Logarithmic Sobolev inequalities of diffusions for the L^2 metric

Mathieu Gourcy[*], Liming Wu[†]

February 2006

Abstract

Under the Bakry-Emery's Γ_2-minoration condition, we establish the logarithmic Sobolev inequality for the Brownian motion with drift in the metric L^2 instead of the usual Cameron-Martin metric. The involved constant is sharp and does not explode for large time. This inequality with respect to the L^2-metric provides us the gaussian concentration inequalities for the large time behavior of the diffusion.

Keywords : Logarithmic Sobolev inequality (LSI), concentration inequality, path space, Malliavin calculus.

AMS Subject Classification 2000: 60E15, 60H07.

1 Introduction

The logarithmic Sobolev inequality (LSI in short), discovered by Gross ([16], 1975) for Gaussian measures and the Wiener measure, plays a prominent role in the infinite dimensional analysis as the Sobolev inequalities do in the finite dimensional analysis. Herbst first, Aida-Masuda-Shigekawa [2] and Aida-Stroock [3] found that LSI implies the Gaussian concentration inequality for Lipschitzian functions, Bobkov-Götze [6] characterized the Gaussian concentration inequality for Lipschitzian functions by means of the so called T_1-transportation inequality, which is equivalent to the Gaussian integrability of the distance function by the recent work of H. Djellout and al. [10]. Otto and Villani [20] and Bobkov-Gentil-Ledoux [7] find that LSI implies the Talagrand T_2-transportation inequality. See Bakry [4], Ledoux [17] and Villani [21] for systematic treatment and further study.

---

[*] Laboratoire de Mathématiques Appliquées, CNRS-UMR 6620, Université Blaise Pascal, 63177 Aubière, France, mathieu.gourcy@math.univ-bpclermont.fr

[†] Laboratoire de Mathématiques Appliquées, CNRS-UMR 6620, Université Blaise Pascal, 63177 Aubière, France, Li-Ming.Wu@math.univ-bpclermont.fr and Department of Mathematics, Wuhan University, 430072 China

For introducing our question for diffusions, let us first recall the LSI of Gross on the Wiener space.

Let μ be the Wiener measure on $W^d = C\big([0,T],\mathbb{R}^d\big)$ and $h \in H$ where H is the Cameron Martin space, i.e, the Hilbert space H of all absolutely continuous functions $h : [0,T] \to \mathbb{R}^d$ such that $h(0) = 0$ and

$$\|h\|_H^2 := \int_0^T |\dot{h}(t)|^2 dt < \infty.$$

The directional derivative along h of a smooth function f at $\gamma \in W^d$ is defined on W^d by :

$$D_h f(\gamma) = \lim_{\varepsilon \to 0} \frac{f(\gamma + \varepsilon h) - f(\gamma)}{\varepsilon}.$$

The linearity and continuity in h of the above expression gives the existence of an element $Df(\gamma) = (Df(\gamma,t))_{t\in[0,T]}$ in H (the Malliavin gradient) such that :

$$\forall h \in H, D_h f(\gamma) = \langle Df(\gamma), h\rangle_H = \int_0^T \dot{\widehat{Df}}(\gamma,t) \cdot \dot{h}(t) dt \tag{1.1}$$

where $\dot{\widehat{Df}}(\gamma,t) = \frac{d}{dt} Df(\gamma,t)$.

For the Wiener measure μ on W^d, Gross ([16], 1975) proved the following (see also the Ph.D of Gentil [13]):

Theorem 1.1. *For all bounded and smooth functions F on W^d,*

$$Ent_\mu(F^2) \le 2C\mathbb{E}^\mu \left(\|DF\|_H^2\right) \tag{1.2}$$

where the constant $C = 1$ is sharp.

Here the entropy of $0 \le f \in L^1(\mu)$ w.r.t. μ is defined as

$$Ent_\mu(f) = \mathbb{E}^\mu \left(f \log \left(\frac{f}{\mathbb{E}^\mu f} \right) \right).$$

When μ satisfies (1.2), we shall say that μ satisfies the LSI(C) w.r.t. the gradient in H.

This inequality w.r.t. the Cameron-Martin metric is extended to the Brownian motion over a Riemannian manifold M by Hsu [15], Aida [1], Capitaine-Hsu-Ledoux [8] under the boundedness of the Ricci curvature, and to general elliptic diffusions under the boundedness of the Bakry-Emery curvature. Particularly Aida [1] proved that the constant in the LSI does not explode for large time T under the uniform positivity of the Ricci curvature. But as explained by the second author in [23], the LSI w.r.t. the Cameron-Martin metric for diffusion (X_t) even with a constant independent of T (as in [1]) does not produce the concentration inequality of correct order in large time T for the functionals

$$F(X) := \frac{1}{\sqrt{T}} \left(\int_0^T g(X_s) ds - \mathbb{E} \int_0^T g(X_s) ds \right), \ g : M \to \mathbb{R} \text{ smooth}.$$

The reason is that the Lipschitzian coefficient of F above w.r.t. the Cameron-Martin metric is of order $\sqrt{T}$ in general, which explodes. To get useful informations for the ergodic behavior (i.e., the large time behavior) of (X_t) or more precisely for producing the Hoeffding's type Gaussian concentration inequality for functionals like $F(X)$ above for large time, as pointed out by Djellout, Guillin and the second author in [10], we should establish the LSI w.r.t. **the L^2-metric** instead of the Cameron-Martin metric, with a constant which does not explode for large time.

The famous Γ_2-minoration criterion of Bakry and Emery [5] gives the LSI for the single time law $P_T(x,\cdot) = \mathbb{P}_x(X_T \in \cdot)$ and for the unique invariant measure m of (X_t). The main purpose of this Note is to extend their LSI to the process-level $X_{[0,T]}$ w.r.t. the L^2-metric. This work is also a continuation of [10] where the Talagrand T_2-transportation inequality for diffusions w.r.t. the L^2-metric is established via the Girsanov transformation, under a weaker condition than the Bakry-Emery condition (see also [24] for T_2-inequality w.r.t. a uniform norm).

This paper is organized as follows. In the next section we first study the Brownian motion with drift, i.e., solution of the SDE

$$dX_t = dW_t + b(X_t)dt, \ X_0 = x \tag{1.3}$$

in $\mathbb{R}^d$, where the drift b satisfies the Bakry-Emery condition:

$$(\nabla b)^\sigma \le -KId, \ K \in \mathbb{R}$$

where $(\nabla b)^\sigma = ((\partial_j b_i + \partial_i b_j)/2)$ is the symmetrized gradient of the vector field b. Our method is very simple: to prove that the mapping Φ from the path W of the Brownian motion equipped with the Cameron-Martin metric to the path X of the diffusion equipped of $L^2([0,T],\mathbb{R}^d)$-norm is Lipschitzian.

The method above does not work if there is a non-constant volatility coefficient $\sigma(X_t)$ in the SDE (1.3) above. In that case, we consider a general elliptic diffusion with generator $\frac{1}{2}\Delta + V$ on a Riemannian manifold. We shall employ the method of the martingale representation formula, as developed by Fang [12] for Poincaré inequality and by Capitaine-Hsu-Ledoux [8] for log-Sobolev inequality. For the Brownian motion over a Riemannian manifold, when the Ricci curvature is bounded, we shall see in §3 how to derive the LSI quickly from the Fang's martingale representation formula. But in the general case where the Bakry-Emery's curvature is not bounded, Fang's version of martingale representation formula seems to be no longer true and that is the main difficulty. We shall use a localization technique to get a martingale representation formula for a class of test functions, and then the LSI. This is the task of §4. Finally in §5 we present several useful consequences of the LSI about the concentration of additive functionals of the diffusions for large time.

In this note we denote the maximum of u and v by $u \vee v$ and their minimum by $u \wedge v$. For $x, y \in \mathbb{R}^d$, $x \cdot y = \sum_{i=1}^d x_i y_i$ and $|x|$ denotes the Euclidian norm of x. When there is no possible confusion, we will use L^2 for $L^2([0,T];\mathbb{R}^d)$.

2 Brownian motion with drift on $\mathbb{R}^d$

2.1 Gradient ∇ in L^2 and the LSI w.r.t. ∇

We are given a standard $\mathbb{R}^d$-valued Brownian motion W_t on $\mathbb{R}^d$, a fixed initial point x in $\mathbb{R}^d$ and a continuously differentiable mapping $b : \mathbb{R}^d \to \mathbb{R}^d$. Let us consider the $\mathbb{R}^d$-valued diffusion $X_t(x)$ defined by the stochastic differential equation :

$$dX_t = dW_t + b(X_t)dt \ , \quad X_0 = x \in \mathbb{R}^d. \tag{2.1}$$

We assume the following condition on the drift b

$$(\nabla b)^\sigma \leq -K \ Id, \ \ K \in \mathbb{R} \tag{2.2}$$

where Id is the identity matrix in $\mathbb{R}^d$, $\nabla b = (\partial_j b_i)_{ij}$ is the Jacobian matrix, and A^σ denotes the symmetrization $(A + A^t)/2$ of the matrix A; and for two symmetric matrices A, B, $A \leq B$ means that $B - A$ is non-negative definite.

