A Weak Interaction Epidemic among Diffusing Particles

by

Ingemar KAJ (*)
Department of Mathematics, Uppsala University
Box 480, S-751 06 Uppsala, Sweden
ingemar.kaj@math.uu.se

Abstract. - A multi-type system of n particles performing spatial motions given by a diffusion process on R^d and changing types according to a general jump process structure is considered. In terms of their empirical measure the particles are allowed to interact, both in the drift of the diffusions as well as in the jump intensity measure for the type motions. In the limit $n \to \infty$ we derive a principle of large deviations from the McKean-Vlasov equation satisfied by the empirical process of the system. The resulting rate function is shown to admit convenient representations.

In particular, the set-up covers a measure-valued model for an epidemic of SIR-type among spatially diffusing individuals. The infection rate is then proportional to the number of infective individuals and their distances to the susceptible one.

(*) Research supported by a Swedish Natural Sciences Research Council grant, contract F-FU 09481.

1. Introduction

1.1 Purpose. The purpose of this report is to provide a multi-type extension, allowing weak interaction in both space and type, of the well-known results of Dawson and Gärtner (1987) [DG] regarding large deviations from the McKean-Vlasov limit for weakly interacting diffusions. This is achieved by integrating more systematically the previous work Djehiche and Kaj (1994) [DK], in which a large deviation result is derived for a class of measure-valued jump processes, with the setting of the Dawson-Gärtner large deviation principle. Necessarilly, some aspects of such an extension will be mere notational rather than substantial. We will try to focus on those parts that are less evident and to point out some techniques from [DK] which can be used as an alternative to those of [DG].

We do not strive for maximal generality, the intention is rather to demonstrate the broad scope of the path-valued approach to large deviations as developed in [DG], in particular to show that it encompasses some natural models for spatial epidemics. As a motivation for this work the further question then arises whether one can apply large deviation techniques to study threshold limit theorems for epidemics, see e.g. Martin-Löf (1987) and Andersson and Djechiche (1994), in a similar manner as Dawson and Gärtner (1989) apply their results to study the phase-transition behaviour of a mean field interaction Curie-Weiss model.

1.2 Model. Fix a set $E \subset R^d$ and let $\mathcal{B}_E$ denote the σ-algebra on E generated by the Borel subset topology. The product space $R^d \times E$ is the one-particle state space and we consider paths of the form $t \mapsto x_t = (u_t, z_t)$, $t \in I := [0, T]$, where u_t is a diffusion process on R^d and z_t a Markov jump process with piecewise constant and rightcontinuous trajectories taking values in E. We interpret a system $x = (x^1, \ldots, x^n)$ of n such particles by saying that at time $t \in I$ the ith particle occupies the position u_t^i and is of type z_t^i. This information is captured by the empirical distribution. Hence, let $\mathcal{M}$ denote the set of probability measures on $R^d \times E$ equipped with the topology of weak convergence and total variation norm $\|\mu\|$. The point measure

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

represents a realization of the state at time t.

To introduce the space and type motions we quote [DG] and [DK] respectively. However, we restrict to the time-homogeneous case. Furthermore, to simplify the presentation we restrict the class of diffusions significantly to the case of bounded drift and covariance coefficients. Let $\mathcal{C}^{2,0}(R^d \times E)$ denote the set of bounded continuous functions $f(x) = f(u, z)$ twice continuously differentiable in u. For a given n-particle state x, let

$$\mathcal{L} f^i(x) = \mathcal{L} f(x^i;\, \mu) = \frac{1}{2} \sum_{k,l=1}^d a^{k,l}(u^i) \frac{\partial^2 f(u^i, z^i)}{\partial u^k \partial u^l} + \sum_{k=1}^d b^k(u^i;\, \mu) \frac{\partial f(u^i, z^i)}{\partial u^k}$$

denote the diffusion generator for $t \mapsto u_t^i$. Assume that the matrix $\{a^{k,l}(u)\}$ is strictly positive definite for each $u \in R^d$ and that for each k, l the map $a^{k,l} : R^d \to R$ is bounded and locally Hölder continuous. Moreover, assume that the functions $(x, \mu) \mapsto b^k(x, \mu) : R^d \times \mathcal{M} \to R$ are bounded and continuous in x and uniformly continuous in μ. Following [DG], introduce the notations

$$(\nabla_a f)^k = \sum_{l=1}^d a^{k,l} \frac{\partial f}{\partial u^k} , \quad (\nabla_a f, \nabla_a g) = \sum_{k,l=1}^d a^{k,l} \frac{\partial f}{\partial u^k} \frac{\partial g}{\partial u^l}$$

and

$$|\nabla_a f|^2 = (\nabla_a f, \nabla_a f).$$

Regarding the type motion, let the jump measure on E have the form

$$\nu^i(x, dy) = \nu(x^i, dy\,; \mu) = \gamma(x^i;\, \mu)\,\pi(x^i, dy), \quad x \in (R^d \times E)^n, \quad y \in E,$$

where the jump intensity measure $\gamma(x^i\,; \mu)\,dt$ gives the rate of a change from type z^i for a particle at u^i when the system is in state x. Moreover, if a change of type occurs then it is of the form $z^i \to y \in B$ with probability $\int_B \pi(x^i, dy)$, all $B \in \mathcal{B}_E$. Assume $\mu \mapsto \gamma(\cdot, \mu)$ is uniformly continuous and assume there is a constant C such that for each n and $1 \le i \le n$

$$\sup_{x \in E^n} \int_E \nu^i(x, dy) \le C.$$

Write

$$\Delta f(x, y) := f(u, y) - f(u, z), \quad x = (u, z) \in R^d \times E,\ y \in E.$$

and define

$$\mathcal{A} f(x^i; \mu) = \int_E \Delta f(x^i, y)\,\nu^i(x, dy).$$

Let $\mathcal{C}^{1,2,0}(I \times R^d \times E)$ denote the set of bounded continuous functions $f_t(x) = f_t(u, z)$ continuously differentiable in t and twice continuously differentiable in u. The total motion generator acting on $\mathcal{C}^{1,2,0}(I \times R^d \times E)$ can then be written

$$(\tfrac{\partial}{\partial t} + \mathcal{L} + \mathcal{A}) f_t = \frac{\partial f_t}{\partial t} + \frac{1}{2}\sum_{k,l=1}^{d} a^{k,l}\frac{\partial^2 f_t}{\partial u^k \partial u^l} + \sum_{k=1}^{d} b^k \frac{\partial f_t}{\partial u^k} + \int_E \Delta f_t\, d\nu.$$

1.3 Martingale problem. We introduce the empirical distribution process

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

Denote the natural duality between $\mathcal{M}$ and the set $\mathcal{C}$ of bounded continuous functions on $R^d \times E$ by

$$\langle \mu, f \rangle = \int f(u, y)\,\mu(du \otimes dy), \quad f \in \mathcal{C}, \quad \mu \in \mathcal{M}.$$

Let $\mathcal{D}(I, \mathcal{M})$ denote the path space of càdlàg functions from I into $\mathcal{M}$ furnished with the usual Skorokhod topology and $\mathcal{C}(I, \mathcal{M})$ its subset of uniformly continuous paths. The trajectories of X are realized in the canonical probability space $[\mathcal{D}(I, \mathcal{M}), (\mathcal{F}_t)_{t \in I}]$, where $\mathcal{F}_t$, $t \in I$, denotes the filtration of σ-algebras generated by the process $t \mapsto x_t$.

