ANNALES DE L'I. H. P., SECTION B

LIMING WU

A deviation inequality for non-reversible Markov processes

Annales de l'I. H. P., section B, tome 36, nº 4 (2000), p. 435-445

<http://www.numdam.org/item?id=AIHPB_2000__36_4_435_0>

Ann. Inst. Henri Poincaré, Probabilités et Statistiques 36, 4 (2000) 435–445

A deviation inequality for non-reversible Markov processes

by

Liming WU [1]
Laboratoire de Mathématiques Appliquées et CNRS-UMR 6620,
Université Blaise Pascal, 63177 Aubiere, France

Article received in 23 September 1998, revised in 25 October 1999

ABSTRACT. – Using the dissipative criterion of Lumer–Philips for the contraction semigroup, we get in this Note a new deviation inequality for $\int_0^t V(X_s)\,ds$ by means of the symmetrized Dirichlet form. A more explicit version is obtained in the case where the logarithmic Sobolev inequality holds. © 2000 Éditions scientifiques et médicales Elsevier SAS

Key words: Dirichlet forms, Deviation inequality, Logarithmic Sobolev inequality

AMS classification: 60F10, 60J25

RÉSUMÉ. – Par le critère de dissipativité de Lumer–Philips pour la contractivité de semigroupes, on obtient une inégalité nouvelle de déviation pour $\int_0^t V(X_s)\,ds$ via la forme de Dirichlet symmetrisée. Une expression plus explicite est obtenue dans le cas où l'inégalité de Sobolev logarithmique est vraie. © 2000 Éditions scientifiques et médicales Elsevier SAS

Mots Clés: Forme de Dirichlet, Inégalité de déviation, Inégalité de Sobolev logarithmique

[1] E-mail: wuliming@ucfma.uni-bpclermont.fr.

1. Let $(\Omega, (\mathcal{F}_t)_{t\in\mathbf{R}^+}, (X_t)_{t\in\mathbf{R}^+}, (\mathbf{P}_x)_{x\in E})$ be a conservative *càdlàg* Markov process with values in a Polish space E, with semigroup of transition probability $(P_t(x,dy))$. We assume that μ is a probability measure on E (equipped with the Borel σ-field $\mathcal{B}$), which is invariant and ergodic with respect to (P_t). For any initial measure ν on E, write $\mathbf{P}_\nu := \int_E \mathbf{P}_x \nu(dx)$.

We denote by $(\mathcal{L}, \mathbf{D}_p(\mathcal{L}))$ the generator of (P_t) acting on $L^p(E,\mu)$ ($\mathbf{D}_p(\mathcal{L})$ being its domain in L^p), where $1 \leqslant p < +\infty$. The *symmetrized Dirichlet form* is given by

$$\mathcal{E}^\sigma(f,g) := \frac{1}{2}\big[\langle -\mathcal{L}f, g\rangle_\mu + \langle -\mathcal{L}g, f\rangle_\mu\big], \quad \forall f,g\in \mathbf{D}_2(\mathcal{L}), \tag{1}$$

where $\langle\cdot,\cdot\rangle_\mu$ is the usual inner product in $L^2(E,\mu)$.

Under the assumption below

$$(\mathcal{E}^\sigma, \mathbf{D}_2(\mathcal{L})) \text{ is } \textit{closable}, \tag{H1}$$

its closure $(\mathcal{E}^\sigma, \mathbf{D}(\mathcal{E}^\sigma))$ corresponds to a symmetric Markov semigroup $(P_t^\sigma)_{t\geqslant 0}$ on $L^2(E,\mu)$.

