ESAIM: PS
January 2008, Vol. 12, p. 12–29
DOI: 10.1051/ps:2007032

ESAIM: Probability and Statistics
www.edpsciences.org/ps

DEVIATION BOUNDS FOR ADDITIVE FUNCTIONALS OF MARKOV PROCESSES

Patrick Cattiaux[1] and Arnaud Guillin[2]

Abstract. In this paper we derive non asymptotic deviation bounds for

$$\mathbb{P}_\nu\left(\left|\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s - \int V\mathrm{d}\mu\right| \geq R\right)$$

where X is a μ stationary and ergodic Markov process and V is some μ integrable function. These bounds are obtained under various moments assumptions for V, and various regularity assumptions for μ. Regularity means here that μ may satisfy various functional inequalities (F-Sobolev, generalized Poincaré etc.).

Mathematics Subject Classification. 60F10, 60J25.

Received March 13, 2006. Revised June 19 and September 26, 2006.

1. Introduction, framework and first results

On some Polish space E, let us consider a conservative (continuous time) Markov process $(X_t, (\mathbb{P}_x)_{x\in E})$ and its associated semi-group $(P_t)_{t\geq 0}$ with infinitesimal generator L (and denote $D(L)$ its domain). Let μ be a probability measure on E which is invariant and ergodic w.r.t. P_t. The celebrated ergodic Theorem tells us that for any V in $\mathbb{L}^1(\mu)$

$$A(t,R,V) := \mathbb{P}_\mu\left(\left|\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s - \int V\mathrm{d}\mu\right| \geq R\right) \to 0$$

as t goes to $+\infty$ for all $R>0$. Level 1 large deviations theory furnishes asymptotic bounds for $\frac{1}{t}\log(A(t,R,V))$ (see *e.g.* [9]). If V is bounded, one may replace the initial measure μ by a μ absolutely continuous probability measure ν.

It is however of major importance in practice to exhibit non asymptotic upper bound but also to ensure practical conditions to verify them; see for example a priori bounds for large and moderate deviations in averaging principle, concentration for particular approximations of granular media equation... It will be the purpose of the present note. In [24], Liming Wu derived such bounds.

Keywords and phrases. Deviation inequalities, functional inequalities, additive functionals.

[1] École Polytechnique, CMAP, 91128 Palaiseau cedex, France, CNRS 756, and Université Paris X Nanterre, équipe MODAL'X, UFR SEGMI, 200 avenue de la République, 92001 Nanterre cedex, France; cattiaux@cmapx.polytechnique.fr

[2] Ceremade, Université Paris IX Dauphine, 75775 Paris cedex, France, CNRS 7534; guillin@ceremade.dauphine.fr

© EDP Sciences, SMAI 2008

Article published by EDP Sciences and available at http://www.esaim-ps.org or http://dx.doi.org/10.1051/ps:2007032

Theorem 1.1 (Wu [24]). *If V is bounded, then for all $t > 0$ and all $R > 0$*

$$\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s - \int V\mathrm{d}\mu \geq R\right) \leq \left|\left|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right|\right|_{\mathbb{L}^2(\mu)} \exp\left\{-tI_V\left(R + \int V\mathrm{d}\mu\right)\right\} \tag{1.2}$$

where $I_V(a) = \sup_{\lambda\geq 0}\{\lambda a - \Lambda(\lambda V)\}$ and

$$\Lambda(V) := \sup\left\{\int Vf^2\mathrm{d}\mu + \langle Lf, f\rangle_\mu; f \in D(L) \text{ and } \int f^2\mathrm{d}\mu = 1\right\}.$$

Of course a similar result holds for $\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s - \int V\mathrm{d}\mu \leq -R\right)$.

The key is that

$$\frac{1}{t}\log\left|\left|P_t^V\right|\right|_{\mathbb{L}^2(\mu)} \leq \Lambda(V)$$

for all $t > 0$, where P_t^V denotes the Feynman-Kac semi-group built from P_t. This result is a consequence of Lumer-Philips theorem. It is worthwhile noticing that, when P_t is μ symmetric, the above bound is asymptotically sharp, according to the spectral radius theorem, but (1.2) is also asymptotically sharp according to large deviations theory (see [9] Th. 5.3.10).

The main difficulty is then to be able to give a precise (and if possible optimal) control of the quantity $I_V(a)$ hence of $\Lambda(\lambda V)$. Our approach mainly relies on the use of functional inequalities to get upper bound on $\Lambda(\lambda V)$. Let us illustrate this approach *via* the use of a Poincaré inequality (or spectral gap inequality), *i.e.* for all nice enough f,

$$\mathrm{Var}_\mu(f) \leq -C_P\langle Lf, f\rangle_\mu. \tag{1.3}$$

Take first V bounded. Of course, by homogeneity, we may only consider the V's satisfying $\int Vd\mu = 0$ and $\sup|V| = 1$, for which the only interesting R's are between 0 and 1. Indeed notice that the bound (1.2) is fortunately 0 if $R > 1$ in this case since $\Lambda(\lambda V) \leq \lambda$ so that $I_V(R) = +\infty$ if $R > 1$. The next result furnishes an explicit bound as soon as a μ satisfies a Poincaré inequality

Proposition 1.4. *Assume that μ satisfies the Poincaré inequality (1.3). Then for all V such that $\sup|V| = 1$, all $0 < R \leq 1$ and all $t > 0$*

$$\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s - \int V\mathrm{d}\mu \geq R\right) \leq \left|\left|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right|\right|_{\mathbb{L}^2(\mu)} \exp\left(-tA\right), \tag{1.5}$$

where

$$A = \frac{1}{C_P}\left(1 - \sqrt{1 - \frac{R}{R + \mathrm{Var}_\mu(V)}}\right)\left(R - \frac{\mathrm{Var}_\mu(V)\left(1 - \sqrt{1 - \frac{R}{R+\mathrm{Var}_\mu(V)}}\right)}{\sqrt{1 - \frac{R}{R+\mathrm{Var}_\mu(V)}}}\right).$$

Remark 1.6. The above bound A is quite intricate. Note that, $A \sim \frac{R^2}{4C_P\mathrm{Var}_\mu(V)}$ when R goes to 0. Actually one can show (see the proof below) that

$$A \geq \frac{R^2}{8C_P\mathrm{Var}_\mu(V)} \text{ if } R \leq 2\mathrm{Var}_\mu(V), \text{ and } A \geq \frac{R - \mathrm{Var}_\mu(V)}{2C_P} \text{ if } R \geq 2\mathrm{Var}_\mu(V).$$

These bounds are similar to those obtained by Lezaud [16] using Kato's perturbation theory. Lezaud's bound depends on the asymptotic variance. Actually the best uniform result contained in [16] (see Rem. 1.2 therein) for small R's is exactly ours. Our proof below is much shorter.

Proof. We may assume that $\int V \mathrm{d}\mu = 0$. If $\int f^2 \mathrm{d}\mu = 1$ we may write

$$f = \frac{1+\varepsilon g}{\sqrt{1+\varepsilon^2}}$$

for some $\varepsilon \geq 0$ and some g satisfying $\int g \mathrm{d}\mu = 0$ and $\int g^2 \mathrm{d}\mu = 1$, and conversely. Thus applying Poincaré inequality with $G_P = 1/C_P$, for any arbitrary λ

$$\begin{aligned}\Lambda(\lambda V) &\leq \sup\left\{\int \lambda V f^2 \mathrm{d}\mu - G_P \mathrm{Var}_\mu(f); f \in D(L) \text{ and } \int f^2 \mathrm{d}\mu = 1\right\} \\ &\leq \sup_{\varepsilon \geq 0}\left(\frac{\varepsilon}{1+\varepsilon^2}\sup\left\{2\lambda \int V g \mathrm{d}\mu + \varepsilon \int (\lambda V - G_P) g^2 \mathrm{d}\mu\right\}\right)\end{aligned}$$

where the second supremum is taken over the set $\left\{g \in D(L), \int g^2 \mathrm{d}\mu = 1, \int g \mathrm{d}\mu = 0\right\}$. It follows, according to Cauchy-Schwarz inequality and our hypotheses, that

$$\Lambda(\lambda V) \leq \sup_{\varepsilon \geq 0} \frac{\varepsilon}{1+\varepsilon^2}\left(2\lambda \mathrm{Var}_\mu(V) + \varepsilon(\lambda - G_P)\right),$$

so that, bounding roughly $1/(1+\varepsilon^2)$ by 1, we finally obtain that for $\lambda < G_P$,

$$\Lambda(\lambda V) \leq \frac{\lambda^2 \mathrm{Var}_\mu(V)}{G_P - \lambda}.$$

Recall that $I_V(R) = \sup_{\lambda \geq 0}\{\lambda R - \Lambda(\lambda V)\}$, so that we finally deduce that

$$I_V(R) \geq \sup_{\lambda \in [0, G_P]}\left\{\lambda R - \frac{\lambda^2 \mathrm{Var}_\mu(V)}{G_P - \lambda}\right\}.$$

This supremum is attained for

$$\lambda = G_P\left(1 - \sqrt{1 - \frac{R}{R + \mathrm{Var}_\mu(V)}}\right)$$

and is equal to A. We can conclude thanks to Wu's theorem.

