Bernoulli **9**(1), 2003, 49–65

A note on the Laplace–Varadhan integral lemma

P. DEL MORAL[1] and T. ZAJIC[2]

[1]*CNRS-UMR C55830, Université P. Sabatier, F-31062 Toulouse, France.*
E-mail: delmoral@cict.fr
[2]*Institute for Mathematics and its Applications, University of Minnesota, 207 Church St SE, Minneapolis MN 55455, USA. E-mail: zajic@ima.umn.edu*

We propose a complement to the Laplace–Varadhan integral lemma arising in the large-deviations literature. We examine a situation in which the state space may depend on the rate of deviations. We use this framework to discuss large-deviations principles for an abstract class of discrete generation interacting particle systems as well as for pure jump and McKean–Vlasov diffusions.

Keywords: interacting particle systems; Laplace–Varadhan integral lemma; large deviations

1. Introduction

The Laplace–Varadhan integral lemma is a powerful change-of-reference-probability technique which enables the transfer of a large-deviations principle (LDP) from one sequence of probability measures $\{\overline{\mathbb{Q}}^N;\ N \geqslant 1\}$ to another $\{\overline{\mathbb{P}}^N;\ N \geqslant 1\}$. A collection of probability measures $\{\overline{\mathbb{Q}}^N;\ N \geqslant 1\}$ on some metric space (M, d) is said to satisfy an LDP with rate function H if there exists a lower semi-continuous function $H : M \to [0, \infty]$ such that, for each open set $\mathcal{A}$ and for each closed subset $\mathcal{B}$,

$$-\inf_{m\in\mathcal{A}} H(m) \leqslant \liminf_{N\to\infty} \log \frac{1}{N}\overline{\mathbb{Q}}^N(\mathcal{A}) \quad \text{and} \quad \limsup_{N\to\infty} \log \frac{1}{N}\overline{\mathbb{Q}}^N(\mathcal{B}) \leqslant -\inf_{m\in\mathcal{B}} H(m). \tag{1}$$

The rate function is good if it has compact level sets; that is, for each $h \in [0, \infty)$ the level set $\{m \in M;\ H(m) \leqslant h\}$ is compact.

The Laplace–Varadhan integral connection consists of a pair $(\overline{\mathbb{P}}^N, \overline{\mathbb{Q}}^N)$ of absolutely continuous measures on some complete separable metric space (M, d) such that

$$\frac{\mathrm{d}\overline{\mathbb{P}}^N}{\mathrm{d}\overline{\mathbb{Q}}^N}(u) = \exp(NC(u)) \quad \overline{\mathbb{Q}}^N\text{-almost everywhere} \tag{2}$$

for some measurable mapping $C : M \to \mathbb{R}$. When C is bounded continuous, the lemma can be stated as follows: if the sequence $\{\overline{\mathbb{Q}}^N;\ N \leqslant 1\}$ satisfies an LDP with good rate function H then $\{\overline{\mathbb{P}}^N;\ N \leqslant 1\}$ satisfies an LDP with good rate function $H - C$. The above sequences are frequently defined in terms of the image measures

$$\overline{\mathbb{P}}^N = \mathbb{P}^N \circ \pi_N^{-1} \quad \text{and} \quad \overline{\mathbb{Q}}^N = \mathbb{Q}^N \circ \pi_N^{-1}$$

1350–7265

for some probabilities $\mathbb{P}^N$ and $\mathbb{Q}^N$ on some measurable space Ω^N which may depend on N and for some measurable mapping $\pi_N : \Omega^N \to M$. Now we start by observing that if $\mathbb{P}^N$ and $\mathbb{Q}^N$ are aboslutely continuous and if, for $\mathbb{Q}^N$-almost every $x \in \Omega^N$,

$$\frac{\mathrm{d}\mathbb{P}^N}{\mathrm{d}\mathbb{Q}^N}(x) = \exp(NC(\pi_N(x))) \tag{3}$$

then the probability images $\overline{\mathbb{P}}^N$ and $\overline{\mathbb{Q}}^N$ are absolutely continuous and their Radon–Nykodim derivative satisfies (2), and the Laplace–Varadhan lemma applies if C is a bounded continuous mapping. In this paper we propose a strategy to relax the analytic representation (3) of the Radon–Nikodym derivative of $\mathbb{P}^N$ with respect to $\mathbb{Q}^N$. Our approach involves replacing $\mathbb{P}^N$ and $\mathbb{Q}^N$ by a pair of sequences $\mathbb{P}^N_{\alpha,m}$ and $\mathbb{Q}^N_m$ indexed respectively by a parameter pair (α, m) with $\alpha \in \mathbb{R}$ and $m \in M$ and by a parameter $m \in M$. Instead of (3) we suppose that for any index pair $(\alpha, m) \in (\mathbb{R} \times M)$ we have $\mathbb{P}^N_{\alpha,m} \sim \mathbb{Q}^N_m$ and, for $\mathbb{Q}^N_m$-almost every $x \in \Omega^N$,

$$\frac{\mathrm{d}\mathbb{P}^N_{\alpha,m}}{\mathrm{d}\mathbb{Q}^N_m}(x) = \exp(N[\alpha S_N(x,\, m) + C_\alpha(\pi_N(x),\, m)]) \tag{4}$$

for some measurable functions $S_N : \Omega^N \times M \to \mathbb{R}$ and $C_\alpha : M \times M \to \mathbb{R}$. We also assume that $\mathbb{P}^N_{1,m}$ is independent of m and we denote the former by $\mathbb{P}^N_1$. For any $(\alpha, m) \in (\mathbb{R} \times M)$ we define the image measures

$$\overline{\mathbb{P}}^N_{\alpha,m} = \mathbb{P}^N_{\alpha,m} \circ \pi_N^{-1} \quad \text{and} \quad \overline{\mathbb{Q}}^N_m = \mathbb{Q}^N_m \circ \pi_N^{-1}.$$

We are now in position to state our main result.

Lemma 1.1 (Integral lemma). *Suppose the sequence of probability measures $\{\overline{\mathbb{Q}}^N_m;\, N \geqslant 1\}$ satisifes an LDP with good rate function $H_m : M \to [0, \infty]$, for each $m \in M$. Also assume that the mappings $\{C_\alpha(\cdot,\, m);\, \alpha \in \mathbb{R}\}$ are continuous at each m, $C_\alpha(m, m) = 0$ and the exponential moment condition*

$$\limsup_{N\to\infty} \frac{1}{N} \log \int_{\Omega^N} \exp nN[S_N(x,\, m) + C_1(\pi_N(x),\, m)]\mathrm{d}\mathbb{Q}^N_m(x) < \infty \tag{5}$$

holds for some $(m, n) \in M \times (1, \infty)$. Then $\{\overline{\mathbb{P}}^N_1;\, N \geqslant 1\}$ satisfies an LDP with good rate function

$$I : m \in M \to I(m) = H_m(m) \in [0, \infty].$$

We remark that (5) holds when the mappings $C_1(\cdot,\, m)$ and $C_n(\cdot,\, m)$ are bounded for some $(m, n) \in M \times (1, \infty)$. To see this, note that

$$
\begin{aligned}
&\int_{\Omega^N} \exp nN[S_N(x,\, m) + C_1(\pi_N(x),\, m)]\mathrm{d}\mathbb{Q}_m^N(x)\\
&\quad = \int_{\Omega^N} \exp N[nC_1(\pi_N(x),\, m) - C_n(\pi_N(x),\, m)]\mathrm{d}\mathbb{P}_{n,\,m}^N(x)\\
&\quad = \int_M \exp N[nC_1(u,\, m) - C_n((u,\, m))]\mathrm{d}\overline{\mathbb{P}}_{n,\,m}^N(u)\\
&\quad \leqslant \exp(\|C_1 - C_n/n\|Nn).
\end{aligned}
$$

