The Annals of Probability
1995, Vol. 23, No. 3, 1414–1438

THE RATE FUNCTION FOR SOME MEASURE-VALUED JUMP PROCESSES

BY BOUALEM DJEHICHE AND INGEMAR KAJ[1]

KTH, Stockholm and Uppsala University

A principle of large deviations from the McKean–Vlasov limit is derived for a class of measure-valued jump processes. It is shown that the associated rate function admits several representations, including the one obtained by entropy methods and the one derived by a pathwise approach.

Introduction. In this study we consider the behavior of some Markovian systems of weakly interacting jump processes as the number of component processes grows to infinity. Our interest focuses on systems which in the limit satisfy a *McKean–Vlasov equation*. This means that the sequence of empirical distribution functions, representing the states of the finite systems, has a deterministic limit of a specific form. The limit, which in the case under consideration is given by the solution of a nonlinear integrodifferential equation, in a sense represents the law of a typical trajectory in the underlying collection of jump processes, subject to the interaction mechanism.

The purpose of this study is to derive the *large deviation principle* in the McKean–Vlasov limit and to obtain several representations of the associated *large deviation rate function*.

The finite systems are of the form $x = (x^1, \dots, x^n) \in E^n$, where to each component corresponds a jump process $t \mapsto x_t^i$, with values in the state space $E \subseteq R^d$. A typical example is that $E = Z_+$, the nonnegative integers, and the x^i represent the number of particles present at each of n reaction sites. Schlögl type models with mean-field interaction, which provide interesting examples of this form, have been studied in depth. Dawson and Zheng (1991) and Feng and Zheng (1992) obtained existence results and studied various properties of the McKean–Vlasov limit, and Feng (1994) considered the large deviation problem.

For another application, think of the x_t^i as some health status indicators, for example, the number of individuals subject to a certain disease, in n possible locations. Léonard (1989, 1990), in this vein, studied epidemic systems with emphasis on limit behavior and large deviations.

Besides the sources already mentioned, our main references are the works by Dawson and Gärtner (1987) and Gärtner (1988) on weakly interacting diffusions. In fact, a motivation for the present work was to unify some results

Received June 1994; revised September 1994.

[1]Research supported by Swedish Natural Sciences Research Council Grant, contract F-FU 09481.

AMS 1991 *subject classifications*. 60J75, 60F10, 60G57, 60K35.

Key words and phrases. McKean–Vlasov limit, large deviations, measure-valued processes, weak interaction, Orlicz space, epidemic process.

obtained by Feng (1994) and Léonard (1989) and possibly extend them in the direction to which the results of Dawson and Gärtner (1987) point.

In Section 1, we present the model. The main results are stated in Section 2. The identification of the rate function is given in Section 3. Then we derive, in Section 4, the large deviation principle for the interaction-free case and, finally, in Section 5 we extend the result to the general case.

1. The jump process.

1.1. *Model.* Let $\mathscr{B}$ denote the Borel sets in R^d and fix a set $E \in \mathscr{B}$. Let $\mathscr{B}_E$ denote the σ-algebra on E generated by the subset topology from $\mathscr{B}$. We consider a system of n components, each a Markovian jump process $t \mapsto x_t^i$, $t \in I := [0, T]$, piecewise constant and with right-continuous trajectories and taking values in E. The component processes $t \mapsto x_t^i$ are given by their characteristics, consisting of a pair (γ_t^i, π^i). Here the jump intensity measure $\gamma_t^i(x)\,dt$ gives the rate of a jump at site i when the system is in state x. The jump size distribution $\pi^i(x, dy)$, $y \in E$, is such that if a jump occurs at site i, then it is of the form $x^i \to y \in B$ with probability $\int_B \pi^i(x, dy)$, for all $B \in \mathscr{B}_E$. Put $n_t^i(x, dy) = \gamma_t^i(x)\,\pi^i(x, dy)$, $x \in E^n$, $y \in E$. We are going to assume throughout that $\pi^i(x, y) = \pi^i(x^i, y)$ does not depend on the state of the system other than as a function of x^i. Consequently, all interaction is guided into the system through the jump rates $\gamma_t^i(x)$. Of course this will be done in a way such that the process $t \mapsto x_t$ is still Markov; see below.

It is convenient to identify the point process system with its empirical distribution and represent a realization of the state at time t with the point measure

$$\mu_t^n = \sum_{i=1}^n \delta_{x_t^{i,n}}.$$

Let $\mathscr{M}$ denote the set of probability measures on E equipped with the topology of weak convergence. We denote the natural duality between $\mathscr{M}$ and the set $\mathscr{C}(E)$ of bounded continuous functions on E by

$$\langle \mu, f \rangle = \int f(y)\,\mu(dy), \qquad f \in \mathscr{C}(E),\ \mu \in \mathscr{M}.$$

Occasionally we consider the subset $\mathscr{C}_0(E)$ of compactly supported functions. Let $\mathscr{D}(I, \mathscr{M})$ denote the path space of cadlag functions from I into $\mathscr{M}$ furnished with the usual Skorokhod topology. The approach to measure-valued jump processes which we adopt here is to study stochastic equations for the *empirical distribution process*

$$t \mapsto X_t^n = \frac{1}{n} \sum_{i=1}^n \delta_{x_t^{i,n}} \in \mathscr{M}, \qquad t \in I.$$

The trajectories are realized in the canonical probability space $[\mathscr{D}(I, \mathscr{M}), (\mathscr{F}_t)_{t \in I}]$, where $\mathscr{F}_t$, $t \in I$, denotes the filtration of σ-algebras generated by the process $t \mapsto x_t$. The limit of $X^n = (X_t^n)_{t\in I}$ as $n \to \infty$ will be an element in the subset $\mathscr{C}(I, \mathscr{M})$ of uniformly continuous paths.

We want to construct next a probability measure $\mathscr{P}^n$ which governs the evolution of the measure-valued interacting jump process X^n.

1.2. *Interaction.* In this section we state the restrictions and assumptions we are going to impose on the jump structure. Of course the *independent case*, when $\gamma_t^i(x) = \gamma_t^i(x^i)$ only depends on x^i and not on x^j, $j \neq i$, is always included. The interaction mechanism is specified as follows.

ASSUMPTION 1. *Assume that the system is weakly interacting in the sense that, for each n,*

$$\gamma_t^i(x) = \gamma_t^i(x^i\,;\, \mu_t^n), \qquad 1 \le i \le n,\ t \in I, \tag{A1}$$

where $\gamma_t^i(\cdot,\cdot)$ is a nonnegative function on $E \times \mathscr{M}$ such that $\mu \mapsto \gamma_t^i(\cdot,\mu)$ is uniformly continuous. Moreover, assume there is a constant C such that, for each n and $1 \le i \le n$,

$$\sup_{x\in E^n} \sup_{t\in I} \int_E n_t^i(x, dy) \le C. \tag{A2}$$

In the sequel we choose whichever is the most convenient notation of

$$n_t^i(x, dy) = n_t^i(x^i, dy\,;\, \mu^n) = \gamma_t^i(x^i\,;\, \mu^n)\, \pi(x^i, dy).$$

Note that since the functions γ^i may include an application of $\langle \mu_t^n, f\rangle$ with $f(y) = y$, the identity function, the assumption covers *mean field interaction* models of the form

$$\gamma_t^i(x) = \gamma_t^i\biggl(x^i\,;\, \frac{1}{n}\sum_{i=1}^n x^{i,n}\biggr).$$

We remark that assumption (A2) requiring bounded jump rates is relatively strong. For example, it rules out the Schlögl type models studied by Feng (1994). However, in order not to burden the presentation with a more complicated technical framework, we have chosen to work first under assumption (A2). Certainly our program works also in other, more general, situations after making the appropriate modifications. The basic line of argument, however, will be the same. See Section 5.4.

1.3. *Martingale problem.* For each set $B \in \mathscr{B}_E$, let $N_t(x^i, B)$ denote the number of jumps of component i during $[0,t]$ which are of the form $x^i \to y \in B$. We will denote by $\mathscr{P}^n$ a distribution of the process which is such that, given any state x of the system and $B, B' \in \mathscr{B}$,

$$\tilde{N}_t^i(x, B) = N_t(x^i, B) - \int_0^t \int_B n_s^i(x, dy)\, ds, \qquad i = 1, \dots, n,$$

is a family of orthogonal (interacting) $(\mathscr{P}^n, \mathscr{F}_t)$-martingales compensating the counting processes $N_t(x^i, B)$, with predictable quadratic variation

$$\langle\langle \tilde{N}^i_\cdot(x,B), \tilde{N}^i_\cdot(x,B')\rangle\rangle_t = \int_0^t \int_{B\cap B'} n^i_s(x,dy)\,ds.$$

A consequence is that under such a measure $\mathscr{P}^n$, the empirical distributions X^n_t obey the following representation. Here the common superindex n in $x = (x^{1,n}, \ldots, x^{n,n})$ is suppressed. For $f \in \mathscr{C}^{1,0}(I \times E)$, the set of bounded continuous functions with one continuous time derivative, we have

$$\begin{aligned}\langle X^n_t, f_t\rangle - \langle X^n_s, f_s\rangle &= \frac{1}{n}\sum_{i=1}^n (f_t(x^i_t) - f_s(x^i_s))\\ &= \frac{1}{n}\sum_{i=1}^n\biggl(\int_s^t \frac{\partial}{\partial r} f_r(x^i_r)\,dr\\ &\qquad + \int_s^t\int_E (f_r(y) - f_r(x^i_r))\,dN_r(x^i_r, dy)\biggr)\\ &= \int_s^t \Bigl\langle X^n_r, \frac{\partial}{\partial r} f_r\Bigr\rangle dr + \int_s^t \Bigl\langle X^n_r, \int_E (f_r(y) - f_r(\cdot))\,dN_r(\cdot, dy)\Bigr\rangle.\end{aligned}$$

The generator $\mathscr{A}_t f(x^i, x) = \mathscr{A}_t f(x^i\,;\mu^n)$ of $t \mapsto x^i_t$ under $\mathscr{P}^n$ is defined by

$$\mathscr{A}_t f(x^i, x) = \int (f(y) - f(x^i))\,n^i_t(x, dy).$$

Then the previous representation of X^n can be rewritten in the form

$$\langle X^n_t, f_t\rangle = \langle X^n_0, f_0\rangle + \int_0^t \Bigl\langle X^n_r, \frac{\partial}{\partial r} f_r\Bigr\rangle dr + \int_0^t \bigl\langle X^n_r, \mathscr{A}_r f_r(\cdot\,; X^n_r)\bigr\rangle dr + M^f_t, \tag{1.1}$$

where

$$M^f_t = \int_0^t \Bigl\langle X^n_{r-}, \int (f_r(y) - f_r(\cdot))\,d\tilde{N}_r(\cdot, dy\,; X^n_r)\Bigr\rangle$$

is a purely discontinuous martingale. For f and g in $\mathscr{C}^{1,0}(I \times E)$ the predictable quadratic variation is given by

$$\begin{aligned}&\langle\langle M^f_\cdot, M^g_\cdot\rangle\rangle_t\\ &\quad = \frac{1}{n}\int_0^t \Bigl\langle X^n_r, \int (f_r(y) - f_r(\cdot))(g_r(y) - g_r(\cdot))\,n_r(\cdot, dy\,; X^n_r)\Bigr\rangle dr.\end{aligned} \tag{1.2}$$

We are interested in processes X^n which obey (1.1) and (1.2) and which have a *McKean–Vlasov limit*, as $n \to \infty$.