Remark that we obtain another lower bound for $I_V(R)$ if we consider the supremum for λ running on $[0, G_P/2]$ only, and replace $G_P - \lambda$ by $G_P/2$. The supremum is then attained for $\lambda = \frac{RG_P}{4\mathrm{Var}_\mu(V)}$ provided this quantity is less than $G_P/2$, and for $\lambda = G_P/2$ otherwise, yielding the lower bounds for A in the remark. □

An interesting feature is that Proposition 1.4 admits a (partial) converse. Indeed

Proposition 1.7. *Assume that μ is diffuse (i.e. for any A and any $0 \leq \alpha \leq \mu(A)$ there exists $B \subseteq A$ such that $\mu(B) = \alpha$).*

Assume that there exist C and λ_0 such that for all V such that $\int V \mathrm{d}\mu = 0$ and $\sup|V| = 1$ and all $0 \leq \lambda \leq \lambda_0$, $\Lambda(\lambda V) \leq C\lambda^2$. Then μ satisfies a Poincaré inequality.

Proof. Let f and g as in the proof of Proposition 1.4, and V such that $\int V \mathrm{d}\mu = 0$ and $\sup|V| = 1$.

We deduce from the hypotheses that for all $\varepsilon \geq 0$ and all g such that $\int g \mathrm{d}\mu = 0$ and $\int g^2 \mathrm{d}\mu = 1$, for $0 \leq \lambda \leq \lambda_0$,

$$0 \leq C(1+\varepsilon^2)\lambda^2 - \lambda\left(\varepsilon^2 \int V g^2 \mathrm{d}\mu + 2\varepsilon \int V g \mathrm{d}\mu\right) - \varepsilon^2 \langle Lg, g\rangle_\mu.$$

We show how to choose appropriate V's to get controls on the variance. Notice that the above quantity reaches its minimum for

$$\lambda = \frac{2\varepsilon \int Vg\mathrm{d}\mu + \varepsilon^2 \int Vg^2\mathrm{d}\mu}{2C(1+\varepsilon^2)}$$

that goes to 0 when ε goes to 0 and to $\int Vg^2\mathrm{d}\mu/2C \leq 1/2C$ when ε goes to $+\infty$. Thus, taking a larger C if necessary, we may assume that $1/2C \leq \lambda_0$, and changing V into $-V$ if necessary, we may assume that $\int Vg^2\mathrm{d}\mu \geq 0$.

For ε small enough the minimum is reached at some $\lambda \leq \lambda_0$ and has to be nonnegative. It follows

$$\left(2\int Vg\mathrm{d}\mu + \varepsilon \int Vg^2\mathrm{d}\mu\right)^2 \leq 4C(1+\varepsilon^2)\langle Lg, g\rangle_\mu,$$

so that letting ε go to 0 we obtain

$$\left(\int Vg\mathrm{d}\mu\right)^2 \leq -C\langle Lg, g\rangle_\mu.$$

We may then choose $V = \mathrm{sign}(g) - \int \mathrm{sign}(g)\mathrm{d}\mu$ in order to obtain

$$\left(\int |g|\mathrm{d}\mu\right)^2 \leq -C\langle Lg, g\rangle_\mu. \tag{1.8}$$

For ε going to $+\infty$ and provided $\int Vg^2\mathrm{d}\mu \geq 0$ we also obtain

$$\left(\int Vg^2\mathrm{d}\mu\right)^2 \leq -4C\langle Lg, g\rangle_\mu. \tag{1.9}$$

We shall now build an appropriate V.

Let $A = \{|g| \leq 1/2\}$. First, if $\mu(A) \leq 1/2$, $\int |g|\mathrm{d}\mu \geq 1/4$, so that (1.8) implies

$$\int g^2\mathrm{d}\mu = 1 \leq -16C\langle Lg, g\rangle_\mu.$$

If $\mu(A) \geq 1/2$, denote by $B = \{|g| \geq 3/4\}$. We have

$$1 = \int g^2\mathrm{d}\mu \leq \int_{B^c} g^2\mathrm{d}\mu + \int_B g^2\mathrm{d}\mu \leq \frac{9}{16} + \int_B g^2\mathrm{d}\mu$$

so that $\int_B g^2\mathrm{d}\mu \geq 7/16$. Choose $A' \subseteq A$ such that $\mu(A') = \mu(B)$, and $V = \mathbb{1}_B - \mathbb{1}_{A'}$. Then

$$\int Vg^2\mathrm{d}\mu \geq \frac{7}{16} - \frac{1}{4}\mu(B) \geq \frac{3}{16} \geq 0,$$

and

$$\int g^2\mathrm{d}\mu = 1 \leq \left(\frac{16}{3}\right)^2 4C\langle -Lg, g\rangle_\mu.$$

Hence there exists some constant K such that $\int g^2\mathrm{d}\mu \leq -K\langle Lg, g\rangle_\mu$ for all g with mean 0 and variance 1, that is Poincaré inequality holds. □

Remark 1.10. It is easy to see that (1.8) (which holds without the assumption of diffusivity) implies the following, $\forall g$ nice enough

$$\mathrm{Var}_\mu^2(g) \leq C\langle Lg, g\rangle_\mu \parallel g \parallel_\infty^2,$$

which is some weak Poincaré inequality. This inequality implies some concentration property for μ (see *e.g.* [21] or [3]) but is quite far from the usual Poincaré inequality. More precisely the aforementioned weak Poincaré inequality on $\mathbb{R}$ implies that μ concentrates like $\alpha(\mathrm{d}s) = c/(1+|s|^3)\mathrm{d}s$, *i.e.* $\mu(|x|>s)$ behaves like $\alpha(|x|>s)$. This weak Poincaré inequality is actually satisfied by $\alpha(\mathrm{d}s)$.

It is not difficult to see that $\Lambda(\lambda V) \le C\lambda^p$ for some $p>2$ (and small λ's) cannot happen (using the same method). This is natural since for very small R (namely $R = r/\sqrt{t}$) we cannot expect a better behaviour as a Gaussian one due to the central limit theorem (see *e.g.* [16] Th. 3.1), while $\Lambda(\lambda V) \le C\lambda^p$ would imply $I_V(R) \ge CR^p$.

We can now state the problems we shall study in the sequel:

- What happens if Poincaré is reinforced, replacing it by stronger functional inequalities? The answer to this question is partly given in [24] for the log-Sobolev inequality, namely we may consider in this case unbounded V having some exponential moments.
- What can be said for bounded V's when Poincaré's inequality does not hold?
- For unbounded V's, how can we obtain (may be rough) deviation bounds in full generality?
- What happens if the initial measure is no more absolutely continuous, or when its density is less integrable?

2. Exponential bounds for unbounded V's and strong functional inequalities

Let us start here with an almost immediate extension of Wu's result, tackling the first question.

Definition 2.1. Let F be defined on $\mathbb{R}^+$. We assume that F is continuous, increasing, concave, goes to $+\infty$ at ∞ and satisfies $F(1)=0$. It follows that F admits an inverse function which is defined on $]F(0),+\infty[$. In addition we assume that F satisfies

$$F(xy) \le F(x) + F(y), \tag{2.2}$$

for all positive x and y.

We say that μ satisfies an F-Sobolev inequality if for all $f \in D(L)$ such that $\int f^2 d\mu = 1$,

$$\int f^2 F(f^2)\mathrm{d}\mu \le -\langle Lf, f\rangle_\mu.$$

Theorem 2.3. *Assume that μ satisfies an F-Sobolev inequality. Then for all V and all $R>0$*

$$\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s - \int V\mathrm{d}\mu \ge R\right) \le \left|\left|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right|\right|_{\mathbb{L}^2(\mu)} \exp\left(-tH^*\left(R + \int V\mathrm{d}\mu\right)\right), \tag{2.4}$$

where

$$H^*(a) := \sup_{0\le\lambda<\lambda_V}\left\{\lambda a - F\left(\int F^{-1}(\lambda V)\mathrm{d}\mu\right)\right\},$$

where λ_V is such that $\lambda V > F(0)$ for all $0\le\lambda<\lambda_V$. In particular $H^*(a)=0$ if $\lambda_V = 0$ or if $\int F^{-1}(\lambda V)d\mu = +\infty$ for all $\lambda>0$.

We also have

$$\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s - \int V\mathrm{d}\mu \ge R\right) \le \left|\left|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right|\right|_{\mathbb{L}^2(\mu)} \exp\left(-tH_c^*(R)\right),$$

with

$$H_c^*(a) := \sup_{0\le\lambda<\lambda'_V}\left\{\lambda a - F\left(\int F^{-1}\left(\lambda\left(V - \int V\mathrm{d}\mu\right)\right)\mathrm{d}\mu\right)\right\},$$

where λ'_V is such that $\lambda(V - \int V\mathrm{d}\mu) > F(0)$ for all $0\le\lambda<\lambda'_V$. This latter bound is better than the previous one when $F(0) = -\infty$.