Notice that when $S_N = 0$ are the null mappings then we have, for any $u \in M$ and $N \geqslant 1$,

$$
\frac{\mathrm{d}\overline{\mathbb{P}}_1^N}{\mathrm{d}\overline{\mathbb{Q}}_m^N}(u) = \exp N[C_1(u,\, m)]
$$

If $C_1(\cdot,\, m)$ is continuous and (5) holds, Varadhan's integral lemma says that the family of probability measures $\overline{\mathbb{P}}_1^N = \mathbb{P}_1^N \circ \pi_N^{-1}$ satisfies the LDP with rate function $I : M \to [0,\, \infty]$ given, for any $u \in M$, by

$$
I(u) = H_m(u) - C_1(u,\, m).
$$

If $C_1(\cdot,\, m)$ is continuous and (5) holds for all m, since $C_1(u,\, u) = 0$ we have that

$$
I(u) = H_u(u).
$$

Note that under our assumptions we only require the continuity of $C_1(\cdot,\, m)$ at the point m and that (5) holds for one m. Therefore, even when $S_N = 0$, our result does not follow from the Laplace–Varadhan integral lemma.

To illuminate the structure of the Radon–Nikodym derivative (4) we discuss the different roles played by the two parameters $(m,\, \alpha) \in M \times \mathbb{R}$. One natural and very useful strategy in many applications of large deviations is to find judicious reference probability measures under which the random sequence at hand satisfies an LDP with a good rate function. The next stage consists of transferring this result to the desired sequence of distributions.

The choice of the reference sequence $\mathbb{Q}_m^N$, $m \in M$, is often dictated by the problem at hand. In the interacting particle system (IPS) context $\mathbb{Q}_m^N$ is often chosen as an N-fold tensor product measure so that the particle are $\mathbb{Q}_m^N$-independent. In this situation Sanov's theorem tells us that an LDP holds with a good rate function. Using the integral lemma the LDP transfer is guaranteed provided that we can find a collection of distributions $\mathbb{P}_{\alpha,m}^N$ satisfying (4). Intuitively speaking, the pair $(\alpha,\, m)$ can be regarded as a deformation parameter of the sequence of measures $\mathbb{P}_1^N$. In the IPS context each $\mathbb{P}_{\alpha,m}^N$ is the distribution law of an N-IPS model with an interaction function depending on the parameters $\alpha \in \mathbb{R}$ and $m \in M$. In some sense $(\alpha,\, m)$ measures the strength of interaction. For instance, in the forthcoming examples, when $\alpha \to 0$ the particles become independent.

2. Proof of the integral lemma

The following technical proposition states the exponential tightness property and two key estimates needed to prove our result.

Proposition 2.1. *Under the assumptions of Lemma* 1.1 *the sequence of probability measures* $\overline{\mathbb{P}}_1^N$ *on* M *is exponentially tight. For any Borel subset* $A \subset M$ *and for any* $1/n + 1/n' = 1$, $1 < n,\ n' < \infty$, *and* $m \in M$, *we have*

$$\overline{\mathbb{P}}_1^N(A) \leqslant \overline{\mathbb{Q}}_m^N(A)^{1/n'}\overline{\mathbb{P}}_{n,m}^N(A)^{1/n}\exp[N\delta_n(m, A)], \tag{6}$$

$$\overline{\mathbb{Q}}_m^N(A) \leqslant \overline{\mathbb{P}}_1^N(A)^{1/n}\overline{\mathbb{P}}_{\alpha,(n),m}^N(A)^{1/n'}\exp[N\delta_{\alpha(n)}(m, A)/n] \tag{7}$$

with $\alpha(n) = -n'/n$, *and, for any* $\alpha \neq 0$,

$$\delta_\alpha(m, A) = \sup_{u \in A}|C_1(u, m) - C_\alpha(u, m)/\alpha|.$$

Since the proof of this proposition is rather technical, we have chosen to present first how it is used to prove our main result. By exponential tightness, we already know from Puhalskii (1991) that there exists a subsequence along which an LDP holds with good rate function.

Proof of Lemma 1.1. If we take A in (6) to be the closure of the ball of radius ϵ and centre $m \in M$, that is,

$$A = \overline{B}(m, \epsilon) = \{u \in M : d(u, m) \leqslant \epsilon\},$$

we find that for any conjugate integers $1/n + 1/n' = 1$, with $1 < n,\ n' < \infty$,

$$\overline{\mathbb{P}}_1^N(\overline{B}(m, \epsilon)) \leqslant \overline{\mathbb{Q}}_m^N(\overline{B}(m, \epsilon))^{1/n'}\exp[N\delta_n(m, \overline{B}(m, \epsilon))].$$

Recalling that $\{\overline{\mathbb{Q}}_m^N;\ N \geqslant 1\}$ satisfies the LDP with good rate function H_m, this implies that

$$\limsup_{N\to\infty}\frac{1}{N}\log\overline{\mathbb{P}}_1^N(\overline{B}(m, \epsilon)) \leqslant -\frac{1}{n'}\inf_{\overline{B}(m,\epsilon)} H_m + \delta_n(m, \overline{B}(m, \epsilon)).$$

Since H_m is a good rate function and $\{\overline{B}(m, \epsilon);\ \epsilon > 0\}$ is a nested family of closed sets, that is,

$$\overline{B}(m, \epsilon) \subset \overline{B}(m, \epsilon') \qquad \text{if } \epsilon < \epsilon',$$

from Lemma 4.1.6 in Dembo and Zeitouni (1993, p. 104) we have

$$H_m(m) = \lim_{\epsilon\to 0}\inf_{\overline{B}(m,\epsilon)} H_m.$$

Since each mapping $C_n(\cdot, m) : M \to \mathbb{R}$ is continuous at the point m and $C_n(m, m) = 0$ by definition of δ_n we also have that

$$\lim_{\epsilon \to 0} \delta_n(m, \bar{B}(m, \epsilon)) = 0.$$

Considering the limit $\epsilon \downarrow 0$, one obtains, for any $n' > 1$,

$$\lim_{\epsilon \to 0} \limsup_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{P}}_1^N (\bar{B}(m, \epsilon)) \leqslant -\frac{1}{n'} H_m(m).$$

Letting $n' \to 1$, we obtain

$$\lim_{\epsilon \to 0} \limsup_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{P}}_1^N (\bar{B}(m, \epsilon)) \leqslant -I(m). \tag{8}$$

Now if we take A in (7) to be the open ball

$$A = B(m, \epsilon) = \{u \in M : d(u, m) < \epsilon\},$$

we obtain

$$\overline{\mathbb{Q}}_m^N (B(m, \epsilon)) \leqslant \overline{\mathbb{P}}_1^N (B(m, \epsilon))^{1/n} \exp[N \delta_{\alpha(n)}(m, B(m, \epsilon))/n].$$

Our assumptions on $\{\overline{\mathbb{Q}}_m^N;\, N \geqslant 1\}$ imply that

$$-H_m(m) \leqslant -\inf_{B(m,\epsilon)} H_m \leqslant \liminf_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{Q}}_m^N (B(m, \epsilon)).$$

