HANS FÖLLMER
MARIANNE ORT

Large deviations and surface entropy for Markov fields

Astérisque, tome 157-158 (1988), p. 173-190
<http://www.numdam.org/item?id=AST_1988__157-158__173_0>

Société Mathématique de France
Astérisque 157-158 (1988)

Large deviations and surface entropy for Markov fields

by Hans FÖLLMER and Marianne ORT

1. Introduction

Let P be a probability measure on a product space $\Omega = S^I$, where I denotes the d-dimensional lattice Z^d, and where S is some finite state space. We assume that P is ergodic with respect to the group of shift transformations θ_i $(i \in I)$ defined by $(\theta_i \omega)(k) = \omega(i+k)$. For $n \geq 1$, let V_n denote the box of all sites $i = (i_1, .., i_d) \in I$ with $|i_l| \leq n$, and let R_n denote the corresponding empirical field

$$R_n(\omega) = \frac{1}{|V_n|} \sum_{i \in V_n} \delta_{\theta_i \omega} .$$

We are interested in large deviations of the empirical field from its ergodic behavior

$$\lim_{n \to \infty} R_n(\omega) = P \qquad P - a.s., \tag{1.1}$$

i.e., in the probability of an event $\{R_n \in A\}$, where A is a set of probability measures on Ω such that P is not contained in the closure of A.

If P is a Gibbs measure with respect to some regular interaction potential U, then

$$P[R_n \in A] \sim \exp[-|V_n| \inf h(Q, P)], \tag{1.2}$$

where the infimum is taken over all stationary measures belonging to A, and where $h(Q,P)$ denotes the specific relative entropy of Q with respect to P; cf. [1, 8, 10]. But the variational principle for Gibbs measures says that $h(Q,P)$ is equal to 0 if and only if both P and Q belong to the same class $\mathcal{G}(U)$ of Gibbs measures specified by U. Thus, the infimum in (1.2) may be equal to 0 in the case of a phase transition $|\mathcal{G}(U)| > 1$. In such a case, the description (1.2) of large deviations needs further refinement. If P has the local Markov property, then it is natural to replace the volume $|V_n|$ in (1.2) by the surface area $|\partial V_n|$ of V_n. For the Ising model, and for large deviations on level (1) where the empirical field R_n is replaced by the average of the values $\omega(i)$ $(i \in V_n)$, it is shown in [14] that lower and upper bounds in terms of surface area do exist. In this paper, our purpose is to derive explicit lower bounds for large deviations of the empirical field. These bounds involve a new entropy quantity $s(Q,P)$, which is obtained by computing relative entropies on surfaces.

The key to the lower bound in (1.2) is a Shannon-McMillan theorem for the entropy $h(P)$ of a stationary random field P and its extension to the relative entropy $h(Q,P)$. This is recalled in section 2 and motivates our approach to the critical case. In section 3 we consider random fields P which satisfy a strong form of the 0-1 law on asymptotic events. We introduce the surface entropy $s(P)$ and prove the corresponding Shannon-McMillan theorem. Its extension to the relative surface entropy $s(Q,P)$ in section 4 is more delicate; here we need further regularity properties of P. In addition to the Markov property, we assume that the interaction U is attractive and that P is the maximal Gibbs measure $P^+ \in \mathcal{G}(U)$. Using monotonicity arguments, we obtain a lower bound of the form

$$\text{(1.3)} \qquad \liminf_{n\to\infty} \frac{1}{|\partial V_n|} \log P^+[R_n \in A] \ge - \inf_{\alpha: P_\alpha \in A} \alpha^{1-1/d} s(P^-, P^+)$$

which involves the mixtures $P_\alpha = \alpha P^- + (1-\alpha)P^+$ of the maximal Gibbs measure P^+ and the minimal Gibbs measure P^-. In the special case of the two-dimensional Ising model, these results on lower bounds form one part of [12]; the other part of [12] contains upper bounds in terms of $s(P^-, P^+)$.

It is a pleasure to thank R. Durrett and R. Schonmann for a useful comment which led to an improved version of (1.3).