Given a measurable function $V: E\to \mathbf{R}$, μ-integrable. In this note we are interested to the probability of deviation of the empirical mean $\frac{1}{t}\int_0^t V(X_s)\,ds$ from its *real* (or asymptotic) mean $m := \int_E V\,d\mu := \langle V\rangle_\mu$, i.e.,

$$\mathbf{P}_\nu\left(\left|\frac{1}{t}\int_0^t V(X_s)\,ds - m\right| > r\right).$$

Introduce

$$J_V(r) := \inf\left\{\mathcal{E}^\sigma(f,f) \,\middle|\, f\in \mathbf{D}(\mathcal{E}^\sigma)\cap L^2(|V|\,d\mu), \int f^2\,d\mu = 1; \text{ and } \int V f^2\,d\mu = r\right\} \tag{2}$$

for every $r\in\mathbf{R}$ (*Convention*: $\inf\emptyset := +\infty$). As is easily seen, J_V is a convex function on $\mathbf{R}$. Then $[J_V < +\infty]^0$ (interior) is some interval (a,b) where $-\infty \leqslant a \leqslant b \leqslant +\infty$.

Define now I_V as the lower semi-continuous (l.s.c. in short) regularization of J_V. Obviously $I_V(m) = J_V(m) = 0$ and $I_V: \mathbf{R}\to[0,+\infty]$ is convex. Then I_V is non-decreasing on $[m,+\infty)$ and non-increasing on

$(-\infty, m]$. Notice that when $a < b$, then for any $r \in \mathbf{R}$,

$$I_V(r) = \begin{cases} J_V(r), & \text{if } r \in (a,b); \\ J_V(a+), & \text{if } r = a; \\ J_V(b-), & \text{if } r = b; \\ +\infty, & \text{otherwise.} \end{cases} \tag{3}$$

When our Markov process (X_t) is μ-reversible (or (P_t) is μ-symmetric), Deuschel and Stroock [4, Theorem 5.3.10, p. 210] (1989) proved essentially the following large deviation estimation (where a general level-2 large deviation lower bound is given)

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbf{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\,ds - m > r\right) = -I_V(m+r), \quad \forall r \geqslant 0, \tag{4}$$

for V bounded. For general unbounded V, (4) is shown in [7] (1993).

In this little note we propose to extend and strengthen (4). Our main observation is

THEOREM 1. – *Assume* (H1). *For any initial measure ν such that $\nu \ll \mu$ and $\frac{d\nu}{d\mu} \in L^2(\mu)$, we have for all $t > 0$, all $r > 0$,*

$$\mathbf{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\,ds - m > r\right) \leqslant \left\|\frac{d\nu}{d\mu}\right\|_{L^2(\mu)} \cdot \exp\left[-t \cdot I_V(m+r)\right], \tag{5}$$

$$\mathbf{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\,ds - m < -r\right) \leqslant \left\|\frac{d\nu}{d\mu}\right\|_{L^2(\mu)} \cdot \exp\left[-t \cdot I_V(m-r)\right]. \tag{6}$$

Remark 2. – In the symmetric case, the deviation inequality (5) is sharp in its exponent for large time t, by (4). The main differences between (4) and (5) are:

(i) The symmetry assumption required in (4) is removed for (5);

(ii) In (5), t and r, being arbitrary, are fixed unlike in (4) which is only an asymptotic relation ($t \to +\infty$). Hence (5) is much more stronger and practical.

However in the non-symmetric case, inequality (5) is no longer asymptotically exact. In fact, when the level-2 large deviation principle of Donsker–Varadhan holds and V is bounded, the limit (4) is given by a contraction form of the Donsker–Varadhan entropy functional, which is different from the expression in terms of Dirichlet form. See Deuschel and Stroock [4, Chapter VI] and Ben Arous and Deuschel [1] (1994).

Nevertheless that last large deviation result requires quite restrictive conditions in the non-symmetric case: indeed there exist geometrically ergodic irreducible Markov processes so that the level-1 large deviation principle fails (see Bryc and Smolenski [2] (1993)). While the deviation inequality (5) requires only (H1), which is satisfied in the most part of interesting cases. Moreover (H1) can be removed in case that V is bounded, see Remarks 3(a) below.