Arguing as above, this implies that

$$-I(m) \leqslant \liminf_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{Q}}_m^N (B(m, \epsilon)) \leqslant \frac{1}{n} \left[\liminf_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{P}}_1^N (B(m, \epsilon)) + \delta_{\alpha(n)}(m, B(m, \epsilon)) \right].$$

Considering the limit $\epsilon \downarrow 0$, one obtains, for any $n > 1$,

$$-nI(m) \leqslant \lim_{\epsilon \to 0} \liminf_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{P}}_1^N (B(m, \epsilon)).$$

Letting $n \to 1$, we obtain from (8) that

$$\lim_{\epsilon \to 0} \limsup_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{P}}_1^N (\bar{B}(m, \epsilon)) \leqslant -I(m) \leqslant \lim_{\epsilon \to 0} \liminf_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{P}}_1^N (B(m, \epsilon)).$$

Since

$$\lim_{\epsilon \to 0} \liminf_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{P}}_1^N (B(m, \epsilon)) \leqslant \lim_{\epsilon \to 0} \limsup_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{P}}_1^N (\bar{B}(m, \epsilon)),$$

it follows that

$$-I(m) = \lim_{\epsilon \to 0} \limsup_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{P}}_1^N (\bar{B}(m, \epsilon)) = \lim_{\epsilon \to 0} \liminf_{N \to \infty} \frac{1}{N} \log \overline{\mathbb{P}}_1^N (B(m, \epsilon)).$$

As noticed in Remark 2.2 in Feng and Kurtz (2000), this statement implies that $\overline{\mathbb{P}}_1^N$ satisfies the weak LDP (i.e. with upper bounds only for compact sets). Since the sequence $\overline{\mathbb{P}}_1^N$ is exponentially tight, the weak LDP is equivalent to the full LDP (see for instance Lemma 1.2.18 in Dembo and Zeitouni 1993, p. 8) and the proof of the lemma is now complete. □

We now come to proof of the technical proposition.

Proof of Proposition 2.1. Fixing $n > 1$ so that (5) holds and denoting the left-hand side of (5) by $nc_n/2$, we have, for N large enough,

$$\int \left(\frac{d\mathbb{P}_1^N}{d\mathbb{Q}_m^N}\right)^n d\mathbb{Q}_m^N = \int_{\Omega^N} \exp nN[S_N(x, m) + C_1(\pi_N(x), m)]\, d\mathbb{Q}_m^N(x)$$

$$\leqslant \exp(nc_nN). \tag{9}$$

Since the sequence $\{\overline{\mathbb{Q}}_m^N;\ N \geqslant 1\}$ satisfies a full LDP, $\{\overline{\mathbb{Q}}_m^N;\ N \geqslant 1\}$ is exponentially tight (see for instance Exercise 4.1.10 in Dembo and Zeitouni 1993, p. 105). For any $a < \infty$ there exists a compact set $K(m, a) \subset M$ such that

$$\limsup_{N\to\infty} \frac{1}{N} \log \overline{\mathbb{Q}}_m^N(K^c(m, a)) < -a \qquad \text{with } K^c(m, a) = M - K(m, a).$$

To prove that $\{\overline{\mathbb{P}}_1^N;\ N \geqslant 1\}$ is exponentially tight, we first notice that

$$\overline{\mathbb{P}}_1^N(K_n^c(m, a)) = \mathbb{P}_1^N(1_{K_n^c(m,a)}(\pi_N(x))),$$

with

$$\frac{1}{n} + \frac{1}{n'} = 1 \quad \text{and} \quad K_n^c(m, a) \equiv K^c(m, n'(c_n + a)).$$

Thus, using Holder's inequality, we verify that

$$\overline{\mathbb{P}}_1^N(K_n^c(m, a)) = \mathbb{Q}_m^N\left((1_{K_n^c(m,a)} \circ \pi_N)\frac{d\mathbb{P}_1^N}{d\mathbb{Q}_m^N}\right)$$

$$\leqslant \overline{\mathbb{Q}}_m^N(K_n^c(m, a))^{1/n'}\mathbb{Q}_m^N\left(\left(\frac{d\mathbb{P}_1^N}{d\mathbb{Q}_m^N}\right)^n\right)^{1/n}$$

$$\leqslant \overline{\mathbb{Q}}_m^N(K_n^c(m, a))^{1/n'}\exp(c_nN).$$

Recalling (9), the above estimate implies that

$$\limsup_{N\to\infty} \frac{1}{N} \log \overline{\mathbb{P}}_1^N(K_n^c(m, a)) < -\frac{1}{n'}[n'(c_n + a)] + c_n = -a.$$

This clearly ends the proof of the exponential tightness of the sequence $\{\overline{\mathbb{P}}_1^N;\ N \geqslant 1\}$. In the same way, for any Borel subset $A \subset M$ and for any $1/n + 1/n' = 1$, $1 < n, n' < \infty$, and $m \in M$, we have

$$\overline{\mathbb{P}}_1^N(A) = \mathbb{Q}_m^N\left((1_A \circ \pi_N)\frac{\mathrm{d}\mathbb{P}_1^N}{\mathrm{d}\mathbb{Q}_m^N}\right)$$

$$\leqslant \overline{\mathbb{Q}}_m^N(A)^{1/n'}\mathbb{Q}_m^N\left((1_A \circ \pi_N)\left(\frac{\mathrm{d}\mathbb{P}_1^N}{\mathrm{d}\mathbb{Q}_m^N}\right)^n\right)^{1/n}$$

$$\leqslant \overline{\mathbb{Q}}_m^N(A)^{1/n'}\mathbb{Q}_m^N((1_A \circ \pi_N)\exp(nN[S_N(x, m) + C_1(\pi_N(x), m)]))^{1/n}.$$

Since we have

$$\mathbb{Q}_m^N((1_A \circ \pi_N)\exp(nN[S_N(\cdot, m) + C_1(\pi_N(\cdot), m)]))$$

$$= \mathbb{Q}_m^N((1_A \circ \pi_N)\exp(N[nS_N(\cdot, m) + C_n(\pi_N(\cdot), m)])$$

$$\times \exp(-N[C_n(\pi_N(\cdot), m) - nC_1(\pi_N(\cdot), m)]))$$

$$= \mathbb{P}_{n,m}^N((1_A \circ \pi_N)\exp(-N[C_n(\pi_N(\cdot), m) - nC_1(\pi_N(\cdot), m)]))$$

$$\leqslant \overline{\mathbb{P}}_{n,m}^N(A)\exp\left(N\sup_{u\in A}|nC_1(u, m) - C_n(u, m)|\right) = \overline{\mathbb{P}}_{n,m}^N(A)\exp(nN\delta_n(m, A)),$$

we find that

$$\overline{\mathbb{P}}_1^N(A) \leqslant \overline{\mathbb{Q}}_m^N(A)^{1/n'}\overline{\mathbb{P}}_{n,m}^N(A)^{1/n}\exp(N\delta_n(m, A)).$$

This establishes (6).

