The Annals of Probability
1992, Vol. 20, No. 2, 1004–1030

ON THE LARGE DEVIATION PRINCIPLE FOR STATIONARY WEAKLY DEPENDENT RANDOM FIELDS

BY WŁODZIMIERZ BRYC

University of Cincinnati

The large deviation principle for the empirical field of a stationary $\mathbb{Z}^d$-indexed random field is proved under strong mixing dependence assumptions. The strong mixing coefficients considered allow us to separate the ratio-mixing condition used in the literature into a part directly responsible for the (nonuniform) large deviation principle and another one, which is used when the state space is noncompact. Results are applied to obtain variants of recent large deviation theorems for Markov chains and for Gibbs fields. The proofs are based on a new criterion for the large deviation principle which is stated in Appendix C.

0. Introduction. In this paper we prove the empirical-process-level large deviation principle for a "weakly dependent" random field on $\mathbb{Z}^d$. Our weak dependence assumptions are similar to, although different from, the strong mixing assumptions usually employed in the literature on limit theorems. The dependence conditions of our paper were suggested by Orey and Pelikan (1988), who consider $d = 1$ only; despite similarity, our assumptions seem not to be directly comparable with theirs (see Remark 5.1 below). Chiyonobu and Kusuoka (1988) consider large deviations under another mixing condition.

Section 1 sets up the notation; the strong mixing coefficients of dependence to be used throughout the paper are defined.

In Section 2 our main results are stated. Theorem 2.1 gives the large deviation principle for a compact-valued random field under strong mixing assumption (2.1). In Theorem 2.2, besides (2.1), an additional strong mixing assumption is used to obtain the large deviation principle for a Polish space-valued random field. In Theorem 2.3 this additional weak dependence assumption is further strengthened and we identify the rate function as the limit of finite dimensional entropies. The quotient of the two strong mixing measures of dependence used in our assumptions is equivalent to the ratio-mixing measure of dependence frequently used in large deviations (see Proposition 5.1). Thus our results can be interpreted as separating the "ratio-mixing" condition into a part directly responsible for the large deviation principle, and another one, which permits us to handle noncompact state spaces and to identify the rate function as entropy.

Large deviation principles are proved in Section 3. The proof is based on Theorem 7.1 in Appendix C and might be of independent interest. The rate function is identified in Section 4. In Appendix B dependence coefficients for

Received October 1989; revised September 1990.
AMS 1980 *subject classification*. Primary 60F10.
Key words and phrases. Large deviations, empirical process, strong mixing.

stationary Gibbs fields with summable interaction potentials are estimated. In Appendix C a criterion for the large deviation principle is stated. The main feature of this criterion is that after establishing exponential tightness, the proof of the large deviation principle is reduced to verifying that limit (7.1) below exists for a large enough class of functions. Similar approaches were employed by Baldi (1988), Comets (1989), Dawson and Gärtner (1987), Ellis (1985), Gärtner (1977), de Acosta (1988), Ney and Nummelin (1987a, b), Plachky (1971) and Sievers (1969); however these authors used the Gateaux differentiability of the limiting expression. In some instances, verifying differentiability is harder than establishing the existence of limit (7.1) for a larger class of functions than those authors needed.

1. Notation. Let $\mathbb{E}$ be a Polish (separable metric complete) space with the Borel σ-field $\mathscr{B}_{\mathbb{E}} = \mathscr{B}$. By $\mathscr{P}(\mathbb{E})$ we denote the set of all probability measures on $(\mathbb{E}, \mathscr{B})$; $\mathscr{P}(\mathbb{E})$ is considered as a separable complete metric space; $\mathbf{C}_b(\cdot)$ denotes the Banach space of all bounded continuous functions on a space $(\cdot)$ with the usual supremum norm. For $\mathbf{p} \in \mathscr{P}(\mathbb{E})$ and $F \in \mathbf{C}_b(\mathbb{E})$, we write $\mathbf{p}(F) = \int_{\mathbb{E}} F(\mathbf{x})\, d\mathbf{p}(\mathbf{x})$. The symbol $\Phi(\mathbf{p})$ is reserved for general continuous (nonlinear) functions of $\mathbf{p} \in \mathscr{P}(\mathbb{E})$; this last notation is used in Section 3 only.

The general framework to be considered is as follows. Throughout the paper we fix $d \geq 1$ and d commuting homeomorphisms $S_1, \dots, S_d$: $\mathbb{E} \to \mathbb{E}$. For $\mathbf{z} = (z(1), \dots, z(d)) \in \mathbb{Z}^d$, let $\mathbf{S}^{\mathbf{z}} = S_1^{z(1)} \circ \cdots \circ S_d^{z(d)}$. By $\mathscr{P}_S(\mathbb{E}) \subset \mathscr{P}(\mathbb{E})$ we shall denote the $\mathbf{S}^{\mathbf{z}}$ stationary probability measures, that is, $\mathbf{p} \in \mathscr{P}_S(\mathbb{E})$ iff $\mathbf{p} \circ S_i = \mathbf{p}$ for each $1 \leq i \leq d$. We fix a probability measure $P \in \mathscr{P}_S(\mathbb{E})$. By $E\{\cdot\}$ we denote the integral with respect to dP; $E^{\mathscr{F}}\{\cdot\}$ denotes the conditional P-expectation with respect to the σ-field $\mathscr{F} \subset \mathscr{B}$ and we write $P_{|\mathscr{F}}$ to denote the restriction of P to $\mathscr{F}$.

Besides P, we shall fix a Polish space $\mathbb{F}$ and a continuous function π: $\mathbb{E} \to \mathbb{F}$. By $\mathbf{C}_b(\pi)$ we denote the subset of $\mathbf{C}_b(\mathbb{E})$ consisting of those functions F which can be written as

$$F(\mathbf{x}) = H(\pi \circ \mathbf{S}^{\mathbf{z}(1)}(\mathbf{x}), \pi \circ \mathbf{S}^{\mathbf{z}(2)}(\mathbf{x}), \dots, \pi \circ \mathbf{S}^{\mathbf{z}(k)}(\mathbf{x})) \tag{1.1}$$

for some $k \geq 1$, $\mathbf{z}(1), \dots, \mathbf{z}(k) \in \mathbb{Z}^d$ and some $H \in \mathbf{C}_b(\mathbb{F}^k)$.

Throughout the paper we assume that π is such that $\mathbf{C}_b(\pi)$ separates points of $\mathscr{P}(\mathbb{E})$. Since $\mathbb{E}$ is a Polish space, this means we assume that $\mathbf{C}_b(\pi)$ separates points of $\mathbb{E}$. The formula

$$\mathbf{X}_{\mathbf{z}}(\mathbf{x}) := \mathbf{S}^{\mathbf{z}}(\mathbf{x}) \tag{1.2}$$

defines on the probability space $(\mathbb{E}, \mathscr{B})$ an $\mathbb{E}$-valued stationary random field indexed by $\mathbf{z} \in \mathbb{Z}^d$. We shall consider empirical distributions $\mu_n \in \mathscr{P}(\mathbb{E})$ associated with $\{\mathbf{X}_{\mathbf{z}}\}$ as follows. Let $\mathscr{C}_n$ be the unit n-cube in $\mathbb{Z}^d$, $n \in \mathbb{N}$, that is, $\mathscr{C}_n = \{\mathbf{z} \in \mathbb{Z}^d \colon 1 \leq z(i) \leq n, \forall\, i\}$. Define random measures μ_n: $\mathbb{E} \to \mathscr{P}(\mathbb{E})$ by

$$\mu_n = n^{-d} \sum_{\mathbf{z} \in \mathscr{C}_n} \delta_{\mathbf{X}_{\mathbf{z}}}, \qquad n \geq 1, \tag{1.3}$$

where $\delta_{\mathbf{x}} \in \mathscr{P}$ denotes the point mass at $\mathbf{x}$. Measures $\{\mu_n\}$ will be referred to as the "empirical fields." This terminology is motivated by the standard example $\mathbb{E} = \mathbb{F}^{\mathbb{Z}^d}$: here $\pi: \mathbb{E} \to \mathbb{F}$ is the projection on the $\mathbf{0}$th coordinate; S_i are shifts along coordinates in $\mathbb{Z}^d$, $i = 1, 2, \dots, d$. In this setup $\mathbf{X}_\mathbf{z} = \{Y_{\mathbf{z}+\mathbf{s}}\}_{\mathbf{s} \in \mathbb{Z}^d}$, where $\{Y_\mathbf{s}\}_{\mathbf{s} \in \mathbb{Z}^d}$ is an $\mathbb{F}$-valued stationary random field; clearly, $\mu_n = n^{-d} \sum_{\mathbf{z} \in \mathscr{C}_n} \delta_{\{Y_{\mathbf{z}+\mathbf{s}}\}_{\mathbf{s} \in \mathbb{Z}^d}}$ is a measure on the "realizations" (trajectories) of $\{Y_\mathbf{s}\}$. This paper could have been written in the notation of this example. However, in such an approach the assumptions of our theorems would be obscured by the particular topological properties of the product probability space. Besides, the less probability-space-dependent form of the large deviation theory is more natural in applications to dynamical systems, where, for example, (1.2) is written for nonproduct spaces $\mathbb{E}$.

We shall be interested in large deviation theorems for $\mathscr{P}(\mathbb{E})$-valued random variables $\{\mu_n\}$ (see Appendix C for a general statement of the large deviation principle). Our assumptions use the strong mixing dependence coefficients defined as follows. Let $\mathscr{F}, \mathscr{G}$ be σ-fields in $\mathscr{B}$. Following Bradley (1983) [see also Blum, Hanson and Koopmans (1963)], consider the following "measures of dependence" between pairs of σ-fields:

$$
\begin{aligned}
\psi_+(\mathscr{F}, \mathscr{G}) &= \sup\{P(A \cap B)/(P(A)P(B)): \\
&\qquad\qquad A \in \mathscr{F},\, B \in \mathscr{G},\, P(A)P(B) > 0\}, \\
\psi_-(\mathscr{F}, \mathscr{G}) &= \inf\{P(A \cap B)/(P(A)P(B)): \\
&\qquad\qquad A \in \mathscr{F},\, B \in \mathscr{G},\, P(A)P(B) > 0\}, \\
\psi(\mathscr{F}, \mathscr{G}) &= \psi_+(\mathscr{F}, \mathscr{G})/\psi_-(\mathscr{F}, \mathscr{G}).
\end{aligned} \tag{1.4}
$$

Clearly, $0 \le \psi_- \le 1 \le \psi_+ \le \infty$.

Let $\{\mathbf{X}_\mathbf{z}\}_{\mathbf{z} \in \mathbb{Z}^d}$ be defined by (1.2). Consider σ-fields

$$
\begin{aligned}
\mathscr{M}(n) &= \sigma\Big\{\pi(\mathbf{X}_\mathbf{z}):\ \min_{1 \le i \le d} z(i) \le n\Big\}, \\
\mathscr{F}(n) &= \sigma\Big\{\pi(\mathbf{X}_\mathbf{z}):\ \min_{1 \le i \le d} z(i) \ge n\Big\}, \qquad n \in \mathbb{N}.
\end{aligned} \tag{1.5}
$$

Notice that the definitions of $\mathscr{M}(n)$ and $\mathscr{F}(n)$ are not symmetric, except when $d = 1$. The dependence coefficients related to the so-called ψ-mixing condition in the strong mixing theory are defined as follows:

$$
\begin{aligned}
\psi_+(n) &= \psi_+(\mathscr{M}(0), \mathscr{F}(n)), \\
\psi_-(n) &= \psi_-(\mathscr{M}(0), \mathscr{F}(n)), \\
\psi(n) &= \psi_+(n)/\psi_-(n), \qquad n \ge 1.
\end{aligned}
$$

In the Markov case condition $\psi(1) < \infty$ is equivalent to the condition used, for example, in Stroock (1984), assumption (6.1), or in Ellis (1988), Hypothesis 1.1(b) (compare Lemma 5.2 below). Condition $\psi(1) < \infty$ can also be verified for the Gibbs field on $\mathbb{Z}^1$ with binary interactions $\Phi_2(k, x, y)$ such that $\sum_{k=1}^{\infty} k \|\Phi_2(k, x, y)\| < \infty$; this can be seen from Bowen (1975) (see the proof of his Proposition 1.14). However, if $d > 1$, then condition $\psi < \infty$ seems to be too restrictive [cf. Bradley (1989)]. To obtain the mixing condition, which is suitable for arbitrary $d \geq 1$ and at the same time easy to verify, we consider a smaller σ-field rather than $\mathscr{F}(n)$. Namely, let

$$\mathscr{G}(n) = \sigma\{\pi(\mathbf{X}_{\mathbf{z}}) : \mathbf{z} \in \mathscr{C}_n\} \tag{1.6}$$

and put

$$\begin{aligned} \psi_L^+(n) &= \psi_+(\mathscr{M}(-L), \mathscr{G}(n)), \\ \psi_L^-(n) &= \psi_-(\mathscr{M}(-L), \mathscr{G}(n)), \\ \psi_L(n) &= \psi_L^+(n)/\psi_L^-(n), \qquad L, n \geq 1. \end{aligned} \tag{1.7}$$

Notice that $1 \leq \psi_L^+(n) \leq \psi^+(L)$. Also $\psi_L^+(n) \geq \psi_{L+1}^+(n)$, which, in particular, implies that limit (2.3) below exists.

Besides the Markov and the m-dependent case, coefficients $\psi_L(n)$ can be effectively estimated for Gibbs fields with summable interactions (see Theorem 6.1 in Appendix B below). A version of coefficient $\psi_L(n)$ was used in the definition of "ratio-mixing" condition (RM) in Orey and Pelikan (1988) (see Remark 5.1 below).