2. Proof of Theorem 1. Consider the Feynman–Kac semigroup

$$P_t^V f(x) := \mathbf{E}^x f(X_t) \cdot \exp\left(\int_0^t V(X_s)\,ds\right) \tag{7}$$

where $f \geqslant 0$ is $\mathcal{B}$-measurable. We shall establish for any μ-integrable function $V : E \to \mathbf{R}$,

$$0 < \|P_t^V\|_2 \leqslant \mathrm{e}^{t\Lambda(V)}, \quad \forall t \geqslant 0, \tag{8}$$

where

$$\begin{aligned}\|P_t^V\|_2 &:= \sup\{\|P_t^V f\|_{L^2(\mu)};\ f \geqslant 0 \text{ and } \langle f^2\rangle_\mu \leqslant 1\}\\ &= \sup\{\langle P_t^V f, g\rangle_\mu;\ f, g \geqslant 0 \text{ and } \langle f^2\rangle_\mu \leqslant 1,\ \langle g^2\rangle_\mu \leqslant 1\},\end{aligned}$$

and

$$\Lambda(V) := \sup\left\{-\mathcal{E}_V^\sigma(f, f) \,\middle|\, f \in \mathbf{D}(\mathcal{E}_V^\sigma),\ \int f^2 d\mu = 1\right\}. \tag{9}$$

Here

$$\mathbf{D}(\mathcal{E}_V^\sigma) := \mathbf{D}(\mathcal{E}^\sigma) \cap L^2(|V|\,d\mu), \qquad \mathcal{E}_V^\sigma(f, f) = \mathcal{E}^\sigma(f, f) - \int V f^2 d\mu.$$

Let us see quickly why (8) implies (5), by a very classical argument borrowed from the Cramèr theorem [4]. In fact set $P(\lambda) := \Lambda(\lambda V)$, $\forall \lambda \in \mathbf{R}$. By Chebychev's inequality, for all $r, t > 0$ fixed,

$$\begin{aligned}&\mathbf{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\,ds - m > r\right)\\ &\quad\leqslant \inf_{\lambda>0} \exp[-\lambda t(m + r)] \cdot \mathbf{E}^\nu \exp\left[\lambda \int_0^t V(X_s)\,ds\right]\end{aligned}$$

$$\leqslant \inf_{\lambda>0} \exp\bigl[-\lambda t(m+r)\bigr] \cdot \left\| \frac{d\nu}{d\mu} \right\|_{L^2(\mu)} \cdot \| P_t^{\lambda V} \|_2 \tag{10}$$

$$\leqslant \left\| \frac{d\nu}{d\mu} \right\|_{L^2(\mu)} \cdot \inf_{\lambda>0} \bigl\{ \exp\bigl[-\lambda t(m+r)\bigr] \cdot \mathrm{e}^{t\Lambda(\lambda V)} \bigr\} \quad \text{(by (8))}$$

$$= \left\| \frac{d\nu}{d\mu} \right\|_{L^2(\mu)} \exp\Bigl\{ -t \cdot \sup_{\lambda>0} \bigl[\lambda(m+r) - P(\lambda) \bigr] \Bigr\}.$$

It remains to identify the exponent in the last term of (10).

Since $\Lambda(\lambda V) \geqslant \lambda m$ by the definition (9), m is a sub-differential of $P(\lambda)$ at $\lambda = 0$. Thus for $r > 0$,

$$\sup_{\lambda>0} \bigl[\lambda(m+r) - P(\lambda) \bigr] = \sup_{\lambda \in \mathbf{R}} \bigl[\lambda(m+r) - P(\lambda) \bigr],$$

which is the Legendre transformation $P^\star(m+r)$ of $P(\lambda)$.

On the other hand, we have by (9)

$$P(\lambda) = \Lambda(\lambda V) = \sup\bigl\{ \lambda z - J_V(z); z \in \mathbf{R} \bigr\} = \sup\bigl\{ \lambda z - I_V(z); z \in \mathbf{R} \bigr\}$$

for all $\lambda \in \mathbf{R}$. Hence the famous Fenchel–Legendre theorem gives us

$$P^\star(m+r) = I_V(m+r).$$

Substituting those into (10), we get (5).

Applying (5) to $-V$, we get (6).