To prove (7) we first use the decomposition

$$1_A(\pi_N(x)) = \left[1_A(\pi_N(x))\exp\left(\frac{N}{n}[S_N(x, m) + C_1(\pi_N(x), m)]\right)\right]$$

$$\times \left[1_A(\pi_N(x))\exp\left(-\frac{N}{n}[S_N(x, m) + C_1(\pi_N(x), m)]\right)\right]$$

and Holder's inequality to prove that

$$\overline{\mathbb{Q}}_m^N(A) \leqslant \mathbb{Q}_m^N(1_A \circ \pi_N \exp(N[S_N(\cdot, m) + C_1(\pi_N(\cdot), m)]))^{1/n}$$

$$\times \mathbb{Q}_m^N\left(1_A \circ \pi_N \exp\left(-N\frac{n'}{n}[S_N(\cdot, m) + C_1(\pi_N(\cdot), m)]\right)\right)^{1/n'} \tag{10}$$

$$= \overline{\mathbb{P}}_1^N(A)^{1/n} \times \mathbb{Q}_m^N\left(1_A \circ \pi_N \exp\left(-N\frac{n'}{n}[S_N(\cdot, m) + C_1(\pi_N(\cdot), m)]\right)\right)^{1/n'}.$$

We now observe that

$$
\begin{aligned}
&\mathbb{Q}_m^N(\left(1_A \circ \pi_N \exp\left(-N\frac{n'}{n}[S_N(\cdot,\ m) + C_1(\pi_N(\cdot),\ m)]\right)\right)\\
&\qquad = \mathbb{Q}_m^N(1_A \circ \pi_N \exp(N\alpha(n)[S_N(\cdot,\ m) + C_1(\pi_N(\cdot),\ m)]))\\
&\qquad = \mathbb{Q}_m^N(1_A \circ \pi_N \exp(N[\alpha(n)S_N(\cdot,\ m) + C_{\alpha(n)}(\pi_N(\cdot),\ m)])\\
&\qquad\quad \times \exp(N[\alpha(n)C_1(\pi_N(\cdot),\ m) - C_{\alpha(n)}(\pi_N(\cdot),\ m)]))\\
&\qquad \leqslant \overline{\mathbb{P}}^N_{\alpha(n),m}(A) \times \exp(N|\alpha(n)|\delta_{\alpha(n)}(m,\ A)).
\end{aligned}
$$

We find from (10) that

$$\overline{\mathbb{Q}}_m^N(A) \leqslant \overline{\mathbb{P}}_1^N(A)^{1/n}\overline{\mathbb{P}}^N_{\alpha(n),m}(A)^{1/n'} \times \exp(N\delta_{\alpha(n)}(m,\ A)/n).$$

This establishes (7) and the proof of the proposition is complete. □

3. Large deviations for interacting particle systems

We illustrate the impact of the integral lemma in the context of IPS models. Our general and abstract context is ideally suited to treating within the same framework discrete generation IPS as well as pure jump and McKean–Vlasov IPS diffusions.

We let $I = \{0,\ 1,\ \ldots,\ T\} \subset \mathbb{N}(I = [0,\ T] \subset \mathbb{R}_+ = [0,\ \infty))$ be a discrete (continuous) time index with a finite time horizon $T \in \mathbb{N}(T \in \mathbb{R}_+)$. For E a complete separable metric space we denote by $\mathcal{P}(E)$ the set of all probability measures on E furnished with the weak topology. By $D(I,\ E)$ we denote the set of all cadlag paths from I into E with the Skorohod metric (when $I = \{0,\ 1,\ \ldots,\ T\}$ the set is simply the product space E^{T+1}). We also denote by $\phi(\eta) = (\eta_t)_{t\in I}$ the distribution flow of the marginals with respect to time of a given measure $\eta \in \mathcal{P}(\mathcal{D}(I,\ E))$. Finally, and with a slight abuse of notation, we denote by $m(x)$ the empirical measure associated with a point in a given product space $x = (x^1,\ \ldots,\ x^N) \in E^N$, that is,

$$m(x) = \frac{1}{N}\sum_{i=1}^N \delta_{x^i},$$

where δ_a stands for the Dirac measure at $a \in E$.

In the discrete time situation we start with a distribution $\eta_0 \in \mathcal{P}(E)$ and a collection of Markov transitions $k = \{K_{n,\eta};\ n \in I,\ \eta \in \mathcal{P}(E)\}$. For any distribution flow $\gamma = (\gamma_n)_{n\in I} \in \mathcal{P}(E)^{T+1}$, we denote by $\mathbb{Q}_\gamma$ the measure on E^{T+1} defined by

$$\mathbb{Q}_\gamma(\mathrm{d}(x_0,\ \ldots,\ x_T)) = \eta_0(\mathrm{d}x_0)K1,\gamma_0(x_0,\ \mathrm{d}x_1)\ \ldots\ K_{T,\gamma T-1}(x_{T-1},\ \mathrm{d}x_T).$$

It is important to notice that the McKean distribution defined by

$$\mathbb{P}(\mathrm{d}(x_0,\ \ldots,\ x_T)) = \eta_0(\mathrm{d}x_0)K_{1,\eta_0}(x_0,\ \mathrm{d}x_1)\ \ldots\ K_T,\eta_{T-1}(x_{T-1},\ \mathrm{d}x_T)$$

and associated with the distribution flow

$$\eta_n = \eta_{n-1} K_{n,\eta_{n-1}} = \int \eta_{n-1}(\mathrm{d}x) K_{n,\eta_{n-1}}(x, \cdot)$$

is a fixed point of the mapping $\eta \in \mathcal{P}(D(I, E)) \to \mathbb{Q}_{\phi(\eta)} \in \mathcal{P}(D(I, E))$. As usual, $\mathbb{P}$ can be interpretated as the law of a time-inmogeneous Markov process X_n with elementary transitions $K_{n,\eta_{n-1}}$ with $\eta_{n-1} = \mathbb{P} \circ X_{n-1}^{-1}$. More precisely, $\mathbb{P}$ is the solution of the following discrete time and time-inhomogeneous martingale problem defined on the canonical space $(\Omega, F, \chi) = (D(I, E), (F_t)_{t \geqslant 0}, (X_t)_{t \in I})$:

- $\mathbb{P} \circ X_0^{-1} = \eta_0$.
- For any bounded measurable test function f on E the sequence

$$M_n(f) = f(X_n) - f(X_0) - \sum_{p=1}^{n} [f(X_p) - K_{p,\eta_{p-1}}(f)(X_{p-1})]$$

  is an F-martingale under $\mathbb{P}$, and the distribution flow $\{\eta_n;\ n \in I\}$ coincides with the set of time marginals of $\mathbb{P}$, that is, $\mathbb{P} \circ X_n^{-1} = \eta_n$.

The N-IPS associated to the collection K and the distribution $\eta_0 \in \mathcal{P}(E)$ is an E^N-valued Markov chain $\xi_n = (\xi_n^1, \ldots, \xi_n^N)$, $n \in I$, with initial distribution $\eta_0^{\otimes N}$ and elementary transitions

$$\Pr(\xi_n \in \mathrm{d}(x^1, \ldots, x^N) | \xi_{n-1}) = \prod_{i=1}^{N} K_{n,m(\xi_{n-1})}(\xi_{n-1}^i, \mathrm{d}x^i).$$

In the continuous time situation we start with a distribution $\eta_0 \in \mathcal{P}(E)$ and a collection of generators $L = \{L_{t,\mu};\ t \in I,\ \mu \in \mathcal{P}(E)\}$ defined on some dense domain $D(L)$ in the space of bounded continuous functions. For any distribution flow $\gamma = (\gamma_t)_{t \in I}$ with $\gamma_t \in \mathcal{P}(E)$, $t \in I$, we suppose there exists a solution $\mathbb{Q}_\gamma \in \mathcal{P}(\mathcal{D}(I, E))$ of the non-homogeneous martingale problem associated with $\{L_{t,\gamma_t};\ t \in I\}$ and starting at η_0. In this framework the McKean measure $\mathbb{P}$ can again be characterized as the fixed point of the mapping $\eta \in \mathcal{P}(D(I, E)) \to \mathbb{Q}_{\phi(\eta)} \in \mathcal{P}(\mathcal{D}(I, E))$. More generally, $\mathbb{P}$ is defined as the solution of the following time-inhomogeneous martingale problem defined on the canonical space $(\Omega, F, \chi) = (D(I, E),\ (F_t)_{t \geqslant 0}, (X_t)_{t \geqslant 0})$:

- $\mathbb{P} \circ X_0^{-1} = \eta_0$.
- For any $f \in D(L)$, the process

$$M_t(f) = f(X_t) - f(X_0) - \int_0^t L_{s,\eta_s}(f)(X_s)\mathrm{d}s$$

  is an F-martingale under $\mathbb{P}$ and the distribution flow $\{\eta_t;\ t \geqslant 0\}$ coincides with the set of time marginals of $\mathbb{P}$, that is, $\mathbb{P} \circ X_t^{-1} = \eta_t$.

