Shannon-MacMillan theorems for random fields along curves and lower bounds for surface-order large deviations

Shannon-MacMillan theorems for random fields along curves and lower bounds for surface-order large deviations

Julia Brettschneider

EURANDOM, P.O.Box 513, 5600 MB Eindhoven, The Netherlands

j.brettschneider@tue.nl[1]

Abstract

Let P be a random field over the two-dimensional lattice $\mathbb{Z}^2$ with finite state space. We introduce the notion of specific entropy $h_c(P)$ of the field along a curve c as the limit of rescaled entropies along the lattice approximations of the blowups of c. We prove a corresponding Shannon-MacMillan theorem. This allows us to represent $h_c(P)$ as a mixture of specific entropies along the tangent lines of c.

The proof is accomplished in three steps. A Shannon-MacMillan theorem along lines is proved for stationarity P. In a second step, we assume a strong 0-1 law for P and extend the result to polygons. Finally, the specific entropy along a curve is obtained by approximation.

As an application, we use the specific entropy along curves to refine Föllmer and Ort's lower bound for the large deviations of attractive Gibbs measures in the phase-transition regime.

1 Introduction

Since the groundbreaking work of Shannon, entropy has played an important role in the analysis of stochastic processes. In particular, entropy is a key concept in the theory of random fields, a meeting point of ergodic theory and statistical mechanics. For example, the variational principle of Lanfort and Ruelle [16] characterizes the Gibbs measures with a given interaction potential as minimizing points of a functional defined in terms of energy and entropy. Alternatively, this characterization can be formulated in strictly information-theoretic terms, by means of the concept of relative entropy (see Föllmer [7]). Such relative entropies are an essential tool in analyzing large deviations of empirical fields from their ergodic behaviour. For this purpose, it is crucial to establish a Shannon-MacMillan theorem, which says that there is an $\mathcal{L}^1$-convergence of suitable rescaled information quantities behind the existence of a relative entropy.

Typically, the entropy of a stationary random field P indexed by a d-dimensional lattice is defined as a limit of entropies on an increasing sequence of boxes, rescaled by the volume of the boxes. In the context of large deviations, however, such volume-order quantities may not provide enough information when a phase transition occurs. For this reason, Föllmer and Ort [10] introduced the concept of surface-order entropy on boxes. They proved corresponding versions of the Shannon-MacMillan theorem and used them to estimate large-deviation probabilities. However, the construction of the Wulff shape by Dobrushin, Kotecky, and Shlosman [5] suggests that such estimates can be improved if boxes are replaced by more general shapes.

Hans Föllmer suggested to investigating the problem of constructing entropies on general surfaces and of proving appropriate versions of the Shannon-MacMillan theorem. In this paper, we

[1] For correspondance after July 2001 use juliab@stat.berkeley.edu

carry out this program in the two-dimensional case, where surfaces reduce to contour curves. In this context, the existence of specific entropy does not simply follow from a subadditivity argument. Instead, we consider directly the problem of proving an appropriate Shannon-MacMillan theorem. Our construction of the specific entropy $h_c(P)$ of a random field P along a curve c involves lattice approximations of the successive blowups of the given curve. We prove a corresponding Shannon-MacMillan theorem, i.e., $L^1(P)$-convergence of rescaled information quantities along these lattice approximations. Our proof relies on extending uniform convergence results in ergodic theory to a suitable skew-product transformation. This leads to an explicit formula for the specific entropy $h_c(P)$ involving the conditional entropy of the random field restricted to the origin, given the σ-algebra of a suitable defined "past" along the curve. Under certain conditions, this construction can be extended to relative entropies of one random field with respect to another. This will be the key to our discussion of refined lower bounds for large deviations in the presence of a phase transition.

We now explain our results in more detail.

Shannon-MacMillan theorems. Consider a random sequence ω of letters from a finite alphabet Υ, modelled by a probability measure P on the space $\Omega := \Upsilon^{\{1,2,\cdots\}}$. For any finite n, the *information* provided by the first n letters can be described by the function

$$-\log P[\omega_{\{1,\ldots,n\}}],$$

where $\omega_{\{1,\ldots,n\}}$ denotes the restriction of ω to $\{1,\ldots,n\}$, and $P\big[(y_1,\ldots,y_n)\big]$ is the probability that the pattern $(y_1,\ldots,y_n)$ appears in the first n trials. In the classical case, when the letters are independent and identically distributed according to a measure μ, the classical Shannon-MacMillan theorem states that the rescaled information functions,

$$-\frac{1}{n}\log P[\omega_{\{1,\ldots,n\}}],$$

converge in $\mathcal{L}^1(P)$ to the *entropy*

$$H(\mu) := -\sum_{y\in\Upsilon}\mu(y)\log\mu(y)$$

of the measure μ. The theorem can be extended to a general ergodic sequence. In the bilateral case, when P is an ergodic measure on $\Upsilon^{\mathbb{Z}}$, the limiting quantity takes the form

$$h(P) := E\Big[H\big(P_0[\cdot\,|\,\mathcal{P}]\big)\Big],$$

where $P_0[\cdot\,|\,\mathcal{P}]$ is the conditional distribution of ω_0, and $\mathcal{P}$ is the "past" σ-algebra generated by the projection of ω to the set $\{-1,-2,\ldots\}$.

These constructions can be extended to a spatial setting when the random field is given by a stationary probability measure P on a configuration space $\Upsilon^{\mathbb{Z}^d}$. Thouvenot [28] and Föllmer [7] proved spatial extensions of Shannon and MacMillan's result. The *specific entropy* is introduced as

$$h(P) := \lim_{n\to\infty}\frac{1}{|V_n|}H_{V_n}(P),$$

where V_n is the set of all lattice sites in $[-n,n]^d$, and $H_{V_n}(P)$ is the entropy of the measure P restricted to V_n. The existence of the limit follows from the subadditivity of H_V with respect to V.

The corresponding Shannon-MacMillan theorem shows that there is an $\mathcal{L}^1(P)$-convergence of the functions

$$-\frac{1}{|V_n|}\log P[\omega_{V_n}]$$

behind the existence of the specific entropy. If P is ergodic then we obtain the formula

$$h(P) = E\big[H\big(P_0[\cdot\,|\,\mathcal{P}^d]\big)\big],$$

where $\mathcal{P}^d$ is a σ-algebra representing a spatial version of the "past". More precisely, $\mathcal{P}^d$ is generated by the projections of ω to the sites preceding 0 in the lexicographical ordering of $\mathbb{Z}^d$.

Surface entropy. Our goal is to derive refined versions of the Shannon-MacMillan theorem, where the information functions are observed along surfaces. This was carried out in [10] for the surfaces of boxes parallel to the axes. In this work we consider the two-dimesional case, and we develop a construction of surface entropy where rectangles are replaced by general curves. More precisely, guided by a suggestion of Hans Föllmer, we introduce the specific entropy along sets generated by lattice approximations to blowups of lines, and then extend this to polygons and piecewise smooth curves.

Our first result is a Shannon-MacMillan theorem for the *specific entropy* $h_\lambda(P)$ of a stationary random field P *along a line* with slope λ (see Theorem 3.6 for rational and Theorem 3.7 for irrational slopes). We prove the $\mathcal{L}^1(P)$-convergence of the rescaled information functions along increasing segments of the line's *lattice approximation*

$$(z, [\lambda z + a]) \qquad (z \in \mathbb{Z}), \tag{1}$$

where $[x]$ denotes the integer part of x. If P fulfills a 0-1 law on the tail field, we obtain the formula

$$h_\lambda(P) = \int_0^1 E\big[H\big(P_0[\,\cdot\,|\,\mathcal{P}_{\lambda,t}]\big)\big]\,dt, \tag{2}$$

where $\mathcal{P}_{\lambda,t}$ is the σ-algebra generated by those approximating sites which precede 0 in the lexicographical ordering of $\mathbb{Z}^2$.

If λ is rational, the mixing condition can be replaced by an ergodicity assumption. Furthermore, the formula (2) can be written as

$$\frac{1}{q}\sum_{\nu=0}^{q-1} E\Big[H\big(P_0[\,\cdot\,|\mathcal{P}_{\frac{p}{q},\frac{\nu p}{q}}]\big)\Big],$$

where $\frac{p}{q}$ is the unique representation of λ by integers $p \in \mathbb{N}$ and $q \in \mathbb{Z}$ having no common divisor. The past σ-algebras $\mathcal{P}_{\frac{p}{q},\frac{\nu p}{q}}$ correspond to the q diffferent possibilities to start the q-periodic pattern of the lattice approximation (1).

The extension to polygons requires a strong form of the 0-1 law on the tail field, which was introduced in [10]. It says that, for any subset J of $\mathbb{Z}^2$, the σ-algebra generated by the sites in J does not increase if we add information about the tail behaviour outside of J; see Definition 2.2. Under this condition, we prove a Shannon-MacMillan theorem along polygons (see Theorem 4.5). In particular, we obtain a representation of the *specific entropy* of P *along a polygon* as a mixture of entropies along lines corresponding to its edges. Finally, we prove in Theorem 4.8 that the *specific entropy along a curve* $c : [0,T] \longrightarrow \mathbb{R}^2$ is a mixture

$$h_c(P) = \int_0^T h_{c'(t)}(P)\,dt \tag{3}$$

of entropies along the tangent lines. Here, $h_{c'(t)}(P)$ denotes the specific entropy along a line having the same slope as the tangent of c in t; see (69) for the exact definition.

With a view toward different types of Markov fields, we could use an alternative lattice approximation of a given line. Instead of approximating the line by the set of sites (1), we can use the sites corresponding to a contour, i.e., a chain of bonds. Corresponding versions of the Shannon-MacMillan theorem are developed in Section 3.

About the proof. On an interval $I \subset \mathbb{R}$, we define the lattice approximation of a line with slope $\lambda \in [0,1]$ and y-intercept a by

$$L_{\lambda,a}(I) = (z, [\lambda z + a]) \qquad (z \in I \cap \mathbb{Z}) \tag{4}$$

We want to prove the convergence of the rescaled informations

$$-\frac{1}{n+1} \log P\big[\omega_{L_{\lambda,a}([0,n])}\big] \quad (n \in \mathbb{N})$$

along sucessively larger segments of the line. To make this problem accessible to ergodic theory, we have to find a transformation which captures the stair climbing pattern along the lattice approximation of the line. If the slope is rational, the steps become periodic, and we proceed by combining a finite number of different transformations. In the case of an irrational slope, this method fails. Here, we need to keep track not only of the integer part but also of the fractional part $\{\lambda z + a\}$ in each step. This suggests the skew-product transformation

$$\begin{aligned} S_\lambda: \quad \mathbb{T} \times \Omega \quad &\longrightarrow \mathbb{T} \times \Omega \\ (t, \omega) \quad &\longmapsto \big(\tau_\lambda(t), \vartheta_{(1,[\lambda z + a])}\omega\big), \end{aligned} \tag{5}$$

where $\mathbb{T}$ is the one-dimensional torus, equipped with the Borel σ-algebra and the Haar measure, and τ_λ is the translation by λ. Using appropriate ergodic theorems for skew products developed in [2] we obtain a Shannon-MacMillan theorem along the lattice approximation of the line.

The second step toward a specific entropy along general contours is a Shannon-MacMillan theorem along polygons. Given a polygon π, parametrized on $[0,T]$, we study the sequence

$$-\Big(\log P\big[\omega_{L_n^\pi}\big]\Big)_{n \in \mathbb{N}}$$

of rescaled informations of P restricted to the lattice approximations L_n^π of the blowups

$$B_n\pi(t) := n\pi\Big(\frac{t}{n}\Big) \qquad (t \in [0, nT])$$

of π. Conditioning site by site, the problem can be reduced to the Shannon-MacMillan theorem along the edges, which is essentially covered by the first step. The difficulty of getting around the corners remains, but it can be overcome by the technique which Föllmer and Ort used in the case of boxes. Here that we need the strong form of a 0-1 law (Definition 2.2). Under this condition, the entropy along a polygon is represented as a mixture of the entropies of its edges.

Our last step is the entropy along a piecewise smooth curve. By approximation with polygons, we obtain our main result (see Theorem 4.8), i.e., the formula (3) which represents the specific entropy along a curve as a mixture of the surface entropies of its tangent vectors.

Relative entropy and large deviations. Shannon-MacMillan theorems for the *specific relative entropy* $h(Q,P)$ of two probability measures Q and P on the sequence space $\Upsilon^{\{1,2,\cdots\}}$ are based on the functions

$$-\log \frac{dQ}{dP}[\omega_{\{1,\ldots,n\}}]$$

describing the *relative information* gained from the first n trials of an experiment. They are a key tool in the search for estimates in the theory of large deviations.

By a *large deviation* we mean a rare event, or an untypical behavior occuring in a random sequence. Consider the *empirical distributions*

$$\mu_n(\omega) := \sum_{i=1}^{n} \delta_{\omega_i} \quad (n \in \mathbb{N})$$

of a stationary random sequence ω_i $(i \in \mathbb{N})$. If the measure P is ergodic then μ_n converges to the marginal distribution μ of P, P-almost surely and in $\mathcal{L}^1(P)$. Large deviations are events like $[\mu_n \in A]$, A being a set in the space of probability measures on Υ whose closure does not contain μ.

The aim of large deviation theory is to find lower and upper bounds which describe the asymptotic decay of the probabilities of such large deviations. In the classical case of a sequence of independent and identically distributed random variables, the decay of large deviations of the empirical distribution is described by Sanov's theorem. Cramér's theorem addresses similar questions for the empirical averages. As a third level for investigating large deviations, Donsker and Varadhan [6] initiated the investigation of large deviations of empirical processes.

In this paper we replace the random sequence by a random field, and the empirical processes by the *empirical fields*

$$R_n(\omega) := \frac{1}{|V_n|} \sum_{i \in V_n} \delta_{\vartheta_i \omega},$$

where ϑ_i $(i \in \mathbb{Z}^d)$ denotes the group the shift transformations.

Comets [3], Föllmer and Orey [9], and Olla [19] found the following large deviation principle for the empirical fields of a stationary Gibbs measure: For any open subset A of the space $\mathcal{M}_1(\Omega)$, of probability measures on $\Omega = \Upsilon^{\mathbb{Z}^d}$,

$$\liminf_{n\to\infty} \frac{1}{|V_n|} \log P(R_n \in A) \geq - \inf_{Q \in A \cap M_1(\Omega)} h(Q,P), \tag{6}$$

and for any closed set $C \in \mathcal{M}_1(\Omega)$,

$$\limsup_{n\to\infty} \frac{1}{|V_n|} \log P(R_n \in C) \leq - \inf_{Q \in C \cap M_1(\Omega)} h(Q,P), \tag{7}$$

where the *rate function* is given in terms of the specific relative entropy

$$h(Q,P) := \lim_{n\to\infty} \frac{1}{|V_n|} H_{V_n}(Q,P).$$

Phase transition. In the case of phase transition, there exists more than one Gibbs measure with respect to the same potential. We are then faced with the following problem. Due to the variational principle for Gibbs measures (see Föllmer [7] and Lanford and Ruelle [16]), the specific relative entropy of P to another stationary Gibbs measure Q with the same interaction potential vanishes. Thus, the relative entropy $h(Q,P)$ appearing in (6) and (7) may be zero even though Q is not contained in the closure of A. This suggests we need a refined analysis of large deviations in terms of surface-order rather than volume-order entropies. Assume in fact that the interaction satisfies the local Markov property. Then

$$H_V(Q,P) = H_{\partial V}(Q,P)$$

for any finite subset V of $\mathbb{Z}^2$, where ∂V is the boundary of V, that is, the set of all sites outside of V which have distance 1 to V (see the end of Section 2 in [10]). Consequently, this relative entropy is in fact a surface-order term, and so it should be rescaled not by the size of the volume $|V|$ but by the size of its surface $|\partial V|$. This observation was the main motivation for introducing the concept of surface entropy, and for proving the corresponding Shannon-MacMillan theorem.

In the context of the two-dimensional Ising model, Schonmann [25] showed the existence of surface-order upper and lower bounds for the large deviations of the emperical means. For attractive models with a totally ordered state space, Föllmer and Ort found a lower bound for the large deviations of the empirical field in terms of the relative surface entropies along boxes (see [10] or Theorem 6.1).

In their detailed analysis of the two-dimensional Ising model, Dobrushin, Kotecky, and Shlosman [5] justified on the basis of local interactions, that the phase-separating curve has the form of a Wulff shape. They proved a large deviation principle with a rate function in terms of the surface tension along the Wulff shape. Using different methods, Ioffe (see [12] and [13]) was able to extend their result up to the critical temperature. The appearance of such shapes suggests to extend the large deviation bounds of Föllmer and Ort from rectangles to general shapes.

This extension is carried out in the last section of this work, where we consider Gibbs measures with attractive interactions on a two-dimensional lattice. We use the generalized surface entropies introduced in Section 4 to refine the lower bound obtained by Föllmer and Ort [10]. Theorem 6.2 gives a lower bound in terms of the specific relative entropies along curves. The probabilistic part of the proof is similar to [10], but we need additional geometrical arguments. In particular, the asymptotic ratio of the length of a line segment and its lattice approximation comes into play. In the lower bound (Theorem 6.2) these quantities merge into a factor involving the derivative of the curve.

As an alternative to the lattice approximation, we can use the contour approximation, which corresponds to a different definition of the boundary of a subset of $\mathbb{Z}^2$. If the Markov property holds only for the contour boundary, we can again prove a lower bound (Theorem 6.5), where the surface entropy is constructed in terms of the contour approximation.

The role of Shannon-MacMillan theorems in the refined analysis of large deviations provided the original motivation for this work. It seems, however, that the study of entropies along surfaces may hold independent interest. This paper lays some of the groundwork for such a general theory of specific entropies along shapes.

