$

where, thanks to (3.23), $O(1)$ is bounded in L^1 uniformly in x and $r \geq 1$. Since Z is continuously distributed, (3.42) gives

$$L_\beta(r) \leq L_\beta(r^{\gamma_1}) + L_\beta(r^{\gamma_2}) \tag{3.44}$$

for all $\gamma_1, \gamma_2 \in (0, \frac{1}{2}(1+\gamma))$ with $\gamma_1 + \gamma_2 = 2\gamma$. This implies convexity of $t \mapsto L_\beta(\mathrm{e}^t)$ and thus continuity of $r \mapsto L_\beta(r)$ and $r \mapsto \phi_\beta(r)$ on $(1, \infty)$.

The next step in the argument is the replacement of a limit along doubly exponentially growing sequences by a plane limit $r \to \infty$. This comes at the cost of reinserting W:

Lemma 19. *Suppose $\widetilde{D}$ and W are independent. Then*

$$\frac{\widetilde{D}\big(0, \lfloor rW \rfloor\big)}{L(r)} \underset{r\to\infty}{\longrightarrow} 1, \quad \textit{in probability and in } L^2. \tag{3.45}$$

Proof. The corresponding statement in [BL19] (see Proposition 3.7 there) is deduced from the fact that, for $X_n(r) := 2^{-n}\widetilde{D}(0, r^{\gamma^{-n}} W)$, the limits $EX_n(r) \to L_\beta(r)$ and $\mathrm{Var}(X_n(r)) \to 0$ are locally uniform in $r \geq 1$. This is in turn proved by noting that, thanks to (3.22), both $EX_n(r)$ and $E(X_n(r)^2)$ are downward monotone in n modulo vanishing additive correction terms. As $r \mapsto X_n(r)$ is continuous in the continuum model, the local uniformity is then extracted from Dini's Theorem.

In order to adapt this reasoning to our setting, we need to supply an argument for continuity. This can be achieved by extending the definition of $x \mapsto \widetilde{D}(0,x)$ to all $x \in \mathbb{R}^d$ as follows: Let dist_∞ denote the ℓ^∞-distance on $\mathbb{R}^d$ and let $h\colon [0,1]^d \times \{0,1\}^d \to [0,1]$ be defined by

$$h(x,\sigma) := \bigl[1-\text{dist}_\infty(x,\sigma)\bigr]\Bigl(\sum_{\sigma'\in\{0,1\}^d}\bigl[1-\text{dist}_\infty(x,\sigma')\bigr]\Bigr)^{-1} \tag{3.46}$$

This function is continuous in x with $h(\sigma,\sigma') = \delta_{\sigma,\sigma'}$ for all $\sigma,\sigma' \in \{0,1\}^d$. Now set

$$\widetilde{D}(0,x) := \sum_{\sigma\in\{0,1\}^d} h\bigl(x-\lfloor x\rfloor,\sigma\bigr)\widetilde{D}\bigl(0,\lfloor x\rfloor+\sigma\bigr) \tag{3.47}$$

The subadditive bound (3.22) (which implied the aforementioned downward monotonicity) holds without any rounding albeit with "1" on the right replaced by a d-dependent constant thanks to $|\widetilde{D}(0,x)-\widetilde{D}(0,\lfloor x\rfloor)| \le d$ for all $x\in\mathbb{R}^d$. This constant is irrelevant in the argument and so we can then proceed as in [BL19]. $\square$

Before we move on, we record a useful consequence of above derivations:

Corollary 3. *For each $\beta \in (0,\infty)$ there is $c \in (0,\infty)$ such that*

$$\forall x \in \mathbb{Z}^d \smallsetminus \{0\}\colon \quad E_\beta\bigl(\widetilde{D}(0,x)\bigr) \le c\bigl[1+(\log|x|)^\Delta\bigr] \tag{3.48}$$

