Journal of Statistical Physics, Vol. 48, Nos. 3/4, 1987

On the Asymptotics of Occurrence Times of Rare Events for Stochastic Spin Systems

J. L. Lebowitz[1] **and R. H. Schonmann**[1,2]

Received February 24, 1987

We consider translation-invariant attractive spin systems. Let $T^{\nu}_{\Lambda,x}$ be the first time that the average spin inside the hypercube Λ reaches the value x when the process is started from an invariant measure ν with density smaller than x. We obtain sufficient conditions for (1) $|\Lambda|^{-1} \log T^{\nu}_{\Lambda,x} \to \varphi(x)$ in distribution as $|\Lambda| \to \infty$, and $|\Lambda|^{-1} \log ET^{\nu}_{\Lambda,x} \to \varphi(x)$ as $|\Lambda| \to \infty$, where $\varphi(x) := -\lim_{\Lambda} |\Lambda|^{-1} \log \nu\{(\text{average spin inside } \Lambda) \geqslant x\}$. And (2) $T^{\nu}_{\Lambda,x}/ET^{\nu}_{\Lambda,x}$ converges to a unit mean exponential random variable as $|\Lambda| \to \infty$. Both (1) and (2) are proven under some type of rapid convergence to equilibrium. (1) is also proven without extra conditions for Ising models with ferromagnetic pair interactions evolving according to an attractive reversible dynamics; in this case φ is a thermodynamic function. We discuss also the case of finite systems with boundary conditions and what can be said about the state of the system at the time $T^{\nu}_{\Lambda,x}$.

KEY WORDS: Interacting spin systems; large deviations; occurrence times; Glauber dynamics.

1. INTRODUCTION

Large deviations from typical behavior play an important role in many physical, biological, and sociological systems. Their systematic investigation is a topic of much current interest in statistical mechanics and probability theory (see, for instance, Refs. 2, 7, 12–15, 27, and 29). In this paper we investigate the asymptotics of first occurrence (hitting) times of some rare events for spin systems or lattice gases on Z^d evolving under certain types of stochastics dynamics, including Glauber dynamics. These

[1] Mathematics Department, Rutgers University, New Brunswick, New Jersey 08903.
[2] On leave from São Paulo University. Present address: Mathematics Department, Cornell University, Ithaca, New York 14853.

0022-4715/87/0800-0727$05.00/0

systems are frequently referred to as interacting particle systems or random cellular automata.

In our analysis the system will be assumed to be in a statistically stationary state which may or may not be a Gibbs state for some interaction. In either case the rare events will be described by large deviations of a macroscopic quantity, i.e., the magnetization in a large box, from its typical (stationary) value. We are therefore studying large fluctuations around equilibrium. We hope, however, and this was one of our motivations, that the methods developed here will also be useful for studying the large fluctuations around a metastable state that are responsible for its decay.[5,16,24,30,32]

The setting of our problem is very general: we have a space E on which there is defined a stochastic (or deterministic) process $(\xi_t^\nu)_{t \geqslant 0}$ with an invariant measure ν, not necessarily unique; $t \in \mathbb{R}^+$ or Z^+. Let $A_n \subset E$ be a sequence of events such that $\lim_n \nu(A_n) = 0$ and define the sequence of hitting times, starting from ν, as

$$T_n = \inf\{t \geqslant 0 \colon \xi_t^\nu \in A_n\} \tag{1.1}$$

We study two basic questions about the asymptotic behavior of T_n as $n \to \infty$:

1. What is the magnitude of T_n? E.g., the behavior of its expectation, median, etc.
2. What is the asymptotic distribution of T_n? In other words, does there exist a sequence of numbers (a_n) such that T_n/a_n converges in distribution to a measure that is not concentrated on 0 or ∞? In this case what can be said about the limiting distribution?

These kinds of questions have been studied in Refs. 2–4, 5, 20, 23, and 25 when (ξ_t) is a Markov process with good recurrence properties. These conditions are unfortunately not satisfied for the interacting particle systems in which we are interested. There are, however, some things that can be said quite generally. Question 1 is clearly related to the ergodic theory fact that the fraction of time the system spends in A_n is equal to $\nu(A_n)$; this indicates that for systems that after visiting A_n return rapidly to equilibrium, T_n should be of the order of $1/\nu(A_n)$. In discrete time it is easy to prove a corresponding lower bound for T_n. Indeed,

$$P(T_n \leqslant a_n) \leqslant a_n \nu(A_n) \tag{1.2}$$

from which it follows that if $a_n \nu(A_n) \to 0$, then T_n is asymptotically concentrated above a_n. In continuous time one can use a similar argument, provided that the system spends a positive minimal amount of time in A_n before leaving it. {If this is not true, it is easy to think of conterexamples; for instance, consider the deterministic evolution $\xi_t = \xi_0 + t (\text{mod } 1)$, $\nu =$ the Lebesgue measure on $[0, 1)$, and $A_n = [0, 1/n]$, or, even worse, A_n equal to the rationals in $[0, 1)$.} In the cases considered in this paper we will use an argument close to the one above by showing that after reaching A_n the system spends there, with large probability, an amount of time much larger than $\nu(A_n)$. Our main problem as far as (1) is concerned will be to obtain an upper bound on T_n that is also of the same order.

To see what may go wrong with the upper bound, we can imagine situations where the system, after reaching A_n for the first time, returns to A_n many times before relaxing back to equilibrium. This is the type of behavior that might be expected to occur when A_n is separated from "typical" states by a "potential barrier," e.g., when A_n represents a metastable state. In these cases the magnitude of T_n should be related to the reciprocal of the height of the barrier, which can be much larger than $\nu(A_n)^{-1}$. This is in fact what happens in certain mean field models of ferromagnetic systems with Glauber type dynamics[5,19] and weakly perturbed dynamical systems in double-well potentials.[16,24] (The analysis in these papers is made with the system starting in the higher well, which corresponds to a metastable state, and one is interested there in the passage to the lower well, which corresponds to equilibrium. Nevertheless, it is clear that the same kind of analysis applies for the reverse phenomenon.)

In the present paper we show that for attractive stochastic spin systems (those that satisfy an FKG-type property[28]) a rapid rate of approach to a stationary state (which can be proven in many cases) leads naturally to an upper bound on the T_n that coincides to leading order with the lower bound. The proof goes via an inequality, which may also be valid in other cases. Assume that $\lim_n n^{-1} \log \nu(A_n) = -\varphi$ and that for all k and $0 \leqslant t_1 \leqslant t_2 \cdots < t_k < t$

$$|P(\xi_t^\nu \in A_n | \xi_{t_i} \in (A_n)^c, i = 1,..., k) - \nu(A_n)| \leqslant f(n)\, e^{-\lambda(t - t_k)} \tag{1.3}$$

with $\lambda > 0$ and $f(n)$ growing only polynomially with n. Then we have the result that

$$n^{-1} \log T_n \to \varphi \text{ in distribution as } n \to \infty \tag{1.4}$$

For attractive spin flip dynamics whose rates satisfy detailed balance with respect to a Gibbs state for an Ising system with translation-invariant interactions (attractive stochastic Ising models in the terminology of Ref. 28 or attractive Glauber dynamics) we will prove results of the type (1.4) without assuming conditions of the type (1.3).

Question 2 is not related directly to good ergodic properties and the answer to it can be expected to be less universal. Assuming, however, that the system makes many "attempts" to reach A_n before it succeeds and that after each failure it comes back to "typical" configurations, then T_n/a_n should converge to an exponential random variable for suitable a_n (see Ref. 3 for more on these heuristics). This is what happens in a much studied variant of our problem obtained by keeping the set A fixed, but modifying the process to obtain a sequence of stationary states ν_n so that $\nu_n(A) \to 0$. This is the setup for the Freidlin–Wentzell theory[15] of perturbations of deterministic evolution. One is interested, then, in the asymptotics, as the random perturbation vanishes of hitting times of the complements of domains that include an attracting point. In this case the results indicated by the heuristics above have been proven to be correct with great generality.[8,15,34] We prove the results in our case under rapid convergence to the stationary state. (In Ref. 15 information about the hitting place was also obtained; we will obtain some information of this type also in our case.)

In the next section we introduce the type of interacting particle systems that we will be considering and state the results. The proofs appear in Sections 3 and 4.