Proposition 1.1. *Suppose that weakly interacting generators $\mathcal{L}f(x;\mu)$ and $\mathcal{A}f(x;\mu)$, $x \in R^d \times E$, are given such that the diffusion coefficients a and b and the jump measure ν satisfies the conditions stated above. For each n and each $\mu_0 \in \mathcal{M}$, there exists a measure $\mathcal{P}^n$ with $\mathcal{P}^n(X_0^n = \mu_0) = 1$ such that for each f in $\mathcal{C}^{1,2,0}(I \times R^d \times E)$,*

$$\langle X_t^n, f_t \rangle = \langle X_0^n, f_0 \rangle + \int_0^t \Big\langle X_r^n, \tfrac{\partial}{\partial r} f_r + \mathcal{L}f_r(\cdot\,; X_r^n) + \mathcal{A}f_r(\cdot\,; X_r^n) \Big\rangle \, dr + M_t^f \,,$$

where $M^f = (M_t^f)_{t \in I}$ is a $(\mathcal{P}^n, \mathcal{F}_t)$-martingale. For f, g in $\mathcal{C}^{1,2,0}(I \times R^d \times E)$ the predictable quadratic variation of M^f and M^g is given by

$$\begin{aligned} \langle\!\langle M_\cdot^f, M_\cdot^g \rangle\!\rangle_t = \frac{1}{n} \int_0^t \Big\langle X_r^n, &\left(\nabla_a f_r(\cdot), \nabla_a g_r(\cdot) \right) \\ &+ \int \varDelta f_r(\cdot, y) \varDelta g_r(\cdot, y) \, \nu(\cdot, dy\,; X_r^n) \Big\rangle \, dr \,. \end{aligned}$$

Suppose X_0^n converges weakly to some $\eta_0 \in \mathcal{M}$. Then X^n converges weakly in $\mathcal{D}(I, \mathcal{M})$ to a deterministic path $\eta \in \mathcal{C}(I, \mathcal{M})$ which is the unique solution of the McKean-Vlasov equation

$$\langle \eta_t, f_t \rangle = \langle \eta_0, f_0 \rangle + \int_0^t \big\langle \eta_r, \tfrac{\partial}{\partial r} f_r + \mathcal{L}f_r(\cdot\,; \eta_r) + \mathcal{A}f_r(\cdot\,; \eta_r) \big\rangle \, dr \,. \tag{1.2}$$

Proof; references: For the diffusion part see [DG], section 5.1, noting that we avoid the main difficulties regarding the inductive topology setting by imposing bounded coefficients. Similarly, the assumption of bounded jump rates considerably simplifies the task of showing that the martingale problem is well-posed. In the pure jump case, a more general situation was studied by Feng (1994) using the inductive topology approach. For complete proofs under general assumptions in closely related models including both continuous and jump processes, see e.g. Oelschläger (1984) and Graham (1992). □

We next consider the general martingale problem

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

for functionals $F(\mu) = F(\langle \mu, f \rangle)$, $F \in \mathcal{C}^2(R)$, $\mu \in \mathcal{M}$ and $f \in \mathcal{C}^{2,0}(R^d \times E)$, where M_t^F is a $(\mathcal{P}^n, \mathcal{F}_t)$-local martingale and the infinitesimal measure generator $\mathcal{G}^n$ acts on probability measures $\mu \in \mathcal{M}$ by

$$\begin{aligned} \mathcal{G}^n F(\mu) = &\big\langle \mu, \mathcal{L}f(\mu) \big\rangle \, F'(\langle \mu, f \rangle) + \frac{1}{2n} \big\langle \mu, |\nabla_a f|^2 \big\rangle \, F''(\langle \mu, f \rangle) \\ &+ n \Big\langle \mu, \int \Big(F(\mu + \frac{1}{n}\delta_y - \frac{1}{n}\delta_\cdot) - F(\mu) \Big) \nu(\cdot, dy\,; \mu) \Big\rangle \,. \end{aligned}$$

The corrresponding Hamiltonian operator is defined as

$$\mathcal{H}^n(\mu, f) = e^{-\langle \mu, f\rangle}\, \mathcal{G}^n\, e^{\langle \mu, f\rangle}\,.$$

Hence

$$\mathcal{H}^n(\mu, f) = \big\langle \mu, \mathcal{L}f(\mu)\big\rangle + \frac{1}{2n}\big\langle \mu, |\nabla_a f|^2\big\rangle + \Big\langle \mu, \int n\big(e^{\Delta f/n} - 1\big)\, d\nu(\mu)\Big\rangle\,.$$

It turns out that the scaled Hamiltonian

$$\begin{aligned}\mathcal{H}(\mu, f) &:= \lim_{n\to\infty} \frac{1}{n}\mathcal{H}^n(\mu, nf) \\ &= \big\langle \mu, \mathcal{L}f(\mu)\big\rangle + \frac{1}{2}\big\langle \mu, |\nabla_a f|^2\big\rangle + \Big\langle \mu, \int \big(e^{\Delta f} - 1\big)\, d\nu(\mu)\Big\rangle\,,\end{aligned} \tag{1.3}$$

is a central quantity for the large deviation result (in this case $\frac{1}{n}\mathcal{H}^n(\mu, nf)$ is independent of n).

2. Large deviations result

2.1 Weak form McKean-Vlasov equation. By considering $\mathcal{M}$ as a subset of the Schwartz space $\mathcal{D}'$ of real distributions on $R^d \times E$, the notion of absolutely continuous maps $t \mapsto \mu_t \in \mathcal{D}'$ can be made precise, see [DG], section 4.1. For a discussion relevant to jump processes, see [DK], section 2.1. We will not recall these matters here, but use freely the time derivatives

$$\dot\mu_t = \lim_{h\to 0} h^{-1}\big(\mu_{t+h} - \mu_t\big) \qquad \text{for almost all } t \in I\,,$$

existing in the distribution sense and satisfying the integration by parts formula

$$\int_0^t \big\langle \dot\mu_r, f_r\big\rangle\, dr = \langle \mu_t, f_t\rangle - \langle \mu_0, f_0\rangle - \int_0^t \big\langle \mu_r, \tfrac{\partial}{\partial r} f_r\big\rangle\, dr\,.$$

Here $f \in \mathcal{C}_0^\infty(I \times R^d \times E)$, the set of smooth test functions compactly supported on $R^d \times E$. Introduce a formal adjoint $\mathcal{L}^*(a, b)$ of the diffusion generator $\mathcal{L}$ with coefficients a and b and a formal adjoint $\mathcal{A}^*(\nu)$ of the type generator $\mathcal{A}$ corresponding to the jump measure ν. The McKean-Vlasov equation (1.2) takes the weak form

$$\dot\eta_t = \mathcal{L}^*(a, b)\eta_t + \mathcal{A}^*(\nu)\eta_t\,, \quad \text{almost all} \quad t \in I\,,$$

where again b and ν may depend interactively on η_t.

