Commun. Math. Phys. 164, 433–454 (1994)

Communications in Mathematical Physics

Large Deviations for $\mathbb{Z}^d$-Actions

A. Eizenberg★, Y. Kifer★, B. Weiss★
Institute of Mathematics The Hebrew University, Jerusalem, Israel

Received: 13 January 1993

Abstract: We establish large deviations bounds for translation invariant Gibbs measures of multidimensional subshifts of finite type. This generalizes [FO] and partially [C, O, and B], where only full shifts were considered. Our framework includes, in particular, the hard-core lattice gas models which are outside of the scope of [FO, C, O, and B].

1. Introduction

In [R1] Ruelle rewrote a part of the general theory of statistical mechanics for the case of a $\mathbb{Z}^d$-action, $d \geqq 1$ on a compact metric space Ω satisfying expansiveness and the specification. The main model for which one constructs translation invariant Gibbs states (see [R2]) consists of a finite set Q taken with the discrete topology and called the alphabet (which may represent, for instance, the spin values etc.), the set $Q^{\mathbb{Z}^d}$ considered with the product topology (making it compact) of all maps (configurations) $\omega : \mathbb{Z}^d \to Q$, the shifts $\theta_m, m \in \mathbb{Z}^d$ of $Q^{\mathbb{Z}^d}$ acting by the formula $(\theta_m \omega)_n = \omega_{n+m}$, where $\omega_k \in Q$ is the value of $\omega \in Q^{\mathbb{Z}^d}$ on $k \in \mathbb{Z}^d$, and a closed in the product topology subset Ω of $Q^{\mathbb{Z}^d}$ called the space of (permissible) configurations which is supposed to be shift invariant, i.e. $\theta_m \Omega = \Omega$ for every $m \in \mathbb{Z}^d$. The pair (Ω, θ) is called a subshift and if $\Omega = Q^{\mathbb{Z}^d}$ it is called the full shift. The construction in [R2] assumes, in fact, that (Ω, θ) is a subshift of finite type (see [Sh]) which means that there exist a finite set $F \subset \mathbb{Z}^d$ and a set $\Xi \subset Q^F$ such that

$$\Omega = \Omega_{(F,\Xi)} = \{\omega \in Q^{\mathbb{Z}^d} : (\theta_m \omega)_F \in \Xi \text{ for every } m \in \mathbb{Z}^d\}, \tag{1.1}$$

where $(\omega)_R = \omega_R$ denotes the restriction of $\omega \in Q^{\mathbb{Z}^d}$ to $R \subset \mathbb{Z}^d$. The set $\Xi \subset Q^F$ is the collection of permissible (allowed) words or configurations on F.

★ Partially supported by US-Israel BSF. Partially sponsored by the Edmund Landau Center for research in Mathematical Analysis, supported by the Minerva Foundation (Germany).

Recall that a sequence $\Lambda_n \subset \mathbb{Z}^d, n = 1, 2, \ldots$ of finite sets tend to infinity in the sense of van Hove, written $\Lambda_n \nearrow \infty$, if $|\Lambda_n| \to \infty$ and $|\partial \Lambda_n| = o(|\Lambda_n|)$, where $|\Lambda|$ denotes the number of points in Λ and $\partial\Lambda$ is the boundary of Λ, i.e. set of points which have neighbors both inside and outside of Λ. If $a = (a_1, \ldots, a_d) \in \mathbb{Z}^d, a_i > 0, 1 \leqq i \leqq d$ and $\Lambda(a) = \{i \in \mathbb{Z}^d : 0 \leqq i_k < a_k, 1 \leqq k \leqq d\}$ then $\Lambda(a) \nearrow \infty$ provided $a_1, \ldots, a_d \to \infty$ which will be written as $a \to \infty$. If limits are the same over all sequences $\Lambda_n \nearrow \infty$ we will write this limit over $\Lambda \nearrow \infty$. We will need also the notion of the weak specification from [R1] which means in our circumstances that there exists $N > 0$ such that for any subsets $R_i \subset \mathbb{Z}^d$ which are N apart and for any permissible configurations ξ_i on R_i one can find $\omega \in \Omega$ such that $\omega_{R_i} = \xi_i$. By [R2] for any shift invariant summable interaction potential Φ defined on finite configurations there exists a Gibbs measure (Gibbs state) on Ω whose conditional probabilities for given configurations outside of finite sets are determined by Φ (see Sect. 2 for the definitions). The probability measures

$$\zeta_\omega^\Lambda = |\Lambda|^{-1} \sum_{m \in \Lambda} \delta_{\theta_m \omega}, \quad \omega \in \Omega, \quad \Lambda \subset \mathbb{Z}^d \text{ is finite}, \tag{1.2}$$

where δ_ω is the unit mass at ω, are called usually the occupational measures.

Theorem A. *Let (Ω, θ) be a subshift of finite type satisfying the weak specification and μ be a Gibbs measure for a shift invariant summable interaction potential Φ. Then for any closed subset K of the space $\mathscr{P}(\Omega)$ of probability measures on Ω,*

$$\limsup_{\Lambda \nearrow \infty} |\Lambda|^{-1} \log \mu\{\omega : \zeta_\omega^\Lambda \in K\} \leqq - \inf_{\nu \in K} I^\Phi(\nu) \tag{1.3}$$

and for any open $G \subset \mathscr{P}(\Omega)$,

$$\liminf_{\Lambda \nearrow \infty} |\Lambda|^{-1} \log \mu\{\omega : \zeta_\omega^\Lambda \in G\} \geqq - \inf_{\nu \in G} I^\Phi(\nu), \tag{1.4}$$

where

$$I^\Phi(\nu) = \begin{cases} P(A^\Phi) - \nu(A^\Phi) - h_\nu & \text{if } \nu \text{ is shift invariant}, \\ \infty, & \text{otherwise}. \end{cases} \tag{1.5}$$

Here h_ν is the entropy of the subshift (Ω, θ) with respect to ν, $P(g)$ is the pressure of a function g on $\Omega, \nu(g) = \int g d\nu$, and

$$A^\Phi(\omega) = -\sum \{\Phi_R(\omega_R) : R \subset \mathbb{Z}^d \text{ is finite and } 0 \text{ is the first element of } R \text{ in the lexicographic order in } \mathbb{Z}^d\}. \tag{1.6}$$

Remark that following the dynamical systems tradition we denote the entropy by h_ν, though in statistical mechanics it is denoted usually by $s(\nu)$ and is called the mean entropy of ν. In the proof of Theorem A we do not use explicitly that (Ω, θ) is a subshift of finite type and this assumption comes in only via [R2] where it was needed for the construction of the Gibbs states and for the proof of their coincidence with the equilibrium states on which the supremum is attained in the Gibbs variational principle. For finite range potentials Φ Theorem A follows essentially from [DSZ] where a completely different method was employed.

Leaving the precise definitions of the quantities appearing in Theorem A till the next section we remark that Theorem A is a generalization of the main result of [FO] where only full shifts were considered and instead of $\Lambda \nearrow \infty$ the limits were taken only over sequences of increasing cubes. This generalization enables us to include physically important models where certain configurations are not allowed, for instance, the "hard core" model (see [R2]) (called in [S] "the golden mean") where $Q = \{0, 1\}$ and two ones are not allowed in any pair of neighboring sites of $\mathbb{Z}^d$. Of course, one can consider such models in a full shift framework prescribing interaction potentials of certain configurations to be infinity but this approach usually does not help much.

We derive the upper bound (1.3) of large deviations following the approach in [Kil] in view of existence of the thermodynamic type limits for the pressure. The lower bound required in [Kil] the uniqueness of Gibbs measures for a sufficiently large class of potentials. This does not hold true for $\mathbb{Z}^d$-actions with $d \geqq 2$ in view of phase transitions. By this reason, in order to obtain the lower large deviations bound (1.4) we apply a modification of the approach from [FO]. An important part in the proof of (1.4) is played by the following result which has an independent interest and holds true in more general circumstances than Theorems A and C.

Theorem B. *Suppose that $\mathbb{Z}^d$-acts continuously on a compact metric space (Ω, d), the action preserves a probability measure ν on Ω, it satisfies the weak specification (see Sect. 2), and the entropy h_η of the $\mathbb{Z}^d$-action as the function on the space of $\mathbb{Z}^d$-invariant probability measures η is upper semicontinuous at the point $\eta = \nu$. Then there exists a sequence of $\mathbb{Z}^d$-invariant ergodic probability measures ν_n on Ω such that*

$$\nu_n \xrightarrow{w} \nu \text{ and } h_{\nu_n} \to h_\nu \text{ as } n \to \infty ,$$

where $\xrightarrow{w}$ denotes the convergence in the weak topology of $\mathscr{P}(\Omega)$.

Following [Ki2] we will obtain also the bounds for large deviations from the set of measures with maximal entropy for occupational measures sitting on periodic orbits. This means the following: For a given $a \in \mathbb{Z}^d, a_i > 0, i = 1, \ldots, d$ let $\mathbb{Z}^d(a)$ be the subgroup of $\mathbb{Z}^d$ generated by $(a_1, 0, \ldots, 0), \ldots, (0, \ldots, 0, a_d)$. The collection

$$\Pi_a = \{\omega \in \Omega : \mathbb{Z}^d(a)\omega = \omega\} , \tag{1.7}$$

which is clearly finite, is called the set of a-periodic point. Define $\nu_a \in \mathscr{P}(\Omega)$ by

$$\nu_a(\Gamma) = |\Pi_a|^{-1} |\Gamma \cap \Pi_a|, \quad \Gamma \subset \Omega , \tag{1.8}$$

which is the uniform distribution on Π_a, and $\zeta^a : \Pi_a \to \mathscr{P}(\Omega)$ by

$$\zeta^a_\omega = \zeta^{\Lambda(a)}_\omega . \tag{1.9}$$

Theorem C. *Suppose that the conditions of Theorem A are satisfied except that the weak specification is replaced by the strong specification (see* [R1] *and Sect. 2). Then for any closed $K \subset \mathscr{P}(\Omega)$,*

$$\limsup_{a \to \infty} |\Lambda(a)|^{-1} \log \nu_a\{\omega : \zeta^a_\omega \in K\} \leqq - \inf_{\eta \in K} J(\eta) \tag{1.10}$$

and for any open $G \subset \mathscr{P}(\Omega)$,

$$\liminf_{a \to \infty} |\Lambda(a)|^{-1} \log \nu_a\{\omega : \zeta^a_\omega \in G\} \geqq - \inf_{\eta \in G} J(\eta) , \tag{1.11}$$

where

$$J(\eta) = \begin{cases} h_{\text{top}} - h_\eta & \text{if } \eta \text{ is shift invariant} \\ \infty, & \text{otherwise} \end{cases}$$

and $h_{\text{top}} = \sup\{h_\eta : \eta$ *is shift invariant* $\}$ *is the topological entropy of the subshift* (Ω, θ).

The final version of this paper was written while the second author visited the University of North Carolina at Chapel Hill whose support he gratefully acknowledges.