PROPOSITION 1. *Suppose $n^i(x, dy)$, $i \geq 1$, is a collection of weakly interacting jump measures which satisfies Assumption* 1. *For each n and each $\nu \in \mathscr{M}$, there exists a measure $\mathscr{P}^n_\nu$ with $\mathscr{P}^n_\nu(X^n_0 = \nu) = 1$ which is the unique solution of the nonlinear martingale problem* (1.1)–(1.2).

Suppose X_0^n converges weakly to some $\eta_0 \in \mathscr{M}$. Then X^n converges weakly in $\mathscr{D}(I, \mathscr{M})$ to a deterministic path $\eta \in \mathscr{C}(I, \mathscr{M})$ such that

$$(1.3) \qquad \langle \eta_t, f_t \rangle = \langle \eta_0, f_0 \rangle + \int_0^t \left\langle \eta_r, \frac{\partial}{\partial r} f_r \right\rangle dr + \int_0^t \langle \eta_r, \mathscr{A}_r f_r(\cdot\,; \eta_r) \rangle dr.$$

For a proof, see, for example, Oelschläger (1984).

1.4. *Infinitesimal measure generator.* Consider functionals $F(\mu) = F(\langle \mu, f \rangle)$, where $F \in \mathscr{C}(R)$, $\mu \in \mathscr{M}$ and $f \in \mathscr{C}^{1,0}(I \times E)$. Then

$$(1.4) \qquad F(X_t^n) = F(X_0^n) + \int_0^t \mathscr{G}_s^n F(X_s^n)\, ds + M_t^F,$$

where M_t^F is a $(\mathscr{P}^n, \mathscr{F}_t)$-local martingale and the infinitesimal generator $\mathscr{G}^n$ of the measure-valued jump process acts on probability measures $\mu \in \mathscr{M}$ by

$$\mathscr{G}_t^n F(\mu) = n \left\langle \mu, \int \left(F\left(\mu + \frac{1}{n}\delta_y - \frac{1}{n}\delta_\cdot \right) - F(\mu) \right) n_t(\cdot, dy\,; \mu) \right\rangle.$$

In particular, for $F(\mu) = e^{\langle \mu, f \rangle}$,

$$\mathscr{G}_t^n\, e^{\langle \mu, f \rangle} = \left\langle \mu, \int n\left(e^{(f(y) - f(\cdot))/n} - 1\right) n_t(\cdot, dy\,; \mu) \right\rangle e^{\langle \mu, f \rangle}.$$

The *Hamiltonian* associated to $\mathscr{G}^n$ is defined as

$$(1.5) \qquad \mathscr{H}_t^n(\mu, f) = e^{-\langle \mu, f \rangle}\, \mathscr{G}_t^n\, e^{\langle \mu, f \rangle} = \left\langle \mu, \int n\left(e^{(f(y) - f(\cdot))/n} - 1\right) n_t(\cdot, dy\,; \mu) \right\rangle.$$

2. Large deviations and main results.

2.1. *Absolutely continuous paths.* It is a well-known feature in the theory of large deviations on path space that the deviations from a deterministic limit occur with exponentially small probabilities and are concentrated on highly regular paths.

The regularity notion relevant in our situation is that of absolutely continuous paths, as developed in Dawson and Gärtner (1987).

Consider the Schwartz space $\mathscr{D}$ of test functions φ with compact support $\operatorname{supp} \varphi$ in R^d and continuous derivatives of all orders. Via the subspaces $\mathscr{D}_K = \{\varphi \in \mathscr{D}\colon \operatorname{supp} \varphi \subseteq K\}$, $K \subset R^d$ compact subsets, $\mathscr{D}$ is furnished with the usual inductive topology. Let $\mathscr{D}'$ denote the dual space of real distributions on R^d. By embedding E into R^d, we may consider $\mathscr{M}$ as a subset of $\mathscr{D}'$. Hence we extend the notation $\varphi \mapsto \langle \vartheta, \varphi \rangle$ used for $\vartheta = \mu \in \mathscr{M}$ and $\varphi = f \in \mathscr{C}$ to any $\vartheta \in \mathscr{D}'$ and $\varphi \in \mathscr{D}$. A map $t \mapsto \vartheta_t \in \mathscr{D}'$ defined on I is said to be *absolutely continuous* if for each compact set $K \subset R^d$ there exists a neighborhood U_K of the zero function 0 in $\mathscr{D}_K$ and an absolutely continuous real-valued function k_K on I such that

$$|\langle \vartheta_t, \varphi \rangle - \langle \vartheta_s, \varphi \rangle| \le |k_K(t) - k_K(s)|, \qquad s, t \in I,\ \varphi \in U_K.$$

For such absolutely continuous maps $t \mapsto \vartheta_t$ the time derivative $(d/dt)\vartheta_t = \dot{\vartheta}_t$ exists in the distribution sense:

$$\dot{\vartheta}_t = \lim_{h\to 0} h^{-1}\big(\vartheta_{t+h} - \vartheta_t\big) \quad \text{for almost all } t \in I.$$

If E is discrete, we make the identification of $\mathscr{C}_0(E)$ with test functions of the form $\psi = \sum_{i=1}^m a_i \delta^{\varepsilon}_{x_i}$, where $x_1, \dots, x_m$ are points in E, δ^{ε}_x is an approximation to the Dirac function at x as $\varepsilon \to 0$ and the a_i's are constants. By taking limits as $\varepsilon \to 0$, we obtain the distributions $\mathscr{D}'(E)$ on E in duality with $\mathscr{C}_0(E)$. Thus we can take $\mu \in \mathscr{M}$ and form, for example, $\dot{\mu} \in \mathscr{D}'(E)$. In particular, $\langle \mu_t, \varphi\rangle$ is absolutely continuous for all $\varphi \in \mathscr{C}_0(E)$ with derivative $\langle \dot{\mu}_t, \varphi\rangle$.

The defining equation (1.3) for the limit process η_t can now be written

$$\begin{aligned}\int_0^t \langle \dot{\eta}_r, f_r\rangle\, dr &= \langle \eta_t, f_t\rangle - \langle \eta_0, f_0\rangle - \int_0^t \Big\langle \eta_r, \frac{\partial}{\partial r} f_r\Big\rangle\, dr \\ &= \int_0^t \langle \eta_r, \mathscr{A}_r f_r(\cdot\,; \eta_r)\rangle\, dr, \qquad f \in \mathscr{C}_0(I \times E).\end{aligned}$$

Writing $\mathscr{A}_t^*(n)$ for the formal adjoint of the generator $\mathscr{A}_t$ corresponding to the jump measure $n = \{n_t^i(x, dy)\}_{1\le i\le n}$, this relation takes the weak form

$$\dot{\eta}_t = \mathscr{A}_t^*(n)\eta_t \quad \text{for almost all } t \in I.$$

2.2. *Orlicz space, admissible paths.* Consider the pair of Young functions

$$\tau(t) = e^t - t - 1, \qquad \tau^*(s) = (s+1)\log(s+1) - s, \qquad s > -1.$$

Fix a jump measure $n_t(x, dy)$ and a path $\mu \in \mathscr{D}(I, \mathscr{M})$. For any measurable function $h_t(\cdot,\cdot)$ on the measurable space $(I \times E^2\,; ds\, \mu_s(dx)\, n_s(x, dy))$, define

$$\|h\|_\theta = \inf\left\{a > 0\colon \int_I \Big\langle \mu_r, \int \theta(h_r(\cdot, y)/a)\, n(\cdot, dy)\Big\rangle\, dr \le 1\right\},$$

where θ equals τ or τ^*. This defines a norm and correspondingly we obtain two *Orlicz spaces* $(L^\tau, \|\cdot\|_\tau)$ and $(L^{\tau^*}, \|\cdot\|_{\tau^*})$ of functions on $I \times E^2$. These are Banach spaces in topological duality. For more details, see Krasnoselskii and Rutickii (1961).

We also introduce the Orlicz space $\mathscr{O}(\mu, n) = \mathscr{O}(I\times E^2\,; ds\, \mu_s(dx)\, n_s(x, dy))$ of strictly positive functions g of the form $g = h + 1$, $h \in L^{\tau^*}$. This is the set of functions for which

$$\|g\| = \inf\left\{a > 0\colon \int_I \Big\langle \mu_r, \int \Big(\frac{g_r(\cdot, y)}{a}\log\frac{g_r(\cdot, y)}{a} - \frac{g_r(\cdot, y)}{a} + 1\Big)\, n_r(\cdot, dy)\Big\rangle\, dr \le 1\right\}$$

is finite.

We now define a subset H_0 of regular paths in $\mathscr{C}(I, \mathscr{M})$ which will constitute all the trajectories possibly deviating from the McKean–Vlasov limit. It is an analog to the "Cameron–Martin" space defined in Fleischmann, Gärtner and Kaj (1993), for regular paths of small perturbations of super-Brownian motion away from the heat flow.

DEFINITION 1. Fix $\eta_0 \in \mathscr{M}$. Let H_0 denote the set of all paths $\mu \in \mathscr{C}(I, \mathscr{M})$ with:

(i) $\mu_0 = \eta_0$.