Next we introduce for any given path $t \to \mu_t \in \mathcal{M}$ two function spaces. First the set $L^2(I, \mu) = L^2\big(I \times R^d \times E;\, dr\, \mu_r(dx)\big)$ of square-integrable functions $f : I \times R^d \times E \to R^d$ with

$$\|f\|_2^2 = \int_0^T \big\langle \mu_r, |f_r|^2\big\rangle\, dr < \infty\,,$$

where $|\cdot|$ refers to the 'Riemannian structure' generated by the matrix a as in section 1.2. Second the Orlicz space

$$\mathcal{O}(I,\mu) = \mathcal{O}\big(I \times R^d \times E^2\,;\, ds\,\mu_s(dx)\,n(x,dy)\big)$$

of strictly positive functions g on $I \times R^d \times E$ for which the norm expression

$$\|g\| = \inf\Big\{\theta > 0 : \int_I \Big\langle \mu_r, \int \big(\frac{g_r}{\theta}\log\frac{g_r}{\theta} - \frac{g_r}{\theta} + 1\big)\,d\nu\Big\rangle\,dr \le 1\Big\}$$

is finite.

Given a, b, ν and the unique solution $t \mapsto \eta_t$, according to Proposition 1.1, we now define a set H consisiting of all deterministic paths $t \mapsto \mu_t$ with the following properties:

(i) $\mu_0 = \eta_0$,

(ii) the map $t \mapsto \mu_t$ defined on I is absolutely continuous,

(iii) there exists $h \in L^2(I,\mu)$ and $g \in \mathcal{O}(I,\mu)$ such that the path μ is the unique solution of the weak McKean-Vlasov equation

$$\dot{\mu}_t = \mathcal{L}^*(a, b + ah_t)\mu_t + \mathcal{A}^*(g_t\,\nu)\mu_t\,, \quad \text{almost all} \quad t \in I\,,$$

that is,

$$\langle \dot{\mu}_t, f\rangle = \Big\langle \mu_t\,, \mathcal{L}f(\mu_t) + h_t\cdot\nabla_a f + \int \Delta f\, g_t\, d\nu(\mu_t)\Big\rangle\,,$$

for $f \in \mathcal{C}_0^\infty(R^d \times E)$ and almost all t. We call H the set of paths admissible for η.

2.2 A large deviation theorem. The following is a version of Theorems I and II in [DK] for the extended interacting multitype diffusions model.

Theorem 2.1. *Suppose a, b and ν are given as above. Fix $\eta_0 \neq 0$ and let η be the McKean-Vlasov limit solution of equation (1.2). For $\mu \in \mathcal{D}(I,\mathcal{M})$ with $\mu_0 = \eta_0$ and such that $t \mapsto \mu_t$ is absolutely continuous define*

$$S(\mu) = \int_0^T \sup_{f\in\mathcal{C}_0^\infty(R^d\times E)} \Big\{\langle\dot{\mu}_t, f\rangle - \mathcal{H}(\mu_t, f)\Big\}\,dt\,,$$

where $\mathcal{H}(\mu,f)$ is the scaled Hamiltonian defined in (1.3). Put $S(\mu) = \infty$ otherwise. For any measurable set A in $\mathcal{D}(I,\mathcal{M})$, put

$$\mathcal{P}^n(A) = \mathcal{P}^n(X^n \in A | X_0^n = \eta_0)\,.$$

Then

(i) for each open subset G of $\mathcal{D}(I,\mathcal{M})$,

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathcal{P}^n(G)\geq -\inf_{\mu\in G}S(\mu)\,,$$

(ii) for each closed subset F of $\mathcal{D}(I,\mathcal{M})$,

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathcal{P}^n(F)\leq -\inf_{\mu\in F}S(\mu)\,,$$

(iii) the level sets $\{\mu\in\mathcal{D}(I,\mathcal{M}):S(\mu)\leq N\}$, $N>0$, are compact.

Moreover,

(2.2)

$$S(\mu)=\sup_f\Big\{\langle\mu_T,f_T\rangle-\langle\mu_0,f_0\rangle-\int_0^T\big\langle\mu_r,\tfrac{\partial}{\partial r}f_r\big\rangle\,dr-\int_0^T\mathcal{H}(\mu_r,f_r)\,dr\Big\}\,,$$

where the supremum is over all $f\in\mathcal{C}^{1,2,0}(I\times R^d\times E)$. Finally, $S(\mu)<\infty$ if and only if $\mu\in H$, the set of paths which are admissible for η, and if $\mu\in H$ then

$$S(\mu)=\int_0^T\Big\langle\mu_r\,,\,\frac{1}{2}|h_r|^2+\int\big(g_r\log g_r-g_r+1\big)\,d\nu\Big\rangle\,dr\,. \tag{2.3}$$

Remark. The basic representation of S in [DG] (when $g=1$) is in terms of the dual form of the squared $L^2(\mu_r)$-norm of h:

$$\begin{aligned}S(\mu)&=\frac{1}{2}\int_0^T\sup_{f\in C_0^\infty(R^d)}\frac{\big\langle\mu_r,h\cdot\nabla_a f\big\rangle^2}{\big\langle\mu_r,|\nabla_a f|^2\big\rangle}\\&=\frac{1}{2}\int_0^T\sup_{f\in C_0^\infty(R^d)}\frac{\big\langle\dot\mu_r-\mathcal{L}^*(a,b)\mu_r,f\big\rangle^2}{\big\langle\mu_r,|\nabla_a f|^2\big\rangle}\,.\end{aligned}$$

We do not know of a similar representation in the multitype case.

2.3 SIR-epidemics. Consider the case when E is discrete and only consists of three states $E=[\text{susc},\text{inf},\text{rem}]$ for the possible types S=*susceptible*, I=*infective* or R=*removed* and the space motion extends with a cemetry position $\dagger$ to $\dot{R}^d=R^d\cup\dagger$. The motion $t\mapsto x_t^i$ signifies that among n individuals the ith follows the type cycle *susc*→*inf*→*rem* while diffusing according to u_t^i, until z_t^i hits *rem* when it is brought to position $\dagger$. Using self-contained notations the empirical process may be written

$$\frac{1}{n}\sum_{i=1}^n\delta_{x^i}=\frac{1}{n}\Big(\sum_{i=1}^{|S|}\delta^{(S)}_{u^i}+\sum_{j=1}^{|I|}\delta^{(I)}_{u^j}+\sum_{k=1}^{|R|}\delta_\dagger\Big)\,.$$

Furthermore, if f is a bounded continuous function on $\dot{R}^d \times E$ put $\varphi_1(u) = f(u, susc)$, $\varphi_2(u) = f(u, inf)$ and extend by letting $f(u, rem) = f(\dagger, rem) = 0$. Then write $\varphi = (\varphi_1, \varphi_2)$ and

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

In general, $\langle \mu, \varphi \rangle = \langle \mu^1, \varphi_1 \rangle + \langle \mu^2, \varphi_2 \rangle$, where μ^1 and μ^2 are measures on R^d.