2. Large deviations and entropy

Let $\mathcal{M}(\Omega)$ denote the compact space of probability measures on $\Omega = S^I$. For $J \subseteq I$ we denote by ω_J the restriction of $\omega \in \Omega$ to J, by $\mathcal{F}_J$ the σ-field generated by the map $\omega \longrightarrow \omega_J$, and by P_J the distribution of ω_J under $P \in \mathcal{M}(\Omega)$. For a stationary measure $P \in \mathcal{M}(\Omega)$, the *specific entropy* is given by

$$h(P) = \lim_{n \to \infty} \frac{1}{|V_n|} H(P_{V_n}) \tag{2.1}$$

and satisfies

$$h(P) = \int H(P_0[\,\cdot \mid \mathcal{F}_{\{i|i<0\}}](\omega))P(d\omega) \tag{2.2}$$

where " $<$ " denotes the lexicographical order on Z^d and $P_0[\,\cdot \mid \mathcal{F}_J]$ is the conditional distribution of $\omega_{\{0\}}$ with respect to $\mathcal{F}_J$ and P. Moreover, there is a Shannon-McMillan theorem behind the existence of the limit in (2.1):

$$\lim_{n \to \infty} \frac{1}{|V_n|} \log P[\omega_{V_n}] = -E[H(P_0[\,\cdot \mid \mathcal{F}_{\{i|i<0\}}]) \mid \mathcal{J}] \qquad \text{in } L^1(P) \tag{2.3}$$

and even $P-a.s.$, where $P[\omega_V]$ denotes the measure of the cylinder set determined by ω_V, and where $\mathcal{J}$ is the σ-field of shift-invariant sets ; cf. [2, 15, 9, 5, 11].

(2.4) **Remark.** For two probability measures μ and ν on some finite set E we use the notation

$$H(\mu) = -\sum_{x \in E} \mu(x) \log \mu(x)$$

for the *entropy* of μ and

$$H(\mu,\nu) = \sum_{x \in E} \mu(x) \log \frac{\mu(x)}{\nu(x)}$$

($:= \infty$ if $\mu \not\ll \nu$) for the *relative entropy* of μ with respect to ν.

Now suppose that P is a stationary Gibbs measure; the precise definition will be given below. Then the *specific relative entropy*

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

exists for any stationary measure $Q \in \mathcal{M}(\Omega)$, and the Shannon-McMillan theorem above can be extended as follows:

$$\lim_{n\to\infty} \frac{1}{|V_n|} \log \frac{Q[\omega_{V_n}]}{P[\omega_{V_n}]} = h(Q_\omega, P) \qquad \text{in } L^1(Q) \tag{2.5}$$

and even Q-a.s., where Q_ω denotes the ergodic component of Q associated to ω via $\mathcal{J}$; cf. [5, 11, 7]. This relative entropy governs the large deviations in (1.2). The precise formulation is as follows [1, 8, 10]:

(2.6) **Theorem.** *If $A \subseteq \mathcal{M}(\Omega)$ is open then*

$$\liminf_{n\to\infty} \frac{1}{|V_n|} \log P[R_n \in A] \geq -\inf h(Q,P)\,, \tag{2.7}$$

where the infimum is taken over all stationary measures $Q \in A$; the corresponding upper bound holds if A is closed.

In the proof of the lower bound (2.7), the key idea is to replace P by a new measure Q such that the large deviation under P becomes normal behavior under Q. More precisely, let $Q \in A$ be ergodic. Since A is open, the ergodic theorem implies

$$\lim_{n\to\infty} Q[R_n \in A] = 1 \tag{2.8}$$

It is no loss of generality to assume that $A \in \mathcal{F}_{V_p}$ for some finite p, and to redefine R_n as the empirical field over V_{n-p} so that $\{R_n \in A\} \in \mathcal{F}_{V_n}$. Let ϕ_n denote the density of Q with respect to P on $\mathcal{F}_{V_n}$ which appears in (2.5). Then we can write

$$\begin{aligned} P[R_n \in A] &\geq P[R_n \in A, \frac{1}{|V_n|} \log \phi_n \leq h(Q, P) + \epsilon] \\ &\geq \exp(-h(Q,P) - \epsilon)\, Q[R_n \in A, \frac{1}{|V_n|} \log \phi_n \leq h(Q,P) + \epsilon], \end{aligned}$$

and so the lower bound

$$\liminf_{n \to \infty} \frac{1}{|V_n|} \log P[R_n \in A] \geq -h(Q,P) \tag{2.9}$$

follows from (2.8) and from the Shannon-McMillan theorem (2.5). In order to complete the proof of (2.7), one has to approximate any stationary measure Q by ergodic measures Q_n in such a way that $\lim_n h(Q_n, P) = h(Q,P)$; cf. [8]. Here again, the assumption that P is a Gibbs measure comes in. Let us now recall the precise definition.