Our main results, Theorems 2.1 and 2.2 below, use a dependence coefficient weaker than ψ_L^-. Namely, put

$$\begin{aligned} \tau_L^-(n) = \inf\Big\{ & L^{-d} \sum_{\mathbf{k} \in \mathscr{C}_L} P(A \cap \mathbf{S}^{-\mathbf{k}} B)/(P(A)P(B)) : \\ & A \in \mathscr{M}(-L),\ B \in \mathscr{G}(n),\ P(A)P(B) > 0 \Big\}. \end{aligned} \tag{1.8}$$

Clearly, $\psi_L^-(n)/L^d \leq \tau_L^-(n) \leq 1$ for each $n, L \geq 1$ [to see the last inequality put $B = \mathbb{E}$ in (1.8)]. It is also easily seen that $\tau_L^-(n+1) \leq \tau_L^-(n)$ and $\tau_{L+1}^-(n) \geq \tau_L^-(n)$, which implies that limit (2.1) below exists.

Definition (1.8) was motivated by Ellis (1988) and is devised to get the condition which can be verified for some periodic Markov chains [see (2.7) below]. Our proofs work as well with arbitrary weights in the sum on the right-hand side of (1.8) as long as a finite convex combination is formed. It would be interesting to know, however, if the sum over $\mathscr{C}_L$ can be replaced in (1.8) by the infinite sum as in Jain (1990) [e.g., (2.7) below implies both Jain (1990) $H_i(1)$ and $H_i(2)$].

2. Large deviation theorems. The following result was motivated by Orey and Pelikan (1988), Theorem 1.1. Our mixing condition allows a multidimensional index set $\mathbb{Z}^d$, $d \geq 1$, and is based on the weaker measure of dependence (see Remark 5.1 below). However, the large deviation principle obtained is nonuniform with respect to conditioning on $\mathscr{M}(0)$, even if $d = 1$.

THEOREM 2.1. *Let $\mathbb{E}$ be a compact metric space. Suppose $(\mathbf{X}_{\mathbf{z}})_{\mathbf{z} \in \mathbb{Z}^d}$ is defined by* (1.2) *and such that $\mathbf{C}_b(\pi)$ separates points of $\mathscr{P}(\mathbb{E})$. Suppose* *furthermore that (for this π) we have*

$$\lim_{L\to\infty}\liminf_{n\to\infty}\left(\tau_L^-(n)\right)^{1/n^d}=1. \tag{2.1}$$

Then family (1.3) *of empirical fields* $\{\mu_n\}$ *satisfies the large deviation principle with the rate function* $\mathbf{I}(\cdot)$ *determined by*

$$\mathbf{I}(\mathbf{p})=\sup\{\mathbf{p}(F)-\mathbb{L}(F)\colon F\in\mathbf{C}_b(\pi)\},\qquad \mathbf{p}\in\mathscr{P}(\mathbb{E}), \tag{2.2}$$

where $\mathbb{L}(F)=\lim_{n\to\infty}n^{-d}\log E\{\exp(n^d\mu_n(F))\}$; *the last limit exists and is finite.*

In our next result the compactness assumption is replaced by the requirement that $\psi_N^+(1)<\infty$ for some $N\geq 1$; to this end we need a stronger separation notion.

DEFINITION. We say that a continuous mapping $\pi\colon\mathbb{E}\to\mathbb{F}$ strongly separates points of $\mathbb{E}$, if for each $\varepsilon>0$ there are $\mathbf{z}=\mathbf{z}(\varepsilon)\in\mathbb{Z}^d$ and $\delta>0$ such that if $d_{\mathbb{E}}(\mathbf{x},\mathbf{y})>\varepsilon$, then $d_{\mathbb{F}}(\pi\circ\mathbf{S}^{\mathbf{z}}\mathbf{x},\pi\circ\mathbf{S}^{\mathbf{z}}\mathbf{y})>\delta$.

Clearly, if $\pi\colon\mathbb{E}\to\mathbb{F}$ strongly separates points of a Polish space $\mathbb{E}$, then $\mathbf{C}_b(\pi)$ separates points of $\mathscr{P}(\mathbb{E})$. Strong separation is a joint property of $\pi,\mathbf{S}$ and is satisfied, for example, if $\mathbb{E}=\mathbb{F}^{\mathbb{Z}^d}$ with shifts along coordinates; here π is a projection on the $\mathbf{0}$th coordinate and the metric is defined by

$$d_{\mathbb{E}}(\mathbf{x},\mathbf{y})=\sum_{\mathbf{k}\in\mathbb{Z}^d}2^{-|\mathbf{k}|}d_{\mathbb{F}}(\pi\circ\mathbf{S}^{\mathbf{k}}\mathbf{x},\pi\circ\mathbf{S}^{\mathbf{k}}\mathbf{y}),$$

where $2^{-|\mathbf{k}|}=2^{-|k(1)|-\cdots-|k(d)|}$ and $d_{\mathbb{F}}(\cdot,\cdot)$ is a bounded metric on $\mathbb{F}$.

THEOREM 2.2. *Let $\mathbb{E}$ be a Polish space. Suppose $\{\mathbf{X}_{\mathbf{z}}\}_{\mathbf{z}\in\mathbb{Z}^d}$ is defined by* (1.2) *and that π strongly separates points of $\mathbb{E}$. Suppose furthermore that $\psi_N^+(1)<\infty$ for some $N\geq 1$ and that* (2.1) *holds. Then family* (1.3) *of empirical fields $\{\mu_n\}$ satisfies the large deviation principle with the rate function $\mathbf{I}(\cdot)$ determined by* (2.2).

The next result complements Olla (1987) and Orey and Pelikan (1988); we consider $d\geq 1$ and we also do not use the continuity of conditional distributions. The latter generalizes Ellis (1988), Remark 2.1, to non-Markovian situations and multidimensional index set. The result is stated and proved for product space $\mathbb{E}=\mathbb{F}^{\mathbb{Z}^d}$ only; this product representation is used in Lemma 4.1 below only.

THEOREM 2.3. *Suppose $\mathbb{E}=\mathbb{F}^{\mathbb{Z}^d}$, where $\mathbb{F}$ is a Polish space and $d\geq 1$. Let π be the projection on the $\mathbf{0}$th coordinate. Assume $P\in\mathscr{P}_S(\mathbb{E})$ is such that* (2.1) *holds and*

$$\lim_{L\to\infty}\limsup_{n\to\infty}\left(\psi_L^+(n)\right)^{1/n^d}=1. \tag{2.3}$$

Then (1.3) *defines measures* $\{\mu_n\}$ *satisfying the large deviation principle with the rate function* $\mathbf{I}\colon \mathscr{P}(\mathbb{E}) \to [0,\infty]$ *given by*

$$\mathbf{I}(\mathbf{p}) = \infty, \quad \text{if } \mathbf{p} \notin \mathscr{P}_S(\mathbb{E}), \tag{2.4}$$

$$\mathbf{I}(\mathbf{p}) = \infty, \tag{2.5}$$

if there is $n \geq 1$ *such that* $\mathbf{p}_{|\mathscr{I}(n)}$ *is not absolutely continuous with respect to* $P_{|\mathscr{I}(n)}$,

$$\mathbf{I}(\mathbf{p}) = \lim_{n\to\infty} n^{-d} \int_{\mathbb{E}} f_n(\mathbf{x}) \log f_n(\mathbf{x})\, dP(\mathbf{x}), \tag{2.6}$$

if $\mathbf{p} \in \mathscr{P}_S(\mathbb{E}), \mathbf{p}_{|\mathscr{I}(n)}$ *is absolutely continuous with respect to* $P_{|\mathscr{I}(n)}$ *for all* $n \geq 1$ *and* $f_n = d\mathbf{p}_{|\mathscr{I}(n)}/dP_{|\mathscr{I}(n)}$.

In particular, the proof shows that the limit on the right-hand side of (2.6) exists, possibly being ∞. Theorems 2.1 and 2.2 are proved in Section 3. Theorem 2.3 is proved in Section 4. Below we state corollaries which illustrate the applicability of our theorems to some recent large deviation results. The following corollary of Theorem 2.1 gives the "nonuniform" large deviation principle for Markov chains. The same result also follows from de Acosta (1990), Section 5 (after dropping the Feller property in the assumption of his theorem 5.3) [see also Jain (1990)]; for a noncompact state space additional assumptions seem to be needed.

COROLLARY 2.1. *Suppose* $P_n(\mathbf{x}, d\mathbf{y})$ *is the family of* n*-step transition probability measures of a time-homogeneous and stationary Markov chain* $\{Y_n\}_{n\geq 0}$ *with a compact state space* $\mathbb{F}$ *and with invariant initial distribution* $p \in \mathscr{P}(\mathbb{F})$. *If for some* $N \geq 1$, $C < \infty$,

$$p(A) \leq C \sum_{k=1}^{N} P_k(\mathbf{x}, A) \tag{2.7}$$

for each $A \in \mathscr{B}(\mathbb{F})$, $\mathbf{x} \in \mathbb{F}$, *then the family of empirical fields* $\{\mu_n\}$ *defined on* $\mathbb{F}^{\mathbb{N}}$ *by*

$$\mu_n = n^{-1} \sum_{k=1}^{n} \delta_{\{Y_{s+k}\}_{s\in\mathbb{N}}} \tag{2.8}$$

satisfies the large deviation principle.

PROOF. Extend $\{Y_n\}_{n\geq 0}$ to $\{Y_n\}_{n\in\mathbb{Z}}$. By (5.3) in Appendix A we have $1 \geq \tau_L^-(n) \geq \tau_N^-(n) \geq (CN)^{-1}$ for each $L \geq N$, $n \geq 1$. Hence $\lim_{n\to\infty}(\tau_L^-(n))^{1/n} = 1$ for each $L \geq N$ and (2.1) holds. The corollary follows from Theorem 2.1 by the contraction principle. □

The following corollary of Theorem 2.2 is known in the uniform (i.e., stronger) version [see Deuschel and Stroock (1989), 4.4.12; see also Ellis and Wayne (1989) and de Acosta (1990), Section 5]. While there is no major difficulty in extending our proof to arbitrary initial probability distributions, in order to make the conclusion hold uniformly over the initial points, revisions in condition (7.1) below would be needed.

COROLLARY 2.2. *Suppose $P_n(\mathbf{x}, d\mathbf{y})$ is the family of n-step transition probability measures of a time-homogeneous and stationary Markov chain $\{Y_n\}_{n\geq 0}$ with a Polish state space $\mathbb{F}$. If there are $C < \infty$ and $N, M \geq 1$ such that for each $A \in \mathscr{B}_{\mathbb{F}}$, $\mathbf{x} \in \mathbb{F}$, $\mathbf{y} \in \mathbb{F}$,*

$$P_M(\mathbf{x}, A) \leq C \sum_{k=1}^{N} P_k(\mathbf{y}, A), \tag{2.9}$$

then empirical measures $\{\mu_n\}$ defined on $(\mathbb{F}^{\mathbb{N}}, \mathscr{B})$ by (2.8) satisfy the large deviation principle.

PROOF. Extend $\{Y_n\}_{n\geq 0}$ to $\{Y_n\}_{n\in\mathbb{Z}}$. By (5.4) and (5.5) (see Appendix A) we have $\psi_M^+(n) \leq a < \infty$ and $\tau_N^-(n) \geq b > 0$, $\forall\ n \geq 1$, for some constants a, b. Hence $\psi_M^-(1) < \infty$ and $\lim_{n\to\infty}(\tau_L^-(n))^{1/n} = 1$ for each $L \geq N$. The result follows from Theorem 2.2 by the contraction principle. □

The following corollary gives a non-Markovian application of Theorem 2.2, and also illustrates the importance of allowing a multidimensional index set. The result is known [see Comets (1986) for a different proof]; Föllmer and Orey (1988) and Olla (1988) have similar results under an additional compactness assumption. Our proof is based on Theorem 6.1 in Appendix B; the latter might be of independent interest [compare Bowen (1975), Proposition 1.14].

COROLLARY 2.3. *Suppose $\{Y_{\mathbf{z}}\}_{\mathbf{z}\in\mathbb{Z}^d}$ is a stationary Gibbs field with a countable state space $\mathbb{F}$ and stationary summable interaction potential [as defined by (6.2) in Appendix B]. Then empirical fields $\{\mu_n\}$, defined on $\mathbb{E} = \mathbb{F}^{\mathbb{Z}^d}$ by (1.3) with $\{\mathbf{X}_{\mathbf{z}}\} = \{Y_{\mathbf{z}+\mathbf{s}}\}_{\mathbf{s}\in\mathbb{Z}^d}$, satisfy the large deviation principle.*

PROOF. The result follows directly from Theorem 2.2 which can be applied by Theorem 6.1 in Appendix B. □

3. Proof of the large deviation principle. Let $F_1, F_2, \dots, F_k \in \mathbf{C}_b(\pi)$ and define Φ: $\mathscr{P}(\mathbb{E}) \to \mathbb{R}$ by

$$\Phi(\mathbf{p}) = \mathbf{p}(F_1) \wedge \cdots \wedge \mathbf{p}(F_k). \tag{3.1}$$

Denote by *Concave* the set of all functions Φ defined by (3.1), $k = 1, 2, \dots$.

The plan of the proof is as follows: In Claim I we show that the distributions of the sequence $\{\mu_n\}$ admit an asymptotic value over *Concave* [for the definition see (3.2) or (7.1) below]; in Claim II we show that the rate function is convex; in Claim III we verify exponential tightness (for the definition see Appendix C). The result will then follow as an application of Theorem 7.1 below.

CLAIM I. *If* $\mathbb{E}$ *is a Polish space and* (2.1) *holds, then the asymptotic value*

$$\mathbb{L}(\Phi) := \lim_{n\to\infty} n^{-d} \log E\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{C}_n} F_i(\mathbf{X}_{\mathbf{z}})\right)\right\} \tag{3.2}$$

exists and is a finite number for each choice of $F_1, \dots, F_k \in \mathbf{C}_b(\pi)$, $k \geq 1$.

Note that $\mathbb{L}(\Phi) = \lim_{n\to\infty} n^{-d} \log E\{\exp n^d \Phi(\mu_n)\}$ indeed depends on Φ only, and (3.2) is consistent with (7.1). Define

$$\underline{M}(n) = \operatorname{ess\,inf} E^{\mathscr{M}(0)}\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{C}_n} F_i(\mathbf{X}_{\mathbf{z}})\right)\right\}, \qquad n \in \mathbb{N}. \tag{3.3}$$

By $\mathbf{C}_{b-}(\pi)$ we denote the subset of $\mathbf{C}_b(\pi)$ consisting of those functions F which can be written using (1.1) with $\mathbf{z}(1), \dots, \mathbf{z}(k) \in \{\mathbf{z}\colon \max_{1\leq i\leq d} z_i \leq 0\}$.

