Stochastic Processes and their Applications 44 (1993) 329–345
North-Holland

Finite Kullback information diffusion laws with fixed marginals and associated large deviations functionals

Marc Brunaud

Equipe de Statistique Appliquée, Université Paris-Sud, Orsay, France

Received 16 May 1990
Revised 13 February 1991

In this article we extend a theorem previously proved by H. Föllmer for the Wiener process on $C([0,1],\mathbb{R}^d)$ to diffusion processes; we therefore straightforward recover, under slightly less general technical assumptions but in its whole generality, a theorem already given by Dawson and Gärtner. The result is intimately related with a Ventcel–Freidlin action functional associated to some N-particle system which is driven according to a non-linear McKean–Vlasov limiting equation.

AMS 1980 Subject Classifications: 60F10, 60J60, 35Q20.

large deviations * diffusion processes * Kullback information * particular equation of mathematical physics

1. Introduction and result

In Dawson and Gärtner (1987) the large deviations from the McKean–Vlasov limit for N particle systems are investigated; the authors underline a strong analogy with the now classical Ventcel–Freidlin theory. Moreover, they derive the following expression of the associated action functional denoted by $S(\mu(\cdot))$:

$$S(\mu(\cdot)) = \int_0^T \|\dot{\mu}(t) - L(\mu(t))^*\mu(t)\|^2_{\mu(t)}\,\mathrm{d}t.$$

Formally speaking, the space of probability measures on $\mathbb{R}^d$, denoted by $\mathcal{M}$ is given a Riemannian manifold structure, for which each tangent space $\mathcal{T}_\mu$ at $\mu \in \mathcal{M}$ is embedded in $\mathcal{D}'$ (the Schwartz distribution space on $\mathbb{R}^d$) and equipped with the norm

$$\|\theta\|^2_\mu = \sup\{(\theta, f) - \tfrac{1}{2}\langle\mu, |\nabla f|^2\rangle \,|\, f \in \mathcal{D}\}.$$

Correspondence to: Dr. M. Brunaud, U.F.R. de Mathématiques, Université Paris VII, 2 Place Jussieu, 75251 Paris Cedex 05, France.

$\mathcal{T}_\mu$ thus coincides with the $\theta \in \mathcal{D}'$ such that the latter sup is finite (here $|\cdot|^2$ denotes the squared Riemannian norm induced by the diffusion matrix $a(\cdot)$). $L(\mu)$ is a diffusion operator of the type $\frac{1}{2}\sum a^{i,j}(\cdot)\partial^2_{i,j}+\sum b^i(\cdot\,;\mu)\partial_i$, of which the drift depends on $\mu \in \mathcal{M}$: this dependance is to feature weak interactions between diffusions in the mean-field model, where μ equals in this case the empirical mean $(1/N)\sum_{1\le j\le N}\delta_{X_j}$ of N particles. Let us finally remark that $\|\cdot\|^2_\mu$ is the Legendre transform of the mapping $f \to \frac{1}{2}\langle \mu, |\nabla f|^2\rangle$ for the dual pairing $(\mathcal{D}, \mathcal{D}')$ and generalizes the well-known formulae when the underlying driving process is, say, Brownian motion or ordinary diffusion processes. On the other hand, the quantity appearing in the integrand of the action functional under squared norm value is related to the mean-field limit dynamics: the so called McKean–Vlasov equation

$$\dot{\mu}(t) = L(\mu(t))^* \mu(t)$$

(to be understood in the weak sense) defines, when existing, some process which is to be the large number limit of the empirical mean $(1/N)\sum_{1\le j\le N}\delta_{X_j(\cdot)}$ of our N particles.

The main difficulty is to derive directly the expression of the action functional. In fact, as H. Föllmer pointed it out and showed in his 'Cours de Saint-Flour', when the diffusion matrix equals everywhere the identity, there exists one single law Q_0 on the Wiener space with prescribed marginals $(\nu_t)_{t\in[0,1]}$ minimizing the Kullback information with respect to the Wiener measure P_0 and for which the minimum is computed by

$$\mathcal{I}(Q_0; P_0) = \mathcal{I}(q_0; p_0) + \int_0^1 \|\dot{\nu}_u - \tfrac{1}{2}\Delta \nu_u\|^2_{\nu_u}\, du.$$

Moreover, Q_0 is Markovian. The proof of Dawson and Gärtner's result contains a contraction argument (Lemma 4.6 of Dawson and Gärtner, 1987) which uses a large deviation principle for the diffusion laws on the path space; their proof consists in giving three different expressions of the large deviations of the flow of one-dimensional marginals and comparing them to each other. In Ben Arous and Brunaud (1990), we met an analogous problem: we were interested in the large deviations for the empirical mean of N particle system driven according to the nonlinear McKean–Vlasov limiting equation; we also looked for characterizing the minima of the large deviations rate function. But there, the crucial point was that the rate function was the sum of the Kullback information w.r.t. some fixed process and an integral functional of one-dimensional marginals. We thus had to look at laws which minimize the Kullback information under prescribed one-dimensional marginals.

We introduce now the principle notations and our main theorem; in Section 2, we give its proof using the notion of h-path processes as indicated in Föllmer (1986).

Let $a:\mathbb{R}^d \times (0,+\infty) \to \mathbb{R}^d \otimes \mathbb{R}^d$ be a diffusion matrix, which defines a pseudo Riemannian metrics on $\mathbb{R}^d$; $\nabla_t, (\cdot,\cdot)_t, |\cdot|_t$ denote respectively the Riemannian gradient, scalar product and norm on the tangent space. In global Euclidean coordinate $(x^1, \dots, x^d)$,

$$(X, Y)_t = \sum_{1\le i,j\le d} a_{i,j}(\cdot, t) X^i Y^j, \tag{1}$$

$$(\nabla_t f)^i = \sum_{j=1}^d a^{i,j}(\cdot, t)\frac{\partial f}{\partial x_j}, \tag{2}$$

$$|X|_t^2 = (X, X)_t, \tag{3}$$

$$\langle X, Y\rangle_t = \sum_{1\le i,j\le d} a^{i,j}(\cdot, t) X^i Y^j, \tag{4}$$

$$\langle\!\langle X\rangle\!\rangle_t^2 = \langle X, X\rangle_t, \tag{5}$$

where the matrix $(a_{i,j}(\cdot, t))_{i,j}$ denotes the inverse of $(a^{i,j}(\cdot, t))_{i,j} \equiv a$. Let us observe that formulae (2) and (4) both imply

$$|\nabla_t f|_t^2 = \sum_{1\le i,j\le d} a^{i,j}(\cdot, t)\frac{\partial f}{\partial x_i}\cdot\frac{\partial f}{\partial x_j}. \tag{6}$$

Let $b:\mathbb{R}^d\times[0,+\infty)\to\mathbb{R}^d$ be a measurable vector field (drift). We shall make the following:

Assumption A1. For any $T>0$, there exist $0<\lambda_T<\Lambda_T<\infty$, $B_T<\infty$ such that, for all $(t, x)\in[0, T]\times\mathbb{R}^d$ and $\theta\in\mathbb{R}^d$, are satisfied:

$$\lambda_T|\theta|^2 \le \langle\!\langle\theta\rangle\!\rangle_t^2(x) \le \Lambda_T|\theta|^2,$$

$$|b(x, t)| \le B_T,$$

$$\limsup_{\delta\to 0}\ \sup_{|x^1-x^2|\le\delta}\ \sup_{0\le s\le T} |a(x^1, s) - a(x^2, s)|_{\mathrm{HS}} = 0.$$