(ii) The map $t \mapsto \mu_t$ defined on I is absolutely continuous.

(iii) $\dot{\mu}_t = \mathscr{A}_t^*(m)\mu_t$ for some m such that $m_t(\cdot, dy)$ is absolutely continuous with respect to $n_t(\cdot, dy)$ for almost all $t \in I$.

(iv) The Radon–Nikodym derivative $t \mapsto g_t(\cdot, y) := m_t(\cdot, dy)/n_t(\cdot, dy) > 0$ belongs to the Orlicz space $\mathscr{O}(\mu, n)$.

2.3. *Main result.* Define the *scaled Hamiltonian* $\mathscr{H}$ by

$$\mathscr{H}_t(\mu, f) = \lim_{n\to\infty} \frac{1}{n} \mathscr{H}_t^n(\mu, nf) = \left\langle \mu, \int (e^{(f(y)-f(\cdot))} - 1) n_t(\cdot, dy\,; \mu) \right\rangle.$$

For $\vartheta \in \mathscr{D}'(E)$ and $\mu \in \mathscr{M}$, define the *Legendre transform* of $\mathscr{H}$ by

$$\mathscr{L}_t(\mu, \vartheta) = \sup_{\varphi \in \mathscr{C}_0(E)} \{\langle \vartheta, \varphi \rangle - \mathscr{H}_t(\mu, \varphi)\}.$$

THEOREM 1. *Consider a fixed (interacting) jump measure $n_t(\cdot, dy)$ which satisfies Assumption 1. Let $\mathscr{H}_t$ be the associated Hamiltonian and $\mathscr{L}_t$ the Legendre transform. Fix $\eta_0 \neq 0$ and let η be the solution of (1.3) for the generator $\mathscr{A}(n)$. For any measurable set A in $\mathscr{D}(I, \mathscr{M})$, put*

$$\mathscr{P}^n(A) = \mathscr{P}^n(X^n \in A \mid X_0^n).$$

Then:

(i) *For each open subset G of $\mathscr{D}(I, \mathscr{M})$,* $\liminf_{n\to\infty} (1/n) \log \mathscr{P}^n(G) \geq -\inf_{\mu \in G} S(\mu)$.

(ii) *For each closed subset F of $\mathscr{D}(I, \mathscr{M})$,* $\limsup_{n\to\infty} (1/n) \log \mathscr{P}^n(F) \leq -\inf_{\mu \in F} S(\mu)$.

(iii) *The level sets $\{\mu \in \mathscr{D}(I, \mathscr{M})\colon S(\mu) \leq N\}$, $N > 0$, are compact,*

where

$$S(\mu) = \begin{cases} \displaystyle\int_0^T \mathscr{L}_t(\mu_t, \dot{\mu}_t)\, dt, & \textit{if } t \mapsto \mu_t \textit{ is absolutely continuous and } \mu_0 = \eta_0, \\ \infty, & \textit{otherwise.} \end{cases}$$

To obtain alternative representations for S, we introduce, for $0 \leq s \leq t \leq T$, the functionals

$$J_{s,t}(\mu, f) = \langle \mu_t, f_t \rangle - \langle \mu_s, f_s \rangle - \int_s^t \left\langle \mu_r, \frac{\partial}{\partial r} f_r \right\rangle dr - \int_s^t \mathscr{H}_r(\mu_r, f_r)\, dr. \tag{2.1}$$

THEOREM 2. *Let η be the solution of* (1.3) *and H_0 the corresponding set of paths which are admissible for η according to Definition* 1. *The representation*

$$S(\mu) = \sup\{J_{0,T}(\mu, f)\colon f \in \mathscr{C}^{1,0}(I \times E)\} \tag{2.2}$$

holds. Moreover,

$$S(\mu) < \infty \quad \textit{if and only if } \mu \in H_0,$$

and in the case $S(\mu) < \infty$, *then*

$$\begin{aligned} S(\mu) &= \int_0^T \left\langle \mu_r, \int (g_r(\cdot, y) \log g_r(\cdot, y) - g_r(\cdot, y) + 1) n_r(\cdot, dy) \right\rangle dr \\ &=: \int_0^T \langle \mu_r \otimes n_r, (g_r \log g_r - g_r + 1) \rangle \, dr. \end{aligned} \tag{2.3}$$

2.4. *Example: Epidemic SIR model.* We give an example of a multitype model with *type interaction*. Let the state space E, informally described as $E = \{\text{susc}, \text{inf}, \text{rem}\} \times \text{position}$, consist of pointers to one of three possible types, $S =$ susceptible, $I =$ infective and $R =$ removed, and in addition of a position variable $r_i \in R$, say, if component number i is of type S or I and a cemetery position $\dagger$ if of type R. Then

$$\frac{1}{n} \sum_{i=1}^{n} \delta_{x^{i,n}} = \frac{1}{n} \left(\sum_{i=1}^{|S|} \delta_{r_i}^{(S)} + \sum_{j=1}^{|I|} \delta_{r_j}^{(I)} + \sum_{k=1}^{|R|} \delta_{\dagger} \right).$$

Suppose $\lambda(r, r')$ is a nonnegative function on R^2 which satisfies, for example,

$$\sup_{r \in R} \int_R \lambda(r, r') \, dr' < \infty.$$

Define jump rates by

$$n^i(\text{susc}, \text{inf}\,; x) = \gamma^i(\text{susc}\,; x) = \frac{1}{n} \sum_{j=1}^{|I|} \lambda(r_i, r_j)$$

for a jump $S \to I$ of particle i at r_i and

$$n^i(\text{inf}, \text{rem}\,; x) = \gamma^i(\text{inf}\,; x) = \rho$$

for $I \to R$ of particle i. Write, for $\varphi = (\varphi_1, \varphi_2)$,

$$\langle X_t^n, \varphi \rangle = \langle S_t^n, \varphi_1 \rangle + \langle I_t^n, \varphi_2 \rangle.$$

Then

$$\langle X_t^n, \varphi \rangle = \langle X_0^n, \varphi \rangle + \int_0^t \langle S_s^n \otimes I_s^n, \lambda(\varphi_2 - \varphi_1) \rangle \, ds + \varrho \int_0^t \langle I_s^n, (0 - \varphi_2) \rangle \, ds + M_t^{\varphi},$$

where the $(\mathscr{P}^n, \mathscr{F}_t)$-martingale M_t^{φ}, $t \in I$, has quadratic variation

$$\langle\!\langle M^{\varphi} \rangle\!\rangle_t = \frac{1}{n} \left(\int_0^t \left\langle S_s^n \otimes I_s^n, \lambda(\varphi_2 - \varphi_1)^2 \right\rangle ds + \varrho \int_0^t \left\langle I_s^n, \varphi_2^2 \right\rangle ds \right).$$

Hence, with $\langle \mu, \varphi \rangle = \langle \mu^1, \varphi_1 \rangle + \langle \mu^2, \varphi_2 \rangle$,

$$\mathscr{H}(\mu, \varphi) = \langle \mu^1 \otimes \mu^2, \; \lambda(e^{\varphi_2 - \varphi_1} - 1) \rangle + \varrho \, \langle \mu^2, (e^{-\varphi_2} - 1) \rangle.$$

Along the extremals of the corresponding variational problem for $\mathscr{L}$ there are functions g^1 and g^2, such that the rate function can be written as

$$S(\mu) = \int_0^T \left(\left\langle \mu_r^1 \otimes \mu_r^2,\ \lambda\left(g_r^1 \log g_r^1 - g_r^1 + 1\right)\right\rangle + \varrho \left\langle \mu_r^2, \left(g_r^2 \log g_r^2 - g_r^2 + 1\right)\right\rangle\right) dr,$$

which is in agreement with (2.3).

3. Representations of the rate function. Frequently we will use in the sequel a simpler notation which we illustrate with the example

$$\begin{aligned}\mathscr{H}_t(\mu, f) &= \left\langle \mu, \int \left(e^{(f(y)-f(\cdot))} - 1\right) n_t(\cdot, dy\,; \mu)\right\rangle \\ &= \left\langle \mu, \int \left(e^{\Delta f} - 1\right) dn_t\right\rangle, \qquad \Delta f(\cdot, y) = f(y) - f(\cdot).\end{aligned}$$

PROOF OF THEOREM 2. We show first that the rate function $S(\mu)$ in Theorem 1, given by

$$S(\mu) = \int_0^T \sup_{\varphi \in \mathscr{C}_0(E)} \{\langle \dot{\mu}_r, \varphi\rangle - \mathscr{H}_r(\mu_r, \varphi)\}\, dr,$$

has the representation (2.2).

For all $f \in \mathscr{C}_0^{1,0}(I \times E)$,

$$S(\mu) \geq \int_I (\langle \dot{\mu}_r, f_r\rangle - \mathscr{H}_r(\mu_r, f_r))\, dr = J_{0,T}(\mu, f).$$

Hence

$$S(\mu) \geq \sup_{f \in \mathscr{C}^{1,0}(I \times E)} J_{0,T}(\mu, f). \tag{3.1}$$

Therefore, if the right side of (2.2) is infinite, then the rate function is infinite.

Next suppose that $\sup_{f \in \mathscr{C}^{1,0}(I\times E)} J_{0,T}(\mu, f) < \infty$. Define, for $0 \leq s \leq t \leq T$,

$$\ell_{s,t}(f) := \langle \mu_t, f_t\rangle - \langle \mu_s, f_s\rangle - \int_s^t \left\langle \mu_r, \frac{\partial}{\partial r} f_r + \mathscr{A}_r f_r\right\rangle dr.$$

Since

$$\mathscr{H}_t(\mu_t, f) = \langle \mu_t, \mathscr{A}_t f\rangle + \left\langle \mu_t, \int \tau(\Delta f)\, dn_t\right\rangle,$$

we have

$$J_{s,t}(\mu, f) = \ell_{s,t}(f) - \int_s^t \left\langle \mu_r, \int \tau(\Delta f_r)\, dn_r\right\rangle dr \leq \sup_{f \in \mathscr{C}^{1,0}(I\times E)} J_{0,T}(\mu, f) < \infty.$$

Hence, following Léonard (1989),

$$\frac{1}{\|\Delta f\|_\tau}\ell_{s,t}(f) = \ell_{s,t}\left(\frac{f}{\|\Delta f\|_\tau}\right)$$
$$\le \sup_{f\in\mathscr{C}^{1,0}(I\times E)} J_{0,T}(\mu,f) + \left\langle \mu_r, \int \tau\left(\frac{\Delta f_r}{\|\Delta f\|_\tau}\right) dn_r\right\rangle dr$$
$$\le \sup_{f\in\mathscr{C}^{1,0}(I\times E)} J_{0,T}(\mu,f) + 1 < \infty.$$

Therefore, $\Delta f \mapsto \ell_{s,t}(f)$ can be extended to a continuous linear functional on L^τ. By the Riesz representation theorem there exists a unique $h \in L^{\tau^*}$ ($h = g - 1$) such that

$$\ell_{s,t}(f) = \int_s^t \left\langle \mu_r, \int \Delta f_r\, h_r\, dn_r\right\rangle dr.$$

Then choosing $f_t(y) = \varphi(y) \in \mathscr{C}_0(E)$, for all $0 \le s < t \le T$,

$$\langle \mu_t, \varphi\rangle = \langle \mu_s, \varphi\rangle + \int_s^t \left\langle \mu_r, \mathscr{A}_r\varphi + \int \Delta\varphi\, h_r\, dn_r\right\rangle dr,$$

which shows that μ is absolutely continuous.