2. Preliminaries and the upper bound

Let $\mathscr{A}$ be the family of all finite nonempty subsets Λ of $\mathbb{Z}^d, (\Omega, \theta)$ be a subshift, and Ω_Λ be the restriction of Ω to Λ. A collection $\Phi = \{\Phi_\Lambda, \Lambda \in \mathscr{A}\}$ of real functions $\Phi_\Lambda : \Omega_\Lambda \to \mathbb{R}$ is called an interaction potential (or just: interaction). Following [R2] we assume that

$$\|\Phi\| = \sum_{\Lambda: 0 \in \Lambda \in \mathscr{A}} |\Phi_\Lambda| < \infty, \tag{2.1}$$

where

$$|\Phi_\Lambda| = \sup_{\xi \in \Omega_\Lambda} |\Phi_\Lambda(\xi)| \tag{2.2}$$

and

$$\Phi_{\Lambda - m}(\theta_m \xi) = \Phi_\Lambda(\xi) \text{ for any } \Lambda \in \mathscr{A} \text{ and } \xi \in \Omega_\Lambda, \tag{2.3}$$

the latter means that we consider only shift invariant interactions. For $\Lambda \subset \mathbb{Z}^d$ set $\Lambda^c = \mathbb{Z}^d \backslash \Lambda$. If $\xi \in \Omega_\Lambda$ and $\eta \in \Omega_{\Lambda^c}$ denote by $\xi \vee \eta$ the point $\zeta \in Q^{\mathbb{Z}^d}$ such that $\zeta_\Lambda = \xi$ and $\zeta_{\Lambda^c} = \eta$. If, in addition, $\xi \vee \eta \in \Omega$ and $\Lambda \subset \mathscr{A}$, then one can define the energy functions

$$U_\Lambda^\Phi(\xi) = \sum_{X \subset \Lambda} \Phi_X(\xi_X), \tag{2.4}$$

$$U_{\Lambda,\eta}^\Phi(\xi) = \sum_{X \in \mathscr{A}: X \cap \Lambda \neq \emptyset} \Phi_X((\xi \vee \eta)_X), \tag{2.5}$$

and the partition functions

$$Z_\Lambda^\Phi = \sum_{\xi \in \Omega_\Lambda} \exp(-U_\Lambda^\Phi(\xi)), \tag{2.6}$$

$$Z_\Lambda^\Phi(\eta) = \sum_{\xi \in \Omega_\Lambda : \xi \vee \eta \in \Omega} \exp(-U_{\Lambda,\eta}^\Phi(\xi)). \tag{2.7}$$

For any $\xi \in \Omega_\Lambda$ set also

$$\Xi_\Lambda(\xi) = \{\omega \in \Omega : \omega_\Lambda = \xi\}. \tag{2.8}$$

Given an interaction Φ we will call $\mu \in \mathscr{P}(\Omega)$ a Gibbs state (or a Gibbs measure) if for any finite $\Lambda \subset \mathbb{Z}^d$ and all $\xi \in \Omega_\Lambda, \eta \in \Omega_{\Lambda^c}$ satisfying $\xi \vee \eta \in \Omega$ μ-almost everywhere

$$\mu(\Xi_\Lambda(\xi) | \mathscr{B}_{\Lambda^c})(\eta) = (Z_\Lambda^\Phi(\eta))^{-1} \exp(-U_{\Lambda,\eta}^\Phi(\xi)), \tag{2.9}$$

where $\mu(\cdot \mid \cdot)$ denotes the conditional probability and $\mathscr{B}_M, M \subset \mathbb{Z}^d$ is the restriction of the Borel σ-field on Ω to Ω_M. This definition is the same as in [F1 and G], and it is equivalent to the definition given in [R2] in view of Theorem 1.8 there.

In particular, (2.9) implies

$$\mu(\Xi_\Lambda(\xi)) = \int_{\{\eta\in\Omega_{\Lambda^c}:\xi\vee\eta\in\Omega\}} \mu_\Lambda(\xi,\eta)d\mu_{\Lambda^c}(\eta)\,, \tag{2.10}$$

where $\mu_\Lambda(\xi,\eta)$ denotes the right-hand side of (2.9) and μ_M, $M\subset\mathbb{Z}^d$ is the projection of $\mu\in\mathscr{P}(\Omega)$ to M.

Recall that a family of finite sets $\Lambda_\gamma\subset\mathbb{Z}^d$ indexed by a directed set Γ tend to infinity in the sense of van Hove (see [R2]) written $\Lambda_\gamma\nearrow\infty$ if

$$\lim_{\gamma\in\Gamma}|\Lambda_\gamma| = \infty \text{ and } \lim_{\gamma\in\Gamma}|(\Lambda_\gamma + a)\backslash\Lambda_\gamma|/|\Lambda_\gamma| = 0 \tag{2.11}$$

for any $a\in\mathbb{Z}^d$. Roughly speaking this means that the "boundary of Λ" becomes negligible in the limit as compared to Λ. In particular, $\Lambda_a\nearrow\infty$ if $a\to\infty$ as in the Introduction. If Γ coincides with the family $\mathscr{A}$ of all nonempty finite subsets of $\mathbb{Z}^d$ ordered by inclusion then we will write just $\Lambda\nearrow\infty$. By Corollary 3.13 from [R2] the limit

$$P^\Phi = \lim_{\Lambda\nearrow\infty}|\Lambda|^{-1}\log Z_\Lambda^\Phi \tag{2.12}$$

exists and is called the pressure. Let $C(\Omega)$ be the space of continuous functions on Ω and $A\in C(\Omega)$. Set

$$Z_\Lambda^*(A) = \sum_{\xi\in\Omega_\Lambda}\exp\left(\sum_{m\in\Lambda}A(\theta_m\omega^\xi)\right) \tag{2.13}$$

where ω^ξ is an arbitrary point of $\Xi_\Lambda(\xi)$, and so the above expression depends on these choices of ω^ξ. Again by Corollary 3.13 from [R2] the limit

$$P(A) = \lim_{\Lambda\nearrow\infty}|\Lambda|^{-1}\log Z_\Lambda^*(A) \tag{2.14}$$

exists, it is independent of choices of ω^ξ, and

$$P(A^\Phi) = P^\Phi\,, \tag{2.15}$$

where A^Φ defined by (1.6) is a continuous function in view of (2.1).

Denote by $\mathscr{P}_I(\Omega)$ the set of θ-invariant probability measures on Ω, i.e. $\nu\in\mathscr{P}_I(\Omega)$ means that $\nu(\theta_m\Gamma) = \nu(\Gamma)$ for any $m\in\mathbb{Z}^d$ and a Borel set $\Gamma\subset\Omega$.

Set

$$H_\Lambda(\nu) = -\sum_{\xi\in\Omega_\Lambda}\nu(\Xi_\Lambda(\xi))\log\nu(\Xi_\Lambda(\xi))\,, \tag{2.16}$$

then for $\nu\in\mathscr{P}_I(\Omega)$ the following limit

$$h_\nu = \lim_{\Lambda\nearrow\infty}|\Lambda|^{-1}H_\Lambda(\nu) = \inf_\Lambda|\Lambda|^{-1}H_\Lambda(\nu) \tag{2.17}$$

exists, is called the (mean) entropy, and it is a nonnegative, affine, and upper semicontinuous function on $\mathscr{P}_I(\Omega)$ (see Theorem 3.10 in [R2]). Moreover ([R2], Theorem 3.12) for all $A\in C(\Omega)$,

$$P(A) = \sup_{\sigma\in\mathscr{P}_I(\Omega)}(\sigma(A) + h_\sigma)\,, \tag{2.18}$$

which is called the variational principle.

Until now we only assumed in this section that (Ω,θ) is a subshift, but further results require (Ω,θ) to be a subshift of finite type. In fact, results of [R2], which we are going to rely on, use another equivalent condition saying that there exists a locally finite collection $\mathscr{F}\subset\mathscr{A}$ (i.e. each $m\in\mathbb{Z}^d$ belongs to a finite number of $\Lambda\in\mathscr{F}$) and a family $\{\Psi_\Lambda,\Lambda\in\mathscr{F}\}$ such that $\Psi_\Lambda\subset Q^\Lambda$, if $\Lambda\in\mathscr{F}$ then $\Lambda-m\in\mathscr{F}$ and $\theta_m\Psi_\Lambda=\Psi_{\Lambda-m}$ for all $m\in\mathbb{Z}^d$, and

$$\Omega=\{\omega\in Q^{\mathbb{Z}^d}:\omega_\Lambda\in\Psi_\Lambda \text{ for all } \Lambda\in\mathscr{F}\}\,. \tag{2.19}$$

If (Ω,θ) is a subshift of finite type defined in the Introduction, then taking $\mathscr{F}=\{F+m,m\in\mathbb{Z}^d\}$ and $\Psi_{F+m}=\theta_{-m}\Xi$ we obtain the above condition with the same Ω as in (1.1). On the other hand, if Ω is defined by (2.19) then taking $F=\bigcup_{\Lambda:0\in\Lambda\in\mathscr{F}}\Lambda$ and $\Xi=\Psi_F$ we obtain the subshift of finite type framework with the same Ω, i.e. both conditions are equivalent.

Following [R2] we will say that a sequence of finite sets $\Lambda_n\subset\mathbb{Z}^d,n=1,2,\ldots$ with $\Lambda_n\nearrow\infty$ satisfies the condition D if there exists another sequence of finite sets $M_n\subset\mathbb{Z}^d,M_n\subset\Lambda_n$ such that

$$\lim_{n\to\infty}|\Lambda_n||M_n|^{-1}=1 \tag{2.20}$$

and for each $\xi,\eta\in\Omega$ there exists $\zeta^{(n)}\in\Omega,n=1,2,\ldots,$ such that $\zeta^{(n)}_{\Lambda_n}=\xi_{\Lambda_n}$ and $\zeta^{(n)}_{M_n^c}=\eta_{M_n^c}$, where again $M^c=\mathbb{Z}^d\backslash M$. In order to prove (1.3) and (1.4) for all sequences $\Lambda_n\nearrow\infty$, in fact for $\Lambda\nearrow\infty$, one has to assume that the condition D is satisfied for all such sequences which is enough for the upper bound (1.3). A large part of the proof of the lower bound (1.4) needs nothing else, as well, but its final step relies on Theorem B which employs the weak specification condition which is a stronger assumption.

We will give the corresponding definitions in the more general framework of Theorem B. The result of action of $m\in\mathbb{Z}^d$ on $\omega\in\Omega$ in this more general set up will be denoted by $m\omega$ which in the subshift case is the same as $\theta_m\omega$. A continuous action is said to satisfy the weak specification if for any $\varepsilon>0$ there exists $N(\varepsilon)>0$ such that for any collection of sets $R_i\subset\mathbb{Z}^d$ that are $N(\varepsilon)$ apart, i.e. for $i\neq j,\xi\in R_i,\eta\in R_j$ one has $\|\xi-\eta\|=\max_k|\xi_k-\eta_k|\geqq N(\varepsilon)$, and any points $\zeta_i\in\Omega$ there is an $\omega\in\Omega$ such that $d(m\omega,m\zeta_i)\leqq\varepsilon$ for all i and all $m\in R_i$. The action of $\mathbb{Z}^d$ is said to satisfy the strong specification if for any $R_i\subset\Lambda(a)$ such that all $R_i+\mathbb{Z}^d(a)$ are $N(\varepsilon)$ apart and any points $\zeta_i\in\Omega$ one can find $\omega\in\Pi_a$ such that $d(m\omega,m\zeta_i)\leqq\varepsilon$ for all i and all $m\in R_i$.