To define the type interaction, let

$$\lambda(u,v) : R^d \times R^d \mapsto R ,$$

be a bounded, continuous and nonnegative function. Then introduce a function $\Phi(u,v) = (\Phi_1(u,v), \Phi_2(u,v))$, $u, v \in R^d$, by setting

$$\Phi_1(u,v) = 0 , \quad \Phi_2(u,v) = \lambda(u,v)$$

Because the type only changes from susceptible to infective and from infective to removed, the jump measure ν is determined by the rate function $\gamma(x^i\,;\mu)$. We choose the following weakly interacting rates

$$\gamma((susc, u^i)\,;\mu) = \langle \mu, \Phi(u^i, \cdot) \rangle = \langle \mu^2, \lambda(u^i, \cdot) \rangle , \quad \gamma((inf, u^i)\,;\mu) = \rho > 0 .$$

Hence, if X^n is the state at some fixed time when there are $|I|$ infective particles at positions u^j, then

$$\nu((susc, u^i), inf\,; X^n) = \langle I^n, \lambda(u^i, \cdot) \rangle = \frac{1}{n} \sum_{j=1}^{|I|} \lambda(u^i, u^j)$$

is the rate at which a susceptible particle at position u^i contracts the disease. By the second assumption we have adopted the standard model of constant removal rate

$$\nu((inf, u^i), rem\,; X^n) = \rho .$$

To concretize the martingale problem for this model one should note that now

$$\int_0^t \left\langle X_r^n, \mathcal{A} f_r(\cdot\,; X_r^n) \right\rangle dr = \int_0^t \left\langle S_s^n \otimes I_s^n, \lambda(\varphi_2 - \varphi_1) \right\rangle ds - \varrho \int_0^t \langle I_s^n, \varphi_2 \rangle \rangle\, ds ,$$

and

$$\int_0^t \left\langle X_r^n, \int \Delta f_r(\cdot, y) \Delta g_r(\cdot, y)\, \nu(\cdot, dy\,; X_r^n) \right\rangle dr$$
$$= \int_0^t \left\langle S_s^n \otimes I_s^n, \lambda(\varphi_2 - \varphi_1)^2 \right\rangle ds + \varrho \int_0^t \left\langle I_s^n, \varphi_2^2 \right\rangle ds .$$

Moreover,

$$\begin{aligned}\mathcal{H}(\mu,\varphi) =&\big\langle\mu,\mathcal{L}\varphi(\mu)\big\rangle + \frac{1}{2}\big\langle\mu,|\nabla_a\varphi|^2\big\rangle \\ &+\big\langle\mu^1\otimes\mu^2,\lambda(e^{\varphi_2-\varphi_1}-1)\big\rangle + \varrho\,\big\langle\mu^2,(e^{-\varphi_2}-1)\big\rangle\,.\end{aligned}$$

Therefore, as a corollary of Theorem 2.1 we obtain a large deviation principle for the SIR-epidemics with rate function

$$\begin{aligned}&S(\mu) = \int_0^T \big\langle\mu_r\,,\,|h_r|^2\big\rangle \\ &+\int_0^T \Big(\big\langle\mu_r^1\otimes\mu_r^2,\lambda\big(g_r^1\log g_r^1-g_r^1+1\big)\big\rangle + \varrho\,\big\langle\mu_r^2,\big(g_r^2\log g_r^2-g_r^2+1\big)\big\rangle\Big)\,dr\,.\end{aligned}$$

Here $h\in L^2(\mu)$ and $g=(g^1,g^2)\in\mathcal{O}(\mu,\nu)$ are such that for smooth φ, μ is a solution of

$$\big\langle\dot\mu_t,\varphi\big\rangle = \big\langle\mu_t\,,\,\mathcal{L}\varphi(\mu_t)+h_t\cdot\nabla_a\varphi\big\rangle + \big\langle\mu_t^1\otimes\mu^2,\lambda\,(\varphi_2-\varphi_1)\,g_t^1\big\rangle - \big\langle\mu_t^2,\varphi_2\,g_t^2\big\rangle\,.$$

3. Proofs

Section 3.1 is devoted to properties of the integral S and the space H which can be derived without reference to large deviation estimates. Then we start to prove that the function $S(\mu)$ actually is a large deviation rate function. According to the general Theorems 5.2 and 5.3 in [DG] it suffices to derive local lower and upper bounds and check the exponential tightness property. In sections 3.2-3.5 we give proofs for the independent case with locally frozen interaction. The change-of-measure applied to "switch on" true interaction is the topic of section 3.6.

3.1 Representation of S. We show first that $S(\mu)$ has the representation (2.2).

For fixed $\mu\in\mathcal{D}(I,\mathcal{M})$, introduce for $f\in\mathcal{C}^{1,2,0}(I\times R^d\times E)$ and $0\le s\le t\le T$ the functionals

$$J_{s,t}(f) = \langle\mu_t,f_t\rangle - \langle\mu_s,f_s\rangle - \int_s^t \big\langle\mu_r,\tfrac{\partial}{\partial r}f_r\big\rangle\,dr - \int_0^t \mathcal{H}(\mu_r,f_r)\,dr\,,$$

with the Hamiltonian $\mathcal{H}$ defined in (1.3). Similarly, put

$$\ell_{s,t}(f) := \langle\mu_t,f_t\rangle - \langle\mu_s,f_s\rangle - \int_s^t \big\langle\mu_r,\tfrac{\partial}{\partial r}f_r+\mathcal{L}f_r+\mathcal{A}f_r\big\rangle\,dr\,.$$

Hence

$$J_{0,T}(f) = \ell_{0,T}(f) - \int_0^T \Big\langle\mu_r\,,\,\frac{1}{2}|\nabla_a f_r|^2 + \int\tau(\Delta f_r)\,d\nu\Big\rangle\,dr\,. \tag{3.1}$$

Observe that for all $f \in \mathcal{C}_0^\infty(I \times R^d \times E)$,

$$\ell_{s,t}(f) = \int_s^t \Big(\langle \dot{\mu}_r, f_r \rangle - \big\langle \mu_r, \mathcal{L} f_r + \mathcal{A} f_r \big\rangle \Big)\, dr\,.$$

By approximation this implies that $\ell_{s,t}(f)$ can be obtained from $\ell_{0,T}(f)$ by restricting to f non-zero on $[s,t]$ only. Now, for any smooth f,

$$\begin{aligned} S(\mu) &= \int_0^T \sup_{f \in \mathcal{C}_0^\infty(R^d \times E)} \Big\{ \langle \dot{\mu}_t, f \rangle - \mathcal{H}(\mu_t, f) \Big\}\, dt \\ &\geq \int_I \big(\langle \dot{\mu}_r, f_r \rangle - \mathcal{H}(\mu_r, f_r) \big)\, dr = J_{0,T}(f)\,. \end{aligned}$$

Hence the right hand side of (2.2) is bounded by $S(\mu)$:

$$\sup_{f \in \mathcal{C}^{1,2,0}(I \times R^d \times E)} J_{0,T}(f) \leq S(\mu)\,.$$

Before proving the opposite inequality we introduce some more notation. The Orlicz space $\mathcal{O}(I,\mu)$ is obtained from the pair of Young functions

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

In fact, τ and τ^* define a pair of topologically dual Orlicz spaces $(L^\tau(I,\mu)\,,\, \|\cdot\|_\tau)$ and $(L^{\tau^*}(I,\mu)\,,\, \|\cdot\|_{\tau^*})$ of functions on $I \times R^d \times E^2$. Then $\mathcal{O}(\mu,\nu)$ consists of all functions $g = 1 + \tilde{h}$ with $\tilde{h}$ in L^{τ^*}.