For some fixed $t>0$ and $\xi\in\mathbb{R}^d$, let $L_t(\xi)$ be a second order differential operator, defined for every $f\in C_K^2(\mathbb{R}^d)$ (the set of compactly supported twice differentiable functions on $\mathbb{R}^d$) by

$$L_t(\xi)(f) = \frac{1}{2}\sum_{i,j=1}^d a^{i,j}(\xi, t)\frac{\partial^2 f}{\partial x_i\partial x_j} + \sum_{j=1}^d b^j(\xi, t)\frac{\partial f}{\partial x_j}. \tag{7}$$

We shall omit ξ in $L_t(\xi)$, whenever it is applied on a progressively measurable process $X(t, \xi)$:

$$L_t(\xi)(f)(X(t, \xi)) = L_t f(X_t(\xi)).$$

Moreover, L_t^* will denote the formal adjoint of L_t acting on probability distributions on $\mathbb{R}^d$. Under Assumption A1, Theorem 9.1.9 of Stroock and Varadhan (1979) yields the existence of a transition probability density $\pi(s, x; t, y)$ for the solution P of the martingale problem associated with L_t. Furthermore, we shall impose:

Assumption A2. For any nonnegative measurable function ϕ on $\mathbb{R}^d$ such that

$$\Pi_t\phi(u, x) = \int \phi(y)\pi(u, x; t, y)\,\mathrm{d}y < \infty, \quad (u, x)\ \mathrm{d}t\otimes m_d \text{ a.s.},$$

$\Pi_t\phi$ is a.s. equal to a continuous mapping from $[0,\infty)\times\mathbb{R}^d$ to $[0,\infty)$, C^∞ in x for a.e. t, such that, for all $t>0$ the continuous modification of $\Pi_t\phi$ is $C_b^{1,2}([0,t]\times\mathbb{R}^d)$ and satisfies

$$\left(\frac{\partial}{\partial s}+L_s\right)\Pi_t\phi=0,\quad 0\leq s<t.$$

Let us now introduce the norm $\|\cdot\|_{\mu,t}$ for some fixed probability measure μ on $\mathbb{R}^d$: if $\mathcal{D}$ is the Schwartz test function space and $\mathcal{D}'$ the Schwartz distribution space on $\mathbb{R}^d$, $\mathcal{T}_{\mu,t}=\{\theta\in\mathcal{D}'\mid\|\theta\|^2_{\mu,t}<\infty\}$, where we defined

$$\|\theta\|^2_{\mu,t}=\sup\{(\theta,f)-\tfrac{1}{2}\langle\mu,|\nabla_t f|_t^2\rangle\mid f\in\mathcal{D}\}.$$

One can easily show that, if $\mathcal{D}_{\mu,t}=\{f\in\mathcal{D}\mid\langle\mu,|\nabla_t f|_t^2\rangle\neq 0\}$,

$$\|\theta\|^2_{\mu,t}=\tfrac{1}{2}\sup\left\{\frac{|(\theta,f)|^2}{\langle\mu,|\nabla_t f|_t^2\rangle}\,\middle|\,f\in\mathcal{D}_{\mu,t}\right\}.$$

Finally, let us recall the definition of Kullback information of a probability measure with respect to another one:

Definition 1.1. Let (X,Σ) be some measurable space and μ, λ be two probability distributions on (X,Σ). The Kullback information of μ with respect to λ, denoted by $\mathcal{I}(\mu;\lambda)$, is given by

$$\mathcal{I}(\mu:\lambda)=\sup\{\langle F,\mu\rangle-\log\langle\exp(F),\lambda\rangle\mid F\in \mathrm{b}L^0(X,\Sigma)\}$$

where $\mathrm{b}L^0(X,\Sigma)$ denotes the space of all bounded, real valued Σ-measurable functions on X. Moreover, when X is Polish and Σ its Borel σ field, we may restrict the sup over the set of all bounded continuous function on X. If we take a sub-σ-field $\mathcal{F}$ of Σ instead of Σ, we shall denote the corresponding Kullback information $\mathcal{I}_{\mathcal{F}}(\mu;\lambda)$.

We then have the following:

Lemma 1.1. *With the preceding notations, the following holds*:

$$\mathcal{I}(\mu;\lambda)=\begin{cases}\displaystyle\int\log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}\right)\mathrm{d}\mu & \text{if } \mu\ll\lambda \text{ and } \dfrac{\mathrm{d}\mu}{\mathrm{d}\lambda}\in L^1(\mu),\\ +\infty & \text{otherwise.}\end{cases}$$

Remark. From now on, we shall always refer to the canonical space of the trajectories $\Omega_s=C([s,\infty);\mathbb{R}^d)$ from time s endowed with the canonical filtration $(\mathcal{F}_{s,t})_{t\in[s,T]}=(\sigma\{x_u\mid s\leq u\leq t\})_{t\in[s,\infty)}$, $(x_u)_{u\in[s,\infty)}$ being the coordinate projections.

Notation. Throughout this paper, $\boldsymbol{W}^d_{s,t}$ will stand for the set of continuous trajectories from $[s, t]$ to $\mathbb{R}^d$.

Our main result consists in the following:

Theorem 1.1. *For fixed $s \in [0, \infty)$, let P_s be the law of an Itô process with variance a and drift b, with some fixed initial law at time s and satisfying both Assumptions* A1 *and* A2. *Some $T > s$ being fixed, denote P the restriction of P_s on $\boldsymbol{W}^d_{s,T}$, endowed with its canonical filtration generated by the coordinate projections $(x_u)_{u\in[s,T]}$. Let $\boldsymbol{q} = (q_u)_{u\in[s,T]} \in M_1^+(\mathbb{R}^d)^{[s,T]}$ be a flow of marginals such that $q_u \sim p_u$ for all $u \in [s, T]$, where $(p_u)_{u\in[s,T]}$ is the corresponding flow for P. Let $K(\boldsymbol{q})$ be the set of all the probability distributions Q on $\boldsymbol{W}^d_{s,T}$ verifying both conditions*

$$\mathscr{I}(Q; P) < \infty,$$

$$Q \circ x_u^{-1} = q_u \quad \textit{for all } u \in [s, T].$$

We shall assume that $K(\boldsymbol{q})$ is non-empty. Then there exists a single law $\bar{Q}$ in $K(\boldsymbol{q})$, minimizing $\mathscr{I}(\cdot\,; P)$. Moreover, $\bar{Q}$ is Markovian and satisfies the following 'Pythagoras' equality:

$$\mathscr{I}(Q; P) = \mathscr{I}(\bar{Q}; P) + \mathscr{I}(Q; \bar{Q}) \quad \textit{for any } Q \in K(\boldsymbol{q})$$

where the minimum is computed by

$$\mathscr{I}(\bar{Q}; P) = \mathscr{I}(q_s; p_s) + \int_s^T \|\dot{q}_u - L_u^* q_u\|^2_{q_u,u}\, \mathrm{d}u.$$

