A note on functional inequalities for some Lévy processes

D. Chafaï & F. Malrieu

University of Toulouse

March 2002, compiled September 27, 2002

Abstract

By using a $\mathbf{\Gamma_2}$ approach like in [AL00], we establish a modified logarithmic Sobolev inequality for the law at time t for some Lévy processes, which yields Poissonian confidence intervals in a Monte-Carlo method by means of a concentration inequality for Lipschitz functions. Actually, this result was already obtained in [Wu00] via a martingale representation (Clark-Ocône formulae).

It is well known that the law at time t of a standard Brownian motion $(B_t)_{t\geqslant 0}$ on $\mathbb{R}^d$ satisfies to a logarithmic Sobolev inequality with constant $2t$ for the usual Euclidean gradient, cf. [Gro75]. The independent increments of the Brownian motion and the tensorisation property of such inequalities allows to extend this result to path spaces for the Wiener measure by replacing the gradient on $\mathbb{R}^d$ with the Malliavin derivative. Recently, such inequalities were extended in a suitable modified form to some continuous time random walks on graphs in [AL00], to the Poisson point process in [AL00] and [Wu00], and to normal martingales in [Pri00]. It is then quite natural to ask if such inequalities hold for diffusions with jumps. In this direction, we are interested in particular Lévy processes $(X_t)_{t\geqslant 0}$ on $\mathbb{R}^d$ with infinitesimal generator $\mathbf{L}$ of the form

$$\mathbf{L} := \sigma\,\mathbf{\Delta} + b\cdot\nabla + \lambda\,(K - I), \tag{1}$$

where $(\sigma,\lambda,b)\in\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}^d$ and K is defined for any $f:\mathbb{R}^d\to\mathbb{R}$ by

$$(Kf)(x) := \int_{\mathbb{R}^d} f(x+y)\,\nu(dy),$$

where ν is a fixed probability measure on $\mathbb{R}^d$. In other words, $\mathbf{L}$ acts on a smooth function $f:\mathbb{R}^d\to\mathbb{R}$ as follows:

$$(\mathbf{L}f)(x) = \sigma\,(\mathbf{\Delta}f)(x) + b\cdot(\nabla f)(x) + \lambda\int_{\mathbb{R}^d}(f(x+y)-f(x))\,\nu(dy).$$

The process $(X_t)_{t\geqslant 0}$ can be viewed as the independent sum $X_t = \sqrt{2\sigma}\,B_t + tb + Y_t$, where $(B_t)_{t\geqslant 0}$ is a standard Brownian motion and $(Y_t)_{t\geqslant 0}$ is a compound Poisson process with jump kernel K and jumping rate λ. In other words, $Y_t = Z_{N_t}$ where $(Z_n)_{n\in\mathbb{N}^*}$ is a Markov chain with i.i.d. increments of law ν and $(N_t)_{t\geqslant 0}$ is a simple Poisson point process of intensity λ

independent of $(Z_n)_{n\in\mathbb{N}^*}$. As for any Markov process, one can define the associated Markov semi-group $(\mathbf{P}_t)_{t\geqslant 0}$ given by

$$\mathbf{P}_t(f)\,(x) := \mathbb{E}(f(X_t)\,|\,X_0 = x),$$

and the "carré du champ" operator $\mathbf{\Gamma}$ given by

$$\begin{aligned}\mathbf{\Gamma}(f)(x) &:= \frac{1}{2}\left(\mathbf{L}(f^2) - 2\,f\mathbf{L}f\right)(x)\\ &= \sigma\,|\nabla f|^2(x) + \frac{\lambda}{2}\int_{\mathbb{R}^d}(f(x+y)-f(x))^2\,\nu(dy).\end{aligned}$$

We establish modified logarithmic Sobolev inequalities for $\mathbf{P}_t(\cdot)\,(x)$ uniformly in $x\in\mathbb{R}^n$ with constant $2t$, as stated in the following theorem.