Consider the product space $L^2(I,\mu) \times L^\tau(I,\mu)$ equipped with the norm

$$\|(f,\tilde{f})\| = \max\big\{ \|f\|_2, \|\tilde{f}\|_\tau \big\}$$

and its dual product $L^2(I,\mu) \times L^{\tau^*}(I,\mu)$ with norm

$$\|(h,\tilde{h})\|_* = \|h\|_2 + \|\tilde{h}\|_{\tau^*}\,.$$

For our application typical elements of $L^2(I,\mu) \times L^\tau(I,\mu)$ will be (equivalence classes) of the form $(\nabla_a f, \Delta f)$, in one-to-one correspondence with $f \in \mathcal{C}^{1,2,0}(I \times R^d \times E)$.

Lemma 3.2.

$$S(\mu) \leq \sup_{f \in \mathcal{C}^{1,2,0}(I \times R^d \times E)} J_{0,T}(f)\,.$$

Proof: For $f \in \mathcal{C}^{1,2,0}(I \times R^d \times E)$ put $c = \|(\nabla_a f, \Delta f)\|$. By (3.1),

$$\begin{aligned} \frac{1}{c}\ell_{0,T}(f) &\leq J_{0,T}(f/c) + \int_0^T \Big\langle \mu_r\,,\, \frac{1}{2}\Big|\frac{\nabla_a f_r}{\|\nabla_a f\|_2}\Big|^2 + \int \tau\Big(\frac{\Delta f_r}{\|\Delta f\|_\tau}\Big)\, d\nu \Big\rangle\, dr \\ &\leq J_{0,T}(f/c) + \frac{1}{2} + 1\,. \end{aligned}$$

We can assume that the supremum on the right side in the lemma is finite. Hence

$$\ell_{0,T}(f) \le \text{const}\ \|(\nabla_a f, \Delta f)\|\,,$$

and therefore $\ell_{0,T}$ can be viewed as a bounded linear functional on the linear space L_{diff} of all pairs $(\nabla_a f, \Delta f)$ in $L^2(I,\mu) \times L^\tau(I,\mu)$.

By the Riesz representation theorem there exists a unique element $(h, \tilde h)$ in the closure in $L^2(I,\mu) \times L^{\tau^*}(I,\mu)$ of the dual linear space of L_{diff}, such that

$$\ell_{0,T}(f) = \int_0^T \left\langle \mu_r\,,\, h_r \cdot \nabla_a f_r + \int \tilde h_r\, \Delta f_r\, d\nu \right\rangle dr\,, \quad f \in \mathcal{C}^{1,2,0}(I \times R^d \times E)\,.$$

In particular, for $f \in \mathcal{C}_0^\infty(R^d \times E)$ and $0 \le s < t \le T$ and putting $g = \tilde h + 1$,

$$\langle \mu_t, f\rangle = \langle \mu_s, f\rangle + \int_s^t \left\langle \mu_r\,,\, \mathcal{L} f + h_r \cdot \nabla_a f + \int g_r\, \Delta f\, d\nu \right\rangle dr\,.$$

From this it follows that μ is absolutely continuous, c.f. [DG]. Moreover, for almost all t, we have the McKean-Vlasov equation

$$\langle \dot\mu_t, f\rangle = \left\langle \mu_t\,,\, \mathcal{L} f + h_t \cdot \nabla_a f + \int \Delta f\, g_t\, dn_r \right\rangle. \tag{3.3}$$

Now, by (3.3) and (1.3), computing Legendre transforms

$$\begin{aligned} S(\mu) = \int_0^T \sup_{f \in \mathcal{C}_0^\infty(R^d \times E)} & \left\langle \mu_r\,,\, h_r \cdot \nabla_a f + \int \tilde h_r \Delta f\, d\nu \right. \\ & \qquad \left. - \frac{1}{2}|\nabla_a f|^2 - \int \tau(\Delta f)\, d\nu \right\rangle dr \\ \le \int_0^T & \left\langle \mu_r\,,\, \frac{1}{2}|h_r|^2 + \int \tau^*(\tilde h_r)\, d\nu \right\rangle dr\,. \end{aligned}$$

However, for smooth f we have also

$$\begin{aligned} J_{0,T}(f) &= \int_0^T \left\langle \mu_r\,,\, \frac{1}{2}|h_r|^2 + \int \tau^*(\tilde h_r)\, d\nu \right\rangle dr \\ &\quad - \int_0^T \left\langle \mu_r\,,\, \frac{1}{2}\big|h_r - \nabla_a f\big|^2 + \int \big(\tau^*(\tilde h_r) + \tau(\Delta f_r) - h_r \Delta f_r\big)\, dn_r \right\rangle dr\,. \end{aligned}$$

By Young's inequality and approximation in $L^2(I,\mu) \times L^\tau(I,\mu)$, therefore

$$\sup_{f \in \mathcal{C}^{1,2,0}(I \times R^d \times E)} J_{0,T}(f) = \int_0^T \left\langle \mu_r\,,\, \frac{1}{2}|h_r|^2 + \int \tau^*(\tilde h_r)\, d\nu \right\rangle dr\,, \tag{3.4}$$

which finishes the proof of the lemma. $\square$

Lemma 3.5. *For any $\mu \in \mathcal{D}(I,\mathcal{M})$, absolutely continuous and with $\mu_0 = \eta_0$, $\mu \in H$ if and only if $S(\mu) < \infty$. If $\mu \in H$ then (2.3) holds.*

PROOF: Suppose $S(\mu) < \infty$. Then the supremum of the functional $J_{0,T}$ in (2.2) is finite. It was then a byproduct of the proof of the previous lemma that μ solves equation (3.3) for h and g such that μ is admissible for η, that is, $\mu \in H$.

