Large deviations for empirical processes of interacting particle systems

2001 J. Phys. A: Math. Gen. 34 1603

(http://iopscience.iop.org/0305-4470/34/8/307)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **34** (2001) 1603–1610 www.iop.org/Journals/ja PII: S0305-4470(01)17331-2

Large deviations for empirical processes of interacting particle systems

Jinwen Chen

Department of Mathematical Sciences, Tsinghua University Beijing 100084, People's Republic of China

Received 19 September 2000, in final form 12 January 2001

Abstract

In this paper we study process-level large deviations for spin-flip particle systems and more general Markov processes. We prove that a modification of a well known Donsker–Varadhan entropy function can be used to govern the large deviation lower bounds for any such process starting from a class of initial distributions related to an extremal invariant measure of the process. In some specific cases, we obtain the full large deviation principle.

PACS numbers: 0250G, 0250F, 0550, 4550

1. Introduction

By interacting particle systems (IPS) we refer to models which describe systems consisting of a large number of individuals, such as microscopic particles, that stochastically interact with each other. For such a system, one of the most important problems is to study its long-time behaviour, including its asymptotic stability and fluctuation around its stable states. Such a stable state is physically termed an equilibrium, and mathematically described as an invariant probability measure, or a Markov measure, of a stochastic process. Thus, in a certain sense, the asymptotic stability can be described by convergence to such a measure of the system. More precisely, if we denote by ν such an invariant measure, P_ν the corresponding Markov measure, L_t and R_t the empirical measure and the process of the system up to time t respectively (all to be formally defined later), then starting from certain initial configurations, $L_t \Rightarrow \nu$ and $R_t \Rightarrow P_\nu$ respectively, as $t \to \infty$, where $\Rightarrow$ denotes the weak convergence of probability measures. That is, if η is such an initial configuration, P_η is the corresponding distribution law of the system, then informally we have

$$P_\eta(L_t \sim \nu) \sim 1 \qquad \text{and} \qquad P_\eta(R_t \sim P_\nu) \sim 1$$

where by '$\sim$' we mean 'close'. Hence if G and O are neighbourhoods of ν and P_ν respectively, then the probabilities for large fluctuation $P_\eta(L_t \notin G)$ and $P_\eta(R_t \notin O)$ go to 0 as $t \to \infty$. Now, for what η will these statements hold and how fast will the fluctuation probabilities go to 0? There are various reasons for considering these problems from different points of view. For example, from the point of view of physics, or from the point of view of ergodic theory, for each equilibrium state, one would like to have knowledge about its domain of attraction, and the metastability of the system before it becomes close to this state. Statistically, one also needs to know starting from what initial points the empirical quantities will converge to an unknown invariant measure and to know the convergence rate, so as to make an estimation of this unknown measure and give the error probability as explicitly as possible. The large deviation (LD) technique is important for studying these problems. Notably, a full large deviation principle (LDP) will give rise to a series of further interesting consequences, including the variational representation of Laplacian integrals, that is,

$$\lim_{t\to\infty}\frac{1}{t}\log\int\exp\{t\Phi(L_t)\}\,\mathrm{d}P_\eta=\sup\left[\int\Phi\,\mathrm{d}\mu-I(\mu)\right]$$

and

$$\lim_{t\to\infty}\frac{1}{t}\log\int\exp\{t\Psi(R_t)\}\,\mathrm{d}P_\eta=\sup\left[\int\Psi\,\mathrm{d}Q-H(Q)\right]$$

where I and H are the LD rate functions for L_t and R_t respectively, and the two sup's are taken over the space of probability measures on the space of configurations and on the path space of the system respectively. Obviously, these are important both in mathematics and physics, since the pressure functional of a particle system can be represented as certain Laplacian integrals.

