The Annals of Probability
1994, Vol. 22, No. 4, 2122–2151

LARGE DEVIATIONS FOR EMPIRICAL PROCESS OF MEAN-FIELD INTERACTING PARTICLE SYSTEM WITH UNBOUNDED JUMPS[1]

By Shui Feng
McMaster University

A large deviation system is established for the empirical processes of a mean-field interacting particle system with unbounded jump rates under assumptions that are satisfied by many interesting models including the first and the second Schlögl models. The action functional obtained has a form that is very useful for applications.

1. Introduction. In this paper we consider a finite particle system interacting via the mean of the system. Such a system is usually called a particle system with mean-field interaction and can be described in the following way.

Let $E = \{0, 1, \dots\}$, and let $E^{\otimes N}$ be the N-fold Cartesian product of E for $N \geq 1$. $M_1(E)$ denotes the set of all probability measures on E. $C_b(E)$ and $C_b(E^{\otimes N})$ will denote the sets of bounded continuous functions on E and $E^{\otimes N}$, respectively. For any $u \in M_1(E)$, $f \in C_b(E)$, we introduce an operator Q_u:

$$Q_u f(x) = \sum_{y \in E} q_{x,y}\big(f(y) - f(x)\big) + \|u\|\big(f(x+1) - f(x)\big), \tag{1.1}$$

where $(q_{x,y})_{x,y \in E}$ is the jump matrix and $\|u\|$ denotes the first moment of u.

Let δ_x denote the Dirac measure at x. Then the N-particle system mentioned previously is a Markov process $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_N^{(N)}(t))$ on $E^{\otimes N}$ with generator

$$\Omega^{(N)}\psi\big(x^{(N)}\big) = \sum_{k=1}^{N} Q^{(k)}_{\varepsilon_{x^{(N)}}}\psi\big(x^{(N)}\big), \qquad \psi \in C_b\big(E^{\otimes N}\big), \tag{1.2}$$

where $x^{(N)} = (x_1^{(N)}, \dots, x_N^{(N)}) \in E^{\otimes N}$, $\varepsilon_{x^{(N)}} = (1/N)\Sigma_{k=1}^{N} \delta_{x_k^{(N)}}$ and $Q_u^{(k)}$ is used instead of Q_u when it acts on the kth variable of ψ.

For a given process $x^{(N)}(t) = (x_1^{(N)}(t), \dots, x_N^{(N)}(t))$ with generator (1.2), let

$$\varepsilon_{x^{(N)}(t)} = \frac{1}{N}\sum_{k=1}^{N} \delta_{x_k^{(N)}(t)}, \qquad t \geq 0, \tag{1.3}$$

denote its empirical process. It is known that for any finite time interval, the empirical process satisfies a law of large numbers as $N \to \infty$ [4] and that the

Received August 1992; revised December 1993.

[1]Partially Supported by a scholarship from the Faculty of Graduate Studies and Research of Carleton University and the NSERC operating grant of D. A. Dawson.

AMS 1991 *subject classifications.* Primary 60F10, 60J75, 60K35; secondary 82C26.

Key words and phrases. Q-process, pure jump Markov process, nonlinear master equation, Schlögl model, mean-field interacting particle system, empirical process, large deviation system.

limiting deterministic dynamics are characterized as a weak solution of the nonlinear master equation

$$\frac{d\langle u(t),f\rangle}{dt} = \langle u(t), Q_{u(t)}f\rangle, \qquad f \in C_k(E), \tag{1.4}$$

where $C_k(E)$ denotes the set of continuous functions on E with compact support.

The nonlinear master equation was proposed by Nicolis and Prigogine [13] as a mean-field model of a chemical reaction with spatial diffusion. Feng and Zheng [8] established the existence and uniqueness of the solution to the nonlinear master equation and proved the existence of at least three equilibrium states for the second Schlögl model.

The motivation for this paper is to investigate the long-time behavior such as tunnelling and metastability of the N-particle system. Consider the case when the nonlinear master equation has more than one equilibrium state. From the law of large numbers we can see that if N is large but finite, then the empirical process will normally follow the path of the dynamical system (1.4) that is attracted by one of the equilibrium states. As time goes on, the particle system will make small fluctuations near the equilibrium states. However, because of ergodicity, the transition from the neighborhood of one equilibrium state to another may eventually occur via large deviations. It is the purpose of the present paper to study the large deviations of the empirical process $\varepsilon_{x^{(N)}(.)}$ from the dynamical system (1.4) and to find a nice representation for the action functional.

Similar problems have been studied extensively by many authors. Freidlin and Wentzell [9] studied the small random perturbations of finite dimensional dynamical systems and developed a large deviation theory to investigate the long-time behavior of such systems. An infinite dimensional generalization of the Freidlin–Wentzell theory was obtained by Dawson and Gärtner [3] for the weakly interacting diffusions in the McKean–Vlasov limit. By identifying three different expressions for the action functional, they got an integral form for the action functional that is an analogue to the finite dimensional case. In [3], the large deviation system is established by a generalized Sanov theorem for the empirical distribution and the contraction principle, whereas a much greater effort was made to obtain the integral form for the action functional.

Our model belongs to the setting of jump processes. Large deviations for this type of process have been studied by several authors in various situations. Comets [2] proved a large deviation principle of the empirical processes for the Curie–Weiss model on the torus. Sugiura [16] studied the pure jump processes with compact state space and established a large deviation principle for both the empirical distributions and the empirical processes. Léonard [12] discussed the large deviations for the empirical processes of the particle systems associated with spatially homogeneous Boltzmann type equations. The special feature of our model is that it is not only a pure jump Markov process with mean-field interaction, but also with unbounded jumps. The main result of the present paper is to establish a large deviation system for the empirical process (1.3) and give a nice expression for the action functional.

The present paper is a continuation of [7] in which the large deviations for the empirical distribution

$$\eta_{x^{(N)}(\cdot)} = \frac{1}{N}\sum_{k=1}^{N} \delta_{x_k^{(N)}(\cdot)}$$

are studied. Using the "inductive topology" and the Cameron–Martin–Girsanov formula, it is proved that the distribution of $\eta_{x^{(N)}(\cdot)}$ satisfies a full large deviation principle with rate function equal to the relative entropy. Consider the map

$$\pi\big(\eta_{x^{(N)}(\cdot)}\big) = \big(t \mapsto \varepsilon_{x^{(N)}(t)}\big).$$

If π were continuous in the appropriate space, as it is in the diffusion case, then the large deviation property for the empirical processes would follow directly from the contraction principle. Unfortunately, this is not true. Even so, it is still possible to get the large deviation principle by using Lemma 2.1.4 of [5] and some continuous exponential approximations of the map π. However, in order to get a nicer expression for the action functional, we choose to use a direct proof that is a combination of the techniques used in [2], [7] and [16].

An outline of the development of the article is as follows. Some notation and preliminary results are given in Section 2. The main result is given in Section 3. Finally, in section 4 we prove the main result through three subsections: subsection 4.1 deals with the lower bound, subsection 4.2 gives the proof of the compactness of the level sets and the upper bound is obtained in subsection 4.3.

2. Notation and preliminary. Let $E = \{0, 1, \ldots\}$ be equipped with the discrete topology, let ρ be the discrete metric, let $E^{\otimes N}$ be the N-fold Cartesian product of E for $N \geq 1$ and let $M_1(E)$ be the set of all probability measures on E with the usual weak topology. For any $T > 0$, $D([0,T], E)$ denotes the space of functions from $[0,T]$ to E that are right continuous and have left limits at each $t \in (0,T]$ and are left continuous at T, furnished with the Skorohod topology. The restriction to $[0,T]$ of the processes in the sequel will have a left continuity modification at time T. It is well known that $D([0,T], E)$ is a Polish space in which the Borel σ-algebra coincides with $\mathcal{F} = \sigma\{x(t)\colon 0 \leq t \leq T\}$, the smallest σ-algebra generated by $\{x(t)\colon 0 \leq t \leq T\}$, where $x(t) = x(t,\omega) = \omega(t)$ for all $t \geq 0$ and $\omega \in D([0,T],E)$ (cf. [6]). Similarly, we can introduce spaces $D([0,T], E^{\otimes N})$ and $D([0,T], M_1(E))$. $C_b(E)$ is the set of all bounded continuous functions on E and $C_k(E)$ is the set of all continuous functions on E with compact support. (*Note*: In fact all functions on E are continuous.) $C_b(E^{\otimes N})$ and $C_k(E^{\otimes N})$ are defined similarly. For any $f \in C_b(E)$, $\mu \in M_1(E)$, we will use $\langle \mu, f\rangle$ to denote the integration of f with respect to μ. For any subset A of E, I_A will be used to denote the indicator function of set A.

Let φ be a function on E defined by $\varphi(x) = 1 + x\log\log(x+3)$. For each $m \geq 1$, let $\mathcal{D}_m = \{\mu(\cdot) \in D([0,T], M_1(E));\ \sup_{0\leq t\leq T}\langle \mu(t), \varphi\rangle \leq m\}$ be equipped with the subspace topology of $D([0,T], M_1(E))$. $\mathcal{D}_\infty = \bigcup_{m\geq 1} \mathcal{D}_m$ is equipped with the "inductive topology." By definition, a set V is open in $\mathcal{D}_\infty$ if and only if $V \cap \mathcal{D}_m$

is open in $\mathcal{D}_m$ for each $m \geq 1$. It is known that a function is continuous on $\mathcal{D}_\infty$ if and only if it is sequentially continuous.

Let $M_1^m(E) = \{\mu \in M_1(E);\ \langle \mu, \varphi \rangle \leq m\}$ with the subspace topology of $M_1(E)$. Then it is not hard to see that $\mathcal{D}_m = D([0,T],\ M_1^m(E))$. Let r denote the metric on $M_1(E)$ defined by $r(u,v) = \sum_{n=0}^{\infty} 2^{-n}|u(n) - v(n)|$, $\forall u,\ v \in M_1(E)$. It is well known that metric r induces the vague topology on $M_1(E)$ and the vague topology coincides with the weak topology on $M_1^m(E)$ for any $m \geq 1$. Hence the two topologies will induce the same "inductive topology" on $\mathcal{D}_\infty$. For any $\mu(\cdot),\ \nu(\cdot) \in \mathcal{D}_\infty$, let $r_{0T}(\mu(\cdot), \nu(\cdot)) = \sup_{0 \leq t \leq T} r(\mu(t), \nu(t))$. Then r_{0T} induces the uniform convergence topology on $\mathcal{D}_\infty$. From the definition of r and the right continuity, we can prove that for any fixed $\nu_0(\cdot) \in \mathcal{D}_\infty$, $F(\mu(\cdot)) = r_{0T}(\mu(\cdot), \nu_0(\cdot))$ is measurable with respect to the Borel σ-algebra of space $\mathcal{D}_\infty$ with the "inductive topology." This fact will be used in proving the upper bound.

Let $Q = (q_{x,y})_{x,y \in E}$ be a totally stable conservative Q-matrix satisfying:

$$\inf_{x \in E}\{q_{x,x+1}\} > 0; \tag{2.1}$$

$$\exists \Lambda > 0, \text{ such that } q_{x,y} = 0 \text{ for } |x - y| \geq \Lambda; \tag{2.2}$$

$$\begin{gathered} \exists \lambda > 0 \text{ such that } \forall x \in E, \\ \sum_{y \in E} q_{x,y}(y - x) \leq \lambda(x + 1), \qquad \sum_{y \in E} q_{x,y}\big(\varphi(y) - \varphi(x)\big) \leq \lambda \varphi(x); \end{gathered} \tag{2.3}$$

$$\begin{aligned} &\exists 0 < c < \infty, \text{ such that } \forall x,\ y \in E,\ y > x, \\ &\sum_{z \neq 0}(q_{y,y+z} - q_{x,x+z})z + 2\sum_{z=1}^{\infty}\big[(q_{y,x-z} - q_{x,2x-y-z}) \vee 0 \\ &\qquad + (q_{x,y+z} - q_{y,2y-x+z}) \vee 0\big]z \leq c(y - x); \end{aligned} \tag{2.4}$$

$$\begin{aligned} &\forall l \geq 0,\ \exists\ \lambda(l) > 0,\ \ni \\ &\sup_{x \in E} \max\Bigg\{ \sum_{y \in E} q_{x,y}\big(\exp(y - x) - 1\big) + (e - 1)x, \\ &\qquad \sum_{y \in E} q_{x,y}\Big(\exp\big(\varphi(y) - \varphi(x)\big) - 1\Big) \\ &\qquad\quad + \Big(\exp\big(\varphi(x+1) - \varphi(x)\big) - 1\Big)l \Bigg\} \leq \lambda(l). \end{aligned} \tag{2.5}$$

REMARK. All these conditions are satisfied by many models including the first and the second Schlögl models. For more detail refer to [7], [8] and [15].

For $x, x', y \in E$, $u \in M_1(E)$, define

$$Q(x, x', y) = \begin{cases} q_{x,x+1} + x', & \text{for } y = x + 1, \\ q_{x,y}, & \text{for } y \neq x, y \neq x + 1, \\ 0, & \text{otherwise}, \end{cases} \tag{2.6}$$

$$Q_u(x,y) = \int_E Q(x,x',y)u(dx'). \tag{2.7}$$

Then we can rewrite (1.1) and (1.2) as

$$Q_u f(x) = \int_E Q_u(x,dy)\big(f(y) - f(x)\big),$$

$$\Omega^{(N)}\psi\big(x^{(N)}\big) = \sum_{i=1}^{N} \int_E Q_{\varepsilon_{x^{(N)}}}\big(x_i^{(N)}, dy\big)\Delta_i^y \psi\big(x^{(N)}\big),$$

where

$$\Delta_i^y \psi\big(x^{(N)}\big) = \psi\big(x_1^{(N)}, \dots, x_{i-1}^{(N)}, y, x_{i+1}^{(N)}, \dots, x_N^{(N)}\big) - \psi\big(x_1^{(N)}, \dots, x_{i-1}^{(N)}, x_i^{(N)}, x_{i+1}^{(N)}, \dots, x_N^{(N)}\big).$$

For every $u \in M_1(E)$, $\overline{\mu}(\cdot) \in D([0,T], M_1(E))$ satisfying

$$\sup_{0 \le t \le T} \langle \overline{\mu}(t), \varphi \rangle < \infty, \qquad \langle u, \varphi \rangle < \infty,$$

let $P_{\bar{\mu}(\cdot), u}$ be the unique solution to the time-inhomogeneous martingale problem for $Q_{\bar{\mu}(\cdot)}$ with initial distribution u.