The proof of Claim I consists of a series of lemmas. The first three lemmas prove that $\lim_{n\to\infty} n^{-d} \log \underline{M}(n)$ exists, if $F_k \in \mathbf{C}_{b-}(\pi)$, $k \geq 1$. The last two lemmas show that the limit gives indeed what is needed, namely $\mathbb{L}(\Phi)$.

The lemmas will also be used in Section 4, where bounded measurable functions are considered; therefore, we give a more general statement than what is needed here. Let $\mathbf{M}_b(\mathbb{E})$ denote the family of bounded measurable functions on $(\mathbb{E}, \mathscr{B})$. In analogy with our notation for the continuous functions, we denote by $\mathbf{M}_b(\pi)$ the set of all bounded functions that are measurable with respect to the σ-field $\mathscr{M}(N) \cap \mathscr{F}(-N)$ for some large enough N; $\mathbf{M}_{b-}(\pi)$ is defined as the set of all $\sigma(\mathbf{C}_{b-}(\pi))$-measurable elements of $\mathbf{M}_b(\pi)$.

LEMMA 3.1. *If* $F_1, \dots, F_k, G_1, \dots, G_k \in \mathbf{M}_{b-}(\pi)$, *then for each* $\mathbf{m} \in \mathbb{N}^d$ *and all finite sets* $\mathscr{U}, \mathscr{V} \in \mathbb{Z}^d$ *such that* $\mathscr{U} \subset \{\mathbf{z}\colon \min_{1\leq i\leq d}\{z(i) - m(i)\} \leq 0\}$, $\mathscr{V} \subset \{\mathbf{z}\colon \min_{1\leq i\leq d}\{z(i) - m(i)\} \geq 1\}$ *we have*

$$\begin{aligned}
&\operatorname{ess\,inf} E^{\mathscr{M}(0)}\left\{\exp\left(\min_{1\leq i\leq k} \sum_{\mathbf{z}\in\mathscr{U}} F_i(\mathbf{X}_{\mathbf{z}}) + \min_{1\leq i\leq k} \sum_{\mathbf{z}\in\mathscr{V}} G_i(\mathbf{X}_{\mathbf{z}})\right)\right\} \\
&\quad \geq \operatorname{ess\,inf} E^{\mathscr{M}(0)}\left\{\exp\left(\min_{i\leq k} \sum_{\mathbf{z}\in\mathscr{U}} F_i(\mathbf{X}_{\mathbf{z}})\right)\right\} \\
&\qquad \times \operatorname{ess\,inf} E^{\mathscr{M}(0)}\left\{\exp\left(\min_{i\leq k} \sum_{\mathbf{z}\in\mathscr{V}-\mathbf{m}} G_i(\mathbf{X}_{\mathbf{z}})\right)\right\}.
\end{aligned}$$

PROOF. For $\mathbf{m} \in \mathbb{Z}^d$ put $\mathscr{M}_{\mathbf{m}} = \mathbf{S}^{-\mathbf{m}}(\mathscr{M}(0))$. Clearly,

$$\begin{aligned}
&E^{\mathscr{M}(0)}\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{U}} F_i(\mathbf{X}_{\mathbf{z}}) + \min_i \sum_{\mathbf{z}\in\mathscr{V}} G_i(\mathbf{X}_{\mathbf{z}})\right)\right\} \\
&\quad = E^{\mathscr{M}(0)}\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{U}} F_i(\mathbf{X}_{\mathbf{z}})\right) E^{\mathscr{M}_{\mathbf{m}}}\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{V}} G_i(\mathbf{X}_{\mathbf{z}})\right)\right\}\right\};
\end{aligned}$$

this follows from the assumption that $F_1, \dots, F_k \in \mathbf{M}_{b-}(\pi)$, so that

$\min_i \sum_{\mathbf{z}\in\mathscr{U}} F_i(\mathbf{X_z})$ is $\mathscr{M}_{\mathbf{m}}$-measurable. Therefore,

$$\begin{aligned} E^{\mathscr{M}(0)}&\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{U}} F_i(\mathbf{X_z}) + \min_i \sum_{\mathbf{z}\in\mathscr{V}} G_i(\mathbf{X_z})\right)\right\} \\ &\geq E^{\mathscr{M}(0)}\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{U}} F_i(\mathbf{X_z})\right) \operatorname{ess\,inf} E^{\mathscr{M}_{\mathbf{m}}}\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{V}} G_i(\mathbf{X_z})\right)\right\}\right\} \\ &= \operatorname{ess\,inf} E^{\mathscr{M}(0)}\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{V}-\mathbf{m}} G_i(\mathbf{X_z})\right)\right\} E^{\mathscr{M}(0)}\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{U}} F_i(\mathbf{X_z})\right)\right\} \end{aligned}$$

and taking the essential infimum ends the proof. □

LEMMA 3.2. *If $F_1, \ldots, F_k \in \mathbf{M}_{b-}(\pi)$ and $\underline{M}(n)$ is defined by* (3.3), *then*

$$\underline{M}(kn) \geq \left(\underline{M}(n)\right)^{k^d} \quad \text{for each } k, n \geq 1.$$

PROOF. Write $\mathscr{C}_{kn} = \bigcup_{0\leq i(1),\ldots,i(d)\leq k-1}\{n\mathbf{i} + \mathscr{C}_n\}$ as the union of k^d disjoint translations of cubes $\mathscr{C}_n$. The result follows from Lemma 3.1 applied k^d times to the finite sets $\mathscr{U}_{\mathbf{m}} = \bigcup_{1\leq i(1)\leq m(1),\ldots,1\leq i(d)\leq m(d)}\{n(\mathbf{i} - \mathbf{1}) + \mathscr{C}_n\}$ and $\mathscr{V}_{\mathbf{m}} = \mathbf{m} + \mathscr{C}_n$ [with the appropriate choices of $\mathbf{m}$; $\mathbf{1} = (1, \ldots, 1)$]. Indeed,

$$\min_{1\leq j\leq k} \sum_{\mathbf{z}\in\mathscr{U}_{\mathbf{m}}\cup\mathscr{V}_{\mathbf{m}}} F_j(\mathbf{X_z}) \geq \min_{1\leq j\leq k} \sum_{\mathbf{z}\in\mathscr{U}_{\mathbf{m}}} F_j(\mathbf{X_z}) + \min_{1\leq j\leq k} \sum_{\mathbf{z}\in\mathscr{V}_{\mathbf{m}}} F_j(\mathbf{X_z}),$$

and Lemma 3.1 applies. □

LEMMA 3.3. *If $F_1, \ldots, F_k \in \mathbf{M}_{b-}(\pi)$ and $\underline{M}(n)$ is defined by* (3.3), *then*

$$n^{-d}\log \underline{M}(n) \to \sup_m m^{-d}\log \underline{M}(m) \quad \text{as } n\to\infty.$$

In particular, $\lim_{n\to\infty} n^{-d}\log\underline{M}(n)$ *exists.*

PROOF. For $d = 1$ this is a well-known consequence of Lemma 3.2 [see, e.g., Dunford and Schwartz (1958), 8.1.1]. Case $d > 1$ is handled similarly. □

LEMMA 3.4. *If $F_1, \ldots, F_k \in \mathbf{M}_{b-}(\pi)$, then*

$$n^{-d}\left(\log\left(E\left\{\exp\left(\min_{1\leq i\leq k} \sum_{\mathbf{z}\in\mathscr{C}_n} F_i(\mathbf{X_z})\right)\right\}\right) - \log \underline{M}(n)\right) \to 0 \quad \text{as } n\to\infty.$$

PROOF. Since trivially $\underline{M}(n) \leq E\{\exp(\min_i \sum_{\mathbf{z}\in\mathscr{C}_n} F_i(\mathbf{X_z}))\}$, we need only to show that

$$\liminf_{n\to\infty}\left(\underline{M}(n)/E\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{C}_n} F_i(\mathbf{X_z})\right)\right\}\right)^{1/n^d} \geq 1.$$

Take $n, N \in \mathbb{N}$. Since $F_1, \ldots, F_k \in \mathbf{M}_{b-}(\pi)$, increasing N if necessary, we may assume that each $F_i(\mathbf{X}_0)$ is $\mathscr{F}(-N)$-measurable, $1 \le i \le k$. Write $\mathbf{N} = (N, \dots, N) \in \mathbb{N}^d$. Clearly, for each $n > 2N$ we have

$$\min_i \sum_{\mathbf{z}\in\mathscr{C}_n} F_i(\mathbf{X}_\mathbf{z}) \ge -\left(n^d - (n-2N)^d\right)\max_i \|F_i\|_\infty + \max_{\mathbf{r}\in\mathscr{C}_N} \min_i \sum_{\mathbf{z}\in 2\mathbf{N}-\mathbf{r}+\mathscr{C}_{n-2N}} F_i(\mathbf{X}_\mathbf{z}).$$

Therefore, taking stationarity into account, we get

$$\begin{aligned}\underline{M}(n) &\ge e^{C_N(n)} \operatorname{ess\,inf}\left\{E^{\mathscr{M}(0)}\left\{\max_{\mathbf{r}\in\mathscr{C}_N} \exp\left(\min_i \sum_{\mathbf{z}\in 2\mathbf{N}-\mathbf{r}+\mathscr{C}_{n-2N}} F_i(\mathbf{X}_\mathbf{z})\right)\right\}\right\} \\ &= e^{C_N(n)} \operatorname{ess\,inf}\left\{E^{\mathscr{M}(-N)}\left\{\max_{\mathbf{r}\in\mathscr{C}_N} \exp\left(\min_i \sum_{\mathbf{z}\in \mathbf{N}-\mathbf{r}+\mathscr{C}_{n-2N}} F_i(\mathbf{X}_\mathbf{z})\right)\right\}\right\} \\ &\ge e^{C_N(n)} N^{-d} \operatorname{ess\,inf}\left\{E^{\mathscr{M}(-N)}\left\{\sum_{\mathbf{r}\in\mathscr{C}_N} \exp\left(\min_i \sum_{\mathbf{z}\in \mathbf{N}-\mathbf{r}+\mathscr{C}_{n-2N}} F_i(\mathbf{X}_\mathbf{z})\right)\right\}\right\},\end{aligned}$$

where $C_N(n) = -(n^d - (n-2N)^d)\max_i \|F_i\|_\infty$. Since, by our choice of N, $\min_i \sum_{\mathbf{z}\in\mathbf{N}+\mathscr{C}_{n-2N}} F_i(\mathbf{X}_\mathbf{z})$ is $\mathscr{G}(n)$-measurable, by (5.2) we obtain

$$\underline{M}(n) \ge e^{C_N(n)} \tau_N^-(n) E\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{C}_{n-2N}} F_i(\mathbf{X}_\mathbf{z})\right)\right\}. \tag{3.4}$$

Trivially, we have again

$$\min_i \sum_{\mathbf{z}\in\mathscr{C}_{n-2N}} F_i(\mathbf{X}_\mathbf{z}) \ge -\left(n^d - (n-2N)^d\right)\max_i \|F_i\|_\infty + \min_i \sum_{\mathbf{z}\in\mathscr{C}_n} F_i(\mathbf{X}_\mathbf{z}),$$

which together with (3.4) implies

$$\underline{M}(n) \ge e^{2C_N(n)} \tau_N^-(n) E\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{C}_n} F_i(\mathbf{X}_\mathbf{z})\right)\right\}. \tag{3.5}$$

Since $C_N(n)/n^d \to 0$ as $n \to \infty$ and $\liminf_{n\to\infty} (\tau_N^-(n))^{1/n^d} = 1 - \varepsilon_N$, therefore

$$\liminf_{n\to\infty} \left(\underline{M}(n) / E\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{C}_n} F_i(\mathbf{X}_\mathbf{z})\right)\right\}\right)^{1/n^d} \ge 1 - \varepsilon_N$$

and the proof is completed, as ε_N can be taken arbitrarily close to 0 [see (2.1)]. □

PROOF OF CLAIM I. Let $F_1, \dots, F_k \in \mathbf{C}_b(\pi)$ be fixed. By stationarity,

$$E\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{C}_n} F_i(\mathbf{X}_\mathbf{z})\right)\right\} = E\left\{\exp\left(\min_i \sum_{\mathbf{z}\in\mathscr{C}_n} G_i(\mathbf{X}_\mathbf{z})\right)\right\},$$

where $G_i(\mathbf{x}) = F_i(\mathbf{S}^{-\mathbf{N}}\mathbf{x})$ for some $\mathbf{N} = (N, \dots, N) \in \mathbb{N}^d$. Taking N large enough, we may ensure that $G_1, \dots, G_k \in \mathbf{C}_{b-}(\pi)$; therefore, by Lemmas 3.4 and 3.3 the sequence $n^{-d} \log E\{\exp(\min_i \sum_{\mathbf{z} \in \mathscr{C}_n} F_i(\mathbf{X}_\mathbf{z}))\}$, $n \geq 1$, converges as $n \to \infty$. Moreover, since $|n^{-d} \log E\{\exp(\min_i \sum_{\mathbf{z} \in \mathscr{C}_n} F_i(\mathbf{X}_\mathbf{z}))\}| \leq \max_i \|F_i\|_\infty$ for each n, therefore the limit is a finite number. □

CLAIM II. *If $\mathbb{E}$ is a Polish space and* (2.1) *holds, then $\mathbf{I}_0(\cdot)$ defined by*

$$\mathbf{I}_0(\mathbf{p}) = \sup\{\Phi(\mathbf{p}) - \mathbb{L}(\Phi) : \Phi \in \textit{Concave}\}$$

is a convex function. [*Recall that* $\Phi(\mathbf{p}) = \min_i \int F_i(\mathbf{x})\, d\mathbf{p}(\mathbf{x})$; *see* (3.1).]