Theorem 1. *For any $t>0$, and any positive smooth function $f:\mathbb{R}^d\to\mathbb{R}$,*

$$\mathbf{P}_t(f\log f) - \mathbf{P}_t(f)\log\mathbf{P}_t(f) \leqslant 2t\,\mathbf{P}_t\left(\frac{\mathbf{\Gamma}f}{f}\right). \tag{2}$$

Notice that the constant $2t$ is not optimal when the jump part is not present (i.e. λ=0) and may be then replaced in this case by t. When ν is compactly supported, we are able to use the techniques developed in [BL98] (see also [Led99]) to deduce the following concentration result for Lipschitz functions of X_t.

Corollary 1. *Let $(X_t)_{t\geqslant 0}$ be the Lévy process on $\mathbb{R}^d$ generated by (1). Assume that ν is compactly supported with* $\mathrm{supp}(\nu)\subset B(0,K)$*, then, for any $t>0$, any $r>0$ and any 1-Lipschitz function f,*

$$\mathbb{P}(|f(X_t)-\mathbb{E}f(X_t)|\geqslant r)\leqslant 2\,\exp\left(-\frac{r}{8K}\log\left(1+\frac{2Kr}{t(\sigma+\lambda K^2)}\right)\right). \tag{3}$$

This inequality for the law of X_t expresses a Gaussian concentration for small values of r and a Poissonian concentration for large values of r. It gives exact confidence intervals for a Monte-Carlo method, as presented in Section 4.

Finally, since $(X_t)_{t\geqslant 0}$ has independent increments, inequality (2) can be tensorised to cylindrical functions of the process $(X_t)_{t\geqslant 0}$, as stated in the following theorem. It can be viewed as an extension of certain results of [AL00] stated to derive local inequalities on path spaces.

Theorem 2. *Let $(X_t)_{t\geqslant 0}$ be the Lévy process generated by (1). For any smooth function F of $(X_{t_1},\ldots,X_{t_n})$ where $0=t_0<t_1<\cdots<t_n$:*

$$\mathbf{E}(F\log F) - \mathbf{E}(F)\log\mathbf{E}(F) \leqslant 2\sum_{i=1}^{n}(t_i-t_{i-1})\,\mathbf{E}\left(\frac{\mathbf{\Gamma}^{i\cdots n}F}{F}\right), \tag{4}$$

where

$$\mathbf{\Gamma}^{i\cdots n}F(x) := \sigma\left|\sum_{j=i}^{n}\nabla^j F(x)\right|^2 + \frac{\lambda}{2}\int_{\mathbb{R}^d}(F\circ\tau_i(y)(x)-F(x))^2\nu(dy),$$

and

$$F\circ\tau_i(y)(x) := F(x_1,\ldots,x_{i-1},x_i+y,\ldots,x_n+y).$$

1 Proof of Theorem 1

Towards the derivation of the local inequality (2), we introduce the $\mathbf{\Gamma_2}$ operator which is constructed from $\mathbf{L}$ and $\mathbf{\Gamma}$:

$$\mathbf{\Gamma_2} f := \frac{1}{2}\left(\mathbf{\Gamma}(f,g) - 2\,\mathbf{\Gamma}(f,\mathbf{L}g)\right),$$

where $\mathbf{\Gamma}(f,g) := (\mathbf{L}(fg) - f\mathbf{L}g - g\mathbf{L}f))/2$ for any smooth functions $f, g : \mathbb{R}^d \to \mathbb{R}$. In the sequel, the exponent $^{(1)}$ (resp. $^{(2)}$) always refers to the continuous part (resp. jump part) of the generator (1). For example, $\mathbf{\Gamma}(f,g)$ is the sum of

$$\mathbf{\Gamma}^{(1)}(f,g)(x) := \sigma\, \nabla f(x) \cdot \nabla g(x)$$

and of

$$\mathbf{\Gamma}^{(2)}(f,g)(x) := \frac{\lambda}{2} \int_{\mathbb{R}^d} (f(x+y) - f(x))(g(x+y) - g(x))\, \nu(dy).$$