Conversely, suppose $\mu \in H$. Then there exists $h \in L^2(I,\mu)$ and $g \in \mathcal{O}(I,\mu)$, with $\tilde{h} = g - 1 \in L^{\tau^*}(I,\mu)$, such that (3.3) and hence (3.4) hold. But since τ^* has the growth property

$$\sup_t \tau^*(at)/\tau^*(t) < \infty \quad \text{for some } a > 1\,,$$

the integral

$$\int_0^T \Big\langle \mu_r, \int \tau^*(\tilde{h}_r)\, d\nu \Big\rangle\, dr = \int_0^T \Big\langle \mu_r, \Big(\int g_r \log g_r - g_r + 1 \Big)\, d\nu \Big\rangle\, dr$$

is finite, see e.g. Neveu (1975), Appendix. Therefore, by (3.4), $S(\mu) < \infty$ and (2.3) holds.

□

3.2 Independent case, Cramér transformation.
We start to prove Theorem 1.1 in the case when the processes $t \mapsto x_t^i$ are independent. Hence fix a path $\widehat{\mu}$ in $\mathcal{D}(I,\mathcal{M})$ and consider the system with locally frozen interaction $\widehat{\mu}$. This means that the diffusion generator $\mathcal{L}$ has variance matrix a and drift vector $b(\widehat{\mu})$ and the jump measure is of the form $\widehat{\nu} = \nu(x^i, y\,;\widehat{\mu})$.

For $\mu \in \mathcal{D}(I,\mathcal{M})$ and $f \in \mathcal{C}^{1,2,0}(I \times R^d \times E)$, define

$$K_t(\mu, f) = \langle \mu_t, f_t \rangle - \langle \mu_0, f_0 \rangle - \int_0^t \langle \mu_r, \tfrac{\partial}{\partial r} f_r \rangle\, dr\,,$$

and consider the signed measure paths

$$f \mapsto K_t^{n,f} := K_t(X^n, f)\,, \quad t \in I.$$

Then for each $\mathcal{F}_t$-predictable and $\mathcal{P}^n$- $a.s.$ bounded function $(c_t)_{t\in I}$,

$$\xi_t^{n,f}(c) := \exp\big\{ n \int_0^t c_r\, dK_r^{n,f} - n \int_0^t \mathcal{H}_r(X_r^n, c_r f_r)\, dr \big\}\,, \quad t \in I,$$

is a $(\mathcal{P}^n, \mathcal{F}_t)$-martingale. The martingale (or for more general c, the local martingale) $\xi_t^{n,f}(c)$ is called the Esscher-Cramér transform or generalized Cramér transform. The relation

$$\xi_t^{n,f}(c) = \frac{d\mathcal{P}_\lambda^n}{d\mathcal{P}^n}\Big|_{\mathcal{F}_t}\,,$$

defines a new probability measure $\mathcal{P}_c^n$ which is equivalent to $\mathcal{P}^n$ and under $\mathcal{P}_c^n$ subsequent martingales (local martingales) can be derived. In fact, introduce the derivatives

$$\mathcal{H}_t'(\mu, cf) = \langle \mu, \mathcal{L}f\rangle + c\,\langle \mu, |\nabla_a f|^2\rangle + \Big\langle \mu, \int \Delta f e^{c\,\Delta f}\, d\widehat{\nu}\Big\rangle,$$

$$\mathcal{H}_t''(\mu, cf) = \langle \mu, |\nabla_a f|^2\rangle + \Big\langle \mu, \int (\Delta f)^2 e^{c\,\Delta f}\, d\widehat{\nu}\Big\rangle.$$

Then, for each $(\beta_t)_{t\in I}$, $\mathcal{F}_t$-predictable and $\mathcal{P}^n$- *a.s.* bounded, the process

$$M_t^n = \int_0^t \beta_r\, dK_r^{n,f} - \int_0^t \beta_r\, \mathcal{H}_r'(X_r, c_r f_r)\, dr$$

is a $(\mathcal{P}_c^n, \mathcal{F}_t)$-martingale on I with predictable quadratic variation

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

compare [DK], Lemmas 4.1 and 4.2.

Next we identify a specific predictable function $c = \lambda^n$ for which $\mathcal{P}_{\lambda^n}^n$ will be used to derive the lower large deviation bound. As usual $f^+ = f\vee 0$ and $f^- = (-f)\vee 0$ denote positive and negative parts of a realvalued function f. The notation $f^\pm$ refers to either of the two functions.

Lemma 3.6. *Let $t\mapsto \mu_t$ be absolutely continuous. Fix a function $f\in \mathcal{C}_0^{1,2,0}(I\times R^d\times E)$ which satisfies*

$$\gamma^\pm(f) := \inf_{t\in I}\Big\{|\nabla_a f_t|^2 + \int (\Delta f_t^\pm)^2\, d\,\widehat{\nu}_t\Big\} \ge c_1 > 0\,,$$

$$|\langle \dot\mu_t, f_t\rangle| \le c_2\,, \quad \text{for almost all} \quad t\in I\,.$$

Then there exists a unique function $\lambda_t = \lambda(\mu_t, \dot\mu_t, f_t, t)$, such that for almost all $t\in I$

$$\langle \dot\mu_t, f_t\rangle = \mathcal{H}'(\mu_t, \lambda_t f_t)\,,$$

and a unique progressively measurable process $t\mapsto \lambda_t^n = \lambda(X_t^n, \dot\mu_t, f_t, t)$, almost all $t\in I$, with

$$\langle \dot\mu_t, f_t\rangle = \mathcal{H}'(X_t^n, \lambda_t^n f_t)\,.$$

Moreover, there is a constant K such that for all n

$$|\lambda_t|\vee|\lambda_t^n| \le \Big(\big|\langle \dot\mu_t, f_t\rangle\big| + \big|\langle \mu_t, \mathcal{A}f_t + \mathcal{L}f_t\rangle\big|\Big)/(\gamma^+\wedge\gamma^-) \le K\,,$$

and λ^n is continuous in X^n.

PROOF: This is a straightforward extension of [DK], section 4.2. We have indicated the inequality which we use together with the uniform bounds on a, b and ν to obtain the constant K. □

Two of the martingales M^n with $c = \lambda^n$ are used in the proofs below. First, taking $\beta_t = \lambda^n_{t-}$,

$$\begin{aligned} M_t^{1,n} &:= \int_0^t \lambda^n_{r-}\big(\, dK_r^{n,f} - \mathcal{H}'_r(X_r^n, \lambda_r^n f_r)\, dr\big) \\ &= \int_0^t \lambda^n_{r-}\big(\, dK_r^{n,f} - \langle \dot\mu_r, f_r\rangle\, dr\big) \qquad (3.7)\\ \langle\!\langle M_\cdot^{1,n}\rangle\!\rangle_t &= \frac{1}{n}\int_0^t (\lambda_r^n)^2\, \mathcal{H}''_r(X_r^n, \lambda_r^n f_r)\, dr\,. \end{aligned}$$

Second, take $\beta_t = 1$ and $f_t = \varphi$. Since then

$$M_t^{2,n} := K_t^{n,\varphi} - \int_0^t \mathcal{H}'_r(X_r^n, \lambda_r^n\varphi)\, dr = \langle X_t^n, \varphi\rangle - \langle \mu_t, \varphi\rangle$$

is a martingale under $\mathcal{P}^n_{\lambda^n}$, we have $E^n_{\lambda^n}\langle X_t, \varphi\rangle = \langle \mu_t, \varphi\rangle$. Also

$$\langle\!\langle M_\cdot^{2,n}\rangle\!\rangle_t = \frac{1}{n}\int_0^t \mathcal{H}''_r(X_r^n, \lambda_r^n\varphi)\, dr\,.$$

Here

$$\mathcal{H}''_r(X_r^n, \lambda_r^n\varphi) = \Big\langle X_r^n\,,\, |\nabla_a\varphi|^2 + \int (\Delta f)^2 e^{\lambda^n\Delta f}\, d\widehat{\nu}\Big\rangle \le K_1\,. \qquad (3.8)$$

3.3 Lower bound in the independent case. The following proposition is the analog of [DK], Proposition 4.6. To demonstrate the method we repeat the proof partially.