In this paper we consider the LD problem for a class of well studied IPS–spin-flip systems. In [1–5] LDs for space–time empirical processes of spin-flip particle systems were studied. The corresponding rate functions H_0 are shown to satisfy a certain variational principle, i.e., $H_0(Q) = 0$ iff Q is a stationary Markov measure corresponding to the spin system under consideration. Applications of the results to ergodicity were also discussed. From the point of view of ergodic behaviour, a natural and reasonable question is: does one have a LDP for the usual time empirical processes? Presently there are few results in this direction. The study turns out to be very complicated due to the infiniteness of the number of particles and their interactions. In the space–time cases, a spin system has a certain local space–time Gibbsian structure, which reduces the LD study to that for a non-interacting system and the control of a certain type of Girsanov transformation. This is one of the standard techniques used in studying the LDP. In the cases of time empirical processes, we do not know how to use this technique. A powerful technique for studying the LD of time empirical processes was introduced by Donsker and Varadhan (DV) (cf [6]), which provides upper bounds for very general Markov processes. As for lower bounds, some conditions on the transition probabilities were imposed. Jain [7] also provided some results for the lower bounds. Some essential conditions needed in [6] or [7] were certain irreducibility or strong ergodicity of the processes; which is reasonable. However, it is generally difficult to verify any of these conditions for an interacting infinite particle system. Nevertheless, their techniques can be modified to study LDs for systems starting 'near' extremal invariant measures of the systems. The content of this paper is as follows. We will obtain LD lower bounds governed by a modified DV entropy function. In some specific cases, this modified entropy function coincides with the exact DV one, so we obtain a full LDP. Since our approach applies to more general Markov processes, we give our results in a more general setting and take spin-flip systems as examples.

We first introduce some general notations. Let E be a Polish space, $M_1(E)$ be the space of all probability measures on E, equipped with the weak topology. $\Omega = \mathrm{D}(R, E)$ and $\Omega_+ = \mathrm{D}([0,\infty), E)$ denote the spaces of cadlag functions from R and $[0,\infty)$ to E, respectively, both equipped with the Skorohod topology and the Borel σ-algebra. Let $\{P_x, x \in E\}$ be a Markov family of probability measures of a time-homogeneous Markov process on Ω_+, with P_x being weakly continuous in x. $\{S(t), t \geqslant 0\}$ denotes the semigroup of the process, $M_{\mathrm{s}}(E)$ the set of invariant probability measures of the semigroup, and $M_{\mathrm{s}}^{\mathrm{e}}(E)$ the set of the extremal elements in $M_{\mathrm{s}}(E)$. For $\mu \in M_1(E)$, $P_\mu = \int P_x\mu(\mathrm{d}x)$. For $-\infty \leqslant s \leqslant t \leqslant \infty$, denote

$\mathcal{F}_s^t = \sigma\{\omega_u, s \leqslant u \leqslant t\}$, $\mathcal{F}_t \equiv \mathcal{F}_0^t$. $M_s(\Omega_+)$ is the set of stationary probability measures on Ω_+, equipped with the weak topology. Now we define the empirical processes. For $\omega \in \Omega_+$ and $t > 0$, let ω^t be the t-periodic element of ω defined by $(\omega^t)_{s+nt} = \omega_s$ for $0 \leqslant s < t$ and $n \in Z_+$. Then define

$$R_t = R_t(\omega) = \frac{1}{t}\int_0^t \delta_{\theta_u \omega^t}\,\mathrm{d}u$$

where θ_u is the shift operator on Ω_+ given by $(\theta_u\omega)_t = \omega_{t+u}$, δ_ω is the usual Dirac measure centred at ω. Then $R_t \in M_s(\Omega_+)$. We want to study LD for $\{P_\mu(R_t \in \cdot),\ t > 0\}$. In [6] one can find the following uniform upper bounds for general Markov processes:

$$\limsup_{t\to\infty} \frac{1}{t} \log \sup_{x\in E} P_x(R_t \in F) \leqslant - \inf_{Q\in F} H(Q) \qquad \text{for compact} \quad F \tag{1.1}$$