Proof. We shall use again Wu's theorem. Assume first that V is bounded. Applying the F-Sobolev inequality, we get

$$\Lambda(\lambda V) \quad \leq \quad \sup\left\{\int \lambda V f^2 \mathrm{d}\mu - \int f^2 F(f^2)\mathrm{d}\mu; f \in D(L) \text{ and } \int f^2 \mathrm{d}\mu = 1\right\}.$$

For all f as above, it thus holds

$$\begin{aligned} F\left(\int F^{-1}(V)\mathrm{d}\mu\right) &= F\left(\int \frac{F^{-1}(V)}{f^2} f^2 \mathrm{d}\mu\right) \\ &\geq \int F\left(\frac{F^{-1}(V)}{f^2}\right) f^2 \mathrm{d}\mu \\ &\geq \int \left(V - F(f^2)\right) f^2 \mathrm{d}\mu \end{aligned} \tag{2.5}$$

where we have successively used the facts that F is non-decreasing, concave and (2.2). It follows that

$$\Lambda(\lambda V) \leq F\left(\int F^{-1}(\lambda V)\mathrm{d}\mu\right),$$

yielding $I_V \geq H^*$.

If V is not bounded just approximate it by $(V \wedge n) \vee -n$.

Finally we may replace V by $V - \int V \mathrm{d}\mu$ and obtain H_c^*. The property $F(xy) \leq F(x) + F(y)$ immediately shows that this bound is better than the previous one except that the authorized set of λ's differ in general, except for $F(0) = -\infty$ where $\lambda_V = \lambda'_V = +\infty$. □

Remark 2.6. The bound obtained in Theorem 2.3 is interesting since, assuming some regularity for F, $\Lambda(\lambda(V - \int V \mathrm{d}\mu))$ behaves like λ^2 for small λ provided it is finite for some $\lambda_0 > 0$. Hence H_c^* is strictly positive and actually behaves like Ca^2 for small a while it behaves like Ca for large a.

Note that if $F(0) > -\infty$ the theorem only applies to the bounded from below V's.

In the examples below $-\langle Lf, f\rangle_\mu = 1/2 \int |\nabla f|^2 \mathrm{d}\mu$, corresponding to diffusion process with constant diffusion term.

Example 2.7.

1) The function $F(x) = C\log(x)$ satisfies all the previous assumptions with $F(0) = -\infty$. The corresponding result is then Corollary 4 in [24]. Gaussian measures satisfy such log-Sobolev inequalities. In this case $F^{-1}(y) = \exp(y/C)$, so that the above result holds as soon as V has some exponential moment.

Note that some converse holds in this case. Indeed if $F = C\log$, and if

$$\Lambda(V) \leq \frac{1}{C}\log\left(\int \mathrm{e}^{CV}\mathrm{d}\mu\right),$$

for all V, then for all f we may choose $V = \frac{1}{C}\log f^2$ and deduce the log-Sobolev inequality.

2) The functions $F_\alpha(x) = \log^\alpha(1+x) - \log^\alpha(2)$ also satisfy all the assumptions as soon as $0 < \alpha \leq 1$ (see the proof of Th. 38 in [4] for instance). The measure $\mu_\beta(\mathrm{d}x) = \frac{\exp(-|x|^\beta)}{Z_\beta}\mathrm{d}x$ satisfies a $c_\alpha F_\alpha$-Sobolev inequality for $\alpha = 2(1 - 1/\beta)$ and some well chosen constant c_α (see [4] Sect. 7). Here $F_\alpha^{-1}(y)$ behaves like $\exp(y^{1/\alpha})$ at infinity. Here again some converse holds, but details are a little bit tedious.

3) Conditions for some F-Sobolev inequalities are discussed in details in [4] and [5]. In particular explicit (and tractable) criteria for absolutely continuous measures on the line are given in [4] Theorem 27, while sufficient

conditions are discussed in [5] Section 8 for a general Riemannian manifold. In these papers the corresponding F-Sobolev inequalities satisfy the tensorization property, hence due to the CLT, F cannot grow faster than a logarithm. That is the described field of measures is between Poincaré and Gross (log-Sobolev) inequalities. The condition $F(xy) \leq F(x) + F(y)$ certainly obliges us to such a restriction.

Example 2.8. It is interesting to see how the previous result applies on simple examples. Consider the standard Ornstein-Uhlenbeck process on $\mathbb{R}$, $\mathrm{d}X_t = \mathrm{d}B_t - \frac{1}{2}X_t\mathrm{d}t$ with its symmetric probability measure γ the standard normal law. γ satisfies a log-Sobolev inequality with constant $C = 4$. Easy calculations yield

$$\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t X_s^2\mathrm{d}s - 1 \geq R\right) \leq \left|\left|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right|\right|_{\mathbb{L}^2(\mu)} \exp\left(-\frac{t}{8}(R - \log(1+R))\right).$$

This result is not asymptotically sharp, since according to a result by Bryc and Dembo [6] the large deviations rate function is $R^2/8(R+1)$ which is greater than our $(R - \log(1+R))/8$. In particular for small R, we are loosing a factor 2. Note that Lezaud obtains in [16] Example 4.2 the correct bound, but that this case is a little bit miraculous since the spectral gap of the Feynman-Kac operator can be explicitly calculated. We shall discuss other explicit examples later on.

Remark 2.9. In the examples above, we have assumed that the diffusion coefficient is constant. And one knows that F-Sobolev inequalities are usually verified with the energy given by $\int |\nabla f|^2$ which could be seen as a limitation on the diffusion process we may consider for our deviation inequalities. However we may easily replace this assumption by some strict ellipticity, namely suppose that there exists $\delta > 0$ such that for all x, y, $\langle \sigma(x)\sigma(x)^* y, y\rangle \geq \delta|y|^2$, then

$$-\langle Lf, f\rangle = \frac{1}{2}\int |\sigma(x)\nabla f(x)|^2\mathrm{d}\mu(x) \geq \frac{\delta}{2}\int |\nabla f|^2\mathrm{d}\mu.$$

It enables us to consider deviation inequalities for strictly elliptic diffusion of the form

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t$$

using "standard" functional inequalities.

Remark 2.10. One strongly suspects that the integrability condition ($\int F^{-1}(\lambda V)\mathrm{d}\mu < +\infty$ for some $\lambda > 0$) is also necessary for an exponential bound to hold. We do not know whether this is true in full generality or not, but one can easily build some examples.

Still in the Gaussian case of Example 2.8, consider $V(x) = x^4$, and define $V_N = V \wedge N$ and $F^{-1} = \exp$. Choosing $f(x) = c\mathrm{e}^{x^2/4}/(1+x^2)$ for some normalizing constant c, one can show by a direct (but tedious) calculation using Laplace method, that for N large and $\lambda \geq 1/N^{\frac{1}{8}}$,

$$\Lambda(\lambda V_N) \geq D\lambda N^{\frac{1}{4}} \geq DN^{\frac{1}{8}}$$

for some nonnegative constant D. It follows that $I_{V_N}(a) \leq a/N^{\frac{1}{8}}$. The bound for I_{V_N} is asymptotically sharp since we are in the symmetric case (recall the remark in the introduction). Thus for all R there exists some t_N (extracting a diagonal subsequence the same t_N can be used for all $R \in \mathbb{Q}$) such that

$$\mathbb{P}_\mu\left(\frac{1}{t_N}\int_0^{t_N} V_N(X_s)\mathrm{d}s - \int V_N\mathrm{d}\mu \geq R\right) \geq \exp\left(-\frac{1}{2}t_N I_{V_N}\left(R + \int V_N\mathrm{d}\mu\right)\right),$$

from which it is easy to deduce (taking $R = R' + \int V - \int V_N$ and using that $V \geq V_N$),

$$\mathbb{P}_\mu\left(\frac{1}{t_N}\int_0^{t_N} V(X_s)\mathrm{d}s - \int V\mathrm{d}\mu \geq R'\right) \geq \exp\left(-\frac{1}{2}t_N N^{-\frac{1}{8}}\left(R' + \int V\mathrm{d}\mu\right)\right).$$

Hence we have no asymptotic exponential bound.

Since Theorem 2.3 is not satisfactory when $F(0) > -\infty$, we shall complete it, at least when a Poincaré inequality also holds.