Let $U = (U_V)_{V finite}$ be a stationary interaction potential with

$$\sum_{V \ni 0} \|U_V\| < \infty;$$

cf. [13]. For $\xi, \eta \in \Omega$ we define the conditional energy of ξ on V given η on V^c as

$$E_V(\xi \mid \eta) = \sum_{W \cap V \neq \emptyset} U_W(\zeta)$$

where $\zeta_V = \omega_V$ and $\zeta_{I-V} = \eta_{I-V}$.

(2.10) **Definition.** P is called a *Gibbs measure* with respect to U, and we write $P \in \mathcal{G}(U)$, if for any finite $V \subseteq I$ the conditional distribution of ω_V under P with respect to $\mathcal{F}_{I-V}$ is given by

$$P[\omega_V = \xi_V \mid \mathcal{F}_{I-V}](\eta) = \pi_V(\xi \mid \eta) := \frac{1}{Z_V(\eta)} \exp[-E_V(\xi \mid \eta)], \tag{2.10}$$

with normalizing factor $Z_V(\eta)$.

(2.11) **Remark.** If U is a nearest neighbor potential then any $P \in \mathcal{G}(U)$ has the *local Markov property*, i.e.,

$$E[\phi \mid \mathcal{F}_{I-V}] = E[\phi \mid \mathcal{F}_{\partial V}] \tag{2.12}$$

for any finite $V \subset I$ and for any $\mathcal{F}_V$-measurable $\phi \geq 0$, where the *boundary* ∂V is defined as the set of all sites $i \in I - V$ which have distance 1 to V. In that case, P is also called a *Markov field*.

The variational principle for Gibbs measures of Lanford and Ruelle can now be stated as follows: Among all stationary measures $Q \in \mathcal{M}(\Omega)$, the Gibbs measures in $\mathcal{G}(U)$ can be characterized by the fact that they have relative entropy

$$h(Q, P) = 0$$

with respect to the given Gibbs measure $P \in \mathcal{G}(U)$; cf. [5, 13].

Let us now reconsider the description (2.6) of large deviations in the case of a phase transition $|\mathcal{G}(U)| \neq 1$. It may then happen that A contains some stationary measure $Q \in \mathcal{G}(U)$ even though P is not contained in the closure of A. Thus, the right side of (2.7) is equal to 0, and so we need a more refined description. The key idea above for proving a lower bound is still valid; the point is that we need a refined version of the Shannon-McMillan theorem. Consider two stationary Gibbs measures P, $Q \in \mathcal{G}(U)$, and assume that the local Markov property (2.12) holds. Since both P and Q have the same conditional distribution on S^V given the configuration on the boundary ∂V, we have

$$H(Q_{V\cup\partial V}, P_{V\cup\partial V}) = H(Q_{\partial V}, P_{\partial V})\,.$$

It is therefore natural to look for a Shannon-McMillan theorem for the specific entropy quantities

$$\frac{1}{|\partial V_n|} H(P_{\partial V_n}) \qquad \text{and} \qquad \frac{1}{|\partial V_n|} H(Q_{\partial V_n}, P_{\partial V_n})$$

on surfaces rather than volumes.

3. A Shannon-McMillan theorem for surface entropy

If P is ergodic then P satisfies a 0-1 law on the σ-field $\mathcal{J}$ of shift-invariant events. $\mathcal{J}$ is contained, modulo P, in the σ-field

$$\mathcal{A} := \bigcap_{V finite} \mathcal{F}_{I-V}$$

of asymptotic events, and if P is also an extreme point in $\mathcal{G}(U)$ then the 0-1 law extends from $\mathcal{J}$ to $\mathcal{A}$. Often one can go even further:

(3.1) **Definition.** Let us say that P satisfies the *strong 0-1 law* if for any $J \subset I$ the σ-field $\mathcal{F}_J$ coincides modulo P with the σ-field

$$\mathcal{F}_J^* := \bigcap_{V finite} \mathcal{F}_{J \cup (I-V)}.$$

(3.2) **Remarks.**1) For $J = \emptyset$ the strong 0-1 law reduces to the 0-1 law on $\mathcal{A}$.