Proof. The above gives $E_\beta\widetilde{D}(0,rW) \le c(\log r)^\Delta$ once r is sufficiently large. Using that $Z \overset{\text{law}}{=} W/|W'|^\gamma$ for $W \perp\!\!\!\perp W'$ with $W' \overset{\text{law}}{=} W$ gives

$$E_\beta\widetilde{D}(0,rZ) \le cE\Bigl(\bigl(\log(r|W|^{-\gamma})\bigr)^\Delta\Bigr) \tag{3.49}$$

Since W has a bounded density and at most Gaussian tails, the expectation on the right is at most $c[1+(\log r)^\Delta]$ once r is sufficiently large. The claim now follows from (3.37) and the bound (3.23). $\square$

3.2.3 Actual distance

We are now ready to start working towards the asymptotics of the actual distance D. Paralleling the approach in [BL19, Section 4], fix $\overline{\gamma} \in (\gamma, 1)$ and extend $\widetilde{D}$ to a family of restricted "distance" functions,

$$\widetilde{D}_k(x,y) := \min \left\{ n \geq 0 \colon \begin{array}{c} \{(x_{i-1}, x_i) \colon i = 1, \dots, n\} \subseteq \mathscr{E},\ x_0 = x,\ x_n = y, \\ \forall i = 1, \dots, n \colon\ |x_i - x| \leq 2|x-y|^{\overline{\gamma}^{-k}} \end{array} \right\}. \tag{3.50}$$

These interpolate between the actual distance and the restricted distance monotonically:

$$\begin{aligned} D(x,y) &\leq \cdots \leq \widetilde{D}_{k+1}(x,y) \leq \widetilde{D}_k(x,y) \\ &\leq \cdots \leq \widetilde{D}_1(x,y) \leq \widetilde{D}_0(x,y) = \widetilde{D}(x,y). \end{aligned} \tag{3.51}$$

Since $k \mapsto \widetilde{D}_k(x,y)$ is non-increasing, non-negative, and takes values in $\mathbb{Z}$, the sequence $\{\widetilde{D}_k(x,y)\}_{k\geq 1}$ must stabilize; i.e., $\widetilde{D}_k(x,y) = \widetilde{D}(x,y)$ for all k sufficiently large, depending on x, y, and on the random edges that determine the distances. A key fact is that, at large scales, this happens uniformly with high probability:

Lemma 20. *Let W be independent of the distances $\widetilde{D}_k$ and D. There is a $k \in \mathbb{N}$ such that*

$$\lim_{r\to\infty} P\Big(\widetilde{D}_k\big(0, \lfloor rW \rfloor\big) = D\big(0, \lfloor rW \rfloor\big)\Big) = 0. \tag{3.52}$$

Proof. This is a lattice version of [BL19, Lemma 4.2] whose proof went through by way of the discrete distances and so can be taken over without change. □

The next result to establish is an analogue of [BL19, Lemma 4.3], which bounds the ratio $E\widetilde{D}_k(0, \lfloor rW \rfloor)/L(r)$ asymptotically by one from below. In [BL19], the proof relied on continuity of $\beta \mapsto \phi_\beta(r)$ which was in turn proved using scaling arguments that do not seem to apply here. However, the above does give us the following:

Lemma 21. *For each $r > 1$, $\beta \mapsto \phi_\beta(r)$ is left-continuous and downward monotone. There exists an (at most) countable set $\Sigma \subseteq (0,\infty)$ such that, for each $r > 1$, the function $\beta \mapsto \phi_\beta(r)$ is continuous at all points $\beta' \in (0,\infty) \smallsetminus \Sigma$.*

Proof. In light of (3.41), the downward monotonicity follows from (3.42) and the fact that, under a monotone coupling of edge sets for two different β, distances are ordered pointwise. Being a downward limit of continuous functions, $\beta \mapsto L_\beta(r)$ is left-continuous, and hence so is $\beta \mapsto \phi_\beta(r)$.