Moreover, for almost all t,

$$\langle \dot{\mu}_t, \varphi\rangle = \left\langle \mu_t, \mathscr{A}_t\varphi + \int \Delta\varphi\, h_t\, dn_r\right\rangle, \tag{3.2}$$

so

$$\begin{aligned} S(\mu) &= \int_0^T \sup_{\varphi\in\mathscr{C}_0(E)} \left\langle \mu_r, \int (h_r\ \Delta\varphi - \tau(\Delta\varphi))\, dn_r\right\rangle dr \\ &\le \int_0^T \left\langle \mu_r, \int \tau^*(h_r)\, dn_r\right\rangle dr. \end{aligned} \tag{3.3}$$

If $f \in \mathscr{C}_0^{1,2}(I\times E)$, then $\Delta f \in L^\tau$. Hence

$$J_{0,T}(\mu,f) = \int_0^T \left\langle \mu_r, \int \tau^*(h_r)\, dn_r\right\rangle dr$$
$$- \int_0^T \left\langle \mu_r, \int (\tau^*(h_r) + \tau(\Delta f_r) - h_r\ \Delta f_r)\, dn_r\right\rangle dr.$$

By Young's inequality and an approximation in L^τ,

$$\sup_{f\in\mathscr{C}^{1,0}(I\times E)} J_{0,T}(\mu,f) = \int_0^T \left\langle \mu_r, \int \tau^*(h_r)\, dn_r\right\rangle dr.$$

Hence, by (3.3) and the last equality,

$$S(\mu) \le \int_0^T \left\langle \mu_r, \int \tau^*(h_r)\, dn_r\right\rangle dr = \sup_{f\in\mathscr{C}^{1,0}(I\times E)} J_{0,T}(\mu,f) < \infty.$$

Together with (3.1) this completes the proof of (2.2).

Furthermore, we have shown that whenever $S(\mu)$ is finite, then

$$S(\mu) = \int_0^T \left\langle \mu_r, \int \tau^*(h_r)\, dn_r \right\rangle dr = \int_0^T \langle \mu_r \otimes n_r,\ (g_r \log g_r - g_r + 1)\rangle\, dr,$$

whence (2.3).

Summing up the above, if $S(\mu) < \infty$, then $\mu \in H_0$. It remains to show the converse. Therefore, assume that $\mu \in H_0$. Then there exists a function $g \in \mathscr{O}(\mu, n)$ such that (3.2) holds with $h = g - 1 \in L^{\tau^*}$. We even have

$$\int_0^T \left\langle \mu_r, \int \tau^*(h_r)\, dn_r \right\rangle dr < \infty,$$

since τ^* satisfies the condition $\sup_t \tau^*(at)/\tau^*(t) < \infty$ for some $a > 0$ [see, e.g., Neveu (1975), Appendix.] Therefore, as in (3.3), $S(\mu) < \infty$. □

4. Large deviation principle for the independent case. In this section we prove Theorem 1 in the case when the interaction is locally "frozen," and hence the processes $t \mapsto x_t^i$ are independent. According to Theorems 5.2 and 5.3 in Dawson and Gärtner (1987), it suffices to derive local lower and upper bounds and check the exponential tightness property. This is done below in Propositions 2, 3 and 4. Sections 4.1–4.4 are similar in approach to Djehiche (1993), adapted to the case of measure-valued processes. Section 4.5, in which we prove compactness of the level sets, is close to Dawson and Gärtner [(1987), Lemma 5.6] and Feng [(1992), Lemma 3.9].

Throughout we let $\hat{\mu}$ denote a fixed path in $\mathscr{D}(I, \mathscr{M})$ and study the empirical process X^n with jump measure

$$\hat{n}^i(x^i, dy) = \gamma_t(x^i, \hat{\mu})\, \pi(x^i, dy), \qquad 1 \le i \le n.$$

4.1. *Change of measure.* For $\mu \in \mathscr{D}(I, \mathscr{M})$ and $f \in \mathscr{C}^{1,0}(I \times E)$, define

$$K_t(\mu, f) = \langle \mu_t, f_t \rangle - \langle \mu_0, f_0 \rangle - \int_0^t \left\langle \mu_r, \frac{\partial}{\partial r} f_r \right\rangle dr. \tag{4.1}$$

Consider the signed measure paths defined by

$$\begin{aligned} f \mapsto K_t^{n,f} &:= K_t(X^n, f) \\ &= \int_0^t \left\langle X_{r-}^n, \int (f_r(y) - f_r(\cdot))\, dN_r(\cdot, dy) \right\rangle dr, \qquad t \in I, \end{aligned} \tag{4.2}$$

where the second equality is from (1.1).

LEMMA 4.1. *Fix $f \in \mathscr{C}^{1,0}(I \times E)$. For each $\mathscr{F}_t$-predictable and $\mathscr{P}^n$-a.s. bounded function a_t,*

$$\xi_t^{n,f}(a) := \exp\left\{ n \int_0^t a_r\, dK_r^{n,f} - n \int_0^t \mathscr{H}_r(X_r^n, a_r f_r)\, dr \right\}, \qquad t \in I,$$

is a $(\mathscr{P}^n, \mathscr{F}_t)$-martingale. In particular,

$$Z_t^{n,f} := \exp\left\{nK_t^{n,f} - n\int_0^t \mathscr{H}_r(X_r^n, f_r)\,dr\right\} = \exp nJ_{0,t}(X^n, f), \qquad t \in I,$$

is a $(\mathscr{P}^n, \mathscr{F}_t)$-martingale.

PROOF. Put

$$U_t^{n,f} = \exp\left\{K_t^{n,f} - \int_0^t \mathscr{H}_r^n(X_r^n, f_r)\,dr\right\}.$$

Since $n\int_0^t a_r\,dK_r^{n,f} = K_t^{n,anf}$ and $\mathscr{H}_t^n(\mu, nf) = n\mathscr{H}_t(\mu, f)$, we have

$$\xi_t^{n,f}(a) = U_t^{n,anf}.$$

Therefore, it is enough to show that $U_t^{n,f}$, $t \geq 0$, is a $(\mathscr{P}^n, \mathscr{F}_t)$-martingale. By (1.4) and (1.5) there is a martingale $M^{\exp}$, such that

$$\exp(\langle X_t^n, f_t\rangle) - \exp(\langle X_0^n, f_0\rangle) - \int_0^t \exp(\langle X_r^n, f_r\rangle)\left(\mathscr{H}_r^n(X_r^n, f_r) + \left\langle X_r^n, \frac{\partial}{\partial r} f_r\right\rangle\right)dr = M_t^{\exp}.$$

It is then easy to check that

$$dU_t^{n,f} = \exp(\langle X_t^n, f_t\rangle)\,dM_t^{\exp},$$

and we are done. □

Define a new probability measure $\mathscr{P}_a^n$ which is equivalent to $\mathscr{P}^n$ by setting

$$\xi_t^{n,f}(a) = \frac{d\mathscr{P}_a^n}{d\mathscr{P}^n}\bigg|_{\mathscr{F}_t}.$$

Introduce the notation

$$\mathscr{H}_t'(\mu, af) = \frac{d}{da}\mathscr{H}_t(\mu, af), \qquad \mathscr{H}_t''(\mu, af) = \frac{d^2}{da^2}\mathscr{H}_t(\mu, af).$$

LEMMA 4.2. *Let b_t, $t \in I$, denote a $\mathscr{F}_t$-predictable and $\mathscr{P}^n$-a.s. bounded function. The process*

$$\begin{aligned} M_t^n &= \int_0^t b_r\,dK_r^{n,f} - \int_0^t b_r\,\mathscr{H}_r'(X_r, a_r f_r)\,dr \\ &= \int_0^t b_r\,dK_r^{n,f} \\ &\quad - \int_0^t b_r\left\langle X_r^n, \int (f_r(y) - f_r(\cdot))\exp(a_r(f_r(y) - f_r(\cdot)))\,\hat{n}_r(\cdot, dy)\right\rangle dr \end{aligned}$$

is a $(\mathscr{P}_a^n, \mathscr{F}_t)$-martingale on I with predictable quadratic variation given by

$$\langle\!\langle M^n\rangle\!\rangle_t = \frac{1}{n}\int_0^t b_r^2\,\mathscr{H}_r''(X_r, a_r f_r)\,dr.$$

PROOF. This follows from the property of Radon–Nikodym derivatives that quotients of the form

$$\xi_t^{n,f}(a+\varepsilon b)/\xi_t^{n,f}(a),$$

are $(\mathscr{P}_a^n, \mathscr{F}_t)$-martingales on I. Differentiate twice at $\varepsilon = 0$ to get the stated properties of M_t^n. □

4.2. *A variational problem.* As usual we write $f^+ = f \vee 0$ and $f^- = (-f) \vee 0$ for the positive and negative parts of a real-valued function f on R^d. The notation $f^\pm$ refers to either of the two functions.