This condition, when $K > 0$, is exactly the Bakry-Emery's Γ_2-minoration condition for the logarithmic Sobolev inequality of the unique invariant measure m on $\mathbb{R}^d$. Details about this condition, and applications to hypercontractive diffusions can be found in [4], [5]. We are interested here in the logarithmic Sobolev inequality on the path space w.r.t. the gradient in $L^2([0,T], \mathbb{R}^d)$. More precisely, regarding $W^d = C([0,T], \mathbb{R}^d)$ as a dense subspace of $L^2([0,T], \mathbb{R}^d)$, we can introduce:

Definition 2.1. *For a function f on W^d, differentiable with respect to the L^2-norm, let $\nabla f(\gamma)$ be the element in $L^2([0,T], \mathbb{R}^d)$ such that*

$$D_g f(\gamma) = \langle \nabla f(\gamma) \ , \ g \rangle_{L^2([0,T],\mathbb{R}^d)} \ , \ \forall g \in H.$$

We shall write $\nabla f(\gamma) = (\nabla_t f(\gamma))_{t \in [0,T]}$.

Its relation with the gradient Df in the Cameron-Martin space H (introduced in (1.1)) is given by

Lemma 2.2. *For a function f on W^d, differentiable with respect to the L^2-norm, we have for every $\gamma \in W^d$, $\ddot{\widehat{D}} f(\gamma)$ is in $L^2([0,T], \mathbb{R}^d)$ in the sense of distribution and*

$$\nabla_t f(\gamma) = -\ddot{\widehat{D}} f(\gamma, t), \quad dt - a.e.$$

Proof. Such function f is differentiable w.r.t. the norm of H. Let $C_0^\infty((0,T), \mathbb{R}^d)$ be the space of all $\mathbb{R}^d$-valued C^∞-functions with compact support in $(0,T)$, which is dense in $L^2([0,T], \mathbb{R}^d)$. For every $h \in C_0^\infty((0,T), \mathbb{R}^d)$, writing $\ddot{\widehat{D}} f(\gamma, t) = \frac{d^2}{dt^2} Df(\gamma, t)$ in the sense of distribution, we have,

$$\int_0^T \nabla_s f(\gamma) \cdot h(s) ds \ = \ \langle \nabla f(\gamma), h \rangle_{L^2} = D_h f$$

$$\begin{aligned} &= \int_0^T \dot{\widehat{D}} f(\gamma, s) \cdot \dot{h}(s) ds \\ &= -\int_0^T \ddot{\widehat{D}} f(\gamma, s) \cdot h(s) ds \end{aligned}$$

where the desired result follows. □

We denote by $C_b^1(W^d/L^2)$ the space of all bounded functions f on W^d, differentiable w.r.t. the $L^2([0,T],\mathbb{R}^d)$-norm, such that ∇f is also continuous and bounded from $(W^d, \|\cdot\|_{L^2})$ to $L^2([0,T],\mathbb{R}^d)$. The aim of this section is to prove

Theorem 2.3. *Let $\mathbb{P}_x$ be the law of the diffusion $X.(x)$ defined by the SDE (2.1). Assume (2.2). Then for all functions $f \in C_b^1(W^d/L^2)$,*

$$Ent_{\mathbb{P}_x}(f^2) \leq 2C(T)\mathbb{E}^{\mathbb{P}_x} \int_0^T |\nabla_t f|^2 dt \tag{2.3}$$

where

$$C(T) := \begin{cases} \frac{1}{K^2}\left(1 - \sqrt{2e^{-KT} - e^{-2KT}}\right), & \text{if } \ 2e^{-KT} - e^{-2KT} > 0, K \neq 0 \\ \frac{T^2}{2}, & \text{if } \ K = 0 \\ \frac{1}{K^2}\left(e^{-2KT} - 2e^{-KT} + 1\right), & \text{if } \ 2e^{-KT} - e^{-2KT} \leq 0, K \neq 0. \end{cases} \tag{2.4}$$

In particular if $K > 0$, $C(T) \leq 1/K^2$ (for all $T > 0$) is non-explosive.

Note that our estimate of $C(T)$ is sharp for $T \to +\infty$ for the Ornstein-Uhlenbeck process $dX_t = dW_t - (X_t/2)dt$ by the discussion in [10, Remarks 5.7].

2.2 Analytic preparation

In this paragraph we recall and prove several known facts, for self-containedness. Let $M_d(\mathbb{R})$ be the space of all real $d \times d$-matrices. For $A \in M_d(\mathbb{R})$, it is clear that $\forall x \in \mathbb{R}^d$, $\langle Ax \ , \ x\rangle = \langle A^\sigma x \ , \ x\rangle$. We first give a Gronwall Lemma of matrix type.

Lemma 2.4. *If $J(t) \in M_d(\mathbb{R})$ satisfies the equation*

$$\dot{J}(t) = A(t)J(t) \quad \text{with} \quad J(0) = Id,$$

and if $A^\sigma \leq BId$ for a real number B in the order of the nonnegative definiteness, then we have

$$|J(t)y| \leq e^{Bt}|y|, \forall y \in \mathbb{R}^d, \ \forall t \geq 0.$$

Proof. Indeed we have :

$$\begin{aligned} \frac{d}{dt}\big(e^{-2Bt}|J(t)x|^2\big) &= e^{-2Bt}\big(-2B|J(t)x|^2 + 2\langle \dot{J}(t)x, J(t)x\rangle\big) \\ &= e^{-2Bt}\big(-2B|J(t)x|^2 + 2\langle A(t)J(t)x, J(t)x\rangle\big) \\ &= e^{-2Bt}\big(-2B|J(t)x|^2 + 2\langle A^\sigma(t)J(t)x, J(t)x\rangle\big) \end{aligned}$$

$$
\begin{aligned}
&= -2e^{-2Bt}\langle (BId - A^{\sigma}(t))\, J(t)x, J(t)x\rangle \\
&\leq 0
\end{aligned}
$$

so,

$$|J(t)x| \leq e^{Bt}|x|.$$

$\square$

We now give a result to control the norm of a bounded self-adjoint operator from L^2 to L^2.

Lemma 2.5. *Let* Γ *be a bounded self-adjoint operator from* $L^2(E,\mu)$ *to* $L^2(E,\mu)$ *where* μ *is* σ*-finite. Assume that*

$$\|\Gamma\|_{1,1} := \sup_{0\neq f\in L^1\bigcap L^2} \frac{\|\Gamma f\|_{L^1(\mu)}}{\|f\|_{L^1(\mu)}} < +\infty.$$

Then $\|\Gamma\|_{2,2}$*(the norm in* L^2*)* $\leq \|\Gamma\|_{1,1}$.

Proof. For any $f \in L^1 \bigcap L^\infty \subset L^2$, we have by the symmetry of Γ,

$$
\begin{aligned}
\|\Gamma f\|_\infty &= \sup_{g\in L^1\bigcap L^\infty, \|g\|_1\leq 1} \langle \Gamma f, g\rangle \\
&= \sup_{g\in L^1\bigcap L^\infty, \|g\|_1\leq 1} \langle f, \Gamma g\rangle \\
&\leq \|\Gamma\|_{1,1}\|f\|_\infty,
\end{aligned}
$$

i.e., $\|\Gamma\|_{\infty,\infty} \leq \|\Gamma\|_{1,1}$. Hence by the Riesz-Thorin's interpolation theorem, $\|\Gamma\|_{2,2} \leq \|\Gamma\|_{1,1}$. $\square$

2.3 Proof of Theorem 2.3

2.3.1 Outline of the proof

The SDE (2.1) has a unique strong solution (the reader is referred to [18], p 194)

$$X = \Phi(W),$$

where $\Phi : W^d \to W^d$. Indeed Φ is continuous on W^d and for each $\gamma \in W^d$, $\eta = \Phi(\gamma)$ is the unique solution of the ordinary differential equation

$$\eta_t = x + \gamma_t + \int_0^t b(\eta_s)ds.$$

We start with the LSI (1.2) of Gross. If $f\circ\Phi \in \mathbb{D}_2(D)$ (the domain of D in $L^2(W^d,\mu)$ as defined in the Malliavin calculus), we have

$$Ent_{\mathbb{P}_x} f^2 = Ent_\mu[(f\circ\Phi)^2] \leq 2\mathbb{E}^\mu\|D(f\circ\Phi)\|_H^2. \tag{2.5}$$