2) The following condition is equivalent to (3.1): For any $J \subset I$, the conditional distribution $P[\cdot|\mathcal{F}_J]$ can be chosen in such a way that, for P-almost all ω, the measure $P_{I-J}[\cdot|\mathcal{F}_J](\omega)$ satisfies a 0-1 law on the σ-field of asymptotic events in S^{I-J}. In this form, the strong 0-1 law has been verified for the maximal and the minimal Gibbs measure with respect to an attractive interaction potential; cf. [6].

3) For a Markov field the strong 0-1 law implies the *global Markov property* , i.e.,

$$E[\phi|\mathcal{F}_{I-J}] = E[\phi|\mathcal{F}_{\partial J}] \tag{3.3}$$

for any $\mathcal{F}_J$-measurable $\phi \geq 0$ and for any $J \subseteq I$. In fact, we may assume that ϕ only depends on the sites in some finite $W \subseteq J$, and then (3.1) implies, P-a.s.,

$$\begin{aligned} E[\phi|\mathcal{F}_{I-J}] = E[\phi|\mathcal{F}^*_{I-J}] &= \lim_{n\to\infty} E[\phi|\mathcal{F}_{(I-J)\cup(I-V_n)}] \\ &= \lim_{n\to\infty} E[\phi|\mathcal{F}_{\partial J\cup(I-V_n)}] \\ &= E[\phi|\mathcal{F}^*_{\partial J}] = E[\phi|\mathcal{F}_{\partial J}] \end{aligned}$$

where we use martingale convergence in the second and the fourth step, and the local Markov property (2.12) for $V_n \supseteq W$ in the third.

Let us now assume that P satisfies the strong $0-1$ law. For $l = 1,\ldots,d$ let $P^{(l)}$ denote the projection of P on the coordinates in $I^{(l)} = \{i \in I | i_l = 0\}$. $P^{(l)}$ can be viewed as an ergodic random field on Z^{d-1}, and we denote by $h(P^{(l)})$ its specific entropy as defined in section 2. Formula (2.2) for $h(P^{(l)})$, rewritten in terms of P, takes the form

$$h(P^{(l)}) = \int H(P_0[\,\cdot\,|\mathcal{F}^{(l)}](\omega))P(d\omega) \tag{3.4}$$

where $\mathcal{F}^{(l)}$ denotes the σ-field generated by those coordinates $i \in I^{(l)}$ which precede 0 in the lexicographical order on $I^{(l)} \cong Z^{d-1}$.

The following theorem introduces the *surface entropy* $s(P)$ of P, computes $s(P)$ in terms of the specific entropies $h(P^{(l)})$, and provides the corresponding Shannon-McMillan theorem:

(3.5) **Theorem.** *The surface entropy*

$$s(P) := \lim_{n\to\infty} \frac{1}{|\partial V_n|} H(P_{\partial V_n}) \tag{3.6}$$

exists and satisfies

$$(3.7) \qquad s(P) = \frac{1}{d} \sum_{l=1}^{d} h(P^{(l)}).$$

Moreover,

$$(3.8) \qquad s(P) = - \lim_{n \to \infty} \frac{1}{|\partial V_n|} \log P[\omega_{\partial V_n}] \qquad \text{in } L^1(P).$$

Proof. In order to keep the notation simple, let us consider the case $d = 2$. For a box V_n we split $P[\omega_{\partial V_n}]$ into a product of successive conditional probabilities along the four sides, and on each side in increasing order; for $d > 2$ we would use the lexicographical order on each side. Thus,

$$\log\, P[\omega_{\partial V_n}] = \sum_{t \in \partial V_n} \log\, P_0[\omega(0)|\omega_{L(k(n,t)) \cup W(n,t)}] \circ \theta_t$$

where we put

$$L(k) = \{i \in I^{(l)} | i < 0\} \cap V_k$$

for those t which lie on a side parallel to $I^{(l)}$,

$$k(n,t) = \min(t_l + n - 1, n - t_l - 1),$$

and where $W(n,t)$ is some finite set contained in $V^c_{k(n,t)}$.