Consequently to conclude this theorem, it remains to show (8). We divide its proof into three cases.

Case 1. – *V bounded.* In this bounded case (P_t^V) is a strongly continuous semigroup of bounded operators on $L^2(\mu)$, whose generator is exactly $(\mathcal{L} + V; \mathbf{D}_2(\mathcal{L} + V) = \mathbf{D}_2(\mathcal{L}))$ by the well known Feynman–Kac formula. By the definition (9) of $\Lambda(V)$,

$$\bigl\langle \bigl(\mathcal{L} + V - \Lambda(V)\bigr) f, f \bigr\rangle_\mu \leqslant 0, \quad \forall f \in \mathbf{D}_2(\mathcal{L}). \tag{11}$$

That means exactly that the generator $\mathcal{L} + V - \Lambda(V)$ with domain $\mathbf{D}_2(\mathcal{L})$ is a dissipative operator on $L^2(E,\mu)$ in the sense of Lumer and Philips [9, Chapter IX, p. 250]. By the Lumer–Philips Theorem [9, Chapter IX, p. 250], the semigroup $(\mathrm{e}^{-t\Lambda(V)} P_t^V)$ generated by $\mathcal{L} + V - \Lambda(V)$ is contractive on $L^2(E,\mu)$. In other words,

$$\|e^{-t\Lambda(V)}P_t^V\|_2 \leqslant 1, \quad \forall t \geqslant 0,$$

which is exactly (8).

Case 2. – *V upper bounded* ($V \leqslant a$). Considering $V - a$ if necessary, we can assume $V \leqslant 0$. Take $V_n = \max\{V, -n\}$ for $n \in \mathbf{N}$. We have by the Case 1,

$$\|P_t^V\|_2 \leqslant \lim_{n\to\infty} \|P_t^{V_n}\|_2 \leqslant \lim_{n\to\infty} e^{t\Lambda(V_n)} = \exp\big(t \cdot \inf_{n\geqslant 1} \Lambda(V_n)\big). \tag{12}$$

Recall that

$$\begin{aligned} -\Lambda(V_n) &= \inf\left\{\mathcal{E}^\sigma(f,f) - \int V_n f^2\,d\mu \,\Big|\, f \in \mathbf{D}(\mathcal{E}^\sigma) \text{ and } \int f^2 d\mu \leqslant 1\right\} \\ &= \inf\left\{F_n(f) \,\Big|\, \int f^2 d\mu \leqslant 1\right\}, \end{aligned}$$

where $F_n : L^2(E,\mu) \to [0, +\infty]$ is given by

$$F_n(f) := \mathcal{E}^\sigma(f,f) - \int V_n f^2\,d\mu, \quad \text{if } f \in \mathbf{D}(\mathcal{E}^\sigma), \text{ and } +\infty \text{ else}.$$

By Kato [5, p. 461, Lemma 3.14a] and our assumption (H1), F_n is lower semicontinuous on $L^2(E,\mu)$ with respect to the strong topology, then with respect to the weak topology $\sigma(L^2, L^2)$ (since F_n, being the sum of two nonnegative quadratic forms, is convex on $L^2(E,\mu)$). Moreover, since the unit ball $\{f \in L^2(\mu); \int f^2 d\mu \leqslant 1\}$ is compact with respect to $\sigma(L^2, L^2)$, by an elementary analytical lemma (see e.g. [8, Proposition 1.2]),

$$\begin{aligned} -\inf_{n\geqslant 1} \Lambda(V_n) &= \sup_{n\geqslant 1} \inf\left\{F_n(f) \,\Big|\, \int f^2\,d\mu \leqslant 1\right\} \\ &= \inf\left\{\sup_n F_n(f) \,\Big|\, \int f^2\,d\mu \leqslant 1\right\} = -\Lambda(V). \end{aligned}$$

Substituting it into (12), we get (8) again.