Proposition 3.9. *Fix $\mu \in \mathcal{D}(I, \mathcal{M})$ and let V be an open neighbourhood of μ. Then*

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

We can suppose $S(\mu) < \infty$ and $\mu_0 = \eta_0$. In the course of the proof of Lemma 3.2 is was seen that μ is then absolutely continuous. Take $f \in \mathcal{C}_0^{1,2,0}(I\times R^d\times E)$ as in Lemma 3.6. The proof of the following lemma is identical to [DK], Lemma 4.7.

Lemma 3.10. *Put*

$$\Sigma_T^n = \int_0^T \Big\{ \big|\mathcal{H}_r(\mu_r, \lambda_r f_r) - \mathcal{H}_r(X_r^n, \lambda_r^n f_r)\big| + |\langle \dot\mu_r, f_r\rangle|\, |\lambda_r^n - \lambda_r| \Big\}\, dr\,.$$

Then

$$\Big|\int_0^T \lambda_r^n\, dK_r^{n,f} - \int_0^T \mathcal{H}_r(X_r^n, \lambda_r^n f_r)\, dr\Big| \le S(\mu) + |M_T^{1,n}| + \Sigma_T^n\,.$$

For $\delta > 0$, set

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

The next two lemmas are also direct adaptions from [DK], Lemma 4.7 and 4.8. The proof of the first lemma uses the continuity of λ^n in X^n and the second lemma is an application of Lenglart-Rebolledo's inequality for $M^{2,n}$, using the bound (3.8).

Lemma 3.11. *For any $\varepsilon > 0$ there is a $\delta_0 > 0$ such that if $\delta \le \delta_0$ and $X^n \in V_\delta$, then $\Sigma_T^n \le \varepsilon$.*

Lemma 3.12. *For every $\delta > 0$,*

$$\lim_{n\to\infty} \mathcal{P}^n_{\lambda^n}\big(X^n \notin V_\delta\big) = 0\,.$$

Proof of Proposition 3.9: By Lemma 3.10,

$$\xi_T^{n,f}(\lambda^n) \le e^{n\,S(\mu)} \exp n\big\{|M_T^{1,n}| + \Sigma_T^n\big\}\,.$$

Therefore

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

where

$$\Delta^n = \int_{X^n\in V} \exp\big\{ -n\,|M_T^{1,n}| - n\,\Sigma_T^n\big\}\, d\mathcal{P}^n_{\lambda^n}\,.$$

Restrict to V_δ and further to the set of paths for which $|M_T^{1,n}| \le \delta$. Then

$$\Delta^n \ge e^{-n\,\delta} \int_{X^n\in V_\delta,\, |M_T^{1,n}|\le\delta} e^{-n\,\Sigma_T^n}\, d\mathcal{P}^n_{\lambda^n}\,.$$

For $\varepsilon > 0$ and $\delta \le \delta_0$, by Lemma 3.11

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

Now

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

By Doob's inequality for the martingale in (3.7) using (3.8),

$$\mathcal{P}^n_{\lambda^n}\big(|M_T^{1,n}| > \delta\big) \le \frac{\delta^{-2}}{n}\,K^2\,K_1\,T\,.$$

By this estimate and Lemma 3.12 we can now choose n_0 and $\delta_n \to 0$, such that

$$\mathcal{P}^n_{\lambda^n}\big(X^n \in V_\delta\,,\ |M_T^{1,n}| \le \delta\big) \ge \frac{1}{2}\,, \quad n \ge n_0\,.$$

Then

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

and finally,

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

Let ε go to zero to finish the proof. $\square$

3.4 Upper bound in the independent case.

Proposition 3.13. *Fix $\mu \in \mathcal{D}(I,\mathcal{M})$. For every $\delta > 0$, there exists an open neighbourhood V of μ such that*

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

provided $S(\mu) < \infty$. If $S(\mu) = \infty$ the assertion holds with $-\delta$ on the right side of the inequality.

Proof: We follow [DK], Proposition 4.7. Fix $\delta > 0$. For $\delta' > 0$, put

$$V = V(\delta') := \Big\{\nu : \sup_{t\in I}\ \sup_{f\in\mathcal{C}^{1,2,0}(I\times R^d\times E)} |\langle \nu_t - \mu_t, f_t\rangle| < \delta'\Big\}\,.$$

In this proof we use the notation $J_{s,t}(\mu,f) = J_{s,t}(f)$. From section 3.2 we know that $\xi_t^{n,f}(1) = \exp n\,J_{0,t}(X^n,f)$ is a $\mathcal{P}^n$-martingale. Use $1 \ge E^n\big[\xi_T^{n,f}(1)\,;\,X^n \in V\big]$ to see

$$\exp -n\,J_{0,T}(\mu,f) \ge E^n\Big[\exp -n\big|J_{0,T}(X^n,f) - J_{0,T}(\mu,f)\big|\,;\,X^n \in V\Big]\,.$$

But simple estimates show that for $X^n \in V$ there is a constant C_1 such that

$$|J_{0,T}(X^n, f) - J_{0,T}(\mu, f)| \leq C_1\, \delta' .$$

Choose $\delta' = \delta / C_1$ to get a set V for which

$$\exp -n\, J_{0,T}(\mu, f) \geq e^{-n\,\delta}\, \mathcal{P}^n(V) .$$

Hence

$$\frac{1}{n} \log \mathcal{P}^n(V) \leq -J_{0,T}(\mu, f) + \delta$$

and therefore, by (2.2) if $S(\mu) < \infty$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathcal{P}^n(V) \leq - \sup_{f \in \mathcal{C}^{1,2,0}(I \times R^d \times E)} J_{0,T}(\mu, f) + \delta = -S(\mu) + \delta .$$

If $S(\mu) = \infty$, by (2.2) we can find $f \in \mathcal{C}^{1,2,0}(I \times R^d \times E)$ such that $J_{0,T}(f) \geq 2\delta$. $\square$

3.5 Exponential tightness.
The basic estimate needed to prove compactness of the level sets for S is covered by the following

Proposition 3.14. *For each $m \geq 1$ there is a compact set E_m in $\mathcal{D}(I, \mathcal{M})$ such that*

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

In [DG], Lemma 5.6, a method was developed to obtain such bounds. For jump processes Feng (1994), Lemma 3.9, gave a similar result. Proposition 4.11 in [DK] follows these ideas very closely. It is tedious but straightforward to extend the proofs to the multitype situation. Since the ideas are clear from the given references we feel it is superfluous to repeat them here.