In this case the marginal distribution flow $(\eta_t)_{t \in I}$ solves the nonlinear evolution equation defined for any $f \in D(L)$ by

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta_t(f) = \eta_t(L_{t,\eta_t}(f)).$$

The N-IPS associated with L and $\eta_0 \in \mathcal{P}(E)$ is an E^N-valued Markov process $\xi_t = (\xi_t^1, \ldots, \xi_t^N)$ having the initial distribution $\eta_0^{\otimes N}$ with generator

$$\mathcal{L}_t(f)(x^1, \ldots, x^N) = \sum_{i=1}^N L_{t,m(x)}^{(i)}(f)(x^1, \ldots, x^N),$$

where $L_{t,\mu}^{(i)}$ is used instead of $L_{t,\mu}$ when it acts on the ith variable of $f(x^1, \ldots, x^N)$.

In the discrete or continuous time situation it is also known that the empirical distribution $(1/N)\sum_{i=1}^N \delta_{(\xi_t^i)_{t\in I}}$ on the path space $D(I, E)$ converges as $N \to \infty$ to the McKean measure $\mathbb{P}$ in the sense that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N F((\xi_t^i)_{t\in I}) = \int_\Omega F(\omega)\mathrm{d}\mathbb{P}(\omega)$$

in probability and for any bounded continuous function F on $D(I, E)$.

In the remainder of this section we use the following notation:

$$\Omega = D(I, E), \qquad M = \mathcal{P}(\Omega).$$

Let $\mathbb{P}^N$ be the probability measure induced by the N-IPS process $(\xi_t)_{t\in I}$ in the product path space $\Omega^N = \Omega \times \ldots \times \Omega$. By $\overline{\mathbb{P}}^N = \mathbb{P}^N \circ \pi_N^{-1}$ we denote the image probability measure of the empirical measure on the path space with

$$\pi_N : \omega \in \Omega^N \to \pi_N(\omega) = \frac{1}{N}\sum_{i=1}^N \delta_{\omega^i} \in M.$$

For each $\eta \in M$ we also denote by $\mathbb{Q}_\eta^N = (\mathbb{Q}_{\phi(\eta)})^{\otimes N}$ the N-fold tensor product of the measure $\mathbb{Q}_{\phi(\eta)}$ and by $\overline{\mathbb{Q}}_\eta^N = \mathbb{Q}_\eta^N \circ \pi_N^{-1}$ the corresponding image measure. For each $\eta \in M$, Sanov's theorem tells us that the sequence $\overline{\mathbb{Q}}_{\phi(\eta)}^N$, $N \geqslant 1$, satisfies an LDP with good rate function

$$\eta' \in M \to H_\eta(\eta') = \mathrm{Ent}(\eta'|\mathbb{Q}_{\phi(\eta)}) \in [0, \infty].$$

Here, $\mathrm{Ent}(\eta'|\mathbb{Q}_{\phi(\eta)})$ denotes the relative entropy of η' with respect to $\phi(\eta)$. To transfer this LDP to the sequence $\overline{\mathbb{P}}^N$, $N \geqslant 1$, using Lemma 1.1 we need to find a collection of measures $\mathbb{P}_{\alpha,m}^N$ of the form (4) with $\overline{\mathbb{P}}^N = \overline{\mathbb{P}}_1^N$. In what follows we indicate a simple way to construct these distributions. The construction will be notationally complicated, but it is a straightforward application of Girsanov's theorem.

A great deal of work has been done on LDPs for interacting particle systems on path space. To motivate our work let us briefly connect our strategy with existing results in the literature on the subject. In the case of jump processes, Feng (1994a) proved that the collection of measures $\overline{\mathbb{P}}^N$ satisfies the LDP with a good rate function

$$\mathcal{I} : \eta \in M \to \mathcal{I}(\eta) = \mathrm{Ent}(\eta|\mathbb{Q}_{\phi(\eta)}), \tag{11}$$

while Feng (1994b) and Léonard (1995) obtained the LDP for the empirical process $(m(\xi_t))_{t\in I}$. In the diffusion case, Dawson and Gärtner (1994) also obtained the LDP for the empirical process. In the discrete time situation Del Moral and Guionnet (1998) obtained the

LDP in path space when the Markov transitions $K_{n,\mu}$ are such that $K_{n,\mu}(x, \cdot) = K_{n,\mu}(y, \cdot)$ for any $x, y \in E$. Here we present a straightforward approach to obtaining the LDP in path space for discrete, pure jump and diffusion IPS based on a direct application of Lemma 1.1. In all situations the good rate function is given by (11).

We begin with discrete time IPSs. Suppose that, for each pair of measures $(\mu, \eta) \in \mathcal{P}(E)^2$ and for each $n \in I$, $x \in E$ and $\alpha \in \mathbb{R}$,

$$K_{n,\mu}(x, \cdot) \sim K_{n,\eta}(x, \cdot) \quad \text{and} \quad Z_n^{(\alpha)}(\mu, \eta)(x) = \int \left[\frac{\mathrm{d}K_{n,\mu}(x, \cdot)}{\mathrm{d}K_{n,\eta}(x, \cdot)}(y)\right]^{\alpha} K_{n,\eta}(x, \mathrm{d}y) \in (0, \infty).$$

We also suppose that the mappings

$$\mu \in \mathcal{P}(E) \to \int \mu(\mathrm{d}x) \log Z_n^{(\alpha)}(\mu, \eta)(x) \tag{12}$$

are bounded and continuous at each $\mu = \eta$. Let η be a fixed distribution on the path space $\Omega(= E^{T+1})$ and let $\mathbb{P}^N_{\alpha,\eta}$, $\alpha \in \mathbb{R}$, denote the collection of distributions on $\Omega^N = (E^{T+1})^N$ defined by

$$\frac{\mathrm{d}\mathbb{P}^N_{\alpha,\eta}}{\mathrm{d}\mathbb{Q}^N_{\eta}}(\omega) = \exp(N[\alpha S_N(\omega, \eta) + C_\alpha(\pi_N(\omega), \eta)])$$

for $\mathbb{Q}^N_\eta$-almost every $\omega \in \Omega^N$ with

$$S_N(\omega, \eta) = \sum_{p=1}^{T} \int m(\omega_{p-1}, \omega_p)(\mathrm{d}(u, v)) \log\left[\frac{\mathrm{d}K_{p,m(\omega_{p-1})}(u, \cdot)}{\mathrm{d}K_{p,\eta_{p-1}}(u, \cdot)}(v)\right],$$

$$C_\alpha(\mu, \eta) = -\sum_{p=1}^{T} \int \mu_{p-1}(\mathrm{d}u) \log\left[\int \left(\frac{\mathrm{d}K_{p,\mu_{p-1})}(u, \cdot)}{\mathrm{d}K_{p,\eta_{p-1}}(u, \cdot)}(v)\right)^{\alpha} K_{p,\eta_{p-1}}(u, \mathrm{d}v)\right].$$