For each $\omega \in D$ and $A \subset E$, let

$$N(t,A;\omega) = \#\{s\colon \omega(s) \in A,\ \omega(s) \ne \omega(s-),\ s \le t\}, \tag{2.8}$$

$$\widetilde{N}(t,A;\omega) = N(t,A;\omega) - \int_0^t Q_{\bar{\mu}(s)}\big(\omega(s), A\big)\, ds, \tag{2.9}$$

where $Q_{\bar{\mu}(s)}(\omega(s), A) = \sum_{y \in A} Q_{\bar{\mu}(s)}(\omega(s), y)$. Then we have the following lemma (cf. [7]).

LEMMA 2.1. *For every finite subset A of E,*

$$\widetilde{N}(t,A;\ \omega) \text{ is a } P_{\bar{\mu}(\cdot),u}\text{-martingale.} \tag{2.10}$$

For every $x, x', y \in E, y \ne x, \mu, \nu \in M_1(E)$, let

$$q_\mu(x,x',y) = \frac{Q(x,x',y)}{Q_\mu(x,y)}, \qquad q_\mu^\nu(x;y) = \frac{Q_\nu(x,y)}{Q_\mu(x,y)} \tag{2.11}$$

and

$$q_\mu(x,x',x) = q_\mu^\nu(x;x) = 0.$$

For any integer $N \ge 1$, let $P_u^{(N)}$ be the unique solution to the martingale problem for $\Omega^{(N)}$ with initial distribution $u^{\otimes N}$. Let $P_{\bar{\mu}(\cdot),u}^{(N)}$ be the N-fold independent product of $P_{\bar{\mu}(\cdot),u}$. For every $x^{(N)}(\cdot) \in D([0,T], E^{\otimes N})$, define

$$\varepsilon_{x^{(N)}(\cdot)} = \frac{1}{N}\sum_{k=1}^{N} \delta_{x_k^{(N)}(\cdot)}.$$

Let $\mathcal{P}_u^{(N)}(\cdot) = P_u^{(N)} \circ \varepsilon_{x^{(N)}(\cdot)}^{-1}(\cdot)$ and $\mathcal{P}_{\bar{\mu}(\cdot),u}^{(N)}(\cdot) = P_{\bar{\mu}(\cdot),u}^{(N)} \circ \varepsilon_{x^{(N)}(\cdot)}^{-1}(\cdot)$. Obviously,

$$\mathcal{P}_u^{(N)}(\cdot),\ \mathcal{P}_{\bar{\mu}(\cdot),u}^{(N)}(\cdot) \in M_1\Big(D\big([0,T], M_1(E)\big)\Big).$$

In fact, they are both concentrated on the space $\mathcal{D}_\infty$.

Then we have the following theorem (cf. [4]).

THEOREM 2.2 (Cameron–Martin–Girsanov). *$P_{\bar{\mu}(\cdot),u}^{(N)}$ and $P_u^{(N)}$ are mutually absolutely continuous and*

$$\frac{dP_u^{(N)}}{dP_{\bar{\mu}(\cdot),u}^{(N)}}\big(x^{(N)}(\cdot)\big) = \exp\Big(H_T^{(N)}\big(x^{(N)}(\cdot)\big)\Big), \tag{2.12}$$

where

$$\begin{aligned} H_T^{(N)}\big(x^{(N)}(\cdot)\big) &= \sum_{i=1}^N \int_{[0,T]\times E} \Big\{ \log q_{\bar{\mu}(s-)}^{\varepsilon_{x^{(N)}(s-)}}\big(x_i^{(N)}(s-);\ y\big)\Big\} N\big(ds,dy;\ x_i^{(N)}(\cdot)\big) \\ &\quad - \sum_{i=1}^N \int_0^T \{\|\varepsilon_{x^{(N)}(s)}\| - \|\overline{\mu}(s)\|\}\,ds \\ &= \sum_{i=1}^N \int_{[0,T]\times E} \Big\{ \log \frac{1}{N} \sum_{j=1}^N q_{\bar{\mu}(s-)}\big(x_i^{(N)}(s-), x_j^{(N)}(s-);\ y\big)\Big\} \\ &\quad \times N\big(ds,dy;\ x_i^{(N)}(\cdot)\big) - \sum_{i=1}^N \int_0^T \Big\{ \Big(\frac{\sum_{j=1}^N x_j^{(N)}(s)}{N}\Big) - \|\overline{\mu}(s)\| \Big\}\,ds. \end{aligned} \tag{2.13}$$

The following lemma was proved in [7] and will be frequently used in the sequel.

LEMMA 2.3. *Let $u \in M_1(E)$ satisfy*

$$\eta = \int_E \exp\big(\varphi(x)\big)u(dx) < \infty. \tag{2.14}$$

Assume conditions (2.1)–(2.5) *are satisfied. Then for any $r > 0$, there exists an $R_0 > 0$ such that for all $R \geq R_0$ and $N \geq 1$, we have*

$$\mathcal{P}_u^{(N)}\{\mathcal{D}_\infty \backslash \mathcal{D}_R\} \leq \exp(-Nr). \tag{2.15}$$

REMARK. Using exactly the same argument, we can show that (2.15) still holds when $\mathcal{P}_u^{(N)}$ is replaced by $\mathcal{P}_{\bar{\mu}(\cdot),u}^{(N)}$.

3. The main result. Let X be a Hausdorff topological space and let $\{P_N\}_{N\geq 1}$ be a sequence of probability measures on X. $\{a_N\}$ is a sequence of positive numbers tending to ∞. $S(\cdot)$ is a function defined on X with values in $[0,\infty]$.

DEFINITION 3.1. (X, P_N, a_N) is said to be a large deviation system with action functional $S(\cdot)$ if:

(i) For every open subset G of X,

$$\liminf_{N\to\infty} a_N^{-1} \log P_N(G) \geq -\inf_{x\in G} S(x). \tag{3.1}$$

(ii) For every closed subset F of X,

$$\limsup_{N\to\infty} a_N^{-1} \log P_N(F) \leq -\inf_{x\in F} S(x). \tag{3.2}$$

(iii) The level sets $\{x\in X: S(x)\leq s\}$ are compact for all $s\geq 0$.

(By convention the infimum of the empty set is ∞.)

REMARK. If (X, P_N, a_N) is a large deviation system with action functional $S(\cdot)$, then we also say that the sequence $\{P_N\}$ satisfies a full large deviation principle with rate function $S(\cdot)$ (cf. [5], Chapter 2).

For fixed $T>0$, let $C_k^{1,0}([0,T]\times E)$ denote the set of all continuous functions on $[0,T]\times E$ with compact support and first order continuous derivative with respect to t. For every $u,\nu\in M_1(E)$, $g\in C_b(E)$, $f\in C_k^{1,0}([0,T]\times E)$ and every $\bar{\mu}(\cdot)$, $\mu(\cdot)\in \mathcal{D}_\infty$, let

$$\begin{aligned} J_{\bar{\mu}(\cdot)}\big(\mu(\cdot); f, T\big) &= \langle \mu(T), f(T)\rangle - \langle \mu(0), f(0)\rangle \\ &\quad - \int_0^T \left\langle \mu(s), \frac{\partial f(s)}{\partial s} + \exp\big(-f(s)\big) Q_{\bar{\mu}(s)} \exp\big(f(s)\big)\right\rangle ds, \\ J\big(\mu(\cdot); f, T\big) &= J_{\mu(\cdot)}\big(\mu(\cdot); f, T\big), \qquad l_u(\nu, g) = \langle \nu, g\rangle - \log\langle u, e^g\rangle \end{aligned}$$

and

$$S_u^{\bar{\mu}(\cdot)}\big(\mu(\cdot)\big) = \sup\Big\{ l_u\big(\mu(0), g\big) + J_{\bar{\mu}(\cdot)}\big(\mu(\cdot); f, T\big);\ f\in C_k^{1,0}\big([0,T]\times E\big), g\in C_b(E)\Big\}, \tag{3.3}$$

$$S_u\big(\mu(\cdot)\big) = S_u^{\mu(\cdot)}\big(\mu(\cdot)\big). \tag{3.4}$$

It is not hard to see that $\forall g\in C_b(E)$, $f\in C_k^{1,0}([0,T]\times E)$, $l_u(\mu(0), g) + J(\mu(\cdot); f, T)$ is continuous on the space $\mathcal{D}_\infty$ and thus $S_u(\cdot)$ is lower semicontinuous. It is also true that $S_u(\mu(\cdot)) = 0$ if $\mu(\cdot)$ is the unique solution to the nonlinear master equation (1.4) with starting point u.

The main result of the present paper is:

THEOREM 3.1. *Let $u \in M_1(E)$ satisfy* (2.14). *Then under assumptions* (2.1)–(2.5), $(\mathcal{D}_\infty, \mathcal{P}_u^{(N)}, N)$ *is a large deviation system with action functional* $S_u(\cdot)$.

4. The proof of theorem 3.1.

4.1. *Lower bound.* In this subsection we will get the lower bound through a series of lemmas. The method we use is similar to that used in [16]. Because of the unboundedness of jumps, some new methods and techniques are needed.

For $\bar{\mu}(\cdot) \in \mathcal{D}_\infty$, $T > 0$, $s \in [0, T]$ and $x \in E$, let $P_{\bar{\mu}(\cdot),(s,x)}$ be the unique solution to the martingale problem for $\{Q_{\bar{\mu}(t)};\ t \in [0, T]\}$ with initial distribution δ_x at time s. For $0 \le s < t \le T$, $u, \nu \in M_1(E)$, $f \in C_b(E)$, we define

$$l_{\bar{\mu}(\cdot)}^{s,t}\colon M_1(E) \times M_1(E) \times C_b(E) \to (-\infty, +\infty),$$

$$l_{\bar{\mu}(\cdot)}^{s,t}(u, \nu; f) = \Big\langle \nu, f\big(x(t)\big)\Big\rangle - \Big\langle u, \log E^{P_{\bar{\mu}(\cdot),(s,x)}}\Big[\exp\Big(f\big(x(t)\big)\Big)\Big]\Big\rangle.$$

For any $M \ge 1$ and $0 = t_0 < t_1 < \cdots < t_M = T$, we denote by $\pi^{(N)}(t_0, t_1, \dots, t_M)$ the joint distribution of $(\varepsilon_{x^{(N)}(t_0)}, \dots, \varepsilon_{x^{(N)}(t_M)})$ on $M_1(E)^{\otimes(M+1)}$ under $P_{\bar{\mu}(\cdot),u}^{(N)}$, $p(t_0, t_1, \dots, t_M)$ the joint distribution of $(x(t_0), \dots, x(t_M))$ on $E^{\otimes(M+1)}$ under $P_{\bar{\mu}(\cdot),u}$. Then we have the following lemma.

LEMMA 4.1. *If $M_1(E)$ is endowed with the weak topology, then for any $M \ge 1$, $0 = t_0 < t_1 < \cdots < t_M = T$, $(M_1(E)^{\otimes(M+1)}, \pi^{(N)}(t_0, t_1, \dots, t_M), N)$ is a large deviation system with action functional $L_{\bar{\mu}(\cdot),u}^{t_0,\dots,t_M}(\cdots)$ defined by*

$$\begin{aligned}
&L_{\bar{\mu}(\cdot),u}^{t_0,\dots,t_M}(\mu_0, \mu_1, \dots, \mu_M)\\
&\qquad = \sup_{g \in C_b(E)} l_u(\mu_0; g) + \sum_{k=0}^{M-1} \sup_{f \in C_k(E)} l_{\bar{\mu}(\cdot)}^{t_k,t_{k+1}}(\mu_k, \mu_{k+1}; f)\\
&\qquad = \sup_{g \in C_b(E)} l_u(\mu_0; g) + \sum_{k=0}^{M-1} \sup_{f \in C_b(E)} l_{\bar{\mu}(\cdot)}^{t_k,t_{k+1}}(\mu_k, \mu_{k+1}; f).
\end{aligned} \tag{4.1}$$

The second equality holds because $C_k(E)$ is pointwise dense in $C_b(E)$.