2. THE MODELS AND RESULTS

2.1. The Models and Terminology

The systems we consider in the paper are interacting spin systems. The state space is $E_d = \{0, 1\}^{Z^d}$ (endowed with the product topology) for some dimension d; elements of this space are called configurations and will be represented by the letters η, ζ, and ξ. For a given configuration η and $i \in Z^d$, $\eta(i) = 0$ or 1 will be called the spin at site i. Measures on E_d will be denoted by μ and ν. We let $\mathscr{T}$ denote that set of measures that are invariant with respect to the translations of Z^d. We let δ_a, $a = 0, 1$, be the measure concentrated on the configuration identically a. Let C be the set of continuous functions from E to R and C_+ be the subset of C of functions that are coordinatewise nondecreasing. μ is said to be stochastically greater than ν if, for any $f \in C_+$,

$$\int f\, d\nu \leqslant \int f\, d\mu$$

We write in this case $\mu \geqslant \nu$. A measure μ is said to have positive correlation or to be FKG if for any pair $f, g \in C_+$,

$$\int fg\, d\mu \geqslant \int f\, d\mu \int g\, d\mu$$

Sometimes it will be more convenient to consider the state space as $\{-1, 1\}^{Z^d}$; for this purpose we introduce the variables

$$\sigma(i) = 2\eta(i) - 1$$

We consider the class of *translation-invariant attractive spin systems*,[28] hereafter denoted by $\mathscr{A}$. These are Markov and Feller processes with state space E_d, whose evolution is given by the flip rates $c(i, \eta)$ [the rate at which $\eta(i)$ flips to $1-\eta(i)$ when the system is in the configuration η]. Translation invariance means that $c(i, \eta) = c(i+j, \eta+j)$, where $(\eta+j)(k) = \eta(k+j)$. Attractiveness means informally that zeros attract zeros and ones attract ones. More precisely, if the configuration η is dominated by ζ, in the sense that $\eta(j) \leqslant \zeta(j)$, $\forall j \in Z^d$, then

$$c(i, \eta) \leqslant c(i, \zeta) \qquad \text{if} \quad \eta(i) = \zeta(i) = 0$$

$$c(i, \eta) \geqslant c(i, \zeta) \qquad \text{if} \quad \eta(i) = \zeta(i) = 1$$

In order for the infinitesimal rates $c(i, \eta)$ to define a unique process, one must assume that they do not depend very strongly on the spins at sites far away from i. A sufficient condition[28] is that $c(0, \cdot)$ be a continuous function and that

$$\sum_{i \in Z^d} \sup_{\eta \in E_d} |c(0, \eta) - c(0, \eta_i)| < \infty \tag{2.1}$$

where η_i is the configuration obtained from η by flipping the spin at the site i. If (2.1) holds, then

$$c = \sup_{\eta} c(0, \eta) < \infty$$

Let $S(t)$ denote the semigroup corresponding to the process above and write $\mu S(t)$ for its action on a measure μ. The set of invariant measures under $S(t)$ will be denoted by $\mathscr{I}$.

The process will be denoted by $(\xi_t^\mu, t \geqslant 0)$ or (ξ_t^μ), where μ is the initial distribution. If μ is concentrated on a confuguration η, we write $(\xi_t^\eta, t \geqslant 0)$. Expectations with respect to these processes will be indicated by $E(\cdot)$.

Some of the fundamental facts about systems in $\mathscr{A}$ are summarized next (for proofs see Ref. 28):

1. $\delta_0 S(t)$ [resp. $\delta_1 S(t)$] converges weakly to a masure ν_- [resp. ν_+] which belongs to $\mathscr{T} \cap \mathscr{I}$ and is called the lower [resp. upper] invariant measure.
2. ν_- and ν_+ are FKG and ergodic with respect to translations.

3. If $\nu \in \mathscr{I}$, then $\nu_- \leqslant \nu \leqslant \nu_+$.
4. Let $\rho_\pm = \nu_\pm\{\eta(0) = 1\}$. Then the following three statements are equivalent: (a) The process is ergodic, i.e., there exists a measure ν such that $\mu S(t)$ converges weakly to ν for any μ; (b) $\nu_- = \nu_+$; $\rho_- = \rho_+$. Clearly in this situation $\nu = \nu_- = \nu_+$.
5. If $\mu \in \mathscr{T}$ is FKG and $\mu S(t)$ converges weakly to a measure ν, then $\nu \in \mathscr{T} \cap \mathscr{I}$ and is FKG.

An important subclass of $\mathscr{A}$ is that of the stochastic Ising models (Glauber dynamics) with attractive flip rates $c(\cdot, \cdot)$ corresponding to systems with ferromagnetic, translation-invariant pair interactions. This class will be denoted by $\mathscr{F}$. To define it, one considers a constant $H \in \mathbb{R}$ and a nonnegative, real-valued function J on Z^d that satisfies $J(i) = J(-i)$, $\forall i \in Z^d$, and

$$\sum_{i \in Z^d} J(i) < \infty$$

J is called the interaction and H the external field. In order to define the dynamics, one imposes on the $c(i, \eta)$ the condition of reversibility (detailed balance):

$$\begin{aligned} &c(i, \eta) \exp\left[\beta \sum_j J(i-j)\, \sigma(i)\, \sigma(j) + \beta H \sigma(i)\right] \\ &\qquad = c(i, \eta_i) \exp\left[-\beta \sum_j J(i-j)\, \sigma(i)\, \sigma(j) - \beta H \sigma(i)\right] \end{aligned} \tag{2.2}$$

where η_i is the configuration obtained from η by flipping the spin at site i, and $\beta > 0$ is called the inverse temperature. For any J, H, and β as above there are rates $c(i, \eta)$ that satisfy (2.1) and (2.2) and are attractive.[22,28] Two such rates that appear frequently in the literature are

$$c(i, \eta) = \exp\left\{-\beta\left[\sum_j J(i-j)\, \sigma(i)\, \sigma(j) + H\sigma(i)\right]\right\}$$

$$c(i, \eta) = \left(1 + \exp\left\{2\beta\left[\sum_j J(i-j)\, \sigma(i)\, \sigma(j) + H\sigma(i)\right]\right\}\right)^{-1}$$

It is known[28] that $\mathscr{T} \cap \mathscr{I}$ is in this case identical to the intersection of $\mathscr{T}$ with the set $\mathscr{G}$ of Gibbs measures for the interaction J and external field H at inverse temperature β, i.e., the set of measures such that a version of the conditional probability

$$\nu\{\eta: \sigma(i) = \tau(i) \mid \sigma(j) = \tau(j)\ \forall j \neq i\}$$

is given by

$$\left\{1+\exp\left[-2\beta\sum_j J(i-j)\,\tau(i)\,\tau(j)-2\beta H\tau(i)\right]\right\}^{-1}$$

for any $\tau\in\{-1,1\}^{Z^d}$.

We are interested in the process $(\xi_t^\nu,\ t\geqslant 0)$ starting from an invariant measure. In order to define the rare events that we consider, let $(\Lambda_n)_{n=1,2,\dots}=(\Lambda)$ be a sequence of d-dimensional hypercubes in Z^d converging to the whole space Z^d. Given $\Omega\subset Z^d$, let $|\Omega|$ be its cardinality and for a given $\eta\in E$ set

$$X_\Omega=X_\Omega(\eta)=|\Omega|^{-1}\sum_{i\in\Omega}\eta(i)$$

Define now

$$A_{\Lambda,x}^+=\{\eta\colon X_\Lambda(\eta)\geqslant x\}$$

Given μ and x, define the hitting times

$$T_{\Lambda,x}^{+,\mu}=\inf\{t\geqslant 0\colon \xi_t^\mu\in A_{\Lambda,x}^+\}$$

Its quantils $\beta_{\Lambda,x}^{+,\mu}(a)$ are defined by

$$P(T_{\Lambda,x}^{+,\mu}\geqslant\beta_{\Lambda,x}^{+,\mu}(a))=a$$

for $a\in(0,1)$. The most important one will be $\beta_{\Lambda,x}^{+,\mu}(e^{-1})=\beta_{\Lambda,x}^{+,\mu}$. There are analogous definitions for $A_{\Lambda,x}^-$, $T_{\Lambda,x}^{-,\mu}$, and $\beta_{\Lambda,x}^{-,\mu}(a)$.