Outline of the paper. The first section reviews the notions of a two-dimensional discrete random field, the boundary of a lattice set, the local and the global Markov property, tail-triviality, short-range correlations, and Föllmer and Ort's 0-1 law. We further recall the definitions of information and entropy, and the Shannon-MacMillan theorem for stationary random fields.

In the next section, we introduce a *specific entropy* of a random field P *along a line*. The idea to investigate such an object has two precursors. The first is a volume-order *directional entropy*, which Milnor (see [17] and [18]) introduced in the context of cellular automata. The second is specific entropy along hyperplanes perpendicular to an axis. Föllmer and Ort [10] defined this as a step toward their surface entropy along boxes. We combine these ideas for a two-dimensional random field P. We prove a Shannon-MacMillan theorem and an explicit representation for the specific entropy $h_\lambda(P)$ of P along a line with slope λ.

The key to our proof is a careful description of the line's lattice approximation. If the slope λ is rational the steps in the lattice approximation become periodic. We can then prove a Shannon-MacMillan theorem by combining a finite number of different transformations. In the case of an irrational slope, no such simplification is possible. Instead, we must use a skew-product transformation, whose second component keeps track of the irrational remainder at each step. It may be noted that a technical distinction between rational and irrational slopes was also made by Sinai in his work [26] on Milnor's directional entropy for cellular automata. This construction was further developed by Park (see [20], [21], and [22]) and Sinai [27]). The original problem of continuity with respect to the direction was finally solved in Park [23].

In Section 4 we construct the *specific entropy* $h_c(P)$ of a random field P *along a curve* c. It will be obtained as the limit of renormalized entropies along lattice approximations of the blowups of c. We prove an underlying Shannon-MacMillan theorem, which states that there is an $\mathcal{L}^1$-convergence of suitable rescaled information quantities behind the existence of the specific entropie. We further prove a formula which represents $h_c(P)$ as a mixture of the specific entropies along its tangent lines.

The proof is divided into three parts. Inspired by the construction of the specific entropy along a line, we start again with a linear shape. Since the blowup procedure moves the shape in space, we cannot immediately apply our result for the entropy along a line, but we will imitate that proof in Section 3. In the second part, we prove a Shannon-MacMillan theorem along polygons. In the last part, we extend the result to curves, approximating them by polygons. Assuming that P fulfills Föllmer and Ort's 0-1 law we can extend the result to polygons. We can then pass by polygonal approximation to general piecewise smooth curves. We conclude by deriving a scaling property of $h_c(P)$.

The last two sections focus on lower bounds for large deviations of Gibbs measures with attractive interactions in the case of a phase transition. The proofs depend upon the surface entropies which we constructed in Chapter 4, and on the corresponding Shannon-MacMillan theorems. We begin by recalling the notions of interaction potential, energy, Gibbs measure, and phase transition. We then turn our attention to attractive potentials, and to the extremal random fields P^+ and P^-. In this context, we introduce the *specific relative entropy* $h_c(P^-, P^+)$ *along a curve* c.

The main result (Theorem 6.2) of this part appears in the final section. It is a lower bound for the large deviations under a Markov assumption. The proof uses the well known strategy of switching to a measure under which the large deviation becomes normal behavior, and then applying a Shannon-McMillan theorem. Making use of the global Markov property, we pass from densities restricted to the lattice points inside of a polygon to densities on the lattice approximations of its boundary. In this context, we prove an appropriate *relative* version of the Shannon-MacMillan theorem, in analogy to the results in Section 4. The second ingredient in the proof are geometrical observations resulting from replacing line segments parallel to the axes by general line segments. For example, Lemma 6.3 computes that the asymptotic contribution of the fraction between a lattice approximation of a line segment and its length equals $\left(\sqrt{1+\lambda^2}\right)^{-1}$, where λ is the slope of the line segment. As an alternative, we prove a similar bound (Theorem 6.5) in the case when the Markov property is only satisfied with respect to contour bound.

2 Random fields

Consider $\Omega := \Upsilon^{\mathbb{Z}^2}$, where Υ is a finite set. For any subset V of $\mathbb{Z}^2$ define $\Omega_V := \Upsilon^V$. Let ω_V be the projection of ω to V, P_V the distribution of ω_V with respect to P, and $\mathcal{F}_V := \sigma(\omega_V)$ the σ-algebra generated by this projection. A probability measure P on $(\Omega, \mathcal{F})$ is called a two-dimensional discrete *random field.* The transformations $(\theta_v)_{v\in\mathbb{Z}^2}$ defined by $\theta_v\omega(u) = \omega(u+v)$ $(u \in \mathbb{Z}^2)$ form the group of transformations on Ω, called *shift transformations.* We assume P is *stationary,* that is, invariant with respect to the shift transformations. The classical case of a random field is a collection $(\omega_u)_{u\in\mathbb{Z}^2}$ of independent random variables.

There are different levels of Markov properties for random fields: when the subset of the lattice which generates the condition has to be *finite,* and when it can be any type of subset of the lattice. They both involve the *boundary*

$$\partial V := \{\, j \in \mathbb{Z}^2 \setminus V \,|\, \mathrm{dist}_V(j) = 1 \,\}. \tag{8}$$

of a subset V of the lattice $\mathbb{Z}^2$.

Definition 2.1. *A random field P has the* local Markov property *if, for any finite $V \subset \mathbb{Z}^d$ and for any nonnegative $\mathcal{F}_V$-measurable ϕ,*

$$E[\phi \,|\, \mathcal{F}_{\mathbb{Z}^d\setminus V}] = E[\phi \,|\, \mathcal{F}_{\partial V}]. \tag{9}$$

A random field P which fulfills the local Markov property is called a Markov field*. If (9) holds for all any $V \subset \mathbb{Z}^d$, then P has the* global Markov property.

In Section 5, we will introduce the class of Gibbs measures in terms of interaction potentials. Any Gibbs measure belonging to a nearest-neighbor potential is a Markov field. Examples of random fields which have the local Markov property but not the global Markov property were given by Weizsäcker [29] and Israel [14].

Note that the boundary (8) is not necessarily a *contour* in the sense of statistical mechanics, that is, a chain of bonds. Closing the gaps, we obtain the *contour boundary,* defined as

$$\widehat{\partial} V := \{\, j \in \mathbb{Z}^2 \setminus V \,|\, \mathrm{dist}_V(j) = 1 \text{ or } \mathrm{dist}_V(j) = \sqrt{2} \,\}. \tag{10}$$

Replacing ∂ by $\widehat{\partial}$ in Definition 2.1 leads to slightly different Markov properties.

P is called *tail-trivial* if it fulfills a 0-1 law on the *tail field*

$$\mathcal{T} := \bigcap_{V\subset\mathbb{Z}^2 \text{ finite}} \mathcal{F}_{\mathbb{Z}^2\setminus V} = \bigcap_{n\in\mathbb{N}} \mathcal{F}_{Z^2\setminus V_n}, \tag{11}$$

where $V_n := \{v \in \mathbb{Z}^2 \,|\, \|v\| \leq n\}$, with the maximum norm $\|\cdot\|$. Due to the spatial structure of a random field, tail-triviality is equivalent (cf. Proposition 7.9 from [11]) to a mixing condition called *short-range correlations*:

$$\sup_{A \in \mathcal{F}_{Z^2 \setminus V_n}} |\, P(A \cap B) - P(A)P(B)\,| \xrightarrow{n \to \infty} 0 \qquad \text{for all } B \in \mathcal{F}. \tag{12}$$

The following condition was introduced by Föllmer and Ort in [10].

Definition 2.2. *P satisfies the* strong 0-1 law *if for any subset J of $\mathbb{Z}^d$ the σ-algebra $\mathcal{F}_j$ coincides modulo P with the σ-algebra*

$$\mathcal{F}_J^* := \bigcap_{V \subset \mathbb{Z}^2 \text{ finite}} \mathcal{F}_{J \cup (\mathbb{Z}^2 \setminus V)}. \tag{13}$$

For $J = \emptyset$ it reduces to the classical 0-1 law on $\mathcal{F}$. Remark 3.2.3 from [10] shows that the strong 0-1 law implies the global Markov property for P provided P has the local Markov property.

Let V and W be subsets of $\mathbb{Z}^2$. The *information* in ω restricted to V, with respect to P, is given by the random variable

$$\mathcal{I}(P_V)(\omega) := -\log P[\omega_V], \tag{14}$$

and the information conditioned on $\mathcal{F}_W$ is defined as

$$\mathcal{I}\big(P_V[\,\cdot\,|\mathcal{F}_W](\omega)\big) := -\log P[\omega_V \,|\, \omega_W].$$

The *entropy* of P restricted to V is

$$H_V(P) := -E[\mathcal{I}(P_V)(\omega)] = -\sum_{\omega \in \Upsilon^{\mathbb{Z}^2}} P[\omega_V] \log P[\omega_V], \tag{15}$$

and the *conditional entropy* of P to $\mathcal{F}_W$ is

$$H_V(P[\,\cdot\,|\mathcal{F}_W]) := -E[\log P[\omega_V \,|\, \omega_W]] = H(P_V[\,\cdot\,|\mathcal{F}_W]).$$

The *specific entropy* of P is defined by

$$h(P) := \lim_{n \to \infty} \frac{1}{|\,V_n|} H(P_{V_n}).$$

Its existence can be proved by a subadditivity property (for instance cf. Theorem 15.12 in [11]), but it also follows from a Shannon-MacMillan theorem. Föllmer [7] and Thouvenot [28] proved that the specific entropy for ergodic P is

$$E\big[H\big(P_0[\,\cdot\,|\mathcal{P}](\omega)\big)\big],$$

where $P_0[\,\cdot\,|\mathcal{P}]$ is the conditional distribution of ω with respect to the σ-algebra $\mathcal{P}$ generated by all sites which are smaller than 0 with respect to the lexicographical ordering on $\mathbb{Z}^2$. Moreover, they showed that the sequence of rescaled information quantities provided by a stationary P, restricted to the lattice sets V_n $(n \in \mathbb{N})$, converges in $\mathcal{L}^1(P)$:

$$\frac{1}{|\,V_n|}\mathcal{I}(P_{V_n}) \xrightarrow{n \to \infty} E\big[H\big(P_0[\,\cdot\,|\mathcal{P}](\omega)\big)\big|\mathcal{J}\big], \tag{16}$$

where $\mathcal{J}$ is the σ-algebra of all sets which are invariant with respect to the transformations θ_v $(v \in \mathbb{Z}^2)$.

3 A Shannon-MacMillan theorem along lines

The function

$$l_{\lambda,a}(x) = \lambda x + a \qquad (x \in \mathbb{R}) \tag{17}$$

describes the line with slope λ and y-intercept a. Using $[x]$ and $\{x\}$ for the integer and the fractional part of x, respectively, the two-sided sequences

$$\big([l_{\lambda,a}(z)]\big)_{z\in\mathbb{Z}} \quad \text{and} \quad \big(\{l_{\lambda,a}(z)\}\big)_{z\in\mathbb{Z}}$$

are the line's integer and fractional parts at the integer points $z \in \mathbb{Z}$. In the case when $0 \le \lambda \le 1$, the *lattice approximation* of $l_{\lambda,a}$ on an interval $I \subset \mathbb{Z}$ is defined as

$$L_{\lambda,a}(I) := \big\{(z, [l_{\lambda,a}(z)]) \mid z \in I \cap \mathbb{Z}\big\}. \tag{18}$$

In the case when $-1 \le \lambda < 0$, we use the lattice approximation $L_{\lambda,a}(I) := -L_{-\lambda,a}(I)$. If $|\lambda| > 1$, we represent the line as a function of the y-axis with the new slope $\frac{1}{\lambda}$ (or 0 in the case of the y-axis itself) and proceed as before.

We want to identify the specific directional surface entropy $h_\lambda(P)$ of P as the limit of the rescaled entropies along successively increased parts of the lattice approximation to the line, that is, as the limit of the sequence

$$\frac{1}{|L_{\lambda,a}([0,n])|} H\big(P_{L_{\lambda,a}([0,n])}\big) \qquad (n \in \mathbb{N}).$$

The convergence of this sequence will follow from a stronger result, the $\mathcal{L}^1(P)$-convergence of the corresponding sequence of rescaled information, that is,

$$\frac{1}{|L_{\lambda,a}([0,n])|} \mathcal{I}\big(P_{L_{\lambda,a}([0,n])}\big) \qquad (n \in \mathbb{N}).$$

This will be the main result of this section (see Theorem 3.6 for rational λ, and Theorem 3.7 for irrational λ).

In order to make our problem accessible to ergodic theory, we need to create a transformation that follows the stair climbing pattern along the lattice approximation of the line. This will be achieved by keeping track at each step not only of the integer part, but also of the fractional part. Let

$$\tau_\lambda(t) := \{t + \lambda\} \qquad (t \in \mathbb{T})$$

be the translation by λ on the torus $\mathbb{T} := [0,1]$ with its ends identified. Consider the product space $\mathbb{T} \times \Omega$, equipped with the product σ-algebra $\overline{\mathcal{F}}$, and the product measure $\overline{P} = \mu \otimes P$. The transformation

$$\begin{aligned} S_\lambda : \quad \mathbb{T} \times \Omega &\longrightarrow \mathbb{T} \times \Omega \\ (t, \omega) &\longmapsto (\tau_\lambda(t), \theta_{(1,[t+\lambda])}\omega) \end{aligned}$$

follows the desired path, as we shall see in Lemma 3.2.

We first develop several technical tools, that assist in constructing specific entropy along lines. The same tools will find application again when we come to defining specific entropy along nonlinear shapes.

Lemma 3.1. *For $\lambda \in \mathbb{R}, z, \tilde{z} \in \mathbb{Z}, a \in \mathbb{T}$ and $I \subset \mathbb{Z}$ we have the following:*

(i) $\tau_\lambda^z(a) = \{a + \lambda z\} = \{l_{\lambda,a}(z)\}$.

(ii) *The function τ_λ^z has a unique zero, at $t = \{-z\lambda\}$. More explicitly, we get: If z and so λ are both positive or both negative, then τ_λ^z has a unique zero, at $t = 1 - \{z\lambda\}$. If one is negative and the other is positive, then τ_λ^z has a unique zero, at $t = -\{z\lambda\}$. If one of them is zero then τ_λ^z has a unique zero, at $t = 0$.*

(iii) $l_{\lambda,a}(z + \tilde{z}) = l_{\lambda,a}(z) + \lambda\tilde{z}$ *and* $l_{\lambda,a+z}(z) = l_{\lambda,a}(z) + \tilde{z}$.

(iv) $[l_{\lambda,a}(z + \tilde{z})] = [l_{\lambda,a}(z)] + [l_{\tau_\lambda^z(a)}(\tilde{z})]$ *and* $[l_{\lambda,a+z}(z)] = [l_{\lambda,a}(z)] + \tilde{z}$.

(v) $L_{\lambda,a}(I + z) = L_{\lambda,\tau_\lambda^z(a)}(I) + (z, [l_{\lambda,a}(z)])$.

(vi) $L_{\lambda,a+z}(I) = L_{\lambda,a}(I) + (0, z)$.

Proof.

(i) The first equality can be seen easily by induction, and the second follows from (17).

(ii) The case $z = 0$ and the case $\lambda = 0$ are trivial. Let $z \in \mathbb{Z} \setminus \{0\}$. By (i), τ_λ^z has a zero at t if and only if $\{t + z\lambda\} = 0$. The latter is equivalent to $t + z\lambda \in \mathbb{Z}$, which means

$$t + \{z\lambda\} \in \mathbb{Z}. \tag{19}$$

If $z \in \mathbb{N}$ and $\lambda > 0$ then $0 < t + \{z\lambda\} < 2$, since $0 \le t < 1$. Therefore, condition (19) is equivalent to $t + \{z\lambda\} = 1$, that is $t = 1 - \{z\lambda\}$. If z and λ are both negative, $\{z\lambda\}$ is again positive and the argument works as well. If one is negative and the other positive, then $-1 < t + \{z\lambda\} < 1$, and (19) is equivalent to $t + \{z\lambda\} = 0$, so $t = -\{z\lambda\}$.

(iii)

$$l_{\lambda,a}(z + \tilde{z}) = \lambda(z + \tilde{z}) + a = l_{\lambda,a}(z) + \lambda\tilde{z},$$

$$\text{and} \quad l_{\lambda,a+z}(z) = \lambda z + a + \tilde{z} = l_{\lambda,a}(z) + \tilde{z}.$$

(iv) Using (iii) we get

$$\begin{aligned}[l_{\lambda,a}(z + \tilde{z})] &= \big[[l_{\lambda,a}(z)] + \{l_{\lambda,a}(z)\} + \lambda\tilde{z}\big] \\ &= [l_{\lambda,a}(z)] + [\tau_\lambda^z(a) + \lambda\tilde{z}] = [l_{\lambda,a}(z)] + [l_{\tau_\lambda^z(a)}(\tilde{z})].\end{aligned}$$

The second equation follows from the second equation in (iii), because z is an integer.