In 2) we are going to show that

$$(3.9) \qquad \lim_k \sup_{W \subseteq V_k^c} \|P_0[\omega(0)|\omega_{L(k) \cup W}] - P_0[\omega(0)|\mathcal{F}^{(l)}](\omega)\| = 0$$

in $L^1(P)$. As in the usual proof of the Shannon-McMillan theorem, one can now pass to the $L^1(P)$-convergence of the corresponding logarithmic terms; for a Gibbs measure P, the conditional probabilities are bounded away from

0 anyway. On any of the four sides, say on $\{(t_1, t_2)| -n \leq t_1 < n, t_2 = n\}$, the contribution to the n-th term on the right side of (3.8) can thus be reduced to the form

$$(3.10) \qquad -\frac{1}{2d}\Big[\frac{1}{2n} \sum_{-n\leq t_1<n} \log\, P_0[\omega(0)|\mathcal{F}^{(1)}] \circ \theta_{t_1}\Big] \circ \theta_n .$$

Applying the $(d-1)$-dimensional ergodic theorem as in the usual proof of the Shannon-McMillan theorem for $d = 2$, and as in [5] for $d > 2$, we obtain the convergence

$$-\lim_n \frac{1}{2n} \sum_{-n\leq t_1<n} \log\, P_0[\omega(0)|\mathcal{F}^{(1)}] \circ \theta_{t_1} = -E[\log\, P_0[\,\cdot\,|\mathcal{F}^{(1)}]] = h(P^{(1)})$$

in $L^1(P)$. This implies the convergence of (3.10) to $(2d)^{-1}h(P^{(1)})$. Summing up over the $2d$ sides we obtain the convergence in $L^1(P)$ of the right side of (3.8) to the right side of (3.7), and this implies (3.8) and (3.7).

2) In order to prove (3.9) it is enough to show

$$\lim_k \sup_{W\subseteq V_k^c} \|P_0[\,s\,|\omega_{L(k)\cup W}] - P_0[\,s\,|\mathcal{F}^{(l)}](\omega)\| = 0$$

for fixed $s \in S$. But this is of the form (3.12) below, with $\mathcal{B}_k := \mathcal{F}_{L(k)}$, $\mathcal{C}_k := \mathcal{F}_{L(k)\cup W}$, $\mathcal{B}_k^* := \mathcal{F}_{(I-V_k)\cup L(k)}$, and $\mathcal{B}_\infty = \mathcal{F}^{(l)}$. The strong 0-1 law for P implies $\mathcal{B}_\infty = \mathcal{B}_\infty^* := \cap_k \mathcal{B}_k^*$ *mod* P, and so it is enough to apply the following lemma.

(3.11) **Lemma.** Consider σ-fields $\mathcal{B}_k \subseteq \mathcal{B}_k^*$ $(k = 1, 2, \ldots)$ increasing to $\mathcal{B}_\infty$ resp. decreasing to $\mathcal{B}_\infty^*$, and assume that

$$\mathcal{B}_\infty = \mathcal{B}_\infty^* \quad mod\, P .$$

Then

$$(3.12) \qquad \lim_{k\to\infty} \sup_{\mathcal{B}_k\subseteq\mathcal{C}_k\subseteq\mathcal{B}_k^*} \|E[\phi|\mathcal{C}_k] - E[\phi|\mathcal{B}_\infty]\| = 0 \qquad \text{in } L^1(P)$$

for any $\phi \in L^1(P)$.

Proof. Put $\phi_k = E[\phi|\mathcal{B}_k]$ and $\phi_k^* = E[\phi|\mathcal{B}_k^*]$ for $k = 1, \dots, \infty$. If $\mathcal{B}_k \subseteq \mathcal{C}_k \subseteq \mathcal{B}_k^*$ then, by projectivity and contraction,

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

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

4. Large deviations and phase transition

In this section we assume the *Markov property* (2.11). We also assume that the interaction is *attractive* with respect to some total order $\le$ on S, in the sense of [13] (9.7). Let P^+ and P^- denote the maximal and the minimal Gibbs measure; a phase transition occurs if and only if $P^+ \neq P^-$. Both P^+ and P^- are ergodic, and they also satisfy the strong 0-1 law (3.1); cf. [6]. In particular, both P^+ and P^- have the global Markov property (3.3).