Warnings: We will sometimes omit the indices $+$, μ, x when the meaning is clear. We may also use a bar instead of the index ν_-, so $\bar{T}_{\Lambda,x}^+=T_{\Lambda,x}^{+,\nu_-}$, etc. And we write $\nu\{\eta(0)=1\}$ instead of $\nu\{\eta\colon\eta(0)=1\}$, etc.

2.2. Main Results

The following was proven in Ref. 27:

Theorem 0. Suppose that ν is an invariant measure for a system in $\mathscr{A}$, which is also FKG and translation-invariant and is neither δ_0 nor δ_1. Then there exists a convex function $\varphi_\nu\colon[0,1]\to[0,\infty)$ such that:

(a) $\varphi_\nu(x)>0$ if $x<\rho_-$ or $x>\rho_+$

(b) $\lim_\Lambda|\Lambda|^{-1}\log\nu(A_{\Lambda,x}^+)=-\varphi_\nu(x)$ if $x>\rho$

$\lim_\Lambda|\Lambda|^{-1}\log\nu(A_{\Lambda,x}^-)=-\varphi_\nu(x)$ if $x<\rho$

where

(c) $\varphi_\nu(x) = \sup_{h \in \mathbb{R}} \{hx - \pi_\nu(h)\}$

where

$$\pi_\nu(h) = \lim_{\Lambda} |\Lambda|^{-1} \log \int \exp(h |\Lambda| X_\Lambda)\, d\nu$$

Under some extra conditions we will prove the following statements for $x > \rho$ [the first three are clearly equivalent and will be referred to as (S1) later]:

(S1a) $\lim_{\Lambda} |\Lambda|^{-1} \log \beta^{+,\nu}_{\Lambda,x}(a) = \varphi_\nu(x), \quad \forall a \in (0, 1)$

(S1b) $\lim_{\Lambda} P\{\exp[|\Lambda| (\varphi_\nu(x) - \varepsilon)] < T^{+,\nu}_{\Lambda,x} < \exp[|\Lambda| (\varphi_\nu(x) + \varepsilon)]\} = 1, \ \forall \varepsilon > 0$

(S1c) $|\Lambda|^{-1} \log T^{+,\nu}_{\Lambda,x}$ converges in law to a degenerate distribution concentrated on $\varphi_\nu(x)$

(S2) $\lim_{\Lambda} |\Lambda|^{-1} \log ET^{+,\nu}_{\Lambda,x} = \varphi_\nu(x)$

(S3) $T^{+,\nu}_{\Lambda,x}/\beta^{+,\nu}_{\Lambda,x}$ converges in law to a unit mean exponential random variable as $|\Lambda| \to \infty$

(S4) $\lim_{\Lambda} ET^{+,\nu}_{\Lambda,x}/\beta^{+,\nu}_{\Lambda,x} = 1$

We will assume some conditions on the system in order to be able to prove these statements. We believe that the results are true with greater generality, but it should be clear that some conditions are needed. For instance, if the system is not ergodic, and we pick $\nu = \frac{1}{2}\nu_- + \frac{1}{2}\nu_+$ and $\frac{1}{2}\rho_- + \frac{1}{2}\rho_+ < x < \rho_+$, then (S3) should be false.

We have two types of conditions. One type [(C.Exp.) and (C.Pol.)] in terms of the rate of convergence to equilibrium and the other [(C.Pr.)] in terms of the approximation of the system by finite systems. Let

$$D_t(\mu) = |\mu S(t)\{\eta(0) = 1\} - \rho_-|$$

The first conditions are then

(C.Exp.) $D_t(\delta_0) \leqslant C \exp(-\alpha t^\delta)$ for some positive C, α, δ

(C.Pol.) $D_t(\delta_0) \leqslant Ct^{-\alpha}$ for some positive C and $\alpha > 1$

Remark 1. Condition (C.Exp.) is clearly true for exponentially ergodic systems.[28] This class includes the extralineal proximity systems,[18] i.e., additive spin systems with a positive rate of spontaneous births (jumps from 0 to 1). It also includes systems with weak interactions, such as stochastic Ising models at high temperature (β small)[1,21] or in one dimension,[21] and systems defined by adding independent flips with sufficiently large rates to a system in $\mathscr{A}$.[28]

Remark 2. The analogue of (C.Exp.) from the other side (i.e., replacing ρ_- by ρ_+ and δ_0 by δ_1) is known to hold for the supercritical basic contact process in one dimension.[9] This result was extended in Ref. 17 to a larger class of systems in $\mathscr{A}$ with nearest neighbor interactions.

Remark 3. An important tool that we will use in our analysis is the basic coupling.[28] For $\mu \leqslant \nu$ one can construct (ξ_t^μ) and (ξ_t^ν) on the same probability space in such a way that the mass is concentrated on paths such that $\xi_t^\mu(i) \leqslant \xi_t^\nu(i)\ \forall i \in Z^d,\ \forall t \geqslant 0$. It follows that for $\mu \leqslant \nu_-$

$$\begin{aligned} P(\xi_t^\mu(0) \neq \xi_t^{\nu_-}(0)) &= P(\xi_t^\mu(0) = 0,\ \xi_t^{\nu_-}(0) = 1) \\ &= P(\xi_t^{\nu_-}(0) = 1) - P(\xi_t^\mu(0) = 1) \\ &= D_t(\mu) \leqslant D_t(\delta_0) \end{aligned} \tag{2.3}$$

In order to specify the other condition, we first need some definitions. For $N = 1, 2, \ldots$ let $\Gamma = \Gamma_N = \{1, \ldots, N\}^d$ and consider the new process on E_d defined by the rates

$$\tilde{c}(i, \eta) = \begin{cases} c(i, \eta) & \text{if } i \in \Gamma \\ 0 & \text{if } i \notin \Gamma \end{cases}$$

One can think of the spins outside Γ as a fixed boundary condition. The new system is also attractive; it is in general not translation-invariant, but as in fact 1 of Section 2.1 one can define the lower and upper invariant measures ν_-^N and ν_+^N by starting with δ_0 or δ_1 and letting $t \to \infty$. Define now

$$\pi_N^\pm(h) = |\Gamma|^{-1} \log \int d\nu_\pm^N \exp(h |\Gamma| X_\Gamma) \tag{2.4}$$

We have the following condition:

(C.Pr.) For any $h \in \mathbb{R}$, $\pi_N^-(h)$ and $\pi_N^+(h)$ converge as $N \to \infty$ to the same limit (which depends on h).

Remark 4. By attractiveness it folows (see Theorem 2.7, Chapter III of Ref. 28) that $\nu_-^N \leqslant \nu_+^N$. Therefore it is clear that (C.Pr.) is equivalent to

$$\limsup_{N\to\infty} \pi_N^+(h) \leqslant \liminf_{N\to\infty} \pi_N^-(h) \qquad \text{if} \quad h \geqslant 0$$

$$\limsup_{N\to\infty} \pi_N^-(h) \leqslant \liminf_{N\to\infty} \pi_N^+(h) \qquad \text{if} \quad h < 0$$

Remark 5. If $\nu \in \mathscr{I}$, it follows as above that $\nu_-^N \leqslant \nu \leqslant \nu_+^N$. So

$$\pi_N^-(h) \leqslant |\Gamma|^{-1} \log \int \exp(h|\Gamma| X_\Gamma)\, d\nu \leqslant \pi_N^+(h) \qquad \text{if} \quad h \geqslant 0$$

$$\pi_N^+(h) \leqslant |\Gamma|^{-1} \log \int \exp(h|\Gamma| X_\Gamma)\, d\nu \leqslant \pi_N^-(h) \qquad \text{if} \quad h < 0$$

Consequently, (C.Pr.) implies that all the invariant measures have the same "pressure" $\pi_\nu(h) = \pi(h)$ and that

$$\lim_{N\to\infty} \pi_N^{\pm}(h) = \pi(h) \tag{2.5}$$

By the uniqueness of the Legendre–Frechet transform, it follows that $\varphi_\nu(x)$ does not depend on ν. Set $\varphi_\nu(x) = \varphi(x)$. Since ν_- is ergodic with respect to translations, it follows that $\varphi(\rho_-) = 0$. Analogously, $\varphi(\rho_+) = 0$. By the convexity of $\varphi(\cdot)$ and part (a) of Theorem 0, we conclude that in this case

$$\varphi(x) = 0 \Leftrightarrow x \in [\rho_-, \rho_+]$$

Remark 6. It is easy to verify that (C.Pr.) holds for the systems in $\mathscr{F}$. The ν_-^N and ν_+^N restricted to Γ are the Gibbs measures with $-$ and $+$ boundary conditions and it is well known[31,33] that in the thermodynamic limit, $N \to \infty$, $\pi_N^{\pm}(h)$ converge to the same limit, which is, up to translations, the thermodynamic pressure. Large deviations for the translation-invariant Gibbs measures (for a much more general class of potentials) were first studied in Ref. 26 and more recently (in a stronger sense) in Refs. 7, 13, 14, and 29. The following result, stronger then part b of Theorem 0, holds for any measure ν in this class:

$$\lim_{\Lambda} |\Lambda|^{-1} \log \nu\{X_\Lambda \in [a, b]\} = -\inf_{a \leqslant x \leqslant b} \varphi(x) \tag{2.6}$$

for any $0 \leqslant a < b \leqslant 1$.