Now for every $h \in H$, we shall prove that

$$A(\gamma)h := \lim_{\varepsilon \to 0} \frac{\Phi(\gamma + \varepsilon h) - \Phi(\gamma)}{\varepsilon} \tag{2.6}$$

holds in the sup-norm of W^d (then in that of L^2). Thus

$$\Phi(\gamma + \varepsilon h) = \Phi(\gamma) + \varepsilon \frac{d\Phi(\gamma + \varepsilon h)}{d\varepsilon}|_{\varepsilon=0} + \circ(\varepsilon) = \Phi(\gamma) + \varepsilon A(\gamma)h + \circ(\varepsilon)$$

and consequently

$$\begin{aligned} D_h(f \circ \Phi)(\gamma) &= \lim_{\varepsilon \to 0} \frac{f\big(\Phi(\gamma + \varepsilon h)\big) - f\big(\Phi(\gamma)\big)}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \frac{f\big(\Phi(\gamma) + \varepsilon A(\gamma)h + \circ(\varepsilon)\big) - f\big(\Phi(\gamma)\big)}{\varepsilon} \\ &= (D_{A(\gamma)h} f)\big(\Phi(\gamma)\big). \end{aligned}$$

If we can prove that

$$\|A(\gamma)h\|_{L^2} \leq \sqrt{C(T)}\|h\|_H \tag{2.7}$$

where $C(T)$ is given in the theorem, we shall get

$$\begin{aligned} \|D(f \circ \Phi)(\gamma)\|_H &= \sup_{\|h\|_H \leq 1} |D_h(f \circ \Phi)(\gamma)| \\ &= \sup_{h \in H\ ,\ \|h\|_H \leq 1} |(D_{A(\gamma)h} f)\big(\Phi(\gamma)\big)| \\ &= \sup_{h \in H\ ,\ \|h\|_H \leq 1} \langle (\nabla f)\big(\Phi(\gamma)\big)\ ,\ A(\gamma)h \rangle_{L^2} \\ &\leq \sup_{h \in H\ ,\ \|h\|_H \leq 1} \|(\nabla f)\big(\Phi(\gamma)\big)\|_{L^2} \|A(\gamma)h\|_{L^2} \\ &\leq \sqrt{C(T)}\|(\nabla f)\big(\Phi(\gamma)\big)\|_{L^2}. \end{aligned}$$

Thus by (2.5), we obtain

$$\begin{aligned} Ent_{\mathbb{P}_x} f^2 &\leq 2\mathbb{E}^\mu \|D(f \circ \Phi)\|_H^2 \\ &\leq 2C(T)\mathbb{E}^\mu \|(\nabla f) \circ \Phi\big)\|_{L^2}^2 \\ &= 2C(T)\mathbb{E}^{\mathbb{P}_x} \int_0^T |\nabla_t f|^2 dt \end{aligned}$$

the desired inequality.

2.3.2 Proof of Theorem 2.3

To make the above proof rigorous, it remains to verify three points: (2.6), (2.7) and $f \circ \Phi \in \mathbb{D}_2(D)$. The last point is a consequence of (2.6), (2.7), because (2.6), (2.7) imply

$$|D_h(f \circ \Phi)(\gamma)| \leq \sqrt{C(T)}\|h\|_H \|(\nabla f)(\Phi(\gamma))\|_{L^2}$$

for all $h \in H$ (this is a well known fact in the Malliavin calculus, cf. Nualart [19]). (2.6) is an elementary fact in the ODE theory, so omitted. We now prove (2.7).

Since $\eta^\epsilon = \Phi(\gamma + \varepsilon h)$ verifies

$$\eta_t^\epsilon = x + \gamma_t + \varepsilon h(t) + \int_0^t b\big(\eta_s^\varepsilon\big) ds$$

then, $g := A(\gamma)h = \frac{d}{d\varepsilon}\eta^\varepsilon$ satisfies the linear differential equation :

$$g(t) = h(t) + \int_0^t \nabla b\big(\eta_s\big) g(s) ds$$

where $\eta = \Phi(\gamma)$. This equation has a unique solution and it can be given explicitly by

$$g(t) = \int_0^t A_t {A_s}^{-1} \dot{h}(s) ds$$

where A_t satisfies :

$$A_0 = Id \ , \ \ \frac{d}{dt} A_t = \nabla b(\eta_t) A_t.$$

Under the condition (2.2), we obtain using Lemma 2.4 :

$$|A_t {A_s}^{-1} x| \le e^{-K(t-s)} |x|, \ \forall 0 \le s \le t.$$

We state now the desired inequality (2.7) as

Lemma 2.6. *We have*

$$\|A(\gamma)h\|_{L^2}^2 = \|g\|_{L^2}^2 \le C(T) \|h\|_H^2$$

where $C(T)$ is given in Theorem 2.3.

Proof. From the following expression $g(t) = \int_0^t A_t A_s^{-1} \dot{h}(s) ds$, we can compute :

$$\begin{aligned} \|g\|_{L^2}^2 &= \int_0^T \left| \int_0^t A_t A_s^{-1} \dot{h}(s) ds \right|^2 dt \\ &\le \int_0^T \left(\int_0^t e^{-K(t-s)} |\dot{h}(s)| ds \right)^2 dt \end{aligned}$$

Now, we have to control the last term. A Cauchy-Schwarz control is easy, but the constant so obtained explodes. In order to avoid this explosion, let us write

$$\begin{aligned} &\int_0^T \left(\int_0^t e^{-K(t-s)} |\dot{h}(s)| ds \right)^2 dt \\ = &\int_0^T \left(\int_0^t e^{-K(t-u)} |\dot{h}(u)| du \right) \left(\int_0^t e^{-K(t-v)} |\dot{h}(v)| dv \right) dt \end{aligned}$$

$$
\begin{aligned}
&= \int_0^T \left(\iint_{0\le u,v\le t} e^{-K(t-u)}|\dot{h}(u)|e^{-K(t-v)}|\dot{h}(v)|du\ dv \right) dt \\
&= \iint_{0\le u,v\le T} \left(\int_{u\vee v}^T e^{-2Kt+K(u+v)}dt \right) |\dot{h}(u)||\dot{h}(v)|du\ dv.
\end{aligned}
$$

Since $2(u \vee v) - (u+v) = |u-v|$, we have if $K \neq 0$,

$$
\begin{aligned}
\Gamma(u,v) := \int_{u\vee v}^T e^{-2Kt+K(u+v)}dt &= \frac{e^{K(u+v)}}{2K}\left(e^{-2K(u\vee v)} - e^{-2KT}\right) \\
&= \frac{e^{-K|u-v|} - e^{-K(2T-u-v)}}{2K}
\end{aligned}
$$

and if $K = 0$,

$$
\Gamma(u,v) := \int_{u\vee v}^T e^{-2Kt+K(u+v)}dt = (T - u \vee v).
$$

Then we obtain :

$$
\|g\|_{L^2}^2 \ \le \ \int_0^T \int_0^T \Gamma(u,v)|\dot{h}(u)||\dot{h}(v)|dudv. \tag{2.8}
$$

Define an operator $\Gamma : L^2(0,T) \to L^2(0,T)$ by :

$$
\Gamma f(v) = \int_0^T \Gamma(u,v) f(u) du.
$$

It is a self-adjoint operator on $L^2(0,T)$. Then the term to control becomes :

$$
\langle \Gamma|\dot{h}|\ ,\ |\dot{h}|\rangle_{L^2}.
$$

But it is clear that (since $\Gamma(u,v) \ge 0$):

$$
\begin{aligned}
\|\Gamma f\|_{L^1} &= \int_0^T \left| \int_0^T \Gamma(u,v) f(u) du \right| dv \\
&\le \left(\sup_{v\in[0,T]} \int_0^T \Gamma(u,v) du \right) \int_0^T |f(v)|dv.
\end{aligned}
$$

Thus Lemma 2.5 yields $\|\Gamma\|_{L^2} \le A(T)$ where

$$
A(T) = \sup_{v\in[0,T]} \int_0^T \Gamma(u,v) du.
$$

Finally by an elementary calculus, we get $A(T) = C(T)$ given in Theorem 2.3.