Case 3. – *General case.* Take $V^N = \min\{V, N\}$ for $N \in \mathbf{N}$. By the monotone convergence theorem,

$$\|P_t^V\|_2 = \sup\{\langle P_t^V f, g\rangle_\mu \,|\, f, g \geqslant 0 \text{ and } \langle f^2\rangle_\mu \leqslant 1, \langle g^2\rangle_\mu \leqslant 1\}$$

$$= \sup_{N\geqslant 1} \sup\left\{\left\langle P_t^{V^N} f, g\right\rangle_\mu \middle| f, g \geqslant 0 \text{ and } \langle f^2\rangle_\mu \leqslant 1,\ \langle g^2\rangle_\mu \leqslant 1\right\}$$
$$\leqslant \sup_{N\geqslant 1} \mathrm{e}^{t\Lambda(V^N)} = \mathrm{e}^{t\Lambda(V)},$$

where the third inequality follows from the Case 2, and the last equality follows from the fact that $\mathbf{D}(\mathcal{E}^\sigma) \cap L^\infty(\mu)$ is a form core for all $\mathcal{E}^\sigma_{V^N}$, $N \geqslant 1$, and for the not necessarily closable quadratic form $\mathcal{E}^\sigma_V$.

The proof of (8) and then that of Theorem 1 are so finished. □

Remark 3. –

(a) When V is bounded, it holds that

$$\left\|P_t^V\right\|_2 \leqslant \exp\left[t \cdot \Lambda^0(V)\right]$$

where

$$\Lambda^0(V) := \sup\left\{\int V f^2\, d\mu + \langle \mathcal{L} f, f\rangle_\mu \middle| f \in \mathbf{D}_2(\mathcal{L}) \text{ and } \langle f^2\rangle_\mu \leqslant 1\right\} \tag{13}$$

without the assumption (H1) about the closability of $(\mathcal{E}^\sigma, \mathbf{D}_2(\mathcal{L}))$, by the proof in the Case 1 above. As in the proof of (8) $\Rightarrow$ (5) above, one can deduce from (13) the deviation inequalities (5) and (6) without (H1), but with I_V substituted by the l.s.c. regularization I_V^0 of

$$J_V^0(r) := \inf\left\{\mathcal{E}^\sigma(f, f) \middle| f \in \mathbf{D}_2(\mathcal{L}),\ \int f^2 d\mu = 1;\ \int V f^2\, d\mu = r\right\}.$$

When (H1) is satisfied and V is bounded, $\Lambda^0(\lambda V) = \Lambda(\lambda V)$, $\forall \lambda \in \mathbf{R}$ (by the fact that $\mathbf{D}_2(\mathcal{L})$, being a form core of $\mathcal{E}^\sigma$, is so for $\mathcal{E}^\sigma_{\lambda V}$ because of the boundedness of V), and then $I_V^0 = I_V$.

(b) Note also the following (indicated by the referee): the inequality (8) implies not only (5) and (6), but also (with the same argument)

$$\mathbf{E}^\mu f(X_0) g(X_t) 1_{\left[\frac{1}{t}\int_0^t V(X_s)\, ds - m > r\right]}$$
$$\leqslant \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)} \cdot \exp\left[-t \cdot I_V(m + r)\right], \quad \forall r, t > 0.$$

(c) Applying the Lumer–Philips theorem to $\mathcal{L} - V$ in $L^p(\mu)$ with $1 \leqslant p < +\infty$, we get, instead of (8), that for any V bounded,

$$\left\|P_t^V\right\|_p \leqslant \exp\left(t \Lambda_p(V)\right)$$

where

$$\Lambda_p(V) := \sup\left\{\int V|f|^p\,d\mu + \left\langle \mathrm{sgn}(f)|f|^{p-1}, \mathcal{L}f\right\rangle_\mu \,\middle|\, f \in \mathbf{D}_p(\mathcal{L}),\ \langle |f|^p\rangle_\mu = 1\right\}.$$

3. In this paragraph we do not require (H1) but we assume the log-Sobolev inequality below: there exists $C > 0$ such that for all $f \in \mathbf{D}_2(\mathcal{L})$,

$$\int_E f^2 \log f^2 - \langle f^2\rangle_\mu \log\langle f^2\rangle_\mu \leqslant C\langle -\mathcal{L}f, f\rangle_\mu. \tag{14}$$