In the more general set up of Theorem B one also has to keep in mind the following expansivity condition on the action of $\mathbb{Z}^d$ on Ω saying that there exists $\delta>0$ such that $d(m\omega,m\tilde\omega)<\delta$ for some $\omega,\tilde\omega\in\Omega$ and all $m\in\mathbb{Z}^d$ implies $\omega=\tilde\omega$. The expansivity is a sufficient condition for the upper semicontinuity of the entropy (see [R1]) and it is, clearly, always satisfied for subshifts with finite alphabets.

Pick up $\beta\in(0,1)$ and define the metric on $\Omega\subset\Omega^{\mathbb{Z}^d}$ by

$$d(\omega,\tilde\omega)=\beta^L,\ \text{where } L=\min\{\|\ell\|:\omega_\ell\neq\tilde\omega_\ell\}\,. \tag{2.21}$$

Then $d(\omega,\tilde\omega)<1$ implies $\omega_0=\tilde\omega_0$, and so if $d(m\omega,m\tilde\omega)<1$ for all $m\in R_i\subset\mathbb{Z}^d$, then $\omega_{R_i}=\tilde\omega_{R_i}$. Thus if ξ_i are permissible configurations on R_i, i.e. $\xi\in\Omega_{R_i}$, then there exist $\zeta^{(i)}\in\Omega$ such that $\zeta^{(i)}_{R_i}=\xi_i$, and if R_i are, say, $N(1/2)$ apart then one can find $\omega \in \Omega$ with $\omega_{R_i} = \zeta^{(i)}_{R_i} = \xi_i$. Thus for subshifts the above definition coincides with the definition of weak specification given in the Introduction and, in particular, the weak specification implies the condition D for all sequences $\Lambda_n \nearrow \infty$ by taking

$$M_n = \{\ell \in \mathbb{Z}^d : \min_{m \in \Lambda_n} \|\ell - m\| \leqq N(1/2)\} . \tag{2.22}$$

Remark that the "hard core" model described in the Introduction is a subshift of finite type satisfying both the weak and the strong specification. The subshift of finite type conditions are checked here trivially and both specifications follow since surrounding any configuration on R_i by zeros we eliminate the influence of this configuration and can continue in any permissible way, and so one can take $N(\varepsilon) = 1$ for all $\varepsilon > 0$.

Next, we will prove the upper bound (1.3). For this one has to assume only that (Ω, θ) is a subshift of finite type and that the condition D is satisfied for all sequences $\Lambda_n \nearrow \infty$. Denote the set of all Gibbs measures for an interaction Φ by K^Φ. By Theorem 1.9 from [R2], $K^\Phi \neq \emptyset, K^\Phi$ is closed, compact, and it is a Choquet simplex. Let $\mu \in K^\Phi$ then by Proposition 4.4 from [R2] for any $g \in C(\Omega)$,

$$\lim_{\Lambda \nearrow \infty} |\Lambda|^{-1} \log \mu \left(\exp \sum_{m \in \Lambda} g \circ \theta_m \right) = P(A^\Phi + g) - P(A^\Phi) \stackrel{\text{def}}{=} P^\Phi(g) , \tag{2.23}$$

where $\mu(q) = \int q d\mu$ and $(q \circ \theta_m)(\omega) = q(\theta_m \omega)$. Then by (2.18),

$$P^\Phi(g) = \sup_{\nu \in \mathscr{P}(\Omega)} (\nu(g) - I^\Phi(\nu)) , \tag{2.24}$$

where $I^\Phi(\nu)$ is given by (1.5). Since the entropy h_ν is affine and upper semicontinuous then $I^\Phi(\nu)$ is convex and lower semicontinuous, and so by the duality theorem in convex analysis (see, for instance, Theorem 3.12 in [R2]),

$$I^\Phi(\nu) = \sup_{g \in C(\Omega)} (\nu(g) - P^\Phi(g)) . \tag{2.25}$$

Finally, (2.23)–(2.25) together with Theorem 2.1 from [Ki1] yield the upper bound of large deviations (1.3).

3. The Kullback–Leibler Information

For each finite nonempty $\Lambda \subset \mathbb{Z}^d$ and any $\mu, \nu \in \mathscr{P}(\Omega)$ define

$$H_\Lambda(\nu|\mu) = \sum_{\xi \in \Omega_\Lambda} \nu(\Xi_\Lambda(\xi)) \log \frac{\nu(\Xi_\Lambda(\xi))}{\mu(\Xi_\Lambda(\xi))} , \tag{3.1}$$

where we assume $0 \log \frac{0}{c} = 0$ for any $c \geqq 0$ and $\log \frac{c}{0} = \infty$ for any $c > 0$. We call $H_\Lambda(\nu|\mu)$ the Kullback–Leibler information and it appears in different variations sometimes under the name of relative entropy in many works (see, for instance, [DS, DV, F, FO, and G]). The following result was proved in [F] in the full shift case with $\Lambda(a) \nearrow \infty$ in place of $\Lambda \nearrow \infty$.

3.1. Proposition. *Suppose that (Ω, θ) is a subshift of finite type and the condition D holds true for any sequence $\Lambda_n \nearrow \infty$. Then for any $\nu \in \mathscr{P}_I(\Omega)$ and $\mu \in K^\Phi$,*

$$\lim_{\Lambda \nearrow \infty} |\Lambda|^{-1} H_\Lambda(\nu|\mu) = P(A^\Phi) - \nu(A^\Phi) - h_\nu . \tag{3.2}$$

First we will prove the following assertion.

3.2. Lemma. *Suppose that the sequence $\Lambda_n \nearrow \infty$ satisfies the condition D and that $M_n \supset \Lambda_n$ is a corresponding sequence satisfying* (2.20). *Then*

(i) *There exists a sequence of numbers $\alpha(n) > 0$ such that*

$$\lim_{n\to\infty} \alpha(n)|M_n|^{-1} = 0 \text{ and } |U^\Phi_{M_n,\eta}(\xi) - U^\Phi_{M_n}(\xi)| \leqq \alpha(n) \tag{3.3}$$

for any $\xi \in \Omega_{M_n}$ and $\eta \in \Omega_{M_n^c}$ satisfying $\xi \vee \eta \in \Omega$.

(ii) *For any $\varepsilon > 0$ there is $n(\varepsilon)$ such that*

$$|U^\Phi_{M_n}(\xi) - U^\Phi_{\Lambda_n}(\xi_{\Lambda_n})| \leqq \varepsilon |\Lambda_n| \tag{3.4}$$

for any $\xi \in \Omega_{M_n}$ and $n \geqq n(\varepsilon)$.

(iii) *For any $\varepsilon > 0$ there is $n^*(\varepsilon)$ such that*

$$|U^\Phi_{\Lambda_n}(\omega_{\Lambda_n}) + \sum_{m\in\Lambda_n} A^\Phi(\theta_m\omega)| \leqq \varepsilon |\Lambda_n| \tag{3.5}$$

for any $\omega \in \Omega$ and $n \geqq n^*(\varepsilon)$.

Proof. Some parts of the following proof are standard and can be found in [F and R2] but for the reader's convenience we give the whole proof here. To obtain (i) notice that by (2.1), (2.2) and (2.4), (2.5) for any integer $k \geqq 1$,

$$\begin{aligned}
|U^\Phi_{M_n,\eta}(\xi) - U^\Phi_{M_n}(\xi)| &= |\sum_X \{\Phi_X(\xi \vee \eta) : X \in \mathscr{A} ,\\
X \cap M_n \neq \emptyset,\quad X \cap M_n^c \neq \emptyset\}| &\leqq \sum_X \{\|\Phi_X\|\\
X \in \mathscr{A},\quad X \cap M_n \neq \emptyset,\quad X \cap M_n^c &\neq \emptyset\} \stackrel{\text{def}}{=} \alpha(n)\\
&\leqq \sum_{m\in M_n} \sum_X \{\|\Phi_X\| : X \in \mathscr{A}; X \ni m, X \not\subset C(k) + m\}\\
&\quad + \sum_{m\in M_n} \sum_X \{\|\Phi_X\| : X \in \mathscr{A}_k(m)\} ,
\end{aligned} \tag{3.6}$$

where $C(k) = \{m \in \mathbb{Z}^d : 0 \leqq |m_i| \leqq k\}$ is the k-cube centered at zero and

$$\mathscr{A}_k(m) = \{X \in \mathscr{A} : m \in X \subset C(k) + m, X \not\subset M_n\} .$$

By (2.1)–(2.3) for any $m \in M_n$,

$$\begin{aligned}
&\sum_X \{\|\Phi_X\| : X \in \mathscr{A},\ X \ni m,\ X \not\subset C(k) + m\}\\
&= \sum_X \{\|\Phi_X\| : X \in \mathscr{A},\ X \ni 0,\ X \not\subset C(k)\}\\
&\stackrel{\text{def}}{=} \beta(k) \to 0 \text{ as } k \to \infty .
\end{aligned} \tag{3.7}$$

On the other hand, for any $X \in \mathscr{A}_k(m)$ there exists $m_1 \in C(k)$ such that $m + m_1 \notin M_n$. Thus

$$|\{m \in M_n : \mathscr{A}_k(m) \neq \emptyset\}| \leqq |M_n(k)| , \tag{3.8}$$

where we set $\Lambda(k)=\{m\in\Lambda:\min_{\ell\in\Lambda^c}\|m-\ell\|\leqq k\}|$ and again $\|a\|=\max_i|a_i|$ for $a\in\mathbb{Z}^d$. Since $M_n\nearrow\infty$ then by (2.11) for any $k\geqq 1$,

$$\lim_{n\to\infty}|M_n(k)||M_n|^{-1}=0\,. \tag{3.9}$$

By (2.1)–(2.3),

$$\sum_X\{\|\Phi_X\|:X\in\mathscr{A}_k(m)\}\leqq\sum\{\|\Phi_X\|:X\in\mathscr{A},\ X\ni 0\}\leqq\|\Phi\|\,. \tag{3.10}$$

Thus by (3.6)–(3.8) and (3.10),

$$\alpha(n)|M_n|^{-1}\leqq\beta(k)+\|\Phi\||M_n(k)||M_n|^{-1}\,. \tag{3.11}$$

Letting, first, $n\to\infty$ and then $k\to\infty$ we derive (3.3) from (3.7) and (3.9). In order to obtain (ii) observe that for any $\xi\in\Omega_{M_n}$,

$$\begin{aligned}|U^\Phi_{M_n}(\xi)-U^\Phi_{\Lambda_n}(\xi_{\Lambda_n})|&\leqq\sum_X\{|\Phi_X(\xi_X)|:X\in\mathscr{A},\ X\subset M_n,\ X\not\subset\Lambda_n\}\\&\leqq\sum_{m\in M_n\setminus\Lambda_n}\ \sum_{X\in\mathscr{A};X\ni m}\|\Phi_X\|\leqq|M_n\setminus\Lambda_n|\|\Phi\|\,.\end{aligned} \tag{3.12}$$