□

Remarks 2.7. In the proof above, we see that $(\gamma, h) \to A(\gamma)h = g$ is continuous from $W^d \times H$ to $L^2([0,T], \mathbb{R}^d)$. Thus for all $f \in C_b^1(W^d/L^2)$, $f \circ \Phi$ is continuously differentiable on W^d w.r.t. the norm of H, and we have shown

$$\|D(f \circ \Phi)(\gamma)\|_H \leq \sqrt{C(T)} \|\nabla f(\Phi(\gamma))\|_H.$$

Now let $Q \in M_d(\mathbb{R})$ and F a continuously differentiable function on W^d w.r.t. the norm of H, applying the Gross theorem to $F(Q\gamma)$, we get

$$Ent(F^2(QW.)) \leq 2\lambda_{max}(Q^tQ)\mathbb{E} \int_0^T |\dot{\widehat{DF}}|^2(QW., t)dt$$

where $\lambda_{max}(Q^tQ)$ is the maximal eigenvalue of Q^tQ. Thus by the estimate above, we see that

$$dX_t = b(X_t)dt + QdW_t, \; X_0 = x$$

satisfies : $\forall f \in C_b^1(W^d/L^2)$,

$$Ent(f^2(X)) \leq 2C(T)\lambda_{max}(Q^tQ)\mathbb{E} \int_0^T |\nabla_t f|^2(X)dt.$$

In other words our approach here is also well adapted for eventually degenerate noise QdW_t.

3 Brownian motion on manifold

When one studies the diffusion $dX_t = \sigma(X_t)dW_t + b(X_t)dt$ with a non-constant volatility coefficient σ, the application Φ mapping the path of the BM to the path of X does not satisfy : $\|\nabla_h \Phi\|_{L^2} \leq \sqrt{C(T)}\|h\|_H$. So the approach in the previous section loses its pertinence. Instead, we shall explore another method, based on the martingale representation formula, as developed in [12] and [8]. We shall study only the elliptic diffusion on a Riemannian manifold generated by $\frac{1}{2}\Delta + V$.

To illustrate clearly the idea, we begin with the BM over a Riemannian manifold with bounded Ricci curvature. In that case, using the martingale representation formula, Fang [12] obtained the Poincaré inequality, Hsu [14, 15] and Capitaine-Hsu-Ledoux [8] obtained the logarithmic Sobolev inequality on the path space, both w.r.t. the Cameron-Martin metric. The constant of the LSI in those works explodes for large time t. Aida [1] obtained the LSI w.r.t. the Cameron-Martin metric with a bounded constant for large time T once the Ricci curvature is bounded from below by a positive constant. But as explained in the Introduction the LSI w.r.t. the Cameron-Martin metric even with a non-explosive constant does not produce concentration inequality of correct order (in large time T) for functionals such as $F(X) = \int_0^T g(X_s)ds$.

3.1 Martingale representation and gradient w.r.t. L^2-norm

We follow the exposition of [8]. Let M be a complete and connected manifold of dimension d equipped with the Levi-Civita connection ∇. Denote by Δ the Laplace-Beltrami operator on M, $O(M)$ the bundle of orthonormal frames, and $\Pi : O(M) \to M$ the canonical projection. Fix a point $x_0 \in M$. Each frame $u \in O(M)$ is a linear isometry $u : \mathbb{R}^d \to T_{\Pi(u)}(M)$. We assume throughout this section the boundedness of the Ricci curvature

$$\sup_{u \in O(M)} \|Ric_u\| < +\infty, \tag{3.1}$$

and denote also Ric_u its scalarization. Let $W_{x_0}(M)$ be the space of continuous paths from $[0,T]$ to M starting at x_0. Let U be the horizontal Brownian motion. We recall that U is solution of the Stratonovich SDE :

$$dU_t = \sum_{i=1}^{d} H_i(U_t) \circ dW_t^i \ , \ U_0 = u_0 \in O(M) \tag{3.2}$$

where W_t is a standard Brownian motion on $\mathbb{R}^d$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $H_i, 1 \leq i \leq n$ are the canonical horizontal vector fields on $O(M)$. Details about this construction can be found in [14] and [15]. Then $X = \Pi(U)$ is the M-valued Brownian motion starting at x_0, whose law μ_M is the Wiener measure on $W_{x_0}(M)$. Hence, for a Brownian path X, and h in the Cameron Martin space, $U_t h_t$ is a tangent vector at X_t. It determines a vector field D_h by $D_h X_t = U_t h_t$.

For smooth cylindrical functions f on $W_{x_0}(M)$, that is $f(\gamma) = f(\gamma_{t_1}, \ldots, \gamma_{t_n})$ with $\gamma \in W_{x_0}(M)$, the Malliavin derivative operator is defined by :

$$Df(X,t) = \sum_{i=1}^{n} U_{t_i}^{-1} \frac{df}{dx_i}(X_{t_1}, \ldots, X_{t_n}) t_i \wedge t.$$

Notice that from the path of the BM X over the Riemannian manifold and $U_0 = u_0$, one can reconstruct (U_t) and (W_t) a.s. (w.r.t. the law of X). Thus the above definition is intrinsic (cf. [14]).

Let $(D, \mathbb{D}_2(D))$ be its closure in $L^2(\mu_M)$. For $h \in H$, $D_h f(\gamma)$ and $Df(\gamma) \in H$ are such that :

$$\forall h \in H,\ D_h f(\gamma) = \langle Df(\gamma)\ ,\ h \rangle_H = \int_0^T \dot{\widehat{D}} f(\gamma, s) \cdot \dot{h}(s) ds.$$

Let $F : W_{x_0}(M) \to \mathbb{R}$ be a real function in the domain $\mathbb{D}_2(D)$ of D. Under the boundedness condition (3.1), the Fang's version of Clark-Ocone-Haussmann formula for the Brownian motion on a manifold is read as

$$F(X) = \mathbb{E}\big(F(X)\big) + \int_0^T \langle H_t\ , dW_t \rangle \tag{3.3}$$

where

$$H_t = \mathbb{E}^{\mu_M} \left(\dot{\widehat{D}} F(t) - \frac{1}{2} A_t^* \int_t^T (A_s^*)^{-1} Ric_{U_s} \dot{\widehat{D}} F(s) ds | \mathcal{B}_t \right). \tag{3.4}$$

In this expression, $(\mathcal{B}_t)$ is the filtration generated by (W_t), $(A_t)_{0\leq t\leq T}$ is a matrix valued process defined by :

$$\frac{dA_t}{dt} - \frac{1}{2}A_t Ric_{U_t} = 0 \quad , \quad A_0 = Id. \tag{3.5}$$

The above formula relies on the Bismut-Driver integration by parts formula : for any $h \in H$, the adjoint D_h^* of D_h in $L^2(\mu_M)$ is given by

$$D_h^* = -D_h + \int_0^T \langle \dot{h}_t + \frac{1}{2}Ric_{U_t}h_t, dW_t\rangle. \tag{3.6}$$

According to Lemma 2.2, we introduce our gradient w.r.t. the L^2-metric.

Definition 3.1. *For $F \in \mathbb{D}_2(D)$ such that $\ddot{\overset{\frown}{D}}F \in L^2([0,T],\mathbb{R}^d)$ we define the gradient of F w.r.t. L^2-norm by*

$$\nabla_t F(\gamma) = -\ddot{\overset{\frown}{D}}F(\gamma,t) \quad dt - a.e.$$

In other words, $\nabla F(\gamma) \in L^2([0,T],\mathbb{R}^d)$ is determined by

$$\langle \nabla F(\gamma) \ , \ h\rangle_{L^2} \ = \ D_h F(\gamma) \ , \ \forall h \in C_0^\infty(]0,T[,\mathbb{R}^d).$$

We assume now

$$\forall u \in O(M) \ , \ \frac{1}{2}Ric_u \geq KId, \ K \in \mathbb{R}. \tag{3.7}$$

The constant of our LSI will depend only on this lower bound on the Ricci curvature.

Lemma 3.2. *Let A solution of (3.5). We have*

$$\forall s \geq t \ , \ \ \|A_t^* A_s^{-1^*}\| \leq \exp\bigl(-K(s-t)\bigr).$$

Proof. With a derivation of $A_t A_t^{-1} = Id$, we obtain

$$\frac{d}{ds}(A_s^{-1}A_t) = -\frac{1}{2}Ric_{U_s}A_s^{-1}A_t,$$

which implies the desired result using Lemma 2.4 and (3.7). □

The key observation is:

Lemma 3.3. *For $F \in \mathbb{D}_2(D)$ such that $\mu_M - a.s$, $\ddot{\overset{\frown}{D}}F(\gamma) \in L^2([0,T],\mathbb{R}^d)$ (then $\dot{\overset{\frown}{D}}F(\gamma)$ can and will be chosen to be absolutely continuous) and $\dot{\overset{\frown}{D}}F(\gamma,T) = 0$, in the representation formula (3.3), we have*

$$H_t = \mathbb{E}^{\mu_M}\left(\int_t^T A_t^* A_s^{*-1}\nabla_s F ds | \mathcal{B}_t\right).$$