(v) By (18) and (iv), we obtain

$$\begin{aligned} L_{\lambda,a}(I + z) &= \big\{(\tilde{z}, [l_{\lambda,a}(\tilde{z})])\big|\, \tilde{z} \in I + z\big\} \\ &= \big\{(\tilde{z} + z, [l_{\lambda,a}(\tilde{z} + z)])\big|\, \tilde{z} \in I\big\} \\ &= \big\{(\tilde{z}, [l_{\tau_\lambda^z(a)}(\tilde{z})]) + (z, [l_{\lambda,a}(z)])\big|\, \tilde{z} \in I\big\} \\ &= L_{\tau_\lambda^z(a)}(I) + L_{\lambda,a}(z). \end{aligned}$$

(vi) By (18) and (iv), we get

$$\begin{aligned} L_{\lambda,a+z}(I) &= \big\{(\widetilde{z},[l_{\lambda,a+z}(\widetilde{z})) \,|\, \widetilde{z}\in I\big\} \\ &= \big\{(\widetilde{z},[l_{\lambda,a+z}(\widetilde{z})) + (0,z) \,|\, \widetilde{z}\in I\big\} = L_{\lambda,a}(I)+(0,z). \end{aligned}$$

$\square$

Lemma 3.2. *The iterates of the transformation* $S_\lambda(t,\omega) = (\tau_\lambda(t), \theta_{(1,[t+\lambda])}\omega)$ $(t\in M, \omega\in\Omega)$ *are given by* $S_\lambda^n(a,\omega) = \big(\tau_\lambda^n(a), \theta_{L_a(n)}\omega\big)$ *for all* $n\in\mathbb{N}_0$.

Proof. With $\kappa(a) := (1,[a+\lambda])$, we have $S_\lambda(a,\omega) = (\tau_\lambda(a), \theta_{\kappa(a)}\omega)$, and we obtain

$$S_\lambda^n(a,\omega) = (\tau_\lambda^n(a), \theta_{\kappa_n(a)}\omega) \qquad \text{where} \qquad \kappa_n = \sum_{i=0}^{n-1} \kappa\circ\tau_\lambda^i.$$

It remains to show that $\kappa_n(a) = (n,[l_{\lambda,a}(n)])$ for all $a\in\mathbb{T}$. For the first component this is obvious. For the second component it follows by induction: It is trivial for $n=0$, and by definition (17), and by Lemma 3.1(iv) we have

$$\begin{aligned} \kappa_{n+1}^{(2)}(a) = \kappa_n^{(2)}(a) + \kappa^{(2)}(\tau_\lambda^n(a)) &= [l_{\lambda,a}(n)] + [\tau_\lambda^n(a)+\lambda] \\ &= [l_{\lambda,a}(n)] + [l_{\lambda,\tau_\lambda^n(a)}] = [l_{\lambda,a}(n+1)]. \end{aligned}$$

$\square$

The next lemma plays a key role in proving Shannon-MacMillan theorems by means of the ergodic theorem from [2]. To avoid double indices we skip λ in the notation, and write $L(a,I)$ in place of $L_{\lambda,a}(I)$. Define the functions

$$F_i(t,\omega) := \mathcal{I}\big(P_0[\,\cdot\,|\mathcal{F}_{L(t,[-i,-1])}]\big)(\omega) \quad (t,\omega)\in\mathbb{T}\times\Omega, \tag{20}$$

for $i\in\mathbb{N}$, and

$$F(t,\omega) := \mathcal{I}\big(P_0[\,\cdot\,|\mathcal{F}_{L(t,(-\infty,-1])}]\big)(\omega) \quad (t,\omega)\in\mathbb{T}\times\Omega. \tag{21}$$

Lemma 3.3. *For all* $a\in\mathbb{T}, \omega\in\Omega$, *and* $n\in\mathbb{N}$,

$$\mathcal{I}(P_{L(a,[0,n])})(\omega) = \sum_{i=1}^n F_i\circ S^i(a,\omega).$$

Proof. Conditioning of the measure of the left-hand side leads to

$$P[\omega_{L(a,[0,n])}] = \prod_{i=1}^n P[\omega_{L(a,\{i\})}|\omega_{L(a,[0,i-1])}].$$

By definition of L and shifting we obtain that the latter expression equals

$$\prod_{i=1}^n P[\omega_{(0,0)}|\omega_{L(a,[0,i-1])-L(a,\{i\})}]\circ\theta_{L(a,\{i\})}.$$

Lemma 3.1(iii), with $I=[-i,-1]$, yields

$$\prod_{i=1}^n P[\omega_{(0,0)}|\omega_{L(\tau^i(a),[-i,-1])}]\circ\theta_{L(a,\{i\})},$$

and by the previous lemma we obtain

$$\mathcal{I}(P_{L(a,[0,\dots,n])})(\omega) = \sum_{i=1}^{n} \mathcal{I}\big(P_0[\,\cdot\,|\mathcal{F}_{L(\tau^i(a),[-1,-i])}]\big)(\theta_{L(a,\{i\})}\omega)$$

for all $a \in \mathbb{T}$ and $n \in \mathbb{N}$. □

To apply ergodic theorems to the right-hand side in 3.3, we need to study the asymptotic behavior of the functions $\mathcal{I}\big(P_0[\,\cdot\,|\mathcal{F}_{L(\tau^i(a),[-1,-i])}]\big)$ as i goes to infinity as well as their dependence on the parameter a.

Lemma 3.4. *Assume that, for all $A \in \mathcal{F}_{\mathbb{Z}^2\setminus\{(0,0)\}}$,*

$$P\big[\omega_{(0,0)}\big|A\big] > 0 \quad \textit{for } P\textit{-almost all } \omega \in \Omega. \tag{22}$$

Then for any $t \in \mathbb{T}$, $F_i(t,\cdot)$ converges to $F(t,\cdot)$ P-almost surely and in $\mathcal{L}^1(P)$ as i goes to infinity.

For any $\omega \in \Omega$, the functions $F_i(\cdot,\omega)$ $(i \in \mathbb{N})$ are piecewise constant in t, and the number of discontinuities is finite. If λ is rational then $F(\cdot,\omega)$ is of this type as well.

Let λ be irrational. Assume that P fulfills the strong 0-1 law and that there is a constant $c > 0$, such that, for all $A \in \mathcal{F}_{\mathbb{Z}^2\setminus\{(0,0)\}}$,

$$P\big[\omega_{(0,0)}\big|A\big] > c \quad \textit{for } P\textit{-almost all } \omega \tag{23}$$

Then $F(\cdot\,,\omega)$ is Riemann-integrable in t.

Remark 3.5. *As can be seen in the proof, the set of points where the function $F_i(\cdot,\omega)$ may be discontinuous is given by $\big\{\{\nu\lambda\}\big|\,\nu = -1,\dots,-i\big\}$. When λ is rational the set of potential points of discontinuities of $F_i(\cdot,\omega)$ and $F(\cdot,\omega)$ is*

$$\Big\{\,\{\nu\frac{p}{q}\}\,\Big|\,\nu = -1,\dots,-(q\wedge i)\Big\},$$

where $\frac{p}{q}$ is the unique representation of λ with integers $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ having no common divisor.

Proof of the Lemma. Fix any $t \in \mathbb{T}$. Since the σ-algebras $\big(\mathcal{F}_{L_{\lambda,t}([-1,-i])}\big)_{i\in\mathbb{N}}$ form an increasing family, $\Big(P\big[\omega_{(0,0)}\,\big|\,\omega_{L_{\lambda,t}([-1,-i])}\big]\Big)_{i\in\mathbb{N}}$ is a martingale, so that we obtain by the convergence theorem for martingales,

$$P\big[\omega_{(0,0)}\,\big|\,\omega_{L_{\lambda,t}([-1,-i])}\big] \xrightarrow{i\to\infty} P\big[\omega_{(0,0)}\,\big|\,\omega_{L_{\lambda,t}([-1,-\infty))}\big], \tag{24}$$

P-almost surely and in $\mathcal{L}^1(P)$. By (22) this remains true when we take logarithms on both sides, and this proves the first assertion of the lemma.

To prove the second part of the lemma, fix $\omega \in \Omega$. Chose any $t, \tilde{t} \in \mathbb{T}$, and find a (sufficient) condition under which $F_i(t,\omega) = F_i(\tilde{t},\omega)$. The only influence that the variable t actually has on F_i, is its effect on the set $L_{\lambda,t}([-1,-i])$ of sites we condition on. By (18), $L_{\lambda,t}([-1,-i]) = L_{\lambda,\tilde{t}}([-1,-i])$ if and only if

$$[l_{\lambda,t}(\nu)] = [l_{\lambda,\tilde{t}}(\nu)] \qquad \text{for all} \quad \nu = -1,\dots,-i. \tag{25}$$

Fix $\nu \in \{1,\dots,i\}$. By Lemma 3.1 (i),

$$[l_{\lambda,t}(\nu)] = l_{\lambda,t}(\nu) - \{l_{\lambda,t}(\nu)\} = -\lambda\nu - \tau_\lambda^\nu(t),$$

so that, the equality in (25) is equivalent to

$$t - \tilde{t} = \tau_\lambda^\nu(t) - \tau_\lambda^\nu(\tilde{t}). \tag{26}$$

We know from Lemma 3.1(ii) that τ_λ^ν has a unique zero in $t_{\lambda,\nu} := \{\nu\lambda\}$. The equality (26) is fulfilled if and only if t and $\tilde{t}$ are both either smaller than $t_{\lambda,\nu}$ or larger than $t_{\lambda,\nu}$. Applying this argument to all $\nu \in \{-1, ..., -i\}$, we see that the function $F_i(\,\cdot\,, \omega)$ is piecewise constant, and the set of possible jumps is given by

$$D_i = \big\{\{\nu\lambda\}\,\big|\,\nu = -1, ..., -i\big\}.$$

If λ is rational, these sets are actually independent of i, for i large enough. Use the unique representation $\lambda = \frac{p}{q}$ given in Remark 3.5. We obtain by periodicity of the sequence $(\{-\nu\frac{p}{q}\})_{\nu\in\mathbb{N}}$ that

$$D_i = \big\{\,\{-\nu\frac{p}{q}\}\,\big|\,\nu = -1, ..., -(q \wedge i)\big\}.$$

It remains to prove that, in the case when λ is irrational, $F(\cdot, \omega)$ is still Riemann-integrable in t. It suffices to show that the set of points where the function is discontinuous has Lebesgue measure zero. We will prove that $F(\cdot, \omega)$ is continuous on $\mathbb{T} \setminus D_\infty$, where $D_\infty := \big\{\{\nu\lambda\}\,\big|\,\nu = -1, -2, ...\big\}$. Fix $t_0 \in \mathbb{T} \setminus D_\infty$ and let be $\varepsilon > 0$. We apply a Lemma from [10] (see Lemma 4.7), with $\mathcal{B}_k = L_{\lambda,t}([-k, -1])$, and $\mathcal{B}_k^* = L_{\lambda,t}([-k, -1]) \cup (\mathbb{Z}^2 \setminus V_k)$, where $V_k = [-k, k]^2$. This gives us a $k_0 \in \mathbb{N}$ such that for all $k \geq k_0$ and $t \in \mathbb{T}$ with

$$L_{\lambda,t_0}([-k, -1]) = L_{\lambda,t}([-k, -1]), \tag{27}$$

we obtain

$$\Big|\, P\big[\omega_{(0,0)}\big|\,\omega_{L_{\lambda,t_0}([-k,-1])}\big] - P\big[\omega_{(0,0)}\big|\,\omega_{L_{\lambda,t}([-k,-1])}\big]\,\Big| < \varepsilon. \tag{28}$$

By definition of t_0, $\delta := \min\big\{|\,t_0 - t_{\lambda,\nu}|\,\big|\,\nu = -1, ..., -k\big\}$ is larger than 0. But by Lemma 3.1(i), (27) is true for all $t \in \mathbb{T}$ for which $|t - t_0| < \delta$, which proves the continuity of $P\big[\omega_{(0,0)}\big|\,\omega_{L_{\lambda,t}([-k,-1])}\big]$ in $t = t_0$. Finally, by taking logarithms and using (23), we obtain that the continuity in t_0 yields as well for $F_k(\cdot, \omega)$. $\square$

Our Shannon-MacMillan theorems will now follow by the ergodic theorems which we proved in Section 4 of [2]. We use the abreviation

$$\mathcal{P}_{\lambda,t} := \mathcal{F}_{L_{\lambda,t}((-\infty,-1])} \tag{29}$$

for the *past* σ-algebra occuring in the limits. In the case of a rational slope we apply Corollary 4.1 in [2], and Maker's version of Birkhoff's ergodic theorem (cf. Theorem 7.4 in Chapter 1 of [15] for Maker's theorem, or see Corollary 2.3.1 in [1] for an explicit version of Corollary 4.1 in [2] in the spirit of Maker). The assumptions on the functions $F_{n,i}$ in these theorems a fulfilled by Lemma 3.4.

Theorem 3.6. *Let λ be rational, and $\frac{p}{q}$ its unique representation with integers $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ having no common divisor. Assume that P fulfills condition (22). Then for all $a \in \mathbb{R}$,*

$$\frac{1}{n+1}\mathcal{I}(P_{L_\lambda(a,[0,n])}) \xrightarrow{n\to\infty} \frac{1}{q}\sum_{i=0}^{q-1} E\Big[H\big(P_0[\,\cdot\,|\mathcal{P}_{\frac{p}{q},\frac{ip}{q}}](\omega)\big)\Big|\mathcal{J}\Big]$$

P-almost surely and in $\mathcal{L}^1(P)$, where $\mathcal{J}$ is the σ-algebra of all sets which are invariant with respect to the transformation $\theta_{(q,p)}$.

In particular, if P is ergodic with respect to $\theta_{(q,p)}$ then the limit simplifies to

$$\frac{1}{q}\sum_{i=0}^{q-1} E\Big[H\big(P_0[\,\cdot\,|\mathcal{P}_{\frac{p}{q},\frac{ip}{q}}](\omega)\big)\Big].$$

In the case of an irrational slope we need stronger assumptions to guarantee the Riemann integrability of the functions (20) and (21).

Theorem 3.7. *Let λ be irrational. Assume that P fulfills (22) and the strong 0-1 law. Then for all $a \in \mathbb{R}$,*

$$\frac{1}{n+1}\mathcal{I}(P_{L_\lambda(a,[0,n])}) \xrightarrow{n\to\infty} \int_0^1 E\big[H\big(P_0[\,\cdot\,|\mathcal{P}_{\lambda,t}](\omega)\big)\big]\,dt$$

in $\mathcal{L}^1(P)$.

Proof. Set $\kappa(t) := (1, [t+\lambda])$. Since $\|\kappa_n(t)\| \geq n$, the sequence tends to infinity as n goes to infinity. As a special case of the strong 0-1 law, P fulfills a 0-1 law on the tail field. Since λ is irrational, τ_λ is ergodic. Corollary 2.6 in [2] with $v_1 = (1,0)$ and $v_2 = (0,1)$ implies the ergodicity of the skew product S from Lemma 3.2. By Lemma 3.3,

$$\frac{1}{n+1}\mathcal{I}(P_{L_\lambda(a,[0,n])}) = \sum_{i=1}^{n} F_i \circ S^i(a,\omega).$$

We are going to apply Corollary 4.14 in [2], Maker's version of Birkhoff's ergodic theorem (cf. Theorem 7.4 in Chapter 1 of [15] for Maker's theorem, or see Corollary 2.3.15 in [1] for an explicit version of Corollary 4.14 in [2] in the spirit of Maker). By Lemma 3.4, the Riemann integrability assumptions on the function F, F_i ($i \in \mathbb{N}$) are fulfilled. We have $\|\kappa_i(t) - \kappa_j(t)\| \geq \|i-j\|$, and

$$\lim_{n\to\infty}\frac{1}{n^2}\big|\{1 \leq i,j \leq n \,\big|\, |i-j| \leq m\}\big| = 0$$

for all $m \in \mathbb{N}$. This implies condition the condition on κ we needed for Corollary 4.14 in [2]. $\square$

4 A Shannon-MacMillan theorem along general shapes

Let P be a stationary random field that satiesfies the strong 0-1 law and the condition (23). The goal of this section is a Shannon-MacMillan theorem for a stochastic field along the lattice approximations of the blowups of a curve c, and a nice formula for the limit $h_c(P)$, the *specific entropy* of P *along* c.

In the first subsection, we considered the linear case, that is, when c is a line segment. The proof of the corresponding Shannon-MacMillan theorem (see Theorems 4.3 and Theorem 4.3) is based on the same ideas as the proofs of Theorem 3.6 and Theorem 3.7. The sequence of blowups of a line segment is moving in space, requiring more careful attention than a single line being progressively revealed. The second subsection proves a Shannon-MacMillan theorem along polygons (see Theorem 4.5), and the third subsection develops the result for curves (see Theorem 4.8).

Another purpose of this section is to introduce an alternative way of approximating a line by a subset of the lattice. A contour in statistical mechanics is a set of sites corresponding unambiguously to a chain of bonds. The lattice approximation we used so far is not a contour in this sense. As an alternative, we introduce the *contour approximation.* It is a contour, and it corresponds to the contour boundary defined in (10). It creates a slightly different formula for the specific entropy along a line. Since the specific entropy along a line is the foundation of the specific entropy along polygons and curves, we obtain alternative formulas for those entropies as well. We will call them *specific contour entropies*, and denote them by $\widehat{h}_\lambda$, $\widehat{h}_\pi$, and $\widehat{h}_c$.

Let $c = \left(c^{(1)}, c^{(2)}\right)$ be a piecewise differentiable planar curve parametrized by $t \in [0, T]$. Suppose that c does not go through the origin, and that it hits the y-axis in $t = 0$. The blowups of the curve c are given by

$$B_n c : [0, nT) \longrightarrow \mathbb{R}^2, \qquad B_n c(t) = n\, c\Big(\frac{t}{n}\Big) \qquad (n \in \mathbb{N}). \tag{30}$$

In particular, $B_1 c = c$.