2. Proof of the theorem

2.1. First step: Existence of an I-projection

Let us first translate the conditions satisfied by any element of $K(\boldsymbol{q})$. If $\{s_j\}_{j\geqslant 1}$ is a dense denumerable subset of $[s, T]$ and $\{\phi_i\}_{i\geqslant 1}$ a total subset of $C_b(\mathbb{R}^d)$, the second condition defining the set $K(\boldsymbol{q})$ is equivalent to

$$\forall (i, j) \in \mathbb{N}^* \times \mathbb{N}^*, \quad \langle \phi_i \circ x_{s_j}, Q\rangle = \langle \phi_i, q_{s_j}\rangle$$

i.e., as the marginals of Q are already prescribed by $\boldsymbol{q}$, to denumerable linear constraints on Q. We shall reindex these constraints and put, if k is related to (i, j), $c_k = \langle \phi_i, q_{s_j}\rangle$, $\psi_k = \phi_i \circ x_{s_j}$, so that

$$\forall k \in \mathbb{N}^*, \quad \langle \psi_k, Q\rangle = c_k. \tag{8}$$

It easily follows from these equalities that $K(\boldsymbol{q})$ is a norm variation closed subset of all the probabilities on $\boldsymbol{W}^d_{s,T}$. From Theorem 2.1 of Csiszar (1975), we directly obtain the existence and unicity of $\bar{Q}$. Let us see now how to recover the solution of our minimization problem according to a discretized procedure. We consider the problem consisting in finding a probability distribution Q which minimizes $\mathscr{I}(\cdot; P)$ under the n first preceding constraints. As a consequence of Theorem (3.3) of Csiszar (1975) and the non-emptiness of $K(\boldsymbol{q})$, this problem admits a unique solution Q_n such that, for any $u \in [s, T]$, there exists a unique vector $\theta^{(n,u)}$ in $\mathbb{R}^n$ solving the normal equation in θ:

$$\nabla_\theta \left[\log \left\{ \int \exp\left(\sum_{j=1}^{n} \theta_j \psi_j \right) \mathrm{d}P \right\} \right] = (c_j)_{1 \leqslant j \leqslant n}, \tag{9}$$

and

$$\left. \frac{\mathrm{d}Q_n}{\mathrm{d}P} \right|_{\mathscr{F}_u} = \exp\left(\sum_{j=1}^{n} \theta_j^{(n,u)} \psi_j \right). \tag{10}$$

We prove that the sequence $(Q_n)_{n \geqslant 1}$ converges in variation norm and weakly to a solution Q of the initial minimization problem (see Föllmer, 1986). $K_n(\boldsymbol{q})$ will denote similarly the set of the Q's such that $\mathscr{I}(Q, P) < \infty$ and Q satisfies the n first linear constraints.

A property of Q_n implies that, for any probability distribution Q in $K_n(\boldsymbol{q})$, the following is verified:

$$\mathscr{I}(Q; P) = \mathscr{I}(Q; Q_n) + \mathscr{I}(Q_n; P). \tag{11}$$

But, if $n \geqslant m$, $Q_n \in K_m(\boldsymbol{q})$, so that

$$\forall m,\ \forall n \geqslant m, \quad \mathscr{I}(Q_n; P) = \mathscr{I}(Q_n; Q_m) + \mathscr{I}(Q_m; P). \tag{12}$$

As $K(\boldsymbol{q}) \subset \bigcap_{n \geqslant 1} K_n(\boldsymbol{q})$ and $K(\boldsymbol{q})$ is non-empty, the preceding relation implies that the sequence $(\mathscr{I}(Q_n : P))_{n \geqslant 1}$ is bounded by $\mathscr{I}(R; P)$, where R is an arbitrary element of $K(\boldsymbol{q})$; moreover, the same sequence is nondecreasing and thus convergent. Let l be its limit. As

$$l - \mathscr{I}(Q_m; P) = \sup_{n \geqslant m} [\mathscr{I}(Q_n; P) - \mathscr{I}(Q_m; P)] = \sup_{n \geqslant m} \mathscr{I}(Q_n; Q_m)$$

we get

$$\lim_{m \to \infty} \sup_{n \geqslant m} \mathscr{I}(Q_n; Q_m) = 0.$$

Using the elementary inequality (where $\|\mu\|_{\mathrm{var}}$ denotes the total variation norm of the measure μ) (see Jacod and Shyriaev, 1988)

$$\|\mu - \nu\|_{\mathrm{var}}^2 \leqslant 2\mathscr{I}(\mu; \nu),$$

we deduce from the above relations that the sequence $(Q_n)_{n \geqslant 1}$ is Cauchy and thus converges in variation norm: its weak limit $\bar{Q}$ therefore exists. By means of the lower semi-continuity of $\mathscr{I}(\cdot; P)$, of the fact that $\bigcap_{n \geqslant 1} K_n(\boldsymbol{q})$ contains $\bar{Q}$, due to the continuity of the mappings $Q \mapsto \langle \psi_k, Q \rangle$, we successively get the inequalities

$$\lim_{n \to \infty} \mathscr{I}(Q_n; P) \geqslant \mathscr{I}(\bar{Q}; P)$$

and

$$\mathcal{I}(Q_n;\, P) \leq \mathcal{I}(\bar{Q};\, P)$$

so that $\mathcal{I}(\bar{Q};\, P) = \lim_{n\to\infty} \mathcal{I}(Q_n;\, P)$ and $\lim_{n\to\infty} \mathcal{I}(\bar{Q};\, Q_n) = 0$. This shows that $\bar{Q}$ is a solution of the initial minimization problem; its unicity follows from the unicity of the solution of each minimization problem and the fact that $K(\boldsymbol{q}) = \bigcap_{n\geq 1} K_n(\boldsymbol{q})$. Nevertheless, the unicity of $\bar{Q}$ may be obtained from direct topological considerations: as we are interested in a minimization problem, we may assume that there exists some constant $C > 0$ such that the minimum is taken on $K(\boldsymbol{q}) \cap \{\mathcal{I}(\cdot\,;\, P) \leq C\}$, which is a compact convex non-void subset of $M_1^+(\boldsymbol{W}^d_{s,t})$; $\mathcal{I}(\cdot\,;\, P)$ is a 'good' rate function (i.e. its level sets are compact), lower semi-continuous and strictly convex on $K(\boldsymbol{q})$; so, $\mathcal{I}(\cdot\,;\, P)$ attains its minimum at a unique point of $K(\boldsymbol{q})$. Finally, taking the limit $n \to \infty$ in (11), we get the first assertion of the theorem, which shows that $\bar{Q}$ is the I-projection of P on $K(\boldsymbol{q})$ in the Csiszar's terminology.

2.2. Second step: Description of $\bar{Q}$

As has been shown in the preceding paragraph, the probability distribution $\bar{Q}$, solution of the minimization problem of the theorem exists and is unique. For two main reasons, we need giving a more descriptive characterization of $\bar{Q}$. On one hand, the preceding one, though simple, is not explicit, as given by a limit of unknown probability distributions. On the other hand, in view of our main result, $\bar{Q}$ should be *Markovian*; this property does not follow the Csiszar's like arguments. This crucial feature is related to the smoothing procedure, which underlies an entropy minimizing technique. Let us define $Q^{(n)}$ to be the unique probability distribution minimizing $\mathcal{I}(\cdot\,;\, P)$ under the n following constraints:

$$\forall j,\ 1 \leq j \leq n, \quad Q \circ x_{s_j}^{-1} = q_{s_j}. \tag{13}$$

One can easily show that the same arguments as above work to prove both its existence and unicity. We first give a precise description of $Q^{(2)}$ as an h-conditioned diffusion law for some precise P-invariant mapping h.

(a) *The case $n = 2$*: Let us fix $t \in [s, T]$ and look for laws Q minimizing $\mathcal{I}_{\mathcal{F}_{s,T}}(\cdot\,;\, P)$ such that

$$Q \circ x_s^{-1} = q_s, \qquad Q \circ x_t^{-1} = q_t,$$

where q_s and q_t are two given probability distributions on $\mathbb{R}^d$. For any (x, y) in $\mathbb{R}^d \times \mathbb{R}^d$ and any law R on $\boldsymbol{W}^d_{s,\infty}$ (also denoted by $R_{s,x}$, when emphasizing on the initial conditions) $R^{t,y}_{s,x}$ will denote the regular conditional probability distribution of R, conditioned to be a.s. equal to x (resp. y) at time s (resp. t).