Now (ii) follows from (2.20) and (3.12). It remains to establish (iii). Consider some total order on $\mathbb{Z}^d$ which is compatible with the translations of $\mathbb{Z}^d$, for instance, the lexicographic order. Then (see Sect. 3.2 in [R2]),

$$A^\Phi(\omega)=A^\Phi_{\mathbb{Z}^d}(0,\omega)\,, \tag{3.13}$$

where for $\Lambda\subset\mathbb{Z}^d, m\in\Lambda$, and $\xi\in\Omega_\Lambda$ we set

$$\begin{aligned}A^\Phi_\Lambda(m,\xi)=-\sum_X\{\Phi_X(\xi_X):X\in\mathscr{A},\ X\ni m,\ X\subset\Lambda\,,\\ \text{and } m \text{ is the first element of } X\}\,.\end{aligned} \tag{3.14}$$

By (2.4) for any $\xi\in\Omega_{\Lambda_n}$,

$$U^\Phi_{\Lambda_n}(\xi)=-\sum_{m\in\Lambda_n}A^\Phi_{\Lambda_n}(m,\xi)\,. \tag{3.15}$$

By (2.1)–(2.3),

$$\left|\sum_{m\in\Lambda_n(k)}A^\Phi_{\Lambda_n}(m,\xi)|\leqq\|\Phi\||\Lambda_n(k)\right|\,. \tag{3.16}$$

Since $\Lambda_n\nearrow\infty$ then by (2.11) for any $k\geqq 1$,

$$\lim_{n\to\infty}|\Lambda_n(k)||\Lambda_n|^{-1}=0\,. \tag{3.17}$$

On the other hand, if $m\in\Lambda_n\setminus\Lambda_n(k)$ then

$$C(k)\subset\Lambda_n-m\,, \tag{3.18}$$

and so by (2.3) and (3.13) for any $\omega\in\Omega$,

$$A^{\Phi}_{\Lambda_n}(m,\omega_{\Lambda_n}) = A^{\Phi}_{\Lambda_{n-m}}(0,(\theta_m\omega)_{\Lambda_{n-m}})$$
$$= A^{\Phi}(\theta_m\omega) - \sum_X\{\Phi_X((\theta_m\omega)_X : X \not\subset \Lambda_{n-m}$$
$$\text{and } 0 \text{ is the first element of } X\} . \tag{3.19}$$

Thus by (2.1)–(2.3), (3.14)–(3.16), (3.18), and (3.19),

$$|U^{\Phi}_{\Lambda_n}(\omega_{\Lambda_n}) + \sum_{m\in\Lambda_n} A^{\Phi}(\theta_m\omega)| \leqq \|\Phi\| |\Lambda_n(k)| + |\Lambda_n|\beta(k) , \tag{3.20}$$

where $\beta(k)$ was defined in (3.7). By (2.1), $\beta(k) \to 0$ as $k \to \infty$, and so (iii) follows in view of (3.17).

Next, we pass to the proof of Proposition 3.1. First, we will show that for any sequence $\Lambda_n \nearrow \infty$ satisfying the condition D, $\nu \in \mathscr{P}_I(\Omega)$, and $\mu \in K^{\Phi}$,

$$\lim_{n\to\infty} \sup|\Lambda_n|^{-1} H_{\Lambda_n}(\nu|\mu) \leqq P(A^{\Phi}) - \nu(A^{\Phi}) - h_\nu . \tag{3.21}$$

By (2.6), (2.7), and Lemma 3.2(i) for any $n \geqq n_1(\varepsilon) = \max\{n : \alpha(n)|M_n|^{-1} > \varepsilon\}$

$$Z^{\Phi}_{M_n}(\eta) \leqq e^{\varepsilon|M_n|} Z^{\Phi}_{M_n} \tag{3.22}$$

for any $\eta \in \Omega_{M_n^c}$, where the sequence $M_n \nearrow \infty$ is given by the condition D. By the assertions (i) and (ii) of Lemma 3.2 for any $n \geqq n_2(\varepsilon) = \max(n(\frac{\varepsilon}{2}), n_1(\frac{\varepsilon}{2}))$ and all $\xi \in \Omega_{M_n}$ and $\eta \in \Omega_{M_n^c}$ satisfying $\xi \vee \eta \in \Omega$,

$$\exp(-U^{\Phi}_{M_n,\eta}(\xi)) \geqq e^{-\varepsilon|M_n|} \exp(-U^{\Phi}_{\Lambda_n}(\xi_{\Lambda_n})) . \tag{3.23}$$

By (2.10), (3.22), and (3.23) for any $\zeta \in \Omega_{\Lambda_n}$ and $n \geqq n_3(\varepsilon) = \max(n_1(\varepsilon), n_2(\varepsilon))$,

$$\mu(\Xi_{\Lambda_n}(\zeta)) = \sum_{\xi\in\Omega_{M_n}, \xi_{\Lambda_n}=\zeta} \mu(\Xi_{M_n}(\xi))$$
$$\geqq (Z^{\Phi}_{M_n})^{-1} \sum_{\xi\in\Omega_{M_n}, \xi_{\Lambda_n}=\zeta} \exp(-U_{\Lambda_n}(\zeta) - 2\varepsilon|M_n|)\mu_{M_n^c}\{\eta \in \Omega_{M_n^c} : \xi \vee \eta \in \Omega\}$$
$$= (Z^{\Phi}_{M_n})^{-1} \exp(-U_{\Lambda_n}(\zeta) - 2\varepsilon|M_n|) , \tag{3.24}$$

since by the condition D,

$$\sum_{\xi\in\Omega_{M_n}, \xi_{\Lambda_n}=\zeta} \mu_{M_n^c}\{\eta \in \Omega_{M_n^c} : \xi \vee \eta \in \Omega\}$$
$$= \mu_{M_n^c}\{\eta \in \Omega_{M_n^c} : \text{there exists } \omega \in \Omega \text{ such}$$
$$\text{that } \omega_{\Lambda_n} = \zeta \text{ and } \omega_{M_n^c} = \eta\} = \mu_{M_n^c}(M_n^c) = 1 . \tag{3.25}$$

Finally, by (2.16), (3.1), (3.5), (3.24), and the shift invariance of ν for any $n \geqq n_4(\varepsilon) = \max(n_3(\varepsilon), n^*(\varepsilon))$,

$$H_{\Lambda_n}(\nu|\mu) \leqq -H_{\Lambda_n}(\nu) + \log Z^{\Phi}_{M_n} + 2\varepsilon|M_n|$$
$$+ \int_\Omega U_{\Lambda_n}(\omega_{\Lambda_n}) d\nu(\omega) \leqq -H_{\Lambda_n}(\nu) + \log Z^{\Phi}_{M_n} + 2\varepsilon|M_n|$$
$$+ \varepsilon|\Lambda_n| - |\Lambda_n|\nu(A^{\Phi}) . \tag{3.26}$$

Now dividing both parts of (3.26) by $|\Lambda_n|$, letting first $n \to \infty$ and then $\varepsilon \to 0$, we obtain (3.21) by (2.12), (2.15), (2.17), (2.20), and (3.26).

Next, we are going to show that

$$\liminf_{n\to\infty} |\Lambda_n|^{-1} H_{\Lambda_n}(\nu|\mu) \geqq P(A^\Phi) - \nu(A^\Phi) - h_\nu \,. \tag{3.27}$$

By (2.6), (2.7), and Lemma 3.2 (i) and (ii) for any $n \geqq n_2(\varepsilon)$ and $\eta \in \Omega_{M_n^c}$,

$$\begin{aligned} Z_{M_n}(\eta) &= \sum_{\zeta\in\Omega_{\Lambda_n}} \sum_{\xi} \{\exp(-U_{M_n,\eta}(\xi)) : \xi \in \Omega_{M_n}, \xi_{\Lambda_n} = \zeta, \xi \vee \eta \in \Omega\} \\ &\geqq e^{-2\varepsilon|M_n|} \sum_{\zeta\in\Omega_{\Lambda_n}} \exp(-U_{\Lambda_n}(\zeta))|\Psi_{\xi,\eta}| \,, \end{aligned} \tag{3.28}$$

where, again, $|\Psi|$ is the number of elements in Ψ and $\Psi_{\zeta,\eta} = \{\xi \in \Omega_{M_n} : \xi_{\Lambda_n} = \zeta$ and $\xi \vee \eta \in \Omega\}$. By the condition D for any $\zeta \in \Omega_{\Lambda_n}$ and $\eta \in \Omega_{M_n^c}$ $\Psi_{\zeta,\eta} \neq \emptyset$, and so $|\Psi_{\zeta,\eta}| \geqq 1$. Thus by (2.6),

$$Z^\Phi_{M_n}(\eta) \geqq e^{-2\varepsilon|M_n|} Z^\Phi_\Lambda \tag{3.29}$$

for $n \geqq n_2(\varepsilon)$. By (2.9), (2.10), (3.3), (3.4), and (3.29) for all $n \geqq n_3(\varepsilon)$ and $\zeta \in \Omega_{\Lambda_n}$,

$$\begin{aligned} \mu(\Xi_{\Lambda_n}(\zeta)) &= \sum_{\xi\in\Omega_{M_n}, \xi_{\Lambda_n}=\zeta} \mu(\Xi_{M_n}(\xi)) \\ &\leqq |Q|^{|M_n\setminus\Lambda_n|} e^{3\varepsilon|M_n|} (Z^\Phi_{\Lambda_n})^{-1} \exp(-U_{\Lambda_n}(\zeta)) \,, \end{aligned} \tag{3.30}$$

where, recall, Q is the alphabet. Thus by (2.16), (3.1), (3.5), (3.30), and the shift invariance of ν for any $n \geqq n_4(\varepsilon)$,

$$\begin{aligned} H_{\Lambda_n}(\nu|\mu) \geqq & - H_{\Lambda_n}(\nu) - 3\varepsilon|M_n| + \log Z^\Phi_{\Lambda_n} \\ & - |M_n\setminus\Lambda_n| \log|Q| - |\Lambda_n|\nu(A^\Phi) - \varepsilon|\Lambda_n| \,. \end{aligned} \tag{3.31}$$

Dividing both parts of (3.31) by $|\Lambda_n|$, letting first $n \to \infty$ and then $\varepsilon \to 0$ we obtain (3.27) by (2.12), (2.15), (2.17), (2.20), and (3.31) completing the proof of Proposition 3.1. □

3.3 Remark. The main part of the proof of Proposition 3.1. is, in fact, the result saying that $\log\mu(\Xi_{\Lambda_n}(\xi))$ is equivalent when $\Lambda_n \nearrow \infty$ to $(\sum_{m\in\Lambda_n} A^\Phi(\theta_m\omega^\xi) - |\Lambda_n| P(A^\Phi))$, where $\omega^\xi \in \Omega$ satisfies $\omega^\xi_{\Lambda_n} = \xi$, which is the "volume lemma" type statement similar to Proposition 3.2 from [Ki1].