We shall need two auxiliary observations.

LEMMA 3.5. *Fix $\mathbf{q}_1, \mathbf{q}_2 \in \mathscr{P}(\mathbb{E})$. For each $\Phi \in$ Concave there are $\Phi_1, \Phi_2 \in$ Concave such that*

$$\Phi\big(\tfrac{1}{2}(\mathbf{p}_1 + \mathbf{p}_2)\big) \geq \tfrac{1}{2}\Phi_1(\mathbf{p}_1) + \tfrac{1}{2}\Phi_2(\mathbf{p}_2) \quad \textit{for all } \mathbf{p}_1, \mathbf{p}_2 \in \mathscr{P}(\mathbb{E}), \tag{3.6}$$

$$\Phi\big(\tfrac{1}{2}(\mathbf{q}_1 + \mathbf{q}_2)\big) = \Phi_1(\mathbf{q}_1) = \Phi_2(\mathbf{q}_2). \tag{3.7}$$

PROOF. Let $F_1, \dots, F_k$ represent Φ [see (3.1)]. Define

$$G_i(\mathbf{x}) = F_i(\mathbf{x}) + \tfrac{1}{2}\big(\mathbf{q}_2(F_i) - \mathbf{q}_1(F_i)\big),$$
$$H_i(\mathbf{x}) = F_i(\mathbf{x}) + \tfrac{1}{2}\big(\mathbf{q}_1(F_i) - \mathbf{q}_2(F_i)\big), \qquad 1 \leq i \leq k.$$

Obviously, $G_1, G_2, \dots, G_k, H_1, H_2, \dots, H_k \in \mathbf{C}_b(\pi)$. Let Φ_1 be determined by (3.1) with functions $\{G_i\}$ and let Φ_2 be determined by (3.1) with functions $\{H_i\}$. Then it is easy to check that both (3.6) and (3.7) hold. Indeed, we have $\mathbf{p}_1(F_i) + \mathbf{p}_2(F_i) = \mathbf{p}_1(G_i) + \mathbf{p}_2(H_i)$ for all $\mathbf{p}_1, \mathbf{p}_2 \in \mathscr{P}$, $1 \leq i \leq k$. Therefore,

$$\begin{aligned}\Phi(\mathbf{p}_1 + \mathbf{p}_2) &= (\mathbf{p}_1 + \mathbf{p}_2)(F_1) \wedge \cdots \wedge (\mathbf{p}_1 + \mathbf{p}_2)(F_k)\\ &\geq \mathbf{p}_1(G_1) \wedge \cdots \wedge \mathbf{p}_1(G_k) + \mathbf{p}_2(H_1) \wedge \cdots \wedge \mathbf{p}_2(H_k)\\ &= \Phi(\mathbf{p}_1) + \Phi_2(\mathbf{p}_2),\end{aligned}$$

which proves (3.6). Equality (3.7) is trivial, as $\mathbf{q}_1(G_i) = \mathbf{q}_2(H_i) = \frac{1}{2}(\mathbf{q}_1 + \mathbf{q}_2)(F_i)$ for each $1 \leq i \leq k$. □

LEMMA 3.6. *Suppose $\Phi, \Phi_1, \Phi_2 \in$ Concave are the concave functions from the conclusion of Lemma* 3.5 *and assume that* (2.1) *holds. Then*

$$\mathbb{L}(\Phi) \geq \tfrac{1}{2}\big(\mathbb{L}(\Phi_1) + \mathbb{L}(\Phi_2)\big).$$

PROOF. Write $\Phi = \Phi_0$ for the next line only. By the proof of Claim I we have

$$\mathbb{L}(\Phi_r) = \lim_{n \to \infty} (2n)^{-d} \log \operatorname{ess\,inf} E^{\mathscr{M}(0)}\left\{\exp\left(\Phi_r\left(\sum_{\mathbf{z} \in \mathscr{C}_{2n}} \delta_{\mathbf{X}_{\mathbf{z}-\mathbf{N}}}\right)\right)\right\}, \qquad r = 0, 1, 2,$$

for some $\mathbf{N} = (N, \dots, N) \in \mathbb{N}^d$. By (3.6), multiplied by $2^d n^d$, we have

$$\Phi\biggl(\sum_{\mathbf{z}\in\mathscr{C}_{2n}} \delta_{\mathbf{X}_{\mathbf{z}-\mathbf{N}}}\biggr) \ge \Phi_1\biggl(\sum_{\mathbf{a}} \sum_{\mathbf{z}\in\mathbf{a}+\mathscr{C}_n} \delta_{\mathbf{X}_{\mathbf{z}-\mathbf{N}}}\biggr) + \Phi_2\biggl(\sum_{\mathbf{b}} \sum_{\mathbf{z}\in\mathbf{b}+\mathscr{C}_n} \delta_{\mathbf{X}_{\mathbf{z}-\mathbf{N}}}\biggr),$$

where $\mathbf{a} \in \{(0, k_1, \dots, k_{d-1})\colon k_i = 0 \text{ or } k_i = n,\ 1 \le i \le d-1\}$, $\mathbf{b} \in \{(n, k_1, \dots, k_{d-1})\colon k_i = 0 \text{ or } k_i = n,\ 1 \le i \le d-1\}$, that is, we split the sum over $\mathscr{C}_{2n}$ into the sum over 2^d disjoint translations of $\mathscr{C}_n$, half of them being assigned to Φ_1 and another half to Φ_2 (clearly, the particular form of the partition given above is of no importance). Therefore,

$$\begin{aligned}
&\operatorname{ess\,inf} E^{\mathscr{M}(0)}\biggl\{\exp\biggl(\Phi\biggl(\sum_{\mathbf{z}\in\mathscr{C}_{2n}} \delta_{\mathbf{X}_{\mathbf{z}-\mathbf{N}}}\biggr)\biggr)\biggr\}\\
&\quad\ge \operatorname{ess\,inf} E^{\mathscr{M}(0)}\biggl\{\exp\biggl(\min_i\biggl(\sum_{\mathbf{a}} \sum_{\mathbf{z}\in\mathbf{a}+\mathscr{C}_n} G_i(\mathbf{X}_{\mathbf{z}-\mathbf{N}})\biggr)\\
&\qquad\qquad + \min_i\biggl(\sum_{\mathbf{b}} \sum_{\mathbf{z}\in\mathbf{b}+\mathscr{C}_n} H_i(\mathbf{X}_{\mathbf{z}-\mathbf{N}})\biggr)\biggr)\biggr\},
\end{aligned}$$

where G_i and H_i, $1 \le i \le k$, represent Φ_1 and Φ_2 respectively [see (3.1)]. Using Lemma 3.1 2^d times, we get by stationarity that

$$\begin{aligned}
&\log \operatorname{ess\,inf} E^{\mathscr{M}(0)}\biggl\{\exp\biggl(\Phi\biggl(\sum_{\mathbf{z}\in\mathscr{C}_{2n}} \delta_{\mathbf{X}_{\mathbf{z}-\mathbf{N}}}\biggr)\biggr)\biggr\}\\
&\quad\ge 2^{d-1} \log \operatorname{ess\,inf} E^{\mathscr{M}(0)}\biggl\{\exp\biggl(\Phi_1\biggl(\sum_{\mathbf{z}\in\mathscr{C}_n} \delta_{\mathbf{X}_{\mathbf{z}-\mathbf{N}}}\biggr)\biggr)\biggr\}\\
&\qquad + 2^{d-1} \log \operatorname{ess\,inf} E^{\mathscr{M}(0)}\biggl\{\exp\biggl(\Phi_2\biggl(\sum_{\mathbf{z}\in\mathscr{C}_n} \delta_{\mathbf{X}_{\mathbf{z}-\mathbf{N}}}\biggr)\biggr)\biggr\}.
\end{aligned}$$

Dividing by $2^d n^d$ and passing to the limit as $n \to \infty$, we end the proof. □

PROOF OF CLAIM II. To prove the convexity of a lower semicontinuous function $\mathbf{I}_0(\cdot)$, it is enough to show that

$$\mathbf{I}_0\bigl(\tfrac{1}{2}(\mathbf{q}_1 + \mathbf{q}_2)\bigr) \le \tfrac{1}{2}\bigl(\mathbf{I}_0(\mathbf{q}_1) + \mathbf{I}_0(\mathbf{q}_2)\bigr) \quad \text{for all } \mathbf{q}_1, \mathbf{q}_2 \in \mathscr{P}(\mathbb{E}) \tag{3.8}$$

[see, e.g., Stroock (1984), page 39]. Fix $\mathbf{q}_1, \mathbf{q}_2 \in \mathscr{P}(\mathbb{E})$ and $\Phi \in$ *Concave*. From Lemmas 3.5 and 3.6 we have

$$\begin{aligned}
\Phi\bigl(\tfrac{1}{2}(\mathbf{q}_1 + \mathbf{q}_2)\bigr) - \mathbb{L}(\Phi) &\le \tfrac{1}{2}\bigl(\Phi_1(\mathbf{q}_1) + \Phi_2(\mathbf{q}_2)\bigr) - \tfrac{1}{2}\bigl(\mathbb{L}(\Phi_1) + \mathbb{L}(\Phi_2)\bigr)\\
&\le \tfrac{1}{2}\bigl(\mathbf{I}_0(\mathbf{q}_1) + \mathbf{I}_0(\mathbf{q}_2)\bigr).
\end{aligned}$$

Since $\Phi \in$ *Concave* was arbitrary, taking the supremum over all Φ ends the proof of (3.8). □

CLAIM III. *If $\mathbb{E}$ is a Polish space, π strongly separates points of $\mathbb{E}$ and $\psi_N^+(1) < \infty$ for some $N \geq 1$, then* (1.3) *defines an exponentially tight family $\{\mu_n\}_{n\geq 1}$ of $\mathscr{P}(\mathbb{E})$-valued random variables.*

We shall first establish a simpler implication.

LEMMA 3.7. *If $\mathbb{E}$ is a Polish space and $\psi_N^+(1) < \infty$ for some $N \geq 1$, then $\mathscr{P}(\mathbb{F})$-valued random variables $\{\nu_n\}_{n\geq 1}$, defined by*

$$\nu_n = n^{-d} \sum_{\mathbf{z}\in \mathscr{C}_n} \delta_{\pi(\mathbf{X}_\mathbf{z})},$$

are exponentially tight.

PROOF. Let $\mathscr{Q}(\mathbb{F})$ denote the set of all nonnegative countably additive measures $\mathbf{m}$ on $(\mathbb{F}, \mathscr{B})$ such that $\mathbf{m}(\mathbb{F}) \leq 1$; $\mathscr{Q}(\mathbb{F})$ is equipped with the weak topology. Clearly, $\nu_n \in \mathscr{P}(\mathbb{F})$ with P-probability 1 for each $n \geq 1$; furthermore, $\mathscr{P}(\mathbb{F}) \subset \mathscr{Q}(\mathbb{F})$ is a closed subset. Therefore, it is enough to prove that $\{\nu_n\}$ is exponentially tight as a $\mathscr{Q}(\mathbb{F})$-valued sequence of random variables.

Since $\mathscr{Q}(\mathbb{F})$ is a Polish space and $\mathscr{Q}(\mathbb{F})$ is a positively balanced and bounded subset of the vector space $\mathbb{V}$ of all signed measures on $(\mathbb{F}, \mathscr{B})$ with the topology of weak convergence, therefore by de Acosta (1985), Theorem 3.1 [see also Bretagnolle (1979), Proposition 3.4], there exists a semi-norm $q\colon \mathscr{Q}(\mathbb{F}) \to \mathbb{R}$ such that $q^{-1}[0,1]$ is compact in $\mathscr{Q}(\mathbb{F})$, convex and

$$E\big\{\exp q\big(\delta_{\pi(\mathbf{X}_0)}\big)\big\} < \infty.$$

We shall show that

$$\sup_n \big(E\{\exp\big(N^{-d}n^d q(\nu_n)\big)\}\big)^{1/n^d} < \infty. \tag{3.9}$$

This implies exponential tightness [see, e.g., Stroock (1984), Corollary 3.27]. It is enough to consider $n > N$ only. Denote $\mathbf{Y}_\mathbf{z} = \pi(\mathbf{X}_\mathbf{z})$. Let $k = [n/N]$. We have

$$\begin{aligned} E\{\exp\big(N^{-d}n^d q(\nu_n)\big)\} &\leq E\left\{\exp\left(N^{-d} \sum_{\mathbf{z}\in\mathscr{C}_{Nk}} q\big(\delta_{\mathbf{Y}_\mathbf{z}}\big)\right)\right\} \\ &= E\left\{\exp\left(N^{-d} \sum_{\mathbf{z}\in\mathscr{C}_N} \sum_{\mathbf{j}\in\mathscr{C}_k} q\big(\delta_{\mathbf{Y}_{\mathbf{z}+N(\mathbf{j}-1)}}\big)\right)\right\}. \end{aligned}$$

Hölder's inequality, stationarity and (5.1) consecutively applied give

$$\begin{aligned} E\{\exp\big(N^{-d}n^d q(\nu_n)\big)\} &\leq E\left\{\exp\left(\sum_{\mathbf{j}\in\mathscr{C}_k} q\big(\delta_{\mathbf{Y}_{N\mathbf{j}}}\big)\right)\right\} \\ &\leq \big(\psi_N^+(1)\big)^{k^d}\Big(E\big\{\exp\big(q\big(\delta_{\mathbf{Y}_0}\big)\big)\big\}\Big)^{k^d}. \end{aligned}$$