Proof. Taking the adjoint in the proof of Lemma 3.2 gives

$$\frac{d}{ds}(A_t^* A_s^{*-1}) = -\frac{1}{2} A_t^* A_s^{*-1} Ric_{U_s}.$$

By integration by parts in the expression (3.4) of H_t, we have $\mu_M - a.s.$,

$$\begin{aligned} -\frac{1}{2} A_t^* \int_t^T (A_s^*)^{-1} Ric_{U_s} \dot{\widehat{DF}}(s) ds &= \int_t^T \frac{d}{ds}(A_t^* A_s^{*-1}) \dot{\widehat{DF}}(s) ds \\ &= \left[A_t^* A_s^{*-1} \dot{\widehat{DF}}(s) \right]_{s=t}^{s=T} - \int_t^T A_t^* A_s^{*-1} \ddot{\widehat{DF}}(s) ds \\ &= -\dot{\widehat{DF}}(t) + \int_t^T A_t^* A_s^{*-1} \nabla_s F ds. \end{aligned}$$

Then the desired result follows by (3.4). ☐

3.2 Poincaré inequality for the L^2 norm

Now, we follow Fang [12]. For each $F \in \mathbb{D}_2(D)$ such that $\mu_M - a.s$, $\ddot{\widehat{DF}}(\gamma) \in L^2([0,T], \mathbb{R}^d)$ and $\dot{\widehat{DF}}(\gamma, T) = 0$, we have by Lemma 3.3,

$$\begin{aligned} Var_{\mu_M}(F) &= \mathbb{E}\left(F(X) - \mathbb{E}F(X)\right)^2 = \mathbb{E}^{\mu_M} \int_0^T |H_t|^2 dt \\ &= \mathbb{E} \int_0^T \left| \mathbb{E}\left[\int_t^T A_t^* A_s^{*-1} \nabla_s F(X) ds | \mathcal{B}_t \right] \right|^2 dt \\ &\leq \mathbb{E} \int_0^T \left[\int_t^T \|A_t^* A_s^{*-1}\| \, |(\nabla_s F)(X)| ds \right]^2 dt \\ &\leq \mathbb{E}^{\mu_M} \int_0^T \left[\int_t^T e^{-K(s-t)} |\nabla_s F| ds \right]^2 dt \end{aligned}$$

where the last inequality is given by Lemma 3.2. We write the last quantity as (by Fubini)

$$\begin{aligned} &\mathbb{E}^{\mu_M} \int_0^T \int_t^T e^{-K(u-t)} |\nabla_u F| \int_t^T e^{-K(v-t)} |\nabla_v F| du dv dt \\ &\quad = \mathbb{E}^{\mu_M} \int_0^T \int_0^T du dv |\nabla_u F| \cdot |\nabla_v F| \int_0^{u \wedge v} e^{-K(u-t)-K(v-t)} dt. \end{aligned}$$

Let $\Gamma(u,v) := \int_0^{u\wedge v} e^{-K(u-t)-K(v-t)} dt$ and $\Gamma f(v) := \int_0^T \Gamma(u,v) f(u) du$. We have as in Lemma 2.6

$$\|\Gamma\|_{2,2} \leq \|\Gamma\|_{1,1} = \sup_{u \in [0,T]} \int_0^T \Gamma(u,v) dv \leq C(T)$$

where $C(T)$ is given in Theorem 2.3. Thus we get,

$$Var_{\mu_M}(F) \leq \|\Gamma\|_{2,2} \mathbb{E}^{\mu_M} \int_0^T |\nabla_t F|^2 dt \leq C(T) \mathbb{E}^{\mu_M} \int_0^T |\nabla_t F|^2 dt,$$

i.e., we have shown

Theorem 3.4. *Assume (3.1) and (3.7). For each* $F \in \mathbb{D}_2(D)$ *such that* $\mu_M - a.s$, $\overset{\ddot{\frown}}{D}F(\gamma) \in L^2([0,T],\mathbb{R}^d)$ *and* $\overset{\dot{\frown}}{D}F(\gamma,T) = 0$,

$$Var_{\mu_M}(F) \leq C(T)\mathbb{E}^{\mu_M}\int_0^T |\nabla_t F|^2 dt$$

where $C(T)$ *is given in Theorem 2.3.*

3.3 The logarithmic Sobolev inequality for L^2-norm

Theorem 3.5. *Assume (3.1) and (3.7). For each* $F \in \mathbb{D}_2(D)$ *such that* $\mu_M - a.s$, $\overset{\ddot{\frown}}{D}F(\gamma) \in L^2([0,T],\mathbb{R}^d)$ *and* $\overset{\dot{\frown}}{D}F(\gamma,T) = 0$,

$$Ent_{\mu_M}(F^2) \leq 2C(T)\mathbb{E}^{\mu_M}\int_0^T |\nabla_t F|^2 dt$$

where $C(T)$ *is given in Theorem 2.3.*

Proof. We follow the ingenious method of Capitaine-Hsu-Ledoux [8]. Assume at first

$$\frac{1}{\varepsilon} \geq F \geq \varepsilon > 0 \tag{3.8}$$

for some $\varepsilon \in (0,1)$. Consider the continuous martingale $M_s = \mathbb{E}\left[F^2(X) \mid \mathcal{B}_s\right]$, such that $M_0 = \mathbb{E}^{\mu_M}F^2$ and $M_T = F^2(X)$. Let

$$H_t = \mathbb{E}^{\mu_M}\left(\int_t^T A_t^* A_s^{*-1}\nabla_s(F^2)ds|\mathcal{B}_t\right).$$

Using the martingale representation in Lemma 3.3 and applying the Itô formula, we get after taking expectation

$$\mathbb{E}\left(M_T \log M_T\right) - \mathbb{E}\left(M_0 \log M_0\right) = \mathbb{E}\frac{1}{2}\int_0^T \frac{d\langle M\rangle_s}{M_s} = \mathbb{E}\frac{1}{2}\int_0^T \frac{|H_s|^2}{M_s}ds.$$

Since $\nabla_s(F^2) = 2F\nabla_s F$, $ds \otimes \mu_M - a.e.$, by the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
Ent_{\mu_M}(F^2) &= \frac{1}{2}\mathbb{E}\int_0^T \frac{\left|\mathbb{E}\left(\int_t^T A_t^* A_s^{*-1}\nabla_s(F^2)(X)ds \mid \mathcal{B}_t\right)\right|^2}{\mathbb{E}\left[F^2(X) \mid \mathcal{B}_t\right]}dt \\
&\leq 2\mathbb{E}\int_0^T \mathbb{E}\left[\left|\int_t^T A_t^* A_s^{*-1}\nabla_s F(X)ds\right|^2 \mid \mathcal{B}_t\right]dt \\
&\leq 2\mathbb{E}^{\mu_M}\int_0^T \left(\int_t^T \|A_t^* A_s^{*-1}\| \cdot |\nabla_s F|ds\right)^2 dt.
\end{aligned}$$

We can conclude by the proof of the Poincaré inequality.

We now remove the restriction (3.8). For general F, let

$$F_{\varepsilon,L} := f_{\varepsilon,L}(F), \text{ where } f_{\varepsilon,L} := L \arctan \frac{\sqrt{\varepsilon + x^2}}{L}.$$

It is easy to see that $F_{\varepsilon,L} \in \mathbb{D}_2(D)$ and

$$\dot{\overset{\frown}{D}} F_{\varepsilon,L}(t) = f'_{\varepsilon,L}(F) \dot{\overset{\frown}{D}} F(t), \ \nabla_t F_{\varepsilon,L} = f'_{\varepsilon,L}(F) \nabla_t F.$$

Then $F_{\varepsilon,L}$ satisfies again the assumption of the theorem and especially (3.8), so the LSI is true for $F_{\varepsilon,L}$. Moreover by the expression above we also have

$$\mathbb{E}^{\mu_M} \int_0^T |\nabla_t F_{\varepsilon,L}|^2 dt \leq \mathbb{E}^{\mu_M} \int_0^T |\nabla_t F|^2 dt.$$

By letting $L \to +\infty$ first $F^2_{\varepsilon,L} \uparrow \varepsilon + F^2$, and $\varepsilon \to 0+$ next, we have $Ent_{\mathbb{P}}\left(F^2_{\varepsilon,L}(X)\right) \to Ent_{\mathbb{P}}\left(F^2(X)\right)$, so the desired LSI for F follows. □

Notice that the smooth cylindrical functions $f(\gamma_{t_1}, \cdots, \gamma_{t_n})$ does not satisfy the assumptions of Theorems 3.4 and 3.5 (this is quite natural as for t fixed, γ_t has no meaning in L^2). However $F(\gamma) := \int_0^T f(t, \gamma_t) dt$ with bounded and smooth $f : [0, T] \times M \to \mathbb{R}$ satisfies the conditions in Theorem 3.4 and Theorem 3.5, because it is easy to prove that $F \in \mathbb{D}_2(D)$ and

$$\dot{\overset{\frown}{D}} F(t, X) = \int_t^T U_s^{-1} \nabla_x f(s, X_s) ds, \ \nabla_t F(X) = U_t^{-1} \nabla_x f(t, X_t)$$

where $\nabla_x f(t, x)$ is the gradient on the space variable $x \in M$.