4. The Lower Bound

We will describe next a more general approach to lower large deviations bounds, than actually needed for Theorem A. Let $(\Omega, \mathscr{B})$ be a measurable space and $\mathscr{P}(\Omega)$ be the space of probability measures defined on elements of the σ-algebra B. Suppose that $\mathscr{F} \subset \mathscr{B}$ is a sub σ-algebra of $\mathscr{B}$ and $\nu, \mu \in \mathscr{P}(\Omega)$. Define the Kullback–Leibler information (cf. [DS, DV, F, G]) by

$$H^{\mathscr{F}}(\nu|\mu) = \mu(p^{\mathscr{F}}_{\nu,\mu} \log p^{\mathscr{F}}_{\nu,\mu}) \tag{4.1}$$

if ν is absolutely continuous with respect to μ on $(\Omega,\mathscr{F})$ (written $\nu \overset{\mathscr{F}}{\prec} \mu$) with Radon–Nikodym derivative $p^{\mathscr{F}}_{\nu,\mu}$, and $H^{\mathscr{F}}(\nu|\mu) = \infty$ otherwise. If Ω is the same as in the previous section and F is the finite σ-algebra of all subsets of Ω_Λ for a finite $\Lambda \subset \mathbb{Z}^d$ then $H^{\mathscr{F}}(\nu|\mu)$ coincides with $H_\Lambda(\nu|\mu)$ definied by (3.1). By Jensen's inequality for all $\nu,\mu \in \mathscr{P}(\Omega)$,

$$H^{\mathscr{F}}(\nu|\mu) \geqq 0 . \tag{4.2}$$

Set $g^+ = \max(0,g)$ and let $L^1_+(\Omega,\mathscr{F},\nu)$ denotes the set of all F-measurable functions g with $\nu(g^+) < \infty$. If $g \in L^1_+(\Omega,\mathscr{F},\nu)$ then $\nu(g) = \nu(g^+) - \nu(g^+ - g)$ is defined though it may be equal to $-\infty$.

4.1. Proposition. *For any* $\nu,\mu \in \mathscr{P}(\Omega)$,

$$H^{\mathscr{F}}(\nu|\mu) = \sup_{g\in L^1_+(\Omega,\mathscr{F},\nu)} (\nu(g) - \log\mu(e^g)) , \tag{4.3}$$

and if $\nu \overset{\mathscr{F}}{\prec} \mu$ *then for each* $C > 0$,

$$\nu\{\omega : p^{\mathscr{F}}_{\nu,\mu}(\omega) \geqq e^C\} \leqq C^{-1}(H^{\mathscr{F}}(\nu|\mu) + \log 2) . \tag{4.4}$$

Proof. Though (4.3) is contained essentially in [DV] and also in [DS, p.68] (with the supremum taken over bounded or bounded continuous functions) we will give the proof here for the sake of completeness.

First, remark that if $H^{\mathscr{F}}(\nu|\mu) = \infty$ then both parts of (4.3) equal ∞. Indeed, if $\nu \overset{\mathscr{F}}{\prec} \mu$ does not hold true then there is a set $\Psi \in \mathscr{F}$ with $\mu(\Psi) = 0$ and $\nu(\Psi) > 0$. Choose $g_n = n\mathbf{1}_\Psi, n = 1,2,\dots$ where $\mathbf{1}_\Psi(\omega) = 1$ if $\omega \in \Psi$ and $= 0$, otherwise. Then the supremum of the right-hand side of (4.3) over $\{g_n, n = 1,\dots\}$ is ∞. If $\nu \overset{\mathscr{F}}{\prec} \mu$ but $\nu(\log^+ p^{\mathscr{F}}_{\nu,\mu}) = \infty$ then define $\Psi_n = \{\omega : 0 \leqq \log p^{\mathscr{F}}_{\nu,\mu}(\Omega) \leqq n\}$ and $g_n = \mathbf{1}_{\Psi_n} \log p^{\mathscr{F}}_{\nu,\mu}$. Then $\mu(e^{g_n}) = \mu(\Omega\backslash\Psi_n) + \nu(\Psi_n) \leqq 2$ and $\nu(g_n) \to \infty$ as $n \to \infty$. Hence, again, the supremum of the right-hand side of (4.3) over $\{g_n, n = 1,\dots\}$ is ∞. Now assume that $\nu \overset{\mathscr{F}}{\prec} \mu, \nu(\log^+ p^{\mathscr{F}}_{\nu,\mu}) < \infty$, and set $\Psi = \{\omega \in \Omega : p^{\mathscr{F}}_{\nu,\mu}(\omega) > 0\}$. Let $g \in L^1_+(\Omega,\mathscr{F},\nu)$ and $q = e^g(p^{\mathscr{F}}_{\nu,\mu})^{-1}\mathbf{1}_\Psi + \mathbf{1}_{\Omega\backslash\Psi}$ then, clearly,

$$\nu(g) - \log\mu(e^g) \leqq H^{\mathscr{F}}(\nu|\mu) + [\nu(\log q) - \log\nu(q)] . \tag{4.5}$$

The expression in the square brackets is nonpositive by Jensen's inequality, and so the lefthand side of (4.5) does not exceed $H^{\mathscr{F}}(\nu|\mu)$. On the other hand, set $g_n = \mathbf{1}_\Psi \log p^{\mathscr{F}}_{\nu,\mu} - n\mathbf{1}_{\Omega\backslash\Psi}$ then $g_n \in L^1_+(\Omega,\mathscr{F},\nu)$ and

$$\lim_{n\to\infty} (\nu(g_n) - \log\mu(e^{g_n})) = H^{\mathscr{F}}(\nu|\mu) ,$$

completing the proof of (4.3).

Next, we prove (4.4). If $\nu(\log^+ p^{\mathscr{F}}_{\nu,\mu}) = \infty$, then $H^{\mathscr{F}}(\nu|\mu) = \infty$ and there is nothing to prove. So assume that $\log p^{\mathscr{F}}_{\nu,\mu} \in L^1_+(\Omega,\mathscr{F},\nu)$ then by (4.3) and the Chebyshev inequality,

$$\begin{aligned}\nu\{ p^{\mathscr{F}}_{\nu,\mu} \geqq e^C\} &\leqq C^{-1}\nu(\log^+ p^{\mathscr{F}}_{\nu,\mu})\\ &\leqq C^{-1}(H^{\mathscr{F}}(\nu|\mu) + \log\mu(\exp(\log^+ p^{\mathscr{F}}_{\nu,\mu})))\\ &= C^{-1}(H^{\mathscr{F}}(\nu|\mu) + \log(\mu\{ p^{\mathscr{F}}_{\nu,\mu} < 1\} + \nu\{ p^{\mathscr{F}}_{\nu,\mu} \geqq 1\})\\ &\leqq C^{-1}(H^{\mathscr{F}}(\nu|\mu) + \log 2)\,, \end{aligned} \tag{4.6}$$

completing the proof of Proposition 4.1. □

Next, let $(\Omega, \mathscr{B})$ be again a measurable space, $\nu, \mu \in \mathscr{P}(\Omega)$, and $\zeta^n : \Omega \to \mathscr{P}(\Omega), n = 1, 2, \dots$ be a sequence of measurable maps, where $\mathscr{P}(\Omega)$ is taken with some measurable structure.

4.2. Proposition. *Suppose that there exists a measurable set $U \subset \mathscr{P}(\Omega)$ and a sequence of σ-algebras $\mathscr{F}_n \subset \mathscr{B}, n = 1, 2, \dots$ such that $\{\omega : \zeta^n_\omega \in U\} \in \mathscr{F}_n$ for all $n = 1, 2, \dots$ and*

$$\lim_{n\to\infty} \nu\{\zeta^n \in U\} = 1\,. \tag{4.7}$$

If $r(n) \to \infty$ as $n \to \infty$ and

$$h = \lim_{n\to\infty} \sup (r(n))^{-1} H^{\mathscr{F}_n}(\nu|\mu)\,, \tag{4.8}$$

then

$$\lim_{n\to\infty} \inf (r(n))^{-1} \log\mu\{\zeta^n \in U\} \geqq -h\,. \tag{4.9}$$

Proof. If $h = \infty$ then there is nothing to prove. So suppose that $h < \infty$, then without loss of generality we may assume that $H^{\mathscr{F}_n}(\nu|\mu) < \infty$ for all n which means that $\nu^{\mathscr{F}_n} \prec \mu$ and $\nu(\log^+ p^{\mathscr{F}_n}_{\nu,\mu}) < \infty$. By (4.7) and (4.8) for any $\varepsilon > 0$ there exists $n(\varepsilon) \geqq 1$ such that

$$\nu\{\zeta^n \in U\} \geqq 1 - \varepsilon \tag{4.10}$$

and

$$H^{\mathscr{F}_n}(\nu|\mu) \leqq r(n)(h + \varepsilon) \tag{4.11}$$

for any $n \geqq n(\varepsilon)$. Applying (4.4) with

$$C = C_{n,\varepsilon} = r(n)(1 - 2\varepsilon)^{-1}(h + \varepsilon)\,, \tag{4.12}$$

we obtain from (4.11) that for all $n \geqq n(\varepsilon)$,

$$\nu\{ p^{\mathscr{F}_n}_{\nu,\mu} \geqq e^{C_{n,\varepsilon}}\} \leqq 1 - 2\varepsilon\,. \tag{4.13}$$

Remark that $\nu\{p^{\mathscr{F}_n}_{\nu,\mu} > 0\} = 1$ which together with (4.10) and (4.13) gives

$$\nu(\Psi_n) \geqq \varepsilon, \text{ where } \Psi_n = \{\zeta^n \in U, p^{\mathscr{F}_n}_{\nu,\mu} > 0, (p^{\mathscr{F}_n}_{\nu,\mu})^{-1} \geqq e^{-C_{n,\varepsilon}}\}\,.$$

Now by (4.13),

$$\mu\{\zeta^n \in U\} \geqq \mu\{\Psi_m\} = \int_{\Psi_n} (p^{\mathscr{F}_n}_{\nu,\mu})^{-1} d\nu \geqq \varepsilon e^{-C_{n,\varepsilon}}\,. \tag{4.14}$$

Taking log in (4.14), dividing by $r(n)$, letting first $n \to \infty$ and then $\varepsilon \to 0$ we derive (4.9). $\square$

Next, we are going to apply Proposition 4.2 in the set up of Theorem A in order to prove (1.4). Let $G \subset \mathscr{P}(\Omega)$ be an open set and $\mu \in K_\Phi$. Set $I_G^\Phi = \inf_{v \in G} I^\Phi(v)$ with $I^\Phi(v)$ given by (1.5). If $I_G^\Phi = \infty$ there is nothing to prove, so we assume $I_G^\Phi < \infty$. Then by Theorem B which will be proved in Sect. 6 for any $\varepsilon > 0$ there exists an ergodic $v_\varepsilon \in \mathscr{P}_I(\Omega)$ such that

$$v_\varepsilon \in G \text{ and } I^\Phi(v_\varepsilon) \leqq I_G^\Phi + \varepsilon\,. \tag{4.15}$$