As will be shown later, it is enough to consider a curve c given by the graph of a function ϕ on a segment of one of the axes. Suppose that ϕ is a function on the interval $[x, \widetilde{x}]$ of the x-axis. The case of the y-axis can be treated analogously. More precisely, $x = c^{(1)}(0)$ and $\widetilde{x} = c^{(1)}(T)$. The interval $[x, \widetilde{x}]$ contains a finite number u of integers $z, z+1, \ldots, z+u$. More precisely, $u = [x - \widetilde{x}]$ or $u = [x - \widetilde{x}] - 1$. In the same way, the blowups $B_n c$ can be represented as graphs of functions ϕ_n of intervals $[x_n, \widetilde{x}_n]$. More explicitly, we obtain by (30)

$$\begin{aligned} x_n &= B_n c^{(1)}(0) = nx = nc^{(1)}(0) \\ \widetilde{x}_n &= B_n c^{(1)}(nT) = n\widetilde{x} = nc^{(1)}(T). \end{aligned}$$

Again, the interval $[x_n, \widetilde{x}_n]$ contains a finite number u_n of integers $z_n, z_n + 1, \ldots, z_n + u_n$, where

$$u_n = [n(\widetilde{x} - x)] \qquad \text{or} \qquad u_n = [n(\widetilde{x} - x)] - 1. \tag{31}$$

In particular, the sequence $(u_n)_{n\in\mathbb{N}}$ goes to infinity.

4.1 Line segments

As a first step, we consider the linear case, that is, when c is a line segment. Then we have for the corresponding functions ϕ_n

$$\begin{aligned} &\phi_n(x) = \lambda x + a_n \quad \text{for} \quad x \in [nc^{(1)}(0), nc^{(1)}(T)], \\ &\qquad\qquad \text{where} \quad \lambda = \frac{c^{(2)}(T) - c^{(2)}(0)}{c^{(1)}(T) - c^{(1)}(0)} \quad \text{and} \quad a_n = n(c^{(2)}(0) - \lambda c^{(1)}(0)). \end{aligned}$$

Assume $0 \le \lambda \le 1$. As explained at the beginning of Section 3, the other cases can be reduced to this case. At first sight it seems we could just apply the results for the specific entropy along a line from the last section. But the blowups of the line segments move in space, which has the following consequences.

(i) There is a sequence $(a_n)_{n\in\mathbb{N}}$ instead of a constant a;

(ii) the sequence is real-valued, as opposed to the constant having values in $\mathbb{T}$;

(iii) the positions of the lattice points in each step, as well as their number, are more difficult to control. In the case of the line we simply looked at approximating points with x-values between 0 and n. Now, we have to deal with lattice points whose x-values lie in the interval $z_n, \ldots, z_n + u_n$.

The last problem forces us to apply, at each step n, an additional shift to ω which brings the line segment close to the origin. These shifts do not affect the $\mathcal{L}^1(P)$-convergence since the limit is shift invariant. The number of points in the nth step is given by $(u_n)_{n\in\mathbb{N}}$, instead of simply $n+1$ as in the last section, but this is irrelevant as long as the sequence goes to infinity. The second problem requires another shift in each step n. The first point is the most delicate. It is here that we really need the convergence of the ergodic averages in *all* t, which was discussed in Section 4 of [2].

We begin with a precise description of the contour approximation. Unless $\lambda = 0$, the lattice approximation $L_{\lambda,a}$ used in the last section does not define a unique contour. The last site before a new step is catercornered from the first site of the step, so not connected by a bond. The *contour approximation* is obtained by adding, at each new step, the site which is one unit below it. A new step begins in $i+1$ if

$$\tau_\lambda^i(\{a\}) \geq 1 - \lambda, \tag{32}$$

and the site we will add in this case is $L_{\lambda,a}(I) - (0,1)$. For $I \subset \mathbb{Z}$ let

$$\begin{aligned}\widehat{L}_{\lambda,a}(I) :=& \big\{ L_{\lambda,a}(i) \,\big|\, i \in I \big\} \cup \\ & \big\{ L_{\lambda,a}(i) - (0,1) \,\big|\, i \in I \ \wedge\ \{\lambda(i) + a\} + \lambda \geq 1 \big\}\end{aligned} \tag{33}$$

be the *contour approximation* of the line segment $l_{\lambda,a}(I)$. Lemma 3.1 and Lemma 3.4 translate immediately to $\widehat{L}$.

Unfortunately, the new sites are not in the orbit of the skew product, and are therefore overlooked by the ergodic averages we want to use to prove the convergence. We will get around this difficulty by taking ergodic averages of two functions: one function is evaluated along the orbit of the skewproduct; the other one is taken at all sites which are one unit to the left of the orbit, but vanishes as long as no new step is reached.

Set $a := a_1$. Using the notation for $[x, \widetilde{x}]$ described above (31), define for $n \in \mathbb{N}$,

$$L_n(a) := L_{\lambda,a_n}(z_n, z_n + 1, \ldots, z_n + u_n), \tag{34}$$

and

$$\widehat{L}_n(a) := \widehat{L}_{\lambda,a_n}(z_n, z_n + 1, \ldots, z_n + u_n). \tag{35}$$

The total number of sites in $L_n(a)$ is $u_n + 1$, and in $\widehat{L}_n(a)$ it is

$$|\widehat{L}_n(a)| = u_n + 1 + \big|\big\{ i \in \mathbb{N}_0 \,\big|\, 0 \leq i \leq u_n,\ \tau_\lambda^{z_n + i - 1}(\{a\}) \geq 1 - \lambda \big\}\big|. \tag{36}$$

We will study the renormalized information functions

$$\frac{1}{|L_n(a)|}\mathcal{I}(P_{L_n(a)})(\omega) \qquad \text{and} \qquad \frac{1}{|\widehat{L}_n(a)|}\mathcal{I}(P_{L_n(a)})(\omega) \tag{37}$$

along the lattice approximation and the contour approximation of the line segment at $\omega \in \Omega$. Our goal is to prove their $\mathcal{L}^1(P)$ convergence.

The proof for the lattice approximation is easier version of the proof for the contour approximation. We will carry out the details of the proof only in the latter case. To transform (37) into some sort of ergodic average we first condition on sucessively smaller parts of $\widehat{L}_n(a)$. A new step begins at i if $\tau_\lambda^{z_n+i-1}(\{a\}) \geq 1-\lambda$. In this case, $\widehat{L}_{\lambda,a}(z_n+i)$ consists of two sites,

$$L_{\lambda,a}(z_n+i)-(0,1) \qquad \text{and} \qquad L_{\lambda,a}(z_n+i). \tag{38}$$

For conditioning, we have to treat these two sites separately. For the first one, the set we have to condition on is just $\widehat{L}_{\lambda,a_n}(z_n+i-1,\dots,z_n)$, but for the second one we need to add the site $L_{\lambda,a_n}(z_n+i)-(0,1)$. For $z,\tilde{z}\in\mathbb{Z}$ with $\tilde{z}\geq z$ define the sets

$$\begin{aligned}&\widehat{L}^\sharp_{\lambda,a}(\tilde{z},\dots,z) := \\ &\begin{cases}\widehat{L}_{\lambda,a}(\tilde{z}-1,\dots,z)\cup\big(L_{\lambda,a}(\tilde{z})-(0,1)\big) & \text{if } \tau_\lambda^{z-1}(\{a\})\geq 1-\lambda,\\ \widehat{L}_{\lambda,a}(\tilde{z}-1,\dots,z) & \text{otherwise.}\end{cases}\end{aligned} \tag{39}$$

We will need the following equalities:

Lemma 4.1. *For $n,i\in\mathbb{N}$, $z\in\mathbb{Z}$, and $a\in\mathbb{R}$ we have*

(i) $L_{\lambda,a}(z+i)-L_{\lambda,a}(z)=L_{\lambda,\tau_\lambda^z(\{a\})}(i)$.

(ii) $\widehat{L}_{\lambda,a}(z+i-1,\dots,z)-L_{\lambda,a}(z+i)=\widehat{L}_{\lambda,\tau_\lambda^{z+i}(\{a\})}(-1,\dots,-i)$.

(iii) $\widehat{L}^\sharp_{\lambda,a}(z+i-1,\dots,z)-L_{\lambda,a}(z+i)=\widehat{L}^\sharp_{\lambda,\tau_\lambda^{z+i}(\{a\})}(-1,\dots,-i)$.

Proof.

(i) Applying the second equation in Lemma 3.1(v), with $a=\{a\}$ and $z=[a]$, and then the first equation, with $\tilde{z}=i$, and finally the definition (34), we obtain

$$\begin{aligned}L_{\lambda,a}(z+i) &= (z+i,[l_{\lambda,\{a\}}(z+i)])+(0,[a])\\ &= (z+i,[l_{\lambda,\{a\}}(z)]+[l_{\lambda,\tau_\lambda^z(\{a\})}(i)])+(0,[a])\\ &= \big(z,[l_{\lambda,a}(z)]\big)+\big(i,[l_{\lambda,\tau_\lambda^z(\{a\})}(i)]\big)\\ &= L_{\lambda,a}(z)+L_{\lambda,\tau_\lambda^z(\{a\})}(i).\end{aligned}$$

(ii) Using Lemma 3.1(vi), with $a=\{a\}$ and $z=[a]$, and then Lemma (v), with $I=\{-1,\dots-i\}$ and $z=z+i$, and finally the first equation in (iv) with $a=\{a\}$ and $z=[a]$, we see that

$$\begin{aligned}\widehat{L}_{\lambda,a}(z+i-1,\dots,z) &= \widehat{L}_{\lambda,\{a\}}(z+i-1,\dots,z)+(0,[a])\\ &= \widehat{L}_{\lambda,\tau_\lambda^{z+i}(\{a\})}(-1,\dots,-i)+L_{\lambda,\{a\}}(z+i)+(0,[a])\\ &= \widehat{L}_{\lambda,\tau_\lambda^{z+i}(\{a\})}(-1,\dots,-i)+L_{\lambda,a}(z+i).\end{aligned}$$

(iii) If $\tau_\lambda^{z+i-1} < 1-\lambda$ $\widehat{L}^\sharp_{\lambda,a}(z+i-1,\dots,z)$ coincides with $\widehat{L}_{\lambda,a}(z+i-1,\dots,z)$. Otherwise we obtain by (i) and (ii),

$$\begin{aligned}
&\widehat{L}^\sharp_{\lambda,a}(z+i-1,\dots,z) - L_{\lambda,a}(z+i)\\
&= \big(\widehat{L}_{\lambda,a}(z+i-1,\dots,z) \cup (L_{\lambda,a}(z+i)-(0,1))\big) - L_{\lambda,a}(z+i)\\
&= \widehat{L}_{\lambda,\tau_\lambda^{z+i}(\{a\})}(-1,\dots,-i) \cup \big(L_{\lambda,a}(z+i)-(0,1)-L_{\lambda,a}(z+i)\big)\\
&= \widehat{L}_{\lambda,\tau_\lambda^{z+i}(\{a\})}(-1,\dots,-i) \cup \{-(0,1)\},
\end{aligned}$$

and applying Definition 39, with $\widetilde{z}=0$, leads to the desired result. □

We calculate the information in (37) by conditioning site by site along $\widehat{L}_n(a)$. To make reading easier we use $\omega(i)$ instead of ω_i. We see that

$$\begin{aligned}
&-\mathcal{I}(P_{L_n(a)})(\omega) \qquad (40)\\
&= \sum_{i=0}^{u_n}\Big(\log P\Big[\omega\big(L_{\lambda,a_n}(z_n+i)\big)\,\Big|\,\omega\big(\widehat{L}^\sharp_{\lambda,a_n}(z_n+i-1,\dots,z_n)\Big]\\
&\qquad + 1_{\{\tau_\lambda^{z_n+i-1}(\{a_n\})\geq 1-\lambda\}}\times\\
&\qquad\quad \log P\Big[\omega\big(L_{\lambda,a_n}(z_n+i)-(0,1)\big)\,\Big|\,\omega\big(\widehat{L}_{\lambda,a_n}(z_n+i-1,\dots,z_n)\big)\Big]\Big)
\end{aligned}$$

Shifting ω to the origin leads to

$$\begin{aligned}
&\sum_{i=0}^{u_n}\log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}^\sharp_{\lambda,a_n}(z_n+i-1,\dots,z_n)-L_{\lambda,a_n}(z_n+i)\big)\Big]\circ\theta_{L_{\lambda,a_n}(z_n+i)}\\
&\quad+\sum_{i=0}^{u_n}1_{\{\tau_\lambda^{z_n+i-1}(\{a_n\})\geq 1-\lambda\}}\times \qquad (41)\\
&\quad\log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}_{\lambda,a_n}(z_n+i-1,\dots,z_n)-(L_{\lambda,a_n}(z_n+i)-(0,1))\big)\Big]\\
&\qquad\qquad\qquad\circ\theta_{L_{\lambda,a_n}(z_n+i)-(0,1)}.
\end{aligned}$$

We examine the two sums separately. Lemma 4.1(iii) transforms the summands of the first sum to

$$\log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}^\sharp_{\lambda,\tau_\lambda^{z_n+i}(\{a_n\})}(-1,\dots,-i)\big)\Big]\circ\theta_{L_{\lambda,a_n}(z_n+i)}. \qquad (42)$$

By Lemma 4.1(i), this equals

$$\log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}^\sharp_{\lambda,\tau_\lambda^{z_n+i}(\{a_n\})}(-1,\dots,-i)\big)\Big]\circ\theta_{L_{\lambda,\tau_\lambda^{z_n}(\{a_n\})}(i)}\circ\theta_{L_{\lambda,a_n}(z_n)}. \qquad (43)$$

Defining the family of functions

$$F_i(t,\omega) = -\log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}^\sharp_{\lambda,t}(-1,\dots,-i)\big)\Big] \qquad ((t,\omega)\in\mathbb{T}\times\Omega) \qquad (44)$$

allows us to rewrite this as

$$F_i\Big(\tau_\lambda^i(\tau_\lambda^{z_n}(\{a_n\})),\theta_{L_{\lambda,\tau_\lambda^{z_n}(\{a_n\})}(i)}\circ\theta_{L_{\lambda,a_n}(z_n)}\,\omega\Big), \qquad (45)$$

which is, by Lemma 3.2, the same as

$$F_i \circ S^i\big(\tau_\lambda^{z_n}(\{a_n\}), \theta_{L_{\lambda,a_n}(z_n)}\,\omega\big). \tag{46}$$

For the second sum in (41) we can proceed similarly. By Lemma 4.1(ii), we get for their addends

$$\begin{aligned}&1_{\{\tau_\lambda^{z_n+i-1}(a_n)\geq 1-\lambda\}} \times \\ &\log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}_{\lambda,\tau_\lambda^{z_n+i}(\{a_n\})}(-1,\ldots,-i)+(0,1)\big)\Big]\circ\theta_{L_{\lambda,a_n}(z_n+i)-(0,1)}.\end{aligned}$$

Using Lemma 4.1(i) again leads to

$$\begin{aligned}1_{\{\tau_\lambda^{z_n+i-1}(a_n)\geq 1-\lambda\}}\,\log P\Big[\omega(0,0)\,\big|\,\omega\big(\widehat{L}_{\lambda,\tau_\lambda^{z_n+i}(\{a_n\})}(-1,\ldots,-i)+(0,1)\big)\Big]&\\ \circ\,\theta_{L_{\lambda,\tau_\lambda^{z_n}(\{a_n\})}(i)}\circ\theta_{L_{\lambda,a_n}(z_n)-(0,1)}.&\end{aligned} \tag{47}$$

Defining the family of functions

$$G_i(t,\omega) = -\log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}_{\lambda,t}(-1,...,-i)+(0,1)\big)\Big] \quad ((t,\omega)\in\mathbb{T}\times\Omega) \tag{48}$$

allows us to write the second addend in (41) in the form

$$1_{\{\tau_\lambda^{i-1}(t)\geq 1-\lambda\}}\,G_i\Big(\tau_\lambda^i\big(\tau_\lambda^{z_n}(\{a_n\})\big), \theta_{L_{\lambda,\tau_\lambda^{z_n}(\{a_n\})}(i)}\circ\theta_{L_{\lambda,a_n}(z_n)-(0,1)}\,\omega\Big), \tag{49}$$

which is, by Lemma 3.2, the same as

$$1_{\{\tau_\lambda^{i-1}(t)\geq 1-\lambda\}}\,G_i\circ S^i\big(\tau_\lambda^{z_n}(\{a_n\}), \theta_{L_{\lambda,a_n}(z_n)-(0,1)}\,\omega\big). \tag{50}$$

Putting (45) and (49) in (40), and renormalizing,

$$\begin{aligned}\frac{1}{\widehat{L}_n(a)}\mathcal{I}(P_{L_n(a)})(\omega) &= \frac{1}{\widehat{L}_n(a)}\sum_{i=0}^{u_n} F_i\circ S^i\big(\tau_\lambda^{z_n}(\{a_n\}), \theta_{L_{\lambda,a_n}(z_n)}\,\omega\big) \\ &+ \frac{1}{\widehat{L}_n(a)}\sum_{i=0}^{u_n} 1_{\{\tau_\lambda^{i-1}(t)\geq 1-\lambda\}} G_i\circ S^i\big(\tau_\lambda^{z_n}(\{a_n\}), \theta_{L_{\lambda,a_n}(z_n)-(0,1)}\,\omega\big).\end{aligned} \tag{51}$$

Finally, we have to divide by the number $u_n + 1 + k_{u_n}$ of summands, in place of the original rescaling factor $\widehat{L}_n(a)$. The next lemma computes the asymptotic contribution of the fraction resulting from the change of renormalization.