We state now our main results for systems in $\mathscr{A}$.

Proposition 1. If $\nu \in \mathscr{T} \cap \mathscr{I}$ and is FKG, then for any $x > \rho = \nu\{\eta(0) = 1\}$

(i) $\liminf_{\Lambda} |\Lambda|^{-1} \log \beta^{+;\nu}_{\Lambda,x}(a) \geqslant \varphi_\nu(x) \qquad \forall a \in (0, 1)$

(ii) $\liminf_{\Lambda} |\Lambda|^{-1} \log ET^{+;\nu}_{\Lambda,x} \geqslant \varphi_\nu(x)$

Theorem 1. Assume that (C.Pol.) holds. Then (S3) and (S4) hold for ν_- and any $x > \rho_-$ such that $\varphi_{\nu_-}(x) > 0$.

Theorem 2. Assume that (C.Exp.) holds. Then (S1) holds for ν_- and any $x > \rho_-$. If $\varphi_{\nu_-}(x) > 0$, then (S2) also holds for ν_-.

Corollary 1. Assume that (C.Exp.) holds and $\nu \in \mathscr{I}$. Then for $x > \rho_-$

(i) $\limsup_{\Lambda} |\Lambda|^{-1} \log \beta^{+;\nu}_{\Lambda,x}(a) \leqslant \varphi_{\nu_-}(x) \qquad \forall a \in (0, 1)$

(ii) $\limsup_{\Lambda} |\Lambda|^{-1} \log ET^{+;\nu}_{\Lambda,x} \leqslant \varphi_{\nu_-}(x) \qquad \text{if} \quad \varphi_{\nu_-}(x) > 0$

Theorem 3. If $\nu \in \mathscr{I}$, then for any $N \in \{1, 2,...\}$ and $x \in (\rho_-, 1)$

(i) $\limsup_{\Lambda} |\Lambda|^{-1} \log \beta^{+;\nu}_{\Lambda,x}(a) \leqslant \varphi^-_N(x) \qquad \text{for any} \quad 0 < a < 1$

(ii) $\limsup_{\Lambda} |\Lambda|^{-1} \log ET^{+;\nu}_{\Lambda,x} \leqslant \varphi^-_N(x)$

where

$$\varphi^-_N(x) = \sup_h \{hx - \pi^-_N(h)\}$$

Corollary 2. Assume that (C.Pr.) holds and $\nu \in \mathscr{I}$. Then for any $x \in (\rho_-, 1)$, (1) and (S2) hold.

Clearly, there are analogous statements for $T^{-;\nu}_{\Lambda,x}$ with ν_+ in place of ν_-, etc. In the case of exponentially ergodic systems (see Remark 1), Theorems 1 and 2 state that (S1)–(S4) hold when ν is the invariant measure. For the supercritical one-dimensional basic contact process (see Remark 2), the analogues of (S1)–(S4) hold for $T^{-;\nu_+}_{\Lambda,x}$, $x < \rho_+$. This should be compared with similar results obtained, when $x = 0$, for the contact process on a finite set $\Lambda = \{1,..., n\}$ (i.e., when there is no infection from Λ^c to Λ) in Refs. 5, 10, 11, and 32.

Proposition 1 is a very general and easy to prove statement that the time $T^{+;\nu}_{\Lambda,x}$ is bounded below by $\exp\{[\varphi(x) - \varepsilon] |\Lambda|\}$ for any $\varepsilon > 0$. In Theorem 2 we give sufficient conditions for a sharp bound in the other direction, but it works only for ν_-. This result has as an immediate consequence Corollary 1, which gives an upper bound for a general $\nu \in \mathscr{I}$, but it is not necessarily the best upper bound. In Theorem 3 we obtain another upper bound for a general $\nu \in \mathscr{I}$, this time without any extra condition.

This bound is sharp when (C.Pr.) holds (Corollary 2); in particular, for any system in $\mathscr{F}$. This last result holds even at critical temperatures, when the hypothesis of Theorem 2 should be false (it is in fact known to be false in the case of the two-dimensional nearest neighbor ferromagnetic Ising model[22]). Furthermore, since $\pi(h)$ is, up to translations, the thermodynamic pressure, one is relating the time to see a large fluctuation in a stochastic Ising model to a purely thermodynamic (equilibrium) quantity. In particular, the result does not depend on which attractive rates $c(i, \eta)$ satisfying (2.1) and (2.2) are chosen. A particularly interesting case is when $H=0$ and β and J are such that $\nu_- \neq \nu_+$, i.e., when there is a phase transition. In this case $\varphi(x)$ is identically zero for $\rho_- < x < \rho_+$ and positive outside this interval (see Remarks 5 and 6). For any $\nu \in \mathscr{I}$, Corollary 2 implies then that $T^{+,\nu}_{\Lambda,x}$ (and $T^{-,\nu}_{\Lambda,x}$) does not grow exponentially with $|\Lambda|$ if $x \in [\rho_-, \rho_+]$, but does grow exponentially otherwise.

2.3. Dynamics with Boundary Conditions

A problem related to those discussed up to now is that in which one considers systems with fixed boundary conditions outside Λ. Then as Λ increases, both the dynamics and the events we are waiting for to happen are being modified. One has a large system and is observing the behavior of the whole system, while in the cases treated before one was observing a large part of a system that is in fact much larger. The techniques used to prove the previous results work only in part for this type of problem. Theorems 1 and 2 have analogues in this case depending on estimates similar to (C.Pol.) and (C.Exp.), which are uniform in the boundary conditions. These types of conditions are verified for extralineal proximity systems.[18] Theorem 3 and Corollary 2, on the other hand, go through with the same proofs. In particular, for attractive stochastic Ising models with any sequence of boundary conditions outside Λ the following result holds. For each Λ consider the process starting from its unique invariant (Gibbs) measure corresponding to the chosen boundary conditions. Let T_Λ be the first time that the average magnetization $M_\Lambda = |\Lambda|^{-1} \sum_{i\in\Lambda} \sigma(i)$ reaches a fixed, nonempty open set $S \subset (-1, 1)$. Define the free energy associated to S (as usual, except for a factor β^{-1}) as

$$F(S) = -\lim_{\Lambda} |\Lambda|^{-1} \log \sum_{\substack{\sigma \in \{-1,+1\}^\Lambda \\ M_\Lambda \in S}} \exp\left\{\beta\left[\sum_{i,j\in\Lambda} J(i-j)\,\sigma(i)\,\sigma(j) + H \sum_{i\in\Lambda} \sigma(i)\right]\right\}$$

Then, as $|\Lambda| \to \infty$,

$$|\Lambda|^{-1} \log T_\Lambda \to F(S) - F([-1, 1]) \qquad \text{in distribution}$$

$$|\Lambda|^{-1} \log ET_\Lambda \to F(S) - F([-1, 1])$$

2.4. Structure of Trajectories

The previous results can also give some information about how the system looks at the time $T^{\nu,+}_{\Lambda,x}$. In fact, with the techniques used to prove them, it is not hard to prove the following result:

Proposition 2. Assume that (S1) holds. Let B_Λ be a sequence of sets of configurations and suppose that

$$\limsup_{\Lambda} |\Lambda|^{-1} \log \nu(B_\Lambda) < -\varphi(x) \tag{2.7}$$

Then

$$\lim_{\Lambda} P(\xi^{\nu}_{T^{\nu,+}_{\Lambda,x}} \in B_\Lambda) = 0$$

This result is particularly useful for systems in $\mathscr{F}$, since for them condition (2.7) can be expressed in terms of restricted partition functions and in many interesting cases one can verify this condition. For these systems an analogue of Proposition 2 holds also when boundary conditions outside Λ are fixed. In a sense to be made precise below, Proposition 2 implies that for systems in $\mathscr{F}$ the system at the time $T^{\nu,+}_{\Lambda,x}$ is very likely to be in a typical configuration of a Gibbs measure that corresponds to a different value of the external magnetic field H. For this purpose consider the function $h(m)$ such that

$$\int \sigma_0 \, d\mu_{h(m)} = m$$

where μ_h is a Gibbs measure for the system with the same interaction $J(\cdot)$ and inverse temperature β and external magnetic field $H+h$. The basic properties of the Gibbs measures of systems in $\mathscr{F}$[12,31,33] assure that $h(m)$ is well defined and unique for $-1 \leqslant m \leqslant 1$. If $J(\cdot)$ and β are such that more than one Gibbs measure exists for $H=0$, then different values of m correspond to the same h. More precisely, in this case, there exists $m^* \in (0, 1)$ such that $h(m) = -H$ for $-m^* \leqslant m \leqslant m^*$. Now we are ready to state:

Proposition 3. Let B_Λ be a sequence of sets of configurations and suppose that

$$\mu_{h(m)}(B_\Lambda) \leqslant Ce^{-\gamma|\Lambda|} \tag{2.8}$$

for some positive C and γ. Let ν be a translation-invariant Gibbs measure and assume $m > \int \sigma(0)\, d\nu$. Set $x = (m+1)/2$. Then

$$\lim_{\Lambda} P(\xi^{\nu}_{T^{\nu,+}_{\Lambda,x}} \in B_\Lambda) = 0$$

To some extent this proposition says that at time $T^{+,\nu}_{\Lambda,x}$ the system looks as if it were in a Gibbs ensemble for the system with modified magnetic field $H + h(m)$. This is the case if $H + h(m) \neq 0$, so that $\mu_{h(m)}$ is unique. In this case it is, for instance, easy to verify from Proposition 3 that the expected value with respct to $\xi^{\nu}_{T^{+,\nu}_{\Lambda,x}}$ of the spatial average of translations of any cylindrical function must be close to the expected value of this function with respect to $\mu_{h(m)}$. It is also easy to prove that if Λ is divided in two parts that grow to infinity, the average $\sigma(i)$ in each one of these regions must be close to m at time $T^{\nu,+}_{\Lambda,x}$.

2.5. Independent Systems

The hypothesis of attractiveness made in all of the above statements is certainly not a necessary one. Nevertheless, as will be seen, it plays a major role in the proofs. We mention now only one very particular case of a process for which part of the proofs can be adapted. Let $W = \{w_1,..., w_r\}$ be a finite set of real numbers. To each $i \in Z^d$ associate a continuous-time irreducible Markov process $\xi_t(i)$ with state space W. These processes are assumed to be mutually indepedent and to have the same transition rates. We will refer to these processes as independent processes and when dealing with them we will use ξ and η for configurations in W^{Z^d} and μ and ν for measures on this space. $A^+_{\Lambda,x}$, $T^+_{\Lambda,x}$, and $\beta^+_{\Lambda,x}(a)$ are defined as before. Systems in this class have clearly only one invariant measure ν, which is a product of measure ν_0 on W. Therefore, ν has large-deviation properties as stated in Theorem 0, with

$$\pi_\nu(h) = \sum_{i=1}^{r} \nu_0(w_i)\, e^{hw_i}$$

$$\varphi_\nu(x) = \sup_h \{xh - \pi_\nu(h)\}$$

For this class of systems we will prove:

Proposition 4. Let ν be the unique invariant measure for an independent process. Then for x greater than the density of ν, (S1) and (S2) hold.

In the case $W=\{0, 1\}$ the independent processes are in fact attractive. In the particular case in which the flips from 0 to 1 and from 1 to 0 occur with the same rate the system is the well-known Ehrenfest model (in continuous time). In this case the convergence to unit mean exponentials for the renormalized hitting times of rare events was first proven in Ref. 4.

3. PROOFS UNDER RAPID CONVERGENCE TO EQUILIBRIUM

Proof of Proposition 1. (i) If $\varphi(x)=0$, the proof is easy; therefore we assume $\varphi(x)>0$.

Set

$$S_A=\inf\{t\geqslant T_A : \xi_t^\nu\notin A_{A,x}^+\}$$

Clearly S_A-T_A has tails that are greater than those of an exponential random variable with mean $(c|A|)^{-1}$.

Given $0<\gamma<\varphi(x)$, let n_A be the integer part of $e^{\gamma|A|}|A|^2$. Then

$$\begin{aligned}
&P(T_A<\exp(\gamma|A|))\\
&\quad\leqslant P(S_A-T_A\leqslant 2|A|^{-2})\\
&\qquad+P(\xi_t^\nu\in A_{A,x}^+ \text{ for some } t\in\{k|A|^{-2}\colon\ k=1, 2,..., n_A+1\})\\
&\quad\leqslant[1-\exp(c|A|\ 2|A|^{-2})]+[|A|^2\exp(\gamma|A|)+1]\ \nu(A_{A,x}^+)
\end{aligned}$$

By Theorem 0, part b, the right-hand side of the above expression vanishes as $|A|\to\infty$. Therefore, for $a\in(0, 1)$ and $|A|$ large enough, $\beta_A(a)>e^{\gamma|A|}$. Hence

$$\liminf_A |A|^{-1}\log\beta_A(a)\geqslant\gamma$$

(ii) $ET_A\geqslant\beta_A P(T_A>\beta_A)=\beta_A\cdot e^{-1}$. Therefore the result follows from (i) above. ∎

The proof of Theorem 1 will be divided into several lemmas. The basic idea is that conditioned on not having hit $A_{A,x}^+$ by time t the process is in a state that is stochastically smaller than ν_- (Lemma 1) and then (C.Pol.) implies that in an appropriate sense it approaches ν_- again very fast (Lemma 2).

Lemma 1. Assume that $\nu \in \mathscr{T} \cap \mathscr{I}$ and is FKG. Fix $t>0$ and x and let μ be the conditional distribution of ξ_t^ν given $T_{\Lambda,x}^{+,\nu} > t$. Then $\mu \leqslant \nu$.

Proof. Let $\{t_1, t_2,...\}$ be an enumeration of the rational numbers in $[0, t]$ such that $t_1 = t$. By Corollary 2.21 of Chapter II of Ref. 28 it follows that for any $n \in N$ and any pair of continuous increasing functions f, g: $(E_d)^n \to \mathbb{R}$

$$E(f(\xi_{t_1}^\nu,..., \xi_{t_n}^\nu) \cdot g(\xi_{t_1}^\nu,..., \xi_{t_n}^\nu)) \\ \geqslant E(f(\xi_{t_1}^\nu,..., \xi_{t_n}^\nu)) \cdot E(g(\xi_{t_1}^\nu,..., \xi_{t_n}^\nu)) \tag{3.1}$$

Take

$$g(\eta_1,..., \eta_n) = \begin{cases} 1 & \text{if} \quad \eta_i \in A_{\Lambda,x}^+ \text{ for some } i = 1,..., n \\ 0 & \text{otherwise} \end{cases}$$

and let f depend only on its first coordinate, $f(\eta_1,..., \eta_n) = f(\eta_1)$. Define the event

$$G_n = \{\xi_s^\nu \notin A_{\Lambda,x}^+, s = t_1,..., t_n\}$$

Then it follows from (3.1) that

$$E(f(\xi_t^\nu); (G_n)^c) \geqslant E(F(\xi_t^\nu)) \, P((G_n)^c)$$

Hence

$$E(f(\xi_t^\nu) \,|\, G_n) \leqslant E(f(\xi_t^\nu))$$

But as $n \to \infty$, G_n decreases to $\{T_\Lambda > t\}$. So

$$\int f \, d\mu = E(f(\xi_t^\nu) \,|\, T_\Lambda > t) \leqslant E(f(\xi_t^\nu)) = \int f \, d\nu \quad \blacksquare$$

Lemma 2. Assume that (C.Pol) holds, and $\mu \leqslant \nu_-$. Fix a $t>0$ and $x > \rho_-$ such that $\varphi_{\nu_-}(x) > 0$. Then

$$|P(T_{\Lambda,x}^{+,\mu} > t\beta_{\Lambda,x}^{+,\nu_-}) - P(T_{\Lambda,x}^{+,\nu_-} > t\beta_{\Lambda,x}^{+,\nu_-})| \to 0 \qquad \text{as} \quad |\Lambda| \to \infty$$