LEMMA 4.3. *Let ν be a Borel measure on R^d and $f \in L^1(\nu)$. Define*

$$g(\lambda) = \int f e^{\lambda f}\, d\nu, \qquad \lambda \in R.$$

Suppose

$$\gamma^\pm := \int (f^\pm)^2\, d\nu > 0.$$

Then for every real ζ, there exists a unique λ such that $\zeta = g(\lambda)$. The map $\nu \mapsto \lambda$ is continuous. Moreover,

$$|\lambda| \le \left|\zeta - \int f\, d\nu\right| / (\gamma^+ \wedge \gamma^-).$$

PROOF. If $\lambda \ge 0$, then $\int f^2 e^{\lambda f}\, d\nu \ge \int (f^+)^2\, d\nu$. If $\lambda \le 0$, then $\int f^2 e^{\lambda f}\, d\nu \ge \int (f^-)^2\, d\nu$. The existence of a unique solution to the equation $\zeta = g(\lambda)$ now follows from the facts that $\lambda \mapsto g(\lambda)$ is continuous and

$$\lambda \mapsto g'(\lambda) = \int f^2 e^{\lambda f}\, d\nu \ge \gamma^+ \wedge \gamma^- > 0,$$

so that the map $\lambda \mapsto g(\lambda)$ is bijective. It follows also that the inverse is continuous, which we can interpret as the continuity of λ in ν.

The last part of the lemma is proved as follows. Fix ζ and λ such that $\zeta = g(\lambda)$. If $\lambda > 0$,

$$\zeta - g(0) = \int f(e^{\lambda f} - 1)\, d\nu \ge \lambda \int (f^+)^2\, d\nu = \lambda \gamma^+ > 0.$$

If $\lambda < 0$,

$$\zeta - g(0) = \int f(e^{\lambda f} - 1)\, d\nu \le \lambda \int (f^-)^2\, d\nu = \lambda \gamma^- < 0.$$

Hence $|\lambda| \le |\zeta - g(0)| / (\gamma^+ \wedge \gamma^-)$. □

COROLLARY 4.4. *Let* $\mu \in H_0$ *and* $f \in \mathscr{C}_0^{1,0}(I \times E)$. *Assume*

$$\gamma^{\pm} = \inf_{t\in I}\left\langle \mu_t, \int (\Delta f_t^{\pm})^2 \, d\,\hat{n}_t \right\rangle > 0,$$

where $\Delta f_t(x,y) = f_t(y) - f_t(x)$. *Then there exists a unique function* $\lambda_t = \lambda(\mu_t, \dot{\mu}_t, f_t, t)$, *such that, for almost all* $t \in I$,

$$\langle \dot{\mu}_t, f_t \rangle = \mathscr{H}_t'(\mu_t, \lambda_t f_t). \tag{4.3}$$

Assume

$$\gamma_n^{\pm} := \inf_{t\in I}\left\langle X_t^n, \int (\Delta f_t^{\pm})^2 \, d\,\hat{n}_t \right\rangle > 0, \qquad n \geq n_0.$$

Then there exists a unique progressively measurable process $t \mapsto \lambda_t^n = \lambda(X_t^n, \dot{\mu}_t, f_t, t)$ *on* I *such that, for* $n \geq n_0$ *and almost all* $t \in I$,

$$\langle \dot{\mu}_t, f_t \rangle = \mathscr{H}_t'(X_t^n, \lambda_t^n f_t). \tag{4.4}$$

Moreover,

$$\begin{aligned} |\lambda_t| &\leq (|\langle \dot{\mu}_t, f_t \rangle| + 2C\|f_t\|)/(\gamma^+ \wedge \gamma^-), \\ |\lambda_t^n| &\leq (|\langle \dot{\mu}_t, f_t \rangle| + 2C\|f_t\|)/(\gamma_n^+ \wedge \gamma_n^-). \end{aligned} \tag{4.5}$$

and λ^n *is continuous in* X^n.

PROOF. Apply the lemma to the measure $d\nu = \mu_t(dx) \otimes n_t(x, dy)$ defined on E^2 (fixed t) and with $\varphi = \Delta f_t$, to obtain λ and λ^n. In addition, we have to show that λ_t^n is $\mathscr{F}_t$ progressive. However, for α real,

$$\begin{aligned} &\{(\zeta, t, \nu) \colon \lambda^n(\mu_t, \zeta, f_t, t) > \alpha\} \\ &\qquad = \{(\zeta, t, \nu) \colon \zeta > \mathscr{H}'(\mu_t, \alpha f_t)\} \in \mathscr{B}_E \times \mathscr{P}(\mathscr{D}(I, \mathscr{M})), \end{aligned}$$

where $\mathscr{P}(\mathscr{D}(I, \mathscr{M}))$ denotes the progressively measurable paths from I into $\mathscr{M}$.

Also by the lemma,

$$(\gamma_n^+ \wedge \gamma_n^-)\,|\lambda_t^n| \leq \big|\langle \dot{\mu}_t, f_t \rangle\big| + \left\langle X_t^n, \int_E |\Delta f_t| \, d\,\hat{n}_t \right\rangle \leq |\langle \dot{\mu}_t, f_t \rangle| + 2C\|f_t\|,$$

where the constant C is from assumption (A2). Similarly for λ_t. □

4.3. *Lower bound.* The purpose of this section is to prove the following proposition.

PROPOSITION 2. *Fix* $\mu \in \mathscr{D}(I, \mathscr{M})$ *and let* V *be an open neighborhood of* μ. *Then*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathscr{P}^n(V) \geq -S(\mu).$$

First note that if $S(\mu)=\infty$, there is nothing to prove. Therefore, throughout this section we assume $S(\mu)<\infty$.

As a further preliminary for the proof, we will choose a test function f, which will then be fixed throughout this subsection. The quantities $\gamma^{\pm}=\gamma^{\pm}(f)$ and $\gamma_n^{\pm}=\gamma_n^{\pm}(f)$ were defined in the previous section.

Take $f\in\mathscr{C}_0^{1,0}(I\times E)$ and n_0, such that

$$\begin{aligned}&\gamma^{\pm}(f)\wedge\gamma_n^{\pm}(f)\ge c_1>0,\quad n\ge n_0,\\ &|\langle\dot{\mu}_t,f_t\rangle|\le c_2\quad\text{for almost all } t\in I.\end{aligned}\tag{4.6}$$

Now, by Corollary 4.4 there exists unique solutions $\lambda_t=\lambda_t(f)$ and $\lambda_t^n=\lambda_t^n(f)$ of the variational problems (4.3) and (4.4), and by (4.5) and (4.6) we can find a constant K such that almost surely

$$|\lambda_t|\vee|\lambda_t^n|\le K,\qquad t\in I,\ n\ge n_0.$$

LEMMA 4.5. *Put*

$$M_t^{1,n}=\int_0^t\lambda_{r-}^n\big(dK_r^{n,f}-\mathscr{H}_r'(X_r^n,\lambda_r^nf_r)\,dr\big)$$

and

$$\underline{\Sigma}_T^n=\int_0^T\{|\mathscr{H}_r(\mu_r,\lambda_rf_r)-\mathscr{H}_r(X_r^n,\lambda_r^nf_r)|+|\langle\dot{\mu}_r,f_r\rangle|\,|\lambda_r^n-\lambda_r|\}\,dr.$$

Then

$$\left|\int_0^T\lambda_r^n\,dK_r^{n,f}-\int_0^T\mathscr{H}_r(X_r^n,\lambda_r^nf_r)\,dr\right|\le S(\mu)+|M_T^{1,n}|+\underline{\Sigma}_T^n.$$

PROOF. We have

$$\begin{aligned}S(\mu)&=\int_0^T\sup_{\varphi\in\mathscr{C}_0(E)}\{\langle\dot{\mu}_r,\varphi\rangle-\mathscr{H}_r(\mu_r,\varphi)\}\,dr\\&=\int_0^T\sup_{\varphi\in\mathscr{C}_0(E)}\sup_{\lambda\in R}\{\lambda\langle\dot{\mu}_r,\varphi\rangle-\mathscr{H}_r(\mu_r,\lambda\varphi)\}\,dr.\end{aligned}$$

Hence

$$\infty>S(\mu)\ge\int_0^T\big(\lambda_r\langle\dot{\mu}_r,f_r\rangle-\mathscr{H}(\mu_r,\lambda_rf_r)\big)\,dr.$$

Therefore,

$$\begin{aligned}&\int_0^T\lambda_r^n\,dK_r^{n,f}-\int_0^T\mathscr{H}_r(X_r^n,\lambda_r^nf_r)\,dr+S(\mu)-S(\mu)\\&\quad\le S(\mu)+\int_0^T\lambda_{r-}^n\,(dK_r^{n,f}-\langle\dot{\mu}_r,f_r\rangle)\,dr+\int_0^T(\lambda_r^n-\lambda_r)\,\langle\dot{\mu}_r,f_r\rangle\,dr\\&\qquad+\int_0^T\big(\mathscr{H}_r(\mu_r,\lambda_rf_r)-\mathscr{H}_r(X_r^n,\lambda_r^nf_r)\big)\,dr.\end{aligned}$$

However, according to (4.4),

$$M_t^{1,n} = \int_0^t \lambda_{r-}^n (dK_r^{n,f} - \langle \dot{\mu}_r, f_r \rangle \, dr). \qquad \square$$

For $\delta > 0$, set

$$V_\delta = \{ \nu \in V: \sup_{t \in I} \| \nu_t - \mu_t \| < \delta \},$$

where $\| \cdot \|$ denotes the total variation norm on $\mathscr{M}$.