Lemma 4.2. *For all* $t\in\mathbb{T}$,

$$\frac{u_n+1}{|\widehat{L}_n(t)|} \xrightarrow{n\to\infty} \frac{1}{1+\lambda}.$$

Proof. Let

$$k_{u_n} := \big|\{i\in\mathbb{N}_0\,|\,1\leq i\leq u_n,\ \tau_\lambda^{z_n+i}(\{a_n\})\geq 1-\lambda\}\big|$$

denote the number of steps between 0 and u_n. Since (36) implies

$$\frac{u_n+1}{|\widehat{L}_n(t)|} = \frac{u_n+1}{u_n+1+k_{u_n}(t)} = \frac{1}{1+\frac{k_{u_n}(t)}{u_n+1}},$$

it remains to show that

$$\frac{k_{u_n}(t)}{u_n+1} \xrightarrow{n\to\infty} \lambda.$$

We have the estimate $k_{u_n} \le [L_{\lambda,a}(z_n+u_n)-L_{\lambda,a}(z_n)]+1 = [\lambda u_n]+1$, and, analogously, $k_{u_n} \ge [\lambda u_n]$. The convergence follows from

$$\frac{[\lambda u_n]+1}{u_n+1} = \lambda\frac{u_n}{u_n+1} - \frac{\{\lambda u_n\}}{u_n+1} + \frac{1}{u_n+1}$$

and

$$\frac{[\lambda u_n]}{u_n+1} = \lambda\frac{u_n}{u_n+1} + \frac{\{\lambda u_n\}}{u_n+1}.$$

□

Now, the sums are in the form required by our ergodic theorems for skew products. To prove the convergence, we have to distinguish the case when λ is rational from the cases when it is irrational, because this determines whether τ_λ is periodic or uniquely ergodic. However, since Lemma 3.4 can be carried over from L to $\widehat{L}$, the assumptions on the functions F_i and G_i were already proven in the last section. In particular,

$$F_i(t,\omega) \xrightarrow{i\to\infty} -\log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}^\sharp_{\lambda,t}(-\mathbb{N})\big)\Big] \tag{52}$$

and

$$G_i(t,\omega) \xrightarrow{i\to\infty} -\log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}_{\lambda,t}(-\mathbb{N})+(0,1)\big)\Big]. \tag{53}$$

For rational λ, Theorem 4.3 follows from Corollary 4.1 in [2], in the way we explained it in the proof of Theorem 3.6. In the case when λ is irrational we need the ergodicity of the skew product (see Theorem 2.5 in [2]), and we apply Corollary 4.14 from [2] as it was done in the proof of Theorem 3.7.

Before we state our theorem, we introduce a shorter notation for *past* σ-algebras arising from the conditionings in (52) and (53).

$$\mathcal{P}^\sharp_{\lambda,t} := \mathcal{F}\big(\widehat{L}^\sharp_{\lambda,t}(-\mathbb{N})\big) \tag{54}$$

corresponds to the condition used for upper sites, and

$$\mathcal{P}^\flat_{\lambda,t} := \mathcal{F}\big(\widehat{L}_{\lambda,t}(-\mathbb{N})+(0,1)\big) \tag{55}$$

correspond to the conditioning needed for the lower sites. Note that there is a step between $\widehat{L}(-1)$ and $\widehat{L}(0)$ if and only if $t<\lambda$, and that, by (39),

$$\mathcal{P}^\sharp_{\lambda,t} = \begin{cases} \mathcal{F}\Big(\big\{(-i,[t-\lambda i])\,\big|\,i\in\mathbb{N}\big\}\cup\{(0,-1)\}\Big) & \text{if } t<\lambda, \\ \mathcal{F}\Big(\big\{(-i,[t-\lambda i])\,\big|\,i\in\mathbb{N}\big\}\Big) & \text{if } t\ge\lambda. \end{cases}$$

Theorem 4.3. *As n goes to infinity,*

$$\frac{1}{|\widehat{L}_n(a)|}\mathcal{I}(P_{L_n(a)})$$

converges in $\mathcal{L}^1(P)$ and uniformly in $a\in\mathbb{R}$ to the specific contour entropy along a line *with slope λ*

$$\widehat{h}_\lambda(P) := \frac{1}{1+\lambda}\bigg(\int_0^1 E\big[H\big(P_0[\cdot\,|\,\mathcal{P}^\sharp_{\lambda,t}](\omega)\big)\big]\,dt + \int_{1-\lambda}^1 E\big[H\big(P_0[\cdot\,|\,\mathcal{P}^\flat_{\lambda,t}](\omega)\big)\big]\,dt\bigg).$$

Replacing the contour approximation by a lattice approximation, we obtain a similar theorem. The proof is an easier version of the proof of Theorem 4.3. We simply set $G_i = 0$, for all $i \in \mathbb{N}$, and switch to the lattice approximation in the conditionings. The limit is just the entropy $h_\lambda(P)$ of P along a line with slope λ. Recall that

$$h_\lambda(P) = \int_0^1 E\big[H\big(P_0[\,\cdot\,|\,\mathcal{P}_{\lambda,t}](\omega)\big)\big]\,dt.$$

Theorem 4.4. *As n goes to infinity,*

$$\lim_{n\to\infty} \frac{1}{|L_n(a)|}\,\mathcal{I}(P_{L_n(a)}) = h_\lambda(P)$$

in $\mathcal{L}^1(P)$, and uniformly in $a \in \mathbb{R}$.

4.2 Polygons

The next step is to define the entropy along a polygon, that is a piecewise linear curve $\pi : [0,T] \to \mathbb{R}^2$ without self-intersections. Without loss of generality, we can assume that $\pi([0,T])$ does not contain the origin. Now, the blowups of π are well-defined. We want to track the construction back to entropies along the lines corresponding to the edges of the polygon. To make use of the construction given above we describe the polygons in the same notation. Let R be the number of edges of π. We can find slopes $\lambda^{(r)} \in (-1,1]$, constants $t^{(r)} \in \mathbb{R}$, and intervals $I^{(r)}$ of the x- or the y-axis such that

$$\pi([0,T]) = \bigcup_{r=1}^{R} l_{\lambda^{(r)},t^{(r)}}(I^{(r)}), \tag{56}$$

with $l_{\lambda,t}$ as defined in (17) as a function of the x- of the y-axis.

Proceeding the same way for the blowups $B_n\pi$ ($n \in \mathbb{N}$) (see 30), we choose $t_n^{(r)} \in \mathbb{R}$ and $I_n^{(r)} \subset \mathbb{R}$, such that

$$B_n\pi([0,T]) = \bigcup_{r=1}^{R} l_{\lambda^{(r)},t_n^{(r)}}(I_n^{(r)}),$$

The contour approximations (33) of the edges of $B_n\pi$ combine to a contour approximation

$$\widehat{L}_n^\pi := \bigcup_{r=1}^{R} \widehat{L}_{\lambda^{(r)},t_n^{(r)}} \tag{57}$$

of $B_n\pi$. And analogously, the lattice approximations of the edges combine to a lattice approximation

$$L_n^\pi := \bigcup_{r=1}^{R} L_{\lambda^{(r)},t_n^{(r)}} \tag{58}$$

of $B_n\pi$.

Theorem 4.5. *For the lattice approximation we have*

$$\frac{1}{\text{length } L_n^\pi}\,\mathcal{I}(P_{L_n^\pi}) \xrightarrow{n\to\infty} \frac{1}{\text{length } \pi} \sum_{r=1}^{R} \text{length } \pi^{(r)}\, h_{\lambda^{(r)}}(P) \tag{59}$$

in $\mathcal{L}^1(P)$, *and for the contour approximation we obtain*

$$\frac{1}{\text{length}\,\widehat{L}_n^\pi}\,\mathcal{I}(P_{L_n^\pi}) \xrightarrow{n\to\infty} \frac{1}{\text{length}\,\pi}\sum_{r=1}^{R}\text{length}\,\pi^{(r)}\,\widehat{h}_{\lambda^{(r)}}(P) \tag{60}$$

in $\mathcal{L}^1(P)$.

Remark 4.6. *The limits in (59) and* (60) *can be written as*

$$\frac{1}{\text{length}\,\pi}\int_0^T h_{\pi'(t)}(P)\,dt, \qquad \textit{respectively} \qquad \frac{1}{\text{length}\,\pi}\int_0^T \widehat{h}_{\pi'(t)}(P)\,dt,$$

where $\pi'(t)$ *denotes the right derivative of* π.

The proof of the theorem requires a lemma from [10], which we recall for the reader's convenience.

Lemma 4.7. (Föllmer and Ort) *Consider* σ*-algebras* $\mathcal{B}_i \subseteq \mathcal{B}_i^*$ $(i \in \mathbb{N})$ *increasing to* $\mathcal{B}_\infty$, *respectively decreasing to* $\mathcal{B}_\infty^*$, *and assume that*

$$\mathcal{B}_\infty = \mathcal{B}_\infty^* \quad \textit{mod } P. \tag{61}$$

Then for any $\phi \in \mathcal{L}^1(P)$,

$$\lim_{i\to\infty}\sup_{\mathcal{B}_i\subseteq\mathcal{C}_i\subseteq\mathcal{B}_i^*}\big\|E[\phi\,|\,\mathcal{C}_k] - E[\phi\,|\,\mathcal{B}_\infty]\big\|_{\mathcal{L}^1(P)} = 0. \tag{62}$$

Proof. Let $\|\cdot\|$ denote the $\mathcal{L}^1(P)$-norm. Put $\phi_i = E[\phi\,|\,\mathcal{B}_i]$ and $\phi_i^* = E[\phi\,|\,\mathcal{B}_i^*]$ for $i = 1,\ldots,\infty$. If $\mathcal{B}_i \subseteq \mathcal{C}_i \subseteq \mathcal{B}_i^*$ then, by projection and contraction,

$$\begin{aligned}\|\,\phi_\infty - E[\phi\,|\,\mathcal{C}_i]\,\| &= \|\,\phi_\infty - E[\phi_i^*\,|\,\mathcal{C}_i]\,\|\\ &\le \|\,\phi_\infty - \phi_i\,\| + \|\,\phi_i - E[\phi_\infty^*\,|\,\mathcal{C}_i]\,\| + \|\,E[\phi_\infty^*\,|\,\mathcal{C}_i] - E[\phi_i^*\,|\,\mathcal{C}_i]\,\|\\ &\le \|\,\phi_\infty - \phi_i\,\| + \|\,\phi_i - \phi_i^*\,\| + \|\,\phi_\infty^* - \phi_i^*\,\|,\end{aligned}$$

and this converges to 0 by forward and backward martingale convergence, since $\phi_\infty = \phi_\infty^*$ by assumption. □

Proof of the theorem. The proof is carried out for the contour approximation. The result for the lattice approximation then follows as Theorem 4.4 from Theorem 4.3. To make the reading easier we use $\omega(i)$ for ω_i and we define the sets

$$E_n^{(r)} := \widehat{L}_{\lambda^{(r)},t_n^{(r)}}(I_n^{(r)}) \qquad (r \in \{1,\ldots,R\}). \tag{63}$$

By conditioning, we obtain for the information function

$$\mathcal{I}(P_{L_n^\pi})(\omega) = \sum_{r=1}^{R}\log P\big[\omega(E_n^{(r)})\,\big|\,\omega(E_n^{(r-1)},\ldots,E_n^{(1)})\big]. \tag{64}$$

Fix $r \in \{1,\ldots,R\}$. To simplify notation we omit the index r when there is no risk of confusion, for example $\lambda := \lambda^{(r)}, t_n := t_n^{(r)}, \widehat{L} := \widehat{L}^{(r)}$ and $\widehat{L}^\sharp := {L^{(r)}}^\sharp$. We also use the short form $\breve{E}_n := E_n^{(r-1)} \cup\cdots\cup E_n^{(1)}$ for the contour approximations of the edges of the polygon which come before

$E_n^{(r)}$ in our enumeration. We will condition sucessively on the elements of $E_n^{(r)}$. Denoting the integers in $I_n^{(r)}$ by $z_n, z_n+1, \ldots, z_n+u_n$ as in (31), and using the modification $\widehat{L}^\sharp$ as defined in (39), yields for the rth addend in (64)

$$\begin{aligned}
&\log P\Big[\omega(E_n^{(r)})\,\Big|\,\omega(\breve{E}_n)\Big]\\
&=\sum_{i=0}^{u_n}\log P\Big[\omega\big(\widehat{L}^\sharp_{\lambda,t_n}(z_n+i)\big)\,\Big|\,\omega\big(\widehat{L}_{\lambda,t_n}(z_n+i-1,\ldots,z_n),\breve{E}_n\big)\Big]\\
&\quad+1_{\left\{\tau_\lambda^{z_n+i-1}(\{t_n\})\geq 1-\lambda\right\}}\times\\
&\qquad\log P\Big[\omega\big(L_{\lambda,t_n}(z_n+i)-(0,1)\big)\,\Big|\,\omega\big(\widehat{L}_{\lambda,t_n}(z_n+i-1,\ldots,z_n),\breve{E}_n\big)\Big]
\end{aligned}$$

Now, we use the same type of calculation as in the last section, when we proved a Shannon-MacMillan theorem along a line. Define

$$v_{ni}(t) := L_{\lambda,t_n}(z_n+i) \qquad \text{and} \qquad v^\flat_{ni}(t) := L_{\lambda,t_n}(z_n+i)-(0,1). \tag{65}$$

Shifting by $v_{ni}(t)$, and $v^\flat_{ni}(t)$ respectively, yields

$$\begin{aligned}
&\sum_{i=0}^{u_n}\log P\Big[\omega(0,0)\Big|\omega\big(\widehat{L}^\sharp_{\lambda,t_n}(z_n+i-1,\ldots,z_n)-v_{ni},\breve{E}_n-v_{ni}(t)\big)\Big]\\
&\quad+1_{\left\{\tau_\lambda^{z_n+i-1}(\{t_n\})\geq 1-\lambda\right\}}\times\\
&\qquad\log P\Big[\omega\big(0,0\big)\,\Big|\,\omega\big(\widehat{L}_{\lambda,t_n}(z_n+i-1,\ldots,z_n)-v^\flat_{ni}(t),\breve{E}_n-v^\flat_{ni}(t)\big)\Big]\circ\theta_{v^\flat_{ni}(t)}.
\end{aligned}$$

We know from Lemma 4.1(iii) that

$$\widehat{L}^\sharp_{\lambda,t_n}(z_n+i-1,\ldots,z_n)-v_{ni}(t)=\widehat{L}^\sharp_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\ldots,-i),$$

from Lemma 4.1(ii) that

$$\begin{aligned}
\widehat{L}^\sharp_{\lambda,t_n}&(z_n+i-1,\ldots,z_n)-v^\flat_{ni}(t)\\
&=\widehat{L}^\sharp_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\ldots,-i)+(0,1),
\end{aligned}$$

and from Lemma 4.1(i) that

$$v_{ni}(t)=L_{\lambda,t_n}(z_n)+L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i),$$

and that

$$v^\flat_{ni}(t)=L_{\lambda,t_n}(z_n)+L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)-(0,1).$$

We obtain

$$\begin{aligned}
&\sum_{i=0}^{u_n}\log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}^\sharp_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\ldots,-i),\breve{E}_n-v_{ni}(t)\big)\Big]\\
&\qquad\qquad\qquad\qquad\circ\,\theta_{L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)}\circ\theta_{L_{\lambda,t_n}(z_n)}\\
&+1_{\left\{\tau_\lambda^{z_n+i-1}(\{t_n\})\geq 1-\lambda\right\}}\times\\
&\quad\log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\ldots,-i)+(0,1),\breve{E}_n-v^\flat_{ni}(t)\big)\Big]\\
&\qquad\qquad\qquad\qquad\circ\,\theta_{L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)-(0,1)}\circ\theta_{L_{\lambda,t_n}(z_n)}.
\end{aligned}\tag{66}$$

These are the same addends as in (43) and (47) except that there are the additional conditionings on the sites $\breve{E}_n - v_{ni}(t)$, and $\breve{E}_n - v^\sharp_{ni}(t)$, respectively. We want to show that these conditions disappear asymptotically. The argument will be given in detail for the first summand; the second proof is similar.

Let α be the minimum angle between any neighboring edges of the polygon π and let d_n be the minimum distance between an edge of the nth blowup $B_n\pi$ of the polygon and any of its nonneighboring edges. Also, let H_n be the hexagon defined as follows: H_n is symmetric around $E_n^{(r)}$, two sides are parallel to $E_n^{(r)}$ at a distance $d_n/2$. The other sides reach from the endpoint of the first two to the endpoints of $E_n^{(r)}$, and they intersect at an angle α. Observe that

$$\breve{E}_n \subset \mathbb{Z}^2 \setminus H_n, \tag{67}$$

and therefore $\breve{E}_n - v_{ni}(t) \subset \mathbb{Z}^2 \setminus \big(H_n - v_{ni}(t)\big)$.