Theorem 2.11. *Let F and μ be as in Theorem 2.3. Assume in addition that μ satisfies some Poincaré inequality with constant C_P. Let V such that $\int V\mathrm{d}\mu = 0$ and $\int V^2\mathrm{d}\mu = m^2 < +\infty$. Assume that $\int F^{-1}(\lambda_0 V^+)\mathrm{d}\mu < +\infty$ for some $\lambda_0 > 0$ and define*

$$\lambda_1 = \sup\left\{0 < \lambda \leq \lambda_0; F\left(\int F^{-1}(2\lambda V \mathbb{1}_{\lambda V > 1/4C_P})\mathrm{d}\mu\right) \leq 1/4C_P\right\}.$$

Then for all $R > 0$

$$\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s \geq R\right) \leq \left|\left|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right|\right|_{\mathbb{L}^2(\mu)} \exp\left\{-t\sup_{0\leq\lambda\leq\lambda_1}\left(R\lambda - 8m^2C_P\lambda^2\right)\right\}.$$

Proof. In the definition of Λ (see Th. 1.1) we write again $f = (1+\varepsilon g)/\sqrt{1+\varepsilon^2}$. It yields

$$\Lambda(\lambda V) = \sup_{\varepsilon\geq 0}\sup_g\left(\frac{\varepsilon}{1+\varepsilon^2}\sup\left\{2\lambda\int Vg\mathrm{d}\mu + \varepsilon\left(\int \lambda V g^2\mathrm{d}\mu + \langle Lg, g\rangle_\mu\right)\right\}\right), \tag{2.12}$$

where the second supremum runs on the $g \in D(L)$ such that $\int g\mathrm{d}\mu = 0$ and $\int g^2\mathrm{d}\mu = 1$. First, thanks to Cauchy-Schwarz inequality, $\int \lambda V g d\mu \leq m\lambda$. Next, the second term is splitting into the sum of

$$\begin{aligned}\int \lambda V g^2 \mathbb{1}_{\lambda V \leq G_P/4}\mathrm{d}\mu + \frac{1}{2}\langle Lg, g\rangle_\mu &\leq \frac{G_P}{4}\int g^2\mathrm{d}\mu + \frac{1}{2}\langle Lg, g\rangle_\mu \\ &\leq -(G_P/4),\end{aligned}$$

according to Poincaré's inequality (recall that $G_P C_P = 1$), and of

$$\int \lambda V g^2 \mathbb{1}_{\lambda V > G_P/4}\mathrm{d}\mu + \frac{1}{2}\langle Lg, g\rangle_\mu \leq \frac{1}{2}F\left(\int F^{-1}(2\lambda V \mathbb{1}_{\lambda V > G_P/4})\mathrm{d}\mu\right).$$

The latter is obtained by applying 2.5 replacing f by g and V by $2\lambda V \mathbb{1}_{\lambda V > G_P/4}$. Using the definition of λ_1 we finally see that for $\lambda \leq \lambda_1$,

$$\Lambda(\lambda V) \leq \sup_{\varepsilon\geq 0}\left(2\varepsilon m\lambda - (1/8C_P)\varepsilon^2\right) \leq 8m^2\lambda^2 C_P.$$

The result follows by Theorem 1.1. □

Remark 2.13.

1) The positivity of λ_1 is ensured by the properties of F and the positivity of λ_0, while the finiteness of the variance of V is ensured by the positivity of λ_0. Once again we obtain a Gaussian bound for small R and an exponential one for large R.

2) As shown by Aida [1] a F-Sobolev inequality together with a weak Poincaré inequality imply the ordinary Poincaré inequality. Since any absolutely continuous measure μ on a manifold with bounded from below Ricci

curvature satisfies a weak Poincaré inequality, the Poincaré inequality is automatically satisfied in this case. In particular Theorem 2.11 completes the picture for the F_α introduced in Example 2.7.

3) One can easily obtain a very rough bound for λ_1. For instance if $F^{-1} = \exp(x^\theta)$ for some $\theta > 1$, the following holds

$$\int \mathbb{1}_{\lambda_1 V > 1/4C_P} \mathrm{e}^{(2\lambda_1 V^+)^\theta} \mathrm{d}\mu \leq \mathrm{e}^{(1/4C_P)^\theta} - 1. \tag{2.14}$$

Recall that, since we are bounding from below some supremum, we may if necessary choose $(2\lambda_1)^\theta \leq \frac{1}{2}(\lambda_0)^\theta$. Applying Cauchy-Schwarz inequality we obtain

$$\mu(\lambda_1 V > 1/4C_P) \leq \frac{\left(\mathrm{e}^{(1/4C_P)^\theta} - 1\right)^2}{\int \mathrm{e}^{(\lambda_0 V^+)^\theta} \mathrm{d}\mu}.$$

Since the left hand side in (2.14) is less than $\left(\int \mathrm{e}^{(\lambda_0 V^+)^\theta} \mathrm{d}\mu\right) \mathrm{e}^{-(\lambda_0/4C_P\lambda_1)^\theta}$ we get an explicit condition for λ_1.

Remark 2.15. In all explicit cases we considered, $x \mapsto xF(x) = G(x)$ is convex. Hence, using the F-Sobolev inequality, $\Lambda(\lambda V) \leq \sup\left(\int \lambda V h \mathrm{d}\mu - \int G(h)\mathrm{d}\mu\right)$ where the supremum is taken over all nonnegative h such that $\int h \mathrm{d}\mu = 1$. This kind of maximization problem is well known in convex analysis since it relies on the calculation of the Fenchel-Legendre transform of an integral which is a convex functional. If we relax both constraints on h, one expects that this supremum is equal to $\int G^*(\lambda V)\mathrm{d}\mu$, where G^* is the usual Fenchel-Legendre conjugate. Actually the situation is a little bit more intricate since $\mathbb{L}^1$ is not reflexive (see [14, 19, 20]). Nevertheless the result we get using this (potential) bound is not interesting. Indeed consider $G(x) = x\log(x)$ ($G(x) = +\infty$ if $x < 0$) so that $G^*(u) = \mathrm{e}^{u-1}$. Since $G^*(0) \neq 0$ we do not obtain any interesting bound for small R. For instance if $V = \mathbb{1}_A - \mathbb{1}_{A^c}$ for some A with $\mu(A) = 1/2$ we obtain $H(R) = R \arg\sinh(eR) - \sqrt{1 + \mathrm{e}^2 R^2}/\mathrm{e}$ which is negative for small R. It seems that the maximisation problem taking into account the (non linear) constraints on h is not easy and we did not find any reference on it (see however [15] for connected problems with linear constraints).

Nevertheless something can be made for still stronger F-Sobolev inequalities. First we introduce some definitions.

Definition 2.16. We shall say that F is a contractive function if $F(x) \to +\infty$ when x goes to $+\infty$ and $x \mapsto xF(x) := G(x)$ is a normalised Young function. This means that G (defined on $\mathbb{R}^+$) is convex, non-decreasing, satisfies $G(0) = 0$ and $G(1) + G^*(1) = 1$ where G^* is the Fenchel-Legendre conjugate of G. We shall denote by N_G the corresponding gauge norm (*i.e.* $N_G(f) = \inf\{u > 0; \int G(f/u)\mathrm{d}\mu \leq G(1)\}$).

Definition 2.17. Let F be a contractive function.

- We shall say that μ satisfies the strong F-Sobolev inequality with constant C_{SF} if for all $g \in D(L)$ such that $\int g \mathrm{d}\mu = 0$ and $\int g^2 \mathrm{d}\mu = 1$ it holds

$$\int g^2 F(g^2)\mathrm{d}\mu \leq -C_{SF}\langle Lg, g\rangle_\mu.$$

- We shall say that μ satisfies the (defective) F-Sobolev inequality with constants C_F and C_b if for all $f \in D(L)$ such that $\int f^2 \mathrm{d}\mu = 1$ it holds

$$\int f^2 F(f^2)\mathrm{d}\mu \leq -C_F\langle Lf, f\rangle_\mu + C_b F(1).$$

  If $C_b = 1$ we say that the inequality is tight.

Before stating the results we have in mind in the above situation, we shall discuss Definition 2.17 and give some examples.

Example 2.18.

1) It is immediate that a F-Sobolev inequality together with a Poincaré inequality imply a strong F-Sobolev inequality with $C_{SF} = C_F + C_b F(1) C_P$.

2) When $F(x) = x^p$ the local version of the strong F-Sobolev inequality is quite useful for studying regularizing effects in p.d.e. theory for elliptic degenerate operators. In particular if L is a sub-elliptic operator in $\mathbb{R}^d$ satisfying some degeneracy conditions, we may associate a natural distance to L and the balls corresponding to the distance. If $\mathrm{d}\mu/\mathrm{d}x$ belongs to some appropriate Muckenhoupt space, then μ will satisfy the strong x^p-Sobolev for some appropriate p in all balls. For precise results in this direction see *e.g.* Franchi [12] Theorem 4.5 or Lu [17] Theorem B.

3) The defective F-Sobolev inequality is much more well known. First we are using "tight" following Bakry, while we used "additive" in [4]. Tight means that we have an equality for $f \equiv 1$. If a non tight inequality holds (*i.e.* replacing $F(1)$ by a larger constant) together with a Poincaré inequality, then modifying the constant C_F we may obtain a tight one. Actually this result is not proved in full generality but is proved in [2] for $F(x) = x^p$ (the usual Sobolev inequality). Hence on a Riemannian manifold with Ricci curvature bounded from below, we may apply Aida's result formerly recalled.