PROOF. Theorem 3.5 in [3] implies that $(M_1(E)^{\otimes(M+1)}, \pi^{(N)}(t_0, t_1, \dots, t_M), N)$ is a large deviation system with action functional

$$\widetilde{L}(\mu_0, \mu_1, \dots, \mu_M) = \sup_{f_0,\dots,f_M \in C_b(E)} \left[\sum_{i=0}^{M} \langle \mu_i, f_i\rangle - H(f_0, \dots, f_M)\right],$$

where

$$H(f_0, \dots, f_M) = \log \int_{E^{\otimes(M+1)}} p(t_0, t_1, \dots, t_M)(dx_0, \dots, dx_M) \exp\left(\sum_{i=0}^{M} f_i(x_i)\right).$$

By the Markov property of $P_{\tilde{\mu}(\cdot),u}$ we can rewrite $H(f_0,\dots,f_M)$ as

$$\log E^{P_{\tilde{\mu}(\cdot),u}}\left[\exp\left\{f_0\big(x(t_0)\big)+h\big(f_1,\dots,f_M\big)\big(x(t_0)\big)\right\}\right],$$

where

$$h(f_1,\dots,f_M)\big(x(t_0)\big)=\log E^{P_{\tilde{\mu}(\cdot),(t_0,x(t_0))}}\left[\exp\left\{\sum_{i=1}^{M} f_i\big(x(t_i)\big)\right\}\right].$$

Because $h(f_1,\dots,f_M)\in C_b(E)$, we conclude that

$$\begin{aligned}
&\tilde{L}(\mu_0,\mu_1,\dots,\mu_M)\\
&\quad=\sup_{f_0,\dots,f_M\in C_b(E)}\left[l_u\big(\mu_0;f_0+h(f_1,\dots,f_M)\big)+\sum_{i=1}^{M}\langle\mu_i,f_i\rangle-\big\langle\mu_0,h(f_1,\dots,f_M)\big\rangle\right]\\
&\quad=\sup_{g\in C_b(E)} l_u(\mu_0;g)+\sup_{f_1,\dots,f_M\in C_b(E)}\sum_{i=1}^{M}\left[\langle\mu_i,f_i\rangle-\big\langle\mu_0,h(f_1,\dots,f_M)\big\rangle\right].
\end{aligned}$$

By induction we get the result. □

The following estimate is crucial in getting the lower bound from the large deviation system obtained in Lemma 4.1.

LEMMA 4.2. *For every $\gamma>0$, $\tilde{\mu}(\cdot)\in\mathcal{D}_\infty\cap C([0,T],M_1(E))$ and open neighborhood V of $\tilde{\mu}(\cdot)$ in $\mathcal{D}_\infty$ with the inductive topology, there exist finitely many $0=t_0<t_1<\cdots<t_M=T$ and open neighborhoods $\tilde{V}_i$ of $\tilde{\mu}(t_i)$, $i=0,\dots,M-1$, in $M_1(E)$, under the vague topology such that*

$$\limsup_{N\to\infty}\frac{1}{N}\log P^{(N)}_{\tilde{\mu}(\cdot),u}\left(\{\varepsilon.\notin V\}\cap\bigcap_{i=0}^{M-1}\{\varepsilon.(t_i)\in\tilde{V}_i\}\right)\le-\gamma. \tag{4.2}$$

PROOF. For any $\tilde{\mu}(\cdot)\in\mathcal{D}_\infty\cap C([0,T],M_1(E))$, let V be any open neighborhood of $\tilde{\mu}(\cdot)$ in $\mathcal{D}_\infty$ with the inductive topology.

Lemma 2.3 implies that $\forall\gamma>0,\exists m>0$, such that

$$\limsup_{N\to\infty}\frac{1}{N}\log P^{(N)}_{\tilde{\mu}(\cdot),u}\big(\mathcal{D}_m^c\big)\le-\gamma. \tag{4.3}$$

The continuity of $\tilde{\mu}(\cdot)$ implies its uniform continuity. Thus

$$\exists\varepsilon,\delta>0\ni\sup_{\substack{0\le s<t\le T,\\ t-s<\delta}} r\big(\tilde{\mu}(t),\tilde{\mu}(s)\big)<\frac{\varepsilon}{4}, \tag{4.4}$$

$$\left\{\mu(\cdot)\in\mathcal{D}_m:\sup_{t\in[0,T]} r\big(\mu(t),\tilde{\mu}(t)\big)<\varepsilon\right\}\subset V. \tag{4.5}$$

Now we choose any partition $0 = t_0 < t_1 < \cdots < t_M = T$ of $[0,T]$ satisfying $t_i - t_{i-1} < \delta$, for $i = 0, \cdots, M$. If we take $\widetilde{V}_i = \{\nu;\ r(\nu, \widetilde{\mu}(t_i)) < \varepsilon/4\}$, then we have

$$\begin{aligned} &P^{(N)}_{\overline{\mu}(\cdot),u}\left(\{\varepsilon.\notin V\}\cap\bigcap_{i=0}^{M-1}\{\varepsilon.(t_i)\in\widetilde{V}_i\}\right)\\ &\quad\le P^{(N)}_{\bar{\mu}(\cdot),u}(\mathcal{D}^c_m)+\sum_{i=1}^{M-1}P^{(N)}_{\bar{\mu}(\cdot),u}\left(\sup_{t\in[t_i,t_{i+1})}r(\varepsilon.(t),\varepsilon.(t_i))\ge\frac{\varepsilon}{2}\right).\end{aligned}\tag{4.6}$$

Let m_ε be an integer such that $\sum_{n=m_\varepsilon+1}^{\infty}2^{-n} < \varepsilon/4$ and

$$q_{m_\varepsilon}=2\left(\sum_{x=0}^{m_\varepsilon+\Lambda}\sum_{y\in E}q_{x,y}+\sup_{0\le t\le T}\{\|\overline{\mu}(t)\|\}\right),$$

$$N([t',t],\ A;\ \omega)=\#\{s\colon \omega(s)\in A,\ \omega(s)\ne\omega(s-),\ s\in[t',\ t]\}.$$

Then we have

$$\begin{aligned} &P^{(N)}_{\bar{\mu}(\cdot),u}\left\{\sup_{t\in[t_i,t_{i+1})}r(\varepsilon.(t),\varepsilon.(t_i))\ge\frac{\varepsilon}{2}\right\}\\ &\quad\le P^{(N)}_{\bar{\mu}(\cdot),u}\left\{\sup_{t\in[t_i,t_{i+1})}\sum_{k=0}^{m_\varepsilon}\frac{1}{N}\left|\sum_{j=1}^{N}\left(I_{\{k\}}(x_j^{(N)}(t_i))-I_{\{k\}}(x_j^{(N)}(t))\right)\right|\ge\frac{\varepsilon}{4}\right\}\\ &\quad\le\sum_{k=0}^{m_\varepsilon}P^{(N)}_{\bar{\mu}(\cdot),u}\left\{\sup_{t\in[t_i,t_{i+1})}\frac{1}{N}\left|\sum_{j=1}^{N}\left(I_{\{k\}}(x_j^{(N)}(t_i))-I_{\{k\}}(x_j^{(N)}(t))\right)\right|\ge\frac{\varepsilon}{4m_\varepsilon}\right\}\end{aligned}\tag{4.7}$$

and

$$\begin{aligned} &P^{(N)}_{\bar{\mu}(\cdot),u}\left\{\sup_{t\in[t_i,t_{i+1})}\left|\left(I_{\{k\}}(x_1^{(N)}(t_i))-I_{\{k\}}(x_1^{(N)}(t))\right)\right|\ge\frac{\varepsilon}{4m_\varepsilon}\right\}\\ &\quad\le P_{\bar{\mu}(\cdot),u}\left\{\sup_{t\in[t_i,t_{i+1})}\left|\left(I_{\{k\}}(x_1^{(N)}(t_i))-I_{\{k\}}(x_1^{(N)}(t))\right)\right|\ge\frac{\varepsilon}{4m_\varepsilon},\right.\\ &\qquad\qquad\left.\exists t\in[t_i,t_{i+1}),\ x_1^{(N)}(t)\ne x_1^{(N)}(t_i)=k\right\}\\ &\qquad+P_{\bar{\mu}(\cdot),u}\left\{\sup_{t\in[t_i,t_{i+1})}\left|\left(I_{\{k\}}(x_1^{(N)}(t_i))-I_{\{k\}}(x_1^{(N)}(t))\right)\right|\ge\frac{\varepsilon}{4m_\varepsilon},\right.\\ &\qquad\qquad\left.\exists t\in[t_i,t_{i+1}),\ x_1^{(N)}(t_i)\ne x_1^{(N)}(t)=k\right\}\\ &\quad\le 2P_{\bar{\mu}(\cdot),u}\left\{N([t_i,t_{i+1}],\{0,1,\ldots,m_\varepsilon+\Lambda\};\ x_1^{(N)}(\cdot))\ge 1\right\}\\ &\quad\le 2E^{P_{\bar{\mu}(\cdot),u}}\left\{N([t_i,t_{i+1}],\ \{0,1,\ldots,m_\varepsilon+\Lambda\};\ x_1^{(N)}(\cdot))\right\}\end{aligned}$$

$$\leq 2E^{P_{\bar{\mu}(\cdot),u}}\left\{\int_{t_i}^{t_{i+1}} Q_{\bar{\mu}(s)}\big(x_1^{(N)}(\cdot);\ \{0,1,\dots,m_\varepsilon+\Lambda\}\big)\,ds\right\}$$
$$\leq q_{m_\varepsilon}(t_{i+1}-t_i).$$

This combined with the independence implies

$$P^{(N)}_{\bar{\mu}(\cdot),u}\left\{\sup_{t\in[t_i,t_{i+1})}\frac{1}{N}\left|\sum_{j=1}^{N}\Big(I_{\{k\}}\big(x_j^{(N)}(t_i)\big)-I_{\{k\}}\big(x_j^{(N)}(t)\big)\Big)\right|\geq\frac{\varepsilon}{4m_\varepsilon}\right\}$$
$$\leq\sum_{n=N_\varepsilon}^{N}\binom{N}{n}\left(2P_{\bar{\mu}(\cdot),u}\Big\{N\big([t_i,t_{i+1}],\ \{0,1,\dots,m_\varepsilon+\Lambda\};\ x_1^{(N)}(\cdot)\big)\geq 1\Big\}\right)^n$$
$$\leq 4^N\big[q_{m_\varepsilon}(t_{i+1}-t_i)\big]^{N_\varepsilon},$$

where $N_\varepsilon=[N\cdot\varepsilon/4m_\varepsilon]-1$.

Thus we have

$$\sum_{i=1}^{M-1}P^{(N)}_{\bar{\mu}(\cdot),u}\left(\sup_{t\in[t_i,t_{i+1})}r\big(\varepsilon.(t),\varepsilon.(t_i)\big)\geq\frac{\varepsilon}{2}\right)\leq Mm_\varepsilon 4^N\Big[q_{m_\varepsilon}\max\big\{(t_{i+1}-t_i)\big\}\Big]^{N_\varepsilon}. \tag{4.8}$$

If we take

$$\max_i\big\{(t_{i+1}-t_i)\big\}\leq\left(\exp\left[\frac{4m_\varepsilon(-\gamma-\log 4)}{\varepsilon}-\log q_{m_\varepsilon}\right]\right)\wedge\delta, \tag{4.9}$$

then we obtain the inequality. □

LEMMA 4.3. *If $\mu(\cdot)\in\mathfrak{D}_\infty, S_u^{\bar{\mu}(\cdot)}(\mu(\cdot))<\infty$, then there exists an integrable function $h_{0,T}$ on $[0,T]$ such that for any $0\leq s<t\leq T$ and $g\in C_k(E)$,*

$$\langle\mu(t),g\rangle-\langle\mu(s),g\rangle=\int_s^t h_{0,T}(r)\,dr+\int_s^t\big\langle\mu(r),Q_{\bar{\mu}(r)}g(r)\big\rangle\,dr. \tag{4.10}$$

This also implies that $\mu(\cdot)$ is continuous with respect to t and $\langle\mu(\cdot),g\rangle$ is absolutely continuous.

PROOF. Let $\mu(\cdot)\in\mathfrak{D}_\infty$. To prove the continuity of $\mu(\cdot)$ it suffices to show that the map $t\to\langle\mu(t),g\rangle$ is continuous for any $g\in C_k(E)$.

For any continuity points $0\leq s<t\leq T$ of $\mu(\cdot)$, we can find a sequence of smooth functions $\{h_n\}$ with $h_n\to I_{[s,t]}$ such that for any $f\in C_k^{1,0}([0,T]\times E)$,

$$J_{\bar{\mu}(\cdot)}\big(\mu(\cdot);h_nf,T\big)=\langle\mu(T),h_nf(T)\rangle-\langle\mu(0),h_nf(0)\rangle$$
$$-\int_0^T\left\langle\mu(r),\frac{\partial h_n f(r)}{\partial r}+\exp\big(-h_nf(r)\big)Q_{\bar{\mu}(r)}\exp\big(h_nf(r)\big)\right\rangle dr$$

$$\to \langle \mu(t), f(t)\rangle - \langle \mu(s), f(s)\rangle - \int_s^t \left\langle \mu(r), \frac{\partial f(r)}{\partial r} + \exp\big(-f(r)\big) Q_{\bar{\mu}(r)} \exp\big(f(r)\big)\right\rangle dr \quad \text{as } n \to 0. \tag{4.11}$$

Thus $S_u^{\bar{\mu}(\cdot)}(\mu(\cdot)) < \infty$ implies that there exists an $R > 0$ such that

$$\sup_{f \in C_k^{1,0}(E)} \left\{ \langle \mu(t), f(t)\rangle - \langle \mu(s), f(s)\rangle - \int_s^t \left\langle \mu(r), \frac{\partial f(r)}{\partial r} + \exp\big(-f(r)\big) Q_{\bar{\mu}(r)} \exp\big(f(r)\big)\right\rangle dr \right\} < R. \tag{4.12}$$

In general, (4.12) is true for arbitrary $0 \le s < t \le T$ by the right continuity of $\overline{\mu}(\cdot)$. For fixed $g \in C_k(E)$, $g \neq 0$, $0 \le s < t \le T$, and $h \in C^1([s,t])$ (the set of all continuous functions on $[s,t]$ with continuous first order derivative), define

$$I^{s,t}(h) = h(t)\langle \mu(t), g\rangle - h(s)\langle \mu(s), g\rangle - \int_s^t \frac{\partial h(r)}{\partial r}\langle \mu(r), g\rangle\, dr - \int_s^t h(r)\langle \mu(r), Q_{\bar{\mu}(r)} g\rangle\, dr.$$