Consider the log-Laplace transformation of $V - m$:

$$H(\lambda) = \log \int_E \mathrm{e}^{\lambda V}\,d\mu - \lambda m \tag{15a}$$

and its Legendre transformation

$$H^\star(r) = \sup\{\lambda r - H(\lambda);\ \lambda \in \mathbf{R}\}. \tag{15b}$$

By the classical Cramèr's theorem [4], $H^\star$ governs the large deviation principle of the i.i.d. sequence of common law $\mu(V - m \in \cdot)$.

The following result says that the log-Sobolev inequality (14) implies a same type of estimation as in the i.i.d. case.

COROLLARY 4. – *Assume* (14) (*not* (H1)). *Then for any* $V \in L^1(\mu)$,

$$\frac{1}{t}\log\|P_t^V\|_2 \leqslant \frac{1}{C}\log\int_E \mathrm{e}^{CV}\,d\mu. \tag{16}$$

In particular for each initial measure $\nu \ll \mu$ *with* $\frac{d\nu}{d\mu} \in L^2(\mu)$ *and for all* $r > 0, t > 0$

$$\mathbf{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\,ds - m > r\right) \leqslant \left\|\frac{d\nu}{d\mu}\right\|_{L^2(\mu)} \cdot \exp\left(-\frac{t}{C}H^\star(r)\right). \tag{17}$$

Proof. – The deviation inequality (17) follows from (16) by Chebychev's inequality as in Theorem 1. To show the key (16), assume at first that V is bounded.

By (13) in Remark 3, we have

$$\frac{1}{t}\log\|P_t^V\|_2$$
$$\leqslant \sup\left\{\int Vf^2\,d\mu + \langle \mathcal{L}f, f\rangle_\mu \,\middle|\, f\in \mathbf{D}_2(\mathcal{L}) \quad \text{and } \langle f^2\rangle_\mu = 1\right\}$$
$$\leqslant \sup\left\{\int Vf^2\,d\mu - \frac{1}{C}\int f^2\log f^2\,d\mu \,\middle|\, f\in \mathbf{D}_2(\mathcal{L}) \quad \text{and } \langle f^2\rangle_\mu = 1\right\} \quad \text{(by (14))}$$
$$= \frac{1}{C}\log\int_E e^{CV}\,d\mu,$$

where the last equality follows from Donsker–Varadhan's variational formula (see e.g. [8]).

Now for V unbounded, set $V_n = \min\{\max\{V, -n\}, n\}$. We have

$$\|P_t^V\|_2 \leqslant \liminf_{n\to+\infty}\|P_t^{V_n}\|_2 \leqslant \lim_{n\to\infty}\left(\int e^{CV_n}\,d\mu\right)^{t/C} = \left(\int e^{CV}d\mu\right)^{t/C}$$

by the bounded case shown above and the dominated convergence (and Fatou's lemma if the last integral is infinite). (16) is hence established. □

Remark 5. – Ledoux [6] (1999) develops systematically the so called Herbst method which consists to derive deviation inequalities from a log-Sobolev inequality. The strategy consists to apply a log-Sobolev inequality to $e^{\lambda F}$ to obtain a differential inequation, from which a control on $\mathbf{E}e^{\lambda F}$ is deduced by comparison lemma. Nevertheless for that strategy works here for $F = \int_0^t V(X_s)\,ds$, we should assume that a log-Sobolev inequality on the path space $(\mathbf{D}([0,t], E), \mathbf{P}_\nu)$ holds, which is in general not the case here.