Under $\mathbb{P}^N_{\alpha,\eta}$ the N-IPS model $(\xi_n)_{n \in I}$ is the N-IPS model associated with the collection of Markov transitions $K^{(\alpha)} = \{K^{(\alpha)}_{n,\mu};\ n \in I, \mu \in \mathcal{P}(E)\}$ on E defined by

$$K^{(\alpha)}_{n,\mu}(x, \mathrm{d}y) = \frac{1}{Z_n^{\alpha}(\mu, \eta_{n-1})(x)} \left[\frac{dK_{n,\mu}(x, \cdot)}{dK_{n,\eta_{n-1}}(x, \cdot)}(y)\right]^{\alpha} K_{n,\eta_{n-1}}(x, \mathrm{d}y).$$

As mentioned above, the parameter α clearly measures the degree of interaction in the system. For instance, for $\alpha = 0$ we have $\mathbb{P}^N_{0,\eta} = \mathbb{Q}^N_\eta = (\mathbb{Q}_{\phi(\eta)})^{\otimes N}$ and under $\mathbb{P}^N_{0,\eta}$ the N-IPS consists of N independent particles with elementary transitions $K_{n,\eta_{n-1}}$. To illustrate this observation, we examine the case of Gaussian mean field transitions on $E = \mathbb{R}$ defined by

$$K_{n,\mu}(x, \mathrm{d}y) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(y - b[x, \mu])^2\right\} \mathrm{d}y$$

for some measurable drift function $b : E \times \mathcal{P}(E) \to E$. In this special case we can verify that, for any $(x, \mu) \in E \times \mathcal{P}(E)$,

$$K^{(\alpha)}_{n,\mu}(x, \mathrm{d}y) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(y - (b[x, \eta_{n-1}] + \alpha(b[x, \mu] - b[x, \eta_{n-1}])))^2\right] \mathrm{d}y.$$

In addition, we can verify that, for any $\mu, \eta \in M$ $(= \mathcal{P}(E^{T+1}))$,

$$C_\alpha(\mu, \eta) = \frac{\alpha(1-\alpha)}{2} \sum_{p=0}^{n-1} \int \mu_p(\mathrm{d}x)(b[x, \eta_p] - b[x, \mu_p])^2.$$

We note that the technical assumption (12) is satisfied if and only if the drift function b is chosen so that the mappings

$$\mu \in \mathcal{P}(E) \to \int \mu(\mathrm{d}x)(b[x, \eta] - b[x, \mu])^2$$

are bounded and continuous at each $\mu = \eta$. To connect this example with the McKean–Vlasov diffusion model examined later let us suppose the drift function $b[x, \mu]$ takes the form

$$b[x, \mu] = \int \mu(\mathrm{d}y) b(x, y)$$

for some measurable function $b(x, y)$ on the product space E^2. In this situation we can verify that

$$\int \mu(\mathrm{d}x)(b[x, \eta] - b[x, \mu])^2 = \mu \otimes [(\mu \otimes \mu) - 2(\mu \otimes \eta) + (\eta \otimes \eta)](B)$$

with $B(x, y, z) = b(x, y)b(x, z)$. The desired regularity property is therefore met if $b(x, y)$ is bounded continuous.

Turning now to continuous time IPSs, we consider two traditional situations: the pure jump generator and the diffusion generator.

For the pure jump generator we suppose that $L = \{L_\mu; \mu \in \mathcal{P}(E)\}$ is a collection of bounded generators and $D(L) = \mathcal{B}_b(E)$ is the set of bounded measurable functions. If $L^+_\mu(x, \cdot)$ and $L^-_\mu(x, \cdot)$ constitute a Jordan decomposition of the signed kernel $L_\mu(x, \cdot)$ then we have $L^+_\mu(x, E) = L^-_\mu(x, E)$ and

$$L_\mu(x, \mathrm{d}y) = \lambda_\mu(x)[Q_\mu(x, \mathrm{d}y) - \delta_x(\mathrm{d}y)]$$

with

$$\lambda_\mu(x) = L^+_\mu(x, E) \quad \text{and} \quad Q_\mu(x, \mathrm{d}y) = L^+_\mu(x, \mathrm{d}y)/L^+_\mu(x, E).$$

In this situation we assume that for any μ, ν and u we have $L^+_\mu(u, \cdot) \sim L^+_\nu(u, \cdot)$ with $\mathrm{d}L^+_\mu(U, \cdot)/\mathrm{d}L^+_\nu(u, \cdot)(v) \geqslant \epsilon > 0$. We also assume there exists some countable subset $\mathcal{H} \subset \mathcal{C}_b(E)$ with

$$|1 - \mathrm{d}L^+_\nu(u, \cdot)/\mathrm{d}L^+_\mu(u, \cdot)(v)| \leqslant \sum_{h \in \mathcal{H}} |\mu(h) - \nu(h)| \quad \text{and} \quad |\mathcal{H}| = \sum_{h \in \mathcal{H}} \|h\| < \infty.$$

For the diffusion generator we suppose that $E = \mathbb{R}^d$, $d \geqslant 1$. If $b : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ is a bounded continuous function then, for each $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $\varphi \in D(L) = \mathcal{C}^2_b(\mathbb{R}^d)$, we set

$$L_\mu(\varphi)(x) = \frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2\varphi}{\partial x_i^2}(x) + \sum_{i=1}^{d} b_i[x,\, \mu]\frac{\partial\varphi}{\partial x_i}(x)$$

with

$$b_i[x,\, \mu] = \int_{\mathbb{R}^d}\mu(\mathrm{d}y)b_i(x,\, y).$$

In both situations it is known that there exists a unique McKean measure $\mathbb{P}$; see, for instance, Graham (1992) in the jump case and Sznitman (1991) for diffusions. In the pure jump situation, with $\eta \in M$ fixed, Girsanov's theorem implies that $\mathbb{P}^N \sim \mathbb{Q}_\eta^N$. In addition, for $\mathbb{Q}_\eta^N$-almost every $\omega \in \Omega^N$, we have that

$$\frac{\mathrm{d}\mathbb{P}^N}{\mathrm{d}\mathbb{Q}_\eta^N}(\omega) = \exp N[S_N(\omega,\, \eta) + C_1(\pi_N(\omega),\, \eta)],$$

with the mappings $S_N : \Omega^N \times M \to \mathbb{R}$ and $C_1 : M \times M \to \mathbb{R}$ defined by

$$S_N(\omega,\, \eta) = \sum_{s\leqslant T}\int_{E\times E} m(\omega_{s-},\, \omega_s)(\mathrm{d}(u,\, v))\log\frac{\mathrm{d}L^+_{m(\omega_{s-})}(u,\, \cdot)}{\mathrm{d}L^+_{\eta_{s-}}(u,\, \cdot)}(v)1_{u\neq v},$$

$$C_1(\mu,\, \eta) = -\int_0^T \mu_{s-}\left(L^+_{\mu_{s-}} - L^+_{\eta_{s-}}\right)(1)\,\mathrm{d}s.$$