Let $\|h\| = \sup_{r \in [s,t]} |h(r)|$. Then for any h with $\|h\| \le 1$ we have

$$\begin{aligned}
|I^{s,t}(h)| &= \max\{I^{s,t}(h), I^{s,t}(-h)\} \\
&\le \max\Bigg\{ I^{s,t}(h) \\
&\qquad - \int_s^t \langle \mu(r), \exp\big(-g(r)h(r)\big) Q_{\bar{\mu}(r)} \exp\big(g(r)h(r)\big) - Q_{\bar{\mu}(r)} h(r) g(r)\rangle\, dr \\
&\qquad + \int_s^t \langle \mu(r), \exp\big(-g(r)h(r)\big) Q_{\bar{\mu}(r)} \exp\big(g(r)h(r)\big) - Q_{\bar{\mu}(r)} h(r) g(r)\rangle\, dr, \\
&\qquad I^{s,t}(-h) - \int_s^t \Big\langle \mu(r), \exp\Big(-g(r)\big(-h(r)\big)\Big) Q_{\bar{\mu}(r)} \exp\Big(g(r)\big(-h(r)\big)\Big) \\
&\qquad\qquad - Q_{\bar{\mu}(r)}\big(-h(r)\big) g(r)\Big\rangle\, dr \\
&\qquad + \int_s^t \Big\langle \mu(r), \exp\Big(-g(r)\big(-h(r)\big)\Big) Q_{\bar{\mu}(r)} \exp\Big(g(r)\big(-h(r)\big)\Big) \\
&\qquad\qquad - Q_{\bar{\mu}(r)}\big(-h(r)\big) g(r)\Big\rangle\, dr \Bigg\} \\
&\le R + M_g(\exp 2\|g\| + 2\|g\| + 1),
\end{aligned}$$

where $M_g = \sup_{r \in [s,t]} [\langle \mu(r), x\rangle + \sum_{x \text{ or } y \in \text{supp}(g),\, x \neq y} q_{x,y}] < \infty$ and $\text{supp}(g)$ denotes the support of g. Because

$$|I^{s,t}(h)| = \|h\| I^{s,t}\left(\frac{h}{\|h\|}\right)$$

and $C^1([s,t])$ is dense in $C([s,t])$, we conclude that $I^{s,t}$ belongs to the dual space $C^*([s,t])$ of $C([s,t])$. Following Riesz–Markov–Kakutani (cf. [14]), there exists a unique Baire measure $m_{s,t}$ (not necessarly nonnegative) on $[s,t]$ such that $I^{s,t}(h) = \int_s^t h(r)m_{s,t}(dr)$. We also have from (4.12) that

$$\left|\int_s^t h(r)m_{s,t}(dr)\right| \leq R + \int_s^t \Big|\langle \mu(r), \exp\big(-g(r)h(r)\big)Q_{\bar{\mu}(r)}\exp\big(g(r)h(r)\big) - Q_{\bar{\mu}(r)}h(r)g(r)\rangle\Big|\,dr.$$

This implies that for every subset A of $[s,t]$ with zero Lebesgue measure and every $h \in C([s,t])$, we have

$$\left|\int_A h(r)m_{s,t}(dr)\right| \leq R.$$

This means that $m_{s,t}$ is absolutely continuous with respect to Lebesgue measure. Thus there exists an integrable function $h_{s,t}$ such that

$$I^{s,t}(h) = \int_s^t h(r)h_{s,t}(r)\,dr.$$

Because

$$\begin{aligned} I^{0,t}(h) &= I^{0,T}(h) - I^{t,T}(h) \\ &= \int_0^T h(r)h_{0,T}(r)\,dr - \int_t^T h(r)h_{t,T}(r)\,dr \\ &= \int_0^t h(r)h_{0,T}(r)\,dr + \int_t^T h(r)\big(h_{0,T}(r) - h_{t,T}(r)\big)dr. \end{aligned}$$

and the left-hand side does not depend on the value of h on $[t,T]$, we get

$$h_{0,T}(r) = h_{0,t}(r), \quad \text{for } r \in [0,t].$$

Thus taking $h = I_{[s,t]}$, we finally get (4.10), which implies the continuity of $\mu(\cdot)$ and absolute continuity of $\langle\mu(\cdot),g\rangle$, $\forall g \in C_k(E)$. □

NOTE. If we adopt the definition of absolute continuity in Section 4.1 of [3], then $\mu(t)$ in Lemma 4.3 is actually absolutely continuous.

LEMMA 4.4. *For every $\widehat{\mu}(\cdot) \in \mathcal{D}_\infty$ and open neighborhood V of $\widehat{\mu}(\cdot)$ in $\mathcal{D}_\infty$,*

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathcal{P}^{(N)}_{\bar{\mu}(\cdot),u}(V) \geq -\widetilde{S}^{\bar{\mu}(\cdot)}_u(\widehat{\mu}), \tag{4.13}$$

where

$$\widetilde{S}^{\bar{\mu}(\cdot)}_u\big(\mu(\cdot)\big) = \sup_{\substack{0=t_0<\cdots<t_M,\\ M\geq 1}} L^{t_0,\ldots,t_M}_{\bar{\mu}(\cdot),u}\big(\mu(t_0),\ldots,\mu(t_M)\big). \tag{4.14}$$

PROOF. Let $\widehat{\mu}(\cdot)$ be any element of $\mathcal{D}_\infty$ and let V be an open neighborhood of $\widehat{\mu}(\cdot)$. For every $m > 0$, $V \cap \mathcal{D}_m$ is an open subset of $\mathcal{D}_m$ by definition. If $\widetilde{S}_u^{\bar{\mu}(\cdot)}(\widehat{\mu}(\cdot)) = \infty$, then (4.13) is obvious. Now we assume $c = \widetilde{S}_u^{\bar{\mu}(\cdot)}(\widehat{\mu}(\cdot)) < \infty$. Then by Lemma 4.3, $\widehat{\mu}(\cdot)$ is continuous. Thus Lemma 4.2 is applicable. By direct calculation we have

$$\begin{aligned}\mathcal{P}^{(N)}_{\bar{\mu}(\cdot),u}(V) &\geq \mathcal{P}^{(N)}_{\bar{\mu}(\cdot),u}\left(V \cap \bigcap_{i=0}^{M-1} V_i\right)\\ &= \mathcal{P}^{(N)}_{\bar{\mu}(\cdot),u}\left(\bigcap_{i=0}^{M-1} V_i\right) - \mathcal{P}^{(N)}_{\bar{\mu}(\cdot),u}\left(V^c \cap \bigcap_{i=0}^{M-1} V_i\right)\\ &= \pi^{(N)}(t_0,\dots,t_M)(\widetilde{V}_0 \times \cdots \times \widetilde{V}_M) - \mathcal{P}^{(N)}_{\bar{\mu}(\cdot),u}\left(V^c \cap \bigcap_{i=0}^{M-1} V_i\right),\end{aligned}$$

where $V_i = \{\mu(\cdot);\ \mu(t_i) \in \widetilde{V}_i\}$, $i = 0,\dots,M-1$, and $\widetilde{V}_M = M_1(E)$.

Because $\widetilde{V}_0 \times \cdots \times \widetilde{V}_M$ is an open subset of $M_1(E)^{\otimes M+1}$ in the weak topology, if we take γ in (4.2) larger than c, then (4.2) combined with Lemma 4.1 implies

$$\begin{aligned}\limsup_{N\to\infty} \frac{1}{N} \log \mathcal{P}^{(N)}_{\bar{\mu}(\cdot),u}(V) &\geq \limsup_{N\to\infty} \frac{1}{N} \log \mathcal{P}^{(N)}_{\mu(\cdot),u}\left(V \cap \bigcap_{i=0}^{M-1} V_i\right)\\ &= \max\left\{ \limsup_{N\to\infty} \frac{1}{N} \log \mathcal{P}^{(N)}_{\bar{\mu}(\cdot),u}\left(V \cap \bigcap_{i=0}^{M-1} V_i\right),\right.\\ &\qquad \left. \limsup_{N\to\infty} \frac{1}{N} \log \mathcal{P}^{(N)}_{\bar{\mu}(\cdot),u}\left(V^c \cap \bigcap_{i=0}^{M-1} V_i\right)\right\}\\ &\geq \limsup_{N\to\infty} \frac{1}{N} \log \pi^{(N)}(t_0,\dots,t_M)(\widetilde{V}_0 \times \cdots \times \widetilde{V}_M)\\ &\geq -\widetilde{S}_u^{\bar{\mu}(\cdot)}(\widehat{\mu}(\cdot)),\end{aligned}$$

which gives (4.13). □

LEMMA 4.5. *For every $\overline{\mu}(\cdot)$, $\mu(\cdot) \in \mathcal{D}_\infty$, we have*

$$\widetilde{S}_u^{\bar{\mu}(\cdot)}(\mu(\cdot)) \leq S_u^{\bar{\mu}(\cdot)}(\mu(\cdot)). \tag{4.15}$$

PROOF. For every $n > 0$, let $E_n = \{0, 1, \dots, n\}$ and

$$q^{(n)}_{x,y} = \begin{cases} q_{x,y}, & \text{for } y \neq x,\ x \in E_n,\\ -\sum_{y\neq x} q^{(n)}_{x,y}, & \text{for } y = x,\ x \in E_n,\\ 0, & \text{otherwise.}\end{cases} \tag{4.16}$$

For any $u \in M_1(E), f \in C_b(E)$, we introduce an operator $Q_u^{(n)}$:

$$Q_u^{(n)} f(x) = \sum_{y \in E} q_{x,y}^{(n)} \big(f(y) - f(x)\big) + \|u\| I_{E_n}(x) \big(f(x+1) - f(x)\big).$$

Let $P^n_{\bar\mu(t),u}$ be the unique solution to the time-inhomogeneous martingale problem for $Q^{(n)}_{\bar\mu(t)}$ with initial distribution u. Then $P^n_{\bar\mu(t),u}$ converges weakly to $P_{\bar\mu(t),u}$, the unique solution to the time-inhomogeneous martingale problem for $Q_{\bar\mu(t)}$ with initial distribution u as n goes to infinity.

Let us define for every $s < t$ and $f \in C_k(E)$,

$$h_t^{(n)}(s,x) = E^{P^{(n)}_{\bar\mu(\cdot),(s,x)}} \Big[\exp\Big(f\big(x(t)\big)\Big) - 1 \Big],$$

$$g_t^{(n)}(s,x) = \log\big(1 + h_t^{(n)}(s,x)\big).$$

Then by the backward equation we get

$$\text{(4.17)} \qquad \frac{\partial h_t^{(n)}(s,x)}{\partial s} + Q^{(n)}_{\bar\mu(s)} h_t^{(n)}(s,x) = 0, \qquad h_t^{(n)}(t,x) = \exp\big(f(x)\big) - 1;$$

$$\text{(4.18)} \qquad \frac{\partial g_t^{(n)}(s,x)}{\partial s} + \exp\big(-g_t^{(n)}(s,x)\big) Q^{(n)}_{\bar\mu(s)} \exp\big(g_t^{(n)}(s,x)\big) = 0, \qquad g_t^{(n)}(t,x) = f(x).$$

If the support of f is contained in E_n, then by definition the supports of $h_t^{(n)}$ and $g_t^{(n)}$ with respect to the second coordinate are both contained in E_n. Hence $h_t^{(n)}$ and $g_t^{(n)}$ are both in space $C_k^{1,0}([0,T] \times E)$.

For all $t_{i+1} > t_i$, let

$$\begin{aligned} J^{t_i,t_{i+1},n}_{\bar\mu(\cdot)}\big(\mu(\cdot); g\big) &= \langle \mu(t_{i+1}), g(t_{i+1})\rangle - \langle \mu(t_i), g(t_i)\rangle \\ &\quad - \int_{t_i}^{t_{i+1}} \left\langle \mu(r), \frac{\partial g(r)}{\partial r} + \exp\big(-g(r)\big) Q^{(n)}_{\bar\mu(r)} \exp\big(g(r)\big) \right\rangle dr, \\ J^{t_i,t_{i+1}}_{\bar\mu(\cdot)}\big(\mu(\cdot); g\big) &= \langle \mu(t_{i+1}), g(t_{i+1})\rangle - \langle \mu(t_i), g(t_i)\rangle \\ &\quad - \int_{t_i}^{t_{i+1}} \left\langle \mu(r), \frac{\partial g(r)}{\partial r} + \exp\big(-g(r)\big) Q_{\bar\mu(r)} \exp\big(g(r)\big) \right\rangle dr, \\ l^{t_i,t_{i+1},n}_{\bar\mu(\cdot)}\big(\mu(\cdot), f\big) &= \langle \mu(t_{i+1}), f(x)\rangle - \langle \mu(t_i), g^{(n)}_{t_{i+1}}(t_i, x)\rangle, \end{aligned}$$

taking n large enough such that the support of f is contained in E_n. If we take $f \le 0$, then we have $g^{(n)}_{t_{i+1}} \le 0$ and

$$Q_{\bar\mu(s)} g^{(n)}_{t_{i+1}}(s,x) - Q^{(n)}_{\bar\mu(s)} g^{(n)}_{t_{i+1}}(s,x) = \begin{cases} 0, & \text{for } x \le n \text{ or } x > n + \Lambda, \\ \sum_{y \in E_n} q_{x,y} g^{(n)}_{t_{i+1}}(s,y), & \text{for } x \in [n+1, n+\Lambda]. \end{cases}$$

Thus by (4.18) we have

$$Q^{(n)}_{\bar\mu(s)} g^{(n)}_{t_{i+1}}(s,x) \ge Q_{\bar\mu(s)} g^{(n)}_{t_{i+1}}(s,x),$$

$$l^{t_i,t_{i+1},n}_{\bar\mu(\cdot)}\big(\mu(\cdot), f\big) \le J^{t_i,t_{i+1}}_{\bar\mu(\cdot)}\big(\mu(\cdot); g^{(n)}_{t_{i+1}}\big),$$

which implies

$$l_{\tilde{\mu}(\cdot)}^{t_i, t_{i+1}, n}\big(\mu(\cdot); f\big) \le \sup_{g \in C_k^{1,0}([t_i, t_{i+1}] \times E)} J_{\tilde{\mu}(\cdot)}^{t_i, t_{i+1}}\big(\mu(\cdot); g\big). \tag{4.19}$$

Letting n go to infinity, we get

$$l_{\tilde{\mu}(\cdot)}^{t_i, t_{i+1}}\big(\mu(\cdot); f\big) \le \sup_{g \in C_k^{1,0}([t_i, t_{i+1}] \times E)} J_{\tilde{\mu}(\cdot)}^{t_i, t_{i+1}}\big(\mu(\cdot); g\big). \tag{4.20}$$

Because we can take $\tilde{f} = f - \|f\|$, $\forall f \in C_k(E)$, then it can be seen that (4.20) is still true for general f.