(4.1) **Definition.** The *relative surface entropy* of P^- with respect to P^+ is defined as

$$s(P^-, P^+) := \frac{1}{d} \sum_{l=1}^{d} \int H(P_0^-[\,\cdot \mid \mathcal{F}^{(l)}](\omega), P_0^+[\,\cdot \mid \mathcal{F}^{(l)}](\omega)) P^-(d\omega) .$$

Let us first explain the definition. For fixed $J \subseteq I$ and $\omega \in \Omega$ we introduce the conditional local specification on S^{I-J} defined by

$$\pi_V^{J,\omega}(\xi \mid \eta) := \pi_V(\xi \mid \zeta)$$

for any finite $V \subseteq I - J$, where ζ coincides with η on $I - J$ and with ω on J. This specification is again attractive. We denote by $P^+[\,\cdot \mid \mathcal{F}_J](\omega)$ the product of the corresponding maximal Gibbs measure on S^{I-J} with δ_{ω_J} on S^J; as shown in [6], $P^+[\,\cdot \mid \mathcal{F}_J](\,\cdot\,)$ is indeed a conditional probability for P^+ with respect to $\mathcal{F}_J$. In particular, the projection $P_0^+[\,\cdot \mid \mathcal{F}^{(l)}](\omega)$ of $P^+[\,\cdot \mid \mathcal{F}^{(l)}](\omega)$ on the coordinate $0 \in I$ is now defined in a canonical manner which does not involve the null sets of P^+. Thus, the integral with respect to P^- in the definition of $s(P^-, P^+)$ does make sense.

We are going to consider large deviations for $P = P^+$ which are of the form $\{R_n \in A\}$, where A is an open set in $\mathcal{M}(\Omega)$ such that $P^+ \notin \overline{A}$ but

$$A \cap \{P_\alpha | 0 < \alpha \leq 1\} \neq \emptyset,$$

with $P_\alpha := \alpha P^- + (1 - \alpha) P^+$. The following theorem gives a lower bound in terms of the relative surface entropy of P^- with respect to P^+:

(4.2) **Theorem.** *If $A \subseteq \mathcal{M}(\Omega)$ is open then*

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

Proof. 1) The key idea is the same as in the proof of the lower bound (2.7): we are going to switch from $P = P^+$ to a new measure such that the large deviation becomes normal behavior under this new measure, and such that the corresponding Radon-Nikodym density will induce the lower bound (4.2).

Suppose that $P_\alpha \in A$ for some $\alpha \in (0, 1]$. Since A is open, we can choose open neighborhoods A^+ and A^- of P^+ resp. P^- such that

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

Without loss of generality we may assume that $A^+, A^- \in \mathcal{F}_{V_p}$ for some finite $p \geq 1$. Now put $C_n = V_{k(n)}$ and $D_n = V_n - V_{l(n)}$ with $k(n) \leq l(n) \leq n$ such that

$$\lim_{n\to\infty}\frac{|C_n|}{|V_n|}=\alpha,\qquad \lim_{n\to\infty}\frac{|D_n|}{|V_n|}=1-\alpha.$$

For $\alpha=1$ we take $C_n=V_n$, for $\alpha<1$ we choose $l(n)$ such that

$$\lim_{n\to\infty}(l(n)-k(n))=\infty. \tag{4.3}$$

Define

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

and

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

with $C_{n,p}=V_{k(n)-p}$ and $D_{n,p}=V_{n-p}-V_{l(n)+p}$. Then

$$\{R_n^-\in A^-\}\in\mathcal{F}_{C_n},\qquad \{R_n^+\in A^+\}\in\mathcal{F}_{D_n}, \tag{4.4}$$

and

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

for large enough n.

2) Let us now introduce the measure

$$Q_n=P_{C_n}^-\otimes P_{I-Cn}^+.$$

In particular, 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,

$$Q_n[\Lambda_n]=P^-[R_n^-\in A^-]\,P^+[R_n^+\in A^+] \tag{4.5}$$

due to (4.4), and the ergodic theorem for P^+ and P^- implies

$$\lim_{n\to\infty}Q_n[\Lambda_n]=1. \tag{4.6}$$

For $c > 0, \epsilon > 0$ and for large enough n we can write

$$\begin{aligned} P^+[R_n \in A] &\geq P^+[\Lambda_n] \\ &\geq \exp[-(c+\epsilon)|\partial V_n|)\, Q_n[\Lambda_n \cap \{\frac{1}{|\partial V_n|} \log \phi_n \leq c+\epsilon\}] \end{aligned}$$

where ϕ_n denotes the density of Q_n with respect to P^+ on $\mathcal{F}_{C_n \cup D_n}$. Due to (4.6), the lower bound

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

follows if c is chosen such that

$$\lim_{n\to\infty} Q_n[\frac{1}{|\partial V_n|} \log \phi_n \leq c+\epsilon] = 1 \tag{4.7}$$

for any $\epsilon > 0$.