LEMMA 4.6. *For any $\varepsilon > 0$ there is a $\delta_0 > 0$ such that if $\delta \le \delta_0$ and $X^n \in V_\delta$, then $\underline{\Sigma}_T^n \le \varepsilon$.*

PROOF. We have

$$\begin{aligned} &|\mathscr{H}_r(\mu_r - X_r^n, \lambda_r^n f_r)| \\ &\quad \le C \| \mu_r - X_r^n \| \, \| \exp(\lambda_r^n \Delta f_r) - 1 \| \le C \| \mu_r - X_r^n \| (\exp(2K \| f_r \|) + 1). \end{aligned}$$

Hence, if $X^n \in V_\delta$, then

$$\begin{aligned} \underline{\Sigma}_T^n \le \int_0^T & \{ \left| \mathscr{H}(\mu_r, \lambda_r f_r) - \mathscr{H}(\mu_r, \lambda_r^n f_r) \right| + c_2 |\lambda_r^n - \lambda_r| \} \, dr \\ & + C\delta \int_0^T (\exp(2K \| f_r \|) + 1) \, dr, \end{aligned}$$

where the second part of (4.6) is used. Since, by the continuity statement in Corollary 4.4, the bounded function λ^n converges to λ as X^n goes to μ, we can choose δ so small that $\underline{\Sigma}_T^n \le \varepsilon$. $\square$

LEMMA 4.7. *For every $\delta > 0$,*

$$\lim_{n \to \infty} \mathscr{P}_{\lambda^n}^n (X^n \notin V_\delta) = 0.$$

PROOF. Let $\delta > 0$. Fix $\varphi \in \mathscr{C}(E)$. Define

$$\sigma = \inf \{ t \le T: \sup_{s \le t} |\langle X_s^n - \mu_s, \varphi \rangle| \ge \delta \}.$$

Then

$$\{ \sigma < T \} \subset \{ \sup_{s \le \sigma} |\langle X_s^n - \mu_s, \varphi \rangle| \ge \delta \}.$$

On the other hand,

$$\begin{aligned} &\{ \sup_{t \le T} \| X_t^n - \mu_t \| > \delta \} \\ &\qquad \subset \{ \sup_{t \le \sigma} \| X_t^n - \mu_t \| > \delta \} \cap \{ \sigma = T \} \cup \{ \sup_{t \le \sigma} |\langle X_t^n - \mu_t, \varphi \rangle| \ge \delta \}. \end{aligned}$$

Hence,

$$\mathscr{P}^n_{\lambda^n}\big(\sup_{t\le T}\|X^n_t-\mu_t\|>\delta\big)\le\mathscr{P}^n_{\lambda^n}\big(\sup_{t\le\sigma}\|X^n_t-\mu_t\|>\delta\big)$$
$$+\,\mathscr{P}^n_{\lambda^n}\big(\sup_{s\le\sigma}|\langle X^n_s-\mu_s,\varphi\rangle|\ge\delta\big).$$

Therefore,

$$\mathscr{P}^n_{\lambda^n}\big(\sup_{t\le T}\|X^n_t-\mu_t\|>\delta\big)\le 2\,\mathscr{P}^n_{\lambda^n}\big(\sup_{s\le\sigma}|\langle X^n_s-\mu_s,\varphi\rangle|\ge\delta\big).$$

By Lemma 4.2, taking $b_t=1$ and $f_t=\varphi$, the process

$$M^{2,n}_t:=K^{n,\varphi}_t-\int_0^t\mathscr{H}'_r(X^n_r,\lambda^n_r\varphi)\,dr=\langle X^n_t,\varphi\rangle-\langle\mu_t,\varphi\rangle$$

is a martingale under $\mathscr{P}^n_{\lambda^n}$ with

$$\langle\!\langle M^{2,n}_{\cdot}\rangle\!\rangle_t=\frac{1}{n}\int_0^t\mathscr{H}''_r(X^n_r,\lambda^n_r\varphi)\,dr.$$

Here

$$\mathscr{H}''_r(X^n_r,\lambda^n_r\varphi)\le 2C\|\varphi\|e^{2K\|\varphi\|}=:K_1.$$

For any $\alpha>0$, using the inequality of Lenglart and Rebolledo, we get

$$\mathscr{P}^n_{\lambda^n}\big(\sup_{s\le\sigma}|\langle X^n_s-\mu_s,\varphi\rangle|\ge\delta\big)\le\frac{\alpha}{\delta^2}+\mathscr{P}^n_{\lambda^n}\big(\langle\!\langle M^{2,n}\rangle\!\rangle_\sigma\ge\alpha\big).$$

Thus

$$\mathscr{P}^n_{\lambda^n}\big(\sup_{t\le T}\|X^n_t-\mu_t\|>\delta\big)\le 2\left(\frac{\alpha}{\delta^2}+\frac{K_1T}{\alpha n}\right).$$

The statement follows by letting first $n\to\infty$ and then $\alpha\to 0$. □

PROOF OF PROPOSITION 2. Apply Lemma 4.1 to the case $a=\lambda^n$. We obtain the exponential $\mathscr{P}^n$-martingale $\xi^{n,f}(\lambda^n)$ and a new probability measure $\mathscr{P}^n_{\lambda^n}$. By Lemma 4.5,

$$\xi^{n,f}_T(\lambda^n)=\exp n\left\{\int_0^T\lambda^n_r\,dK^{n,f}_r-\int_0^T\mathscr{H}_r(X^n_r,\lambda^n_rf_r)\,dr\right\}$$
$$\le\exp(n\,S(\mu))\exp n\{|M^{1,n}_T|+\underline{\Sigma}^n_T\}.$$

Therefore,

$$\mathscr{P}^n(V)=\int_{X^n\in V}\xi^n_T(\lambda^n)^{-1}\,d\mathscr{P}^n_{\lambda^n}\ge\exp\{-n\,S(\mu)\}\,\Delta^n,$$

where

$$\Delta^n=\int_{X^n\in V}\exp\big\{-n\,|M^{1,n}_T|-n\,\underline{\Sigma}^n_T\big\}\,d\mathscr{P}^n_{\lambda^n}.$$

In the definition of Δ^n, restrict the domain of integration to V_δ and further to the set of paths for which $|M_T^{1,n}| \le \delta$. This gives

$$\Delta^n \ge \exp(-n\,\delta) \int_{X^n \in V_\delta,\, |M_T^{1,n}| \le \delta} \exp(-n \textstyle\sum_T^n)\, d\mathscr{P}_{\lambda^n}^n.$$

Let $\varepsilon > 0$ and choose δ_0 according to Lemma 4.6. For all $\delta \le \delta_0$,

$$\Delta^n \ge e^{-n(\delta+\varepsilon)}\, \mathscr{P}_{\lambda^n}^n\big(X^n \in V_\delta,\ |M_T^{1,n}| \le \delta\big). \tag{4.7}$$

Now,

$$\mathscr{P}_{\lambda^n}^n\big(X^n \in V_\delta,\ |M_T^{1,n}| \le \delta\big) \ge 1 - \mathscr{P}_{\lambda^n}^n(X^n \notin V_\delta) - \mathscr{P}_{\lambda^n}^n(|M_T^{1,n}| > \delta).$$

Lemma 4.2, with $b_t = \lambda_{t-}^n$ and f the function chosen in (4.6), shows that $M_t^{1,n}$ is a $(\mathscr{P}_{\lambda^n}^n, \mathscr{F}_t)$-martingale on I with predictable quadratic variation

$$\langle\!\langle M^{1,n} \rangle\!\rangle_t = \frac{1}{n} \int_0^t (\lambda_r^n)^2\, \mathscr{H}_r''(X_r, \lambda_r^n f_r)\, dr.$$

Thus, by Doob's inequality,

$$\mathscr{P}_{\lambda^n}^n(M_T^{1,n} > \delta) \le \frac{\delta^{-2}}{n} E_{\lambda^n}^n\left[\int_0^T (\lambda_r^n)^2\, \mathscr{H}_r''(X_r^n, \lambda_r^n f_r)\, dr\right] \le \frac{\delta^{-2}}{n} K^2 K_1 T.$$

Hence, summing up,

$$\mathscr{P}_{\lambda^n}^n\big(X^n \in V_\delta,\ |M_T^{1,n}| \le \delta\big) \ge 1 - \mathscr{P}_{\lambda^n}^n\big(X^n \notin V_\delta\big) - \frac{\delta^{-2}}{n} K^2 K_1 T.$$

By Lemma 4.8, we can now choose n_0 and a sequence $\delta_n \to 0$, such that

$$\mathscr{P}_{\lambda^n}^n\big(X^n \in V_\delta,\ |M_T^{1,n}| \le \delta\big) \ge \tfrac{1}{2}, \qquad n \ge n_0.$$

Then by (4.7),

$$\frac{1}{n} \log \Delta^n \ge -(\delta_n + \varepsilon) - \frac{1}{n} \log 2 \to -\varepsilon, \qquad n \to \infty.$$

Finally,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathscr{P}^n(V) \ge -S(\mu) + \limsup_{n\to\infty} \frac{1}{n} \log \Delta^n = -S(\mu) - \varepsilon.$$

Let ε go to zero to complete the proof. □

4.4. *Upper bound.*

PROPOSITION 3. *Fix $\mu \in \mathscr{D}(I, \mathscr{M})$. For every $\delta > 0$, there exists an open neighborhood V of μ such that*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathscr{P}^n(V) \le \begin{cases} -S(\mu) + \delta, & \textit{provided that}\quad S(\mu) < \infty, \\ -\delta, & \textit{otherwise}. \end{cases}$$

PROOF. Fix $\delta > 0$. For any fixed $f \in \mathscr{C}^{1,0}(I \times E)$, by the continuity of the mapping

$$\mathscr{D}(I, \mathscr{M}) \ni \mu \mapsto J_{0,T}(\mu, f),$$

there exists an open neighborhood $V = V(\delta, f)$ of μ such that

$$|J_{0,T}(X^n, f) - J_{0,T}(\mu, f)| \le \delta/2, \quad \text{whenever } X^n \in V.$$

Recall the $\mathscr{P}^n$-martingale $Z^{n,f}$ in Lemma 4.1. We have

$$\begin{aligned} 1 \ge E^n Z_T^{n,f} &\ge E^n[Z_T^{n,f}; X^n \in V] \\ &= E^n[\exp n\{J_{0,T}(X^n, f) - J_{0,T}(\mu, f) + J_{0,T}(\mu, f)\}; X^n \in V]. \end{aligned}$$

Hence,

$$\begin{aligned} \exp(-nJ_{0,T}(\mu, f)) &\ge E^n[\exp(-n|J_{0,T}(X^n, f) - J_{0,T}(\mu, f)|); X^n \in V] \\ &\ge \exp(-n\,\delta/2)\mathscr{P}^n(V). \end{aligned} \tag{4.8}$$

If $S(\mu) < \infty$, then, by (2.2), we can find $f_0 \in \mathscr{C}^{1,0}(I \times E)$ such that

$$J_{0,T}(\mu, f_0) > S(\mu) - \delta/2.$$

Apply (4.8) with f_0 and $V = V(\delta, f_0)$ to obtain

$$\frac{1}{n} \log \mathscr{P}^n(V) \le -J_{0,T}(\mu, f_0) + \delta/2 \le -S(\mu) + \delta.$$

If, instead, $S(\mu) = \infty$, then in view of (2.2) we can find a sequence $f^{(k)}$ in $\mathscr{C}^{1,0}(I \times E)$ such that $J_{0,T}(\mu, f^{(k)}) \to \infty$. In particular, $J_{0,T}(\mu, f^{(k)}) \ge 3\delta/2$ for $k \ge k_0$. By (4.8) applied to the open set $V = V(\delta, f^{(k_0)})$,

$$\frac{1}{n} \log \mathscr{P}^n(V) \le -J_{0,T}(\mu, f^{(k_0)}) + \delta/2 \le -\delta. \qquad \square$$

4.5. *Exponential tightness.* This section is devoted to the proof of the following proposition.