The convexity of $t \mapsto L_\beta(\mathrm{e}^t)$ show above guarangees that, for any $0 < \beta_0 < \beta_1 < \infty$ and $1 < r_0 < r_1 < \infty$, the family of functions

$$\bigl\{r \mapsto L_\beta(r) \colon \beta \in [\beta_0, \beta_1]\bigr\} \tag{3.53}$$

is uniformly equicontinuous on $[r_1, r_2]$. This implies that, if $\beta \mapsto \phi_\beta(r)$ is continuous at some $\beta' \in (0, \infty)$ for all $r \in \mathbb{Q} \cap [r_1, r_2]$, then it is continuous at β' for all $r \in [r_1, r_2]$. Invoking the log-log-peridicity (3.15), $\beta \mapsto \phi_\beta(r)$ is continuous for all $r > 1$ as soon as β does not belong to

$$\Sigma := \bigcup_{r \in \mathbb{Q} \cap [\mathrm{e}^\gamma, \mathrm{e}]} \Bigl\{\beta \in (0, \infty) \colon \lim_{\beta' \downarrow \beta} \phi_{\beta'}(r) > \lim_{\beta' \uparrow \beta} \phi_{\beta'}(r)\Bigr\}. \tag{3.54}$$

This set is (at most) countable, since for each $r \in \mathbb{Q} \cap [\mathrm{e}^\gamma, \mathrm{e}]$ the set of jump discontinuities of $\beta \mapsto \phi_\beta(r)$ is at most countable. $\square$

All that continuity of $\beta \mapsto \phi_\beta$ was needed for in [BL19] is condensed into:

Lemma 22. *Let Σ be as in Lemma 21. Then for each $\beta \notin \Sigma$,*

$$\lim_{\beta' \downarrow \beta} \inf_{r > 1} \frac{\phi_{\beta'}(r)}{\phi_\beta(r)} = 1. \tag{3.55}$$

Proof. We will prove the contrapositive. First observe that, by the log-log-periodicity (3.15), we may restrict the infimum to $r \in [\mathrm{e}^\gamma, \mathrm{e}]$ without changing the result. Next, since the ratio is non-increasing in β', we can take β' down to β along any decreasing sequence $\beta_n \downarrow \beta$. The continuity and boundedness imply existence of a minimizer; call it r_n for $\beta' = \beta_n$. By compactness of $[\mathrm{e}^\gamma, \mathrm{e}]$ we may assume $r_n \to r_\infty \in [\mathrm{e}^\gamma, \mathrm{e}]$ as $n \to \infty$. But then the uniform equicontinuity of (3.53) implies

$$\lim_{\beta' \downarrow \beta} \inf_{r \in [\mathrm{e}^\gamma, \mathrm{e}]} \frac{\phi_{\beta'}(r)}{\phi_\beta(r)} = \lim_{n \to \infty} \frac{\phi_{\beta_n}(r_\infty)}{\phi_\beta(r_\infty)}. \tag{3.56}$$

If the latter limit is not equal to one, then $\beta' \mapsto \phi_{\beta'}(r_\infty)$ is not continuous at β, thus forcing $\beta \in \Sigma$. Hence $\beta \notin \Sigma$ implies (3.55) as desired. □

Let us henceforth write P_β for the probability and E_β for the expectation associated with edge probabilities $\mathfrak{p}_\beta$. We then have:

Proposition 2. *Let $\beta \notin \Sigma$ and let W be as in (3.38) for Z with law (3.20) for $\eta := 1$. Then*

$$\forall k \geq 1\colon \qquad \liminf_{r\to\infty} \frac{E_\beta \otimes E_W\, \widetilde{D}_k(0, rW)}{\phi_\beta(r)(\log r)^\Delta} \geq 1, \tag{3.57}$$

where the product of expectation indicates that W and $\widetilde{D}_k$ are independent.