where H is the DV entropy function defined by $H(Q) = E^Q \log(\frac{\mathrm{d}Q_\omega}{\mathrm{d}P_{\omega_0}}|_{\mathcal{F}_1})$, and $Q_\omega = Q(\cdot/\mathcal{F}_{-\infty}^0)(\omega)$. It has been proved that H possesses good properties: H is affine with compact level sets, i.e., for $Q_i \in M_s(E)$ and $\lambda_i \geqslant 0$, $i = 1, \ldots, k$, with $\sum_{i=1}^k \lambda_i = 1$, $H(\sum_{i=1}^k \lambda_i Q_i) = \sum_{i=1}^k \lambda_i H(Q_i)$, and for any $a \geqslant 0$, $\{H(Q) \leqslant a\}$ is compact in $M_s(\Omega_+)$. Furthermore, under certain exponential tightness, (1.1) holds for every closed F. In particular, if E is compact, then H governs the LD upper bounds. In this paper, we will study the associate LD lower bounds, i.e., for open $G \subset M_s(\Omega_+)$,

$$\liminf_{t\to\infty} \frac{1}{t} \log P_\mu(R_t \in G) \geqslant - \inf_{Q\in G} H(Q). \tag{1.2}$$

In [6, 7] one finds some conditions for (1.2) to hold. But for an interacting infinite particle system, we do not know how to check when these conditions are satisfied. Consequently, we will modify the entropy function H and restrict the initial distributions to an important and interesting class which is intimately related to $M_s^e(E)$. For $Q \in M_s(\Omega_+)$, let μ_Q be its single time marginal. Given $\nu \in M_s(E)$, let $M_\nu^\perp(E)$ be the set of measures in $M_s(E)$ that are singular w.r.t. ν. Then define

$$H_{\nu,0}(Q) = \begin{cases} +\infty & \text{if} \quad \mu_Q \in M_\nu^\perp(E) \\ H(Q) & \text{otherwise} \end{cases}$$

and then define its lower semicontinuous version by $H_\nu(Q) = \lim_{\delta\to 0} \inf_{d(Q,Q')<\delta} H_{\nu,0}(Q')$, where $d(\cdot,\cdot)$ is any metric on $M_s(\Omega_+)$ which generates the weak topology. The main result of this paper is the following:

Theorem 1.1. *Under the above notations,*

(1) For each $\nu \in M_s(E)$, H_ν has compact level sets.

(2) If ν is in $M_s^e(E)$, then for ν almost all x, for each open $G \subset M_s(\Omega_+)$,

$$\liminf_{t\to\infty} \frac{1}{t} \log P_x(R_t \in G) \geqslant - \inf_{Q\in G} H_\nu(Q). \tag{1.3}$$

(3) Let $\nu \in M_s^e(E)$. For $\mu \in M_s(E)$, if there is a $T \geqslant 0$ such that $\mu_T \equiv \mu S(T) \notin M_\nu^\perp(E)$, then (1.3) holds with P_x replaced by P_μ, where the measure $\mu S(T) \in M_1(E)$ is defined by $\int f\,\mathrm{d}\mu\,S(T) = \int S(T) f\,\mathrm{d}\mu$ for $f \in C_b(E)$. In particular, (1.3) holds with P_x replaced by P_ν.

Remark 1. From the definition of H_ν we know that if $H(Q) < \infty$ implies $\mu_Q \notin M_\nu^\perp(E)$, then $H_\nu = H$, and hence (1.1) and theorem 1.1 give the full LDP.

The theorem will be proved in section 2. In section 3, we make some further discussion, including applications of our theorem to particle systems and comparison of our results with that obtained in [4, 5].

2. Proof of theorem 1.1

The proof of conclusion (1) is simple, we only need to use the definition of H_ν and the compactness of the level sets of H.