Since $k/n \leq 1$, this establishes (3.9) and ends the proof. □

PROOF OF CLAIM III. Denote $2^{-|\mathbf{k}|} = 2^{-|\mathbf{k}(1)|-\cdots-|\mathbf{k}(d)|}$, $\mathbf{k} \in \mathbb{Z}^d$. By Lemma 3.7 and stationarity, for each $\mathbf{z} \in \mathbb{Z}^d$ and each $L > 0$ there is a compact set $K_\mathbf{z} \subset \mathscr{P}(\mathbb{F})$ such that

$$P(\underline{\pi} \circ \underline{\mathbf{S}}^\mathbf{z}(\mu_n) \notin K_\mathbf{z}) \le 2^{-|\mathbf{z}|-d} e^{-Ln},$$

where $\underline{\pi} \circ \underline{\mathbf{S}}^\mathbf{z}(\cdot)$: $\mathscr{P}(\mathbb{E}) \to \mathscr{P}(\mathbb{F})$ is defined by $\underline{\pi} \circ \underline{\mathbf{S}}^\mathbf{z}(\mathbf{p})(A) = \mathbf{p}(\mathbf{S}^{-\mathbf{z}}(\pi^{-1}(A)))$. Define $K \subset \mathscr{P}(\mathbb{E})$ by $K = \bigcap_{\mathbf{z} \in \mathbb{Z}^d}\{\mathbf{p} \in \mathscr{P}(\mathbb{E})\colon \underline{\pi} \circ \underline{\mathbf{S}}^\mathbf{z}(\mathbf{p}) \in K_\mathbf{z}\}$. Then K is closed (as an intersection of closed sets) and $P(\mu_n \notin K) \le \sum_\mathbf{z} P(\underline{\pi} \circ \underline{\mathbf{S}}^\mathbf{z}(\mu_n) \notin K_\mathbf{z}) \le \sum_\mathbf{z} 2^{-|\mathbf{z}|-d} e^{-Ln} = e^{-Ln}$. To check that K is compact, one needs only to verify that K is a tight set of measures. This follows from the fact that $\underline{\pi} \circ \underline{\mathbf{S}}^\mathbf{z}(K)$ are tight sets of measures and π strongly separates points of $\mathbb{E}$: Strong separation implies that if $C_\mathbf{z} \subset \mathbb{F}$, $\mathbf{z} \in \mathbb{Z}^d$, are compact sets, then $C = \bigcap_{\mathbf{z} \in \mathbb{Z}^d} \mathbf{S}^{-\mathbf{z}} \circ \pi^{-1}(C_\mathbf{z})$ is a compact subset of $\mathbb{E}$. Indeed, if $\{\mathbf{x}_n\} \in C$, then one can select a subsequence n' such that $\pi \circ \mathbf{S}^\mathbf{z}(\mathbf{x}_{n'})$ converges in $\mathbb{F}$ for each $\mathbf{z} \in \mathbb{Z}^d$. This subsequence $\{\mathbf{x}_{n'}\}$ has to converge in $\mathbb{E}$, too. For if not, then by strong separation $d_\mathbb{E}(\mathbf{x}_n, \mathbf{x}_m) > \varepsilon$ for some $\varepsilon > 0$ and all n, m in our subsequence. Hence $d_\mathbb{F}(\pi \circ \mathbf{S}^\mathbf{z}(\mathbf{x}_n), \pi \circ \mathbf{S}^\mathbf{z}(\mathbf{x}_m)) > \delta > 0$ for some $\mathbf{z}$, which is impossible. □

PROOF OF THEOREMS 2.1 AND 2.2. We apply Theorem 7.1 to the closed and convex subset $\mathbb{X} = \mathscr{P}(\mathbb{E})$ of the locally convex Hausdorff topological vector space $\mathbb{V}$ of all signed measures on $\mathbb{E}$, with the topology of weak convergence. We take $\mathscr{I} = \{1, 2^d, 3^d, \dots\}$ and $\Lambda = \mathbf{C}_b(\pi)$ is considered as a subset of $\mathbb{V}^*$, with the usual identification of the bounded continuous functions $F \in \mathbf{C}_b(\pi)$ as the bounded linear functionals, which act on $\mathbb{V} \ni \mu$ by $\int F\,d\mu$. Clearly, $\mathscr{G} =$ *Concave*. By assumption $\mathbf{C}_b(\pi)$ separates points of $\mathbb{E}$. By Claims I and III, the large deviation principle under the assumptions of either Theorem 2.1 or 2.2 follows from Theorem 7.1 below; clearly, Claim III is not needed in the proof of Theorem 2.1. Formula (2.2) is a variant of (7.3) and holds true by Claim II. □

4. Rate function identification. Throughout this section all the assumptions of Theorem 2.3 are supposed to hold even if not written explicitly. Since by Theorem 2.2 the large deviation principle holds, it remains only to show that formulas (2.4)–(2.6) hold. The idea of the proof is to check that particular functions (logarithms of the conditional densities) are among the functions at which the maximum of the expression on the right-hand side of (2.2) is attained. However, one needs to write (2.2) with noncontinuous functions allowed; this is accomplished in lemmas below.

The first lemma takes care of (2.4).

LEMMA 4.1. $\mathbf{I}(\mathbf{p}) = \infty$ *for* $\mathbf{p} \notin \mathscr{P}_S(\mathbb{E})$.

The proof is similar to Orey and Pelikan (1988), page 1487, and is omitted.

For bounded measurable F define

$$\mathbb{L}_n(F) = n^{-d} \log E\{\exp(n^d \mu_n(F))\}.$$

Clearly, $\mathbb{L}(F) = \lim_{n\to\infty}\mathbb{L}_n(F)$, provided that the asymptotic value $\mathbb{L}(F)$ exists; the proof of Claim I shows in particular that $\mathbb{L}(F)$ exists if $F \in \mathbf{M}_{b-}(\pi)$.

LEMMA 4.2. *If $F \in \mathbf{M}_{b-}(\pi)$, then for all $n \in \mathbb{N}$ and all $L \geq 1$ large enough,*

$$\begin{aligned} &|\mathbb{L}_n(F) - n^{-d}\log E^{\mathscr{M}(0)}\{\exp(n^d\mu_n(F))\}| \\ &\quad \leq 4\big(1-(1-2L/n)^d\big)\|F\|_\infty \\ &\qquad + n^{-d}\log(\psi_L^+(n)) - n^{-d}\log(\tau_L^-(n)), \end{aligned} \tag{4.1}$$

P-almost surely.

PROOF. By (3.5), we need only to establish the following upper bound:

$$\begin{aligned} \log\big(\operatorname{ess\,sup} E^{\mathscr{M}(0)}\{\exp(n^d\mu_n(F))\}\big) &\leq \log\big(E\{\exp(n^d\mu_n(F))\}\big) \\ &\quad + 2\big(n^d - (n-2L)^d\big)\|F\|_\infty \\ &\quad + \log(\psi_L^+(n)) \quad \text{a.s.} \end{aligned} \tag{4.2}$$

To prove (4.2), let $\mathbf{L} = (L, \dots, L)$, where $L \geq 0$ is such that $\sum_{\mathbf{z}\in\mathbf{L}+\mathscr{C}_{n-L}}F(\mathbf{X}_\mathbf{z})$ is $\mathscr{F}(1)$-measurable. Observe that as in the proof of Lemma 3.4 we have

$$\sum_{\mathbf{z}\in\mathscr{C}_n}F(\mathbf{X}_\mathbf{z}) \leq \big(n^d-(n-2L)^d\big)\|F\|_\infty + \sum_{\mathbf{z}\in 2\mathbf{L}+\mathscr{C}_{n-2L}}F(\mathbf{X}_\mathbf{z}) \quad \text{for each } n > L.$$

Therefore, by stationarity we get

$$E^{\mathscr{M}(0)}\{\exp(n^d\mu_n(F))\} \leq e^{C_L(n)}\operatorname{ess\,sup} E^{\mathscr{M}(-L)}\left\{\exp\left(\sum_{\mathbf{z}\in\mathbf{L}+\mathscr{C}_{n-2L}}F(\mathbf{X}_\mathbf{z})\right)\right\},$$

where $C_L(n) = (n^d-(n-2L)^d)\|F\|_\infty$. Since $\sum_{\mathbf{z}\in\mathbf{L}+\mathscr{C}_{n-2L}}F(\mathbf{X}_\mathbf{z})$ is $\mathscr{G}(n)$-measurable, by (5.1), we obtain

$$E^{\mathscr{M}(0)}\{\exp(n^d\mu_n(F))\} \leq e^{C_L(n)}\psi_L^+(n)E\left\{\exp\left(\sum_{\mathbf{z}\in\mathscr{C}_{n-2L}}F(\mathbf{X}_\mathbf{z})\right)\right\}. \tag{4.3}$$

Trivially, we have again

$$\sum_{\mathbf{z}\in\mathscr{C}_{n-2L}}F(\mathbf{X}_\mathbf{z}) \leq \big(n^d-(n-2L)^d\big)\|F\|_\infty + \sum_{\mathbf{z}\in\mathscr{C}_n}F(\mathbf{X}_\mathbf{z}).$$

This together with (4.3) implies

$$E^{\mathscr{M}(0)}\{\exp(n^d\mu_n(F))\} \leq e^{2C_L(n)}\psi_L^+(n)E\left\{\exp\left(\sum_{\mathbf{z}\in\mathscr{C}_n}F(\mathbf{X}_\mathbf{z})\right)\right\},$$

which proves (4.2). Inequalities (3.5) and (4.2) end the proof of (4.1). $\square$

LEMMA 4.3. *Put* $\overline{M}(n) = \operatorname{ess\,sup} E^{\mathscr{M}(0)}\{\exp(n^d\mu_n(F))\}$. *If* $F \in \mathbf{M}_{b-}(\pi)$, *then*

$$\lim_{n\to\infty} n^{-d}\log\overline{M}(n) = \inf_{n\geq 1} n^{-d}\log\overline{M}(n); \tag{4.4}$$

in particular, the limit exists.

The proof of Lemma 4.4 is essentially a re-run of a portion of the proof of Claim I (see the proof of Lemma 3.3). Observe that $\overline{M}(kn) \leq \overline{M}(n)^{k^d}$, that is, $\log\overline{M}(n)$ is a "subadditive" (and finite) function of n and use this to get (4.4) (compare Lemma 3.2).

The following lemma gives a variant of (2.2) with the supremum over the bounded continuous functions replaced by the supremum over the bounded measurable functions.

LEMMA 4.4. *If* $\mathbf{p} \in \mathscr{P}_s(\mathbb{E})$, *then*

$$\mathbf{I}(\mathbf{p}) = \sup\{\mathbf{p}(F): F \in \mathbf{M}_b(\pi), \mathbb{L}(F) < 0\}, \tag{4.5}$$

$$\mathbf{I}(\mathbf{p}) = \sup\{\mathbf{p}(F) - \mathbb{L}(F): F \in \mathbf{M}_{b-}(\pi)\}. \tag{4.6}$$

PROOF. If $\mathbf{p} \in \mathscr{P}_S(\mathbb{E})$, then

$$\mathbf{I}(\mathbf{p}) = \sup\{\mathbf{p}(F): F \in \mathbf{C}_{b-}(\pi), \mathbb{L}(F) < 0\}. \tag{4.7}$$

Indeed, it is easily seen that stationarity implies $\mathbb{L}(F) = \mathbb{L}(F\circ\mathbf{S}^{\mathbf{k}})$ for all $\mathbf{k} \in \mathbb{Z}^d$; also, if $\mathbf{p} \in \mathscr{P}_S(\mathbb{E})$, then $\mathbf{p}(F) = \mathbf{p}(F\circ\mathbf{S}^{\mathbf{k}})$ for all $\mathbf{k} \in \mathbb{Z}^d$. Hence from (2.2) we get $\mathbf{I}(\mathbf{p}) = \sup\{\mathbf{p}(F) - \mathbb{L}(F)\colon F \in \mathbf{C}_{b-}(\pi)\}$. To end the proof, notice that since $F - \mathbb{L}(F) \in \mathbf{C}_{b-}(\pi)$, we have

$$\begin{aligned}\mathbf{I}(\mathbf{p}) &= \sup\{\mathbf{p}(F): F \in \mathbf{C}_{b-}(\pi), \mathbb{L}(F) = 0\}\\ &\leq \sup\{\mathbf{p}(F): F \in \mathbf{C}_{b-}(\pi), \mathbb{L}(F) \leq 0\}.\end{aligned}$$

On the other hand, if $\mathbb{L}(F) \leq 0$, then obviously $\mathbf{p}(F) - \mathbb{L}(F) \geq \mathbf{p}(F)$, which proves that

$$\begin{aligned}&\sup\{\mathbf{p}(F): F \in \mathbf{C}_{b-}(\pi), \mathbb{L}(F) \leq 0\}\\ &\quad\leq \sup\{\mathbf{p}(F) - \mathbb{L}(F): F \in \mathbf{C}_{b-}(\pi), \mathbb{L}(F) \leq 0\}\\ &\quad\leq \sup\{\mathbf{p}(F) - \mathbb{L}(F): F \in \mathbf{C}_{b-}(\pi)\} \leq \mathbf{I}(\mathbf{p}).\end{aligned}$$

Finally, notice that since both $\mathbf{p}$ and $\mathbb{L}$ as mappings $\mathbf{C}_b \to \mathbb{R}$ are continuous, replacing in (4.7) $\mathbb{L}(F) < 0$ by $\mathbb{L}(F) \leq 0$ makes no difference.

The argument also shows that taking the supremum in (4.5) over $F \in \mathbf{M}_b(\pi)$ is the same as taking the supremum over $F \in \mathbf{M}_{b-}(\pi)$, provided $\mathbf{p} \in \mathscr{P}_S(\mathbb{E})$. Using (4.7) and (2.2), each of the inequalities $\leq$ in (4.5) and (4.6) is obvious. To verify $\geq$, take $F \in \mathbf{M}_{b-}(\pi)$ with $\limsup_{n\to\infty}(E\{\exp(n^d\mu_n(F))\})^{1/n^d} < 1$.