Proof. As in [BL19], the statement will be deduced from the fact (to be proved) that, for each $k \geq 1$, $\beta > 0$ and $\varepsilon \in (0, 1/2)$ there is $c = c(k, \beta, \varepsilon) \in (0, \infty)$ such that

$$E_\beta \otimes E_W\, \widetilde{D}_k\bigl(0, \lfloor r\varepsilon^{-\frac{1}{2d-s}} W\rfloor\bigr) \leq 2E_{\beta(1-2\varepsilon)} \otimes E_W\, \widetilde{D}_{k+1}\bigl(0, \lfloor r^\gamma \varepsilon^{-\frac{1}{2d-s}} W\rfloor\bigr) + c. \tag{3.58}$$

Indeed, dividing both sides $\phi_{\beta(1-2\varepsilon)}(r)(\log r)^\Delta$ and taking $r \to \infty$ shows

$$\liminf_{r\to\infty} \frac{E_{\beta'} \otimes E_W\, \widetilde{D}_{k+1}(0, rW)}{\phi_{\beta'}(r)(\log r)^\Delta} \geq \left[\inf_{r>1} \frac{\phi_\beta(r)}{\phi_{\beta'}(r)}\right] \liminf_{r\to\infty} \frac{E_\beta \otimes E_W\, \widetilde{D}_k(0, rW)}{\phi_\beta(r)(\log r)^\Delta}, \tag{3.59}$$

where $\beta' := \beta(1-2\varepsilon)$. Since (3.57) holds for $k := 0$ and all $\beta > 0$ by Lemma 19, this bounds (3.57) inductively for any $k \geq 1$ by

$$\prod_{j=1}^k \inf_{r>1} \frac{\phi_{\beta(1-2\varepsilon)^{-j}}(r)}{\phi_{\beta(1-2\varepsilon)^{1-j}}(r)} \geq \left[\inf_{r>1} \frac{\phi_{\beta(1-2\varepsilon)^{-k}}(r)}{\phi_\beta(r)}\right]^k, \tag{3.60}$$

where the inequality follows from downward monotonicity of $\beta \mapsto \phi_\beta$. Taking $\varepsilon \downarrow 0$ and applying Lemma 22 we then get (3.57) for all $\beta \notin \Sigma$.

As in [BL19], the proof of (3.58) is based on a variant of the argument from Proposition 1. Let μ, resp., μ' be as in (3.24–3.25) for $\eta := \varepsilon$ and let $\mathscr{I}$, resp., $\mathscr{I}'$ be independent Poisson processes with intensity measures μ, resp., μ'. For any $(\tilde{x}, \tilde{y}) \in \mathbb{R}^d \times \mathbb{R}^d$ satisfying $|\tilde{x}|_2 \leq |\tilde{y}|_2$, $|\tilde{x}| \leq |x|^{\overline{\gamma}}$, and $|\tilde{y} - x| \leq |x|^{\overline{\gamma}}$ we have

$$\mu\Bigl(\bigl(\lfloor\tilde{x}\rfloor + [0,1)^d\bigr) \times \bigl(\lfloor\tilde{y}\rfloor + [0,1)^d\bigr)\Bigr) + \mathfrak{p}_{\beta(1-2\varepsilon)}\bigl(\lfloor\tilde{x}\rfloor, \lfloor\tilde{y}\rfloor\bigr) \leq \mathfrak{p}_\beta\bigl(\lfloor\tilde{x}\rfloor, \lfloor\tilde{y}\rfloor\bigr) \tag{3.61}$$

provided $|x|$ is sufficiently large. Letting $\mathscr{E}'$ be a sample of edge configuration with probabilities $\mathfrak{p}_{\beta(1-2\varepsilon)}$ which we assume independent of $\mathscr{I}$ and $\mathscr{I}''$, we can couple the above processes to a sample $\mathscr{E}$ of edge configurations with probabilities $\mathfrak{p}_\beta$ so that

$$\bigl\{(\lfloor\tilde{x}\rfloor, \lfloor\tilde{y}\rfloor)\colon (\tilde{x},\tilde{y})\in\mathscr{I}\bigr\}\cup\mathscr{E}'\subseteq\mathscr{E}. \tag{3.62}$$

We then use $\mathscr{I}\cup\mathscr{I}'$ to pick a pair (X,Y) minimizing (3.30), define (Z,Z') from these as in (3.32) and $A(x)$ as in (3.34) unless $|x|$ is small, in which case we set $A(x)$ to the whole probability space.