3) We want to show that (4.7) holds with $c = \alpha^{1-1/d}\, s(P^-, P^+)$. Since $Q_n = P^+$ on D_n, and since both P^- and P^+ belong to $\mathcal{G}(U)$, we can write

$$\phi_n(\omega) = \frac{P^-[\omega_{C_n}]}{P^+[\omega_{C_n} \mid \omega_{D_n}]} = \frac{P^-[\omega_{B_n}]}{P^+[\omega_{B_n} \mid \omega_{D_n}]}$$

where $B_n = \partial V_{k(n)-1}$ is the boundary of the interior of C_n. Going around the sides of B_n and using the lexicographical order on each side as in the proof of (3.5) we obtain

$$\frac{1}{|\partial V_n|} \log \phi_n(\omega) = \frac{1}{|\partial V_n|} \sum_{t\in B_n} Z_{n,t} \circ \theta_t \tag{4.8}$$

with

$$Z_{n,t} = X_{n,t} - Y_{n,t}$$

and

$$\begin{aligned} X_{n,t}(\omega) &= \log P_0^-[\omega(0) \mid \omega_{L(k(n,t))\cup A(n,t)}], \\ Y_{n,t}(\omega) &= \log P_0^+[\omega(0) \mid \omega_{L(k(n,t))\cup B(n,t)}]. \end{aligned}$$

For a site t which lies on a side parallel to $I^{(l)} = \{i \in I \mid i_l = 0\}$, $k(n,t)$ denotes the minimum of $l(n) - k(n)$ and the distance from t to the boundary of that side, $L(k)$ is the set of all sites in $I^{(l)}$ which precede 0 in the lexicographical order and which also belong to V_k, and $A(n,t)$ and $B(n,t)$ are finite sets such that

$$A(n,t) \subseteq B(n,t) \subseteq V(k(n,t))^c.$$

$A(n,t)$ is obtained by shifting a subset of $B_n \subseteq C_n$, and so the behavior of $X_{n,t}$ under Q_n is the same as under P^-. But the proof of (3.5) shows that

$$\lim_{n\to\infty} \frac{1}{|B_n|} \sum_{t\in B_n} X_{n,t} \circ \theta_t = -s(P^-) \qquad \text{in } L^1(P^-). \tag{4.9}$$

4) We still have to control the behavior of $Y_{n,t}$ under Q_n, and it is here that we are going to use the monotonicity properties of an attractive interaction. To begin with, the proof of (3.5) shows that the convergence in (4.9) is not changed if we replace $X_{n,t}$ by

$$X_{n,t}^- = \log P_0^-[\omega(0) \mid \omega_{L(k(n,t))}^-]$$

where $\omega_{L(k)}^-$ is equal to ω on $L(k)$ and assumes the minimal state in S outside of V_k. Put

$$Z_{n,t}^- = X_{n,t}^- - Y_{n,t}.$$

Now we use the law of large numbers for martingales with bounded increments in its L^2-form in order to replace

$$\frac{1}{|B_n|} \sum_{t\in B_n} Z_{n,t}^- \circ \theta_t$$

by

$$\frac{1}{|B_n|} \sum_{t\in B_n} E_{Q_n}[Z_{n,t}^- \circ \theta_t \mid \mathcal{A}_{n,t}]$$

where $\mathcal{A}_{n,t}$ is the σ-field generated by the sites in D_n and by those sites in B_n which precede t in our ordering of B_n. But these conditional expectations

$$E_{Q_n}[Z^-_{n,t} \circ \theta_t \mid \mathcal{A}_{n,t}](\omega)$$

can be written as

$$H(P_0^-[\,\cdot \mid \omega^-_{L(k(n,t))}], P_0^+[\,\cdot \mid \omega_{L(k(n,t))\cup B(n,t)}]) \circ \theta_t. \tag{4.10}$$