Then there is $\delta > 0$ and n_0 such that

$$(E\{\exp(n^d\mu_n(F))\})^{1/n^d} < e^{-3\delta} \quad \text{for each } n > n_0. \tag{4.8}$$

Increasing n_0 if necessary, we can also assume that

$$\begin{aligned} 4\big(1-(1-2L/n)^d\big)\|F\|_\infty + n^{-d}\log(\psi_L^+(n)) \\ - n^{-d}\log(\tau_L^-(n)) < \delta \quad \text{for some } L \text{ and each } n > n_0. \end{aligned}$$

Using (4.8) and the Luzin theorem, for each $\varepsilon > 0$ one can find $G_\varepsilon \in \mathbf{C}_{b-}(\pi)$ such that

$$\|G_\varepsilon\|_\infty = \|F\|_\infty, \qquad \mathbf{p}(F) \le \mathbf{p}(G_\varepsilon) + \varepsilon \quad \text{and} \quad \Big(E\Big\{\exp\Big(n_0^d\mu_{n_0}(G_\varepsilon)\Big)\Big\}\Big)^{1/n_0^d} \le e^{-2\delta}.$$

Hence from (4.2) we get

$$n_0^{-d}\log\Big(\operatorname{ess\,sup} E^{\mathscr{M}(0)}\Big\{\exp\Big(n_0^d\mu_{n_0}(G_\varepsilon)\Big)\Big\}\Big) \le -\delta.$$

By (4.4) the same inequality holds in the limit, that is,

$$\lim_{n\to\infty} n^{-d}\log\big(\operatorname{ess\,sup} E^{\mathscr{M}(0)}\{\exp(n^d\mu_n(G_\varepsilon))\}\big) \le -\delta.$$

By Lemma 4.2 the limit is $\mathbb{L}(G_\varepsilon)$ and hence we have $\mathbb{L}(G_\varepsilon) \le -\delta < 0$. Therefore, (4.7) gives $\mathbf{p}(F) \le \mathbf{p}(G_\varepsilon) + \varepsilon \le \mathbf{I}(\mathbf{p}) + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, this proves the " $\ge$ " inequality in (4.5). The right-hand side of (4.5) dominates the right-hand side of (4.6); thus (4.6) is established, too. □

PROOF OF THEOREM 2.3. Lemma 4.1 proves (2.4). Let $\mathbf{p} \in \mathscr{P}_S(\mathbb{E})$. If there is $n \ge 1$ such that $\mathbf{p}_{|\mathscr{I}(n)}$ is not absolutely continuous with respect to $P_{|\mathscr{I}(n)}$, then there is $A \in \mathscr{I}(n)$ such that $P(A) = 0$ and $\mathbf{p}(A) > 0$. Hence there is a bounded measurable function $F \in \mathbf{M}_b(\pi)$ such that $\mathbf{p}(F)$ is arbitrarily large, while $\mathbb{L}(F) < 0$; take, for example, $F(\mathbf{x}) = -1 + CI_A(\mathbf{x})$ with a large enough constant C. By (4.5) this ends the proof of (2.5).

Suppose now $\mathbf{p}_{|\mathscr{I}(n)}$ is absolutely continuous with respect to $P_{|\mathscr{I}(n)}$ for all $n \ge 1$. Then (2.6) follows from the following two observations.

CLAIM A. *For each $F \in \mathbf{M}_{b-}(\pi)$ we have*

$$\mathbf{p}(F) - \mathbb{L}(F) \le \liminf_{n\to\infty}\mathbf{p}\big(n^{-d}\log f_n(\mathbf{x})\big).$$

CLAIM B.

$$\limsup_{n\to\infty}\mathbf{p}\big(n^{-d}\log f_n(\mathbf{x})\big) \le \mathbf{I}(\mathbf{p}).$$

Indeed, Claim A and (4.6) give $\mathbf{I}(\mathbf{p}) \le \liminf_{n\to\infty} \mathbf{p}(n^{-d}\log f_n(\mathbf{x}))$, which together with Claim B, guarantees that the limit exists and equality (2.6) holds. □

It remains only to prove the claims.

Proof of Claim A. Fix $F \in \mathbf{M}_{b^-}(\pi)$. Let $L \ge 1$ be such that F is $\mathscr{F}(-L)$-measurable. Then

$$
\begin{aligned}
&E\left\{\exp\left(\sum_{\mathscr{C}_n} F(\mathbf{X}_\mathbf{z})\right)\right\}\\
&\quad = \int_{f_{L+n}\circ \mathbf{S}^{-\mathbf{L}}\neq 0} \exp\left(-\log f_{L+n}\circ \mathbf{S}^{-\mathbf{L}}(\mathbf{x}) + \sum_{\mathscr{C}_n} F(\mathbf{x}_\mathbf{z})\right) f_{L+n}\circ \mathbf{S}^{-\mathbf{L}}(\mathbf{x})\, dP(\mathbf{x})\\
&\qquad + \int_{f_{L+n}\circ \mathbf{S}^{-\mathbf{L}}= 0} \exp\left(\sum_{\mathscr{C}_n} F(\mathbf{x}_\mathbf{z})\right) dP(\mathbf{x})\\
&\quad \ge \int_{f_{L+n}\circ \mathbf{S}^{-\mathbf{L}}\neq 0} \exp\left(-\log f_{L+n}\circ \mathbf{S}^{-\mathbf{L}}(\mathbf{x}) + \sum_{\mathscr{C}_n} F(\mathbf{x}_\mathbf{z})\right) d\mathbf{p}(\mathbf{x}), \qquad n \ge 1.
\end{aligned}
$$

Since $\{\mathbf{x}\colon f_{L+nN}\circ \mathbf{S}^{-\mathbf{L}}(\mathbf{x}) = 0\}$ is of $\mathbf{p}$-measure 0, the last integral can be taken over $\mathbb{E}$. By the convexity of $\exp(\cdot)$ and the stationarity of $\mathbf{p}$, we get therefore

$$
E\left\{\exp\left(\sum_{\mathscr{C}_n} F(\mathbf{X}_\mathbf{z})\right)\right\} \ge \exp\left(-\mathbf{p}(\log f_{L+n}) + n^d\mathbf{p}(F)\right).
$$

This implies

$$
\mathbf{p}(F) - \left(\frac{n}{n+L}\right)^d n^{-d}\log E\left\{\exp\left(\sum_{\mathscr{C}_n} F(\mathbf{X}_\mathbf{z})\right)\right\} \le (n+L)^{-d}\mathbf{p}(\log f_{L+n}).
$$

Passing to the limit as $n\to\infty$ over a suitable subsequence ends the proof. □

Proof of Claim B. Fix $N \ge 1$. Let $G(\mathbf{x}) = \log f_N(\mathbf{x})$. For $C \ge 0$ put $G_C(\mathbf{x}) = (0 \vee G(\mathbf{x})) \wedge C$. Clearly, $\mathbf{p}(G^-) < \infty$ and hence $-\infty < \mathbf{p}(G) \le \infty$ exists. (Indeed, $f\log f \to 0$ as $f \searrow 0$.) Hence

$$
\mathbf{p}(G) \le \mathbf{p}(G^+) = \lim_{C\to\infty}\mathbf{p}(G_C).
$$

In particular, for each $\varepsilon > 0$ there is $C = C(N,\varepsilon)$ such that for all $L \ge 1$,

$$
(N+L)^{-d}\mathbf{p}(G_C) \ge (N+L)^{-d}\mathbf{p}(G) - \varepsilon. \tag{4.9}
$$

We also have

$$
\mathbb{L}\left((N+L)^{-d}G_C\right) \le (N+L)^{-d}\log\psi_L^+(N) + (N+L)^{-d}\log 2. \tag{4.10}
$$

Indeed, given $N, L \geq 1$, by Hölder's inequality and stationarity we have

$$
\begin{aligned}
&E\left\{\exp\left(\sum_{\mathbf{z}\in\mathscr{C}_{n(N+L)}}(N+L)^{-d}G_C(\mathbf{X}_{\mathbf{z}})\right)\right\}\\
&\qquad= E\left\{\exp\left(\sum_{\mathbf{y}\in\mathscr{C}_{n-1}}\sum_{\mathbf{z}\in\mathscr{C}_{N+L}}(N+L)^{-d}G_C(\mathbf{X}_{\mathbf{y}(N+L)+\mathbf{z}})\right)\right\}\\
&\qquad\leq E\left\{\exp\left(\sum_{\mathbf{y}\in\mathscr{C}_{n-1}}G_C(\mathbf{X}_{\mathbf{y}(N+L)})\right)\right\}.
\end{aligned}
$$

Therefore, (5.1) applied recursively gives

$$
\begin{aligned}
&E\left\{\exp\left(\sum_{\mathbf{z}\in\mathscr{C}_{n(N+L)}}(N+L)^{-d}G_C(\mathbf{X}_{\mathbf{z}})\right)\right\}\\
&\qquad\leq \left(\psi_L^+(N)\right)^{n^d}\left(E\{\exp(G_C(\mathbf{X}_{\mathbf{0}}))\}\right)^{n^d}.
\end{aligned}
\tag{4.11}
$$

To end the proof of (4.10), it remains only to observe that

$$
\begin{aligned}
E\{\exp(G_C(\mathbf{X}_{\mathbf{0}}))\} &\leq E\{\exp(\log^+ f_N(\mathbf{X}_{\mathbf{0}}))\} = \int_{f_N\geq 1} f(\mathbf{x})\,dP + P(f_N\leq 1)\\
&= \mathbf{p}(f_N\geq 1) + P(f_N\leq 1) \leq 2.
\end{aligned}
$$

Therefore, (4.10) follows from (4.11) by taking the logarithm, dividing by n^d and passing to the limit as $n\to\infty$.

From (4.9) and (4.10) we have

$$
\begin{aligned}
(N+L)^{-d}\mathbf{p}(\log f_N) &\leq \mathbf{p}\left((N+L)^{-d}G_C\right) - \mathbb{L}\left((N+L)^{-d}G_C\right)\\
&\qquad + \varepsilon + (N+L)^{-d}\log\psi_L^+(N) + (N+L)^{-d}\log 2\\
&\leq \mathbf{I}(\mathbf{p}) + \varepsilon + (N+L)^{-d}\log\psi_L^+(N) + (N+L)^{-d}\log 2.
\end{aligned}
$$

Hence, passing to the limit as $N\to\infty$ over a suitable subsequence, we get

$$
\limsup_{n\to\infty} n^{-d}\mathbf{p}(\log f_n) \leq \mathbf{I}(\mathbf{p}) + \varepsilon + \limsup_{n\to\infty} n^{-d}\log\psi_L^+(n).
$$

Since $\varepsilon > 0$ and $L \geq 1$ are arbitrary and (2.3) holds, this ends the proof of Claim B. $\square$

APPENDIX A

5. Inequalities for the strong mixing coefficients. In this appendix we state some "weak independence" inequalities needed in the paper. Proposition 5.1 below relates the more frequently used "ratio mixing" measure of dependence (5.6) to the coefficient ψ. The inequalities obtained separate the ratio-mixing condition into the part directly responsible for large deviations

($\psi_- > 0$) and the part used for handling noncompact state spaces ($\psi_+ < \infty$) (compare Theorems 2.1 and 2.2); ψ_+ is also useful for rate function identification (cf. Theorem 2.3).

The following lemma states basic inequalities; proofs are omitted, since more subtle cases are well known.

LEMMA 5.1. *If $\xi \geq 0$ is $\mathscr{G}(n)$-measurable and $\eta \geq 0$ is $\mathscr{M}(-L)$-measurable* [*see* (1.5) *and* (1.6)], *then*

$$\psi_L^-(n)E\{\xi\} \leq E^{\mathscr{M}(-L)}\{\xi\} \leq \psi_L^+(n)E\{\xi\} \quad a.s. \tag{5.1}$$

for each $L, n \geq 1$;

$$\tau_L^-(n)E\{\xi\} \leq E^{\mathscr{M}(-L)}\left\{\sum_{\mathbf{k}\in\mathscr{C}_L} \xi\circ\mathbf{S}^{-\mathbf{k}}\right\}\Big/L^d \quad a.s. \tag{5.2}$$

for each $L, n \geq 1$.

If $\{Y_n\}_{n\geq 0}$ is an $\mathbb{F}$-valued stationary Markov chain, we extend it first to $\{Y_n\}_{n\in\mathbb{Z}}$ (take the weak limit of $\{Y_{-n}, Y_{-n+1}, \dots, Y_0, Y_1, \dots\}$ as $n \to \infty$). Define $\mathbb{E} = \mathbb{F}^{\mathbb{Z}}$ with π being a projection on the 0th coordinate and P being the (stationary by assumption) distribution of $\{Y_n\}_{n\in\mathbb{Z}}$ on $\mathbb{E}$.

The following result is an exercise in the use of the Markov property and the formula

$$\inf\left\{\sum P(A_k \cap B)\big/P(B) \colon B \in \mathscr{G},\ P(B) > 0\right\} = \operatorname{ess\,inf} \sum P(A_k|\mathscr{G})$$

[see Blum, Hanson and Koopmans (1963), Lemma 8].

LEMMA 5.2. *Suppose $P_n(\mathbf{x}, d\mathbf{y})$ is the family of n-step transition probability measures of a stationary Markov chain $\{Y_n\}_{n\in\mathbb{Z}}$ with a state space $\mathbb{F}$ and the invariant distribution $p \in \mathscr{P}(\mathbb{F})$. If* (2.7) *holds for some $N \geq 1$, $C < \infty$, then*

$$\tau_N^-(n) \geq (NC)^{-1}, \qquad n = 1, 2, \dots. \tag{5.3}$$

If there are $C < \infty$ and $N, M \geq 1$ such that (2.9) *holds for all $A \in \mathscr{B}$, $\mathbf{x}, \mathbf{y} \in \mathbb{F}$, then*

$$\tau_N^-(n) \geq (CN)^{-1}, \qquad n = 1, 2, \dots, \tag{5.4}$$

$$\psi_M^+(n) \leq CN, \qquad n = 1, 2, \dots. \tag{5.5}$$

For a pair of σ-fields $\mathscr{F}, \mathscr{G}$ the ratio-mixing coefficient λ is defined by

$$\lambda(\mathscr{F}, \mathscr{G}) = \sup_{B\in\mathscr{G}}\left\{\frac{\operatorname{ess\,sup} P(B|\mathscr{F})}{\operatorname{ess\,inf} P(B|\mathscr{F})}\right\}. \tag{5.6}$$

With specific choices of σ-fields $\mathscr{F}, \mathscr{G}$, the assumption $\lambda(\mathscr{F}, \mathscr{G}) < \infty$ was used in several large deviation results [see, e.g., Stroock (1984), assumption (6.1), Ellis (1988), Hypothesis 1.1(b), and Olla (1987), page 398, assumption (I); see also Pelikan and Orey (1988)]. The following result shows that for those purposes λ is equivalent to ψ. Proposition 5.1 is also used in the proof of Theorem 6.1 below.