At this stage, an easy calculus gives that,

$$\mathbf{\Gamma_2} f = \mathbf{\Gamma_2}^{(1)}(f) + \mathbf{\Gamma_2}^{(2)}(f) + \sigma\lambda \int_{\mathbb{R}^d} |\nabla f(x+y) - \nabla f(x)|^2\, \nu(dy),$$

where $\mathbf{\Gamma_2}^{(1)} f := \sigma^2 \left\|\mathrm{Hess}(f)\right\|_2^2$ and

$$\mathbf{\Gamma_2}^{(2)}(f)(x) := \frac{\lambda^2}{4} \int_{\mathbb{R}^d} \left[f(x+y+z) - f(x+y) - f(x+z) + f(x)\right]^2 \nu(dy)\nu(dz).$$

One can notice that $\mathbf{\Gamma}$ is always non negative, which is a general property of Markov processes. On the other hand, the positivity of $\mathbf{\Gamma_2}$ yields local Poincaré inequalities for $\mathbf{P}_t$,

$$\mathbf{P}_t\!\left(f^2\right) - \mathbf{P}_t(f)^2 \leqslant 2t\, \mathbf{P}_t(\mathbf{\Gamma} f)\,. \tag{5}$$

More generally, the celebrated Bakry-Émery criterion $\mathbf{\Gamma_2} \geqslant \rho\mathbf{\Gamma}$ implies a Poincaré inequality with constant $\rho^{-1}\left(1 - \exp(-2\rho t)\right)$, cf. [Bak97]. In our case, since linear functions have a null $\mathbf{\Gamma_2}$ and a positive $\mathbf{\Gamma}$, the best constant ρ is 0. Nevertheless, we are able to prove the more precise bound

$$\mathbf{\Gamma_2} \geqslant \mathbf{\Gamma}\left(\sqrt{\mathbf{\Gamma}}\right), \tag{6}$$

which yields the local modified logarithmic Sobolev inequality (2), improving in this way the local Poincaré inequality (5).

Proof of (6). The chain rule formula $\mathbf{\Gamma}^{(1)}(\Phi(f)) = \Phi'^2(f)\, \mathbf{\Gamma}^{(1)} f$ and the Cauchy-Schwarz inequality yield the bound

$$\mathbf{\Gamma_2}^{(1)} \geqslant \mathbf{\Gamma}^{(1)}\left(\sqrt{\mathbf{\Gamma}^{(1)}}\right). \tag{7}$$

Unfortunately, there is no chain rule formula for $\mathbf{\Gamma}^{(2)}$. Nevertheless, we are able to show that the bound (6) still holds. The quantity to estimate is given by :

$$\mathbf{\Gamma}\left(\sqrt{\mathbf{\Gamma} f}\right)(x) = \sigma \left|\nabla \sqrt{\mathbf{\Gamma} f(x)}\right|^2 + \frac{\lambda}{2} \int \left(\sqrt{\mathbf{\Gamma} f(x+y)} - \sqrt{\mathbf{\Gamma} f(x)}\right)^2 \nu(dy).$$

The triangle inequality yields for any α and β

$$\begin{aligned}&\left|\sqrt{\mathbf{\Gamma} f(\alpha)}-\sqrt{\mathbf{\Gamma} f(\beta)}\right|\\&\qquad\leqslant\sqrt{\sigma\left|\nabla f(\alpha)-\nabla f(\beta)\right|^{2}+\frac{\lambda}{2}\int\left[f(\alpha+z)-f(\alpha)-f(\beta+z)+f(\beta)\right]^{2}\nu(dz)}.\end{aligned}$$

Hence, the second part of $\mathbf{\Gamma}\big(\sqrt{\mathbf{\Gamma} f}\big)(x)$ is bounded above by

$$\frac{\sigma\lambda}{2}\int\left|\nabla f(x+y)-\nabla f(x)\right|^{2}\nu(dy)+\mathbf{\Gamma}_{\mathbf{2}}^{(2)}(f)(x).$$

On the other hand, the first part of $\mathbf{\Gamma}\big(\sqrt{\mathbf{\Gamma} f}\big)(x)$ is equal to

$$\frac{\sigma}{4}\,\frac{\left|\nabla\mathbf{\Gamma}^{(1)}f(x)+\lambda\int(\nabla f(x+y)-\nabla f(x))(f(x+y)-f(x))\,\nu(dy)\right|^{2}}{\mathbf{\Gamma}^{(1)}f(x)+\mathbf{\Gamma}^{(2)}f(x)}.$$