Proof. Consider a numerical sequence (b_Λ), which will be chosen later. From the basic coupling

$$|P(T_\Lambda^\mu > t\beta_\Lambda) - P(\bar{T}_\Lambda > t\beta_\Lambda)| \\ \leqslant P(\bar{T}_\Lambda \leqslant b_\Lambda) + P(\bar{S}_\Lambda - \bar{T}_\Lambda < 2|\Lambda|^{-2}) \\ + P(\bar{\xi}_\Lambda(i) \neq \xi_s^\mu(i) \text{ for some } i \in \Lambda \text{ and some } \\ s \in \{b_\Lambda + k|\Lambda|^{-2}: k = 0, 1,...\} \cup \{t\beta_\Lambda\}) \tag{3.2}$$

where, as in the proof of Proposition 1,

$$\bar{S}_A = \inf\{t \geqslant \bar{T}_A : \xi_t \notin A^+_{A,x}\}$$

We know already that the second term on the rhs of (3.2) vanishes as $|A| \to \infty$. Therefore, we need to find b_A such that the other two terms also vanish. The last one is smaller than

$$|A| \left\{ \sum_{k=0}^{\infty} C(b_A + k|A|^{-2})^{-\alpha} + C(t\bar{\beta}_A)^{-\alpha} \right\}$$

$$\leqslant |A|^3 C \sum_{i=0}^{\infty} (b_A + i)^{-\alpha} + |A| C(t\bar{\beta}_A)^{-\alpha}$$

$$\leqslant |A|^3 C \int_{b_A - 1}^{\infty} x^{-\alpha}\, dx + |A| C(t\bar{\beta}_A)^{-\alpha}$$

$$= |A|^3 C(\alpha - 1)^{-1} (b_A - 1)^{-\alpha+1} + |A| C(t\bar{\beta}_A)^{-\alpha}$$

Using Proposition 1 and the fact that $\varphi_{v_-}(x) > 0$, it is easy to see now that the choice $b_A = |A|^{4/(\alpha-1)}$ suffices for us. ∎

The next lemma will be needed to prove (S4).

Lemma 3. Assume that (C.Pol) holds and that $\varphi_{v_-}(x) > 0$. Then there exists a function h: $[0, \infty) \to \mathbb{R}$ such that $\int_0^\infty h(s)\, ds < \infty$ and for A large enough

$$P(T^{+,v_-}_{A,x} > s \cdot \beta^{+,v_-}_{A,x}) \leqslant h(s)$$

Proof. By attractiveness and the Markov property

$$P(\bar{T}_A > s\bar{\beta}_A) \leqslant [P(T^{\delta_0}_A > \bar{\beta}_A)]^{s-1}$$

But by Lemma 2, $P(T^{\delta_0}_A > \bar{\beta}_A) \to e^{-1}$ as $|A| \to \infty$. Therefore it is clear that $h(s)$ can be chosen as b^{s-1}, for some $b \in (e^{-1}, 1)$. ∎

Proof of Theorem 1. To prove (S3), it is enough to verify that for any $t, s > 0$

$$|P(\bar{T}_A > \bar{\beta}_A(t+s)) - P(\bar{T}_A > \bar{\beta}_A t) \cdot P(\bar{T}_A > \bar{\beta}_A s)| \to 0 \tag{3.3}$$

as $|A| \to \infty$. Then it follows by induction that $P(\bar{T}_A > \bar{\beta}_A t) \to e^{-t}$ when t is of the form $p2^{-q}$ for positive integers p, q. By monotonicity and density the same must then be true for any $t > 0$.

By the Markov property

$$P(\bar{T}_A > \bar{\beta}_A(t+s) \mid \bar{T}_A > \bar{\beta}_A \cdot s) = P(T^{\mu}_A > \bar{\beta}_A \cdot t)$$

where μ is the conditional distribution of $\xi^{\nu}_{\bar\beta_\Lambda \cdot s}$ given that $\bar T_\Lambda > \bar\beta_\Lambda \cdot s$. By Lemma 1, $\mu \leqslant \nu_-$. Therefore by Lemma 2

$$|P(T^\mu_\Lambda > \bar\beta_\Lambda t) - P(\bar T_\Lambda > \bar\beta_\Lambda t)| \to 0 \qquad \text{as} \quad |\Lambda| \to \infty$$

from which (3.3) follows immediately, finishing the proof of (S3).

In order to prove (S4), observe that

$$\frac{E\bar T_\Lambda}{\bar\beta_\Lambda} = \frac{1}{\bar\beta_\Lambda} \int_0^\infty P(\bar T_\Lambda > s)\, ds = \int_0^\infty P(\bar T_\Lambda > \bar\beta_\Lambda t)\, dt$$

Using Lemma 3, dominated convergence, and (S3),

$$\lim_\Lambda \frac{E\bar T_\Lambda}{\bar\beta_\Lambda} = \int_0^\infty dt \lim_\Lambda P(\bar T_\Lambda > \bar\beta_\Lambda t) = \int_0^\infty e^{-t}\, dt = 1 \quad \blacksquare$$

Proof of Theorem 2. For simplicity we present the proof only in the case $\delta = 1$. The proof in the more general case is analogous.

Given $\gamma > \varphi(x)$, let l be the integer part of $|\Lambda|^{-2} e^{\gamma|\Lambda|}$. Then by attractiveness and the Markov property

$$\begin{aligned} P(\bar T_\Lambda > e^{\gamma|\Lambda|}) &\leqslant P(\bar\xi_{k|\Lambda|^2} \notin A^+_{\Lambda,x},\ k = 0, 1,..., l) \\ &\leqslant [P(\xi^{\delta_0}_{|\Lambda|^2} \notin A^+_{\Lambda,x})]^l \end{aligned}$$

From (2.3), translation invariance, and (C.Exp.),

$$\begin{aligned} P(\xi^{\delta_0}_{|\Lambda|^2} \notin A^+_{\Lambda,x}) &\leqslant P(\bar\xi_{|\Lambda|^2} \notin A^+_{\Lambda,x}) + C|\Lambda|\, e^{-\alpha|\Lambda|^2} \\ &= \nu_-((A^+_{\Lambda,x})^c) + C|\Lambda|\, e^{-\alpha|\Lambda|^2} \end{aligned}$$

Therefore, using the inequality $\log z \leqslant z - 1$,

$$\begin{aligned} \log P(\bar T_\Lambda > e^{\gamma|\Lambda|}) &\leqslant l \log\{\nu_-[(A^+_{\Lambda,x})^c] + C|\Lambda|^{-\alpha|\Lambda|^2}\} \\ &\leqslant l\{\nu_-[(A^+_{\Lambda,x})^c] + C|\Lambda|\, e^{-\alpha|\Lambda|^2} - 1\} \\ &\leqslant [|\Lambda|^{-2}\, e^{\gamma|\Lambda|} + O(1)][C||\Lambda|\, e^{-\alpha|\Lambda|^2} - \nu_-(A^+_{\Lambda,x})] \end{aligned}$$

Using Theorem 0, part b, it is clear that the rhs of the last expression goes to $-\infty$ as $|\Lambda| \to \infty$. Therefore $\lim_\Lambda P(T^{\nu_-}_\Lambda > e^{\gamma|\Lambda|}) = 0$, from which it follows that

$$\limsup_\Lambda \bar\beta_\Lambda(a) \leqslant \gamma$$

Together with Proposition 1, this implies (S1).

Now (S2) follows from (S4), which was proven in Theorem 1 under the weaker assumption (C.Pol.), provided $\bar{\varphi}(x) > 0$. ▮

We show now how the proof of Theorem 2 can be modified in order to prove Proposition 4.