Before moving to our true object, let us make some comments on the condition (3.1) of the boundedness of the Ricci curvature, whereas the final results in this section seem not to depend on it "in appearance".

Remarks 3.6. *If the Ricci curvature is only lower bounded but not bounded, we have several technical questions related with the stochastic calculus used in this section:*

(a) The Bismut-Driver integration by parts formula (3.6) has a question with the unbounded term $Ric_{U_t} h_t$.

(b) Without the integration by parts formula (3.6), we have neither the Fang's martingale representation formula (which also contains an unbounded term), nor the closability of D in $L^2(\mu_M)$ (more exactly we do not know).

The last point is particularly annoying, since even $\mathbb{D}_2(D)$ is no longer well defined.

Now let us see how to bypass those delicate difficulties for general diffusions generated by $\frac{1}{2}\Delta + V$, whose Bakry-Emery's curvature is only assumed to be lower bounded.

4 Brownian motion with drift over a Riemannian manifold

In this section we extend the preceding logarithmic Sobolev inequality to the law of a diffusion process (X_t) with generator $L = \frac{1}{2}\Delta + V$ over a complete connected Riemannian manifold M with lower bounded Bakry-Emery curvature, i.e.,

$$\frac{1}{2}Ric_u - (\overline{\nabla V})^\sigma \geq KId, \ K \in \mathbb{R}. \tag{4.1}$$

Here V is a smooth vector field and $(\nabla V)^\sigma$ is the symmetrized gradient. Recall that Wang [22] has obtained the Talagrand "T_2"-transportation inequality.

We assume that the BM is non-explosive. In that situation, the diffusion with generator $L = \frac{1}{2}\Delta + V$ is non-explosive (under (4.1)) and the law μ^V of the diffusion is absolutely continuous w.r.t. the law of the BM μ_M.

4.1 A space of test-functions $\mathcal{D}$

We introduce now a space $\mathcal{D}$ of test-functions which is rich enough for applications in the concentration phenomena of the diffusion for large time. It allows us to circumvent the technical problems related to the fact mentioned in Remarks 3.6 that $\mathbb{D}_2(D)$ is ill-posed.

Definition 4.1. *Let*

$$\mathcal{D} = \left\{ \ F(\gamma) \ = \ \Phi\left(\int_0^T f_1(s,\gamma_s)ds \ , \ \ldots \ , \ \int_0^T f_m(s,\gamma_s)ds\right) \ \right\}$$

where $m \in \mathbb{N}^*$, $\Phi \in \mathcal{C}_b^\infty(\mathbb{R}^m, \mathbb{R})$ and $f_i \in \mathcal{C}_0^\infty([0,T] \times M)$.

We define the gradient ∇ w.r.t. L^2-norm by

$$\nabla_t \int_0^T f(s,\gamma_s)ds = U_t^{-1}\nabla_x f(t,\gamma_t)$$

and the fact that it obeys the rule of composition:

$$\nabla_t \Phi\left(\int_0^T f_1(s,\gamma_s)ds \ , \ \ldots \ , \ \int_0^T f_m(s,\gamma_s)ds\right) := \sum_i \partial_i \Phi(\cdots) U_t^{-1}\nabla_x f_i(t,\gamma_t).$$

Here $\nabla_x f(t,x)$ is the gradient w.r.t. the space variable $x \in M$.

Notice that the stochastic parallel transport $U_t = U_t(\gamma)$ is $\mu_M - a.s.$-well defined, then $\mu^V - a.s.$ well defined.

4.2 Martingale representation for test-functions in $\mathcal{D}$

For $u \in O(M)$, the bundle of orthonormal frames, let $\overline{V}(u) = u^{-1}V\left(\Pi(u)\right) \in \mathbb{R}^d$, and let $\overline{\nabla V}(u) = u^{-1}\nabla V\left(\Pi(u)\right) \in \mathbb{R}^d \otimes \mathbb{R}^d$. They are the scalarization of the tensors V and ∇V respectively.

Recall that our diffusion (X_t) with $X_0 = x_0$ can be constructed as follows: $X_t = \Pi(U_t)$ where (U_t) is the $O(M)$-valued diffusion, solution of

$$dU_t = \sum_{i=1}^{n} H_i(U_t) \circ dW_t^i + \tilde{V}(U_t)dt \tag{4.2}$$

where $\tilde{V}$ is the horizontal lift of V, defined on $(\Omega, (\mathcal{B}_t), \mathbb{P})$ and $(\mathcal{B}_t)$ is the filtration of the $\mathbb{R}^d$-valued Brownian motion (W_t). Notice that solution U_t of (4.2) coincides in law with $U_t(X)$ where this last U_t is used in the definition of $\nabla_t F$.

Lemma 4.2. *(due to [8]) Assume that Ric_u and ∇V are both bounded. Then for any smooth cylindrical functional $F = F(\gamma_{t_1}, \cdots, \gamma_{t_n})$ on $W_{x_0}(M)$,*

$$F(X) = \mathbb{E}F(X) + \int_0^T \langle H_t, dW_t \rangle, \tag{4.3}$$

where

$$\begin{aligned} H_t &= \mathbb{E}^{\mu_V}\left(\dot{\widehat{D}}F(t) - \int_t^T A_t^*(A_s^*)^{-1}M_s^*\dot{\widehat{D}}F(s)ds|\mathcal{B}_t\right) \\ M_t &= \frac{1}{2}Ric_{U_t} - \overline{\nabla V}(U_t) \\ A_0 &= I,\ \frac{dA_t}{dt} = A_t M_t. \end{aligned} \tag{4.4}$$

In the actual bounded case, it is easy to see that $F \in \mathcal{D}$ satisfies again (4.3). By integration by parts, we obtain as in Lemma 3.3,

Lemma 4.3. *Assume that Ric_u and ∇V are both bounded. Then for any $F \in \mathcal{D}$,*

$$F(X) = \mathbb{E}F(X) + \int_0^T \langle \mathbb{E}\left(\int_t^T A_t^*(A_s^*)^{-1}\nabla_s F(X)|\mathcal{B}_t\right), dW_t \rangle,$$

where (A_t) is given in Lemma 4.2.

Now we remove the boundedness condition.

Theorem 4.4. *Assume (4.1). Then the conclusion of Lemma 4.3 remains true.*

As corollaries of the above martingale representation, we obtain by repeating the arguments in §3:

Theorem 4.5. *For any $F \in \mathcal{D}$,*

$$Var_{\mu^V}(F) \leq C(T)\mathbb{E}^{\mu_V}\int_0^T |\nabla_t F|^2 dt, \tag{4.5}$$

$$Ent_{\mu^V}(F^2) \leq 2C(T)\mathbb{E}^{\mu_V}\int_0^T |\nabla_t F|^2 dt, \tag{4.6}$$

where $C(T)$ is given in Theorem 2.3.

4.3 Proof of Theorem 4.4: technique of localization

If M is compact, this result is contained in Lemma 4.3. Assume below that M is non-compact.

Step 1: localization. Let $F = \Phi\left(\int_0^T f_1(s,\gamma_s)ds\ ,\ \dots\ ,\ \int_0^T f_m(s,\gamma_s)ds\right) \in \mathcal{D}$ and fix some compact $K \subset M$ such that the supports of $f_i,\ i = 1, \cdots, m$ are all contained in $[0,T] \times K$. Take next a sequence of compact subsets $(K_n)_{n\geq 1}$ of M such that

$$K \subset K_1^o(\text{ interior of } K_1),\ K_n \subset K_{n+1}^o,\ \bigcup_n K_n = M.$$

For each $n \geq 1$, choose

(1) a complete connected Riemannian manifold $M_n \supset K_n$ such that the Riemannian metric of M_n restricted on K_n coincides with the original one of M and the Ricci curvature of M_n is globally bounded;

(2) a smooth vector field V_n on M_n such that $V_n(x) = V(x)$ for $x \in K_n$ and $V_n, \nabla V_n$ are globally bounded;

(3) $(1/2)Ric - \overline{\nabla V_n} \geq (K-1)Id$ over M_n for all n.

On the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with the $\mathbb{R}^d$-valued Brownian motion W_t, let (U_t^n) be the solution of

$$dU_t^n = \sum_{i=1}^d H_i^n(U_t^n) \circ dW_t^i + \tilde{V}_n(U_t^n)dt$$

and $X_t^n = \Pi(U_t^n)$, where $H_i^n, i = 1, \cdots, d$ are the canonical horizontal vector fields on $O(M_n)$, $\tilde{V}_n$ is the horizontal lift of V_n.