Define the σ-algebras

$$\begin{aligned}
\mathcal{B}_i(t) &:= \mathcal{F}\big(\widehat{L}_{\lambda,\tau^i_\lambda(\{t\})}(-1,\ldots,-i)\big)\\
\mathcal{B}_\infty(t) &:= \mathcal{F}\big(\widehat{L}_{\lambda,\tau^i_\lambda(\{t\})}(-1,-2,\ldots)\big)\\
\mathcal{B}^*_i(t) &:= \mathcal{F}\big(\widehat{L}_{\lambda,\tau^i_\lambda(\{t\})}(-1,\ldots,-i) \cup \mathbb{Z}^2 \setminus \big(H_n - v_{ni}(t)\big)\big).
\end{aligned}$$

For any $t \in \mathbb{R}$, the sequence $\big(\mathcal{B}_i(t)\big)_{i\in\mathbb{N}}$ is increasing to $\mathcal{B}_\infty(t)$, and the sequence $\big(\mathcal{B}^*_i(t)\big)_{i\in\mathbb{N}}$ is decreasing to

$$\mathcal{B}^*_\infty(t) := \bigcap_{i\in\mathbb{N}} \mathcal{B}^*_i(t).$$

By the strong 0-1 law, $\mathcal{B}^*_\infty(t) = \mathcal{B}_\infty(t)$ mod P. By Lemma 4.7, for any $t \in \mathbb{T}$,

$$\begin{aligned}
\lim_{i\to\infty} \Big\| \log P\Big[\omega(0,0)\,\Big|\,\omega\big(\widehat{L}_{\lambda,\tau^i_\lambda(t)}(-1,\ldots,-i), \breve{E}_n - v_{ni}(t)\big)\Big]&\\
- \log P\Big[\omega(0,0)\,\Big|\,\mathcal{F}\big(\widehat{L}_{\lambda,\tau^i_\lambda(\{t\})}(-1,\ldots,-i)\big)\Big](\omega)\Big\|_{\mathcal{L}^1(P)} &= 0.
\end{aligned}$$

Proceeding with (66) as in (44) to (51), and using that

$$\lim_{n\to\infty} \frac{\text{length } E_n^{(r)}}{\text{length } \widehat{L}_n^\pi} = \frac{\text{length } \pi^{(r)}}{\text{length } \pi}$$

for all $r \in \{1,\ldots,R\}$ concludes the proof. □

4.3 Curves

The last step is to pass from polygons to curves. Assume that $c : [0,T] \longrightarrow \mathbb{R}^2$ is piecewise differentiable and parametrized by arc length. If ambiguous, c' denotes the right derivative of c. First of all, we need to relate the derivatives of the curve with a slope of a line. Let $v \in S^1 = \{w \in \mathbb{R}^2 \mid \; \| w \| = 1\}$, and α the angle from the positive x-axis to the vector v. If $|\alpha| \le \frac{\pi}{4}$ or $|\alpha| \ge \frac{3\pi}{4}$ then describe the line in the direction of v by a function of the x-axis; otherwise describe it as a function of the y-axis. Let $\lambda(v)$ be the slope of that line, that is,

$$\lambda(v) := \min(|tg\,\alpha|, |ct\,\alpha|). \tag{68}$$

By this correspondence, we can assign any $v \in S^1$ a specific entropy. We will write

$$h_v(P) := h_{\lambda(v)}(P). \tag{69}$$

Theorem 4.8. *Let $c : [0,T] \longrightarrow \mathbb{R}^2$ be a piecewise continuously differentiable curve, and c' its derivative, and let $\pi_n : [0, nT] \longrightarrow \mathbb{R}^2$ $(n \in \mathbb{N})$ be a sequence of polygons such that*

$$\frac{1}{\text{length } \pi_n} \sup_{t \in [0,nT]} \left|(B_n^{-1}\pi_n)'(t) - c'(t)\right| \xrightarrow{n\to\infty} 0. \tag{70}$$

Then we have in $\mathcal{L}^1(P)$,

$$\frac{1}{\text{length } \pi_n} \mathcal{I}(P_{L_n^\pi}) \xrightarrow{n\to\infty} \frac{1}{\text{length } c} \int_0^T h_{c'(t)}(P)\, dt, \tag{71}$$

and

$$\frac{1}{\text{length } \pi_n} \mathcal{I}(P_{L_n^\pi}) \xrightarrow{n\to\infty} \frac{1}{\text{length } c} \int_0^T \widehat{h}_{c'(t)}(P)\, dt. \tag{72}$$

Proof. We carry out the proof of (71). The proof of (72) is similar. We have to show that

$$\lim_{n\to\infty} \left\| \frac{1}{\text{length } \pi_n} \mathcal{I}(P_{L_n^\pi}) - \frac{1}{\text{length } c} \int_0^T h_{c'(t)}(P)\, dt \right\|_{\mathcal{L}^1(P)} = 0. \tag{73}$$

As can be seen by the construction of the entropy for polygons,

$$\left\| \frac{1}{\text{length } \pi_n} \mathcal{I}(P_{L_n^\pi}) - h_{\pi_n}(P) \right\|_{\mathcal{L}^1(P)}$$

converges to 0. By the representation formula in Remark 4.6 and since

$$\pi_n'(t) = ((B_n \pi_n)^{-1})'(t/n) \quad \text{for all } t \in [0, nT],$$

we obtain

$$h_{\pi_n}(P) = \frac{1}{\text{length } \pi_n} \int_0^{nT} h_{\pi_n'(r)}(P)\, dr = \frac{n}{\text{length } \pi_n} \int_0^T h_{((B_n\pi_n)^{-1})'(t)}(P)\, dt,$$

and by (70) and Lemma 3.4, the integral converges to

$$\int_0^T h_{c'(t)}\, dt.$$

Use $\|\cdot\|$ for the euclidian norm in $\mathbb{R}^2$. Since

$$\frac{1}{n}\text{length } \pi_n = \int_0^T \|\pi_n'(nt)\|\, dt = \int_0^T \|(B_n^{-1})'(t)\|\, dt,$$

we obtain by (70)

$$\lim_{n\to\infty} \frac{1}{n}\text{length } \pi_n = \int_0^T \|c'(t)\|\, dt = \text{length } c.$$

This implies that

$$\left\| h_{\pi_n}(P) - \frac{1}{\text{length } c} \int_0^T h_{c'(t)}(P)\, dt \right\|_{\mathcal{L}^1(P)}$$

converges to 0, too, and (73) follows by the triangle inequality. □

Note that the limits do not depend on the sequence of polygons we used to approximate the curve. Observe, further, that any appropriate approximation of the curve by lattice points can be described by a lattice approximation of a suitable polygon. This justifies

Definition 4.9. *Let P and be as in Theorem 4.8. Then*

$$h_c(P) := \frac{1}{\text{length } c} \int_0^T h_{c'(t)}(P)\, dt$$

is called specific entropy of P along c, *and*

$$\widehat{h}_c(P) := \frac{1}{\text{length } c} \int_0^T \widehat{h}_{c'(t)}(P)\, dt$$

is called specific contour entropy of P along c.

Note that the following property for the entropies of the blowups of a curve.

Corollary 4.10. *Let $c : [0,T] \longrightarrow \mathbb{R}^2$ be a piecewise differentiable curve, and let $B_\eta c : [0, \eta T] \longrightarrow \mathbb{R}^2$, $B_\eta c(t) := \eta c(\frac{t}{\eta})$ $(\eta > 0)$ be the family of its blowups. Then*

$$h_{B_\eta c}(P) = h_c(P) \quad \textit{for all } \eta > 0.$$

Proof.

$$h_{B_\eta c}(P) = \frac{1}{\text{length } B_\eta c} \int_0^{\eta T} h_{(B_\eta c)'(t)}(P)\, dt = \frac{1}{\eta \,\text{length } c} \int_0^{\eta T} h_{c'(\frac{t}{\eta})}(P)\, dt = h_c(P).$$

□

5 Gibbs measures

We will define Gibbs measures in terms of interaction potentials. A collection $(U_V)_{V \subset \mathbb{Z}^2 \text{ finite}}$ of functions on Ω is called *stationary summable interaction potential* if the following three conditions are fulfilled:

(i) U_V is measurable with respect to $\mathcal{F}_V$ for all $V \subset \mathbb{Z}^2$.

(ii) For all $i \in \mathbb{N}$ and all finite $V \subset \mathbb{Z}^2$, $U_{V+i} = U_V \circ \theta_i$.

(iii)
$$\sum_{V \subset \mathbb{Z}^2 \text{ finite}:\, 0 \in V} \| U_V \|_\infty < \infty$$

Let $\xi, \eta \in \Omega$. The *conditional energy* of ξ on V given the environment η on $\mathbb{Z}^2 \setminus V$ is defined as

$$E_V(\xi|\eta) = \sum_{A \subset \mathbb{Z}^2 \text{ finite}:\, A \cap V \neq \emptyset} U_A\big((\xi,\eta)_V\big), \tag{74}$$

where $(\xi,\eta)_V$ is the element of Ω given by

$$(\xi,\eta)_V(i) := \begin{cases} \xi(i) & i \in V, \\ \eta(i) & i \in \mathbb{Z}^d \setminus V. \end{cases} \tag{75}$$

P is called *Gibbs measure* with respect to U if for any finite subset V of $\mathbb{Z}^2$ the conditional distribution of ω_V under P with respect to $\mathcal{F}_{\mathbb{Z}^2 \setminus V}$ is given by

$$P[\omega_v = \xi_V \mid \mathcal{F}_{\mathbb{Z}^2 \setminus V}](\eta) = \frac{1}{Z_V(\eta)} e^{-E_V(\xi|\eta)}, \tag{76}$$

where

$$Z_V(\eta) := \int_\Omega e^{-E_V(\xi|\eta)} P(d\xi) \tag{77}$$

is called *partition function*. We say that there is a *phase transition* if there is more than one Gibbs measure with respect to the same interaction potential. Gibbs measure with respect to nearest neighbor potentials are Markov fields.

Assume that Υ is furnished with a total order $\leq$, and denote by $-$ the minimal and by $+$ the maximal element in Υ. Suppose that U is *attractive* with respect to the order on Υ, in the sense of (9.7) in [24]. Let P^- and P^+ denote the minimal and the maximal Gibbs measure with respect to U, and let $P^\alpha = \alpha P^- + (1-\alpha)P^+$ $(0 < \alpha < 1)$ be their mixtures. Both P^- and P^+ are ergodic and, as follows from [8], they fulfill the global Markov property and the strong 0-1 law, and we can define *relative surface entropies.* Using the past σ-algebras (29), (54) and (55), and the correspondance (68) between directions and slopes, they have the following form.

Definition 5.1. *Let $v \in S^1$. Then*

$$h_v(P^-, P^+) := \int_0^1 \int_\Omega H\Big(P_0^-[\,\cdot\,|\mathcal{P}_{\lambda(v),t}](\omega), P_0^+[\,\cdot\,|\mathcal{P}_{\lambda(v),t}](\omega)\Big)\, P^-(d\omega)dt$$

is the specific relative entropy *of P^- with respect to P^+ in along v, and*

$$\begin{aligned}\widehat{h}_v(P^-, P^+) := \frac{1}{1+\lambda(v)}\Bigg[&\int_0^1 \int_\Omega H\Big(P_0^-[\,\cdot\,|\mathcal{P}^\sharp_{\lambda(v),t}](\omega), P_0^+[\,\cdot\,|\mathcal{P}^\sharp_{\lambda(v),t}](\omega)\Big)\, P^-(d\omega)\,dt \\ &+ \int_{1-\lambda(v)}^1 \int_\Omega H\Big(P_0^-[\,\cdot\,|\mathcal{P}^\flat_{\lambda(v),t}](\omega), P_0^+[\,\cdot\,|\mathcal{P}^\flat_{\lambda(v),t}](\omega)\Big)\, P^-(d\omega)\,dt\Bigg]\end{aligned}$$

is the specific relative contour entropy *of P_0^- with respect to P_0^+ along v.*

Let $c : [0,T] \longmapsto \mathbb{R}^2$ a piecewise differentiable curve parametrized by arc length and c' its right derivative. Then

$$h_c(P^-, P^+) := \frac{1}{\text{length } c}\int_0^T h_{c'(t)}(P^-, P^+)\,dt$$

is the specific relative entropy *of P^- with respect to P^+* along c, *and*

$$\widehat{h}_c(P^-, P^+) := \frac{1}{\text{length } c}\int_0^T \widehat{h}_{c'(t)}(P^-, P^+)\,dt$$

is the specific relative contour entropy *of P^- with respect to P^+* along c.

The order on Υ induces an order on the set $\mathcal{M}_1(\Upsilon)$ of probability measures on Υ : We say that μ is *larger* then ν if the density $\frac{d\mu}{d\nu}$ is an increasing function with respect to the order on Υ, and in this case we write $\mu \geq \nu$. In particular, ν is absolutely continuous with respect to μ.

The following inverse triangle inequality for relative entropies was shown in the proof of Theorem 4.2 in [10]. For the reader's convenience we state it in the following form.

Lemma 5.2. *Let $\lambda \geq \mu \geq \nu$, and assume that μ is bounded below by a positive constant. Then*

$$H(\nu, \lambda) \geq H(\nu, \mu) + H(\mu, \lambda).$$

To prove it, they use Theorem 9.4 from [24]:

Theorem 5.3. *(Preston) Let Λ be a finite subset of $\mathbb{Z}^d$, and m_Λ the product measure on Ω_Λ. Let f_1 and f_2 be nonnegative measurable functions on $(\Omega_\Lambda, \mathcal{F}_\Lambda)$ with*

$$\int_{\Omega_\Lambda} f_1 \, dm_\Lambda = \int_{\Omega_\Lambda} f_2 \, dm_\Lambda = 1.$$

Suppose that

$$f_1(\omega \vee \widetilde{\omega}) f_2(\omega \wedge \widetilde{\omega}) \geq f_1(\omega) f_2(\widetilde{\omega}) \qquad \text{for all} \quad \omega, \widetilde{\omega} \in \Omega_\Lambda \tag{78}$$

Then for any bounded measurable increasing function g on $(\Omega_\Lambda, \mathcal{F}_\Lambda)$ we have

$$\int_{\Omega_\Lambda} g f_1 \, dm_\Lambda \geq \int_{\Omega_\Lambda} g f_2 \, dm_\Lambda.$$

Proof of the lemma. By $\mu \geq \nu$, $f_1 := 1$ and $f_2 := \frac{d\nu}{d\mu}$ fulfill condition (78), and by $\lambda \geq \mu$ and the boundedness of μ, $g := -\log \frac{d\mu}{d\lambda}$ is increasing and bounded. By Preston's theorem we obtain

$$\int_\Upsilon \log \frac{d\mu}{d\lambda} \frac{d\nu}{d\mu} \, d\mu \geq \int_\Upsilon \log \frac{d\mu}{d\lambda} \, d\mu,$$

and this yields

$$H(\nu, \lambda) = H(\nu, \mu) + \int_\Upsilon \log \frac{d\mu}{d\lambda} \, d\nu \geq H(\nu, \mu) + H(\mu, \lambda).$$

□

6 Lower bounds

Assume that P^- and P^+ have the local Markov property (Definition 2.1). Then the strong 0-1 law implies that they have the global Makov property as well. Föllmer and Ort [10] introduced the *specific relative surface entropy*

$$s(P^-, P^+) = \frac{1}{2} \sum_{l=1}^{2} \int_\Omega H\Big(P_0^-[\,\cdot\,|\mathcal{F}^{(l)}](\omega), P_0^+[\,\cdot\,|\mathcal{F}^{(l)}](\omega)\Big) \, P^-(d\omega), \tag{79}$$

where $\mathcal{F}^{(l)}$ is the σ-algebra generated by those coordinates in $\{(i^{(1)}, i^{(2)} \in \mathbb{Z}^2 \,|\, i^{(l)} = 0\}$ which precede 0 in the lexicographical order on $\mathbb{Z}^2$. Then they proved the following lower bound for the large deviations of the *empirical field* of P^+,

$$R_n(\omega) = \sum_{i \in V_n} \delta_{\theta_i \omega}.$$

Theorem 6.1. *(Föllmer and Ort) For any open $A \in \mathcal{M}_1(\Omega)$,*

$$\liminf_{n\to\infty} \frac{1}{|\partial V_n|} P^+\big[R_n \in A\big] \geq - \inf_{\alpha: P_\alpha \in A} \sqrt{\alpha}\, s(P^-, P^+).$$

Recall that the boundary of a subset V of $\mathbb{Z}^2$ is defined as

$$\partial V = \{ i \in \mathbb{Z}^2 \setminus V \,|\, \mathrm{dist}(i, V) = 1 \}. \tag{80}$$

The aim of this section is to improve the lower bound by replacing the boxes by more general shapes and using the corresponding Shannon-MacMillan theorems of Section 4. For a closed curve c let $\operatorname{int} c$ be the subset of $\mathbb{R}^2$ surrounded by c. Define the set

$$\begin{aligned} C_\alpha := \{ c \,|\, c : [0,T] \longrightarrow \mathbb{R}^2 & \text{ closed piecewise } C^1\text{-curve parametrized by arc} \\ & \text{length, without self-intersections, and with area} \operatorname{int} c = \alpha \}. \end{aligned} \tag{81}$$

Theorem 6.2. *For any open $A \in \mathcal{M}_1(\Omega)$,*

$$\liminf_{n\to\infty} \frac{1}{|\,\partial V_n\,|} P^+\big[R_n \in A\big] \geq - \inf_{\alpha : P_\alpha \in A} \inf_{c \in \mathcal{C}_\alpha} \frac{1}{4} \int_0^T \frac{dt}{\sqrt{1 + \lambda(c'(t))^2}}\; h_c(P^-, P^+).$$

Replacing the class C_α by squares with area α this bound coincides with the bound in Theorem 6.1: Let π be a square parametrized by arc length and with area $\operatorname{int} \pi = \alpha$. Then the length of every edge is $\sqrt{\alpha}$. For the two horizontal edges of the square the slope λ (c.f. (68)) is 0 with respect to the x-axis, and for the horizontal edges it is 0 with respect to the y-axis. Therefore, the integral equals $4\sqrt{\alpha}$. The entropy $h_\pi(P^-, P^+)$ equals $s(P^-, P^+)$, since the σ-algebras $\mathcal{P}_{0,t}$ coincide with $\mathcal{F}^{(2)}$ for the horizontal edges and with $\mathcal{F}^{(1)}$ for the vertical edges.

The proof of Theorem 6.2 follows the lines of Föllmer and Ort's proof with some adaptations to the different geometry. To begin with, we restate explicitly the global Markov property for random fields in the case when the conditioning is concentrated on a set of sites surrounded by a closed polygon. We use the notation $\Gamma(c) := \operatorname{int} c \cap \mathbb{Z}^2$ to indicate the set of lattice points surrounded by a closed curve c. Let π be a closed polygon without self-intersections. By (80) and Definition 57, $\partial\big(\mathbb{Z}^2 \setminus \Gamma(\pi)\big) = L^\pi$, and the global Markov property (see Definition 2.1), with $V = \mathbb{Z}^2 \setminus \Gamma(\pi)$, yields for any $\mathcal{F}(\mathbb{Z}^2 \setminus \Gamma(\pi))$-measurable nonnegative function Φ,

$$E\big[\phi \,\big|\, \mathcal{F}_{\Gamma(\pi)}\big] = E\big[\phi \,\big|\, \mathcal{F}_{L^\pi}\big]. \tag{82}$$

We will further need two lemmata that compute the asymptotic fractions of the lengths of a line segment, or a polygon, and the sizes of their lattice approximation.