For $0 = t_0 < t_1 < \cdots < t_M = T$, let

$$\frac{1}{4} \min_{0 \le i \le M-1} \big\{|t_{i+1} - t_i|\big\} > \varepsilon,$$

$$\sup_{g \in C_k^{1,0}([t_i, t_{i+1}] \times E)} J_{\tilde{\mu}(\cdot)}^{t_i, t_{i+1}}\big(\mu(\cdot); g\big) \le J_{\tilde{\mu}(\cdot)}^{t_i, t_{i+1}}\big(\mu(\cdot); g_i\big) + \varepsilon/M$$

$$\text{for some } g_i \in C_k^{1,0}\big([t_i, t_{i+1}] \times E\big).$$

Then each g_i has an extension on $C_k^{1,0}([0,T] \times E)$, which also will be denoted by g_i. For $0 < i < M-1$, we define

$$h_i^\varepsilon(t) = \begin{cases} 1, & \text{for } t \in [t_i, t_{i+1}], \\ 0, & \text{for } t \le t_i - \varepsilon/2 \text{ or } t \ge t_{i+1} + \varepsilon/2, \\ \text{smooth}, & \text{otherwise}, \end{cases}$$

$$h_0^\varepsilon(t) = \begin{cases} 1, & \text{for } \in [t_0, t_1], \\ 0, & \text{for } t \ge t_1 + \varepsilon/2, \\ \text{smooth}, & \text{otherwise}, \end{cases}$$

$$h_{M-1}^\varepsilon(t) = \begin{cases} 1, & \text{for } t \in [t_{M-1}, t_M], \\ 0, & \text{for } t \le t_{M-1} - \varepsilon/2, \\ \text{smooth}, & \text{otherwise}. \end{cases}$$

Let

$$\tilde{g}^\varepsilon(t,x) = \sum_{i=0}^{M-1} h_i^\varepsilon(t) g_i(t,x).$$

Then we have

$$J_{\tilde{\mu}(\cdot)}\big(\mu(\cdot); \tilde{g}^\varepsilon, T\big) = \langle \mu(T), \tilde{g}^\varepsilon(T)\rangle - \langle \mu(0), \tilde{g}^\varepsilon(0)\rangle$$
$$- \int_0^T \left\langle \mu(r), \frac{\partial \tilde{g}^\varepsilon(r)}{\partial r} + \exp\big(-\tilde{g}^\varepsilon(r)\big) Q_{\tilde{\mu}(r)} \exp\big(\tilde{g}^\varepsilon(r)\big) \right\rangle dr$$

$$= \sum_{i=0}^{M-1} \left\{ J_{\bar\mu(\cdot)}\big(\mu(\cdot);\, h_i^\varepsilon g_i, T\big) \right\}$$

$$+ \int_0^T \Big\{ \Big\langle \mu(r), \exp\big(-\tilde g^\varepsilon(r)\big) Q_{\bar\mu(r)} \exp\big(\tilde g^\varepsilon(r)\big)$$

$$- \sum_{i=0}^{M-1} \exp\big(-h_i^\varepsilon(r) g_i(r)\big) Q_{\bar\mu(r)} \exp\big(h_i^\varepsilon(r) g_i(r)\big) \Big\rangle \Big\} \, dr.$$

It is not hard to see that the second term will go to zero as $\varepsilon \to 0$. Thus by (4.11) we get that for any pair of continuity points $t_i < t_{i+1}$ of $\mu(\cdot)$,

$$J_{\bar\mu(\cdot)}\big(\mu(\cdot);\, \tilde g^\varepsilon\big) \to \sum_{i=0}^{M-1} J_{\bar\mu(\cdot)}^{t_i, t_{i+1}}\big(\mu(\cdot);\, g_i\big).$$

This combined with the right continuity implies that for any $\eta > 0$ and any $t_i < t_{i+1}$, there exists an $\varepsilon > 0$ such that

$$\sum_{i=0}^{M-1} \sup_{g \in C_k^{1,0}([t_i, t_{i+1}] \times E)} \left\{ J_{\bar\mu(\cdot)}^{t_i, t_{i+1}}\big(\mu(\cdot); g\big) \right\} \le J_{\bar\mu(\cdot)}\big(\mu(\cdot);\, \tilde g^\varepsilon\big) + \eta$$

$$\le \sup_{g \in C_k^{1,0}([0,T] \times E)} \left\{ J_{\bar\mu(\cdot)}\big(\mu(\cdot);\, g\big) \right\} + \eta.$$

This combined with (4.1) and (4.20) implies (4.15). □

Now we turn to prove the lower bound for the nonindependent case. The main result of this subsection is the following lemma.

LEMMA 4.6. *For every $\bar\mu(\cdot) \in \mathcal{D}_\infty$ and any open neighborhood V of $\bar\mu(\cdot)$, we have*

$$\liminf_{N \to \infty} \frac{1}{N} \log \mathcal{P}_u^{(N)}(V) \ge -S_u\big(\bar\mu(\cdot)\big). \tag{4.21}$$

PROOF. Let $\bar\mu(\cdot)$ be any element of $\mathcal{D}_\infty$ and let V be an open neighborhood of it. If $S_u(\bar\mu(\cdot)) = \infty$, the result is obvious. Now we assume $S_u(\bar\mu(\cdot)) < \infty$. For any $p > 1$, $p' > 3$, $1/p + 1/p' = 1$ and $\mu(\cdot) \in \mathcal{D}_\infty$, let

$$K_T^{(N)}\big(x^{(N)}(\cdot), \mu(\cdot)\big) = \sum_{i=1}^N \int_{[0,T] \times E} \log q_{\bar\mu(s-)}^{\mu(s-)}\big(x_i^{(N)}(s-);\, y\big) N\big(ds, dy;\, x_i^{(N)}(\cdot)\big),$$

$$G_T^{(N)}\big(\mu(\cdot)\big) = -N \int_0^T \big\{ \|\mu(s)\| - \|\bar\mu(s)\| \big\}\, ds,$$

$$L_{T,p}^{(N)}\big(x^{(N)}(\cdot), \mu(\cdot)\big) = \sum_{i=1}^N \int_{[0,T] \times E} \Big[\Big(q_{\bar\mu(s-)}^{\mu(s-)}\big(x_i^{(N)}(s-), y\big) \Big)^{-p'/p} - 1 \Big]$$

$$\times\, Q_{\bar\mu(s-)}\big(x_i^{(N)}(s-), dy\big)\, ds.$$

Choose m large enough such that $\overline{\mu} \in \mathcal{D}_m$ and

$$\lim_{N\to\infty} \frac{1}{N} \log \mathcal{P}^{(N)}_{\overline{\mu}(\cdot),u}(\mathcal{D}^c_m) \le -2S_u(\overline{\mu}(\cdot)).$$

Then for any $\delta > 0$, we can find an open neighborhood W_1 of $\overline{\mu}$ such that $W_1 \subset V$ and

$$|G^{(N)}_T(\mu(\cdot))| \le \delta N \quad \text{uniformly on } W_1 \cap \mathcal{D}_m. \tag{4.22}$$

Let $M_\varphi(E) = \{u \in M_1(E)\colon \langle u, \varphi\rangle < \infty\}$ equipped with the weak topology. Then we have that for any $\nu \in M_\varphi(E)$, $y \neq x+1$, $[(q^\mu_\nu(x,y))^{-p'/p} - 1]Q_\nu(x,y) = 0$, whereas for $y = x+1$ it is a continuous function with respect to μ uniformly in x on $M_\varphi(E)$. This implies that there is an open neighborhood $W_2 \subset V$ such that

$$|L^{(N)}_{T,p}(x^{(N)}(\cdot), \mu(\cdot))| \le N\delta r(p,T) \quad \text{uniformly on } W_2 \cap \mathcal{D}_m, \tag{4.23}$$

where $r(p)$ is a constant associated with p.

All these combined with Theorem 2.2 imply that for $W = W_1 \cap W_2$,

$$\begin{aligned}\mathcal{P}^{(N)}_u(V) &\ge \mathcal{P}^{(N)}_u(W \cap \mathcal{D}_m)\\ &\ge E^{P^{(N)}_{\bar\mu,u}}\Big[\exp\Big(K^{(N)}_T(x^{(N)}(\cdot), \varepsilon_{x^{(N)}(\cdot)}) + G^{(N)}_T(\varepsilon_{x^{(N)}(\cdot)})\Big);\ \varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m\Big]\\ &\ge \exp(-N\delta) E^{P^{(N)}_{\bar\mu,u}}\Big[\exp\Big(K^{(N)}_T(x^{(N)}(\cdot), \varepsilon_{x^{(N)}(\cdot)})\Big);\ \varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m\Big].\end{aligned} \tag{4.24}$$

By applying Hölder's inequality, we get

$$\begin{aligned}&E^{P^{(N)}_{\bar\mu,u}}\Big[\exp\Big(K^{(N)}_T(x^{(N)}(\cdot), \varepsilon_{x^{(N)}(\cdot)})\Big); \varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m\Big]\\ &\quad\ge E^{P^{(N)}_{\bar\mu,u}}\Big[\exp\Big(-\frac{p'}{p}K^{(N)}_T(x^{(N)}(\cdot), \varepsilon_{x^{(N)}(\cdot)}) - L^{(N)}_{T,p}(x^{(N)}(\cdot), \varepsilon_{x^{(N)}(\cdot)})\\ &\qquad + L^{(N)}_{T,p}(x^{(N)}(\cdot), \varepsilon_{x^{(N)}(\cdot)})\Big)\Big]^{-p/p'} \times \Big[P^{(N)}_{\bar\mu,u}(\varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m)\Big]^p.\end{aligned} \tag{4.25}$$

Then (4.24) combined with (4.25) implies

$$\begin{aligned}\mathcal{P}^{(N)}_u(V) &\ge \mathcal{P}^{(N)}_u(W \cap \mathcal{D}_m)\\ &\ge \exp(-N\delta)\exp(-Nr(p,T)\delta p/p')\Big[P^{(N)}_{\bar\mu,u}(\varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m)\Big]^p\\ &\quad\times E^{P^{(N)}_{\bar\mu,u}}\Big[\exp\Big(-p'/pK^{(N)}_T(x^{(N)}(\cdot), \varepsilon_{x^{(N)}(\cdot)})\\ &\qquad - L^{(N)}_{T,p}(x^{(N)}(\cdot), \varepsilon_{x^{(N)}(\cdot)})\Big)\Big]^{-p/p'}.\end{aligned} \tag{4.26}$$

Let

$$Y_t\big(x^{(N)}(\cdot)\big) = \exp\Big(-p'/pK_t^{(N)}\big(x^{(N)}(\cdot), \varepsilon_{x^{(N)}(\cdot)}\big) - L_{t,p}^{(N)}\big(x^{(N)}(\cdot), \varepsilon_{x^{(N)}(\cdot)}\big)\Big).$$

Because $\{x_i^{(N)}(t)\}_{i=1}^N$ have no common jumps $P_u^{(N)}$-almost surely, by a local procedure and Itô's formula (cf. [11], page 66), we have that $Y_t(x^{(N)}(\cdot))$ is a $P_{\bar{\mu},u}^{(N)}$-local martingale and thus a supermartingale. This implies

$$\frac{1}{N}\log \mathcal{P}_u^{(N)}(V) \geq -\big(1 + r(p,T)p/p'\big)\delta + \frac{p}{N}\log P_{\bar{\mu},u}^{(N)}\big(\varepsilon_{x^{(N)}(\cdot)} \in W \cap \mathcal{D}_m\big). \tag{4.27}$$

By Lemmas 4.4 and 4.5, we get

$$\frac{1}{N}\log \mathcal{P}_{\bar{\mu},u}^{(N)}(W) \geq -S_u(\bar{\mu}). \tag{4.28}$$

Letting $\delta \to 0$ and then letting $p \to 1$, we finally get (4.21), which is just the lower bound. □

4.2. *Compactness of level sets.* In this subsection we will prove the compactness of the level sets by using the lower bound.

LEMMA 4.7. *Let us assume* (2.1)–(2.5) *and* (2.14). *Then for any* $\gamma > 0$, *the level set* $\Phi_u(\gamma) = \{\mu(\cdot) \in \mathcal{D}_\infty;\ S_u(\mu(\cdot)) \leq \gamma\}$ *is compact in* $\mathcal{D}_\infty$.

PROOF. Let $\gamma > 0$ and $\Phi_u(\gamma)$ be given. Because $C([0,T],\ M_1(E)) \cap \mathcal{D}_\infty$ is a closed subset of $\mathcal{D}_\infty$ in the "inductive topology," to prove the lemma it suffices to prove the compactness of $\Phi_u(\gamma)$ in $C([0,T],\ M_1(E)) \cap \mathcal{D}_\infty$. In order to do this, it suffices to verify that

$$\sup_{\mu(\cdot) \in \Phi_u(\gamma)} \sup_{0 \leq t \leq T} \langle \mu(t), \varphi\rangle < \infty \tag{4.29}$$

and $\Phi_u(\gamma)$ is compact in $C([0,T],\ M_1(E))$ in the uniform topology.