We also need the following definition:

Definition 2.1. Let ω and ω' be two paths in $\boldsymbol{W}^d_{0,\infty}$. Let us pick some $t > 0$. The t-splice of ω and ω', denoted by $\omega \otimes_t \omega'$ is the path defined by

$$(\omega \otimes_t \omega')(s) = \begin{cases} \omega(s)\mathbf{1}_{[0,t]}(s) + \omega'(s)\mathbf{1}_{[t,\infty]}(s) & \text{if } \omega(t) = \omega'(t), \\ \omega(s \wedge t) & \text{otherwise.} \end{cases}$$

Moreover, if P and Q are two probability distributions on $\boldsymbol{W}^d_{0,\infty}$ and $\omega \in \boldsymbol{W}^d_{0,\infty}$, $\omega \otimes_t Q$ denotes the image of Q under the measurable mapping $\omega' \mapsto \omega \otimes_t \omega'$ and $P \otimes_t Q$ the probability distribution $P \otimes_t Q = \int P(\mathrm{d}\omega)\delta_{\omega \otimes_t Q}$.

We now quote the following useful lemma from Varadhan (1984, Lemma 10.3):

Lemma 2.1. *Let (X, Σ) be a Polish space and $\mathscr{F}_1 \subset \mathscr{F}_2 \subset \Sigma$ be sub-σ-fields. Denote also by λ_ω, μ_ω the regular conditional distribution of λ and μ given $\mathscr{F}_1$. Then*

$$\mathscr{I}_{\mathscr{F}_2}(\mu; \lambda) = \mathscr{I}_{\mathscr{F}_1}(\mu; \lambda) + E^{\mu}(\mathscr{I}_{\mathscr{F}_2}(\mu_\omega; \lambda_\omega)). \qquad \square$$

We use the preceding notions to get the following:

Proposition 2.1. *Let Q be any probability distribution on $\boldsymbol{W}^d_{s,T}$ and pick some $t > 0$ in (s, T). $Q^{s,t}_\omega$ (resp. $Q^{t,y}_{s,x}$) will denote a regular conditional probability distribution of Q given $\mathscr{F}_{s,t}$ (resp. $\sigma\{x_s, x_t\}$). It then holds that:*

$$\mathscr{I}_{\mathscr{F}_{s,T}}(Q; P) = \mathscr{I}(\nu_Q; \nu_P) + \int \nu_Q(\mathrm{d}x \otimes \mathrm{d}y)\mathscr{I}_{\mathscr{F}_{s,t}}(Q^{t,y}_{s,x}; P^{t,y}_{s,x})$$
$$+ E^Q \mathscr{I}_{\mathscr{F}_{t,T}}(Q^{s,t}_\omega; P_{t,\omega(t)}),$$

where ν_Q (resp. ν_P) is the law of (x_s, x_t) under Q (resp. P).

Proof. We shall use twice Lemma 2.1. We plug $\mathscr{F}_1 = \sigma\{x_s, x_t\}$, $\mathscr{F}_2 = \mathscr{F}_{s,t}$, $\mathscr{F}_3 = \mathscr{F}_{s,T}$, such that $\mathscr{F}_1 \subset \mathscr{F}_2 \subset \mathscr{F}_3$. Using (2.1) with the pair $(\mathscr{F}_1, \mathscr{F}_2)$, we first get

$$\mathscr{I}_{\mathscr{F}_{s,T}}(Q; P) = \mathscr{I}_{\mathscr{F}_{s,t}}(Q; P) + E^Q \mathscr{I}_{\mathscr{F}_{s,T}}(Q^{s,t}_\omega; P^{s,t}_\omega), \tag{14}$$

and with the pair $(\mathscr{F}_2, \mathscr{F}_3)$,

$$\mathscr{I}_{\mathscr{F}_{s,t}}(Q; P) = \mathscr{I}_{\mathscr{F}_{\sigma\{x_s,x_t\}}}(Q; P) + E^Q \mathscr{I}_{\mathscr{F}_{s,t}}(Q^{\sigma\{x_s,x_t\}}_\omega; P^{\sigma\{x_s,x_t\}}_\omega). \tag{15}$$

As an obvious consequence of the definition of $\mathscr{F}_1$, the first term of (15) equals $\mathscr{I}(\nu_Q; \nu_P)$; for the second term, we notice that, assumed to be finite, the Kullback information of the conditional law $Q^{\sigma\{x_s,x_t\}}_\omega$ w.r.t. $P^{\sigma\{x_s,x_t\}}_\omega$ is computed by taking the expectation of some $\sigma\{x_s, x_t)$ measurable r.v., which justifies to be rewritten as said in the proposition. We deal now with the second term of (14). It is obvious that

$$Q^{s,t}_\omega = \delta_\omega \otimes_t Q^{s,t}_\omega$$

and, because of the Markovian property of the family $(P_{s,x})$ where $s \geq 0$ and $x \in \mathbb{R}^d$,

$$P^{s,t}_\omega = \delta_\omega \otimes_t P_{t,\omega(t)}.$$

We can now deduce that

$$\mathscr{I}_{\mathscr{F}_{s,T}}(Q^{s,t}_\omega; P^{s,t}_\omega) = \mathscr{I}_{\mathscr{F}_{s,T}}(\delta_\omega \otimes_t Q^{s,t}_\omega; \delta_\omega \otimes_t P_{t,\omega(t)})$$
$$= \mathscr{I}_{\mathscr{F}_{t,T}}(Q^{s,t}_\omega; P_{t,\omega(t)}).$$

The latter equality holds for the probability distributions $\delta_\omega \otimes_t R$ are nonrandom until time t and to compute the Kullback information of the considered conditional laws it is sufficient to take $\mathscr{F}_{t,T}$-measurable functions. $\square$

We can now get the characterization of $Q^{(2)}$; we first minimize over all Q knowing $\omega \mapsto Q_\omega^{s,t}$. Therefore the restriction of $Q_\omega^{s,t}$ on $\mathscr{F}_{t,T}$ equals that of $P_{t,\omega(t)}$. Taking care of this, we see that

$$Q = \left(\int \nu_Q(\mathrm{d}x \otimes \mathrm{d}y) P_{s,x}^{t,y} \underset{t}{\otimes} P_{t,y} \right), \tag{16}$$

whenever ν_Q minimizes $\mathscr{I}(\cdot\,; \nu_P)$ having its first (resp. second) marginal identical to q_s (resp. q_t). Following Csiszar (1975) and using Lemma 2.5 from Donsker and Varadhan (1976), the latter finite dimensional problem has a unique solution ν_Q such that there exist two positive measurable functions f_s (resp. f_t) p_s (resp. p_t) a.e. strictly positive and finite with

$$\log(f_s) \in L^1(q_s), \quad \text{idem for } f_t,$$

and

$$\frac{\mathrm{d}\nu_Q}{\mathrm{d}\nu_P}(x, y) = f_s(x) f_t(y), \quad \nu_P\text{-a.s.}$$

We then deduce by a direct calculation the expression of the marginals

$$\frac{\mathrm{d}q_s}{\mathrm{d}p_s}(x) = f_s(x) \int f_t(y) \pi(s, x; t, y)\, \mathrm{d}y, \tag{17}$$

$$\frac{\mathrm{d}q_t}{\mathrm{d}p_t}(y) = f_t(y) \frac{\int p_s(\mathrm{d}x) f_s(x) \pi(s, x; t, y)}{\int p_s(\mathrm{d}x) \pi(s, x; t, y)}. \tag{18}$$

Let us define

$$h(u, x) = \int f_t(y) \pi(u, x; t, y)\, \mathrm{d}y, \quad s \leqslant u < t.$$