In addition. we can choose $N = N(v_\varepsilon) > 0$ so large that

$$U = U(v_\varepsilon) \stackrel{\text{def}}{=} \{v \in \mathscr{P}(\Omega) : \max_{\xi \in \Omega_{C(N)}} |v(\Xi_{C(N)}(\xi)) - v_\varepsilon(\Xi_{C(N)}(\xi))| < \exp(-N^{2d})\} \subset G\,, \tag{4.16}$$

where, recall, $C(N)$ is the N-cube centered at 0. Now let $\Lambda_n \nearrow \infty$. Denote

$$\Lambda_n(N) = \bigcup_{m \in \Lambda_n} (C(N) + m)\,.$$

If $\omega, \tilde{\omega} \in \Xi_{\Lambda_n(N)}(\eta)$ for some $\eta \in \Omega_{\Lambda_n}(N)$, then clearly $(\theta_m \omega)_{C(N)} = (\theta_m \tilde{\omega})_{C(N)}$ for any $m \in \Lambda_n$, and so for any $\xi \in \Omega_{C(N)}$,

$$\zeta_\omega^{\Lambda_n}(\Xi_{C(N)}(\xi)) = \zeta_{\tilde{\omega}}^{\Lambda_n}(\Xi_{C(N)}(\xi))\,.$$

Thus for some $\Psi_{n,\varepsilon} \subset \Omega_{\Lambda_n(N)}$,

$$\{\zeta^{\Lambda_n} \in U(v_\varepsilon)\} = \bigcup_{\eta \in \Phi_{n,\varepsilon}} \Xi_{\Lambda_n(N)}(\eta)\,. \tag{4.17}$$

Let $\mathscr{F}_n$ be the finite σ-algebra of all subsets of $\Omega_{\Lambda_n(N)}$, then by (4.17),

$$\{\zeta^{\Lambda_n} \in U(v_\varepsilon)\} \in \mathscr{F}_n\,. \tag{4.18}$$

Since v_ε is ergodic it follows from the mean ergodic theorem (see [Kr, Sect. 6.4 or [T], Sect. 3 of Chapter 3) and the Chebyshev inequality that

$$\lim_{n \to \infty} v_\varepsilon\{\zeta^{\Lambda_n} \in U(v_\varepsilon)\} = 1\,. \tag{4.19}$$

Remark that $H^{\mathscr{F}_n}(v_\varepsilon|\mu) < \infty$, in particular, $v_\varepsilon^{\mathscr{F}_n} \prec \mu$ since $\mu \in K^\Phi$ implies that $v(\Xi_{\Lambda_n(N)}(\eta)) > 0$ for any $\eta \in \Omega_{\Lambda_n(N)}$. Since $\Lambda_n(N) \nearrow \infty$ as $n \to \infty$, then by (1.5) and (3.2),

$$\lim_{n \to \infty} |\Lambda_n(N)|^{-1} H^{\mathscr{F}_n}(v_\varepsilon|\mu) = I^\Phi(v_\varepsilon)\,. \tag{4.20}$$

Remark that by (2.11),

$$\lim_{n \to \infty} |\Lambda_n||\Lambda_n(N)|^{-1} = 1\,,$$

and so we can apply Proposition 4.2 with $U = U(v_\varepsilon), \zeta^n = \zeta^{\Lambda_n}$, and $r(n) = |\Lambda_n|$ which together with (4.15) and (4.16) yields

$$\begin{aligned} &\liminf_{n \to \infty} |\Lambda_n|^{-1} \log \mu\{\zeta^{\Lambda_n} \in G\} \\ &\quad \geqq \liminf_{n \to \infty} |\Lambda_n|^{-1} \log \mu\{\zeta^{\Lambda_n} \in U(v_\varepsilon)\} \geqq -I^\Phi(v_\varepsilon) \geqq -I_G^\Phi - \varepsilon\,. \end{aligned} \tag{4.21}$$

Since $\varepsilon > 0$ is arbitrary this implies (1.4). $\square$

4.3. Remark. Note that in order to prove (1.4) one needs (4.3) and (4.4) only for finite σ-algebras $\mathscr{F}$ which amounts to

$$H(p|q) = \sum_{i=1}^{k} p_i \log(p_i/q_i) = \sup_{a_1,\dots,a_k} \left(\sum_{i=1}^{k} p_i a_i - \log\left(\sum_{i=1}^{k} q_i e^{a_i} \right)\right) \tag{4.22}$$

and

$$\sum_i \{p_i : p_i q_i^{-1} \geqq e^C\} \leqq C^{-1}(H(p|q)+2), \tag{4.23}$$

being true for any probability vectors $p=(p_1,\dots,p_k)$ and $q=(q_1,\dots,q_k)$ provided $0/0 = 1$, $0\log\frac{0}{C}=0$ for $C\geqq 0$, and $\log\frac{C}{0}=\infty$ for $C>0$.

5. Large Deviations from the Equidistribution

Set

$$\mathscr{P}_{\max}(\Omega) = \{\nu \in \mathscr{P}_I(\Omega) : h_\nu = h_{\text{top}}\},$$

where h_{top} is the topological entropy of the subshift (Ω,θ) and $h_{\text{top}} = P(0)$, i.e. it satisfies (2.18) with $A\equiv 0$. Since h_ν is upper semicontinuous in ν then $\mathscr{P}_{\max}(\Omega) \neq \emptyset$. Recall that the measures on which the supremum in (2.18) is attained are called the equilibrium states (measures) for A. Thus elements of $\mathscr{P}_{\max}(\Omega)$ are equilibrium states for the function $A\equiv 0$. Since the strong specification implies the weak specification and the latter yields the condition D from Sect. 2 then by Theorem 4.2 from [R2] saying that sets of all Gibbs states for an interaction Φ and all equilibrium states for A^Φ coincide it follows that $\mathscr{P}_{\max}(\Omega)$ is the set of all Gibbs measures for the interaction $\Phi\equiv 0$, i.e. $\mathscr{P}_{\max}(\Omega) = K^0$. Remark that when $d=1$ the set $\mathscr{P}_{\max}(\Omega)$ is a singleton and the corresponding measure $\nu_{\max}$ is called the measure with maximal entropy. In this case Theorem C was proved in [K2]. It implies, in particular, that long periodic orbits with the distribution close to $\nu_{\max}$ form a set whose proportion is close to one. This result was proved by Bowen and was called by him the equidistribution of closed orbits. As far as we know there are no general conditions available in $d\geqq 2$ case implying that $\mathscr{P}_{\max}(\Omega)$ is a singleton. Thus the upper bound (1.9) yield only that the collection of "long" periodic orbits with distributions close to the convex set $\mathscr{P}_{\max}(\Omega)$ have proportion close to one.

Let now $g\in C(\Omega)$. Then under the strong specification condition by (1.8) and Theorem 2.2 from [R1],

$$\begin{aligned}
&\lim_{a\to\infty} |\Lambda(a)|^{-1}\log \nu_a\left(\exp \sum_{m\in\Lambda(a)} g\circ\theta_m\right)\\
&\quad= \lim_{a\to\infty} |\Lambda(a)|^{-1}\log \sum_{\omega\in\Pi_a} \exp\left(\sum_{m\in\Lambda(a)} g(\theta_m\omega)\right)\\
&\qquad - \lim_{a\to\infty} |\Lambda(a)|^{-1}\log|\Pi_a| = P(g)-P(0) = P(g)-h_{\text{top}}\,.
\end{aligned} \tag{5.1}$$

But by (1.12) and (2.18),

$$P(g) - h_{\text{top}} = \sup_{\nu\in\mathscr{P}(\Omega)} (\nu(g) - J(\nu)) \tag{5.2}$$

with $J(v)$ being convex and lower semicontinuous in v. Now the upper large deviations bound (1.10) follows from Theorem 2.1 of [Ki1] by the same routine as in the end of Sect. 2.

Since the uniqueness of Gibbs measures for a large class of interactions as required in Theorem 2.1 from [Ki1] does not hold true in our circumstances we cannot use this theorem in order to derive the lower bound (1.11) and we will use for it the lower bound from Theorem A.

Let $\mu \in \mathscr{P}_{\max}(\Omega)$, and so v is a Gibbs measure for the interaction $\Phi \equiv 0$. Let $Z^d \supset M \supset \Lambda$, where Λ and M are finite sets such that

$$\min_{\ell \in \Lambda, m \in M^c} \|\ell - m\| \geqq N \tag{5.3}$$

and N is the constant from the definition of the weak specification which means that for any $\xi \in \Omega_\Lambda$ and $\eta \in \Omega_{M^c}$ there exists $\omega \in \Omega$ with $\omega_\Lambda = \xi$ and $\omega_{M^c} = \eta$. Then by (2.5) and (2.7) for $\eta \in \Omega_{M^c}$,

$$\begin{aligned} Z_M(\eta) = Z_M^0(\eta) &= |\{\xi \in \Omega_M : \xi \vee \eta \in \Omega\}| \\ &= \sum_{\xi \in \Omega_\Lambda} |\{\zeta \in \Omega_M : \zeta|_\Lambda = \xi, \zeta \vee \eta \in \Omega\}| \geqq |\Omega_\Lambda| \,. \end{aligned} \tag{5.4}$$

Thus by (2.9) and (2.10) for any $\zeta \in \Omega_M$,

$$\mu(\Xi_M(\zeta)) \leqq |\Omega_\Lambda|^{-1} \,. \tag{5.5}$$

It follows from here that for any $\xi \in \Omega_\Lambda$,

$$\mu(\Xi_\Lambda(\xi)) \leqq |\Omega_\Lambda|^{-1} |Q|^{|M \setminus \Lambda|} \tag{5.6}$$

where, recall, Q is the alphabet.

In our circumstances we can reformulate the definition of the strong specification given in Sect..2 in the following way: there exists $K > 0$ such that for any subsets $R_i \subset \Lambda(a)$ satisfying $\min\{\|\ell - m\| : \ell \in \Lambda_i, m \in \Lambda_j + Z^d(a)\} \geqq K, i \neq j$ and for any $\xi_i \in \Omega_{R_i}$ there exists $\omega \in \Pi_a$ such that $\omega_{R_i} = \xi_i$ for all i.