Consider the stopping time

$$\tau_n := \inf\{t \geq 0; X_t \notin K_n\}$$

that satisfies $\mathbb{P}(\tau_n \uparrow +\infty) = 1$ by the non-explosion of our diffusion under condition (4.1). Then by the uniqueness of the SDE, we have with probability one,

$$U_t^n = U_t,\ X_t^n = X_t,\ \forall 0 \leq t \leq \tau_n.$$

Regarding f_i as function on $[0,T] \times M_n$ with the convention that $f(t,x) = 0$ for $x \notin K$ and $F(\gamma)$ as a functional on $W_{x_0}(M_n)$, we can apply Lemma 4.2 to get

$$F(X^n) - \mathbb{E}F(X^n) = \int_0^T \langle \mathbb{E}\left(\int_t^T (A_t^{(n)})^*[(A_s^{(n)})^*]^{-1}\nabla_s^{(n)}F(X^n)|\mathcal{B}_t\right), dW_t\rangle \qquad (4.7)$$

where $\nabla_s^{(n)}F$ is the gradient w.r.t. L^2-norm on $W_{x_0}(M_n)$, and

$$M_t^n = \frac{1}{2}Ric_{U_t^n} - \overline{\nabla V_n}(U_t^n)$$

$$A_0^{(n)} = I,\ \frac{dA_t^{(n)}}{dt} = A_t^{(n)}M_t^n.$$

Step 2: convergence in (4.7).

At first, for the left-hand side , let us make a last localization on the functional F :

$$F_n(X) := \Phi\left(\int_0^{T\wedge\tau_n} f_1(s,X_s)ds\ ,\ \dots\ ,\ \int_0^{T\wedge\tau_n} f_m(s,X_s)ds\right).$$

It is clear that $\mathbb{P}(\tau_n \uparrow +\infty) = 1$ implies $F_n(X) \to F(X)$ a.s., remarking that

$$|F_n(X) - F(X)| \le \sum_{i=1}^m \|\partial_i\Phi\|_\infty \|f_i\|_\infty (T - T\wedge\tau_n) \to 0.$$

Since we have $F_n(X) = F_n(X^n)$, we obtain that $F(X^n) \to F(X)$ a.s. and consequently in $L^p(\mathbb{P})$ for all $p \in [1,+\infty)$. Thus the left-hand side of (4.7) converges to $F(X) - \mathbb{E}F(X)$ as desired.

For the right-hand side of (4.7), denoting $\mathbb{E}(\cdot|\mathcal{B}_t)$ by $\mathbb{E}_{\mathcal{B}_t}$, we have :

$$\begin{aligned}
& \int_0^T \langle \mathbb{E}_{\mathcal{B}_t}\left(\int_t^T (A_t^{(n)})^*[(A_s^{(n)})^*]^{-1}\nabla_s^{(n)}F(X^n)ds\right), dW_t\rangle \\
= \ & \int_0^T \langle \mathbb{E}_{\mathcal{B}_t}\left(\int_t^T (A_t)^*[(A_s)^*]^{-1}\nabla_s F(X)\ 1_{T<\tau_n}\right), dW_t\rangle \\
& + \int_0^T \langle \mathbb{E}_{\mathcal{B}_t}\left(\int_t^T (A_t^{(n)})^*[(A_s^{(n)})^*]^{-1}\nabla_s^{(n)}F(X^n)ds\ 1_{T\ge\tau_n}\right), dW_t\rangle \\
:= \ & (I)_n + (II)_n.
\end{aligned} \tag{4.8}$$

Letting $n \to +\infty$, we have obviously in $L^2(\mathbb{P})$,

$$(I)_n \to (I) := \int_0^T \langle \mathbb{E}_{\mathcal{B}_t}\left(\int_t^T (A_t)^*[(A_s)^*]^{-1}\nabla_s F(X)ds\right), dW_t\rangle$$

the desired r.h.s. Indeed, we have by the same proof as Lemma 3.2,

$$\|{A_t}^*[(A_s)^*]^{-1}\| \le e^{-K(s-t)},\ \forall s \ge t$$

and for the functionals in $\mathcal{D}$

$$|\nabla_s F(\gamma)| \le \sum_{i=1}^m \|\partial_i\Phi\|_\infty \|\nabla f_i\|_\infty =: C,$$

so that

$$\begin{aligned}
\mathbb{E}|(I)_n - (I)|^2 &= \mathbb{E}\int_0^T \left|\int_t^T (A_t)^*[(A_s)^*]^{-1}\nabla_s F(X)ds\ 1_{T\ge\tau_n}\right|^2 dt \\
&\le \int_0^T \left|\int_t^T Ce^{-(K)(s-t)}ds\right|^2 dt\ \mathbb{P}(\tau_n \le T) \\
&\le C^2T^3e^{2|K|T}\ \mathbb{P}(\tau_n \le T) \to 0
\end{aligned}$$

We proceed in the same way for $(II)_n$, since by our assumption (3) in Step 1 and the proof of Lemma 3.2,

$$\|A_t^{(n)^*}[(A_s^{(n)})^*]^{-1}\| \le e^{-(K-1)(s-t)}, \ \forall s \ge t$$

and also

$$|\nabla_s^{(n)} F(\gamma)| \le \sum_{i=1}^m \|\partial_i \Phi\|_\infty \|\nabla f_i\|_\infty =: C$$

we have

$$\begin{aligned}\mathbb{E}(II)_n^2 &\le \mathbb{E} \int_0^T \left| \int_t^T (A_t^{(n)})^* [(A_s^{(n)})^*]^{-1} \nabla_s^{(n)} F(X^n) ds \ 1_{T \ge \tau_n} \right|^2 dt \\ &\le C^2 T^3 e^{2(|K|+1)T} \mathbb{P}(\tau_n \le T) \to 0.\end{aligned}$$

Consequently letting n go to infinity in (4.7) for $F(X^n)$, we get the desired martingale representation for $F(X)$. □

5 Some applications

Throughout this section let X be the diffusion generated by $\frac{1}{2}\Delta + V$ as given in §4 and the Bakry-Emery minoration condition (4.1) be satisfied with $K > 0$.

5.1 Hoeffding's type concentration implied by LSI

Let $g \in C_b^1([0,T] \times M)$ and consider the functional related to the central limit theorem:

$$F_T(\gamma) = \frac{1}{\sqrt{T}} \int_0^T (g(t,\gamma_t) - \mathbb{E}g(t,X_t))\, dt.$$

If $g \in C_0^\infty([0,T] \times M)$, then by the log-Sobolev inequality in Theorem 4.5 with $C(T) \le 1/K^2$ and by the Herbst method developed in Ledoux [17], we have

$$\mathbb{E}e^{\lambda F_T(X)} \le \exp\left(\frac{\lambda^2 \|\nabla_x g\|_\infty^2}{2K^2}\right), \ \forall \lambda \in \mathbb{R}.$$

Approximating $g \in C_b^1([0,1] \times M)$ by $g_n \in C_0^\infty([0,1] \times M)$, we see that the above inequality holds again for $g \in C_b^1([0,1] \times M)$. Thus by Chebychev inequality and an optimization over λ, we obtain in the standard way the following Hoeffding's type concentration inequality:

$$\mathbb{P}(F_T(X) > r) \vee \mathbb{P}(F_T(X) < -r) \le \exp\left(-\frac{r^2 K^2}{2\|\nabla_x g\|_\infty^2}\right), \ \forall r > 0. \tag{5.1}$$

5.2 Talagrand's T_2-transportation inequality and Tsirel'son's inequality

When treating the invariance principle (or functional central limit theorem) or functional moderate deviations ([9]), we have to study the functionals of type

$$F_{T,\delta}(X) := \frac{1}{\sqrt{T}} \sup_{0\le s\le t\le s+\delta\le 1} \int_{Ts}^{Tt} (g(u, X_u) - \mathbb{E}g(u, X_u))du \tag{5.2}$$

where $g \in C_b^1([0,T] \times M)$. The concentration inequality of $F_{T,\delta}$ could be deduced from

Corollary 5.1. *Let $g \in C_b^1([0,T] \times M \to \mathbb{R}^m)$ and $Y_t = g(t, X_t)$. Then Y satisfies the log-Sobolev inequality on $L^2([0,T], \mathbb{R}^m)$:*

$$Ent(F^2(Y)) \le 2\frac{\|\nabla_x g\|_\infty^2}{K^2} \mathbb{E} \int_0^T |\nabla_t F|^2(Y)dt \tag{5.3}$$

for all functions $F \in C_b^1(L^2([0,T], \mathbb{R}^m))$, where ∇ is the gradient on $L^2([0,T], \mathbb{R}^m)$. In particular,

(a) the law $\mathbb{P}_Y$ of Y satisfies on $L^2([0,T], \mathbb{R}^m)$ the Talagrand T_2-transportation inequality with constant $\frac{\|\nabla_x g\|_\infty^2}{K^2}$, i.e.,

$$W_2^2(\mathbb{Q}, \mathbb{P}_Y) \le 2\frac{\|\nabla_x g\|_\infty^2}{K^2} Ent\left(\frac{d\mathbb{Q}}{d\mathbb{P}_Y}\right)$$

for all probability measures $\mathbb{Q}$ on $L^2([0,T], \mathbb{R}^m)$ such that $\mathbb{Q} \ll \mathbb{P}_Y$, where

$$W_2(\mathbb{Q}, \mathbb{P}_Y) := \left(\inf_\pi \iint_{(L^2([0,T],\mathbb{R}^m))^2} \|h - \tilde{h}\|_{L^2([0,T],\mathbb{R}^m)}^2 \pi(dh, d\tilde{h})\right)^{1/2},$$

is the Wasserstein L^2-distance between $\mathbb{Q}$ and $\mathbb{P}_Y$, here the infimum is taken over all couplings π of $(\mathbb{Q}, \mathbb{P}_Y)$, i.e., over all the probability measures on $(L^2([0,T], \mathbb{R}^m))^2$ such that their marginal laws are respectively $\mathbb{Q}$ and $\mathbb{P}_Y$.