PROPOSITION 4. *For each $m \ge 1$ there is a compact set E_m in $\mathscr{D}(I, \mathscr{M})$ such that*

$$\mathscr{P}(X^n \notin E_m) \le \exp -n\,m.$$

PROOF. In this proof we let $\{f_j, j \ge 1\}$ denote a dense subset of $\mathscr{C}(E)$ such that $\|f_j\| \le j$. Define, for $\mu \in \mathscr{D}(I, \mathscr{M})$,

$$w_k(\mu, \delta, I) = \sup\left\{\sum_{j=1}^k 4^{-j} |\langle \mu_t, f_j\rangle - \langle \mu_s, f_j\rangle| \colon s, t \in I,\ 0 \le t - s < \delta\right\},$$

$$w_k'(\mu, \delta, I) = \inf_{\{t_i\}} \max_i \sup_{t_{i-1} \le s < t < t_i} \sum_{j=1}^k 4^{-j} |\langle \mu_t, f_j\rangle - \langle \mu_s, f_j\rangle|,$$

where the infimum ranges over all partitions $0 = t_0 < t_1 < \cdots < t_{m-1} < T \le t_m$ of I such that $t_i - t_{i-1} > \delta$ for $1 \le i \le m$. For arbitrary null sequences δ_k and ϱ_k, put

$$F_k = \{\mu \in \mathscr{D}(I, \mathscr{M}):\ w'_k(\mu, \delta_k, I) \le \varrho_k\}.$$

The set defined by

$$E_m := \bigcap_{k \ge m} F_k$$

is relatively compact in $\mathscr{D}(I, \mathscr{M})$ [see Ethier and Kurtz (1986), Lemma 6.1]. We will prove that ϱ_k and δ_k can be chosen such that, for each m,

$$\mathscr{P}^n(E_m^c) \le \exp -n\, m,$$

where E_m^c denotes the complement set of E_m.

Since

$$w'_k(\mu, \delta, T) \le w_k(\mu, 2\delta, T),$$

we have

$$\begin{aligned}\mathscr{P}^n(F_k^c) &= \mathscr{P}^n\big(w'_k(X^n, \delta_k, T) > \varrho_k\big) \le \mathscr{P}^n\big(w_k(X^n, 2\delta_k, T) > \varrho_k\big)\\ &= \mathscr{P}^n\Bigg(\sup_{\substack{s,t \in I\\ 0 \le t-s < \delta_k}} \sum_{j=1}^{k} 4^{-j} |\langle X_t^n, f_j\rangle - \langle X_s^n, f_j\rangle| > \varrho_k\Bigg).\end{aligned}$$

Continue with

$$\begin{aligned}\mathscr{P}^n(F_k^c) &\le \sum_{l=0}^{[T/\delta_k]-1} \mathscr{P}^n\Bigg(\sup_{l\delta_k \le t < ((l+2)\delta_k) \wedge T} \sum_{j=1}^{k} 4^{-j} |\langle X_t^n, f_j\rangle - \langle X_{l\delta_k}^n, f_j\rangle| > \frac{\varrho_k}{2}\Bigg)\\ &\le \frac{T}{\delta_k} \mathscr{P}^n\Bigg(\sup_{t < 2\delta_k} \sum_{j=1}^{k} 4^{-j} |\langle X_t^n, f_j\rangle - \langle X_0^n, f_j\rangle| > \frac{\varrho_k}{2}\Bigg).\end{aligned}$$

The next step is to prove the inequality

$$\begin{aligned}&\mathscr{P}\Bigg(\sup_{t < 2\delta_k} \sum_{j=1}^{k} 4^{-j} \big|\langle X_t^n, f_j\rangle - \langle X_0^n, f_j\rangle\big| > \frac{\varrho_k}{2}\Bigg)\\ &\qquad \le 2 \exp\left\{-2n\delta_k C \tau^*\left(\frac{\varrho_k}{4C\delta_k} - 1\right)\right\},\end{aligned} \tag{4.9}$$

where C is the constant in Assumption A2.

To this end, we use again the exponential $\mathscr{P}^n$-martingale $Z_t^{n,f}$ in Lemma 4.1. Introduce

$$g_k = \sum_{j=1}^{k} 4^{-j} f_j \quad \text{with } \|g_k\| \le \sum_{j=1}^{k} 4^{-j} \|f_j\| \le \sum_{j=1}^{k} 4^{-j} j \le \frac{1}{2}.$$

For $\beta > 0$,

$$\sum_{j=1}^{k} 4^{-j}(\langle X_t^n, f_j\rangle - \langle X_0^n, f_j\rangle) = \langle X_t^n, g_k\rangle - \langle X_0^n, g_k\rangle$$
$$= \frac{1}{n}\log Z_t^{n,g_k} + \int_0^t \mathscr{H}_r(X_r^n, g_k)\,dr$$
$$= \frac{1}{\beta n}\log Z_t^{n,\beta g_k} + \frac{1}{\beta}\int_0^t \mathscr{H}_r(X_r^n, \beta g_k)\,dr.$$

We have

$$\mathscr{H}_t(\mu, \beta g_k) \le C(e^{2\beta\|g_k\|} - 1) \le C(e^{\beta} - 1),$$

for any μ, t and k. Hence,

$$\mathscr{P}\left(\sup_{t<2\delta_k}\langle X_t^n, g_k\rangle - \langle X_0^n, g_k\rangle > \frac{\varrho_k}{2}\right)$$
$$\le \mathscr{P}\left(\sup_{t<2\delta_k}\log Z_t^{n,\beta g_k} > 2n\delta_k C\left\{\beta\left(\frac{\varrho_k}{4C\delta_k} - 1\right) - \tau(\beta)\right\}\right),$$

for any $\beta > 0$. Apply Doob's inequality, use $E^n Z_t^{n,\beta g_k} \le 1$ and minimize the right-hand side over $\beta > 0$. Then repeat the same steps for the function $-g_k$. This gives the same bound for the probability $\mathscr{P}(\sup_{t<2\delta_k}\langle X_t^n, g_k\rangle - \langle X_0^n, g_k\rangle < -\varrho_k/2)$. By adding the two estimates we obtain (4.9).

Now take $\delta_k = T/2k^2$. We can find a sequence $\varrho_k \to 0$ with $\varrho_k > 4C\,\delta_k$ and such that $TC\tau^*(\varrho_k/4C\,\delta_k - 1) > 4k^3$. With these choices the right-hand side in (4.9) is less than $2\exp -4n\,k$ and hence

$$\mathscr{P}^n(E_m^c) \le \sum_{k\ge m}\mathscr{P}^n(F_k^c) \le \sum_{k\ge m}\frac{T}{\delta_k}2\exp -4n\,k$$
$$\le \sum_{k\ge m}4k^2\exp -4n\,k \le \sum_{k\ge m}\exp -2n\,k$$
$$\le \exp -n\,m\,(e^{nm} - 1)^{-1} \le \exp -nm. \qquad \square$$

5. Large deviations for the dependent case. In the previous section we proved Theorem 1 for jump measures of the form $\hat{n}(x, dy) := n(x^i, dy\,;\,\hat{\mu})$, where $\hat{\mu}$ was a fixed measure. The purpose of this section is to lift this restriction and obtain the result for a general jump measure with interaction which satisfies Assumption 1. To distinguish the two cases, we let $\hat{\mathscr{P}}^n$ denote the probability law corresponding to $\hat{n}$, for which we have shown that the large deviation principle is true, and we let $\mathscr{P}^n$ denote now the probability law in the general case.

5.1. *A Girsanov transformation.* Let $\hat{N}_t$ denote the point process defined in Section 1.3 corresponding to the jump measure $\hat{n}$ and put

$$\hat{N}_t^E(x^i) = \int_E \hat{N}_r(x_r^i, dy).$$

For any n and state $x = (x_t^1, \dots, x_t^n)_{t\in I}$ of the finite collection of jump processes, define

$$H_t^n(x) = \sum_{i=1}^n \left\{ \int_0^t \log \frac{\gamma_r(x_{r-}^i\,;\, X_{r-}^n)}{\gamma_r(x_{r-}^i\,;\, \hat{\mu}_{r-})}\, d\hat{N}_r^E(x_r^i) - \int_0^t \left(\gamma_r(x_r^i\,;\, X_r^n) - \gamma_r(x_r^i\,;\, \hat{\mu}_r)\right) dr \right\}.$$

LEMMA 5.1. *The change of measure from $\hat{\mathscr{P}}^n$ to $\mathscr{P}^n$ is absolutely continuous and the Radon–Nikodym derivative is given by*

$$\left.\frac{d\mathscr{P}^n}{d\hat{\mathscr{P}}^n}\right|_{\mathscr{F}_t} = \exp H_t^n$$

PROOF. Same as in Shiga and Tanaka (1985). □

We introduce a separate notation for the first term in H_t^n as

$$K_t^n(x) = \sum_{i=1}^n \int_0^t \left(\log \gamma_r(x_{r-}^i\,;\, X_{r-}^n) - \log \gamma_r(x_{r-}^i\,;\, \hat{\mu}_{r-})\right) d\hat{N}_r^E(x_r^i).$$

Moreover, for any real number α, let

$$L_t^\alpha(x) = \sum_{i=1}^n \int_0^t \left[\left(\frac{\gamma_r(x_r^i, X_r^n)}{\gamma_r(x_r^i, \hat{\mu}_r)} \right)^\alpha - 1 \right] dr. \tag{5.1}$$

LEMMA 5.2. *For any $\alpha \in R$,*

$$Y_t := \exp\{\alpha K_t^n - L_t^\alpha\}, \qquad t \in I,$$

is a $(\hat{\mathscr{P}}^n, \mathscr{F}_t)$-local martingale and thus a supermartingale.