Let $L > \max(N, K)$ and

$$G_L(a) = \left\{ \ell \in \Lambda(a) : \min_{m \in \Lambda^c(a)} \|\ell - m\| \geqq L \right\} \,.$$

The strong specification implies that for any $\xi \in G_L(a)$ there exists $\omega \in \Pi_a$ such that $\omega_{G(a)} = \xi$. Thus defining

$$D_a^L(\omega) = \{\tilde{\omega} \in \Omega : \omega|_{G_{L(a)}} = \tilde{\omega}|_{G_{L(a)}}\}$$

for all $\omega \in \Pi_a$ we obtain

$$\Omega = \bigcup_{\omega \in \Pi_a} D_a^L(\omega) \,. \tag{5.7}$$

It is clear that either $D_a^L(\omega) = D_a^L(\tilde{\omega})$ for $\omega, \tilde{\omega} \in \Pi_a$ or $D_a^L(\omega) \cap D_a^L(\tilde{\omega}) = \emptyset$. Let Π_a^* be a maximal collection of points $\omega \in \Pi_a$ such that $\omega, \tilde{\omega} \in \Pi_a^*, \omega \neq \tilde{\omega}$ imply $D_a^L(\omega) \cap D_a^L(\tilde{\omega}) = \emptyset$. Then by (5.7),

$$\Omega = \bigcup_{\omega \in \Pi_a^*} D_a^L(\omega) \,. \tag{5.8}$$

Taking in (5.6) $\Lambda = G(a)$ and $M = \Lambda(a)$ we obtain that for each $\omega \in \Pi_a$,

$$\mu(D_a^L(\omega)) \leqq |\Omega_{G_L(a)}|^{-1} e^{\rho(a)} \tag{5.9}$$

with $\rho_L(a) = |\Lambda(a) \backslash G_L(a)| \log|Q| = o(|\Lambda(a)|)$ as $a \to \infty$. Since $\Omega_{\Lambda(a)} \neq \tilde{\Omega}_{\Lambda(a)}$ for any $\Omega, \tilde{\Omega} \in \Pi_a, \Omega \neq \tilde{\Omega}$ then

$$|\Pi_a| \leqq |\Omega_{\Lambda(a)}| \leqq |\Omega_{G_L(a)}| e^{\rho(a)} , \tag{5.10}$$

and so by (5.9) for each $\omega \in \Pi_a$,

$$\mu(D_a^L(\omega)) \leqq |\Pi_a|^{-1} e^{2\rho(a)} . \tag{5.11}$$

Next, let $G \subset \mathscr{P}(\Omega)$ be an open set and $J_G = \inf_{\eta \in G} J(\eta)$ with $J(\eta)$ given by (1.12). By the convexity of the function $-x \log x$, $x \in (0,1)$ and Jensen's inequality one derives from (2.12) the well known inequality $H_\Lambda(\nu) \leqq \log|Q|$, and so by (2.17) and (2.19), $h_{\text{top}} = P(0) \leqq \log|Q|$. Thus both $J(\eta)$ and J_G are bounded by $\log|Q|$, as well. Given $\varepsilon > 0$ pick up $\nu_\varepsilon \in \mathscr{P}_I(\Omega)$ such that

$$\nu_\varepsilon \in G \text{ and } J(\nu_\varepsilon) \leqq J_G + \varepsilon . \tag{5.12}$$

Similarly to (4.16) choose $L > 0$ so large that

$$U_{L,\delta} \stackrel{\text{def}}{=} \{\nu \in \mathscr{P}(\Omega) : \max_{\xi \in \Omega_{C(L)}} |\nu(\Xi_{C(L)}(\xi)) - \nu_\varepsilon(\Xi_{C(L)}(\xi))| < \delta \exp(-N^{2d})\} \subset G . \tag{5.13}$$

If $\tilde{\omega} \in D_a^L(\omega)$ then $\omega_{G_L(a)} = \tilde{\omega}_{G_L(a)}$, and so

$$(\theta_m \omega)_{C(L)} = (\theta_m \tilde{\omega})_{C(L)} \text{ for any } m \in G_{2L}(a) . \tag{5.14}$$

Thus for any $\tilde{\omega} \in D_a^L(\omega)$ and $\xi \in \Omega_{C(L)}$,

$$\begin{aligned} &|\zeta_\omega^a(\Xi_{C(L)}(\xi)) - \zeta_{\tilde{\omega}}^a(\Xi_{C(L)}(\xi))| \\ &\leqq |\Lambda(a)|^{-1} |\Lambda(a) \backslash G_{2L}(a)| \stackrel{\text{def}}{=} r_L(a) \to 0 \text{ as } \varepsilon \to \infty . \end{aligned} \tag{5.15}$$

Pick up $n(\delta) > 0$ so that if $a = (a_1, \ldots, a_d)$ and $a_{\min} \stackrel{\text{def}}{=} \min_{1 \leqq i \leqq d} a_i \geqq n(\delta)$ then $r_L(a) \leqq \delta/2$.

This together with (5.15) imply that if $a_{\min} \geqq n(\delta), \tilde{\omega} \in D_a^L(\omega)$, and $\zeta_{\tilde{\omega}}^{\Lambda(a)} \in U_{L,\frac{1}{2}\delta}$ then $\zeta_\omega^a \in U_{L,\delta}$, and so by (5.8),

$$\{\omega \in \Omega : \zeta_\omega^{\Lambda(a)} \in U_{L,\frac{1}{2}\delta}\} \subset \bigcup_\omega \{D_a^L(\omega) : \omega \in \Pi_a^*, \zeta_\omega^a \in U_{L,\delta}\} . \tag{5.16}$$

Taking into account that $D_a^L(\omega)$ are disjoint for different $\omega \in \Pi_a$ we derive from (5.11) and (5.16) that

$$\begin{aligned} |\{\omega \in \Pi_a : \zeta_\omega^a \in G\}| &\geqq |\{\omega \in \Pi_\omega^* : \zeta_\omega^a \in U_{L,\delta}\}| \\ &\geqq |\Pi_a| e^{-2\rho(a)} \sum_\omega \{\mu(D_a^L(\omega)) : \omega \in \Pi_a^*, \zeta_\omega^a \in U_{L,\delta}\} \\ &= |\Pi_a| e^{-2\rho(a)} \mu\left(\bigcup_\omega \{D_a^L(\omega) : \omega \in \Pi_a^*, \zeta_\omega^a \in U_{L,\delta}\}\right) \\ &\geqq |\Pi_a| e^{-2\rho(a)} \mu\{\omega \in \Omega : \zeta_\omega^{\Lambda(a)} \in U_{L,\frac{1}{2}\delta}\} . \end{aligned} \tag{5.17}$$

Since $\mu \in K^0, I^0(v) = J(v)$, and $\rho(a) = o(|\Lambda(a)|)$ then by (1.4), (1.8), (5.12), (5.13), and (5.17),

$$\begin{aligned}
&\lim_{a\to\infty} |\Lambda(a)|^{-1} \log v_a\{\omega : \zeta^a_\omega \in G\} \\
&\quad = \lim_{a\to\infty} |\Lambda(a)|^{-1} \log(|\Pi_a|^{-1} |\{\omega \in \Pi_a : \zeta^a_\omega \in G\}|) \\
&\quad \geqq \lim_{a\to\infty} |\Lambda(a)|^{-1} \log \mu\{\zeta^{\Lambda(a)}_\omega \in U_{L,\frac{1}{2}\delta}\} \geqq -J_G - \varepsilon\,. \qquad (5.18)
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary we obtain (1.11). □

6. Approximation by Ergodic Measures

In this section we will prove Theorem B in the set up of a continuous action of $\mathbb{Z}^d$ on a compact metric space (Ω, d) assuming that the weak specification and the upper semicontinuity of the entropy conditions introduced in Sect. 2 are satisfied. Recall that the expansivity which is always satisfied in the framework of Theorems A and C is a sufficient condition for the upper semicontinuity of the entropy (see [R1]). Remark that usual proofs for the approximation of a $\mathbb{Z}^d$-invariant measure v by a weakly converging to v sequence of ergodic measures v_n takes v_n sitting on periodic orbits thus having zero entropy, and so does not enable us to approximate v in entropy, as well (unless $h_v = 0$), which is the main point here.

In fact, we will produce more than just ergodic measures approximating $v \in \mathscr{P}(\Omega)$. Invariant sets $\Psi \subset \Omega$ will be constructed such that all of their invariant measures are close to v, and such that the topological entropy of the action restricted to Ψ is close to the entropy of $(\Omega, \mathbb{Z}^d, v)$. This can be thought of as an approximate version of the Jewett–Krieger theorem (see [W]) in the non-ergodic setting where, of course, the theorem itself fails to hold true.

Recall that for $\varepsilon > 0$ and $\Lambda \subset \mathbb{Z}^d$ a set $E \subset \Omega$ is said to be (ε, Λ)-separated if for any $\omega, \tilde\omega \in E, \omega \neq \tilde\omega$ there exists $m \in \Lambda$ such that $d(m\omega, m\tilde\omega) \geqq \varepsilon$, where $m\omega$ denotes the result of the action of $m \in \mathbb{Z}^d$ on $\omega \in \Omega$. We will need a general result on the relation between a measure entropy and the topological entropy. Here no specification and expansivity assumptions are necessary. Denote by C_n the cube $\{m = (m_1, \ldots, m_d) : 0 \leqq m_i < n, 1 \leqq i \leqq d\}$.

Proposition 6.1. *Let $\eta \in \mathscr{P}(\Omega)$ be an ergodic invariant measure of a continuous action of $\mathbb{Z}^d$ on (Ω, d). Given continuous functions $\{f_1, \ldots, f_k\}$ and positive constants $\alpha > 0, \beta > 0$ there exist $n_0, \varepsilon > 0$ such that for all $n \geqq n_0$ one can find a (ε, C_n)-separated set $S \subset \Omega$ satisfying:*

(i) $|S| \geqq \exp(n^d(h_\eta - \alpha))$ *where, again, h_η is the entropy of this $\mathbb{Z}^d$-action on Ω with respect to η;*

(ii) $\left| |C_n|^{-1} \sum_{m \in C_n} f_j(m\omega) - \int f_j d\eta \right| < \beta$ *for all $\omega \in S$ and $j = 1, \ldots, k$.*

Proof. By the definition of the entropy h_η (see [R1, Kr, or T]) we can find some finite partition D of Ω into regular sets D_i such that $\eta(\partial D_i) = 0$ and

$$|h_\eta - h_\eta(\mathscr{D})| < \alpha/4\,, \qquad (6.1)$$

where $h_\eta(\mathscr{D})$ is the entropy of the partition D. Next, by the ergodic theorem (see [Kr] or [T]) if n is sufficiently large then the set E_n of x which satisfy

$$\left| |C_n|^{-1} \sum_{m \in C_n} f_j(m\omega) - \int f_j dv \right| < \beta, \quad 1 \leqq j \leqq k \tag{6.2}$$

has measure at least 1/2. Choose now an auxiliary $\delta > 0$ (the precise choice will be specified later) and find an open set V containing $\bigcup_i \partial D_i$ such that

$$\eta(V) < \delta \, . \tag{6.3}$$

Finally, let 2ε be the minimum distance between the disjoint closed sets $(\bar{D}_i \setminus V)$. We will say that points $\omega, \tilde{\omega} \in \Omega$ are "well" (ε, n)-separated if $d(m\omega, m\tilde{\omega}) \geqq \varepsilon$ for some $m \in C_n$ and in addition $m\omega$ and $m\tilde{\omega}$ do not belong to V. Applying again the ergodic theorem and taking n still larger we can ensure that

$$F_n = E_n \cap \{\omega : |C_n|^{-1} \sum_{m \in C_n} \mathbf{1}_V(m\omega) \leqq 2\delta\}$$

has η-measure at least 1/4. Now let S be a maximal "well" (ε, n)-separated subset of F_n. We associate now with each point $\omega \in S$ a collection of atoms $\mathscr{A}(\omega)$ of $\bigvee_{m \in C_n} m^{-1}\mathscr{D}$ in the following way. Put into A(ω) any atom containing a point $\tilde{\omega}$ such that

$$|C_n|^{-1} |\{m \in C_n : m\omega \text{ and } m\tilde{\omega} \text{ are in the same } \mathscr{D}\text{-atom}\}| \geqq 1 - 4\delta \, .$$