Now we prove conclusion (2). In addition, by the definition of H_ν we see that we only need to prove that for $\nu \in M_s^e(E)$, if $Q \in M_s(\Omega_+)$ satisfies $H(Q) = H_{\nu,0}(Q) < \infty$ and G is an open set containing Q, then for ν almost all x,

$$\liminf_{t\to\infty} \frac{1}{t} \log P_x(R_t \in G) \geqslant -H(Q). \tag{2.1}$$

The main ingredient of its proof is contained in several lemmas. To prove them, we need the following lemma which was proved in [6].

Lemma 2.0. *Define* $\phi(t,\omega) = \log(\frac{dQ_\omega}{dP_{\omega_0}}|_{\mathcal{F}_t})$ *and* $H_t(Q) = E^Q\phi(t,\omega)$*. If* $H(Q) < \infty$*, then* $H_t(Q) = tH(Q)$ *for* $t > 0$ *and* $H(Q) = \lim_{t\to\infty} \frac{\phi(t,\omega)}{t}$ Q — *a.s.*

The following lemma is a consequence of the well known ergodic theorem.

Lemma 2.1. *If* $\nu \in M_s^e(E)$ *and* $\mu \in M_1(E)$ *with* $\mu \notin M_\nu^\perp(E)$*, then for every bounded measurable function* f *on* E*, there is a constant* $c > 0$ *such that*

$$\liminf_{t\to\infty} \frac{1}{t}\int_0^t E^{P_\mu} f(\omega_u)\, du \geqslant c\nu(f).$$

In particular, if $\mu \ll \nu$*, then*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(\omega_u)\, du = \nu(f) \quad P_\mu \quad \text{— a.s.}$$

Proof. Since $\nu \in M_s^e(E)$, it follows that P_ν is ergodic, see, e.g., [9, theorem B52]. Hence for bounded f on E, there is $E_0 \subset E$ with $\nu(E_0) = 1$, such that $\forall x \in E_0$,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(\omega_u)\, du = \nu(f) \quad P_x \quad \text{— a.s.}$$

Note that if $\mu \ll \nu$, then $\mu(E_0) = 1$, and we obtain the second conclusion of the lemma. For the first one we have

$$\begin{aligned}\liminf_{t\to\infty} \frac{1}{t}\int_0^t E^{P_\mu} f(\omega_u)\, du &\geqslant \liminf_{t\to\infty} \int_{E_0} \mu(dx) E^{P_x} \frac{1}{t}\int_0^t f(\omega_u)\, du \\ &\geqslant \int_{E_0} \mu(dx) \liminf_{t\to\infty} E^{P_x} \frac{1}{t}\int_0^t f(\omega_u)\, du \\ &= \mu(E_0)\nu(f)\end{aligned}$$

If $\mu \notin M_\nu^\perp(E)$, then $\nu(E_0) = 1$ implies $\mu(E_0) > 0$, the desired conclusion follows. □

Lemma 2.2. *Let* $\nu \in M_s^e(E)$*,* $Q \in M_s(\Omega_+)$ *be ergodic with* $H(Q) < \infty$ *and* $\mu_Q \notin M_\nu^\perp(E)$*,* G *an open set containing* Q*,* A *measurable on* E *with* $\nu(A) > 0$*. Then for fixed* $\epsilon > 0$ *and* $\sigma > 0$*, there exists* $T > 0$ *such that* $\forall \delta > 0$ *there exists measurable* A_0 *with* $\mu_Q(A_0) > 1 - \delta$*, such that*

$$\liminf_{t\to\infty} \frac{1}{t} \log \inf_{x\in A_0} \frac{1}{\sigma t}\int_0^{\sigma t} du\, \frac{1}{T}\int_0^T dv\, P_x(R_t \in G, \omega_{t+u+v} \in A) \geqslant -H(Q) - \epsilon.$$