As $h(t, \cdot) \equiv f_t$, we straightforward verify that h is invariant for the diffusion of law P, i.e. $h(u, x) = \int \pi(u, x; t, y) h(t, y)\, \mathrm{d}y$ for all $t \in [s, T]$. As

$$\left.\frac{\mathrm{d}Q}{\mathrm{d}P}\right|_{\mathscr{F}_{s,t}} = f_s(x_s) h(t, x_t), \quad P\text{-a.s.},$$

$$\frac{\mathrm{d}q_s}{\mathrm{d}p_s}(x_s) = f_s(x_s) h(s, x_s), \quad P\text{-a.s.},$$

we see that $h(s, x_s) > 0$, P-a.s., and we finally obtain:

$$\left.\frac{\mathrm{d}Q}{\mathrm{d}P}\right|_{\mathscr{F}_{s,t}} = \frac{\mathrm{d}q_s}{\mathrm{d}p_s}(x_s) \frac{h(t, x_t)}{h(s, x_s)}, \quad P\text{-a.s.} \tag{19}$$

Notice that it is in no way assumed that the function h is strictly positive $\mathrm{d}t \otimes m_d(\mathrm{d}y)$ a.e. For this reason, let H denote the first hitting time of the closed subset $h^{-1}(\{0\})$ of $[0, \infty) \times \mathbb{R}^d$,

$$H = \inf\{\sigma > s \mid h(\sigma, x_\sigma) = 0\},$$

($\inf\{\emptyset\}=0$). Similarly, let us define the stopping times $\tilde{T}_n$ and $\tilde{T}$ by

$$\tilde{T}_n = \inf\left\{ t > s \;\middle|\; \int_s^t \left|\frac{\nabla_x h}{h}\right|^2 (u, x_u)\, \mathrm{d}u \geq n \right\},$$

$$\tilde{T} = \sup \tilde{T}_n = \lim \tilde{T}_n.$$

Obviously, $\tilde{T} \geq H$, P-a.s. Let us write an Itô formula for $u \mapsto \log h(\cdot, X_\cdot)$ $(u \wedge H \wedge \tilde{T}_n)$,

$$\log \frac{h(u \wedge H \wedge \tilde{T}_n, x_{u\wedge H\wedge \tilde{T}_n})}{h(s, x_s)} = \int_s^{u\wedge H\wedge \tilde{T}_n} \frac{\nabla_x h}{h}(v, x_v)\, \mathrm{d}\beta_v$$
$$- \frac{1}{2}\int_s^{u\wedge H\wedge \tilde{T}_n} \left\langle\!\!\left\langle \frac{\nabla_x h}{h}(v, x_v) \right\rangle\!\!\right\rangle_v^2 (x_v)\, \mathrm{d}v$$
$$+ \int_s^{u\wedge H\wedge \tilde{T}_n} \frac{1}{h}\left(\frac{\partial}{\partial v} + L_v\right)(h)(v, x_v)\, \mathrm{d}v.$$

This latter relation may be written as

$$\frac{h(t \wedge H \wedge \tilde{T}_n, x_{t\wedge H\wedge \tilde{T}_n})}{h(s, x_s)} = \exp\left(\int_s^{t\wedge H\wedge \tilde{T}_n} \frac{1}{h}\left(\frac{\partial}{\partial v} + L_v\right)(h)(v, x_v)\, \mathrm{d}v\right)$$
$$\times \mathscr{E}_s(c\cdot\beta)(t \wedge H \wedge \tilde{T}_n),$$

where β denotes the P adapted martingale

$$\beta_t = x_t - x_s - \int_s^t b(x_u, u)\, \mathrm{d}u, \quad t \in [s, T],$$

c the mapping defined on the complementary of $h^{-1}(\{0\})$ by

$$c(u, x) = \nabla_x \log h(u, x), \quad s \leq u \leq T,$$

and $\mathscr{E}_s(c\cdot\beta)$ is the exponential martingale associated with $c\cdot\beta$ and starting from s. Taking the limit $n \to \infty$, recalling that $\tilde{T} \geq H$, P-a.s., we get

$$\mathbf{1}_{\{t<H\}} \frac{h(t, x_t)}{h(s, x_s)} \exp\left(-\int_s^t \frac{1}{h}\left(\frac{\partial}{\partial v} + L_v\right)(h)(v, x_v)\, \mathrm{d}v\right) = \mathbf{1}_{\{t<H\}} \mathscr{E}_s(c\cdot\beta)(t).$$

(Whenever $t \geq H$, this latter equality boils down to $0=0$.)

If we already know that h is strictly positive everywhere, then using Theorem V.5.2 of Azencott (1983) and the regularity Assumption A2 on the kernel π of the diffusion P, the preceding equality would straightforward imply that the restriction of Q on $[s, T]$ is a diffusion law with the extra-drift: $V_s(x) = a(s, x)\nabla_x \log h(s, x)$ and new kernel

$$\tilde{\pi}(s, x; t, y) = \frac{h(t, y)}{h(s, x)} \pi(s, x; t, y).$$

Then, as a consequence of Assumption A2, we see that

$$\frac{h(t, x_t)}{h(s, x_s)} = \mathscr{E}_s(c \cdot \beta)(t) \quad \text{on } \{t < H\}.$$

Thus

$$\left.\frac{\mathrm{d}Q}{\mathrm{d}P}\right|_{\mathscr{F}_{s,u}} = \frac{\mathrm{d}q_s}{\mathrm{d}p_s}(x_s)\mathscr{E}_s(c \cdot \beta)(u, x_u). \tag{20}$$

(b) *The general case*: We now fix $n+1$ points $(s_i)_{0 \leqslant i \leqslant n}$ (by convention, $s_0 = s$) of the interval $[s, T]$ and $n+1$ probability measures on $\mathbb{R}^d$, $(q_i)_{0 \leqslant i \leqslant n}$. We look for the law $Q^{(n)}$ minimizing $I_{\mathscr{F}_{s,T}}(\,\cdot\,; P)$ under the $n+1$ constraints

$$Q \circ x_{s_i}^{-1} = q_i, \quad 0 \leqslant i \leqslant n.$$

A chain use of the above Varadhan's lemma shows us that

$$I_{\mathscr{F}_{s,t}}(Q; P) = \sum_{i=0}^{n-1} I(\nu_Q^{(i)}; \nu_P^{(i)}) + \int \nu_Q^{(i)}(\mathrm{d}x \otimes \mathrm{d}y) I_{\mathscr{F}_{s_i, s_{i+1}}}(Q_{s_i,x}^{s_{i+1},y}; P_{s_i,x}^{s_{i+1},y})$$
$$+ E^Q I_{\mathscr{F}_{s_n,T}}(Q^{s,s_n}; P_{s_n,\omega(s_n)}),$$

where $\nu_Q^{(i)}$ (resp. $\nu_P^{(i)}$) is the law of $(x_{s_i}, x_{s_{i+1}})$ under Q (resp. P).

Similarly, $Q^{(n)}$ is given, at least formally, by

$$\left(\bigotimes_{i=0}^{n-1} \bigotimes_{s_{i+1}} \int \nu_Q^{(i)}(\mathrm{d}x_i \otimes \mathrm{d}x_{i+1}) P_{s_i,x_i}^{s_{i+1},x_{i+1}}\right) \bigotimes_{s_n} P_{s_n,y}, \tag{21}$$

which is a kind of convatenation of diffusion laws, each of them being respectively restricted on $\boldsymbol{W}_{s_i,s_{i+1}}$ and such that each measure $\nu_Q^{(i)}$ is the unique solution of an analogous problem as for ν_Q but with q_s (resp. q_t) changed into q_i (resp. q_{i+1}).