In the same way, for any bounded Borel function $f \in \mathcal{B}_b(I \times \Omega^N \times \Omega^N)$ the process defined for any $\omega \in \Omega^N$ by

$$\mathcal{E}_t(f)(\omega) = \exp\left[\sum_{\substack{s\leqslant t\\ \omega_s\neq\omega_{s-}}} f(s,\, \omega_{s-},\, \omega_s) - \int_0^t\int_{E^N}(\mathrm{e}^{f(s,\omega_{s-},x)} - 1)\mathcal{L}^+_{\phi(\eta)}(\omega_{s-},\, \mathrm{d}x)\mathrm{d}s\right]$$

with, for any $\varphi \in \mathcal{B}_b(\Omega^N)$,

$$\mathcal{L}^+_{\phi(\eta)}(\varphi)(\omega_{s-}) = \sum_{i=1}^{N}\int_E L^+_{\eta_{s-}}(x^i,\, \mathrm{d}u)\varphi(\theta^i(\omega_{s-},\, u)),$$

is a local $\mathbb{Q}_{\phi(\eta)}$-martingale. Here, for $1 \leqslant i \leqslant N$, $u \in E$ and $x = (x_1,\, \ldots,\, x_N) \in E^N$, $\theta^i(x,\, u) = (\theta^i_j(x,\, u))_{1\leqslant j\leqslant N}$ is the element of E^N given by

$$\theta^i_j(x,\, u) = \begin{cases} x_j, & \text{if } j \neq i,\\ u, & \text{if } j = i.\end{cases}$$

Applying this result to the collection of functions

$$f(s,\, x,\, y) = \alpha\sum_{i=1}^{N} 1_{y^i\neq x^i}\log\frac{\mathrm{d}L^+_{m(x)}(x^i,\, \cdot)}{\mathrm{d}L^+_{\eta_{s-}}(x^i,\, \cdot)}(y^i), \qquad \alpha \in \mathbb{R},$$

we define a family of probability measures $\mathbb{P}^N_{\alpha,\eta}$, with $\alpha \in \mathbb{R}$, by setting

$$\frac{d\mathbb{P}^N_{\alpha,\eta}}{d\mathbb{Q}^N_\eta}(\omega) = \exp N[\alpha S_N(\omega, \eta) + C_\alpha(\pi_N(\omega), \eta)],$$

with the mappings $C_\alpha : M \times M \to \mathbb{R}$ defined by

$$C_\alpha(\mu, \eta) = -\int_0^T \int_{E\times E} \mu_{s-}(du) L^+_{\eta_{s-}}(u, dv) \left(\left[\frac{dL^+_{\mu_{S-}}(u, \cdot)}{dL^+_{\eta_{s-}}(u, \cdot)}(v) \right]^\alpha - 1 \right) ds.$$

Note that under $\mathbb{P}^N_{\alpha,\eta}$ the N-IPS model $(\xi_t)_{t\in I}$ becomes the N-IPS model associated with a signed kernel $L^{(\alpha)}_{t,\mu}$ with Jordan decomposition positive part

$$L^{(\alpha)+}_{t,\mu}(x, dy) = \left(\frac{dL^+_\mu(x, \cdot)}{dL^+_{\eta t-}(x, \cdot)}(y) \right)^\alpha L^+_{\eta_t -}(x, dy).$$

In view of Lemma 1.1 above and Sanov's theorem (Dembo and Zeitouni 1993, Theorem 6.2.10), it remains to prove that, for each $m \in M$ and $\alpha \in \mathbb{R}$, the mapping $C_\alpha(\cdot, m)$ is bounded and continuous on M. First, we note that, for any $\epsilon \leqslant x \leqslant 1/\epsilon$ and $\alpha \in \mathbb{R}$,

$$|1 - x^\alpha| = |1 - \exp(\alpha \log x)| \leqslant |\alpha|\ |\log x|\ |1 + x^\alpha|$$

and therefore

$$|1 - x^\alpha| \leqslant \theta(\alpha, \epsilon)|1 - x| \qquad \text{with } \theta(\alpha, \epsilon) = |\alpha|(\epsilon^{-1} + \epsilon^{-(1+|\alpha|)})(\leqslant 2|\alpha|/\epsilon^{1+|\alpha|}).$$

Hence, under our assumptions we obtain

$$\begin{aligned} |C_\alpha(\mu, \eta)| &\leqslant \theta(\alpha, \epsilon) \int_0^T \int_{E\times E} \mu_{s-}(du) L^+_{\eta_{s-}}(u, dv) \left| \frac{dL^+_{\mu_{s-}}(u, \cdot)}{dL^+_{\eta_{s-}}(u, \cdot)}(v) - 1 \right| ds \\ &\leqslant \theta(\alpha, \epsilon) \sum_{h\in\mathcal{H}} \int_0^T [\mu_{s-} L^+_{\eta_{s-}}(1)] |\mu_{s-}(h) - \eta_{s-}(h)| ds, \end{aligned}$$

and finally

$$|C_\alpha(\mu, \eta)| \leqslant \theta_\eta(\alpha, \epsilon) \sum_{h\in\mathcal{H}} \int_0^T |\mu_t(h) - \eta_t(h)| ds,$$

with $\theta_\eta(\alpha, \epsilon) = \theta(\alpha, \epsilon)\ [\sup_{0\leqslant t\leqslant T} \|\lambda_{\eta t}\|] < \infty$. On the other hand, noting that

$$\int_0^T |\mu_t(h) - \eta_t(h)|^2\, dt = \int_0^T \mu_t(h)^2\, dt + \int_0^T \eta_t(h)^2\, dt - 2\int_0^T \mu_t(h)\eta_t(h) dt$$

and, for any $h \in \mathcal{C}_b(E)$ and $\mu, \eta \in M$,

$$\int_0^T \mu_t(h)\eta_t(h) dt = \int_0^T \int_{\Omega_T\times\Omega_T} h(\omega_t) h(\omega'_t) \mu(d\omega) \eta(d\omega') dt = (\mu \otimes \eta)(H_h)$$

with

$$H_h(\omega, \omega') = \int_0^T (h \otimes h)(\omega_t, \omega'_t)\mathrm{d}t = \int_0^T h(\omega_t)h(\omega'_t)\mathrm{d}t,$$

we have

$$C_\alpha^2(\mu, \eta) \leqslant [\theta_\eta^2(\alpha, \epsilon)|\mathcal{H}|] \sum_{h \in \mathcal{H}} ((\mu \otimes \mu) + (\eta \otimes \eta) - 2(\mu \otimes \eta))(H_h).$$

The space $\Omega = D([0, T], E)$ with the Skorohod metric is topologically complete and separable. Thus, by Lemma 1.1 in Parthasarathy (1967), if μ^n and η^n weakly converge to μ and η then $(\mu^n \otimes \eta^n)$ weakly converges to $(\mu \otimes \eta)$. For each continuous mapping h the integral function

$$\omega \in \Omega \to \int_0^T h(\omega_t)\mathrm{d}t$$

is continuous for the Skorohod metric; see, for instance, Example 8.2 in Kurtz and Protter (1995, p. 32). Since $\mathcal{H} \subset \mathcal{C}_b(E)$, we conclude that $H_h \in \mathcal{C}_b(\Omega)$ for any $h \in \mathcal{H}$ and

$$\mu^n \text{ weakly converges to } \mu \Rightarrow \lim_{n \to \infty} C_\alpha(\mu^n, \eta) = C_\alpha(\mu, \eta).$$