It is well known in the symmetric case (see [8] Cor. 2.4.3) that a x^p-Sobolev inequality for $p > 2$ is equivalent to the ultracontractive bound $\| P_t f \|_\infty \leq c t^{-s} \| f \|_2$ for $0 < t \leq 1$ with $s = p/2(p-2)$. We shall use this bound in the next section.

More generally F-Sobolev inequalities are related to super-Poincaré inequalities. A precise discussion is done in [23] (also see [13]). In particular it is shown therein that for $\beta \geq 1$ the measure $\mu_\beta(\mathrm{d}x) = \frac{\exp(-|x|^\beta)}{Z_\beta}\mathrm{d}x$ satisfies a defective F_α-Sobolev inequality with $F_\alpha(x) = (\log(1+x))^\alpha$ and $\alpha = 2(1 - 1/\beta)$. According to 1), μ_β thus satisfies the strong F_α-Sobolev inequality (since it satisfies Poincaré).

4) In the previous points we did not take care on the normalization assumption for G. It is known that if G is not normalized one can find some k such that $G(kx)$ is normalized. If G is moderate (*i.e.* $G(2x) \leq cG(x)$ for some c and all x) we may replace it by its normalized equivalent, up to a change in C_{SF}. Hence in the cases we discussed before, the normalization hypothesis is not really relevant.

We conclude this section with the analogue of Theorem 2.11 in the contractive situation.

Theorem 2.19. *Assume that μ satisfies a strong F-Sobolev inequality for some contractive function F. Let V such that $\int V \mathrm{d}\mu = 0$ and $\int V^2 \mathrm{d}\mu = m^2 < +\infty$. Then for all $R > 0$*

$$\mathbb{P}_\nu \left(\frac{1}{t} \int_0^t V(X_s) \mathrm{d}s \geq R \right) \leq \left\| \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \right\|_{\mathbb{L}^2(\mu)} \exp \left\{ -t \sup_{0 \leq \lambda \leq \lambda_2} \left(R\lambda - (2m^2 C_{SF}/G(1))\lambda^2 \right) \right\},$$

where $\lambda_2 = G(1)/(2C_{SF} N_{G^*}(V))$. In particular for this bound to be interesting one needs $N_{G^*}(V) < +\infty$.

Proof. Recall (2.12)

$$\Lambda(\lambda V) = \sup_{\varepsilon \geq 0} \sup_g \left(\frac{\varepsilon}{1+\varepsilon^2} \sup \left\{ 2\lambda \int V g \mathrm{d}\mu + \varepsilon \left(\int \lambda V g^2 \mathrm{d}\mu + \langle Lg, g \rangle_\mu \right) \right\} \right),$$

for well chosen g's. According to the Hölder-Orlicz inequality (for normalized Young functions)

$$1 = \int g^2 \mathrm{d}\mu \leq N_G(g^2) N_{G^*}(1) = N_G(g^2).$$

Using the convexity of G and $G(0)=0$, it follows that $\int G(g^2)\mathrm{d}\mu \geq G(1)N_G(g^2)$. Using Hölder-Orlicz inequality again and the strong F-Sobolev inequality, it holds

$$\begin{aligned}\int \lambda V g^2 \mathrm{d}\mu + \langle Lg, g\rangle_\mu &\leq N_{G^*}(\lambda V)N_G(g^2) - (1/C_{SF})\int G(g^2)\mathrm{d}\mu \\ &\leq \left(\lambda N_{G^*}(V) - (G(1)/C_{SF})\right) N_G(g^2) \\ &\leq \lambda N_{G^*}(V) - (G(1)/C_{SF}),\end{aligned}$$

provided $((G(1)/C_{SF}) - \lambda N_{G^*}(V)) > 0$. It follows that

$$\Lambda(\lambda V) \leq \sup_{\varepsilon\geq 0}\left\{2m\lambda\varepsilon - \left((G(1)/C_{SF}) - \lambda N_{G^*}(V)\right)\varepsilon^2\right\} \leq m^2\lambda^2/\left((G(1)/C_{SF}) - \lambda N_{G^*}(V)\right).$$

Once again, we may bound I_V by taking the supremum on $[0, G(1)/(2C_{SF}N_{G^*}(V))]$ and replacing $(G(1)/C_{SF}) - \lambda N_{G^*}(V)$ by $G(1)/(2C_{SF}N_{G^*}(V))$. The proof is completed. □

The previous proof is certainly simpler than the one of Theorem 2.11, while both theorems apply to similar measures, for instance the μ_β's for $\beta > 1$. It is quite difficult to compare the bounds in both Theorems on this example, but the one in Theorem 2.19 has to be worse in general since it lies on the rough use of Hölder and Orlicz norms.

Remark 2.20. We have not discussed here the use of another type of functional inequalities called transportation cost inequalities (in path space), namely

$$\forall \nu, \qquad W_p(\nu,\mu) \leq \sqrt{2C\int \log(\mathrm{d}\nu/\mathrm{d}\mu)\mathrm{d}\nu}$$

where W_p is the usual Wasserstein distance, leading to Gaussian type of deviation inequalities for Lipschitz test function V. But the proofs are very different in spirit as they rely on the verification of some square exponential integrability for some norm on the path space and a (dependent) tensorization property. We refer to Djellout-Guillin-Wu [10]. Note that the results obtained there are reminiscent of an assumption of a logarithmic Sobolev inequality to hold. However they do not rely on the knowledge of the invariant measure but on the conditions on the drift and diffusion coefficient. In the same spirit, one may also use Poincaré inequalities or logarithmic Sobolev inequalities on path space combined with Herbst's argument, but they are much more difficult to prove and do not give good bounds for large time asymptotic.

3. Polynomial and sub-exponential bounds

3.1. The case of bounded V's

In this subsection we shall assume that the semi-group $(P_t)_{t\geq 0}$ satisfies the following decay property:

Assumption 3.1. *There exists some non increasing function η defined on $[0,+\infty[$ such that for all bounded f and all t,*

$$\mathrm{Var}_\mu(P_t f) \leq \eta(t) \parallel f - \int f\mathrm{d}\mu \parallel_\infty^2 .$$

It is known (see [21] Ths. 2.1 and 2.3) that Assumption 3.1 is formally equivalent to a weak Poincaré inequality (WPI). More precisely, if μ satisfies a (WPI), *i.e.* for all $s>0$ and all bounded g,

$$\mathrm{Var}_\mu(g) \leq -\beta(s)\langle Lg, g\rangle_\mu + s \parallel g - \int g\mathrm{d}\mu \parallel_\infty^2, \tag{3.2}$$

for some non increasing β, then Assumption 3.1 holds with

$$\eta(t) \leq 2 \inf \{s > 0; \beta(s) \log(1/s) \leq 2t\}.$$

Conversely, in the symmetric case (ore more generally if L is a normal operator), if η is decreasing with inverse function η^{-1}, then Assumption 3.1 implies a (WPI) with

$$\beta(t) = 2t \inf_{s>0} \left(\frac{1}{s} \eta^{-1}(s \exp[1 - s/t]) \right).$$

In particular, if $\eta(t) \leq e^{-\delta t}$ for some $\delta > 0$, Assumption 3.1 implies the (true) Poincaré inequality. In the sequel we may thus assume that η is decaying slower than an exponential. We shall give explicit examples later.

Assumption 3.1 is clearly connected to **mixing** properties. Indeed recall the

Definition 3.3. Let $\mathcal{F}_s$ (resp. $\mathcal{G}_s$) be the σ-field generated by $(X_u)_{u \leq s}$ (resp. $(X_u)_{u \geq s}$). The strong mixing coefficient $\alpha(r)$ is defined as $\alpha(r) = \sup_{s,F,G}\{|Cov(F,G)|\}$ where the supremum runs over s and F (resp. G) $\mathcal{F}_s$ (resp. $\mathcal{G}_{s+r}$) measurable, non-negative and bounded by 1. The process is then said to be strongly mixing with α.

We then have

Proposition 3.4. *If Assumption 3.1 holds then the stationary process is strongly mixing with $\alpha(r) \leq \sqrt{\eta(r)}$. If in addition μ is symmetric we may choose $\alpha(r) \leq \eta(r/2)$.*

Conversely, if the stationary process is strongly mixing, then Assumption 3.1 holds with $\eta(r) \leq \alpha(r)$.

Proof. Let F and G are centered and bounded by 1, respectively $\mathcal{F}_s$ and $\mathcal{G}_{s+r}$ measurable. We may apply the Markov property to get

$$\mathbb{E}_\mu[FG] = \mathbb{E}_\mu[F\mathbb{E}_\mu[G|X_{s+r}]] = \mathbb{E}_\mu[FP_r g(X_s)]$$

where g is centered and bounded by 1. Indeed since the state space E is Polish, we may find a measurable g such that $\mathbb{E}_\mu[G|X_{s+r}] = g(X_{s+r})$ (desintegration of measure).