Lemma 6.3. *Let I be a real interval, $l(x) = \lambda x + a$ be a linear function with slope λ, and B_k $(k \in \mathbb{N})$ be the sequence of its blowups restricted to I. If L_k is the lattice approximation of B_k then*

$$\lim_{k\to\infty} \frac{|\,L_k\,|}{\text{length } B_k} = \frac{1}{\sqrt{1+\lambda^2}}. \tag{83}$$

Proof. We consider only the case when $0 \leq \lambda \leq 1$, that is, when the lattice approximation is given by $L(z) = (z, [l(z)])$ $(z \in I \cap \mathbb{Z})$. In the other cases we use similar functions to describe the lattice approximations, and the proof of (83) is analoguous.

For any $k \in \mathbb{N}$, $|L_k|$ is either $[\text{length}\, b_k]$ or $[\text{length}\, b_k] + 1$, where b_k is the projection of B_k to the x-axis. We can ignore the second case, since the additional point does not matter for the limit in (83). Observe that $(\text{length}\, B_k)^2 = (\text{length}\, b_k)^2 + (\lambda\, \text{length}\, b_k)^2$. Consequently,

$$\text{length}\, b_k = \frac{\text{length } B_k}{\sqrt{1+\lambda^2}}, \tag{84}$$

which proves the convergence in (83). $\square$

Lemma 6.4. *Let π be a polygon with edges $\pi_1, \ldots, \pi_R$, $B_k\pi$ $(k \in \mathbb{N})$ its blowups, and $L_k\pi$ $(k \in \mathbb{N})$ their lattice aproximations. Then*

$$\lim_{k\to\infty} \frac{|\, L_k\pi \,|}{\text{length}\, B_k\pi} = \sum_{r=1}^{R} \frac{1}{\sqrt{1+\lambda_r^2}} \frac{\text{length}\,\pi_r}{\text{length}\,\pi}, \tag{85}$$

where λ_r is the slope corresponding to π_r' as defined in (68).

Proof. We have

$$\frac{|\, L_k\pi \,|}{\text{length}\, B_k\pi} = \sum_{r=1}^{R} \frac{|\, L_k\pi \,|}{\text{length}\, B_k\pi_r} \frac{\text{length}\, B_k\pi_r}{\text{length}\, B_k\pi},$$

and using $\text{length}\, B_k\pi_r = k\,\text{length}\,\pi_r$, we obtain

$$\sum_{r=1}^{R} \frac{|\, L_k\pi \,|}{\text{length}\, B_k\pi_r} \frac{\text{length}\,\pi_r}{\text{length}\,\pi}.$$

By the previous lemma applied to the individual sides, the first factors converge to $\left(\sqrt{1+\lambda_r^2}\right)^{-1}$, which implies (85). $\square$

Proof of the theorem. Let be $0 < \alpha \leq 1$, such that $P_\alpha \in A$. Since A is open, we can choose open neigborhoods A^- and A^+ of P^- respectively P^+ in $\mathcal{M}_1(\Omega)$ such that

$$\alpha A^- + (1-\alpha)A^+ \subseteq A.$$

Without loss of generality we may assume that A^- and A^+ are in $\mathcal{F}_{V_p}$ for some $p \in \mathbb{N}$. Define the set

$$\Pi_\alpha := \left\{\pi \,\middle|\, \pi \text{ closed polygon without self-intersections, area}\,\text{int}\,\pi = \alpha\right\}.$$

Let $\pi \in \Pi_{4\alpha}$ with $0 \in \text{int}\,\pi$, and let $(B_n\pi)_{n\in\mathbb{N}}$ be the sequence of blowups of π. For $\alpha = 1$ take $C_n := V_n$. Otherwise, define

$$C_n := \Gamma(B_{k(n)}\pi) \quad \text{and} \quad D_n := V_n \setminus \Gamma(B_{l(n)}\pi),$$

where $k(n)$ and $l(n)$ are chosen such that $k(n) \leq l(n)$, $l(n) - k(n) \xrightarrow{n\to\infty} \infty$,

$$\lim_{n\to\infty} \frac{|\, C_n \,|}{|\, V_n \,|} = \alpha \quad \text{and} \quad \lim_{n\to\infty} \frac{|\, D_n \,|}{|V_n\,|} = 1-\alpha. \tag{86}$$

To see that such sequences exist, we give explicit examples:

$$k(n) := \left[\sqrt{\frac{\alpha|V_n|}{\text{area}\,\text{int}\,(\pi)}}\,\right] \quad \text{and} \quad l(n) := \left[k(n) + \sqrt{n}\,\right].$$

Obviously, both the sequences and their difference tend to infinity as n goes to infinity. Using $\text{area}\,\text{int}\,(B_k\pi) = k^2 \text{area}\,\text{int}\,\pi$,

$$\frac{|\, \Gamma(B_k\pi) \,|}{\text{area}\,\text{int}\, B_k\pi} \xrightarrow{k\to\infty} 1,$$

and

$$\frac{k(n)^2}{|\, V_n \,|} \xrightarrow{n\to\infty} \frac{\alpha}{\text{area}\,\text{int}\,\pi},$$

we obtain for the expression in (86)

$$\lim_{n\to\infty} \frac{|\, C_n \,|}{|\, V_n \,|} = \lim_{n\to\infty} \frac{\text{area int}\,(B_{k(n)}\pi)}{|\, V_n \,|} = \lim_{n\to\infty} \frac{k(n)^2 \,\text{area int}\,\pi}{|\, V_n \,|} = \alpha.$$

For the second expression in (86) we note that

$$\frac{l(n)^2}{|\, V_n \,|} = \frac{k(n)^2}{|\, V_n \,|} + \frac{2k(n)\sqrt{n} + n}{|\, V_n \,|},$$

and that the second summand tends to 0 for n going to infinity. Now, doing the same type of calculation as for $k(n)$, we obtain

$$\lim_{n\to\infty} \frac{|\, D_n \,|}{|\, V_n \,|} = 1 - \lim_{n\to\infty} \frac{|\, \Gamma(B_{l(n)}) \,|}{|\, V_n \,|} = 1 - \alpha.$$

Define

$$R_n^- = \frac{1}{|\, C_{n,p} \,|} \sum_{i \in C_{n,p}} \delta_{\theta_i \omega} \qquad \text{and} \qquad R_n^+ = \frac{1}{|\, D_{n,p} \,|} \sum_{i \in D_{n,p}} \delta_{\theta_i \omega},$$

where $C_{n,p} := V_{k(n)-p}$ and $D_{n,p} := V_{n-p} \setminus V_{l(n)+p}$. Then

$$\{R_n^- \in A^-\} \in \mathcal{F}_{C_n}, \quad \{R_n^+ \in A^+\} \in \mathcal{F}_{D_n}, \tag{87}$$

and for large enough n,

$$\{R_n \in A\} \supseteq \{R_n^- \in A^-\} \cap \{R_n^+ \in A^+\} := \Lambda_n.$$

Define the measures $Q_n = P_{C_n}^- \otimes P_{\mathbb{Z}^2 \setminus C_n}^+ \quad (n \in \mathbb{N})$. Q_n coincides with P^- on $\mathcal{F}_{C_n}$ and with P^+ on $\mathcal{F}_{D_n}$, and makes these σ-fields independent. Thus, and by (87), we obtain $Q_n[\Lambda_n] = P^-[R_n^- \in A^-]\, P^+[R_n^+ \in A^+]$, and by the ergodic behaviour of P^- and P^+ we have

$$Q_n[\Lambda_n] \xrightarrow{n\to\infty} 1. \tag{88}$$

Let ϕ_n denote the density of Q_n with respect to P^+ on $\mathcal{F}_{C_n \cup D_n}$. Then for $\gamma > 0, \varepsilon > 0$, and for large enough n,

$$\begin{aligned} P^+[R_n \in A)] &\geq P^+[\Lambda_n] \\ &\geq \int 1_{\Lambda_n \cap \{\frac{1}{|\partial V_n|} \log \phi_n \leq \gamma + \varepsilon\}} \phi_n^{-1}\, dQ_n \\ &\geq \exp(-(\gamma + \varepsilon)\, |\partial V_n|)\, Q_n \Big[\Lambda_n \cap \Big\{\frac{1}{|\partial V_n|} \log \phi_n \leq \gamma + \varepsilon\Big\}\Big]. \end{aligned}$$

By (88), the lower bound

$$\liminf_{n\to\infty} \frac{1}{|\partial V_n|} \log P^+[R_n \in A] \geq -\gamma$$

follows if γ is chosen such that, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} Q_n \Big[\frac{1}{|\partial V_n|} \log \phi_n \leq \gamma + \varepsilon\Big] = 1. \tag{89}$$

It remains to show that (89) holds with

$$\gamma = \sum_{r=1}^{R} \frac{1}{\sqrt{1 + \lambda_r^2}} \frac{\text{length}\,\pi_r}{8}\; h_\pi(P^-, P^+).$$

Since $Q_n = P^+$ on D_n, and the fact that both P^- and P^+ are Gibbs measures with respect to the same potential we obtain

$$\phi_n(\omega) = \frac{P^-[\omega_{C_n}]P^+[\omega_{D_n}]}{P^+[\omega_{C_n \cup D_n}]} = \frac{P^-[\omega_{C_n}]}{P^-[\omega_{C_n} \,|\, \omega_{D_n}]}.$$

Let L_n be the lattice approximation (see Definition 35) of $B_n\pi$. By (82),

$$\begin{aligned} P^-[\omega_{C_n} \,|\, \omega_{D_n}] &= P^-[\omega_{D_n} \,|\, \omega_{C_n}] \frac{P^-[\omega_{D_n}]}{P^-[\omega_{C_n}]} \\ &= P^-[\omega_{D_n} \,|\, \omega_{L_{k(n)}}] \frac{P^-[\omega_{D_n}]}{P^-[\omega_{C_n}]} = P^-[\omega_{L_{k(n)}} \,|\, \omega_{D_n}] \frac{P^-[\omega_{L_{k(n)}}]}{P^-[\omega_{C_n}]}, \end{aligned}$$

and thus

$$\phi_n(\omega) = \frac{P^-(\omega_{L_{k(n)}})}{P^-(\omega_{L_{k(n)}} \,|\, \omega_{D_n})} = \frac{P^-(\omega_{L_{k(n)}})}{P^+(\omega_{L_{k(n)}} \,|\, \omega_{D_n})}. \tag{90}$$

Going around the R sides of $L_{k(n)}$, and conditioning site by site as in the proof of Theorem 4.5, we obtain

$$\frac{1}{|\,V_n\,|} \log \phi_n(\omega) = \frac{1}{|\,V_n\,|} \sum_{r=1}^{R} \Psi^{(r)},$$

where the $\Psi^{(r)}$ corresponds to the r-th side of the polygon. Similar to the calculation between (64) to (66) we obtain

$$\Psi^{(r)} = \sum_{i=0}^{u_n} Z_{n,i,t} \circ \theta_{L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)} \circ \theta_{L_{\lambda,t_n}(z_n)},$$

where λ is the slope of the rth side of the polygon, t_n and r_n are as in Subsection 4.2, and $Z_{n,i,t} = X_{n,i,t} - Y_{n,i,t}$, with

$$\begin{aligned} X_{n,i,t} &= \log P_0^-\Big(\omega(0,0) \,\Big|\, \omega\big(\widehat{L}^\sharp_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\ldots,-i) \cup A_{n,i,t}\big)\Big), \\ \text{and } Y_{n,i,t} &= \log P_0^+\Big(\omega(0,0) \,\Big|\, \omega\big(\widehat{L}^\sharp_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\ldots,-i) \cup B_{n,i,t}\big)\Big). \end{aligned}$$

To simplify notation we have omitted the index r. For the sets in the conditional expectations we have $A_{n,i,t} \subseteq B_{n,i,t} \subseteq \mathbb{Z}^2 \setminus \big(H_n - L_{\lambda,t_n}(z_n+i)\big)$. $A_{n,i,t}$ is obtained by shifting a subset of $L_n \subseteq C_n$. H_n is constructed as in the paragraph above (67), but using the minimum of the diameter d_n and the distance $l(n) - k(n)$ in place of d_n.

To prove convergence, we study the X and Y-parts separately. By construction of the sets $A_{n,i,t}$, the behavior of $X_{n,i,t}$ under Q_n is the same as under P^-. But the proof of Theorem 4.5 shows that

$$\sum_{i=0}^{u_n} X_{n,i,t} \circ \theta_{L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)} \circ \theta_{L_{\lambda,t_n}(z_n)} \tag{91}$$

converges to $-h_\pi(P^-)$ in $\mathcal{L}^1(P^-)$. The convergence remains true when we replace $X_{n,i,t}$ by

$$X^-_{n,i,t} := \log P_0^-\Big(\omega(0,0) \,\Big|\, \omega\big(L_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\ldots,-i)\big)^-\Big),$$

where, for a subset L of $\mathbb{Z}^2$, the element $\omega\big(L\big)^-$ equals ω on L and assumes the minimal state in Υ outside of $H_n - L_{\lambda,t_n}(z_n+i)$.

We still need to control the behavior of $Y_{n,i,t}$ under Q_n. Put

$$Z^-_{n,i,t} = X^-_{n,i,t} - Y_{n,i,t},$$

and use the law of large numbers for martingales with bounded increments in its $\mathcal{L}^2$-form in order to replace

$$\frac{1}{|\, L_{k(n)}\,|} \sum_{i=0}^{u_n} Z^-_{n,i,t} \circ \theta_{L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)} \circ \theta_{L_{\lambda,t_n}(z_n)}$$

by

$$\frac{1}{|\, L_{k(n)}\,|} \sum_{i=0}^{u_n} E\big[Z^-_{n,i,t} \circ \theta_{L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)} \circ \theta_{L_{\lambda,t_n}(z_n)} \,\big|\, \mathcal{A}_{n,i,t}\big],$$

where $\mathcal{A}_{n,i,t}$ is the σ-field generated by the sites in D_n and those sites of $L_{k(n)}$ which precede i in the canonical ordering of $L_{k(n)}$. These conditional expectations can be written as the relative entropy $H(\nu,\mu)$, with the random measures

$$\begin{aligned} \mu(\omega) :=& P_0^-\big[\;\cdot\;\big|\,\omega\big(L_{\lambda,t_\lambda^{z_n+i}(\{t_n\})}(-1,\ldots,-i)\big)^-\big] \\ \text{and} \qquad \nu(\omega) :=& P_0^+\big[\;\cdot\;\big|\,\omega\big(L_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\ldots,-i) \cup B_{n,i,t}\big)\big]. \end{aligned}$$

Our next step is to replace μ by a measure η for which

$$\frac{1}{|\, L_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\ldots,-i)\,|} \sum_{i=0}^{u_n} H(\nu,\eta) \circ \theta_{L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)} \circ \theta_{L_{\lambda,t_n}(z_n)} \tag{92}$$

converges to $h_\lambda(P^-,P^+)$, in $\mathcal{L}^1(P^-)$, as n goes to infinity. Define $\omega(L)^+$ in analogy to $\omega(L)^-$. Since for *all* ω,

$$P_0^-\Big[\;\cdot\;\Big|\,\omega\big(L_{\lambda,t}(-1,\ldots,-i)\big)^+\Big] \xrightarrow{i\to\infty} P_0^-\Big[\;\cdot\;\Big|\,\mathcal{P}_{\lambda,t}\Big],$$

we obtain (92) by taking

$$\eta(\omega) := P_0^-\Big[\;\cdot\;\Big|\,\omega\big(L_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\ldots,-i)\big)^+\Big].$$

By Lemma 5.2 $H(\nu(\omega),\mu(\omega)) \le H(\nu(\omega),\eta(\omega))$. Summing over $r = 1,\ldots,R$, and passing from convergence in $\mathcal{L}^1(P^-)$ to stochastic convergence with respect to Q_n yields

$$\lim_{n\to\infty} Q_n\left[\frac{1}{|\, L_{k(n)}\,|}\phi_n > h_\pi(P^-,P^+) + \varepsilon\right] = 0 \tag{93}$$

for any $\varepsilon > 0$.

To derive (89) with

$$\gamma = \sum_{r=1}^{R} \frac{1}{\sqrt{1+\lambda_r^2}} \frac{\text{length}\ \pi_r}{8}\ h_\pi(P^-,P^+),$$

it remains to show that

$$\lim_{n\to\infty} \frac{|\, L_{k(n)}\,|}{|\, \partial V_n\,|} = \sum_{r=1}^{R} \frac{1}{\sqrt{1+\lambda_r^2}} \frac{\text{length}\ \pi}{8}. \tag{94}$$

The fraction on the left-hand side can be written as

$$\frac{|\,L_{k(n)}\,|}{|\,\partial V_n\,|} = \frac{|\,L_{k(n)}\,|}{\text{length } B_{k(n)}\pi}\,\frac{k(n)\,\text{length } B_{k(n)}\pi}{\text{area int } B_{k(n)}\pi}\,\frac{\text{area int } B_{k(n)}\pi}{k(n)\,|\,\partial V_n\,|}.$$

Apply Lemma 6.4 to the first factor, and observe that

$$\frac{k(n)\,\text{length } B_{k(n)}\pi}{\text{area int } B_{k(n)}\pi} = \frac{k(n)^2\,\text{length } \pi}{k(n)^2\,\text{area int } \pi} = \frac{\text{length } \pi}{4\alpha}. \tag{95}$$

Then using (86) to see that

$$\frac{\text{area int } B_{k(n)}\pi}{k(n)\,|\,\partial V_n\,|} = \frac{k(n)^2\,\text{area int } \pi}{k(n) 4\sqrt{|\,V_n\,|}} = \alpha\sqrt{\frac{k(n)^2}{|\,V_n\,|}} \tag{96}$$

converges to $\frac{1}{2}\alpha$, (94) follows.