The measure

$$\mu := P_0^+[\,\cdot \mid \omega_{L(k(n,t))\cup B(n,t)}]$$

is larger than the measure

$$\nu := P_0^-[\,\cdot \mid \omega^-_{L(k(n,t))}],$$

in the sense that the density $d\mu/d\nu$ is an increasing function with respect to the order on S; cf. [13] Th. (9.4). The relative entropy in (4.10) increases if we replace μ by the even larger measure

$$\lambda := P_0^+[\,\cdot \mid \omega^+_{L(k(n,t))}]$$

where $\omega^+_{L(k)}$ is defined in analogy to $\omega^-_{L(k)}$:

$$\begin{aligned} H(\nu,\lambda) &= H(\nu,\mu) + \int \log \frac{d\mu}{d\lambda}\, d\nu \\ &\geq H(\nu,\mu) + \int \log \frac{d\mu}{d\lambda}\, d\mu \\ &= H(\nu,\mu) + H(\mu,\lambda) \\ &\geq H(\nu,\mu)\,. \end{aligned}$$

Now we use the *sure* convergence of $P_0^+[\,\cdot \mid \omega^+_{L(k(n,t))}]$ to $P_0^+[\,\cdot \mid \mathcal{F}^{(l)}](\omega)$ in order to obtain

$$\lim_{n\to\infty} \frac{1}{|B_n|} \sum_{t\in B_n} H(P_0^-[\,\cdot \mid \omega^-_{L(k(n,t))}], P_0^+[\,\cdot \mid \omega^+_{L(k(n,t))}]) \circ \theta_t = s(P^-, P^+)$$

in $L^1(P^-)$. Putting all these steps together, and passing from convergence in $L^1(P^-)$ to stochastic convergence in terms of Q_n, we obtain

$$\lim_{n\to\infty} Q_n[\frac{1}{|B_n|}\sum_{t\in B_n} Z_{n,t}\circ\theta_t \geq s(P^-,P^+)+\epsilon] = 0$$

for any $\epsilon > 0$. Since

$$\lim_{n\to\infty} \frac{|B_n|}{|\partial V_n|} = \alpha^{1-1/d},$$

this is equivalent to (4.7) with $c = \alpha^{1-1/d}s(P^-,P^+)$.

5. References

[1] F. Comets: Grandes déviations pour des champs de Gibbs sur Z^d. *C.R.A.S. Paris* **303**, sér. 1, no. 11, 511 (1986)

[2] J. P. Conze: Entropie d'un groupe abélien de transformations. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **25**, 11-30 (1972)

[3] M. D. Donsker and S. R. S. Varadhan: Asymptotic Evaluation of certain Markov Expectations for Large Time IV. *Comm. Pure Appl. Math.* **36**, 183-212 (1983)

[4] R. S. Ellis: Entropy, Large Deviations and Statistical Mechanics. *Grundlehren* **271**, Springer (1985)

[5] H. Föllmer: On Entropy and Information Gain in Random Fields. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **26**, 207-217 (1973)

[6] H. Föllmer: On the Global Markov Property. In: *Quantum fields, algebras, processes (ed. L. Streit)*, 207-217, Springer (1980)

[7] H. Föllmer: Random Fields and Diffusion Processes. In: *Ecole d'Eté de Probabilités de St. Flour* **XVI**–1986, *Lecture Notes in Mathematics*, Springer (to appear)

[8] H. Föllmer and S. Orey: Large Deviations for the Empirical Field of a Gibbs Measure. Annals of *Probability* (to appear)

[9] J. Fritz: Generalization of McMillan's Theorem to Random Set Functions. *Studia Sci. Math. Hung.* **5,** 369-394 (1970)

[10] S. Olla: Large Deviations for Gibbs Random Fields. *Probability Th. Rel. Fields* (to appear)

[11] D. Ornstein and B. Weiss: The Shannon-McMillan-Breiman Theorem for a Class of Amenable Groups. *Israel J. Math.* **44,** 53-60 (1983)

[12] M. Ort: Grosse Abweichungen und Oberflächen-Entropie im Ising Modell. *Diss. ETHZ* (1988)

[13] C. Preston: Random Fields. *Lecture Notes in Mathematics* **534,** Springer (1976)

[14] R. Schonmann: Second Order Large Deviation Estimates for Ferromagnetic Systems in the Phase Transition Region. *Comm. Math. Physics* **112,** 409-422 (1987)

[15] J. P. Thouvenot: Convergence en moyenne de l'information pour l' action de Z^2. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **24,** 134-137 (1972)

Hans Föllmer
Marianne Ort

Mathematikdepartement
ETH-Zentrum

CH-8092 Zürich