We therefore get, for every $u \in [s, T]$, such that $s_i \leqslant u < s_{i+1}$ for some $0 \leqslant i \leqslant n-1$,

$$\left.\frac{\mathrm{d}Q^{(n)}}{\mathrm{d}P}\right|_{\mathscr{F}_{s,u}} = \frac{\mathrm{d}q_{s_i}^{(n)}}{\mathrm{d}p_{s_i}}(x_{s_i})\mathscr{E}_{s_i}(c^{(i)} \cdot \beta^{(i)})(u, x_u), \tag{22}$$

where $c^{(i)}$ is the same mapping obtained as in (20) and $\beta^{(i)}$ is the restriction of the Brownian motion β on $\boldsymbol{W}_{s_i,s_{i+1}}$.

2.3. *Third step: The lower bound*

Let us choose in $K(\boldsymbol{q})$ one Q, to which we shall refer all along this sub-section. As $Q \ll P$, it is well known that the process $\boldsymbol{Z} = (Z_t)_{t \geqslant s}$ where $Z_t = (\mathrm{d}Q/\mathrm{d}P)|_{\mathscr{F}_{s,t}}$ defines a positive uniformly integrable martingale. Under Assumption A1, Theorem 7.2.1 of Stroock and Varadhan (1979) yields the unicity of the martingale problem associated with L_t. Theorems 1.5 and 4.3 of Jacod and Yor (1977) thus show that every adapted $L^1(P)$-martingale admits the decomposition as a sum of its initial value and of a stochastic integral of the elementary characteristic martingale. In particular, there exists one $(\mathscr{F}_{s,t})_{t\geq s}$ previsible process $(\phi_u)_{s\leq u\leq T}$ such that

$$Z_t = Z_s + \int_s^t \phi_u \, \mathrm{d}M_u \tag{23}$$

where $(M_u)_{u\geq s}$ denotes the $(\mathscr{F}_{s,u})_{s\leq u\leq T}$ adapted martingale

$$M_u^i = x_u^i - x_s^i - \int_s^u b^i(x_\sigma, \sigma)\,\mathrm{d}\sigma, \quad 1\leq i\leq d,$$

with quadratic variation process

$$\mathrm{d}\langle M_u^i, M_u^j\rangle = a^{i,j}(x_u, u)\,\mathrm{d}u, \quad 1\leq i, j\leq d.$$

Thus, $M.$ being continuous, so is $Z.$.

Let us quickly show that $Z_s = (\mathrm{d}q_s/\mathrm{d}p_s)(x_s)$. By definition, Z_s is the Radon-Nikodym derivative of Q restricted to $\sigma\{x_s\}$ with respect to P restricted to $\sigma\{x_s\}$. Thus, for any bounded borel map f,

$$\begin{aligned}\int f(x_s)Z_s\,\mathrm{d}P &= \int f(x_s)\,\mathrm{d}Q = \int f(y)q_s(\mathrm{d}y)\\ &= \int f(y)\frac{\mathrm{d}q_s}{\mathrm{d}p_s}(y)p_s(\mathrm{d}y) = \int f(x_s)\frac{\mathrm{d}q_s}{\mathrm{d}p_s}(x_s)\,\mathrm{d}P.\end{aligned}$$

Noticing that Z_s is strictly positive P-a.s. (because p_s and q_s are equivalent), we can define the following stopping times:

$$T_n = \inf\{t > s \mid Z_t < 1/n\}, \quad n\geq 1,$$
$$T^* = \inf\{t > s \mid Z_t = 0\},$$

such that T_n converges increasingly towards T^* P-a.s. The T_n-stopped martingale $Z^{T_n} = Z_{\cdot\wedge T_n}$ is a nonnegative margingale, so that the P-martingale Y^{T_n} can be defined by

$$Y_t^{T_n} = \int_0^{t\wedge T_n} \frac{\mathrm{d}Z_u}{Z_u}.$$

Moreover, this one fulfills the property that

$$Y_t^{T_n} - Y_s^{T_n} - \int_s^t \frac{\mathrm{d}\langle Z^{T_n}, Y^{T_n}\rangle_u}{Z_u^{T_n}}$$

is a Q-martingale. Solving the latter equation, we see that

$$Z_t^{T_n} = Z_s^{T_n}\mathscr{E}(Y_\cdot^{T_n})_t,$$

i.e.

$$Z_{t\wedge T_n} = Z_{s\wedge T_n}\exp\{Y_{t\wedge T_n} - \tfrac{1}{2}\langle Y\rangle_{t\wedge T_n}\}, \quad t\in[s, T].$$

Taking the $n \to \infty$ limit in this relation and using the continuity of Z, we deduce

$$Z_{t \wedge T^*} = Z_{s \wedge T^*} \exp\{ Y_{t \wedge T^*} - \tfrac{1}{2}\langle Y \rangle_{t \wedge T^*} \}, \quad t \in [s, T].$$

Let us show that $Q(T^* < +\infty) = 0$. First of all, $Q(T_n < \infty) = E_P(Z_{T_n} \mathbf{1}_{T_n < \infty}) \leq 1/n$ and this implies

$$Q\left(\bigcap_{n \geq 1} \{T_n < \infty\} \right) = 0 = Q(T^* < \infty).$$

Thus, $T^* = +\infty$, Q-a.s. Using Theorem 1.11 of Sharpe (1980) for the P-martingale Y^{T^*}, we see that

$$\{T^* < \infty\} \cap \{\langle Y \rangle_{T^*} < \infty\} = \{T^* < \infty\} \cap \{ Y_{T^*} \text{ exists and is finite}\}$$

up to P-negligible events. As $Z_{T^*} = 0$ on $\{T^* < \infty\}$ P-a.s., by continuity of Z and the above expression of Z_{T^*} (put $t = s$), we also deduce that these events are P-a.s. equal to

$$\{T^* < \infty\} \cap \{Z_{T^*} > 0\},$$

and that $\langle Y \rangle_{T^*} = +\infty$ on $\{T^* < \infty\}$ P-a.s. It is also well known that $Z = 0$ P-a.s. on $[\![T^*, \infty [\![$ (Lemma III.3.6 in Jacod and Shyriaev, 1988). This enables us to compute $\mathcal{I}_{\mathcal{F}_{s,t}}(Q; P)$ for every $t \geq s$ (recall that $\mathcal{I}_{\mathcal{F}_{s,t}}(Q; P) = E_P(Z_t \log Z_t)$ whenever $Z_t(\log Z_t)^+ \in L^1(P)$, and $+\infty$ otherwise; we shall put $0 \log 0 = 0$). We easily get via the Girsanov's theorem the following expression:

$$\mathcal{I}_{\mathcal{F}_{s,T_n}}(Q; P) = \mathcal{I}_{\mathcal{F}_s}(q_s; p_s) + \tfrac{1}{2} E_Q(\langle Y \rangle_{T_n})$$

which implies via Lemma 2.1 that

$$\tfrac{1}{2} E_Q \langle Y \rangle_{T_n} \leq \mathcal{I}_{\mathcal{F}_{s,T_n}}(Q; P) \leq \mathcal{I}(Q; P) < \infty.$$

Taking the limit, we finally obtain

$$\tfrac{1}{2} E_Q\{\langle Y \rangle_\infty;\ T^* = \infty\} + \tfrac{1}{2} E_Q\{\langle Y \rangle_{T^*};\ T^* < \infty\} < \infty. \tag{24}$$

But, as $Q(T^* < +\infty) = 0$, we deduce

$$E_Q\{\langle Y \rangle_{T^*}\} < \infty,$$

$$\mathcal{I}_{\mathcal{F}_{s,t}}(Q; P) = \mathcal{I}_{\mathcal{F}_s}(q_s; p_s) + \tfrac{1}{2} E_Q(\langle Y \rangle_t).$$

Similarly, on $\{T^* = +\infty\}$, $Z_\infty > 0$, P-a.s. and the restrictions of Q and P on $\{T^* = +\infty\}$ are equivalent. Let $J(\phi)$ denote the function defined on $\mathbb{R}^d \times [0, +\infty]$ by

$$J(\phi)(x, t) = \phi \cdot b(x, t) + \tfrac{1}{2} \langle\!\langle \phi \rangle\!\rangle_t^2(x).$$