On $A(x)^{\mathrm{c}}$ we are guaranteed $(X,Y)\in\mathscr{I}$ and so $(\lfloor X\rfloor,\lfloor Y\rfloor)\in\mathscr{E}$. Next note that, once $|x|$ is sufficiently large (which is a restriction that is made part of the definition of $A(x)$), the fact that $\gamma<\overline{\gamma}<1$ implies

$$B\bigl(0,2|X|^{\gamma\overline{\gamma}^{-(k+1)}}\bigr)\cup B\bigl(x,2|x-Y|^{\gamma\overline{\gamma}^{-(k+1)}}\bigr)\subseteq B\bigl(0,2|x|^{\overline{\gamma}^{-k}}\bigr) \tag{3.63}$$

whenever $A(x)$ occurs. Writing $\widetilde{D}'_k$ for the distances generated by $\mathscr{E}'$ and $\widetilde{D}_k$ for those generated by $\mathscr{E}$, concatenating a path minimizing $\widetilde{D}_{k+1}\bigl(0,\lfloor X\rfloor\bigr)$ with edge $(\lfloor X\rfloor,\lfloor Y\rfloor)$ and the path minimizing $\widetilde{D}_{k+1}\bigl(x,\lfloor Y\rfloor\bigr)$ produces a path contributing to the optimization underlying $\widetilde{D}_k\bigl(0,x\bigr)$. Thanks to (3.62) we thus get

$$\widetilde{D}_k\bigl(0,x\bigr)\le\widetilde{D}'_{k+1}\bigl(0,\lfloor X\rfloor\bigr)+\widetilde{D}'_{k+1}\bigl(x,\lfloor Y\rfloor\bigr)+1+|x|_1 1_{A(x)}. \tag{3.64}$$

Rewriting X and Y using Z and Z', plugging for W for x and taking expectation, this yields (3.58) except with W defined using $\eta:=\varepsilon$. As a calculation shows, the change in normalization effectively replaces W by $\varepsilon^{-\frac{1}{2d-s}}W$. $\square$

We are now ready to give:

Proof of Theorem 11. Let $\beta\notin\Sigma$. Summarizing the above developments, for W (defined using $\eta:=1$) independent of D we have

$$\frac{D\bigl(0,\lfloor rW\rfloor\bigr)}{L_\beta(r)}\underset{r\to\infty}{\longrightarrow}1\qquad\text{in probability and }L^2. \tag{3.65}$$

Indeed, the upper bound is supplied by Lemma 19 and $D(0,x) \le \widetilde{D}(0,x)$, while the lower bound follows from Lemma 20 and Proposition 2.

Fix $\delta \in (0,1)$. Using that W admits a probability f, the expectation of the quantity on the left of (3.16) is bounded by

$$\frac{1}{r^d}\#B(0,\delta r^d) + \frac{c_r(\delta)}{\varepsilon} E_\beta \otimes E_W\left(\left|\frac{D(0,\lfloor rW \rfloor)}{L_\beta(r)} - 1\right|\right), \tag{3.66}$$

where

$$c_r(\delta) := \max_{x \in B(0,r) \smallsetminus B(0,\varepsilon r)} \left(\int_{x+[0,1]^d} f(z/r)\,\mathrm{d}z\right)^{-1}. \tag{3.67}$$

Since f is continuous positive on $R^d \smallsetminus \{0\}$ we have $\sup_{r \ge 1} c_r(\delta) < \infty$. The second term thus tends to zero as $r \to \infty$ by (3.65). Noting that $r^{-d}\#B(0,\delta r) \le c\delta^d$, the claim follows by taking $r \to \infty$ and $\delta \downarrow 0$. $\square$

REFERENCES

[AK81] Mustafa A. Akcoglu and Ulrich Krengel. "Ergodic Theorems for Superadditive Processes." *J. Reine Angew. Math.*, **323**:53–67, 1981.