For every $n \geq 1$, let $\varphi_n(x) = \varphi(x \wedge n)$. Then $\{\varphi_n\}$ is a sequence of bounded continuous functions on E. By the definition of $S_u(\mu(\cdot))$, we get

$$\sup_{\substack{\mu(\cdot) \in \Phi_u(\gamma),\\ n \geq 1}} \Big\{\big\langle \mu(0), \varphi_n\big\rangle - \log\langle u, e^{\varphi_n}\rangle\Big\} \leq \gamma. \tag{4.30}$$

This combined with (2.14) implies

$$\sup_{\substack{\mu(\cdot) \in \Phi_u(\gamma),\\ n \geq 1}} \Big\{\big\langle \mu(0), \varphi_n\big\rangle\Big\} \leq \eta + \gamma. \tag{4.31}$$

By Fatou's lemma we get

$$\sup_{\mu(\cdot)\in\Phi_u(\gamma)}\{\langle\mu(0),\varphi\rangle\}\le\eta+\gamma. \tag{4.32}$$

For every $n\ge 1$, let $f_n(x)=\varphi(x)I_{E_n}\in C_k(E)$. By (2.5), we have

$$\exp\bigl(-f_n(x)\bigr)Q_{\mu(t)}\exp\bigl(f_n(x)\bigr)\le\lambda(0)+\|\mu(t)\|(e-1),\quad \text{for any } n\ge 1. \tag{4.33}$$

On the other hand, by (4.12) we have

$$\begin{aligned}\sup_{\substack{\mu(\cdot)\in\Phi_u(\gamma),\\ n\ge 1}}&\Bigl\{J^{0,t}_{\mu(\cdot)}(\mu(\cdot);f_n)\Bigr\}\\ &=\sup_{\substack{\mu(\cdot)\in\Phi_u(\gamma),\\ n\ge 1}}\biggl\{\langle\mu(t),f_n\rangle-\langle\mu(0),f_n\rangle\\ &\qquad-\int_0^t\Bigl\langle\mu(s),\exp\bigl(-f_n(x)\bigr)Q_{\mu(s)}\exp\bigl(f_n(x)\bigr)\Bigr\rangle\,ds\biggr\}\le\gamma.\end{aligned} \tag{4.34}$$

Fatou's lemma combined with (4.32)–(4.34) implies that for any $t\in[0,T]$,

$$\sup_{\mu(\cdot)\in\Phi_u(\gamma)}\Bigl\{\bigl\langle\mu(t),\varphi\bigr\rangle\Bigr\}\le\lambda(0)T+\eta+\gamma+2\int_0^t\sup_{\mu(\cdot)\in\Phi_u(\gamma)}\langle\mu(s),\varphi\rangle\,ds. \tag{4.35}$$

Let $R_0=(\lambda(0)T+\eta+\gamma)e^{2T}$. Then by Gronwall's lemma we conclude that

$$\Phi_u(\gamma)\subset\mathcal{D}_{R_0}. \tag{4.36}$$

By Lemma 4.3, $\Phi_u(\gamma)\subset C([0,T],\ M_1(E))$. To finish the proof, it suffices to show the relative compactness of $\Phi_u(\gamma)$ in $C([0,T],\ M_1(E))$. By Lemma 1.3 of [10], it suffices to show the compactness of $\{\langle\mu(\cdot),f\rangle;\mu(\cdot)\in\Phi_u(\gamma)\}$ in $C([0,T],(-\infty,+\infty))$ for every $f\in C_k(E)$. Let $f\in C_k(E)$ be arbitrarily given. Clearly, we have $\sup_{0\le t\le T}|\langle\mu(t),f\rangle|<\infty$. By Ascoli's theorem, it remains to prove the equicontinuity of $\{t\mapsto\langle\mu(t),f\rangle:\ \mu(\cdot)\in\Phi_u(\gamma)\}$. For any $\mu(\cdot)\in\mathcal{D}_\infty$, $\delta>0, p>0$, let

$$D^f_{\delta,p}=\left\{\mu(\cdot)\in\mathcal{D}_\infty;\ \sup_{\substack{s,t\in[0,T],\\ t-s<\delta}}\Bigl|\bigl\langle\mu(t),f\bigr\rangle-\bigl\langle\mu(s),f\bigr\rangle\Bigr|\le p\right\}. \tag{4.37}$$

For fixed γ we can find an $R\ge R_0$ by Lemma 2.3 such that

$$\limsup_{N\to 0}\frac{1}{N}\log\mathcal{P}^{(N)}_u\bigl(\mathcal{D}^c_R\bigr)\le-2\gamma. \tag{4.38}$$

Applying the Markov property, we derive for any $s > 0$, $\delta \in [0, T/2]$ and $p > 0$ the estimates (cf. [1], page 56)

$$
\begin{aligned}
&\mathcal{P}_u^{(N)}\Big(\mathcal{D}_R \cap (D^f_{\delta,p})^c\Big) \\
&\quad \le \mathcal{P}_u^{(N)}\left(\sup_{\substack{s,t\in[0,T],\\ t-s<\delta}} \big|\langle \mu(t), f\rangle - \langle \mu(s), f\rangle\big| > p,\ \mu(\cdot)\in \mathcal{D}_R \right) \\
&\quad \le \sum_{l=0}^{[T/\delta]-1} P_u^{(N)}\left(\sup_{l\delta\le t\le ((l+2)\delta)\wedge T} \left|\frac{1}{N}\sum_{i=1}^N \Big(f\big(x_i^{(N)}(t)\big) - f\big(x_i^{(N)}(l\delta)\big)\Big)\right| \right. \\
&\qquad\qquad\qquad \left. > \frac{p}{2}, \varepsilon_{x^{(N)}(\cdot)} \in \mathcal{D}_R \right) \\
&\quad \le \frac{T}{\delta} \sup_{\nu\in M_\varphi(E)} P_\nu^{(N)}\left(\sup_{0\le t\le 2\delta} \left|\frac{1}{N}\sum_{i=1}^N \Big(f\big(x_i^{(N)}(t)\big) - f\big(x_i^{(N)}(0)\big)\Big)\right| \right. \\
&\qquad\qquad\qquad \left. > \frac{p}{2},\ \varepsilon_{x^{(N)}(\cdot)} \in \mathcal{D}_R \right).
\end{aligned}
\tag{4.39}
$$

For any $x^{(N)}(\cdot) \in D([0,T],\ E^{\otimes N})$, let

$$
\begin{aligned}
K^N\big(f,t;x^{(N)}(\cdot)\big) = \sum_{i=1}^N \bigg\{ & f\big(x_i^{(N)}(t)\big) - f\big(x_i^{(N)}(0)\big) \\
& - \int_0^t \exp\Big(-f\big(x_i^{(N)}(s)\big)\Big) Q_{\varepsilon_{x^{(N)}(s)}} \exp\Big(f\big(x_i^{(N)}(s)\big)\Big)\, dr \bigg\}.
\end{aligned}
\tag{4.40}
$$

Then for every $\beta > 0$, $\nu \in M_\varphi(E)$, $\exp(K^N(\beta f, t;\, x^{(N)}(\cdot)))$ is a $P_\nu^{(N)}$-martingale. (4.39) combined with (4.40) and Doob's inequality implies

$$
\begin{aligned}
&P_u^{(N)}\left(\sup_{0\le t\le 2\delta} \frac{1}{N}\sum_{i=1}^N \Big(f\big(x_i^{(N)}(t)\big) - f\big(x_i^{(N)}(0)\big)\Big) > \frac{p}{2},\ \varepsilon_{x^{(N)}(\cdot)} \in \mathcal{D}_R \right) \\
&\quad \le \sup_{\nu\in M_\varphi(E)} P_\nu^{(N)} \left\{ \sup_{0\le t\le 2\delta} K^N(\beta f, t;\ \omega) > N\bigg(\frac{\beta p}{2} - 2c(f,R)\delta \right. \\
&\qquad\qquad \left. \times \Big(\exp(2\beta\|f\|) - 1\Big)\bigg), \varepsilon_{x^{(N)}(\cdot)} \in \mathcal{D}_R \right\} \\
&\quad \le \exp\left[-N\left\{ \frac{\beta p}{2} - 2c(f,R)\delta\Big(\exp(2\beta\|f\|) - 1\Big) \right\} \right],
\end{aligned}
\tag{4.41}
$$

where $c(f,R) = \sup\{\sum_{y\in E} q_{x,y};\ x \in \operatorname{supp}(f)\} + R$. This implies

$$
\begin{aligned}
&\mathcal{P}_u^{(N)}\Big(\mathcal{D}_R \cap (D^f_{\delta,p})^c\Big) \\
&\qquad \le \frac{2T}{\delta} \exp\left[-N\left\{ \frac{\beta p}{2} - 2c(f,R)\delta\Big(\exp(2\beta\|f\|) - 1\Big) \right\} \right].
\end{aligned}
\tag{4.42}
$$

Minimizing the expression on the right of (4.42) with respect to $\beta > 0$, we have for $[p/(8c(f,R)\|f\|\delta)] > 1$,

$$\mathcal{P}_u^{(N)}\left(\mathcal{D}_R \cap (D^f_{\delta,p})^c\right) \le \frac{2T}{\delta}\exp\left[-N\left\{\frac{p}{4\|f\|}\left(\log\left(\frac{p}{8c(f,R)\|f\|\delta}\right)-1\right)+2c(f,R)\delta\right\}\right] \tag{4.43}$$

For the given γ, let $n_0 = [1/4\gamma\|f\|]$. If we choose $p_n = 1/n$, $\delta_n = 1/(8\|f\|c(f,R)n)$ $e^{-12\|f\|\gamma n}$, then we have

$$\begin{aligned}\limsup_{N\to\infty}\frac{1}{N}\log\mathcal{P}_u^{(N)}&\left(\mathcal{D}_R \cap (D^f_{\delta_n,p_n})^c\right)\\ &\le \limsup_{N\to\infty}\frac{1}{N}\log\frac{2T}{\delta_n}\exp\left[-N\left\{\frac{p_n}{4\|f\|}\left(\log\left(\frac{p_n}{8c(f,R)\|f\|\delta_n}\right)-1\right)\right.\right.\\ &\qquad\left.\left.+\,2c(f,R)\delta_n\right\}\right]\\ &\le -3\gamma+\frac{1}{4\|f\|n}\le -2\gamma, \qquad \forall n > n_0.\end{aligned} \tag{4.44}$$

This combined with (4.38) implies

$$\begin{aligned}\limsup_{N\to\infty}\frac{1}{N}\log\mathcal{P}_u^{(N)}&\left\{(D^f_{\delta_n,p_n})^c\right\}\\ &\le \max\left\{\limsup_{N\to\infty}\frac{1}{N}\log\mathcal{P}_u^{(N)}\{\mathcal{D}_R^c\},\right.\\ &\qquad\left.\limsup_{N\to\infty}\frac{1}{N}\log\mathcal{P}_u^{(N)}\left\{\mathcal{D}_R\cap(D^f_{\delta_n,p_n})^c\right\}\right\}\\ &\le -2\gamma.\end{aligned} \tag{4.45}$$

Because $D^f_{\delta_n,p_n}$ is a closed subset of $\mathcal{D}_\infty$, we have by (4.21),

$$\inf_{\mu(\cdot)\in(D^f_{\delta_n,p_n})^c} S_u(\mu(\cdot)) > \gamma. \tag{4.46}$$

Thus for any $n > n_0$, we have

$$\Phi_u(\gamma) \subset D^f_{\delta_n,p_n}. \tag{4.47}$$

This means that

$$\Phi_u(\gamma) \subset C([0,T],\, M_1(E)) \bigcap_{n>n_0} D^f_{\delta_n,p_n}. \tag{4.48}$$

Thus we obtain the relative compactness of $\Phi_u(\gamma)$. Because $S_u(\cdot)$ is lower semicontinuous we, in fact, obtain the compactness of $\Phi_u(\gamma)$. □

4.3. *Upper bound.* In this subsection we obtain the large deviation upper bound. The method we use is a generalization of the method used by Comets [2] in proving the large deviation principle for the Curie–Weiss model on the torus.