Proof. For $\omega \in \Omega_+$, by the Markov property we have

$$\begin{aligned}\Delta(t, T, \omega_0) &\equiv \frac{1}{\sigma t}\int_0^{\sigma t} \mathrm{d}u \, \frac{1}{T}\int_0^T \mathrm{d}v \, P_{\omega_0}(R_t \in G, \ \omega'_{t+u+v} \in A) \\ &= \int_{R_t \in G} \frac{1}{\sigma t}\int_0^{\sigma t}\left[\frac{1}{T}\int_0^T P_{\omega'_{t+u}}(\omega''_v \in A)\,\mathrm{d}v\right]\mathrm{d}u\,\mathrm{d}P_{\omega_0} \\ &= \int_{R_t \in G} \frac{1}{\sigma t}\int_t^{t+\sigma t} g_T(\omega'_u)\,\mathrm{d}u\,\mathrm{d}P_{\omega_0} \\ &\geqslant \frac{c\nu(A)}{2}\exp\left\{-t\left[H(Q)+\frac{\epsilon}{2}\right]\right\} Q_\omega(\Gamma_t) \end{aligned} \tag{2.2}$$

where $g_T(\eta) \equiv \frac{1}{T}\int_0^T P_\eta(\omega_v \in A)\,\mathrm{d}v$ and

$$\Gamma_t \equiv \left\{R_t \in G, \log\left(\left.\frac{\mathrm{d}Q_\omega}{\mathrm{d}P_{\omega_0}}\right|_{\mathcal{F}_t}\right) \leqslant t\left[H(Q)+\frac{\epsilon}{2}\right], \frac{1}{\sigma t}\int_t^{t+\sigma t} g_T(\omega_u)\,\mathrm{d}u \geqslant \frac{c\nu(A)}{2}\right\}.$$

By the ergodicity of Q we know that

$$\lim_{t\to\infty}\frac{1}{\sigma t}\int_0^{\sigma t} g_T(\omega_u)\,\mathrm{d}u = \frac{1}{T}\int_0^T P_{\mu_Q}(\omega_v \in A)\,\mathrm{d}v \quad Q \quad \text{—} \quad \text{a.s.}$$

Furthermore, since $\nu \in M_s^e(E)$ and $\mu_Q \notin M_\nu^\perp(E)$, from lemma 2.1 we see that for some $c > 0$, if T is sufficiently large, then

$$\frac{1}{T}\int_0^T P_{\mu_Q}(\omega_v \in A)\,\mathrm{d}v > \frac{c\nu(A)}{2}.$$

Combining this with (2.2), the stationarity and ergodicity of Q and lemma 2.0 we see that $\lim_{t\to\infty} Q_\omega(\Gamma_t) = 1$ Q—a.s. and hence for μ_Q almost all x:

$$\liminf_{t\to\infty}\frac{1}{t}\log\Delta(t, T, x) \geqslant -H(Q) - \frac{\epsilon}{2}.$$

The desired inequality easily follows from this. □

Lemma 2.3. *Let $\nu \in M_s^e(E), Q = \sum_{i=1}^k \lambda_i Q_i$ with Q_i ergodic, $H(Q_i) < \infty$, $\mu_i \equiv \mu_{Q_i} \notin M_\nu^\perp(E)$, $\lambda_i > 0$ and $\sum_{i=1}^k \lambda_i = 1$, G be an open set containing Q. Then for ν almost all x, (2.1) holds.*

Proof. For simplicity, we only consider the case $k = 3$, the general cases can be treated in the same way. Fix $\epsilon > 0$. Choose the neighbourhood G_i of $Q_i (1 \leqslant i \leqslant 3)$ and $t_0 > 0$ such that if $Q'_i \in G_i$ and $||Q''_i - Q'_i||_{\mathcal{F}_{t_0}}$ is sufficiently small for $1 \leqslant i \leqslant 3$, then $\sum_{i=1}^3 \lambda_i Q''_i \in G$, where $||\cdot||_{\mathcal{F}_{t_0}}$ denotes the total variation norm of probability measures restricted on $\mathcal{F}_{t_0}$. Write $t_i = \lambda_i t$. Then we can choose sufficiently small $\sigma > 0$, such that for sufficiently large t and measurable A_i with $\mu_i(A_i) \equiv \mu_{Q_i}(A_i) > 0$ $(1 \leqslant i \leqslant 3)$ (to be chosen later), we have for $x \in E$