[AP96] Peter Antal and Agoston Pisztora. "On the chemical distance for supercritical Bernoulli percolation." *Ann. Probab.*, **24**(2):1036–1048, 1996.

[Bar16] Albert-László Barabási. *Network Science*. Cambridge Univ. Press, 2016.

[BB01] Itai Benjamini and Noam Berger. "The diameter of long-range percolation clusters on finite cycles." *Random Struct. Alg.*, **19**(2):102–111, 2001.

[BCG16] Alexei Borodin, Ivan Corwin, and Vadim Gorin. "Stochastic Six-Vertex Model." *Duke Math. J.*, **165**(3):563–624, February 2016.

[Bei77] Henk van Beijeren. "Exactly Solvable Model for the Roughening Transition of a Crystal Surface." *Phys. Rev. Lett.*, **38**(18), 1977.

[Ber04] Noam Berger. "A lower bound for the chemical distance in sparse long-range percolation models." `arXiv:math/0409021`, 2004.

[Bis04] Marek Biskup. "On the scaling of the chemical distance in long-range percolation models." *Ann. Probab.*, **32**(4):2938–2977, 2004.

[Bis11a] Marek Biskup. "Graph diameter in long-range percolation." *Random Struct. Alg.*, **39**(2):210–227, 2011.

[Bis11b] Marek Biskup. "Recent progress on the random conductance model." *Probab. Surv.*, **8**:294–373, 2011.

[BK21] Marek Biskup and Andrew Krieger. "Arithmetic oscillations for the chemical distance in long-range percolation on $\mathbb{Z}^d$." In preparation, 2021.

[BKP04] Itai Benjamini, Harry Kesten, Yuval Peres, and Oded Schramm. "Geometry of the uniform spanning forest: transitions in dimensions 4,8,12,. . . ." *Ann. Math.*, **160**:465–491, January 2004.

[BL19] Marek Biskup and Jeffrey Lin. "Sharp asymptotic for the chemical distance in long-range percolation." *Random Struct. Alg.*, **55**(3):560–583, 2019.

[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] Raphaël Cerf. *The Wulff Crystal in Ising and Percolation Models*, volume 1878 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006.

[CGS02] Don Coppersmith, David Gamarnik, and Maxim Sviridenko. "The diameter of a long-range percolation graph." *Random Struct. Alg.*, **21**(1):1–13, 2002.

[CKP01] Henry Cohn, Richard Kenyon, and James Propp. "A Variational Principle for Domino Tilings." *J. Amer. Math. Soc.*, **14**(2):297–346 (electronic), 2001.

[CL90] John H Conway and Jeffrey C Lagarias. "Tiling with Polyominoes and Combinatorial Group Theory." *J. Comb. Theory Ser. A*, **53**(2):183–208, March 1990.

[CS16] Filippo Colomo and Andrea Sportiello. "Arctic Curves of the Six-Vertex Model on Generic Domains: The Tangent Method." *J. Stat. Phys.*, **164**(6):1488–1523, 2016.

[Des98] Nicolas Destainville. "Entropy and Boundary Conditions in Random Rhombus Tilings." *J. Phys. A*, **31**(29):6123–6139, 1998.

[DKS92] Roland Lvovich Dobrushin, Roman Kotecký, and Senya 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.

[DS15] Jian Ding and Allan Sly. "Distances in critical long range percolation." `arXiv:1303.3995`, 2015.

[Dur10] Rick Durrett. *Probability: Theory and Examples.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge Univ. Press, fourth edition, 2010.