For every $g \in C_b(E)$, $f \in C_k^{1,0}([0,T] \times E)$, $u \in M_1(E)$, $x^{(N)}(\cdot) \in D([0,T], E^{\otimes N})$ and $\mu(\cdot) \in \mathcal{D}_\infty$, let

$$H_u(\mu(\cdot); f, g, T) = l_u(\mu(0), g) + J(\mu(\cdot); f, T),$$

$$B(t, x^{(N)}(\cdot)) = \sum_{i=1}^{N} \int_0^t \Bigg(\frac{\partial f(s, x_i^{(N)}(s))}{\partial s} + \exp\Big(-f(s, x_i^{(N)}(s))\Big) Q_{\varepsilon_{x^{(N)}(s)}} \exp\Big(f(s, x_i^{(N)}(s))\Big) \Bigg) ds,$$

$$F(t, x^{(N)}(t)) = \exp\left[\sum_i^N \Big(f(t, x_i^{(N)}(t)) - f(0, x_i^{(N)}(0)) \Big) - B(t, x^{(N)}(\cdot)) \right].$$

By Itô's formula we have

$$\begin{aligned}
&F(t, x^{(N)}(t)) - 1 \\
&\quad = \int_0^t F_s'(s, x^{(N)}(s))\, ds - \int_0^t F(s, x^{(N)}(s))\, dB(s, x^{(N)}(\cdot)) \\
&\qquad + \int_0^{t+} \int_E F(s-, x^{(N)}(s-)) \sum_{i=1}^N \Big[\exp\Big(f(s-, y) - f(s-, x_i^{(N)}(s-))\Big) - 1 \Big] \\
&\qquad \times N(ds, dy;\, x_i^{(N)}) \\
&\quad = \int_0^t F(s, x^{(N)}(s)) \sum_{i=1}^N \frac{\partial f(s, x^{(N)}(s))}{\partial s}\, ds - \int_0^t F(s, x^{(N)}(s))\, dB(s, x^{(N)}(\cdot)) \\
&\qquad + \int_0^{t+} \int_E F(s-, x^{(N)}(s-)) \sum_{i=1}^N \Big[\exp\Big(\big(f(s-, y) - f(s-, x_i^{(N)}(s-)) \big) \Big) - 1 \Big] \\
&\qquad \times \widetilde{N}(ds, dy;\, x_i^{(N)}) \\
&\qquad + \int_0^t \sum_{i=1}^N \exp\Big(-f(s-, x_i^{(N)}(s-)) \Big) Q_{\varepsilon_{x^{(N)}(s)}} \exp\Big(f(s-, x_i^{(N)}(s-)) \Big)\, ds \\
&\quad = \int_0^{t+} \int_E F(s-, x^{(N)}(s-)) \sum_{i=1}^N \Big[\exp\Big(\big(f(s-, y) - f(s-, x_i^{(N)}(s-)) \big) \Big) - 1 \Big] \\
&\qquad \times \widetilde{N}(ds, dy;\, x_i^{(N)}).
\end{aligned}$$

Thus $F(t, x^{(N)}(t))$ is a martingale with respect to $P_u^{(N)}$.

Let $Y_u^{(N)}(\mu(\cdot); f, g, T) = \exp[N H_u(\mu(\cdot); f, g, T)]$. Then we have the following lemma.

LEMMA 4.8. $E^{\mathcal{P}_u^{(N)}}[Y_u^{(N)}(\mu(\cdot); f,g,T)] = E^{P_u^{(N)}}[Y_u^{(N)}(\varepsilon_{x^{(N)}(\cdot)}; f,g,T)] = 1.$

PROOF. By direct calculation we have $E^{P_u^{(N)}} \exp[Nl_u(\varepsilon_{x^{(N)}(0)},g)] = 1$. This combined with the fact that $F(t,x^{(N)}(t))$ is a martingale with respect to $P_u^{(N)}$ implies the result. □

The next theorem is crucial for obtaining the upper bound.

THEOREM 4.9. *Assume* (2.1)–(2.4) *and* (2.14) *are satisfied. For all* $\gamma > 0$, $\varepsilon > 0$, $s \geq 0$, *there exists an integer* N_0 *such that for all* $N \geq N_0$,

$$\mathcal{P}_u^{(N)}\Big\{r_{0T}\big(\mu(\cdot), \Phi_u(s)\big) \geq \varepsilon\Big\} \leq \exp\big[-N(s-\gamma)\big], \tag{4.49}$$

where $r_{0T}\big(\mu(\cdot), \Phi_u(s)\big) = \inf_{\nu(\cdot)\in\Phi_u(s)} \sup_{0\leq t\leq T} r\big(\mu(t), \nu(t)\big)$.

In order to prove this theorem we need the following two results:

LEMMA 4.10. *For any* $\overline{\mu}(\cdot) \in \mathcal{D}_\infty$ *and* $I < S_u(\overline{\mu}(\cdot))$ *there exists* $\delta' > 0$ *and an integer* N_1 *such that for all* $N \geq N_1$,

$$\mathcal{P}_u^{(N)}\Big\{r_{0T}\big(\mu(\cdot), \overline{\mu}(\cdot)\big) < \delta'\Big\} \leq \exp[-NI]. \tag{4.50}$$

PROOF. For any $\overline{\mu}(\cdot) \in \mathcal{D}_\infty'$ and $I < S_u(\overline{\mu}(\cdot))$, there exists $\gamma > 0$ such that $I + \gamma < S_u(\overline{\mu}(\cdot))$. Assume $S_u(\overline{\mu}(\cdot)) < \infty$. By Lemma 2.3, there exists $R_0 \geq 1$ such that

$$\mathcal{P}_u^{(N)}\Big\{\big(\mathcal{D}_{R_0}\big)^c\Big\} \leq \exp\big[-N(I+\gamma)\big]. \tag{4.51}$$

By the definition of the action functional, there exist $f \in C_k^{1,0}([0,T] \times E)$, $g \in C_b(E)$, such that

$$S_u\big(\overline{\mu}(\cdot)\big) \leq H_u\big(\overline{\mu}(\cdot); f,g,T\big) + \gamma/3.$$

Choose $R \geq R_0$ such that $\overline{\mu}(\cdot) \in \mathcal{D}_R$.

Because $H_u(\mu(\cdot); f,g,T)$ is continuous at $\overline{\mu}(\cdot)$ on $\mathcal{D}_R$ in the uniform topology, there exists an $\delta' > 0$ such that

$$H_u\big(\overline{\mu}(\cdot); f,g,T\big) \leq \inf_{\mu(\cdot)\in V} H_u\big(\mu(\cdot); f,g,T\big) + \gamma/3,$$

where $V = \{\mu(\cdot) \in \mathcal{D}_R;\ r_{0T}(\mu(\cdot), \overline{\mu}(\cdot)) < \delta'\}$.

By Chebyshev's inequality and Lemma 4.8, we get

$$\begin{aligned}\mathcal{P}_u^{(N)}(V) &\leq \mathcal{P}_u^{(N)}\Big\{\mu(\cdot);\ Y_u^{(N)}\big(\mu(\cdot); f,g,T\big) \in Y_u^{(N)}(V; f,g,T)\Big\}\\ &\leq \frac{\int_{\mathcal{D}_\infty} Y_u^{(N)}\big(\mu(\cdot); f,g,T\big)\mathcal{P}_u^{(N)}(d\mu(\cdot))}{\inf_{\mu(\cdot)\in V} Y_u^{(N)}\big(\mu(\cdot); f,g,T\big)} = \frac{1}{\inf_{\mu(\cdot)\in V} Y_u^{(N)}\big(\mu(\cdot); f,g,T\big)}.\end{aligned}$$

Thus

$$\begin{aligned}
\mathcal{P}_u^{(N)}(V) &\le \exp\Big[-N \inf_{\mu(\cdot)\in V} H_u\big(\mu(\cdot); f, g, T\big)\Big] \\
&\le \exp\Big[-N\Big(H_u\big(\overline{\mu}(\cdot); f, g, T\big) - \gamma/3\Big)\Big] \\
&\le \exp\Big[-N\Big(S_u\big(\overline{\mu}(\cdot)\big) - 2\gamma/3\Big)\Big] \\
&\le \exp\big[-N(I + \gamma/3)\big].
\end{aligned} \tag{4.52}$$

This combined with (4.51) implies (4.50). The case of $S_u(\overline{\mu}(\cdot)) = \infty$ can be proved similarly. □

LEMMA 4.11. *For all $a \ge 0$, there exists an $R \ge 1$ and a compact subset $K_a \subset \mathcal{D}_R$ in both the uniform and the Skorohod topology such that for every $\delta > 0$, there exists $N_2 \ge 1$ such that for all $N \ge N_2$,*

$$\mathcal{P}_u^{(N)}\Big\{r_{0T}\big(\mu(\cdot), K_a\big) \ge \delta\Big\} \le \exp[-Na]. \tag{4.53}$$

PROOF. If $a = 0$, the result is trivial. Suppose $a > 0$. By Lemma 2.3, there exists $R \ge 1$ such that

$$\mathcal{P}_u^{(N)}\Big\{\big(\mathcal{D}_R\big)^c\Big\} \le \exp\big[-N(a+1)\big]. \tag{4.54}$$

For fixed a and R we construct a sequence $\{\Delta_j\}_{j=1,\dots}$ satisfying

$$\begin{aligned}
&T/\Delta_j \in \mathbb{N}, \qquad \Delta_j/\Delta_{j+1} \in \mathbb{N}\backslash\{0,1\}, \\
&\Delta_j \le (a+1)/2\widehat{c}(j,R)(\exp[16(a+1)j] - 1), \\
&j\Delta_j \text{ is decreasing},
\end{aligned}$$

where $\mathbb{N}$ denotes the set of integers and $\widehat{c}(j,R)$ is defined later and is a positive increasing function of j.

Let K_a' be defined as

$$K_a' = \left\{\mu(\cdot) \in \mathcal{D}_R;\ \sup_{\substack{t,s\in[0,T],\\ |t-s|\le\Delta_j}} r\big(\mu(t), \mu(s)\big) < \frac{1}{j},\ j = 1, 2, \dots\right\}.$$

By definition and the Ascoli–Arzela theorem, we have that K_a' is a relatively compact subset of the space $\mathcal{D}_R \cap C([0,T], M_1(E))$. Let K_a be the closure of K_a' in $\mathcal{D}_R \cap C([0,T], M_1(E))$. Then K_a is a compact subset of $\mathcal{D}_R \cap C([0,T], M_1(E))$. Because $\mathcal{D}_R \cap C([0,T], M_1(E))$ is a closed subset of $\mathcal{D}_R$ in the uniform topology, we have that K_a is a compact subset of $\mathcal{D}_R$ in both the uniform and Skorohod topology. For any $\delta > 0$ choose j_0 such that $1/j_0 < \delta$. Let

$$D(j_0) = \left\{\mu(\cdot) \in \mathcal{D}_R;\ \sup_{\substack{t,s\in[0,T],\\ |t-s|\le\Delta_{j_0}}} r\big(\mu(t), \mu(s)\big) < \frac{1}{j_0}\right\}.$$

For any positive integer j, let $n(j) \in \mathbb{N}$ be defined as

$$n(j) = \inf\left\{m;\ \sum_{n>m} \frac{1}{2^n} < \frac{1}{2j}\right\}.$$

By an argument similar to the estimates (4.39) and (4.42), we have for any $\beta > 0$,

$$\begin{aligned}
&\mathcal{P}_u^{(N)}\big(\mathcal{D}_R \cap D(j_0)^c\big)\\
&\quad\le \sum_{k=0}^{n(j_0)} \sum_{l=0}^{[T/\Delta_{j_0}]-1} P_u^{(N)}\Bigg\{\varepsilon_{x^{(N)}(\cdot)} \in \mathcal{D}_R,\ \sup_{l\Delta_{j_0} \le t \le ((l+2)\Delta_{j_0})\wedge T}\\
&\qquad\qquad \times \left|\frac{1}{N}\sum_{i=1}^{N}\Big(I_{\{k\}}\big(x_i^{(N)}(t)\big) - I_{\{k\}}\big(x_i^{(N)}(l\Delta_{j_0})\big)\Big)\right| \ge \frac{1}{4j_0}\Bigg\} \qquad (4.55)\\
&\quad\le \frac{2T}{\Delta_{j_0}} \sum_{k=0}^{n(j_0)} \exp\left[-N\left(\frac{\beta}{4j_0} - 2c(I_{\{k\}}, R)\Delta_{j_0}\big(\exp^{(2\beta)} - 1\big)\right)\right]\\
&\quad\le \frac{2T}{\Delta_{j_0}} n(j_0) \exp\left[-N\left(\frac{\beta}{4j_0} - 2\widehat{c}(j_0, R)\Delta_{j_0}\big(e^{2\beta} - 1\big)\right)\right],
\end{aligned}$$

where $\widehat{c}(j_0, R) = \sup_{k \le n(j_0)} c(I_{\{k\}}, R)$. Let $\beta = 8j_0(a+1)$. Then by the construction of Δ_j, we conclude that

$$\mathcal{P}_u^{(N)}\big(\mathcal{D}_R \cap D(j_0)^c\big) \le \frac{2T}{\Delta_{j_0}} n(j_0) \exp\big[-N(a+1)\big]. \qquad (4.56)$$

Now let $m_0 = T/\Delta_{j_0}$, $t_k = k\Delta_{j_0}$, $k = 1, \dots, m_0$. For any $\mu(\cdot) \in \mathcal{D}_\infty$, we define $\mu^0(\cdot)$ as follow:

$$\mu^0(t) = \begin{cases} \mu(t_k), & \text{if } t = t_k, \text{ for } 1 \le k \le m_0,\\ \text{linear}, & \text{otherwise.}\end{cases} \qquad (4.57)$$

Obviously $r_{0T}(\mu(\cdot), \mu^0(\cdot)) < \delta$ on $D(j_0)$. Thus

$$\begin{aligned}
&\mathcal{P}_u^{(N)}\Big\{\mu(\cdot) \in D(j_0);\ r_{0T}\big(\mu(\cdot), K_a\big) \ge \delta\Big\}\\
&\qquad \le \mathcal{P}_u^{(N)}\Big\{D(j_0) \cap \big\{\mu^0(\cdot) \notin K_a'\big\}\Big\}.
\end{aligned} \qquad (4.58)$$

Using an argument similar to that in Comets ([2], pages 28 and 29), we can verify that on $D(j_0)$, the slope of $\mu^0(\cdot)$ satisfies

$$\begin{aligned}
\frac{r\big(\mu^0(t), \mu^0(s)\big)}{|t-s|} &< (j_0\Delta_{j_0})^{-1}\\
&\le (j\Delta_j)^{-1}, \text{ if } j \ge j_0.
\end{aligned}$$

This implies that for any $j \geq j_0$ and $\mu(\cdot) \in D(j_0)$,

$$\sup_{\substack{t,s \in [0,T],\\ |t-s| \leq \Delta_j}} r\big(\mu^0(t), \mu^0(s)\big) < \frac{1}{j}.$$

For $j < j_0$ we can show

$$\sup_{\substack{t,s \in [0,T],\\ |t-s| \leq \Delta_j}} r\big(\mu^0(t), \mu^0(s)\big) = \max_{|t_k - t_l| \leq \Delta_j} r\big(\mu^0(t_k), \mu^0(t_l)\big), \tag{4.59}$$

where t (resp. s) belongs to some interval $[t_k, t_{k+1}]$ (resp. $[t_l, t_{l+1}]$). For any $\mu,\ \nu \in M_1(E)$, we define $|\mu| = \sum_{n=0}^{\infty} 1/2^n |\mu(n)|$ and $r(\mu, \nu) = |\mu - \nu|$. Then the argument used in [2], (Section V) also works here. To complete the argument note that for $h \in [t_k \vee (t_l + t - s),\ t_{k+1} \wedge (t_{l+1} + t - s)]$, $h \to \mu^0(h) - \mu^0(h + s - t)$ is an affine function, and use the fact that $\mu \to |\mu|$ is a convex function to conclude that $\mu^0(h) - \mu^0(h + s - t)$ achieves its maximum on the boundary of the interval. Thus in order to prove (4.59) it suffices to verify that the $|\mu^0(t) - \mu^0(t_k)|$ is less than or equal to the right-hand side of (4.59) when $t \in (t_l, t_{l+1})$ and $|t - t_k| \leq \Delta_j$. Because Δ_j/Δ_{j_0} is an integer, we get that $|t_l - t_k| \vee |t_{l+1} - t_k| \leq \Delta_j$. This combined with the convexity of $h \to |\mu^0(h) - \mu^0(t_k)|$ implies the result.