(b) (Tsirel'son's inequality) For any non-empty subset $A \subset L^2([0,T], \mathbb{R}^m)$,

$$G(Y) := \sup_{k\in A} \langle Y - \mathbb{E}Y, k\rangle_{L^2([0,T],\mathbb{R}^m)}$$

satisfies

$$\begin{aligned}&\mathbb{E}\exp\left(\frac{K^2}{\|\nabla_x g\|_\infty^2} \sup_{k\in A}\left[\langle Y - \mathbb{E}Y, k\rangle_{L^2([0,T],\mathbb{R}^m)} - \frac{1}{2}\|k\|_{L^2([0,T],\mathbb{R}^m)}^2\right]\right)\\ &\le \exp\left(\frac{K^2}{\|\nabla_x g\|_\infty^2}\mathbb{E}G(Y)\right).\end{aligned}$$

(c) In particular for $F_{T,\delta}(X)$ given in (5.2), we have

$$\mathbb{E}e^{\lambda F_{T,\delta}(X)} \le \exp\left(\lambda \mathbb{E}F_{T,\delta}(X) + \frac{\delta\lambda^2\|\nabla_x g\|_\infty^2}{2K^2}\right), \ \forall \lambda \in \mathbb{R};$$

$$\mathbb{P}\left(F_{T,\delta}(X) - \mathbb{E}F_{T,\delta}(X) > r\right) \le \exp\left(-\frac{r^2K^2}{2\delta\|\nabla_x g\|_\infty^2}\right), \ \forall r > 0.$$

Proof. At first for $F(h) = f(\langle h, e_1\rangle, \cdots, \langle h, e_n\rangle)$ where $e_i \in C^\infty([0,T], \mathbb{R}^m)$ and $f \in C_b^1(\mathbb{R}^n)$, then $F(Y) = \tilde{F}(X)$ where

$$\tilde{F}(\gamma) = f\left(\int_0^T e_1(t)g(t,\gamma_t)dt, \cdots, \int_0^T e_n(t)g(t,\gamma_t)dt\right)$$

belongs to our space $\mathcal{D}$ of test-functions if $g \in C_0^\infty([0,T] \times M)$. Noting that

$$\nabla_t \tilde{F}(X) = (\nabla_t F)(\cdots)U_t^{-1}\nabla_x g(t, X_t)$$

the log-Sobolev inequality (5.3) follows from Theorem 4.5. Approximating g by $g_n \in C_0^\infty([0,T] \times M)$, we obtain (5.3) for all such smooth cylindrical functions F.

Next for $F \in C_b^1(L^2([0,T], \mathbb{R}^m))$, let $(e_i)_{i\in\mathbb{N}} \subset C^\infty([0,T], \mathbb{R}^m)$ be an orthonormal basis of $L^2([0,T], \mathbb{R}^m)$ and $F_n(h) := F(P_n h)$ where P_n is the orthogonal projection to the subspace spanned by $(e_i)_{0\le i\le n}$. By the continuous differentiability of F, $\nabla F_n(h) \to \nabla F(h)$ in $L^2([0,T], \mathbb{R}^m)$ for each h. Hence we obtain (5.3) for such F by the dominated convergence.

Both parts (a) and (b) are consequences of the log-Sobolev inequality (5.3), as shown by Bobkov-Gentil-Ledoux [7] (certainly they established them only on $\mathbb{R}^n$, but their arguments work for the actual Hilbert space setting).

Finally in the case of part (c), $m = 1$, letting

$$A := \left\{\frac{\lambda C}{\sqrt{T}} 1_{[Ts,Tt]};\ 0 \le s \le t \le s + \delta \le 1\right\},\ C := \frac{\|\nabla_x g\|_\infty^2}{K^2}$$

we get part (c) from the Tsirel'son inequality in part (b). □

References

[1] S. Aida, Gradient estimates of harmonic functions and the asymptotics of spectral gaps on path spaces. *Interdiscip. Inform. Sci.* **2***, no. 1, 75-84 (1996).*

[2] S. Aida, T. Masuda, I. Shigekawa, Logarithmic Sobolev inequalities and exponential integrability. *J. Funct. Anal.* **126***, no. 1, 83-101 (1994).*

[3] S. Aida, D.Stroock, Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. *Math. Res. Lett.* **1***, no. 1, 75-86 (1994).*

[4] D. Bakry, *L'hypercontractivité et son utilisation en théorie des semigroupes,* Ecole d'Eté de Probabilités de Saint-Flour (1992), LNM 1581, Springer-Verlag 1994.

[5] D. Bakry, M. Emery. Diffusions hypercontractives. Seminaire de probabilités XIX, *Lecture Notes in Math, Volume* **1123**, *177-206 (1985).*

[6] S.G. Bobkov, F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163**, *1-28 (1999).*

[7] S. Bobkov, I. Gentil, M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pure Appl.,* **80**, *no. 7, 669-696 (2001).*

[8] M. Capitaine, E.P. Hsu and M. Ledoux. Martingale representation and a simple proof of logarithmic Sobolev inequality on path spaces. *Elect. Comm. Probab.,* **2**: *paper 7 (1997).*

[9] H. Djellout, A. Guillin, Moderate deviation for Markov chains with atom. *Stochastic Processes and Their Applications,* **95**, *203-217 (2001).*

[10] H. Djellout, A. Guillin, L. Wu, Transportation cost-information inequalities and application to random dynamical systems and diffusions, *Ann. Probab.,* **32**, *2702-2732 (2004).*

[11] H. Djellout, A. Guillin, L. Wu, Moderate deviations for non-linear functionals of moving average processes. Preprint 2004, to appear in *Ann. Inst. H. Poincaré.*

[12] S. Fang, Inégalité du type de Poincaré sur l'espace des chemins riemanniens, *CRAS,* **318**, *257-260 (1994).*

[13] Gentil, Inégalités de Sobolev logarithmiques et hypercontractivité en mécanique statistique et en EDP, *Thèse de Doctorat 2001, Université de Toulouse III.*

[14] E.P. Hsu, *Stochastic Analysis on Manifolds.* Graduate Studies in Mathematics, Volume 38, 2001.

[15] E.P. Hsu, Analysis on Path and Loop Spaces. *Probability theory and applications, Princeton, NJ, 1996, 277-347.*

[16] L. Gross, Logarithmic Sobolev inequalities, *Amer. J. Math.,***97**, *1061-1083 (1975).*

[17] M. Ledoux. *Concentration of measure and logatithmic Sobolev inequalities.* Séminaire de Probabilités XXXIII. *Lecture Notes in Math. Springer*, 1709, 120-216 (1999).

[18] N. Ikeda, S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*, North Holland, 1989.

[19] D. Nualart. *The Malliavin Calculus and related Topics*, Springer Verlag, 1995.

[20] F. Otto, C. Villani Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality *J. Funct. Anal.,* **173**,*361-400 (2000).*

[21] C. Villani. *Topics in optimal transportation.* Grad. Stud. Math. (58), American Mathematical Society, 2003.

[22] F.Y. Wang. Transportation cost inequalities on path spaces over Riemannian manifolds. *Illinois J maths* **46**, *1197-1206 (2002).*

[23] L. Wu. A deviation inequality for non reversible Markov processes. *Ann. Inst. Henri Poincaré, Probabilités et Statistiques* **36**, *4, 435-445 (2000).*

[24] L. Wu and Z.L. Zhang, Talagrand's T_2-transportation inequality w.r.t. a uniform metric for diffusions, *Acta Math. Appl. Sinica, English Series, Vol.* **20**, *No.3, 357-364 (2004).*