Even in case that such a path level log-Sobolev inequality holds, it seems that the Herbst method does not give directly better estimation than (17). For instance, let (B_t) be the Brownian motion on a Riemannian manifold E, with generator $\Delta/2$, where Δ is the Laplace–Beltrami operator. Assume that the Ricci curvature satisfies $|Ric_u| \leqslant K$ for all $u\in O(E)$ (the bundle of orthonormal frames on E). By Capitaine–Hsu–Ledoux [3, (6)], the path level log-Sobolev inequality below holds:

$$\mathbf{E}^x\left(F^2\log F^2\right) - \mathbf{E}^x F^2\log \mathbf{E}^x F^2 \leqslant 2e^{Kt}\mathbf{E}^x|DF|_H^2 \qquad (18)$$

for any $x \in E$ and $F : C([0,t]; E) \to \mathbf{R}$ provided that the right side term above is finite, where $|DF|_H$ is the norm in the Cameron–Martin subspace of the Malliavin derivative DF on the path space. Now the Herbst method developed in [6, §2.3] yields: if $|DF|_H^2 \leqslant \sigma^2$, $\mathbf{P}_x$-a.s., then

$$\mathbf{P}_x(F - \mathbf{E}^x F > r) \leqslant \exp\left(-\frac{r^2}{2e^{Kt}\sigma^2}\right). \tag{19}$$

Using the notations of [3], we can easily prove that for $F = \int_0^t V(B_s)\,ds$ with $\|\nabla V\|_\infty := \sup_{x\in E} |\nabla V(x)| < +\infty$ (where $|\nabla V(x)|$ is the Riemannian norm of the gradient of V at x),

$$|DF|_H^2 \leqslant \int_0^t \left(\int_s^t |\nabla V|(B_u)\,du\right)^2 ds \leqslant \|\nabla V\|_\infty^2 \cdot \frac{t^3}{3}, \quad \mathbf{P}_x\text{-a.s.}$$

We then obtain by (19),

$$\mathbf{P}_x\left(\int_0^t V(B_s)\,ds - \mathbf{E}^x \int_0^t V(B_s)\,ds > rt\right) \leqslant \exp\left(-\frac{3r^2}{2te^{Kt}\|\nabla V\|_\infty^2}\right). \tag{20}$$

That estimation is quite interesting and sharp for small t, but not so for large t. On the other hand, when E is compact, the log-Sobolev inequality (14) holds (a well known fact), then (17) is valid and it gives a much better estimation than (20) for large t.

Our approach in Corollary 4 consists to apply log-Sobolev inequality after obtaining the control of $\|P_t^V\|_2$ (in Theorem 1), not before, unlike in the Herbst method. One can regard it as another application of log-Sobolev inequality, complementing those amply developed by Ledoux [6].

ACKNOWLEDGEMENTS

This work is done when the author visits the Center of Mathematics, Academie Sinica during May–June 1998. I am grateful to Professor Ma Zhi-Ming for his kind invitation. The warm hospitality of that Center is acknowledged. My thanks go especially to the referee for his careful comments on the first version of this note.

REFERENCES

[1] Ben Arous G., Deuschel J.D., The rate function of hypoelliptic diffusions, Comm. Pure Appl. Math. XLVII (6) (1994) 843–860.

[2] Bryc W., Smolenski W., On the convergence of averages of mixing sequences, J. Theoret. Probab. 6 (3) (1993) 473–484.

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

[4] Deuschel J.D., Stroock D.W., Large deviations, Pure and Appl. Math. 137 (1989).

[5] Kato T., Perturbation Theory for Linear Operators, 2nd ed., Springer, Berlin, 1984; (2nd corrected printing).

[6] Ledoux M., Concentration of measure and logarithmic Sobolev inequalities, in: Séminaire de Probab. XXXIII, Lecture Notes in Math., Vol. 1709, Springer, 1999, pp. 120–216.

[7] Wu L.M., Feynman–Kac semigroups, ground state diffusions and large deviations, J. Funct. Anal. 123 (1) (1994) 202–231.

[8] Wu L.M., An introduction to large deviations, in: Yan J.A., Peng S., Fang S., Wu L. (Eds.), Several Topics in Stochastic Analysis, Academic Press of China, Beijing, 1997, pp. 225–336; (in chinese).

[9] Yosida K., Functional Analysis, 3rd edn., Springer, 1971.