Re-using the representation of the P-martingales, we finally get the existence of an adapted process ψ such that, for every $t \geq s$,

$$Z_t = \mathbf{1}_{t < T^*} Z_s \exp\left[\int_s^t \psi_u \, dx_u - \int_s^t J(\psi_u)(x_u, u) \, du \right].$$

As a consequence of the representation of $Q^{(n)}$ already obtained in the second step (cf. (22)), we observe that $Q^{(n)}$ is (locally) equivalent to P, which implies, as before, the existence of an adapted process $(\psi_t^{(n)})_{t\geqslant s}$ such that, for all $t\in[s,T]$,

$$\frac{\mathrm{d}Q^{(n)}}{\mathrm{d}P}\bigg|_{\mathscr{F}_{s,t}}=\frac{\mathrm{d}q_s^{(n)}}{\mathrm{d}p_s}(x_s)\exp\left[\int_s^t \psi_u^{(n)}\,\mathrm{d}x_u-\int_s^t J(\psi_u^{(n)})(x_u,u)\,\mathrm{d}u\right]. \tag{25}$$

We then deduce the expression of the Kullback information of Q (resp. $Q^{(n)}$) with respect to P,

$$\mathscr{I}_{\mathscr{F}_{s,t}}(Q;P)=\mathscr{I}_{\mathscr{F}_s}(q_s;p_s)+E_Q\left(\frac{1}{2}\int_s^t \langle\!\langle\psi_u\rangle\!\rangle_u^2(x_u)\,\mathrm{d}u\right), \tag{26}$$

and

$$\mathscr{I}_{\mathscr{F}_{s,t}}(Q^{(n)};P)=\mathscr{I}_{\mathscr{F}_s}(q_s^{(n)};p_s)+E_{Q^{(n)}}\left(\frac{1}{2}\int_s^t \langle\!\langle\psi_u^{(n)}\rangle\!\rangle_u^2(x_u)\,\mathrm{d}u\right), \tag{27}$$

using the fact that, if $H_.$ is a P-martingale, the process

$$K_t=H_t-\left\langle H,\int_s^{\cdot}\psi_u\,\mathrm{d}M_u\right\rangle_t$$

(resp. $K^{(n)}$ similarly defined but with ψ replaced by $\psi^{(n)}$) is a Q- (resp. $Q^{(n)}$-) martingale.

Let us now give the following:

Definition 2.2. Let Q be an element of $K(\boldsymbol{q})$ and ψ the adapted process defined as above. Then, for any $t\in[s,T]$, $\tilde{\psi}_t$ denotes a regular conditional probability version of ψ_t knowing x_t, i.e.

$$\tilde{\psi}_t(x)=E_Q[\psi_t\,|\,x_t=x].$$

Thus, as $E_Q[\langle\!\langle\psi_u\rangle\!\rangle_u^2(x_u)\,|\,x_u]=\langle\!\langle\tilde{\psi}_u\rangle\!\rangle_u^2(x_u)+E_Q[\langle\!\langle\psi_u-\tilde{\psi}_u\rangle\!\rangle_u^2(x_u,u)]$, we get the lower bound

$$\mathscr{I}_{\mathscr{F}_{s,t}}(Q;P)\geqslant\mathscr{I}_{\mathscr{F}_s}(q_s;p_s)+\tfrac{1}{2}E_Q\left(\int_s^t\int_{\mathbb{R}^d}\langle\!\langle\tilde{\psi}_u\rangle\!\rangle_u^2(y)q_u(\mathrm{d}y)\,\mathrm{d}u\right). \tag{28}$$

Let us now show how this lower bound is related to the norms $\|\cdot\|_{q_u,u}$. Because of our assumptions, P is the unique solution of the (a,b)-martingale problem, such that, for all $f\in C_K^\infty(\mathbb{R}^d)$, the process

$$\left(f(x_t)-f(x_s)-\int_s^t L_uf(x_u)\,\mathrm{d}u\right)_{t\geqslant s}$$

is an adapted P-martingale. Consequently, thanks to the representation of the martingale Z, we compute for any Q in $K(\boldsymbol{q})$,

$$\begin{aligned}E_Q(f(x_t)-f(x_s))&=\langle f,q_t\rangle-\langle f,q_s\rangle\\&=\int_s^t\langle\langle\nabla_uf,\tilde{\psi}_u\rangle_u,q_u\rangle\,\mathrm{d}u+\int_s^t\langle L_uf,q_u\rangle\,\mathrm{d}u.\end{aligned}$$

We thus have, in the distributional sense,

$$\langle \dot{q}_t - L_t^* q_t, f\rangle = \langle q_t, \langle \nabla_t f, \tilde{\psi}_t\rangle_t\rangle. \tag{29}$$

But

$$\tfrac{1}{2}\langle q_t, \langle\!\langle \tilde{\psi}_t \rangle\!\rangle_t^2\rangle = \tfrac{1}{2}\langle q_t, \langle\!\langle \tilde{\psi}_t - \nabla_t f\rangle\!\rangle_t^2\rangle - \tfrac{1}{2}\langle q_t, \langle\!\langle \nabla_t f\rangle\!\rangle_t^2\rangle + \langle q_t, \langle \nabla_t f, \tilde{\psi}_t\rangle_t\rangle$$
$$\geqslant \langle \dot{q}_t - L_t^*(q_t), f\rangle - \tfrac{1}{2}\langle q_t, \langle\!\langle \nabla_t f\rangle\!\rangle_t^2\rangle.$$

Because of the definition of $\|\cdot\|_{q_t,t}$, given just before Definition 1.1, the preceding inequality yields

$$\|\dot{q}_t - L_t^* q_t\|_{q_t,t}^2 \leqslant \tfrac{1}{2}\langle q_t, \langle\!\langle \tilde{\psi}_t\rangle\!\rangle_t^2\rangle. \tag{30}$$

We summarize these results in:

Proposition 2.2. *For all Q in $K(\boldsymbol{q})$ and all $t\in[s, T]$,*

$$\mathcal{I}_{\mathcal{F}_{s,t}}(Q; P) \geqslant \mathcal{I}_{\mathcal{F}_s}(q_s; p_s) + \int_s^t \|\dot{q}_t - L_t^* q_u\|_{q_u,u}^2 \, \mathrm{d}u. \qquad \square$$

2.4. Fourth step: The lower bound is attained at $\bar{Q}$

We have shown at the second step that $\mathcal{I}_{\mathcal{F}_{s,t}}(\bar{Q}; Q)^{(n)})$ tends to 0 as n tends to ∞, for each $t\in[s, T]$. Using the above 'Pythagoras' equalities (12), we compute

$$\mathcal{I}_{\mathcal{F}_{s,t}}(\bar{Q}; Q^{(n)}) = \mathcal{I}_{\mathcal{F}_{s,t}}(\bar{Q}; P) - E_{\bar{Q}}\left[\log\left(\frac{\mathrm{d}Q^{(n)}}{\mathrm{d}P}\right)\right],$$

$$\mathcal{I}_{\mathcal{F}_{s,t}}(\bar{Q}; Q^{(n)}) = \mathcal{I}_{\mathcal{F}_s}(\bar{q}_s; p_s) + \tfrac{1}{2}E_{\bar{Q}}\left(\int_s^t \langle\!\langle \psi_u\rangle\!\rangle_u^2(x_u)\,\mathrm{d}u\right)$$
$$-\left[\mathcal{I}_{\mathcal{F}_s}(q_s^{(n)}; p_s) + \tfrac{1}{2}E_{\bar{Q}}\left(\int_s^t \langle\!\langle \psi_u^{(n)}\rangle\!\rangle_u^2(x_u)\,\mathrm{d}u\right)\right].$$