PROPOSITION 5.1. *For all σ-fields $\mathscr{F}, \mathscr{G} \subset \mathscr{B}$,*

$$\lambda(\mathscr{F}, \mathscr{G}) \le \psi(\mathscr{F}, \mathscr{G}) \le (\lambda(\mathscr{F}, \mathscr{G}))^2.$$

PROOF. Throughout this proof $\mathscr{F}, \mathscr{G} \subset \mathscr{B}$ are fixed, $A \in \mathscr{F}$, $B \in \mathscr{G}$ and $P(A)P(B) > 0$. From (5.1) it follows immediately that $\operatorname{ess\,sup} P(B|\mathscr{F}) \le \psi_+ P(B)$ and $\operatorname{ess\,inf} P(B|\mathscr{F}) \ge \psi_- P(B)$, which proves the left-hand side inequality. The right-hand side inequality follows from the fact that

$$\psi_+ \le \lambda, \tag{5.7}$$

$$1/\psi_- \le \lambda. \tag{5.8}$$

Indeed, to prove (5.7), observe that

$$\frac{P(A \cap B)}{P(A)P(B)} = \frac{1}{P(B)} \frac{\int_A P(B|\mathscr{F})\,dP}{P(A)} \le \frac{\operatorname{ess\,sup} P(B|\mathscr{F})}{\operatorname{ess\,inf} P(B|\mathscr{F})}.$$

Taking the supremum over all $A \in \mathscr{F}$, $B \in \mathscr{G}$ proves (5.7). Similarly,

$$\frac{P(A)P(B)}{P(A \cap B)} = P(B) \frac{P(A)}{\int_A P(B|\mathscr{F})\,dP} \le \frac{\operatorname{ess\,sup} P(B|\mathscr{F})}{\operatorname{ess\,inf} P(B|\mathscr{F})}.$$

Taking the supremum over all $A \in \mathscr{F}$, $B \in \mathscr{G}$ proves (5.8). Inequalities (5.7) and (5.8) end the proof. □

REMARK 5.1. If $d = 1$, then, after taking into account Proposition 5.1 and the trivial inequality $\psi_L(n) \ge 1$, the (RM) condition of Orey and Pelikan (1988) reads

$$\liminf_{n \to \infty} \left(\psi_{m(n)}(n - m(n))\right)^{1/n} = 1, \tag{RM}$$

where $m(n)/n \to 0$ and $m(n) \le n$. If $m(n) = m$ does not depend on m, then (RM) implies $\lim_{L \to \infty} \liminf_{n \to \infty} (\psi_L(n))^{1/n} = 1$; on the other hand, (RM) with any $m(n) \to \infty$ is implied by $\lim_{L \to \infty} \liminf_{n \to \infty} (\psi_L(n))^{1/n} = 1$. Note, however, that (2.1) uses τ_L^-, which is a weaker measure of dependence than ψ_L.

APPENDIX B

6. Uniform strong mixing for Gibbs fields. In this appendix the weak dependence properties of Gibbs fields on $\mathbb{Z}^d$ are analyzed. Let $\mathbb{E} = \mathbb{N}^{\mathbb{Z}^d}$, $\mathbb{F} = \mathbb{N}$ and let $\pi = \pi_0$ be the projection on the $\mathbf{0}$th coordinate; shifts S_i are along coordinates in $\mathbb{Z}^d$, $1 \le i \le d$. For a finite set $F \subset \mathbb{Z}^d$ denote by π_F the projection $\pi_F\colon \mathbb{N}^{\mathbb{Z}^d} \to \mathbb{N}^F$. (In particular, $\pi_{\mathbf{0}} = \pi_{\{\mathbf{0}\}}$.) Let $\mathbf{x}_{|F}$ denote $(x_{\mathbf{z}}\colon \mathbf{z} \in F)$ and for $\mathbf{x}, \mathbf{y} \in \mathbb{N}^{\mathbb{Z}^d}$ define $(\mathbf{x}_{|F} \searrow \mathbf{y}) \in \mathbb{N}^{\mathbb{Z}^d}$, by $(\mathbf{x}_{|F} \searrow \mathbf{y}) = (a_{\mathbf{z}})$, where $a_{\mathbf{z}} = x_{\mathbf{z}}$ if $\mathbf{z} \in F$ and $a_{\mathbf{z}} = y_{\mathbf{z}}$ otherwise. Recall that P is a Gibbs measure (i.e., an $\mathbb{N}$-valued random field $\mathbf{X} = \{Y_{\mathbf{z}}\}_{\mathbf{z} \in \mathbb{Z}^d}$ is a Gibbs field) with the interaction potentials $\{\mathbf{V}_F: \mathbb{N}^F \to \mathbb{R}\}$, where $F \subset \mathbb{Z}^d$ are finite sets, if for each finite set $G \subset \mathbb{Z}^d$ and each set $A \subset \mathbb{N}^G$,

$$P(\{Y_{\mathbf{z}}\}_{\mathbf{z}\in G} \in A|\sigma\{Y_{\mathbf{z}}\}_{\mathbf{z}\notin G}) \tag{6.1}$$
$$= Z^{-1} \sum_{\mathbf{y}\in A} \exp\Big(\sum_{F:\, F\cap G\neq\varnothing} \mathbf{V}_F \circ \pi_F(\mathbf{y}_{|G} \searrow \mathbf{X})\Big),$$

where

$$Z = \sum_{\mathbf{y}\in\mathbb{N}^G} \exp \sum_{F:\, F\cap G\neq\varnothing} \mathbf{V}_F \circ \pi_F(\mathbf{y}_{|G} \searrow \mathbf{X})$$

is the (random) normalizing constant. We shall assume that $\mathbf{V}_F$ is stationary, that is, $\mathbf{V}_{F+\mathbf{k}} \circ \pi_{F+\mathbf{k}} = \mathbf{V}_F \circ \pi_F \circ \mathbf{S}^{-\mathbf{k}}$. In the case of a finite state space $\mathbb{F}$ rather than $\mathbb{F} = \mathbb{N}$, Preston (1976), Proposition 5.4, gives suitable sufficient conditions for the existence of a stationary Gibbs measure P determined by the stationary potentials $\{\mathbf{V}_F\}$. Put $\|\mathbf{V}_F\| = \sup_{\mathbf{y}} |\mathbf{V}_F(\mathbf{y})|$.

THEOREM 6.1. *If P is a stationary Gibbs measure with the stationary interaction potentials $\{\mathbf{V}_F\}$ such that*

$$\sum_{F\ni\mathbf{0}} \|\mathbf{V}_F\| < \infty, \tag{6.2}$$

then

$$\lim_{L\to\infty} \limsup_{n\to\infty} (\psi_L(n))^{1/n^d} = 1.$$

[*$\psi_L(n)$ is defined by* (1.7); (6.2) *is understood as* (6.7) *below.*]

NOTATION FOR THE PROOF. For $n, N \geq 1$ put $\mathscr{R}(n, N) = \mathscr{C}_{n+2N} - \mathbf{N}$, where $\mathbf{N} = (N, \dots, N)$, $\mathscr{D}(n, N) = \mathscr{R}(n, N) \setminus \mathscr{C}_n$. Let $\mathscr{S}_{N,n} = \sigma\{Y_{\mathbf{z}}: \mathbf{z} \notin \mathscr{R}(n, N)\}$, that is, $\mathscr{S}_{N,n}$ is the σ-field generated by the random variables in sites of distance $\geq N$ from $\mathscr{C}_n$. By (6.1) we have

$$P(\{Y_{\mathbf{z}}\}_{\mathbf{z}\in\mathscr{C}(n)} \in A|\mathscr{S}_{N,n}) = S(A,\mathbf{X})/Z(\mathbf{X}), \tag{6.3}$$

where

$$S(A,\mathbf{X}) = \sum_{\mathbf{y}\in A\times\mathbb{N}^{\mathscr{D}(n,N)}} \exp\Big(\sum_{F:\, F\cap\mathscr{C}(n)\neq\varnothing} \mathbf{V}_F \circ \pi_F(\mathbf{y}_{|\mathscr{R}(n,N)} \searrow \mathbf{X})\Big),$$

$$Z(\mathbf{X}) = S(\mathbb{N}^{\mathscr{C}(n)},\mathbf{X})$$
$$= \sum_{\mathbf{y}\in\mathbb{N}^{\mathscr{R}(n,N)}} \exp\Big(\sum_{F:\, F\cap\mathscr{C}(n)\neq\varnothing} \mathbf{V}_F \circ \pi_F(\mathbf{y}_{|\mathscr{R}(n,N)} \searrow \mathbf{X})\Big).$$

Define $\gamma_N(n) = \psi(\mathscr{S}_{N,n}, \mathscr{S}(n))$ [see (1.4)]. We shall show that

$$\lim_{N\to\infty} \limsup_{n\to\infty} (\gamma_N(n))^{1/n^d} \leq 1. \tag{6.4}$$

Since $\mathscr{M}(-N) \subset \mathscr{S}_{N,n}$, we have $1 \le \psi_N(n) \le \gamma_N(n)$. Therefore, to end the proof of Theorem 6.1, it is enough to establish (6.4).

The following lemma reduces (6.4) to a property of potentials $\mathbf{V}_F$.

LEMMA 6.1.

$$\left(\gamma_N(n)\right)^{1/4} \le \sup_{\mathbf{x},\mathbf{y} \in \mathbb{N}^{\mathbb{Z}^d}} \frac{\exp\left(\sum_{F:\, F\cap\mathscr{C}(n)\neq\varnothing,\, F\setminus\mathscr{B}(n,N)\neq\varnothing} \mathbf{V}_F \circ \pi_F(\mathbf{x})\right)}{\exp\left(\sum_{F:\, F\cap\mathscr{C}(n)\neq\varnothing,\, F\setminus\mathscr{B}(n,N)\neq\varnothing} \mathbf{V}_F \circ \pi_F(\mathbf{y})\right)}.$$

PROOF. Fix $n, N \ge 1$. It is easily seen that

$$(6.5)\qquad \left(\gamma_N(n)\right)^{1/2} \le \sup_{\mathbf{x}} \frac{\operatorname{ess\,sup} P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} = \mathbf{x} \mid \mathscr{S}_{N,n}\right)}{\operatorname{ess\,inf} P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} = \mathbf{x} \mid \mathscr{S}_{N,n}\right)}.$$

Indeed, if

$$M = \sup_{\mathbf{x}} \frac{\operatorname{ess\,sup} P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} = \mathbf{x} \mid \mathscr{S}_{N,n}\right)}{\operatorname{ess\,inf} P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} = \mathbf{x} \mid \mathscr{S}_{N,n}\right)},$$

then for each $A \in \mathbb{N}^{\mathscr{C}(n)}$ we have

$$\begin{aligned}
&P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} \in A \mid \mathscr{S}_{N,n}\right)\\
&\quad = \sum_{\mathbf{x}\in A} P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} = \mathbf{x} \mid \mathscr{S}_{N,n}\right) \le M \sum_{\mathbf{x}\in A} \operatorname{ess\,inf} P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} = \mathbf{x} \mid \mathscr{S}_{N,n}\right)\\
&\quad \le M \operatorname{ess\,inf} P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} \in A \mid \mathscr{S}_{N,n}\right).
\end{aligned}$$

Hence by Proposition 5.1,

$$\left(\gamma_N(n)\right)^{1/2} \le \sup_{A} \frac{\operatorname{ess\,sup} P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} \in A \mid \mathscr{S}_{N,n}\right)}{\operatorname{ess\,inf} P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} \in A \mid \mathscr{S}_{N,n}\right)} \le M,$$

which proves (6.5).

Let

$$K = \sup_{\mathbf{x},\mathbf{y} \in \mathbb{N}^{\mathbb{Z}^d}} \frac{\exp\left(\sum_{F:\, F\cap\mathscr{C}(n)\neq\varnothing,\, F\setminus\mathscr{B}(n,N)\neq\varnothing} \mathbf{V}_F \circ \pi_F(\mathbf{x})\right)}{\exp\left(\sum_{F:\, F\cap\mathscr{C}(n)\neq\varnothing,\, F\setminus\mathscr{B}(n,N)\neq\varnothing} \mathbf{V}_F \circ \pi_F(\mathbf{y})\right)}.$$

Formula (6.3) gives

$$(6.6)\qquad \begin{aligned}
&\sup_{\mathbf{x}} \frac{\operatorname{ess\,sup} P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} = \mathbf{x} \mid \mathscr{S}_{N,n}\right)}{\operatorname{ess\,inf} P\left(\{Y_\mathbf{z}\}_{\mathbf{z}\in\mathscr{C}(n)} = \mathbf{x} \mid \mathscr{S}_{N,n}\right)}\\
&\qquad = \sup_{\mathbf{x},\mathbf{y},\mathbf{z}} \left\{ \frac{S(\{\mathbf{x}\},\mathbf{y})}{Z(\mathbf{y})} \frac{Z(\mathbf{z})}{S(\{\mathbf{x}\},\mathbf{z})} \right\},
\end{aligned}$$

where

$$S(\{\mathbf{x}\},\mathbf{y}) = \sum_{\mathbf{u}\in\mathbb{N}^{\mathscr{D}(n,N)}} \mathsf{e}(\mathbf{x},\mathbf{u},\mathbf{y}),$$

$$Z(\mathbf{y}) = \sum_{\mathbf{v}\in\mathbb{N}^{\mathscr{B}(n,N)}} \mathsf{e}(\mathbf{v}_{|\mathscr{C}(n)},\mathbf{v}_{|\mathscr{D}(n,N)},\mathbf{y}),$$

$$\mathsf{e}(\mathbf{x},\mathbf{u},\mathbf{y}) = \exp\biggl(\sum_{F:\,F\cap\mathscr{C}(n)\neq\varnothing} \mathbf{V}_F\circ\pi_F\bigl(\mathbf{x}_{|\mathscr{C}(n)}\searrow(\mathbf{u}_{|\mathscr{D}(n,N)}\searrow\mathbf{y})\bigr)\biggr);$$

in particular,

$$\mathsf{e}(\mathbf{v}_{|\mathscr{C}(n)},\mathbf{v}_{|\mathscr{D}(n,N)},\mathbf{y}) = \exp\biggl(\sum_{F:\,F\cap\mathscr{C}(n)\neq\varnothing} \mathbf{V}_F\circ\pi_F\bigl(\mathbf{v}_{|\mathscr{B}(n,N)}\searrow\mathbf{y}\bigr)\biggr).$$