$$\begin{aligned} P_x(R_t \in G) \geqslant{}& \frac{1}{\sigma t_1}\int_0^{\sigma t_1}\mathrm{d}u\,\frac{1}{\sigma t_1}\int_0^{\sigma t_1}\mathrm{d}u_1\,\frac{1}{T}\int_0^T \mathrm{d}v_1\,\frac{1}{\sigma t_2}\int_0^{\sigma t_2}\mathrm{d}u_2 \\ &\times\frac{1}{T}\int_0^T \mathrm{d}v_2\,P_x\big(R_{t_1}(\theta_u\omega) \in G_1, \omega_{u+t_1+u_1+v_1} \in A_2, R_{t_2}(\theta_{u+t_1+u_1+v_1}\omega) \in G_2, \\ &\omega_{u+t_1+u_1+v_1+t_2+u_2+v_2} \in A_3, R_{t_3}(\theta_{u+t_1+u_1+v_1+t_2+u_2+v_2}\omega) \in G_3\big) \end{aligned}$$

$$
\begin{aligned}
&\geqslant [\inf_{r\in A_3} P_r(R_{t_3}\in G_3)]\frac{1}{\sigma t_1}\int_0^{\sigma t_1}\mathrm{d}u\,\frac{1}{\sigma t_1}\int_0^{\sigma t_1}\mathrm{d}u_1\,\frac{1}{T}\int_0^{T}\mathrm{d}v_1\,\frac{1}{\sigma t_2}\int_0^{\sigma t_2}\mathrm{d}u_2\\
&\quad\times\frac{1}{T}\int_0^{T}\mathrm{d}v_2\,P_x\big(R_{t_1}(\theta_u\omega)\in G_1,\,\omega_{u+t_1+u_1+v_1}\in A_2,\\
&\quad R_{t_2}(\theta_{u+t_1+u_1+v_1}\omega)\in G_2,\,\omega_{u+t_1+u_1+v_1+t_2+u_2+v_2}\in A_3\big)\\
&\geqslant \Gamma_{1,t}\left[\inf_{y\in A_2}\frac{1}{\sigma t_2}\int_0^{\sigma t_2}\mathrm{d}u_2\,\frac{1}{T}\int_0^{T}\mathrm{d}v_2\,P_y(R_{t_2}\in G_2,\,\omega_{t_2+u_2+v_2}\in A_3)\right]\frac{1}{\sigma t_1}\\
&\quad\times\int_0^{\sigma t_1}\mathrm{d}u\,\frac{1}{\sigma t_1}\int_0^{\sigma t_1}\mathrm{d}u_1\,\frac{1}{T}\int_0^{T}\mathrm{d}v_1\,P_x(R_{t_1}(\theta_u\omega)\in G_1,\,\omega_{u+t_1+u_1+v_1}\in A_2)\\
&\geqslant \Gamma_{1,t}\Gamma_{2,t}\left[\inf_{z\in A_1}\frac{1}{\sigma t_1}\int_0^{\sigma t_1}\mathrm{d}u_1\,\frac{1}{T}\int_0^{T}\mathrm{d}v_1\,P_z(R_{t_1}\in G_1,\,\omega_{t_1+u_1+v_1}\in A_2)\right]\\
&\quad\times\frac{1}{\sigma t_1}\int_0^{\sigma t_1}P_x(\omega_u\in A_1)\,\mathrm{d}u
\end{aligned}
\tag{2.3}
$$

where

$$
\Gamma_{1,t}=\inf_{r\in A_3}P_r(R_{t_3}\in G_3)
$$

$$
\Gamma_{2,t}=\inf_{y\in A_2}\frac{1}{\sigma t_2}\int_0^{\sigma t_2}\mathrm{d}u_2\,\frac{1}{T}\int_0^{T}\mathrm{d}v_2\,P_y(R_{t_2}\in G_2,\,\omega_{t_2+u_2+v_2}\in A_3)
$$

and denote

$$
\Gamma_{3,t}=\inf_{z\in A_1}\frac{1}{\sigma t_1}\int_0^{\sigma t_1}\mathrm{d}u_1\,\frac{1}{T}\int_0^{T}\mathrm{d}v_1\,P_z(R_{t_1}\in G_1,\,\omega_{t_1+u_1+v_1}\in A_2).
$$