We finally get the following expression for our null convergent sequence:

$$\mathcal{I}_{\mathcal{F}_{s,t}}(\bar{Q}; Q^{(n)}) = \mathcal{I}_{\mathcal{F}_s}(\bar{q}_s; q_s^{(n)}) + \tfrac{1}{2}E_{\bar{Q}}\left(\int_s^t \langle\!\langle \psi_u^{(n)} - \psi_u\rangle\!\rangle_u^2(x_u)\,\mathrm{d}u\right)$$

(where we have used the facts that $\bar{Q} \ll Q^n$ and, consequently that $\bar{q}_s \ll q_s^{(n)}$). But, as proved in the first step, $Q^{(n)}$ converges weakly and in variation norm to $\bar{Q}$ such that, for all $n\geqslant 1$, $\bar{Q}\ll Q^{(n)}$ and $\mathcal{I}_{\mathcal{F}_s}(\bar{q}_s; q_s^{(n)})$ tends to 0. We then conclude from the last equality that

$$(u, \omega) \mapsto \langle\!\langle \psi_u^{(n)} - \psi_u\rangle\!\rangle_u(x_u)$$

converges to 0 in $L^2(\Omega\times[s, T], \bar{Q}\otimes \mathrm{d}u)$.

The main result of the second step shows that for all $t\geqslant s$, the process $\psi_t^{(n)}$ is $\sigma(x_t)$-measurable and that there exists a piecewise continuous C^∞ in x function $h^{(n)}$ such that

$$\psi_u^{(n)}(x) = \nabla_x \log h^{(n)}(u, x) \quad \text{for almost all } (u, x) \text{ w.r.t. } \mathrm{d}u\otimes m_d(\mathrm{d}x).$$

Due to Assumption A1, we see that, for almost all $u \in [s, T]$, $\psi_u^{(n)}$ converges to ψ_u in $L^2(\mathbb{R}^d, q_u)$, and, as $\psi_u^{(n)}$ is $\sigma(x_u)$ measurable, that ψ_u is too (one can easily verify that if $(Z_n)_{n\geq 0}$ is some L^2 convergent sequence of $\sigma(X)$ measurable random variables, its limit is $\sigma(X)$ measurable). This implies that $\tilde{\psi}_u = \psi_u$ for almost all $u \in [s, T]$ and that $\bar{Q}$ is Markovian. Recall (cf. Section 2.2(a) above) that $H > t$ iff $Z_t > 0$: thus, $T^* = H$ and $H = +\infty$, $\bar{Q}$-a.s. We may also use the results contained in Nagasawa (1989) which are especially suited for general h. According to the result in Section 2.2(b), one can find a sequence of functions $k^{(n)}: \mathbb{R}^d \times [O, +\infty)$, C_K^∞ in x such that, for almost all $u \in [s, T]$,

$$\tilde{\psi}_u = \lim_{n\to\infty} \nabla_x k^{(n)}(u, x)$$

in $L^2(\mathbb{R}^d, q_u)$ norm. (In fact, before taking smooth approximations, the functions match with $\nabla_x \log h^{(i)}$ on each subinterval $[s_i, s_{i+1})$ and that is why the preceding equality takes place for all u but a denumerable union of finite subsets of $[s, T]$.)

In fact, by means of the relation following (29), we have

$$\tfrac{1}{2}\langle q_u, \langle\!\langle \tilde{\psi}_u \rangle\!\rangle_u^2 \rangle = \tfrac{1}{2}\langle q_u, \langle\!\langle \tilde{\psi}_u - \nabla_u k_u^{(n)} \rangle\!\rangle_u^2 + \langle q_u, \langle \tilde{\psi}_u, \nabla_u k_u^{(n)} \rangle_u \rangle - \tfrac{1}{2}\langle q_u, \langle\!\langle \nabla_u k_u^{(n)} \rangle\!\rangle_u^2 \rangle.$$

The first term of the right side of this equality tends to 0; the second term is equal to $\langle \dot{q}_u - L_u^* q_u, k_u^{(n)} \rangle$. By definition of the norm $N_u = \|\dot{q}_u - L_u^* q_u\|_{q_u,u}^2$, we have

$$N_u \geq \limsup_{n\to\infty} (\langle \dot{q}_u - L_u^* q_u, k_u^{(n)} \rangle - \tfrac{1}{2}\langle q_u, |\nabla_u k_u^{(n)}|_u^2 \rangle).$$

The quantity of which we are taking the limit sup equals

$$o(1) + \tfrac{1}{2}\langle q_u, \langle\!\langle \tilde{\psi}_u \rangle\!\rangle_u^2 \rangle,$$

so that $N_u \geq \frac{1}{2}\langle q_u, \langle\!\langle \tilde{\psi}_u \rangle\!\rangle_u^2 \rangle$. Using the inequality (30), we conclude that the latter inequality is in fact an equality. Thus

$$\begin{aligned} \mathcal{I}_{\mathcal{F}_{s,t}}(\bar{Q}; P) &= \mathcal{I}_{\mathcal{F}_s}(\bar{q}_s; p_s) + \frac{1}{2}\int_s^t \langle q_u, \langle\!\langle \tilde{\psi}_u \rangle\!\rangle_u^2 \rangle \, \mathrm{d}u \\ &= \mathcal{I}_{\mathcal{F}_s}(\bar{q}_s; p_s) + \int_s^t \|\dot{q}_u - L_u^* q_u\|_{q_u,u}^2 \, \mathrm{d}u. \end{aligned}$$

And, as $\bar{Q} \in K(\boldsymbol{q})$, we get an equality in Proposition 2.2. We conclude that the lower bound is really achieved at $\bar{Q}$, and at this point only. □

Acknowledgement

The author would like to thank Pr. M. Yor for fruitful discussions.

References

R. Azencott, Densité des diffusions en temps petit: développement asymptotique, in: Séminaire de Probabilités XVIII, 1982–1983, Lecture Notes in Math. No. 1059 (Springer, New York, 1983).

G. Ben Arous and M. Brunaud, Méthode de Laplace: applications à l'étude variationnelle des fluctuations de diffusions de type "champ moyen", to appear in: Stochastics (1990).

I. Csiszar, I-divergence geometry of probability distributions and minimizations problems, Ann. Probab. 3(1) (1975) 146-158.

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

M.D. Donsker and S.R.S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time III, Comm. Pure Appl. Math. 29 (1976) 389-461.

H. Föllmer, Random fields and diffusion processes, Ecole d'été de Probabilités de Saint-Flour XVI, 1988, Lecture Notes in Math. No. 1362 (Springer, New York, 1986).

J. Jacod and A.N. Shyriaev, Limit theorems for stochastic processes, Grundlehren der Math. Wissenschaften No. 288 (Springer, New York, 1988).

J. Jacod and M. Yor, Etude des solutions extrémales et représentation intégrale des solutions pour certains problèmes de martingales, Z. Wahrsch. Verw. Gebiete 38 (1977) 83-125.

M. Nagasawa, Transformation of diffusion and Schrödinger processes, Probab. Theory Rel. Fields 82 (1989) 109-136.

M.J. Sharpe, Local times and singularities of continuous martingales, Séminaire de Probabilités XIV, 1978-79, Lecture Notes in Math. No 784 (Springer, New York, 1980).

D.W. Stroock and S.R.S. Varadhan, Multidimensional diffusion processes, Grundlehren der Math. Wissenschaften No. 233 (Springer, New York, 1979).

S.R.S. Varadhan, Large deviations and applications, C.B.M.S. 46 (SIAM, Philadelphia, PA, 1984).