PROOF. Apply the exponential formula to get

$$Y_t = 1 + \sum_{i=1}^n \int_0^t Y_{r-} \left[\left(\frac{\gamma_r(x_r^i, X_r^n)}{\gamma_r(x_r^i, \hat{\mu}_r)} \right)^\alpha - 1 \right] \left(d\hat{N}_r^E(x_r^i) - \gamma_r^i(x^i, \hat{\mu}_r)\, dr \right). \qquad \square$$

5.2. *Proof of lower bound.* We shall prove that, for any $\hat{\mu} \in \mathscr{D}(I, \mathscr{M})$ with $S(\hat{\mu}) < \infty$ and open neighborhood V of $\hat{\mu}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathscr{P}^n(V) \ge -S(\hat{\mu}).$$

Set

$$G_t^n = \sum_{i=1}^n \int_0^t \left(\gamma_r(x_r^i\,;\, X_r^n) - \gamma_r(x_r^i\,;\, \hat{\mu}_r)\right) dr.$$

By the uniform continuity of γ we can find for all $\delta > 0$ a neighborhood W of $\hat{\mu}$ with $W \subset V$ such that

$$|G_T^n| \le n\delta \quad \text{uniformly on } W. \tag{5.2}$$

Then, since $H_T^n = K_T^n - G_T^n$,

$$\begin{aligned} \mathscr{P}^n(V) &\ge \mathscr{P}^n(W) \\ &= \hat{E}^n[\exp H_T^n;\ X^n \in W] \\ &\ge \exp(-n\delta)\,\hat{E}^n[\exp K_T^n;\ X^n \in W] \\ &= \exp(-n\delta)\,\hat{E}^n\left[\exp\left\{K_T^n - \frac{p}{q}L_T^{-q/p} + \frac{p}{q}L_T^{-q/p}\right\};\ X^n \in W\right]. \end{aligned} \tag{5.3}$$

Here $L_T^{-\alpha}$ is defined as in (5.2) with $\alpha = p/q > 0$, for a pair of conjugate exponents $p, q > 1$.

Now the uniform continuity of γ yields that

$$|L_T^{-\alpha}| \le C(\alpha)\,\delta\, n \quad \text{uniformly on } W,$$

where $C(\alpha)$ is a constant depending on α only. Hence

$$\mathscr{P}^n(V) \ge \exp(-n\delta)\exp\left(\frac{-n\delta C(\alpha)}{\alpha}\right)\hat{E}^n\left[\exp\left\{K_T^n + \frac{p}{q}L_T^{-q/p}\right\};\ X^n \in W\right],$$

and thus, by Hölder's inequality,

$$\mathscr{P}^n(V) \ge \exp\left(-n\delta\frac{1+C(\alpha)}{\alpha}\right)\hat{E}^n\left[\exp\left\{-\frac{q}{p}K_T^n - L_T^{-q/p}\right\}\right]^{-p/q}\hat{\mathscr{P}}^n(W)^p$$

However, according to Lemma 5.2,

$$\exp\left\{-\frac{q}{p}K_t^n - L_t^{-q/p}\right\}, \qquad t \in I,$$

is a $(\hat{\mathscr{P}}^n)$-supermartingale. Hence

$$\hat{E}^n\left[\exp\left\{-\frac{q}{p}K_T^n - L_T^{-q/p}\right\}\right] \le 1$$

and so

$$\frac{1}{n}\log\mathscr{P}^n(V) \ge -\delta\left(\frac{1+C(\alpha)}{\alpha}\right) + p\,\frac{1}{n}\log\hat{\mathscr{P}}^n(W).$$

From Section 4 we know that Theorem 1 holds for $\hat{\mathscr{P}}^n$. Hence

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathscr{P}^n(V) \ge -\delta\left(\frac{1+C(\alpha)}{\alpha}\right) - p\,S(\hat{\mu}).$$

To complete the proof, take $\delta \to 0$ and then $p \to 1$.

5.3. *Proof of upper bound.*

LEMMA 5.3. *For each $\delta > 0$ there is an open neighborhood $V = V(\delta)$ of $\hat{\mu}$, such that*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathscr{P}^n(V) \le -S(\hat{\mu}) + \delta,$$

whenever $S(\hat{\mu}) < \infty$. If $S(\hat{\mu}) = \infty$, the assertion holds with $-\delta$ on the right side in the inequality.

PROOF. Let $p, q > 1$ be a pair of conjugate exponents and let L_T^p be defined as in (5.1). Choose V small enough such that

$$|L_T^p| \le n\delta C(p)/2, \qquad |G_T^n| \le n\delta/2,$$

uniformly in V. Then, similarly as in (4.3),

$$\mathscr{P}^n(V) \le \exp\left(\frac{n\delta}{2}\right) \hat{E}^n\left[\exp\left\{K_T^n - \frac{1}{p}L_T^p + \frac{1}{p}L_T^p\right\};\ X^n \in W\right].$$

Again by Hölder's inequality,

$$\mathscr{P}^n(V) \le \exp\left(\frac{n\delta}{2}\right) \hat{E}^n\left[\exp\left\{p\, K_T^n - L_T^p\right\}\right]^{1/p} \hat{E}^n\left[\exp\left\{\frac{q}{p}L_T^p\right\};\ X^n \in W\right]^{1/q}.$$

By (5.2) and the supermartingale property in Lemma 5.2,

$$\mathscr{P}^n(V) \le e^{n\delta/2}\, e^{n\delta C(p)q/2p}\ \hat{\mathscr{P}}^n(V)^{1/q},$$

and hence

$$\frac{1}{n}\log \mathscr{P}^n(V) \le \delta\left(\frac{1 + qC(p)}{p}\right)/2 + \frac{1}{nq}\log \hat{\mathscr{P}}^n(V).$$

Finally, by Proposition 4.9,

$$\limsup_{n\to\infty} \frac{1}{n}\log \mathscr{P}^n(V) \le \delta\left(\frac{1 + qC(p)}{p}\right)\Big/2 + \frac{1}{q}\left(-S(\hat{\mu}) + \frac{\delta}{2}\right),$$

with an obvious modification in the case $S(\hat{\mu}) = \infty$. Take $p \to \infty$, that is, $q \to 1$, to complete the proof of the lemma. □

5.4. *Further extensions.* As mentioned in the Introduction, our assumptions on the jump measure are somewhat strong. However, it is not difficult to modify our program to include the case of unbounded jump rates. Then it is natural to consider so-called inductive topologies on $\mathscr{M}$. More exactly, $\{t \mapsto X_t\} \in \bigcup_{m\ge 1} \mathscr{D}_m$, where $\mathscr{D}_m$ is the subspace of $\mathscr{D}(I, \mathscr{M}^{\text{weak}})$ for which $\sup_{t\in I} \langle \mu_t, \varphi_0\rangle \le m$, where φ_0 is some Lyapunov function. For example, analyzing the Schlögl models, Feng (1994) uses $\varphi_0(y) = 1 + y\log\log(3+y)$.

It is likely that one can also apply our methods for other technical setups. We mention the case when $\mathscr{M}$ is equipped with the "Lipschitz-norm" $\|\mu - \nu\| = \sup\{\langle \mu - \nu, f\rangle\colon\ |f(y) - f(z)| \le |y - z|\}$ and $\{t \mapsto X_t\} \in \mathscr{D}(I, \mathscr{M})$ has the

Skorokhod topology. Compare, for example, Oelschläger (1984) and Graham (1992).

For another example, let $(\mathscr{C}, \|\cdot\|)$ be the set of continuous functions on E and $\mathscr{C}_I$ the set of continuous maps $t \mapsto \psi_t \in \mathscr{C}$ with $\sup_{t\in I} \|\psi_t\| < \infty$. Equip $\mathscr{D}(I, \mathscr{M}) \subset \Phi_I^*$ with the relative weak* topology. This yields the coarsest topology on $\mathscr{D}(I, \mathscr{M})$ such that all maps $\nu \mapsto \int_I \langle \nu_t, \psi_t \rangle \, dt$, $\psi \in \Phi_I$, are continuous. Compare Fleischmann and Kaj (1994) and Fleischmann, Gärtner and Kaj (1993).

REFERENCES

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

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

DJEHICHE, B. (1993). A large deviation estimate for ruin probabilities. *Scand. Actuarial J.* **1** 42–59.

ETHIER, S. and KURTZ, T. (1986). *Markov Processes, Characterization and Convergence*. Wiley, New York.

FENG, S. (1994). Large deviations for empirical process of mean field interaction particle system with unbounded jump. *Ann. Probab.* **22** 2122–2151.

FENG, S. and ZHENG, X. (1992). Solutions of a class of nonlinear master equations. *Stochastic Process Appl.* **42** 65–84.

FLEISCHMANN, K. and KAJ, I. (1994). Large deviation probabilities for some rescaled superprocesses. *Ann. Inst. H. Poincaré* **30** 607–645.

FLEISCHMANN, K., GÄRTNER, J. and KAJ, I. (1993). A Schilder type theorem for super-Brownian motion. *Canad. J. Math.* To appear.

GRAHAM, C. (1992). McKean–Vlasov Itô–Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Process. Appl.* **40** 69–82.

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

KRASNOSELSKII, M. and RUTICKII, Y. (1961). *Convex Functions and Orlicz Spaces*. Noordhoff, Groningen.

LÉONARD, C. (1989). Grandes déviation pour des systèmes de processus de Markov avec interaction à longue portée. *C. R. Acad. Sci. Paris Sér. 1* **308** 425–428.

LÉONARD, C. (1990). Some epidemic systems are long range interacting particle systems. *Stochastic Processes in Epidemic Systems. Lecture Notes in Biomath.* **86**. Springer, New York.

NEVEU, J. (1975). *Discrete-Parameter Martingales*. North-Holland, Amsterdam.

OELSCHLÄGER, K. (1984). A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.* **12** 458–479.

SHIGA, T. and TANAKA, H. (1985). Central limit theorem for a system of Markovian particles with mean field interaction. *Z. Wahrsch. Verw. Gebiete* **69** 439–459.

DEPARTMENT OF MATHEMATICS
KTH
S-100 44 STOCKHOLM
SWEDEN

DEPARTMENT OF MATHEMATICS
BOX 480
S-751 06 UPPSALA
SWEDEN