Since

$$\exp\biggl(\sum_{F:\,F\cap\mathscr{C}(n)\neq\varnothing,\ F\subset\mathscr{B}(n,N)} \mathbf{V}_F\circ\pi_F\bigl(\mathbf{x}_{|\mathscr{C}(n)}\searrow(\mathbf{u}_{|\mathscr{D}(n,N)}\searrow\mathbf{y})\bigr)\biggr)$$

does not depend on $\mathbf{y}$ and

$$\begin{aligned}\mathsf{e}(\mathbf{x},\mathbf{u},\mathbf{y}) &= \exp\biggl(\sum_{F:\,F\cap\mathscr{C}(n)\neq\varnothing,\ F\setminus\mathscr{B}(n,N)\neq\varnothing} \mathbf{V}_F\circ\pi_F\bigl(\mathbf{x}_{|\mathscr{C}(n)}\searrow(\mathbf{u}_{|\mathscr{D}(n,N)}\searrow\mathbf{y})\bigr)\biggr)\\ &\quad\times\exp\biggl(\sum_{F:\,F\cap\mathscr{C}(n)\neq\varnothing,\ F\subset\mathscr{B}(n,N)} \mathbf{V}_F\circ\pi_F\bigl(\mathbf{x}_{|\mathscr{C}(n)}\searrow(\mathbf{u}_{|\mathscr{D}(n,N)}\searrow\mathbf{y})\bigr)\biggr),\end{aligned}$$

therefore, by the definition of K, for every $\mathbf{x},\mathbf{u},\mathbf{y},\mathbf{z}$ we have

$$\mathsf{e}(\mathbf{x},\mathbf{u},\mathbf{y}) \le K\mathsf{e}(\mathbf{x},\mathbf{u},\mathbf{z}).$$

Hence $S(\{\mathbf{x}\},\mathbf{y}) \le KS(\{\mathbf{x}\},\mathbf{z})$ and $Z(\mathbf{z}) \le KZ(\mathbf{y})$ for all $\mathbf{x},\mathbf{y},\mathbf{z}$; thus by (6.6) we get

$$\sup_{\mathbf{x}} \frac{\operatorname{ess\,sup} P\bigl(\{Y_{\mathbf{z}}\}_{\mathbf{z}\in\mathscr{C}(n)} = \mathbf{x}|\mathscr{S}_{N,n}\bigr)}{\operatorname{ess\,inf} P\bigl(\{Y_{\mathbf{z}}\}_{\mathbf{z}\in\mathscr{C}(n)} = \mathbf{x}|\mathscr{S}_{N,n}\bigr)} \le K^2.$$

This, together with (6.5), ends the proof of the lemma. □

PROOF OF THEOREM 6.1. By Lemma 6.1 and the stationarity of $\{\mathbf{V}_F\}$, we have

$$\begin{aligned}\bigl(\gamma_N(n)\bigr)^{1/n^d} &\le \exp\biggl(4n^{-d}\sum_{\mathbf{z}\in\mathscr{C}(n)}\ \sum_{F:\,F\ni\mathbf{z},\ F\setminus\mathscr{B}(n,N)\neq\varnothing} \|\mathbf{V}_F\|\biggr)\\ &\le \exp\biggl(4\sum_{F:\,F\ni\mathbf{0},\ F\setminus\mathscr{B}(0,N)\neq\varnothing} \|\mathbf{V}_F\|\biggr).\end{aligned}$$

Assumption (6.2) means that

$$\lim_{N \to \infty} \sum_{F:\, F \ni \mathbf{0},\, F \setminus \mathscr{R}(0,N) \neq \varnothing} \|\mathbf{V}_F\| = 0. \tag{6.7}$$

Therefore,

$$\limsup_{N \to \infty} \limsup_{n \to \infty} \left(\gamma_N(n)\right)^{1/n^d} \leq 1.$$

This proves (6.4). □

APPENDIX C

7. Large deviation principle criterion. In this appendix the criterion for the large deviation principle is stated and related definitions are given. Let $\{P_\nu\}_{\nu \in \mathscr{I}}$ be a family of probability measures, that is, $P_\nu \in \mathscr{P}(\mathbb{X})$, $\nu \in \mathscr{I}$. Here $\mathbb{X}$ is a metric space and $\mathscr{I}$ is a fixed unbounded (not necessarily countable) subset of real numbers $\nu \geq 1$. To simplify the notation, below we shall write $\{\nu \geq 1\}$ instead of $\mathscr{I}$.

DEFINITION. We shall say that a family $\{P_\nu\}$ of probability measures satisfies the large deviation principle with a rate function $\mathbf{I}$: $\mathbb{X} \to [0, \infty]$, if the following conditions are satisfied:

$$-\inf\{\mathbf{I}(\mathbf{x}) : \mathbf{x} \in A\} \leq \liminf_{\nu \to \infty} 1/\nu \log P_\nu(A)$$

for each open set $A \subset \mathbb{X}$;

$$\limsup_{\nu \to \infty} 1/\nu \log P_\nu(A) \leq -\inf\{\mathbf{I}(\mathbf{x}) : \mathbf{x} \in A\}$$

for each closed set $A \subset \mathbb{X}$.

Following Varadhan (1984), we shall also require $\mathbf{I}(\cdot)$ to be lower semicontinuous and to have compact level sets $\mathbf{I}^{-1}([0, a])$, $a \geq 0$.

DEFINITION. We shall say that a family $\{P_\nu\}$ of probability measures is *exponentially tight* if for each $M > 0$ there exists a compact set $K \subset \mathbb{X}$ such that $\sup_{\nu \in \mathscr{I}} 1/\nu \log P_\nu(K^c) \leq -M$.

DEFINITION. We say that family $\{P_\nu\}_{\nu \geq 1}$ of probability measures on a metric space $\mathbb{X}$ *admits an asymptotic value over a class* $\mathscr{F}$ of measurable functions if

$$\mathbb{L}(F) = \lim_{\nu \to \infty} 1/\nu \log \int_{\mathbb{X}} \exp(\nu F(\mathbf{x}))\, dP_\nu(\mathbf{x}) \tag{7.1}$$

exists and is a finite number for each function $F \in \mathscr{F}$.

The following criterion for the large deviation principle was used in Section 3.

THEOREM 7.1. *Let $\mathbb{V}$ be a locally convex Hausdorff topological linear space with the conjugate space $\mathbb{V}^*$. Suppose $\mathbb{X} \subset \mathbb{V}$ is a metric space in the relative topology. Suppose $\Lambda \subset \mathbb{V}^*$ separates points of $\mathbb{X}$ and define $\mathscr{G} = \{\mathbf{g}: \mathbf{g}(\mathbf{x}) = \min_i\{\lambda_i(\mathbf{x}) + c_i\}, c_i \in \mathbb{R}, \lambda_i \in \Lambda, 1 \le i \le n, n \in \mathbb{N}\}$. Suppose furthermore that $\{P_\nu\}$ is an exponentially tight family of probability measures and admits an asymptotic value over $\mathscr{G}$.*

Then $\{P_\nu\}$ satisfies the large deviation principle with the rate function $\mathbf{I}(\mathbf{x})$ defined by

$$\mathbf{I}(\mathbf{x}) = \sup\{\mathbf{g}(\mathbf{x}) - \mathbb{L}(\mathbf{g}): \mathbf{g} \in \mathscr{G}\}. \tag{7.2}$$

Moreover, if $\mathbb{X} \subset \mathbb{V}$ is closed and convex, $\mathbf{I}(\cdot)$ defined by (7.2) is convex and $\sup_\nu[\int_{\mathbb{X}} \exp(\nu\lambda(\mathbf{x}))\, dP_\nu(\mathbf{x})]^{1/\nu} < \infty$ for each $\lambda \in \mathbb{V}^*$, then

$$\mathbf{I}(\mathbf{x}) = \sup\{\lambda(\mathbf{x}) - \mathbb{L}(\lambda): \lambda \in \mathbb{V}^*\}. \tag{7.3}$$

Formula (7.3) under the convexity assumption is well known [see, e.g., Deuschel and Stroock (1989), Theorem 2.2.21]. The large deviation principle follows from the lattice form of the Stone–Weierstrass theorem; the latter implies that $\{P_\nu\}$ admits an asymptotic value over $\mathbf{C}_b(\mathbb{X})$, which in turn is equivalent to the large deviation principle. For a detailed exposition, see Bryc (1990), C.2.1.

Acknowledgments. The author would like to thank R. Bradley, A. de Acosta, N. Jain, S. Pelikan and W. Smolenski for helpful conversations, references and/or for copies of their unpublished papers. In particular, R. Bradley pointed out that Gibbs fields on $\mathbb{Z}$ are ψ-mixing, provided suitable references and made available the preprint of Bradley (1989); thanks to S. Pelikan's remarks, Theorems 2.1 and 2.2 were extended from empirical measures to empirical fields. The referee(s) pointed out some errors and helped to improve the presentation.

REFERENCES

BALDI, P. (1988). Large deviations and stochastic homogenization. *Ann. Mat. Pura Appl.* **151** 161–177.

BLUM, J. R., HANSON, D. L. and KOOPMANS, L. H. (1963). On the strong law of large numbers for a class of stochastic processes. *Z. Wahrsch. Verw. Gebiete* **2** 1–11.

BOWEN, R. (1975). *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms. Lecture Notes in Math.* **470**. Springer, New York.

BRADLEY, R. C. (1983). On the ψ-mixing condition for stationary random sequences. *Trans. Amer. Math. Soc.* **276** 55–66.

BRADLEY, R. C. (1989). A caution on mixing conditions for random fields. *Statist. Probab. Lett.* **8** 489–491.

BRETAGNOLLE, J. (1979). Formule de Chernoff pour les lois empiriques de variables á valeurs dans des espaces généraux. *Grandes Déviations et Applications Statistiques. Astérisque* **68** 33–52.

BRYC, W. (1990). Large deviations by the asymptotic value method. In *Diffusion Processes and Related Problems in Analysis* **1** (M. Pinsky, ed.) 447–472. Birkhäuser, Boston.

CHIYONOBU, T. and KUSUOKA, S. (1988). The large deviation principle for hypermixing processes. *Probab. Theory Related Fields* **78** 627–649.

COMETS, F. (1986). Grandes deviations pour des champs de Gibbs sur Z^d. *C. R. Acad. Sci. Paris Sér. I* **303** 511–513.

COMETS, F. (1989). Large deviation estimates for a conditional probability distribution. Application to random interaction Gibbs measures. *Probab. Theory Related Fields* **80** 407–432.

DAWSON, D. A. and GÄRTNER, J. (1987). Large deviations from the McKean–Vlasov limit for weakly interacting diffusions. *Stochastics* **20** 247–308.

DE ACOSTA, A. (1985). Upper bounds for large deviations of dependent random vectors. *Z. Wahrsch. Verw. Gebiete* **69** 551–565.

DE ACOSTA, A. (1988). Large deviations for vector-valued functionals of a Markov chain: Lower bounds. *Ann. Probab.* **16** 925–960.

DE ACOSTA, A. (1990). Large deviations for empirical measures of Markov chains. *J. Theory Probab.* **3** 395–431.

DEUSCHEL, J. D. and STROOCK, D. W. (1989). Large deviations. *Pure Appl. Math.* **137**.

DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators I*. Interscience, New York.

ELLIS, R. S. (1985). *Entropy, Large Deviations and Statistical Mechanics*. Springer, New York.

ELLIS, R. S. (1988). Large deviations for the empirical measure of a Markov chain with an application to the multivariate empirical measure. *Ann. Probab.* **16** 1496–1508.

ELLIS, R. S. and WAYNE, A. D. (1989). Uniform large deviation property of the empirical process of a Markov chain. *Ann. Probab.* **17** 1147–1151.

FÖLLMER, H. and OREY, S. (1988). Large deviations for the empirical field of a Gibbs measure. *Ann. Probab.* **16** 961–977.

GÄRTNER, J. (1977). On large deviations from the invariant measure. *Theory Probab. Appl.* **22** 24–39.

JAIN, N. C. (1990). Large deviation lower bounds for additive functionals of Markov processes. *Ann. Probab.* **18** 1071–1098.

NEY, P. and NUMMELIN, E. (1987a). Markov additive processes. I. Eigenvalue properties and limit theorems. *Ann. Probab.* **15** 561–592.

NEY, P. and NUMMELIN, E. (1987b). Markov additive processes. II. *Ann. Probab.* **15** 593–609.

OLLA, S. (1987). Large deviations for almost Markovian processes. *Probab. Theory Related Fields* **76** 395–409.

OLLA, S. (1988). Large deviations for Gibbs random fields. *Probab. Theory Related Fields* **77** 395–409.

OREY, S. and PELIKAN, S. (1988). Large deviation principles for stationary processes. *Ann. Probab.* **16** 1481–1495.

PLACHKY, D. (1971). On a theorem by G. L. Sievers. *Ann. Math. Statist.* **42** 1442–1443.

PRESTON, C. (1976). *Random Fields. Lecture Notes in Math.* **534**. Springer, New York.

SIEVERS, G. L. (1969). On the probability of large deviations and exact slopes. *Ann. Math. Statist.* **40** 1908–1921.

STROOCK, D. W. (1984). *An Introduction to the Theory of Large Deviations*. Springer, New York.

VARADHAN, S. R. S. (1984). *Large Deviations and Applications*. SIAM, Philadelphia.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CINCINNATI
CINCINNATI, OHIO 45221-0025