Hence we have

$$D(j_0) \cap \{\mu^0(\cdot) \notin K'_a\} \subset \bigcup_{k,l;\, j<j_0} \left\{\mu(\cdot) \in D(j_0);\ r\big(\mu(t_k), \mu(t_l)\big) \geq \frac{1}{j}\right\}, \tag{4.60}$$

where the union extends to all pairs (k,l) with $0 \leq k < l \leq m_0 \wedge (T/\Delta_j + \Delta_j/\Delta_{j_0})$ and $j < j_0$. For any fixed k, l, j, we have

$$\begin{aligned}
&\mathcal{P}_u^{(N)}\left\{\mu(\cdot) \in D(j_0);\ r\big(\mu(t_k), \mu(t_l)\big) \geq \frac{1}{j}\right\} \\
&\quad \leq \sum_{m=0}^{n(j)} P_u^{(N)}\left\{\left|\frac{1}{N}\sum_{i=1}^{N}\Big(I_{\{m\}}\big(x_i^{(N)}(t_k)\big) - I_{\{m\}}\big(x_i^{(N)}(t_l)\big)\Big)\right| \geq \frac{1}{4j}, \varepsilon_{x^{(N)}(\cdot)} \in \mathcal{D}_R\right\} \\
&\quad \leq \sum_{m=0}^{n(j)} \exp\left[-N\left(\frac{\beta}{4j} - 2c(I_{\{m\}}, R)\Delta_j(e^{2\beta} - 1)\right)\right] \\
&\quad \leq n(j) \exp\left[-N\left(\frac{\beta}{2j} - 2\hat{c}(j, R)\Delta_j(e^{2\beta} - 1)\right)\right].
\end{aligned}$$

Let $\beta = 8j(a+1)$. We conclude

$$\mathcal{P}_u^{(N)}\left\{\mu(\cdot) \in D(j_0);\ r\big(\mu(t_k), \mu(t_l)\big) \geq \frac{1}{j}\right\} \leq n(j)\exp\big[-N(a+1)\big].$$

Using the rough upper bound $m_0\Delta_j/\Delta_{j_0}$ on the pairs (k,l) we get

$$\mathcal{P}_u^{(N)}\Big\{D(j_0)\cap\{\mu^0(\cdot)\notin K_a'\}\Big\}\le m_0\left(\sum_{j=2}^{j_0-1}n(j)\Delta_j/\Delta_{j_0}\right)\exp\big[-N(a+1)\big]. \tag{4.61}$$

This combined with (4.54), (4.56) and (4.58) implies

$$\begin{aligned}
&\mathcal{P}_u^{(N)}\Big\{r_{0T}\big(\mu(\cdot),K_a\big)\ge\delta\Big\}\\
&\quad\le\mathcal{P}_u^{(N)}\Big\{\mu(\cdot)\in\mathcal{D}_R;\ r_{0T}\big(\mu(\cdot),K_a\big)\ge\delta\Big\}+\mathcal{P}_u^{(N)}\{(\mathcal{D}_R)^c\}\\
&\quad\le\mathcal{P}_u^{(N)}\{\mathcal{D}_R\cap D(j_0)^c\}+\mathcal{P}_u^{(N)}\{(\mathcal{D}_R)^c\}\\
&\qquad+\mathcal{P}_u^{(N)}\Big\{\mu(\cdot)\in D(j_0);\ r_{0T}\big(\mu(\cdot),K_a\big)\ge\delta\Big\}\\
&\quad\le\frac{2T}{\Delta_{j_0}}n(j_0)\exp\big[-N(a+1)\big]+\exp\big[-N(a+1)\big]\\
&\qquad+m_0\left(\sum_{j=2}^{j_0-1}n(j)\frac{\Delta_j}{\Delta_{j_0}}\right)\exp\big[-N(a+1)\big].
\end{aligned}$$

Because all the coefficients depend only on j_0, and j_0 is independent of N, there exists an integer N_2 such that for all $N\ge N_2$, (4.53) is true. □

THE PROOF OF THEOREM 4.9. For any $\gamma>0$, $\varepsilon>0$ and $s>0$, choose $R\ge 1$ and a compact subset K_s of $\mathcal{D}_R$ from Lemma 4.11, with $a=s$, such that there exists an integer $\widetilde{N}$ for all $N\ge\widetilde{N}$,

$$\mathcal{P}_u^{(N)}\{(\mathcal{D}_R)^c\}\le\exp[-Ns],\qquad \Phi_u(s)\subset\mathcal{D}_R.$$

Let $K_s^\varepsilon=K_s\cap\{r_{0T}(\mu(\cdot),\Phi_u(s))\ge\varepsilon/2\}$. Then K_s^ε is also compact in $\mathcal{D}_R$ in the uniform topology and thus the Skorohod topology.

For every $\overline{\mu}(\cdot)\in K_s^\varepsilon$, by applying Lemma 4.10 with $I=s$, we have that there exists $N_{\bar\mu(\cdot)}$ and $\delta_{\bar\mu(\cdot)}<\varepsilon$ such that for any $N\ge N_{\bar\mu(\cdot)}$,

$$\mathcal{P}_u^{(N)}\Big\{r_{0T}\big(\mu(\cdot),\overline{\mu}(\cdot)\big)<\delta_{\bar\mu(\cdot)}\Big\}\le\exp[-Ns]. \tag{4.62}$$

Let $\{\{\mu(\cdot)\in\mathcal{D}_R;\ r_{0T}(\mu(\cdot),\overline{\mu}_i(\cdot))<\delta_{\bar\mu_i(\cdot)}\};\ i=1,2,\dots,L\}$ be a finite open covering of K_s^ε from all the balls satisfying (4.62), $\delta=\min\{\frac{1}{2}\delta_{\bar\mu_i(\cdot)};\ i=1,2,\dots,L\}$. Then for any $\mu(\cdot)\in\mathcal{D}_R$ satisfying $r_{0T}(\mu(\cdot),K_s)\le\delta$ and $r_{0T}(\mu(\cdot),\Phi_u(s))\ge\varepsilon$ there exists $\widehat{\mu}(\cdot)\in K_s$ such that $r_{0T}(\mu(\cdot),\widehat{\mu}(\cdot))\le\delta$. This implies that $r_{0T}(\widehat{\mu}(\cdot),\Phi_u(s))\ge\varepsilon/2$, that is, $\widehat{\mu}(\cdot)\in K_s^\varepsilon$. Hence

$$\begin{aligned}
&\mathcal{P}_u^{(N)}\Big\{\mu(\cdot)\in\mathcal{D}_R;\ r_{0T}\big(\mu(\cdot),K_s\big)<\delta,\ r_{0T}\big(\mu(\cdot),\Phi_u(s)\big)\ge\varepsilon\Big\}\\
&\quad\le\sum_{i=1}^{L}\mathcal{P}_u^{(N)}\Big\{\mu(\cdot)\in\mathcal{D}_R;\ r_{0T}\big(\mu(\cdot),\overline{\mu}_i(\cdot)\big)<\delta_{\bar\mu_i(\cdot)}\Big\}\le L\exp[-Ns].
\end{aligned} \tag{4.63}$$

Finally we have

$$\begin{aligned}(4.64)\qquad &\mathcal{P}_u^{(N)}\Big\{r_{0T}\big(\mu(\cdot),\Phi_u(s)\big)\geq\varepsilon\Big\}\\ &\quad\leq \mathcal{P}_u^{(N)}\{(\mathcal{D}_R)^c\}+\mathcal{P}_u^{(N)}\Big\{\mu(\cdot)\in\mathcal{D}_R;\ r_{0T}\big(\mu(\cdot),K_s\big)\geq\delta\Big\}\\ &\quad\quad+\mathcal{P}_u^{(N)}\Big\{\mu(\cdot)\in\mathcal{D}_R; r_{0T}\big(\mu(\cdot),K_s\big)<\delta,\ r_{0T}\big(\mu(\cdot),\Phi_u(s)\big)\geq\varepsilon\Big\}\\ &\quad\leq (L+2)\exp[-Ns].\end{aligned}$$

Letting N go to infinity, we get (4.49). □

Now we are ready to prove the main result of this subsection.

THEOREM 4.12 (Upper bound on $\mathcal{D}_\infty$). *Under assumptions* (2.1)–(2.5) *and* (2.14), *for any closed subset A of $\mathcal{D}_\infty$ in the "inductive topology," we have*

$$(4.65)\qquad \limsup_{N\to\infty}\frac{1}{N}\log\mathcal{P}_u^{(N)}(A)\leq-\inf_{\mu(\cdot)\in A}S_u\big(\mu(\cdot)\big).$$

PROOF. Let A be any closed subset of $\mathcal{D}_\infty$. If $\inf_{\mu\in A}S_u(\mu)=0$, then the result is trivial. Now we assume that $s=\inf_{\mu\in A}S_u(\mu)>0$. Choose $R\geq 1$ and $\widetilde{N}$ such that for any $N\geq\widetilde{N}$,

$$(4.66)\qquad \mathcal{P}_u^{(N)}\{(\mathcal{D}_R)^c\}\leq\exp[-Ns],\qquad \Phi_u(s)\subset\mathcal{D}_R.$$

Because for any $s>\gamma>0$, $\Phi_u(s-\gamma)$ is compact in $\mathcal{D}_\infty$ in the uniform topology, we get that $A_1=A\cap\mathcal{D}_R$ and $\Phi_u(s-\gamma)$ are disjoint closed subsets of $\mathcal{D}_R$ in both the Skorohod and the uniform topology. Thus there exists a strictly positive constant $\varepsilon>0$ such that $r_{0T}(A_1,\Phi_u(s-\gamma))\geq\varepsilon$. This implies

$$(4.67)\qquad \mathcal{P}_u^{(N)}\{\mathcal{D}_R\cap A\}\leq\mathcal{P}_u^{(N)}\Big\{r_{0T}\big(\mu(\cdot),\Phi_u(s-\gamma)\big)\geq\varepsilon\Big\}.$$

Because γ is arbitrary, (4.66) combined with (4.67) and Theorem 4.9 implies (4.65). □

Acknowledgment. This result is part of the author's Ph.D. dissertation. He wishes to thank his supervisor, D. A . Dawson, for guidance, encouragement and stimulating discussions. The author also wishes to thank the referee and an Associate Editor for very useful comments and suggestions that improved the exposition.

REFERENCES

[1] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[2] COMETS, F. (1987). Nucleation for a long range magnetic model. *Ann. Inst. H. Poincaré* **23** 137–178.

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

[4] DAWSON, D. A. and ZHENG, X. (1991). Law of large numbers and a central limit theorem for unbounded jump mean-field models. *Adv. in Appl. Math.* **12** 293–326.

[5] DEUSCHEL, J-D. and STROOCK, D.W. (1989). *Large Deviations.* Academic Press, New York.

[6] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York.

[7] FENG, S. (1992). Large deviations for Markov processes with mean field interaction and unbounded jumps. *C. R. Math. Rep. Acad. Sci. Canada* **XIV**(1) 37–42.

[8] FENG, S. and ZHENG, X. (1992). Solutions of a class of nonlinear master equations. *Stochastic Process. Appl.* **43** 65–84.

[9] FREIDLIN, M. I. and WENTZELL, A. D. (1984). *Random Perturbations of Dynamical Systems*. Springer, New York.

[10] GÄRTNER, J. (1988). On the McKean–Vlasov limit for interacting diffusions. *Math. Nachr.* **137** 197–248.

[11] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam.

[12] LÉONARD, C. (1990). On Large deviations for particle systems associated with spatially homogeneous Boltzmann type equations. Preprint.

[13] NICOLIS, G. and PRIGOGINE, I. (1977). *Self-Organization in Non-Equilibrium Systems*. Wiley, New York.

[14] REED, M. and SIMON, B. (1972). *Methods of Modern Mathematical Physics* **1**. Academic Press, New York.

[15] SCHLÖGL, F. (1972). Chemical reaction models for non-equilibrium phase transitions. *Z. Phys.* **253** 147–161.

[16] SUGIURA, M. (1990). Large deviations for Markov processes of jump type with mean-field interactions. Preprint.

DEPARTMENT OF MATHEMATICS AND STATISTICS
MCMASTER UNIVERSITY
1280 MAIN STREET WEST
HAMILTON, ONTARIO L8S 4K1
CANADA