By lemma 2.2 and the assumption that $\mu_3\notin M_\nu^{\perp}(E)$, we first choose A_3 with $\mu_3(A_3)$ close to 1, such that $\nu(A_3)>0$ and

$$
\liminf_{t\to\infty}\frac{1}{t_3}\log\Gamma_{1,t}\geqslant -H(Q_3)-\epsilon.
$$

Then we can apply lemma 2.2 again to choose A_2 with $\mu_2(A_2)$ close to 1, such that $\nu(A_2)>0$ and

$$
\liminf_{t\to\infty}\frac{1}{t_2}\log\Gamma_{2,t}\geqslant -H(Q_2)-\epsilon.
$$

In the same way we can choose A_1 with $\nu(A_1)>0$, such that

$$
\liminf_{t\to\infty}\frac{1}{t_1}\log\Gamma_{3,t}\geqslant -H(Q_1)-\epsilon.
$$

Furthermore, since $\nu(A_1)>0$, from lemma 2.1 we know that for ν almost all x,

$$
\frac{1}{\sigma t_1}\int_0^{\sigma t_1}P_x(\omega_u\in A_1)\,\mathrm{d}u>\frac{\nu(A_1)}{2}
$$

for sufficiently large t. Combining these with (2.3) we obtain (2.1). □

Now a standard argument shows that (2.1) holds for general Q with $H_{\nu,0}(Q)=H(Q)<\infty$, completing the proof of conclusion (2) of theorem 1.1.

To prove conclusion (3), we note that for $\mu_T=\mu S(T)$, we can find neighbourhood $G_1\subset G$ of Q such that for sufficiently large t,

$$
P_\mu(R_t\in G)\geqslant P_\mu(R_{t-T}(\theta_T\omega)\in G_1)=P_{\mu_T}(R_{t-T}\in G_1).
$$

Consequently, without loss of generality we can assume $\mu \notin M_\nu^\perp(E)$. Then the proof of (3) is the same as that of (2), only in the last step we replace P_x by P_μ and apply lemma 2.1 to obtain that for some $c > 0$ and every large t

$$\frac{1}{\sigma t_1}\int_0^{\sigma t_1} P_\mu(\omega_u \in A_1)\,\mathrm{d}u \geqslant c\nu(A_1).$$

3. Further discussion

In this section we first briefly discuss the application of theorem 1.1 to spin-flip systems. By such a system we refer to a Feller–Markov process $\{P_x, x \in E\}$ on $\Omega_+ = D([0,\infty), E)$ with $E = \{0,1\}^{Z^d}$, determined by a family of spin-flip rates $\{c(i,\cdot), i \in Z^d\}$, where for each i, $c(i,\cdot)$ is a non-negative continuous function on E satisfying certain conditions (cf [8, chapter 3]). As we have already stated, to obtain a full LDP, a sufficient condition is that $H(Q) < \infty$ implies $\mu_Q \notin M_\nu^\perp(E)$. In some cases where there is certain strong ergodicity, we indeed have a stronger implication, i.e., $H(Q) < \infty$ implies $\mu_Q \ll \nu$, where ν is the unique invariant probability measure of the system. Hence $H = H_\nu$ and we have a full LDP under P_ν or under P_x for ν almost all x. For a simple example in this direction, we consider a spin-flip system $\{P_x, x \in E_0\}$ with strictly positive interactions, where $E_0 \equiv \{x \in E, \sum_{i\in Z^d} x(i) < \infty\}$. More precisely, we assume that for each $i \in Z^d$, $c(i,\eta) > 0$ if $\eta \in E_0$; $= 0$ otherwise. Then $\{P_x, x \in E_0\}$ is in fact a continuous-time irreducible Markov chain on E_0. If this chain is ergodic, then every probability measure on E is absolutely continuous w.r.t. the unique invariant probability measure ν of the system.