Clearly, the number of elements in A(ω_0) does not exceed $\sum_{m \leqq 4\delta |C_n|} \binom{|C_n|}{m} |\mathscr{D}|^m$, and so if δ is small enough then

$$|\mathscr{A}(\omega)| < \exp(|C_n|\alpha/4) \, . \tag{6.4}$$

Observe that the maximality of S implies that $\bigcup_{\omega \in S} \mathscr{A}(\omega)$ covers F_n. Indeed, if this were not true we would have a point $\tilde{\omega} \in F_n$ such that for all $\omega \in S$ $m\tilde{\omega}$ and $m\omega$ are in different $\mathscr{D}$-atoms for more than $4\delta|C_n|$ points m in C_n. Then by the definition of F_n there is at least one point m_0 among these m such that both $m_0\omega$ and $m_0\tilde{\omega}$ does not belong to V, and so $\omega, \tilde{\omega} \in F_n$ are "well" (ε, n)-separated for any $\omega \in S$ in contradiction to the maximality of S. Thus we get a collection of atoms from $\bigvee_{m \in C_n} m^{-1}\mathscr{D}$ whose cardinality is at most $|S|\exp(|C_n|\alpha/4)$ covering a set of η-measure 1/4. By the Shannon–McMillan theorem (see [Kr], Sect. 9.2, [T], Sect. 5 of Chapter 8, or [OW]) for n large enough those atoms from the above collection whose measure is smaller than $\exp(-|C_n|(h_{\eta(\mathscr{D})} - \alpha/4))$ will cover a set of measure of at least 1/8, and so

$$|S|\exp(|C_n|\alpha/4)\exp\left(-|C_n|\left(h_\eta(\mathscr{D}) - \frac{\alpha}{4}\right)\right) \geqq 1/8 \tag{6.5}$$

which together with (6.1) imply the lower bound (i) provided n is big enough. The property (ii) is also satisfied by (6.2). □

Next, we will pass to the proof of Theorem B. Using the ergodic decomposition of v and the fact that the entropy is an affine function on the space of $\mathbb{Z}^d$-invariant measures (see [Kr, T, R2]) it follows that with no loss of generality we may assume that v is a finite sum of ergodic measures $v^{(k)}$, $k = 1, \ldots, K$, i.e.

$$\nu = K^{-1} \sum_{1 \leqq k \leqq K} \nu^{(k)}, \tag{6.6}$$

where the $\nu^{(k)}$'s are not necessarily distinct. So we will complete the proof of Theorem B under this assumption. Our next objective will be, given continuous functions f_j with $\sup_\omega |f_j| \leqq 1, j = 1, \dots, J$ and $\beta > 0$, to find a large enough cube C so that the set

$$\Psi = \{\omega \in \Omega : \left| |C|^{-1} \sum_{l \in C} f_j((l+m)\omega) - \int f_j d\nu \right| \leqq \beta \text{ for all } m \in \mathbb{Z}^d \text{ and all } j = 1, \dots, J\}, \tag{6.7}$$

which is clearly closed and $\mathbb{Z}^d$-invariant has topological entropy at least $h_\nu - \alpha$.

We may assume that the $f_1, \dots, f_J$ and $\beta > 0$ are chosen in such a way that for any large enough cube C the topological entropies of the corresponding sets Ψ will not exceed $h_\nu + \alpha$. This can be done since for otherwise by the variational principle (see [M, R1]), the affine nature of the entropy function, and the ergodic theorem we will have a sequence of $\mathbb{Z}^d$-invariant ergodic probability measures $\eta_n \xrightarrow{w} \nu$ as $n \to \infty$ such that $h_{\eta_n} \geqq h_\nu + \alpha$. But this contradicts the upper semicontinuity of the entropy at the point ν. Thus if we will make the above construction for any $\alpha, \beta > 0$ and will show that the topological entropy of Ψ is at least $h_\nu - \alpha$, then by the variational principle and the affine character of the entropy function we will deduce that in any open neighborhood of ν there are ergodic measures on Ψ with the entropy sandwiched between $h_\nu - \alpha$ and $h_\nu + \alpha$. This would yield Theorem B.

Denote $h = h_\nu$ and $h_k = h_{\nu^{(k)}}, k = 1, \dots, K$. Find $\delta > 0$ so that for all the functions $f_j, d(\omega, \tilde{\omega}) < \delta$ implies $|f_j(\omega) - f_j(\tilde{\omega})| < \beta/3$. Then let $N(\delta)$ be the separation needed in the weak specification property. Applying Proposition 6.1 with $\alpha/2$ and $\beta/3$ in place of α and β to each of the ergodic measures $\nu^{(k)}$ we can find an $\varepsilon > 0$ and an n large enough good for all the measures $\nu^{(k)}$, and such that $N(\delta)/n$ is also sufficiently small. Fix such n, and fix (ε, C_n)-separated sets $S_k, 1 \leqq k \leqq K$ with the following properties:

$$|S_k| \geqq \exp(n^d(h_k - \alpha/2)), \tag{6.8}$$

$$\left| |C_n|^{-1} \sum_{m \in C_n} f_j(m\omega) - \int f_j d\nu^{(k)} \right| < \beta/3 \tag{6.9}$$

for any $\omega \in S_k$ and $j = 1, \dots, J$.

Let E_0 represent a parallelepiped in $\mathbb{Z}^d$ that contains disjoint translates $C_n^{(1)}, \dots, C_n^{(K)}$ of C_n such that

$$E_0 = \{m = (m_1, \dots, m_d) \in \mathbb{Z}^d : 0 \leqq m_1 < (n + 2N(\delta))K \text{ and } 0 \leqq m_i < n + 2N(\delta) \quad \text{for } i > 1\}$$

and $C_n^{(k)} = n_k + C_n$ where $n_k = \mathbf{N}(\delta) + (n + 2N(\delta))(k-1)e_1$. $e_1 = (1, 0, \dots, 0) \in \mathbb{Z}^d$ and $\mathbf{N}(\delta) \in \mathbb{Z}^d$ is the vector with all coordinates equal to $N(\delta)$. Let M be the set of those $m = (m_1, \dots, m_d) \in \mathbb{Z}^d$ for which m_1 is an integer multiple of $(n + 2N(\delta))K$

and other coordinates $m_i, i > 1$ are integer multiples of $(n + 2N(\delta))$. Then the disjoint translates $E_m = m + E_0, m \in M$ of E_0 tile $\mathbb{Z}^d$. For each cube C denote by M_C the subset of M so that if $m \in M_C$ then E_m is contained entirely in C. For the C in the definition of the set Ψ we will take any sufficiently large cube so that any translate $\tilde{C}$ of C may have at most $\beta|C|$ points which do not belong to E_m's with $m \in M$. Now assign to each $m + C_n^{(k)}, m \in M_C, k = 1, \ldots, K$ a point $\omega^{k,m} \in S_k$, in any way at all, and then using the weak specification find a point $\omega \in \Omega$ such that

$$d(l\omega, (l - m - n_k)\, \omega^{k,m}) \leqq \delta \text{ for all } \ell \in m + C_n^{(k)} . \tag{6.10}$$

Then $\omega \in \Psi$ provided $N(\delta)n^{-1}$ is sufficiently small. Suppose that $\delta \leqq \frac{1}{4}\varepsilon$, then the points ω chosen for different collections of the points $\omega^{k,m} \in S_k$ give rise to a $(\varepsilon/2, C)$-separated set S in Ψ. By (6.8),

$$|S| \geqq \exp\left((1 - \beta)|C||E_0|^{-1} n^d \left(\sum_{k=1}^{K} h_k - \frac{1}{2}\alpha K \right) \right) , \tag{6.11}$$

and by the definition of E_0.

$$|E_0| = K(n + 2N(\delta))^d . \tag{6.12}$$

Since we can produce this construction for arbitrarily large cubes C with the same set Ψ we will end up with a sequence of cubes $C^{(i)} \nearrow \infty$ as $i \to \infty$ and a corresponding sequence of $(\varepsilon/2, C^{(i)})$-separated sets $S(i), i = 1, 2, \ldots$, such that

$$\limsup_{i \to \infty} |C(i)|^{-1} \log |S(i)| \geqq (1 - \beta)|E_0|^{-1} n^d \left(\sum_{k=1}^{K} h_k - \frac{1}{2}\alpha K \right) . \tag{6.13}$$

The construction above goes through for any small β and large enough n, and so by (6.12) choosing β and n appropriately we can make the right-hand side of (6.13) to be not less than $h - \alpha$. This yields that the topological entropy of Ψ is at least $h - \alpha$.

References

[B] Bryc, W.: On the large deviation principle for stationary weakly dependent random fields. Ann. Prob. **20**, 1004–1030 (1992)

[C] Comets, F.: Grandes deviations pour des champs de Gibbs sur $\mathbb{Z}^d$. CRAS t. **303**, No. 11 511–513 (1986)

[DS] Deuschel, J.-D. Stroock, D.W.: Large Deviations, Boston: Acad. Press, 1989

[DSZ] Deuschel, J.-D. Stroock, D.W. Zessin, H.: Microcanonical distributions for lattice gases. Commun. Math. Phys. **139**, 83–101 (1991)

[DV] Donsker, M.D. Varadhan, S.R.S.: Asymptotic evaluation of certain Markov processes expectations for large time, IV. Commun. Pure. Appl. Math. **36**, 183–212 (1983)

[F] Föllmer, H.: On entropy and information gain in random fields. Z. Wahrsch. verw. Geb. **26**, 207–217 (1973)

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

[G] Georgii, H-O.: Gibbs Measures and Phase Transitions. Berlin: W. de Gruyter, 1988

[Ki1] Kifer, Y.: Large deviations in dynamical systems and stochastic processes. Trans. Am. Math. Soc. **321**, 505–524 (1990)

[Ki2] Kifer, Y.: Large deviations, averaging, and periodic orbits of dynamical systems, Commun. Math. Phys. **162**, 33–46 (1994)

[Kr] Krengel, U.: Ergodic Theorems. Berlin: W. de Gruyter, 1985

[M] Misiurewicz, M.: A short proof of the variational principle for a $\mathbb{Z}_+^N$ action on a compact space. Soc. Math. France Astérisque **40**, 147–157 (1976)

[O] Olla, S.: Large deviations for Gibbs random fields. Prob. Theory and Rel. Field, **77**, 343–357 (1988)

[OW] Ornstein, D. Weiss, B.: The Shannon-McMillan-Breinmann theorem for a class of amenable groups. Israel J. Math. **44**, 53–60 (1983)

[R1] Ruelle, D.: Statistical mechanics on a compact set with Z^ν action satisfying expansiveness and specification. Trans. Am. Math. Soc. **185**, 237–251 (1973)

[R2] Ruelle, D.: Thermodynamics Formalism. Reading, Mass.: Addison-Wesley, 1978

[S] Schmidt, K.: Algebraic Ideas in Ergodic Theory. CBMS Lecture Notes v. **76**, AMS, 1990

[T] Tempelman, A.: Ergodic Theorems for Group Actions. Dordrecht: Kluwer, 1992

[W] Weiss, B.: Strictly ergodic models for dynamical systems. Bull. AMS **13**:2, 143–146 (1985)

Communicated by J.-P. Eckmann