Hence

$$|\mathbb{E}_\mu[FG]| \leq \mathbb{E}_\mu[|P_r g(X_s)|] = \int |P_r g| \mathrm{d}\mu \leq \sqrt{\eta(r)}.$$

In the symmetric case

$$\mathbb{E}_\mu[FP_r g(X_s)] = \mathbb{E}_\mu[F(X_{s-.})P_r g(X_0)] = \mathbb{E}_\mu[f(X_0)P_r g(X_0)] = \int P_{r/2} f P_{r/2} g \mathrm{d}\mu$$

and we conclude using Cauchy-Schwarz again.

For the converse, take $F = P_r f(X_0)$ and $G = f(X_r)$ for f centered and bounded by one. □

The point is that moment bounds for sums of strongly mixing sequences (extending Rosenthal's inequalities in the independent case) are known. A large part of them are due to Doukhan and his coauthors and may be found in Doukhan's book [11]. However we found the most refined version we shall use in Rio's book [18].

Proposition 3.5. *Assume that μ satisfies Assumption 3.1 for some η satisfying for some integer k, $M_k(\alpha) := \sup_r (1+r)^k \alpha(r) < +\infty$ with α as in Proposition 3.4.*

Then there exists a constant $C(k)$ such that for all V with $\sup |V| = 1$, all $0 < R \leq 1$ and all $t < [t]/(1-R)$ where $[t]$ is the integer part of t

$$\mathbb{P}_\mu \left(\frac{1}{t} \int_0^t V(X_s) \mathrm{d}s - \int V \mathrm{d}\mu \geq R \right) \leq \frac{C(k) M_k(\alpha)}{t^k \left(R - (1 - ([t]/t))\right)^{2k}}. \tag{3.6}$$

Proof. Denote by $Y_j = \int_{j-1}^{j} V(X_s)\mathrm{d}s - \int V\mathrm{d}\mu$. Then Y_j is a $(\mathbb{P}_\mu)$ stationary sequence of strongly mixing centered random variables with mixing coefficient $\alpha(r-1)$. Thanks to our hypothesis on α we may apply Theorem 2.2 in [18] (see (2.23) p. 40 therein) which yields $\mathbb{E}_\mu[(\sum_1^n Y_j)^{2k}] \le C(k)M_k(\alpha)n^k$. The result follows by using Markov inequality and the fact that V is bounded by 1. □

In the previous result one can obtain explicit bounds for $C(k)M_k(\alpha)$ as shown by Doukhan and Portal (see [11] Chap. 1.4).

In the examples below again, $-\langle Lf, f\rangle_\mu = 1/2\int |\nabla f|^2\mathrm{d}\mu$.

Example 3.7. 1) If $\mu(\mathrm{d}x) = c(1+|x|)^{-(\mathrm{d}+p)}\mathrm{d}x$ $(p>0)$ on $\mathbb{R}^d$ it is shown in [3] that (WPI) holds with $\beta(s) = c(\mathrm{d})s^{-2/p}$. Actually this result is shown for $d=1$ but extends to the d dimensional case since the tensorized 1-dimensional measure is equivalent to the d-dimensional one (of course all constants depend on d). Hence we may choose $\eta(t) = c(d,p)(\log t/t)^{p/2}$. The bound in Proposition 3.5 is thus available for $p > 2k$ in the symmetric case, and $p > 4k$ in the non symmetric-one.

2) If $\mu(\mathrm{d}x) = c\mathrm{e}^{-|x|^p}\mathrm{d}x$ for some $1 \ge p > 0$, we obtain similarly $\eta(t) = c(d,p)\mathrm{e}^{-c't^{\frac{p}{2-p}}}$. We can thus obtain any polynomial bound. Of course, in this case one expects a better bound. We shall see how to get such a bound below.

In order to get sub-exponential bounds, we recall the following moment inequality from [18] Theorem 2.5, that holds for a $(\mathbb{P}_\mu)$ stationary sequence of strongly mixing centred random variables Y_j bounded by 1

$$\mathbb{E}_\mu[|\sum_1^n Y_j|^{2k}] \le (4nk)^k \int_0^1 (\alpha^{-1}(u)\wedge n)^k \mathrm{d}u. \tag{3.8}$$

Note that (3.8) allows us to give an explicit bound for $C(k)$ in Proposition 3.5, but with a slightly worse speed. Indeed if $\alpha(n) = cn^{-k}$ we get

$$\mathbb{E}_\mu\left[\left|\sum_1^n Y_j\right|^{2k}\right] \le (4nk)^k c(1+\log c + k\log n)$$

recovering (3.6) with an extra logarithm.

Recall now the following elementary

$$\limsup_{q\to+\infty} q^{-1}\left(\int_0^1 \log^q(1/u)\mathrm{d}u\right)^{1/q} \le 1/\mathrm{e}.$$

If $\alpha(n) = c\mathrm{e}^{-c'n^{p/(2-p)}}$ it follows that there exists some k_0 depending only on c such that

$$\int_0^1 (\alpha^{-1}(u)\wedge n)^k\mathrm{d}u \le \left(\frac{(2-p)k}{\mathrm{e}pc'}\right)^{\frac{(2-p)k}{p}},$$

for all $k \ge k_0$. Hence $\mathbb{E}_\mu[|\sum_1^n Y_j|^{2k}] \le \left(4n\left(\frac{2-p}{\mathrm{e}pc'}\right)^{\frac{2-p}{p}} k^{2/p}\right)^k$.

Using Markov inequality it thus holds

$$\mathbb{P}_\mu\left(\int_0^n V(X_s)\mathrm{d}s - n\int V\mathrm{d}\mu \ge S\sqrt{n}\right) \le \mathrm{e}^{-\frac{2-p}{p}k}\left(2\left(\frac{2-p}{pc'}\right)^{\frac{2-p}{2p}} k^{1/p}(1/S)\right)^{2k}.$$

We then choose $k^{1/p} = (pc'/2-p)^{2-p/2p}S/2$ provided it is greater than k_0^p. Finally we have obtained,

Proposition 3.9. *Assume that μ satisfies Assumption 3.1 with $\alpha(s) = c\mathrm{e}^{-c's^{p/(2-p)}}$ as in Proposition 3.4, for some $0 < p \leq 1$.*

Then there exists a constant k_0 depending on c, such that for all V with $\sup|V| = 1$, all $0 < R \leq 1$ and all $t < [t]/(1-R)$ where $[t]$ is the integer part of t

$$\mathbb{P}_\mu\left(\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s - \int V\mathrm{d}\mu \geq R\right) \leq \exp\left\{-c(p)\left((R-(1-([t]/t)))\sqrt{t}\right)^p\right\}, \tag{3.10}$$

with $c(p) = \frac{2-p}{p}(1/2)^p(pc'/2-p)^{2-p/2}$, provided $(R-(1-([t]/t)))\sqrt{t} \geq 2k_0^{1/p}(2-p/pc')^{2-p/2p}$.

Remark 3.11. If the previous result is in accordance with the C.L.T. (that holds as soon as $\int_0^\infty \alpha(s)\mathrm{d}s < +\infty$), it is of course worse than the ones we obtained in the first section. Indeed for $p = 1$ we recover a convergence rate $\mathrm{e}^{-C\sqrt{t}}$ (for some fixed R) while we know that (at least in the symmetric case) a Poincaré inequality holds, hence Proposition 1.4 gives a convergence rate e^{-Ct}.

This fact suggests that we may loose something in the time discretization. At the same time we may ask whether it is possible to use the semi-group structure to calculate

$$G_{2k}(t) := \mathbb{E}_\mu\left[\left|\int_0^t V(X_s)\mathrm{d}s\right|^{2k}\right]$$

or not. If $k \in \mathbb{N}$ it is possible to study the variations of G_{2k}, at least in the symmetric diffusion case. We assume now that $\int V\mathrm{d}\mu = 0$. Then

$$G_2'(t) = 2\mathbb{E}_\mu\left[V(X_t)\int_0^t V(X_s)\mathrm{d}s\right] = 2\int_0^t\int\left(P_{s/2}V\right)^2\mathrm{d}\mu\mathrm{d}s \leq 2\int_0^t \eta(s)\mathrm{d}s.$$

Hence if $\eta \in \mathbb{L}^1(\mathbb{R}^+, dt)$ we obtain that $G_2(t) \leq c_2 t$.