Finally, we replace the polygon π by the polygon $\widetilde{\pi} = B_{\frac{1}{2}}\pi$. Since length $\widetilde{\pi}_r = \frac{1}{2}$length π_r, and since, by Corollary 4.10, $h_\pi(P^-,P^+) = h_\pi(P^-,P^+)$, γ transforms into

$$\sum_{r=1}^{R} \frac{1}{\sqrt{1+\lambda_r^2}}\,\frac{\text{length } \widetilde{\pi}_r}{4}\,h_\pi(P^-,P^+).$$

Finally, by Lemma 3.4, the infimum of that function over all polygons $\widetilde{\pi} \in \Pi_\alpha$ equals the infimum over all curves $c \in C_\alpha$, and we obtain the bound (6.2). □

If we have the Markov property only with respect to the contour boundary we can prove a bound similar to Theorem 6.2 by replacing the lattice approximation by the contour aproximation. Recall that, for any $V \subseteq \mathbb{Z}^2$, the contour boundary is defined as $\widehat{\partial}V = \{z \in \mathbb{Z}^2 \setminus V \,|\, \text{dist}(z,V) = 1 \text{ or } \text{dist}(z,V) = \sqrt{2}\}$. Note that the two boundaries for a box differ only in the four corners. Thus it does not matter whether we divide by ∂V_n or by $\widehat{\partial}V_n$ on the left-hand side.

Theorem 6.5. *For any open $A \in \mathcal{M}_1(\Omega)$,*

$$\liminf_{n\to\infty} \frac{1}{|\partial V_n|} P^+\big[R_n \in A\big] \geq -\inf_{\alpha : P_\alpha \in A}\,\inf_{c \in C_\alpha}\,\frac{1}{4}\int_0^T \frac{1+\lambda(c'(t))}{\sqrt{1+\lambda(c'(t))^2}}\,dt\,\widehat{h}_c(P^-,P^+).$$

The proof is the same as for Theorem 6.2 except for the changes due to the weaker Markov property. First of all, we need a contour version of (82). For a closed polygon π without self-intersections we obtain by Definition 57,

$$\widehat{\partial}\big(\mathbb{Z}^2 \setminus \Gamma(\pi)\big) = \widehat{L}^\pi. \tag{97}$$

The global Markov property from Definition 2.1 with respect to the contour boundary, yields $V = \mathbb{Z}^2 \setminus \Gamma(\pi)$, yields for any $\mathcal{F}(\mathbb{Z}^2 \setminus \Gamma(\pi))$-measurable nonnegative function Φ,

$$E\big[\phi \,\big|\, \mathcal{F}_{\Gamma(\pi)}\big] = E\big[\phi \,\big|\, \mathcal{F}_{L^\pi}\big]. \tag{98}$$

As in the proof of Theorem 6.2 we need lemmata computing the asymptotic ratio of the length of a line segment or a polygon, and its contour approximation.

Lemma 6.6. *Let I be a real interval, let $l(x) = \lambda x + a$ be a linear function with slope λ, and let B_k ($k \in \mathbb{N}$) be the sequence of its blowups restricted to I. Then we have for the contour approximation $\widehat{L}_k$ of B_k,*

$$\lim_{k\to\infty} \frac{|\,\widehat{L}_k\,|}{\text{length } B_k} = \frac{1}{\sqrt{1+\lambda^2}}. \tag{99}$$

Proof. This proof is similar to the proof of Lemma 6.3. As before, we consider only the case when $0 \le \lambda \le 1$. For any $k \in \mathbb{N}$, the number of points in L_k equals $[\text{length}\, b_k] + s_k$, where b_k is the projection of B_k to the x-axis, and s_k is the number of steps in $\widehat{L}_k$, that is, $s_k = |\{z \in b_k \mid \tau_\lambda^{z-1}(t_k) \ge 1 - \lambda\}|$. As in the proof of Lemma 6.3, $\widehat{L}_k$ may contain one or two more points, but they do not matter for the asymptotics in (99). By (84), we have

$$\begin{aligned} \frac{|\,\widehat{L}_k\,|}{\text{length}\, B_k} &= \frac{1}{\text{length}\, B_k}\left[\text{length}\, b_k \Big(1 + \frac{s_k}{\text{length}\, b_k}\Big)\right] \\ &= \frac{1}{\text{length}\, B_k}\left[\frac{\text{length}\, B_k}{\sqrt{1+\lambda^2}} \Big(1 + \frac{s_k}{\text{length}\, b_k}\Big)\right], \end{aligned}$$

and since

$$\lim_{k\to\infty} \frac{s_k}{\text{length}\, b_k} = \lambda,$$

we obtain the convergence in (99). □

Lemma 6.7. *Let π be a polygon with edges $\pi_1, \dots, \pi_R$, $B_k\pi$ $(k \in \mathbb{N})$ its blowups, and $\widehat{L}_k\pi$ $(k \in \mathbb{N})$ their contour aproximations. Then*

$$\lim_{k\to\infty} \frac{|\,\widehat{L}_k\pi\,|}{\text{length}\, B_k\pi} = \sum_{r=1}^{R} \frac{1+\lambda_r}{\sqrt{1+\lambda_r^2}} \frac{\text{length}\, \pi_r}{\text{length}\, \pi}. \tag{100}$$

Proof. The proof is exactly like the proof of Lemma 6.4, but applying Lemma 6.6 in place of Lemma 6.3. □

Proof of the theorem. We indicate only the changes compared to the proof of Theorem 6.2. We want to show that (89) holds with

$$\gamma = \sum_{r=1}^{R} \frac{1+\lambda_r}{\sqrt{1+\lambda_r^2}} \frac{\text{length}\, \pi_r}{8}\, \widehat{h}_\pi(P^-, P^+). \tag{101}$$

By (97), we have in place of (90),

$$\phi_n(\omega) = \frac{P^-(\omega_{L_{k(n)}})}{P^+(\omega_{L_{k(n)}} \,|\, \omega_{D_n})}. \tag{102}$$

Again, conditioning site by site, we go around the R sides of $\widehat{L}_{k(n)}$. Any time the contour approximation has a step, we have to consider the lower point $\downarrow$ as well as the upper point $\uparrow$. This leads to

$$\begin{aligned} \Psi^{(r)} = \sum_{i=0}^{u_n} & Z^{\uparrow}_{n,i,t} \circ \theta_{L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)} \circ \theta_{L_{\lambda,t_n}(z_n)} + \\ & 1_{\{\tau_\lambda^{z_n+i-1}(\{t_n\}) \ge 1-\lambda\}} Z^{\downarrow}_{n,i,t} \circ \theta_{L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)-(0,1)} \circ \theta_{L_{\lambda,t_n}(z_n)}, \end{aligned}$$

with

$$Z^{\uparrow}_{n,i,t} = X^{\uparrow}_{n,i,t} - Y^{\uparrow}_{n,i,t}, \qquad Z^{\downarrow}_{n,i,t} = X^{\downarrow}_{n,i,t} - Y^{\downarrow}_{n,i,t},$$

where

$$\begin{aligned} X^{\uparrow}_{n,i,t} &= \log P_0^-\Big(\omega(0,0)\,\Big|\,\omega\big(\widehat{L}^\sharp_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\dots,-i) \cup A^{\uparrow}_{n,i,t}\big)\Big), \\ Y^{\uparrow}_{n,i,t} &= \log P_0^+\Big(\omega(0,0)\,\Big|\,\omega\big(\widehat{L}^\sharp_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\dots,-i) \cup B^{\uparrow}_{n,i,t}\big)\Big), \end{aligned}$$

and

$$\begin{aligned} X^{\urcorner}_{n,i,t} &= \log P_0^- \Big(\omega(0,0) \,\Big|\, \omega\big(\{\widehat{L}_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\dots,-i) - (0,1)\} \cup A^{\urcorner}_{n,i,t}\big)\Big), \\ Y^{\urcorner}_{n,i,t} &= \log P_0^+ \Big(\omega(0,0) \,\Big|\, \omega\big(\{\widehat{L}_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\dots,-i) - (0,1)\} \cup B^{\urcorner}_{n,i,t}\big)\Big). \end{aligned}$$

To simplify notation, we omit the index r. The sets $A^{\ulcorner}_{n,i,t}$ and $A^{\urcorner}_{n,i,t}$ are obtained by shifting subsets of $\widehat{L}_n \subseteq C_n$, and they fulfill

$$\begin{aligned} A^{\ulcorner}_{n,i,t} &\subseteq B^{\ulcorner}_{n,i,t} \subseteq \mathbb{Z}^2 \setminus \big(H_n - L_{\lambda,t_n}(z_n+i)\big) \\ A^{\urcorner}_{n,i,t} &\subseteq B^{\urcorner}_{n,i,t} \subseteq \mathbb{Z}^2 \setminus \big(H_n - L_{\lambda,t_n}(z_n+i) + (0,1)\big). \end{aligned}$$

H_n is constructed as in the paragraph above (67) but using the minimum of d_n and $l(n)-k(n)$ in place of d_n. Without loss of generality we can assume that the H_n for the lowstep case ($\urcorner$) is the same as in the upstep case ($\ulcorner$).

Instead of (91) we have that

$$\begin{aligned} \Psi^{(r)} = \sum_{i=0}^{u_n} & X^{\ulcorner}_{n,i,t} \circ \theta_{L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)} \circ \theta_{L_{\lambda,t_n}(z_n)} + \\ & 1_{\{\tau_\lambda^{z_n+i-1}(\{t_n\}) \geq 1-\lambda\}} X^{\urcorner}_{n,i,t} \circ \theta_{L_{\lambda,\tau_\lambda^{z_n}(\{t_n\})}(i)-(0,1)} \circ \theta_{L_{\lambda,t_n}(z_n)} \end{aligned}$$

converges to $-h_\pi(P^-)$ in $\mathcal{L}^1(P^-)$, and, by the same argument as before, the convergence remains true when we replace $X^{\ulcorner}_{n,i,t}$ by

$$\big(X^{\ulcorner}_{n,i,t}\big)^- := \log P_0^- \Big[\omega(0,0) \,\Big|\, \omega\big(\widehat{L}^\sharp_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\dots,-i)\big)^-\Big],$$

and $X^{\urcorner}_{n,i,t}$ by

$$\big(X^{\urcorner}_{n,i,t}\big)^- := \log P_0^- \Big[\omega(0,0) \,\Big|\, \omega\big(\{\widehat{L}_{\lambda,\tau_\lambda^{z_n+i}(\{t_n\})}(-1,\dots,-i) - (0,1)\}\big)^-\Big].$$

The behaviour of $Y^{\urcorner}$ and $Y^{\ulcorner}$ under Q_n can be controlled in the same way as before. Applying the law of large numbers twice, that is, to

$$Z^-_{n,i,t} := \big(X^{\urcorner}_{n,i,t}\big)^- - Y_{n,i,t} \qquad \text{and} \qquad Z^-_{n,i,t} = \big(X^{\ulcorner}_{n,i,t}\big)^- - Y_{n,i,t},$$

and using the entropy estimate yields, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} Q_n \left[\frac{1}{|\,\widehat{L}_{k(n)}\,|} \phi_n > \widehat{h}_\pi(P^-,P^+) + \varepsilon \right] = 0. \tag{103}$$

To derive (89) with γ as in (101) it remains to show that

$$\lim_{n\to\infty} \frac{|\,\widehat{L}_{k(n)}\,|}{|\,\partial V_n\,|} = \sum_{r=1}^{R} \frac{1+\lambda_r}{\sqrt{1+\lambda_r^2}} \frac{\text{length}\,\pi_r}{8}. \tag{104}$$

The fraction on the left-hand side can be written as

$$\frac{|\,\widehat{L}_{k(n)}\,|}{|\,\partial V_n\,|} = \frac{|\,\widehat{L}_{k(n)}\,|}{\text{length}\,B_{k(n)}\pi} \, \frac{k(n)\,\text{length}\,B_{k(n)}\pi}{\text{area int}\,B_{k(n)}\pi} \, \frac{\text{area int}\,B_{k(n)}\pi}{k(n)\,|\,\partial V_n\,|}. \tag{105}$$

Applying Lemma 6.7 to the first factor, and using (95) and (96) yields (104).

Finally, to obtain the lower bound Theorem 6.5, we replace the polygon π by the polygon $\widetilde{\pi} = B_{\frac{1}{2}}\pi$ as in the conclusion of the proof of Theorem 6.2, and take the infimum over all polygons $\widetilde{\pi}$ with area int $\widetilde{\pi} = \alpha$. $\square$

Acknowledgements

This work was a part of my PhD thesis, and I would like to thank my supervisor Hans Föllmer for support and guidance. It was his idea to extend the concept of surface entropies to general shapes, and to use them to improve the lower bounds for large deviations of attractive Gibbs measures in the phase-transition region. He also suggested constructing the specific surface entropies first for lines, then for polygons, and finally for curves. I am glad to acknowledge as well my debt to Hans-Otto Georgii for noting and correcting mistakes in the constants in the large deviation bounds (Theorems 4.3.2 and 4.3.5), and in the formula for the specific contour entropy (Theorem 3.3.3). Furthermore, I am grateful to Dima Ioffe for interesting discussions.

References

[1] J. Brettschneider. *Shannon-MacMillan theorems for random fields along curves and lower bounds for surface-order large deviations.* PhD thesis, Humboldt University Berlin, 2000.

[2] J. Brettschneider. Mixing properties for a class of skew products and uniform convergenc in ergodic theorems. EURANDOM *Technical Report*, 2001.

[3] F. Comets. Grandes deviations pour des champs de gibbs sur F^d. *C. R. Acad. Sci., Paris, Ser. I*, 303:511–513, 1986.

[4] Deuschel, J.-D. and Stroock, D.W. *Large deviations.* Academic Press Inc., Boston, MA, 1989.

[5] R. Dobrushin, R. Kotecky, and S. Shlosman. *Wulff construction. A global shape from local interaction*, volume 104 of *Translations of mathematical mongraphs.* AMS, 1992.

[6] M.D. Donsker and S.R.S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. II. *Commun. pure appl. Math.*, 28:279–301, 1975.

[7] H. Föllmer. On entropy and information gain in random fields. *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, 26:207–217, 1973.

[8] H. Föllmer. On the global Markov property. In *Quantum fields – algebras, processes, Proc. Symp., Bielefeld 1978*, pages 293–302. Springer-Verlag, 1980.

[9] H. Föllmer and S. Orey. Large deviations for the empirical field of a Gibbs measure. *Ann. Probab.*, 16(3):961–977, 1988.

[10] H. Föllmer and M. Ort. Large deviations and surface entropy for Markov fields. *Astérisque*, 288:173–190, 1988.

[11] H.-O. Georgii. *Gibbs measures and phase transitions.* W. de Gruyter, Berlin, 1988.

[12] D. Ioffe. Large deviations for the 2D Ising model: a lower bound without cluster expansion. *J. Stat. Phys.*, 74:411–432, 1994.

[13] D. Ioffe. Exact large deviation bounds up to T_c for the Ising model in two dimensions. *Probab. Theory Relat. Fields*, 102(3):313–330, 1995.

[14] R. B. Israel. Some examples concerning the global Markov property. *Commun. Math. Phys.*, 105:669–673, 1986.

[15] U. Krengel. *Ergodic theorems.* W. de Gruyter, Berlin, 1985.

[16] O. E. Lanford, III and D. Ruelle. Observables at infinity and states with short range correlations in statistical mechanics. *Comm. Math. Phys.*, 13:194–215, 1969.

[17] J. Milnor. Directional entropies of cellular automaton-maps. In *Dirorderd systems and biological organization (Les Houches, 1985)*, number 20 in NATO Adv. Sci. Inst. Ser. F: Comut. Systems Sci., pages 113–115, New York, Heidelberg, Berlin, 1986.

[18] J. Milnor. On the entropy geometry of cellular automata. *Complex Syst.*, 2(3):357–385, 1988.

[19] S. Olla. Large deviations for Gibbs random fields. *Probab. Theory Relat. Fields*, 77(3):343–357, 1988.

[20] K.K. Park. Continuity of directional entropy. *Osaka J. Math.*, 31(3):613–628, 1994.

[21] K.K. Park. Continuity of directional entropy for a class of Z^2-actions. *J. Korean Math. Soc.*, 32, No.3:573–582, 1995.

[22] K.K. Park. Entropy of a skew product with a $\mathbb{Z}^2$-action. *Pac. J. Math*, 172(1):227–241, 1996.

[23] K.K. Park. On directional entropy functions. *Israel J. Math.*, 113:243–267, 1999.

[24] C. Preston. *Random fields*, volume 534 of *LNM.* Springer-Verlag, New York, Heidelberg, Berlin, 1976.

[25] R. Schonmann. Second order large deviation estimates for ferromagnetic systems in the phase coexistence region. *Commun. Math. Phys.*, 112(3):409–422, 1987.

[26] Ya.G. Sinai. An answer to a question by J. Milnor. *Comment. Math. Helv.*, 60:173–178, 1985.

[27] Ya.G. Sinai. *Topics in ergodic theory.* Princeton University Press, Princeton, NJ, 1994.

[28] J.-P. Thouvenot. Convergence en moyenne de l'information pour l'action de Z^2. *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, 24:135–137, 1972.

[29] H. V. Weizsäcker. Exchanging the order of taking suprema and countable intersections of σ-algebra. *Ann. Inst. Henri Poincaré B*, 19(1):91–100, 1983.