Proof of Proposition 4. Let γ and l be as in the proof above. Now

$$\begin{aligned} P(T_A > e^{\gamma|A|}) &\leqslant P(\xi^v_{k|A|^2} \notin A^+_{A,x},\ k = 0, 1,..., l) \\ &= P(\xi^v_0 \notin A^+_{A,x})\, P(\xi^v_{|A|^2} \notin A^+_{A,x} \mid \xi^v_0 \notin A^+_{A,x}) \cdots \\ &\quad \times P(\xi^v_{l|A|^2} \notin A^+_{A,x} \mid \xi^v_0 \notin A^+_{A,x},..., \xi^v_{(l-1)|A|^2} \notin A^+_{A,x}) \end{aligned} \tag{3.4}$$

By the finiteness of W, there exists C, $\alpha > 0$ such that for any measure μ

$$P(\xi^v_t(i) \neq \xi^\mu_t(i)) \leqslant Ce^{-\alpha t} \tag{3.5}$$

for an appropriate coupling of (ξ^v_t) and (ξ^μ_t). (Let them evolve independently until they hit. See, for instance, Section 1 of Chapter 2 of Ref. 28). Therefore

$$|P(\xi^v_{|A|^2} \notin A^+_{A,x}) - P(\xi^\mu_{|A|^2} \notin A^+_{A,x})| \leqslant C\,|A|\; e^{-\alpha|A|^2}$$

From this and the Markov property applied to each term on the rhs of (3.4) it follows that

$$P(T_A > e^{\delta|A|}) \leqslant [P(\xi^v_{|A|^2} \notin A^+_{A,x}) + C\,|A|\; e^{-\alpha|A|^2}]^l$$

The proof of (S1) can now be finishes as was done in the proof of Theorem 2 (of course, the analogue of Proposition 1 holds without assuming attractiveness).

We canot prove (S2) as we did in Theorem 2, since we did not prove (S4) in the present case. Using the coupling for which (3.5) holds, the same proof used for Lemma 2 shows that given $\varepsilon > 0$ there exists V large enough such that if $|A| > V$, then for any measure μ

$$P(T^\mu_A > \beta^v_A) \leqslant P(T^v_A > \beta^v_A) + \varepsilon = e^{-1} + \varepsilon \tag{3.6}$$

But for integer K

$$\begin{aligned} P(T^v_A > K\beta^v_A) &= P(T^v_A > \beta^v_A)\, P(T^v_A > 2\beta^v_A \mid T^v_A > \beta^v_A) \cdots \\ &\quad \times P(T^v_A > K\beta^v_A \mid T^v_A > (K-1)\,\beta^v_A) \end{aligned}$$

By the Markov property,

$$P(T^v_A > K\beta^v_A) = \prod_{k=1}^{K} P(T^{\mu^k}_A > \beta^v_A) \tag{3.7}$$

where $\mu^1 = \nu$ and, for $k \geq 2$, μ^k is the conditional distribution of $\xi^{\nu}_{(k-1)\beta^{\nu}_{\Lambda}}$ given that $T^{\nu}_{\Lambda} > (k-1)\beta^{\nu}_{\Lambda}$. From (3.6) and (3.7) it follows that for $|\Lambda|$ large enough

$$P(T^{\nu}_{\Lambda} > t\beta^{\nu}_{\Lambda}) \leq b^{t-1}$$

for some $b \in (e^{-1}, 1)$. Now

$$\begin{aligned} ET^{\nu}_{\Lambda} &= \int_0^{\infty} P(T^{\nu}_{\Lambda} > t)\, dt = \beta^{\nu}_{\Lambda} \int_0^{\infty} P(T^{\nu}_{\Lambda} > \beta^{\nu}_{\Lambda} t)\, dt \\ &\leq \beta^{\nu}_{\Lambda} \int_0^{\infty} b^{t-1}\, dt = C\beta^{\nu}_{\Lambda} \end{aligned}$$

from which point it is easy to finish the proof. ∎

4. PROOFS UNDER GOOD APPROXIMATION BY FINITE SYSTEM

Proof of Theorem 3. Fix an $N \in \{1, 2, 3,...\}$. For each $k \in Z^d$ set

$$\Gamma(k) = \{i \in Z^d: i - kN \in \{1,..., N\}^d\}$$
$$\Gamma(0) = \Gamma = \{1,..., N\}^d$$

We will now define a new spin system (which will be attractive, but not translation-invariant) by the rates

$$c_N(i, j) = c_N(i, \eta^i)$$

where

$$\eta^i = \begin{cases} \eta(j) & \text{if there exists } k \text{ such that } i, j \in \Gamma(k) \\ 0 & \text{otherwise} \end{cases}$$

Let (ζ_t) be the process defined by these rates and $S_N(t)$ the corresponding semigroup. Clearly under this dynamics the spins inside different cubes $\Gamma(k)$ do not interact. Define μ_N as the weak limit of $\delta_0 S_N(t)$ as $t \to \infty$, which exists by attractiveness. Then the restriction of μ_N to each cube $\Gamma(k)$ is an independent copy of ν^N_- [see the definition of this measure before the statement of (C.Pr)].

Set

$$\tau_{\Lambda} = \inf\{t \geq 0: \zeta^{\mu_N}_t \in A^+_{\Lambda,x}\}$$

and define $\delta_\Lambda(a)$, $a \in (0, 1)$, by

$$P(\tau_\Lambda > \delta_\Lambda(a)) = a$$

It follows from Theorem 1.5 of Chapter 3 of Ref. 28 that it is possible to couple $\{\xi_t^\nu)$ and $(\zeta_t^{\mu_N})$ in such a way that the mass is concentrated on paths such that $\zeta_t^{\mu_N} \leqslant \xi_t^\nu$ for every $t \geqslant 0$. Therefore

$$\beta_\Lambda(a) \leqslant \delta_\Lambda(a), \qquad \forall a \in (0, 1)$$

$$ET_\Lambda \leqslant E\tau_\Lambda$$

On the other hand, if the sequence (Λ) is such that each Λ is the union of cubes $\Gamma(k)$, then Proposition 4 implies that

$$\lim_\Lambda |\Lambda|^{-1} \log \delta_\Lambda(a) = \lim_\Lambda |\Lambda|^{-1} \log E\tau_\Lambda = \varphi_N^-(x)$$

[In order to have the hypothesis of Proposition 4 verified consider for each $\Gamma(k)$ a Markov process whose states are only those that can be reached from δ_0.] This finishes the proof in this case. For a general sequence (Λ), let $\Omega_\Lambda = \Omega$ be the largest union of cubes $\Gamma(k)$ contained in Λ. Then $\lim_\Lambda(|\Omega|/|\Lambda|) = 1$ and therefore given $\varepsilon > 0$, $T_{\Lambda,x} \leqslant T_{\Omega,x+\varepsilon}$ for large enough Λ. It is now easy to finish the proof, remembering that $\varphi_N^-(x)$ is convex and therefore continuous on $(0, 1)$. ∎

In order to prove Corollary 2 below, we will need the following lemma. This lemma may be well known, but since we did not find it in the literature, we prove it below.

Lemma 4. Let f_n: $\mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a sequence of convex functions that converges pointwise to a function f. Let

$$g_n(x) = \sup_h \{hx - f_n(h)\}$$

$$g(x) = \sup_h (hx - f(h)\}$$

$$A = \inf\{x \in \mathbb{R}: g(x) < \infty\}$$

$$B = \sup\{x \in \mathbb{R}: g(x) < \infty\}$$

Then $g_n(x) \to g(x)$ for any $x \in (A, B)$.

Proof. Since $x \in (A, B)$, $\lim_{h \to \infty}[hx - f(h)] = -\infty$. Otherwise for $y \in (x, B)$, $\lim_{h \to \infty}[hy - f(h)] = \infty$, which is absurd. Analogously, $\lim_{h \to -\infty}[hx - f(h)] = -\infty$. Therefore, there exists $h_0 \in \mathbb{R}$ such that $g(x) = h_0 x - f(h_0)$.

Now, for fixed $x \in (A, B)$, it is possible to find $a < b$ such that $hx - f(h) \leqslant g(x) - 1$ if $h \notin [a, b]$.

It is known that a sequence of convergent convex functions converges uniformly on compact sets (see Theorem VI.3.3 of Ref. 12). Therefore, given $\varepsilon > 0$, there exists n_0 such that for $n > n_0$

$$|f(h) - f_n((h)| \leqslant \varepsilon/3 \qquad \text{for any } h \in [a, b]$$

Hence, for $n > n_0$, $h_0 x - f_n(h_0)$ is greater than both $ax - f_n(a)$ and $bx - f_n(b)$. By the concavity in h of $hx - f_n(h)$, it follows that

$$\sup_{h \in \mathbb{R}} \{hx - f_n(x)\} = \sup_{h \in [a,b]} \{hx - f_n(x)\}$$

But the distance between the rhs above and $g(x)$ cannot be greater than $\varepsilon/3$, finishing the proof. ∎

Proof of Corollary 2. By Lemma 4, (C.Pr.), and Theorem 0, part c,

$$\varphi_N^-(x) = \sup_h \{hx - \pi_N^-(h)\} \to \sup_h \{hx - \pi(h)\} = \varphi(x)$$

as $N \to \infty$, for any $x \in (0, 1)$.