Using integration by parts and symmetry one can show that

$$\begin{aligned}(G_4)''(t) &= 12\mathbb{E}_\mu\left[V(X_0)V(X_t)\left(\int_0^t V(X_s)\mathrm{d}s\right)^2\right]\\ &= 24\mathbb{E}_\mu\left[V(X_0)V(X_t)\left(\int_0^{t/2} V(X_s)\mathrm{d}s\right)^2\right]\\ &\quad +24\mathbb{E}_\mu\left[V(X_0)V(X_t)\left(\int_{t/2}^t V(X_s)\mathrm{d}s\right)\left(\int_0^{t/2} V(X_s)\mathrm{d}s\right)\right].\end{aligned}$$

The first term in the above sum can be bounded by $6t^2\sqrt{\eta(t/2)}$, so that if this last quantity is in $\mathbb{L}^1(\mathrm{d}t)$ it furnishes a contribution $c_4 t$ again to G_4. The second term in the sum can be rewritten with the help of the function

$$H(x) = \mathbb{E}_x\left[\int_0^{t/2} V(X_{t/2})V(X_s)\mathrm{d}s\right] = \int_0^{t/2} P_s\left(VP_{t/2-s}V\right)(x)\mathrm{d}s.$$

It yields

$$\mathbb{E}_\mu\left[V(X_0)V(X_t)\left(\int_{t/2}^t V(X_s)\mathrm{d}s\right)\left(\int_0^{t/2} V(X_s)\mathrm{d}s\right)\right] = \mathbb{E}_\mu\left[V(X_0)\left(\int_0^{t/2} V(X_s)\mathrm{d}s\right)H(X_{t/2})\right]$$
$$= \mathbb{E}_\mu\left[H(X_0)V(X_{t/2})\left(\int_{t/2}^t V(X_s)\mathrm{d}s\right)\right]\left(=\int H^2\mathrm{d}\mu\right)$$
$$= \int_0^{t/2}\int VP_s(VP_{t/2-s}H)\mathrm{d}\mu\mathrm{d}s = \int_0^{t/2}\int (P_sV)V(P_{t/2-s}H)\mathrm{d}\mu\mathrm{d}s$$
$$= \int_0^{t/2}\int (P_sV)V(P_{t/2-s}(H-\int H\mathrm{d}\mu))\mathrm{d}\mu\mathrm{d}s + \left(\int H\mathrm{d}\mu\right)\int_0^{t/2}\int (P_sV)V\mathrm{d}\mu\mathrm{d}s$$
$$\le \int_0^{t/2}\sqrt{\eta(s)}\sqrt{\eta(t/2-s)}\mathrm{Var}_\mu^{1/2}(H)\mathrm{d}s + \left(\int H\mathrm{d}\mu\right)\int_0^{t/2}\eta(s/2)\mathrm{d}s.$$

But

$$\int H\mathrm{d}\mu = \int_0^{t/2}\int VP_{t/2-s}V\mathrm{d}\mu\mathrm{d}s \le \int_0^{t/2}\eta(s/2)\mathrm{d}s$$

is assumed to be bounded by d_4 (see the control of G_2). Since $\mathrm{Var}_\mu(H) \le \int H^2\mathrm{d}\mu$ it follows that

$$\mathrm{Var}_\mu^{1/2}(H) \le \int_0^{t/2}\sqrt{\eta(s)}\sqrt{\eta(t/2-s)}\mathrm{d}s + d_4 \le (t/2)\sqrt{\eta(t/4)} + d_4.$$

Since we formerly assumed that $t^2\sqrt{\eta(t/2)}$ goes to 0, the variance is bounded and consequently so is $(G_4)''$ yielding a bound c_4t^2 for G_4.

Unfortunately, it seems difficult to iterate the procedure and to get explicit expressions for the constants. Furthermore, one suspects that a clever study will yield $G_{2k}(t) \le c_{2k}t^k$, that is the same behaviour as in the discrete case. It does not seem necessary to go further.

3.2. **Unbounded V's**

If V is no more bounded, or does not fulfill the hypotheses of one of the result in the second section, one can get some bound by truncating V. We shall briefly indicate how to do on an example.

For instance for a centered V such that $\int |V|\mathrm{d}\mu < +\infty$, and all $K > 0$

$$\mathbb{P}_\mu\left(\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s \ge R\right)$$
$$\le \mathbb{P}_\mu\left(\frac{1}{t}\int_0^t (V\wedge K\vee -K)(X_s)\mathrm{d}s - \int (V\wedge K\vee -K)\mathrm{d}\mu \ge R/2 - \int (V\wedge K\vee -K)\mathrm{d}\mu\right)$$
$$+\mathbb{P}_\mu\left(\frac{1}{t}\int_0^t |V|\mathbb{1}_{|V|\ge K}(X_s)\mathrm{d}s \ge R/2\right) = A + B.$$

If $\int |V|^S\mathrm{d}\mu < +\infty$, B can be bounded by $K^{m-S}R^{-m}2^m\mathbb{E}_\mu[|V|^S]$ for all $1 \le m < S$. If $\int \mathrm{e}^{u|V|}\mathrm{d}\mu < +\infty$ for some $u > 0$, we have $B \le \mathrm{e}^{-\lambda}\int \mathrm{e}^{\lambda|V|\sqrt{2/KR}}\mathrm{d}\mu$ as soon as $\lambda\sqrt{2/KR} \le u$ (just summing up the previous bounds for $m = S/2$).

In order to obtain a bound for A we may use the appropriate results in Section 1 or in the previous subsection.

If we assume for example that μ satisfies a Poincaré inequality, and that K is such that $R/4 \geq \mathbb{E}_\mu[|V|^S]/K^{S-m}$ we may apply Proposition 1.4 and obtain a bound for A in the form $\exp\{-tR^2/128C_PK^2\}$.

Choosing $m = S/2$ (provided $S \geq 2$), it is not difficult to see that the (almost) optimal choice is given by $K = cR\sqrt{t}/\sqrt{\log(2tR^4/S)}$ with t large enough for this expression to be meaningful and the previous constraint between R and K to be satisfied.

We thus obtain a bound

$$C(S)\log^{S/4}(2tR^4/S)R^{-S/2}t^{-S/4},$$

for t large enough, μ satisfying a Poincaré inequality and $\mathbb{E}_\mu[|V|^S] < +\infty$ for some $S \geq 2$.

In the same way, if $\int \mathrm{e}^{u|V|}\mathrm{d}\mu < +\infty$, we first choose $\lambda = u\sqrt{KR/2}$, and a similar method yields a bound

$$C(u)\mathrm{e}^{-c(u)t^{1/5}R^{4/5}},$$

for t large enough.

4. About the initial measure

In this final section we shall see what can be said for the initial measure ν. As for the latter subsection, we shall not state general results, but give some hints in various situations. Of course we shall discuss how to get deviation bounds for $\mathbb{P}_\nu(\mathcal{F})$ which are not simply given by $\mathbb{P}_\mu^{1/p}(\mathcal{F}) \parallel \mathrm{d}\nu/\mathrm{d}\mu \parallel_q$.

A. We have seen in Sections 2 and 1 that we may take some initial measure ν such that $\mathrm{d}\nu/\mathrm{d}\mu \in \mathbb{L}^2(\mu)$. As remarked by Wu [24] pp. 441–442, we may replace this assumption by $\mathrm{d}\nu/\mathrm{d}\mu \in \mathbb{L}^q(\mu)$ for $1 \leq q < +\infty$, provided we replace Λ by

$$\Lambda_p(V) := \sup\left\{\int V|f|^p\mathrm{d}\mu + \langle \mathrm{sgn}(f)|f|^{p-1}, Lf\rangle_\mu; f \in D_p(L) \text{ and } \int |f|^p\mathrm{d}\mu = 1\right\},$$

where p and q are conjugate. If L admits a "*carré du champ*" Γ, one can integrate by parts and get

$$\langle \mathrm{sgn}(f)|f|^{p-1}, Lf\rangle_\mu = -(4(p-1)/p^2)\int \Gamma(|f|^{p/2})\mathrm{d}\mu,$$

so that defining $g = |f|^{p/2}$ we obtain that

$$\Lambda_p(V) = (4(p-1)/p^2)\Lambda((p^2/4(p-1))V)$$

at least for a bounded V (remark that $(p-1)/p^2 = (q-1)/q^2$).

Hence all the results in Sections 2 and 1 are still true, up to the constants, for $1 < q < +\infty$. For instance we get an additional constant $4(p-1)/p^2$ in Proposition 1.4. Since the interesting q's are less than 2, the interesting p's are greater than 2 and this bound is better than the $1/p$ obtained *via* Hölder.

B. If μ is symmetric we may argue as follows: let $\mathcal{A}$ be a $\sigma(X_s, u \leq s \leq t)$ measurable subset and denote by R_t the time reversal at time t. Then

$$\mathbb{P}_\nu(\mathcal{A}) = \mathbb{E}_\mu\left[\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(X_0)\mathbb{1}_{\mathcal{A}}\right] = \mathbb{E}_\mu\left[\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(X_t)\mathbb{1}_{\mathcal{A}}\circ R_t\right] = \mathbb{E}_\mu\left[\left(P_u\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right)(X_{t-u})\mathbb{1}_{\mathcal{A}}\circ R_t\right].$$

If V is centered and bounded by 1, the set $\left\{\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s \geq R\right\}$ is included in

$$\mathcal{A} := \left\{\frac{1}{t}\int_u^t V(X_s)\mathrm{d}s \geq (R-(u/t))\right\}$$

to which we may apply the previous trick.