[JPS98] William Jockusch, James Propp, and Peter Shor. "Random Domino Tilings and the Arctic Circle Theorem." `arXiv:math/9801068`, 1998.

[Kah01] Jeff Kahn. "Range of cube-indexed random walk." *Isr. J. Math.*, **124**:189–201, 2001.

[Kas63] Pieter W Kasteleyn. "Dimer Statistics and Phase Transitions." *J. Math. Phys.*, **4**:287–293, 1963.

[Kin68] J. F. C. Kingman. "The Ergodic Theory of Subadditive Stochastic Processes." *J. Royal Stat. Soc. Ser. B*, **30**(3):499–510, 1968.

[Kir34] Mojżesz D Kirszbraun. "Über die zusammenziehende und Lipschitzsche Transformationen." *Fundam. Math.*, **22**(1):77–108, 1934.

[KMT20a] Andrew Krieger, Georg Menz, and Martin Tassy. "Deducing a variational principle with minimal a priori assumptions." *Electron. J. Comb.*, **27**(4), 2020.

[KMT20b] Andrew Krieger, Georg Menz, and Martin Tassy. "Homogenization of the variational principle for discrete random maps." `arXiv:1710.11330`, 2020.

[Kri21] Andrew Krieger. “The concentration inequality for a discrete height function model perturbed by random potential.” `arXiv:2111.05907`, 2021.

[Lig85] Thomas M. Liggett. “An Improved Subadditive Ergodic Theorem.” *Ann. Probab.*, **13**(4):1279–1285, November 1985.

[LRS01] Michael Luby, Dana Randall, and Alistair Sinclair. “Markov Chain Algorithms for Planar Lattice Structures.” *SIAM J. Comput.*, **31**(1):167–192, 2001.

[LS77] Benjamin F Logan and Larry A Shepp. “A Variational Problem for Random Young Tableaux.” *Adv. Math.*, **26**(2):206–222, 1977.

[LT20] Piet Lammers and Martin Tassy. “Macroscopic behavior of Lipschitz random surfaces.” `arXiv:2004.15025`, 2020.

[MT20a] Georg Menz and Martin Tassy. “A Variational Principle for a Non-Integrable Model.” *Probab. Theory Relat. Fields*, **177**(3):747–822, 2020.

[MT20b] Georg Menz and Martin Tassy. “A Variational Principle for a Non-Integrable Model.” *Probab. Theory Relat. Fields*, **177**(3):747–822, 2020.

[New18] Mark Newman. *Networks*. Oxford Univ. Press, 2018.

[PR07] Boris Pittel and Dan Romik. “Limit Shapes for Random Square Young Tableaux.” *Adv. Appl. Math.*, **38**(2):164–209, 2007.

[RS17] Nicolai Reshetikhin and Ananth Sridhar. “Integrability of limit shapes of the six vertex model.” *Commun. Math. Phys.*, **356**(2):535–565, 2017.

[RS18] Nicolai Reshetikhin and Ananth Sridhar. “Limit Shapes of the Stochastic Six Vertex Model.” *Commun. Math. Phys.*, **363**(3):741–765, 2018.

[Sch14] Jean Van Schaftingen. “Approximation in Sobolev Spaces by Piecewise Affine Interpolation.” *J. Math. Anal. Appl.*, **420**(1):40–47, 2014.

[She01] Scott Sheffield. “Ribbon Tilings and Multidimensional Height Functions.” *Trans. Am. Math. Soc.*, **354**, August 2001.

[She05] Scott Sheffield. *Random Surfaces*, volume 304 of *Astérisque*. Société mathématique de France, 2005.

[Sri16] Ananth Sridhar. *Limit Shapes in the Six Vertex Model.* PhD thesis, University of California, Berkeley, 2016.

[VK77] Anatolii Moiseevich Veršik and Sergei 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.

[Wil04] David Bruce Wilson. "Mixing Times of Lozenge Tiling and Card Shuffling Markov Chains." *Ann. Appl. Probab.*, **14**(1):274–325, February 2004.