We now examine mean field diffusions. We first notice that, for any $\eta \in M$, $\omega \in \Omega$ and b bounded continous, the stochastic process $(B_t^\eta)_{t \in I}$ defined by

$$t \in I \to B_t^\eta(\omega) = \omega_t - \omega_0 - \int_0^t b[\omega_t, \eta_t]\mathrm{d}s$$

is an E-valued Brownian motion with respect to $\mathbb{Q}_{\phi(\eta)}$. By Girsanov's theorem we have $\mathbb{P}^N \sim \mathbb{Q}_\eta^N$ and, for $\mathbb{Q}_\eta^N$-almost every $\omega = (\omega^1, \ldots, \omega^N) \in \Omega^N$,

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{P}^N}{\mathrm{d}\mathbb{Q}_\eta^N}(\omega) &= \sum_{i=1}^N \int_0^T (b[\omega_t^i, m(\omega_t)] - b[\omega_t^i, \eta_t])^{\mathrm{T}}\, \mathrm{d}B_t^\eta(\omega_i) \\
&\quad - \frac{1}{2}\sum_{i=1}^N \int_0^T \|b[\omega_t^i, m(\omega_t)] - b[\omega_t^i, \eta_t]\|^2\, \mathrm{d}s \\
&= N[S_N(\omega, \eta) + C_1(\pi_N(\omega), \eta)]
\end{aligned}$$

with, for any $\omega \in \Omega^N$ and $\mu, \eta \in M$,

$$S_N(\omega, \eta) = \int_0^T \int m(\omega_t)(\mathrm{d}u)m(\omega_t)(\mathrm{d}v)(b(u, v) - b[u, \eta_t])^{\mathrm{T}}\, \mathrm{d}B_t^\eta(u)$$

$$C_1(\mu, \eta) = -\frac{1}{2}\int_0^T \int \mu_t(du)\|b[u, \mu_t] - b[u, \eta_t]\|^2\, \mathrm{d}s.$$

Here, the superscript T denotes transpose. Arguing as in the pure jump situation, we define a family of probability measures $\mathbb{P}_{\alpha,\eta}^N$, with $\alpha \in \mathbb{R}$, by setting

$$\frac{d\mathbb{P}^N_{\alpha,\eta}}{d\mathbb{Q}^N_\eta}(\omega) = \exp N[\alpha S_N(\omega, \eta) + C_\alpha(\pi_N(\omega), \eta)], \qquad \text{with } C_\alpha = \alpha^2 C_1.$$

Note that under $\mathbb{P}^N_{\alpha,\eta}$ the N-IPS model $(\xi_t)_{t\in I}$ is the N-IPS model associated with the time-inhomogeneous generators $L^{(\alpha)}_{t,\mu}$ defined, for any $t \in I$ and $\varphi \in D(L) = \mathcal{C}^2_b(\mathbb{R}^d)$, by

$$L^{(\alpha)}_{t,\mu}(\varphi)(x) = L_{\eta t}(\varphi)(x) + \alpha \sum_{i=1}^d (b_i[x, \mu] - b_i[x, \eta_t]) \frac{\partial \varphi}{\partial x_i}(x).$$

For any pair $(\mu, \eta) \in \mathcal{P}(E)^2$, we have that

$$\int \mu(du)\|b[u, \mu] - b[u, \eta]\|^2 = \sum_{i=1}^d \int \mu(du)[\mu(b_i(u, \cdot)) - \eta(b_i(u, \cdot))]^2. \tag{13}$$

Noting that

$$[\mu(b_i(u, \cdot)) - \eta(b_i(u, \cdot))]^2 = \mu(b_i(u, \cdot))^2 + \eta(b_i(u, \cdot))^2 - 2\mu(b_i(u, \cdot))\eta(b_i(u, \cdot))$$

and

$$\begin{aligned}\mu(b_i(u, \cdot))\eta(b_i(u, \cdot)) &= \int (\mu \otimes \eta)(dv, dv')b_i(u, v)b_i(u, v') \\ &= (\mu \otimes \eta)(b_i(u, \cdot) \otimes b_i(u, \cdot)),\end{aligned}$$

we can rewrite (13) as

$$\begin{aligned}&\int \mu(du)\|b[u, \mu] - b[u, \eta]\|^2 \\ &\quad = \sum_{i=1}^d \mu(du)[(\mu \otimes \mu) + (\eta \otimes \eta) - 2(\mu \otimes \eta)](b_i(u, \cdot) \otimes b_i(u, \cdot)) \\ &\quad = \mu \otimes [(\mu \otimes \mu) + (\eta \otimes \eta) - 2(\mu \otimes \eta)](B),\end{aligned}$$

with $B \in \mathcal{C}_b(E^3)$ given by

$$B(u, v, v') = \sum_{i=1}^d b_i(u, v)b_i(u, v').$$

This implies that, for any pair $(\mu, \eta) \in M^2$,

$$\begin{aligned}C_\alpha(\mu, \eta) &= -\frac{\alpha^2}{2}\int_0^T \mu_t \otimes [(\mu_t \otimes \mu_t) + (\eta_t \otimes \eta_t) - 2(\mu_t \otimes \eta_t)](B)ds \\ &= -\frac{\alpha^2}{2}\mu \otimes [(\mu \otimes \mu) + (\eta \otimes \eta) - 2(\mu \otimes \eta)](\mathcal{B}),\end{aligned}$$

with $\mathcal{B} \in \mathcal{C}_b(\Omega^3)$ given by

$$\mathcal{B}(\omega, \omega', \omega'') = \int_0^T B(\omega_t, \omega'_t, \omega''_t)\mathrm{d}t.$$

Arguing as in the jump case, the desired properties of the mapping $C_a(\cdot,\ m)$ are clear.

References

Dawson, D. and Gärtner, J. (1994) Multilevel large deviations and interacting diffusions. *Probab. Theory Related Fields*, **98**, 423–487.

Del Moral, P. and Guionnet, A. (1998) Large deviations for interacting particle systems: Applications to non-linear filtering problems. *Stochastic Process. Appl.*, **78**, 69–95.

Dembo, A. and Zeitouni, O. (1993) *Large Deviations Techniques and Applications.* Boston: Jones and Bartlett.

Feng, J. and Kurtz, T. (2000) Large deviation for stochastic processes. *http://www.math.umass.edu/~feng/Research.html.*

Feng, S. (1994a) Large deviations for Markov processes with mean field interaction and unbounded jumps. *Probab. Theory Related Fields*, **100**, 227–252.

Feng, S. (1994b) Large deviations for empirical processes of mean-field interacting particle system with unbounded jumps. *Ann. Probab.*, **22**, 2122–2151.

Graham, C. (1992) Nonlinear diffusion with jumps. *Ann. Inst. H. Poincaré Probab. Statist.*, **28**, 393–402.

Kurtz, T. and Protter, P.E. (1995) Weak convergence of stochastic integrals. In D. Talay and L. Tubaro (eds), *Probabilistic Models for Nonlinear Partial Differential Equations*, Lecture Notes in Math. 1627. Berlin: Springer-Verlag.

Léonard, C. (1995) On large deviations for particle systems associated with spatially homogeneous Boltzmann type equations. *Probab. Theory Related Fields*, **101**, 1–44.

Parthasarathy, K.R. (1967) *Probability Measures on Metric Spaces*. New York: Academic Press.

Puhalskii, A. (1991) On functional principle of large deviations. In V. Sazonov and T. Shervashidze (eds), *New Trends in Probability and Statistics*, *Vol. 1*, pp. 198–219. Utrecht: VSP.

Sznitman, A.S. (1991) Topics in propagation of chaos. In P.L. Hennequin (ed.) *École d'Été de Probabilités de Saint-Flour XIX – 1989*, Lecture Notes in Math. 1464. Berlin: Springer-Verlag.

Received September 2001 and revised June 2002