In particular, if the semi-group is ultracontractive (*i.e.* there exists some $u > 0$ such that P_u is mapping continuously $\mathbb{L}^1(\mu)$ in $\mathbb{L}^\infty(\mu)$) we obtain a nice bound. Notice that P_v is also mapping continuously $\mathbb{L}^1$ in $\mathbb{L}^2$ for some $v \le u$, so that using reversibility again we may directly use Section 1, with a possible better constant.

If the semi-group is only hypercontractive, *i.e.* if μ satisfies some log-Sobolev inequality, we know that relative entropy is exponentially decaying. Denote by $H(h) := \int h \log h \mathrm{d}\mu$ for any density of probability, and by $h_\nu := \mathrm{d}\nu/\mathrm{d}\mu$. If $H(h_\nu) < +\infty$ it holds $H(P_u h_\nu) \le \mathrm{e}^{-u/C_{LS}} H(h_\nu)$. It is easily seen that

$$\int \exp\left(\mathbb{1}_{\mathcal{A}} - (\mathrm{e}-1)\mathbb{P}_\mu(\mathcal{A})\right) \mathrm{d}\mathbb{P}_\mu \le 1$$

so that using the variational definition of H and reversibility again we get

$$\mathbb{P}_\nu(\mathcal{A}) = \mathbb{E}_\mu\left[P_u h_\nu(X_u)(\mathbb{1}_{\mathcal{A}} - (e-1)\mathbb{P}_\mu(\mathcal{A}) + (e-1)\mathbb{P}_\mu(\mathcal{A}))\right] \le H(P_u h_\nu) + (e-1)\mathbb{P}_\mu(\mathcal{A}).$$

Choosing $u = Rt/2$ we thus obtain

$$\mathbb{P}_\nu(\mathcal{A}) \le (\mathrm{e}-1)\mathrm{e}^{-\frac{tR^2}{32 C_P \mathrm{Var}_\mu(V)}} + H(h_\nu)\mathrm{e}^{-\frac{tR}{2C_{LS}}}.$$

If the semi-group is only Orlicz-hypercontractive in the sense of [4] we do not know whether it is possible to extend the argument to a little bit more integrable initial densities or not. Indeed we did not find the ad-hoc quantity replacing relative entropy.

C. Finally we shall see on a family of examples what can happen when ν is no more absolutely continuous with respect to μ. Actually we shall consider on $\mathbb{R}^d$ a diffusion process

$$X_t^x = x + B_t - \int_0^t \nabla U(X_s^x)\mathrm{d}s,$$

where $x \in \mathbb{R}^d$ and $B_.$ is a standard Brownian motion. We shall assume that U is C^3, and that there exists some function ψ going to $+\infty$ when $|x| \to \infty$ so that $\frac{1}{2}\Delta\psi - \nabla U.\nabla\psi$ is bounded from above. These assumptions ensure the existence of an unique non explosive strong solution. Furthermore the underlying Markov process $X_.$ is μ symmetric for $\mathrm{d}\mu = Z^{-1}e^{-2U}\mathrm{d}x$ where Z is a normalizing constant.

For such a process it is known that the law of X_t^x is absolutely continuous w.r.t. μ. We shall denote by h_t^x its density.

If $|\nabla U|^2(y) - \Delta U(y) \ge -C_m > -\infty$ for all y, one can show that $\int h_t^x \log_+^p(h_t^x)\mathrm{d}\mu < +\infty$ for all $p \ge 1$ (see [7] Prop. 5.1), so that in particular, if the semi-group is hypercontractive (or ultracontractive) we may apply the ideas in **B**.

Actually one can expect a much better integrability and it is shown in [7] Section 5.2 that for $U(y) = |y|^q$ with $1 \le q \le 2$, $h_t^x \in \mathbb{L}^\infty(\mu)$ for all $t > 0$ (U is not C^3 but all the previous discussion is still available).

Indeed we discovered with the help of P.A. Zitt that actually, with our previous assumptions, $h_t^x \in \mathbb{L}^2(\mu)$.

To prove it, as in [7] we follow the idea of [22] Theorem 3.2.7. Replacing the convex γ therein by $\gamma(y) = y^2$ we obtain

$$\int (h_t^x)^2 \mathrm{d}\mu \le Z\mathrm{e}^{2U(x)}\mathbb{E}\left[\mathrm{e}^{-2v(B_t)}\mathrm{e}^{-\frac{1}{2}\int_0^t [|\nabla U|^2 - \Delta U](B_s)\mathrm{d}s}\right] \le Z\mathrm{e}^{2U(x)}\mathrm{e}^{\frac{1}{2}C_m t}$$

where $\mathrm{e}^{-2v(y)} = (2\pi t)^{-\mathrm{d}/2}\ \mathrm{e}^{-|y-x|^2/2t}$. Hence we may directly apply the results in the first two sections.

Acknowledgements. We thank an anonymous referee whose comments have improved the presentation of the paper.

References

[1] S. Aida, Uniform positivity improving property, Sobolev inequalities and spectral gaps. *J. Funct. Anal.* **158** (1998) 152–185.

[2] D. Bakry, L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on Probability theory. École d'été de Probabilités de St-Flour 1992*, *Lect. Notes Math.* **1581** (1994) 1–114.

[3] F. Barthe, P. Cattiaux and C. Roberto, Concentration for independent random variables with heavy tails. *AMRX* **2005** (2005) 39–60.

[4] F. Barthe, P. Cattiaux and C. Roberto, Interpolated inequalities between exponential and Gaussian, Orlicz hypercontractivity and isoperimetry. *Rev. Mat. Iber.* **22** (2006) 993–1067.

[5] F. Barthe, P. Cattiaux and C. Roberto, *Isoperimetry between exponential and Gaussian. EJP* **12** (2007) 1212–1237.

[6] W. Bryc and A. Dembo, Large deviations for quadratic functionals of gaussian processes. *J. Theoret. Prob.* **10** (1997) 307–332.

[7] P. Cattiaux, I. Gentil and G. Guillin, Weak logarithmic-Sobolev inequalities and entropic convergence. *Prob. Theory Related Fields* **139** (2007) 563–603.

[8] E.B. Davies, *Heat kernels and spectral theory.* Cambridge University Press (1989).

[9] J.D. Deuschel and D.W. Stroock, *Large Deviations.* Academic Press, London, *Pure Appl. Math.* **137** (1989).

[10] H. Djellout, A. Guillin and L. Wu, Transportation cost information inequalities for random dynamical systems and diffusions. *Ann. Prob.* **334** (2002) 1025–1028.

[11] P. Doukhan, *Mixing, Properties and Examples.* Springer-Verlag, *Lect. Notes Statist.* **85** (1994).

[12] B. Franchi, Weighted Sobolev-Poincaré inequalities and pointwise estimates for a class of degenerate elliptic equations. *T.A.M.S.* **327** (1991) 125–158.

[13] F.Z. Gong and F.Y. Wang, Functional inequalities for uniformly integrable semigroups and applications to essential spectrums. *Forum Math.* **14** (2002) 293–313.

[14] C. Léonard, Convex conjugates of integral functionals. *Acta Math. Hungar.* **93** (2001) 253–280.

[15] C. Léonard, Minimizers of energy functionals. *Acta Math. Hungar.* **93** (2001) 281–325.

[16] P. Lezaud, Chernoff and Berry-Eessen inequalities for Markov processes. *ESAIM Probab. Statist.* **5** (2001) 183–201.

[17] G. Lu, Weighted Poincaré and Sobolev inequalities for vector fields satisfying Hörmander's condition and applications. *Rev. Mat. Iber.* **8** (1992) 367–439.

[18] E. Rio, *Théorie asymptotique des processus aléatoires faiblement dépendants.* Springer-Verlag, *Math. Appl.* **31** (2000).

[19] R.T. Rockafellar, Integrals which are convex functionals. *Pacific J. Math.* **24** (1968) 525–539.

[20] R.T. Rockafellar, Integrals which are convex functionals II. *Pacific J. Math.* **39** (1971) 439–469.

[21] M. Röckner and F.Y. Wang, Weak Poincaré inequalities and L^2-convergence rates of Markov semigroups. *J. Funct. Anal.* **185** (2001) 564–603.

[22] G. Royer, *Une initiation aux inégalités de Sobolev logarithmiques.* S.M.F., Paris (1999).

[23] F.Y. Wang, Functional inequalities for empty essential spectrum. *J. Funct. Anal.* **170** (2000) 219–245.

[24] L. Wu, A deviation inequality for non-reversible Markov process. *Ann. Inst. Henri Poincaré. Prob. Stat.* **36** (2000) 435–445.