An example of infinite system, along these lines, is the case when $c(i,\cdot) \equiv 1\ \forall i \in Z^d$. This is a non-interacting system. If we denote by ν the product probability measure on E with marginal density $1/2$, then it can be shown that $H(Q) < \infty$, hence $I(\mu_Q) < \infty$, implies $\mu_Q \ll \nu$, where for $\mu \in M_1(E)$, $I(\mu) \equiv \inf\{H(Q), \mu_Q = \mu\}$.

A non-trivial example of interacting infinite system is the one-dimensional stochastic Ising model with finite range potentials. It can be shown that if ν is the unique Gibbs state, then $H(Q) < \infty$ implies $\mu_Q \ll \nu$. For this model, it has been known that under $P\nu$, the occupation times of the system satisfies a full LDP. Our theorem 1.1 means that the empirical processes, and hence the empirical measures, satisfies a full LDP. This may be the first non-trivial LDP result for empirical processes of spin-flip systems.

As one can see, theorem 1.1 should be more applicable for ergodic systems. For example, if a system has a unique invariant probability measure ν and for every initial measure μ, $\lim_{t\to\infty} \mu S(t) = \nu$ in the τ-topology, then a slight modification of the proof of theorem 1.1 shows that (2.3) holds.

Now we make a comparison of our time empirical LD process with the space–time empirical LD process studied in [4, 5]. There are at least three negligible differences we see between the two: (a) the space–time empirical processes are both time stationary and space translation invariant. Thus the corresponding LD rate functions were used to characterize the Markov measures of the system that are space translation invariant, whereas the time empirical LD process can be used to describe all Markov measures of the system. (b) The respective LD rate functions may have quite different features. For example, as we discussed at the beginning of this section, for some strong ergodic system, $H(Q) < \infty$ implies $\mu_Q \ll \nu$. However, even for the same system, one may have $H_0(Q) < \infty$ but μ_Q is singular w.r.t. ν, where H_0 is the space–time LD rate function. The non-interacting system we discussed in the second paragraph of this section is such an example. (c) As we stated in section 1, the technique used in [4, 5] heavily relies on the space–time local Gibbsian structure of the system and requires the spin-flip rates to be strictly positive. Our results may apply to systems with possibly vanishing spin-flip rates: what we need to do is to check that $H(Q) < \infty$ implies that μ_Q is not singular w.r.t. ν, and we think this should not be too restrictive.

Acknowledgment

This work is supported by the NSF of China. The author would like to thank a referee for many helpful comments and suggestions that led to an improvement in the presentation.

References

[1] Chen J W 1998 Large deviations and ergodicity for spin particle systems *J. Stat. Phys.* **91** 369–93
[2] Chen J W 1999 A variational principle for Markov processes *J. Stat. Phys.* **96** 1359–64
[3] Chen J W 2000 Large deviations and occupation times for spin particle systems with long range interactions at press
[4] Dai Pra P 1993 Space–time large deviations for interacting particle systems *Commun. Pure Appl. Math.* 387–422
[5] Dai Pra P 1993 Large deviation and stationary measures for interacting particle systems *Stoch. Process. Appl.* **48** 9–30
[6] Donsker M D and Varadhan S R S 1983 Asymptotic evaluation of certain Markov process expectations for large timeIV *Commun. Pure Appl. Math.* **36** 183–212
[7] Jain N 1990 Large deviation lower bounds for additive functionals of Markov processes *Ann. Probab.* **18** 1071–98
[8] Liggett T M 1985 *Interacting Particle systems* (New York: Springer)
[9] Liggett T M 1999 *Stochstic Interacting Systems: Contact, Voter, and Exclusion Processes* (Berlin: Springer)