The sharp upper bound for $\beta_\Lambda^\nu(a)$ and ET_Λ^ν follows now from Theorem 3. To verify the lower bound, we can use Proposition 1 for ν_+, since T_Λ^ν dominates $T_\Lambda^{\nu_+}$ from above. Another way to prove the lower bound is to repeat the arguments of Theorem 3 using + boundary conditions outside each $\Gamma(k)$ and then using Lemma 4 as above. This approach is necessary when one considers the generalization of the present corollary to finite systems with boundary conditions, as discussed in Section 2.3. ∎

Proof of Proposition 2. Define $\theta_\Lambda = \inf\{t \geqslant 0 \colon \xi_t^\nu \in B_\Lambda\}$. Then, by the hypothesis (2.7) and the argument used to prove Proposition 1, there exists $\varepsilon > 0$ such that

$$\lim_\Lambda P(\theta_\Lambda < \exp\{|\Lambda| [\varphi(x) + \varepsilon]\}) = 0 \tag{4.1}$$

But

$$\begin{aligned} P(\xi_{T_\Lambda}^\nu \in B_\Lambda) &\leqslant P(\theta_\Lambda \leqslant T_\Lambda) \\ &\leqslant P(T_\Lambda \geqslant \exp\{|\Lambda| [\varphi(x) + \varepsilon]\}) \\ &\quad + P(\theta_\Lambda < \exp\{|\Lambda| [\varphi(x) + \varepsilon]\}) \end{aligned}$$

which goes to zero as $|\Lambda| \to \infty$, by (S1) and (4.1). ∎

Proof of Proposition 3. We consider a system with finite-range interaction $J(\cdot)$. In the case of infinite-range interactions the inequalities (4.3) below have to be slightly modified; we leave these details to the reader.

Define the sets of configurations

$$C_\Lambda^\delta = \{\eta: 2X_\Lambda(\eta) - 1 \in [m, m+\delta]\}$$
$$= \left\{\eta: |\Lambda|^{-1} \sum_{i\in\Lambda} \sigma(i) \in [m, m+\delta]\right\}$$

For any $\delta > 0$, it is clear that at the time T_Λ the system must be in the set C_Λ^δ if Λ is large enough. Therefore, from Proposition 2, it is enough to show that

$$\limsup_\Lambda |\Lambda|^{-1} \log \nu(B_\Lambda \cap C_\Lambda^\delta) < -\varphi(x) \tag{4.2}$$

But [we abbreviate $h = h(m)$]

$$\begin{aligned}\frac{\nu(B_\Lambda \cap C_\Lambda^\delta)}{\nu(A_{\Lambda,x}^+)} &\leqslant \frac{\nu(B_\Lambda \cap C_\Lambda^\delta)}{\nu(C_\Lambda^\delta)} \\ &\leqslant \frac{\mu_h(B_\Lambda \cap C_\Lambda^\delta)}{\mu_h(C_\Lambda^\delta)} \exp[2\delta|\Lambda| h + o(|\Lambda|)] \\ &\leqslant \frac{\mu_h(B_\Lambda)}{\mu_h(C_\Lambda^\delta)} \exp[2\delta|\Lambda| h + o(|\Lambda|)]\end{aligned} \tag{4.3}$$

where the $o(|\Lambda|)$ appears because of boundary effects, since the field h also modifies the configuration outside Λ. By (2.6) and the definition of $h(m)$

$$\lim_\Lambda |\Lambda|^{-1} \log \mu_h(C_\Lambda^\delta) = 0 \tag{4.4}$$

From (4.3), (4.4), and (2.8) it follows then that

$$\limsup_\Lambda |\Lambda|^{-1} \log \nu(B_\Lambda \cap C_\Lambda^\delta) \leqslant -\varphi(x) - \gamma + 2\delta h$$

for every $\delta > 0$. This implies (4.2).

ACKNOWLEDGMENTS

We are very much indebted to Antonio Galves for having proposed to us the types of questions treated in this paper and to E. Presutti for many illuminating comments and suggestions. We thank also D. Aldous, M. Bramson, R. Durrett, S. Goldstein, S. Olla, and S. R. S. Varadhan for discussions.

R. H. S. would like to thank the warm hospitality of the Rutgers and Cornell Mathematics Departments and of the Mathematical Sciences Institute of Cornell.

This work was supported in part by NSF-DMR81-14726 (U.S.A.), CNPq (Brazil), and the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University.

REFERENCES

1. M. Aizenman and R. Holley, Rapid convergence to equilibrium of stochatic Ising models in the Dobrushin–Shlosman regime, in Proceedings of the IMA Workshop on Percolation and Infinite Particle Systems (February 1986), H. Kesten, ed.
2. D. J. Aldous, *Stochastic Process. Appl.* **13**:305 (1982).
3. D. J. Aldous, Probability approximations via the Poisson clumping heuristics, Lecture notes (unpublished).
4. R. Bellman and T. E. Harris, Recurrence times for the Ehrenfest model, *Pac. J. Math.* **1**:179 (1951).
5. M. Cassandro, A. Galves, E. Olivieri, and M. E. Vares, *J. Stat. Phys.* **35**:603 (1984).
6. R. Cogburn, *Ann. Prob.* **13**:1219 (1985).
7. F. Comets, Grandes déviations pour des champs de Gibbs sur Z^d, *C. R. Acad. Sci.*, to appear.
8. M. V. Day, *Stochastics* **8**:297 (1983).
9. R. Durrett and D. Griffeath, *Ann. Prob.* **8**:890 (1980).
10. R. Durrett and Liu Xiu-fang, The contact process on a finite set, preprint, Cornell (1986).
11. R. Durrett and R. H. Schonmann, The contact process on a finite set II, preprint, Cornell (1987).
12. R. Ellis, *Entropy, Large Deviations and Statistical Mechanics* (Springer, 1985).
13. R. Ellis, Large deviations for the spin per site in ferromagnetic models, preprint, University of Massachusetts (1986).
14. H. Fölmer and S. Orey, Large deviations for the empirical field of a Gibs measure, preprint (1986).
15. M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer, 1984).
16. A. Galves, E. Olivieri, and M. E. Vares, Metastability for a class of dynamical systems subject to small random perturbations, preprint, IHES (1984).
17. L. Gray, Duality for general attractive spin systems with applications in one dimension, *Ann. Prob.* **14**:371 (1986).
18. D. Griffeath, *Additive and Cancelative Interacting Particle Systems* (Springer, 1979).
19. R. Griffiths, C.-Y. Weng, and J. S. Langer, *Phys. Rev.* **149**:301 (1966).
20. T. E. Harris, *Trans. Am. Math. Soc.* **73**:471 (1952).
21. R. Holley, *Ann. Prob.* **13**:72 (1985).
22. R. Holley, *Contemp. Math.* **41**:215 (1985).
23. J. Keilson, *Markov Chain Models—Rarity and Exponentiality* (Springer, 1979).
24. C. Kipnis and C. Newman, *SIAM J. Appl. Math.* **45**:972 (1985).
25. D. V. Korolyuv and D. S. Sil'vestrov, *Theory Prob. Appl.* **29**:432 (1984).

26. O. E. Lanford, in *Lecture Notes in Physics*, Vol. 20 (Springer, 1973), p. 1–113.
27. J. L. Lebowitz and R. H. Schonmann, Pseudo free energies and large deviations for non-Gibbsian FKG measures, preprint Rutgers (1986).
28. T. M. Liggett, *Interacting Particle Systems* (Springer, 1985).
29. S. Olla, Large deviations for Gibbs random fields, *Prob. Theory Related Fields*, to appear.
30. O. Penrose and J. L. Lebowitz, in *Fluctuation Phenomena*, E. W. Montroll and J. L. Lebowitz, eds. (North-Holland, 1979).
31. D. Ruelle, *Statistical Mechanics Rigorous Results* (Benjamin, 1969).
32. R. H. Schonmann, *J. Stat. Phys.* **41**:445 (1985).
33. Ya. G. Sinai, *Theory of Phase Transitions* (Pergamon Press, 1982).
34. M. Williams, *SIAM J. Appl. Math.* **42**:149 (1982).