3.6 Large deviation estimates for the dependent case.
We have proved the theorem for jump measures $\widehat{n}(x, dy)$ with respect to some fixed $\widehat{\mu}$. Let $\widehat{\mathcal{P}}^n$ denote the law corresponding to the frozen interaction $\widehat{n}$ and let $\mathcal{P}^n$ now denote the probability law in the general case. The common method to lift this restriction and obtain the result for the interaction model is to follow Shiga and Tanaka (1985) and apply a Girsanov tranformation to go from $\widehat{\mathcal{P}}^n$ to $\mathcal{P}^n$. Again we are somewhat sketchy and refer to [DK], section 5, and the references there. However, we give the form of the derivative $d\mathcal{P}^n / d\widehat{\mathcal{P}}^n$ and a rather complete proof of the lower bound.

For an $n-$particle motion $x = (x_t^1, \ldots, x_t^n)_{t \in I}$, let

$$\xi_t^i = u_t^i - u_0^i - \int_0^t b(u_s^i, \widehat{\mu}_s^n)\, ds\,, \quad \langle\langle \xi_\cdot^i, \xi_\cdot^i \rangle\rangle_t = \int_0^t a(u_s^i)\, ds\,,$$

denote the $d-$dimensional continuous $\widehat{\mathcal{P}}^n$-martingale corresponding to the diffusion part and $N_t(x^i)$ the counting process for the total number of jumps in E starting at x^i, i.e. $N_t(x^i)$ is compensated under $\widehat{\mathcal{P}}^n$ by $\int_0^t \int_E \widehat{\nu}(x_s^i, dy)\, ds$. Define

$$A_t^n(x) = \sum_{i=1}^n \Big\{ \int_0^t \big(b(x_r^i\,;\,X_r^n) - b(x_r^i\,;\,\widehat{\mu}_r)\big) \cdot d\xi_r^i + \int_0^t \log \frac{\gamma(x_{r-}^i\,;\,X_{r-}^n)}{\gamma(x_{r-}^i\,;\,\widehat{\mu}_{r-})}\, dN_r(x_r^i) \Big\},$$

and

$$B_t^n(x) = \sum_{i=1}^n \int_0^t \Big\{ \frac{1}{2} \big| b(x_r^i\,;\,X_r^n) - b(x_r^i\,;\,\widehat{\mu}_r) \big|^2 + \big(\gamma(x_r^i\,;\,X_r^n) - \gamma(x_r^i\,;\,\widehat{\mu}_r)\big) \Big\}\, dr\,.$$

Moreover, for any real α, put

$$B_t^{n,(\alpha)}(x) = \sum_{i=1}^n \int_0^t \Big\{ \frac{\alpha^2}{2} \big| b(x_r^i\,;\,X_r^n) - b(x_r^i\,;\,\widehat{\mu}_r) \big|^2 + \Big(\frac{\gamma_r(x_r^i, X_r^n)}{\gamma_r(x_r^i, \widehat{\mu}_r)}\Big)^\alpha - 1 \Big\}\, dr\,.$$

Then by standard theory the change of measure from $\widehat{\mathcal{P}}^n$ to $\mathcal{P}^n$ is absolutely continuous with Radon-Nikodym derivative

$$\frac{d\mathcal{P}^n}{d\widehat{\mathcal{P}}^n}\Big|_{\mathcal{F}_t} = \exp\big\{ A_t^n - B_t^n \big\}\,,$$

and, for any $\alpha \in R$,

$$Y_t^{(\alpha)} := \exp\{\alpha A_t^n - B_t^{n,(\alpha)}\}\,, \quad t \in I\,,$$

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

To prove the lower bound Proposition 3.9 in the general case use first the uniform continuity in μ of b and γ to find for all $\delta > 0$ a neighborhood W of $\widehat{\mu}$ with $W \subset V$ such that

$$|B_T^n| \le n\delta \quad \text{uniformly on} \quad W$$

and

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

where $C(\alpha)$ is a constant depending on α only. Then, for a pair of conjugate exponents $p, q > 1$,

$$\begin{aligned}
\mathcal{P}^n(V) \ge \mathcal{P}^n(W) &= \widehat{E}^n\big[\exp\big\{A_T^n - B_T^n\big\}\,;\, X^n \in W\big] \\
&\ge e^{-n\delta}\, \widehat{E}^n\big[\exp A_T^n\,;\, X^n \in W\big] \\
&\ge e^{-n\delta}\, e^{-n\delta C(\alpha)/\alpha}\, \widehat{E}^n\Big[\exp\big\{A_T^n + \frac{p}{q} B_T^{n,(-q/p)}\big\}\,;\, X^n \in W\Big]\,.
\end{aligned}$$

Thus, by Hölder's inequality and the supermartingale property of $Y_t^{(-q/p)}$,

$$\frac{1}{n}\log\mathcal{P}^n(V) \geq -\delta(1+C(\alpha)/\alpha) + p\,\frac{1}{n}\log\widehat{\mathcal{P}}^n(W)\,.$$

We know that Theorem 1.1 holds for $\widehat{\mathcal{P}}^n$. Hence

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathcal{P}^n(V) \geq -\delta(1+C(\alpha)/\alpha) - p\,S(\widehat{\mu})\,.$$

Take $\delta \to 0$ and then $p \to 1$ to finish the proof of Proposition 3.9 in the interacting case. The proof of Proposition 3.13 is carried out similarly, compare [DK], Lemma 5.3.

References

[1] H. Andersson, B. Djehiche, Limit theorems for spatial epidemics. (Preprint Trita 1994:0019, Royal Institute of Technology, Stockholm, 1994).

[2] D.A. Dawson, J. Gärtner, Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics* **20** (1987) 247-308.

[3] D.A. Dawson, J. Gärtner, Large deviations, free energy functional and quasi-potential for a mean field model of interacting diffusions. *Mem. Am. Math. Soc.* **78** (1989).

[4] B. Djehiche, I. Kaj, The rate function for some measure-valued jump processes. Preprint, Uppsala University 1994. To appear: *Ann. Probab.*

[5] S. Feng, Large deviations for empirical process of mean field interaction particle system with unbounded jump. To appear: *Ann. Probab.* 1994.

[6] C. Graham, McKean-Vlasov Ito-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stoch. Proc. Appl.* **40** (1992) 69-82.

[7] A. Martin-Löf, Symmetric sampling procedures, general epidemic processes and their threshold limit theorems. *J. Appl. Probab.* **23** (1986) 265-282.

[8] J. Neveu, Discrete-parameter martingales. (North-Holland, Amsterdam, 1975).

[9] K. Oelschläger, A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.* **12** (1984) 458-479.

[10] T. Shiga, H. Tanaka, Central limit theorem for a system of Markovian particles with mean field interaction. *Z. Wahrscheinlichk. verw. Gebiete* **69** (1985) 439-459.