The Cauchy-Schwarz inequality yields that the numerator of the latter is bounded above by

$$\left(\sqrt{\mathbf{\Gamma}^{(1)}f(x)}\,\frac{\left|\nabla\mathbf{\Gamma}^{(1)}f(x)\right|}{\sqrt{\mathbf{\Gamma}^{(1)}f(x)}}+\sqrt{\mathbf{\Gamma}^{(2)}f(x)}\,\sqrt{2\lambda\int\left|\nabla f(x+y)-\nabla f(x)\right|^{2}\nu(dy)}\right)^{2}.$$

Again by the Cauchy-Schwarz inequality, this expression is bounded above by

$$\left(\mathbf{\Gamma}^{(1)}f(x)+\mathbf{\Gamma}^{(2)}f(x)\right)\left(\frac{\left|\nabla\mathbf{\Gamma}^{(1)}f(x)\right|^{2}}{\mathbf{\Gamma}^{(1)}f(x)}+2\lambda\int\left|\nabla f(x+y)-\nabla f(x)\right|^{2}\nu(dy)\right).$$

But now, observe that by virtue of (7),

$$\frac{\sigma}{4}\,\frac{\left|\nabla\mathbf{\Gamma}^{(1)}f(x)\right|^{2}}{\mathbf{\Gamma}^{(1)}f(x)}=\mathbf{\Gamma}^{(1)}\left(\sqrt{\mathbf{\Gamma}^{(1)}}\right)(x)\leqslant\mathbf{\Gamma}_{\mathbf{2}}^{(1)}(f)(x),$$

and therefore the first term of $\mathbf{\Gamma}\big(\sqrt{\mathbf{\Gamma} f}\big)(x)$ is bounded above by

$$\mathbf{\Gamma}_{\mathbf{2}}^{(1)}(f)(x)+\frac{\sigma\lambda}{2}\int_{\mathbb{R}^d}\left|\nabla f(x+y)-\nabla f(x)\right|^{2}\nu(dy).$$

Summarising, we get that $\mathbf{\Gamma}\big(\sqrt{\mathbf{\Gamma} f}\big)(x)$ is bounded above by

$$\mathbf{\Gamma}_{\mathbf{2}}^{(1)}(f)(x)+\mathbf{\Gamma}_{\mathbf{2}}^{(2)}(f)(x)+\sigma\lambda\int_{\mathbb{R}^d}\left|\nabla f(x+y)-\nabla f(x)\right|^{2}\nu(dy),$$

which is exactly $\mathbf{\Gamma}_{\mathbf{2}}(f)(x)$. □

For the sake of completeness, let us recall the derivation of (2) from (6), which is taken from [AL00].

Proof of (6) $\Rightarrow$ (2). Since $\mathbf{P}_t(f\log f) - \mathbf{P}_t(f)\log \mathbf{P}_t(f) = \alpha(t) - \alpha(0)$, where

$$\alpha(s) := \mathbf{P}_s(\mathbf{P}_{t-s}f \log \mathbf{P}_{t-s}f)$$

for any s in $[0,t]$, we just have to control α'. With the notation $g = \mathbf{P}_{t-s}f$, we get

$$\alpha'(s) = \mathbf{P}_s(\mathbf{L}(g\log g) - (1+\log g)\mathbf{L}g)\,.$$

Now, $\mathbf{L}(g\log g)(x) - (1+\log g(x))\mathbf{L}g(x)$ is equal to

$$\sigma\,\frac{|\nabla g(x)|^2}{g(x)} + \lambda \int \{g(x+y)[\log g(x+y) - \log g(x)] - [g(x+y) - g(x)]\}\,\nu(dy).$$

Then, using the inequality $\log b - \log a \leqslant (b-a)/a$, which is true for $a, b > 0$, we get

$$\mathbf{L}(g\log g) - (1+\log g)\mathbf{L}g \leqslant 2\,\frac{\mathbf{\Gamma} g}{g},$$

(notice that the absolute constant 2 above may be removed when there is no jump part, i.e. when $\lambda = 0$, and thus, the preceding inequality is not optimal in the continuous part.). As a consequence,

$$\alpha'(s) \leqslant 2\,\mathbf{P}_s\left(\frac{\mathbf{\Gamma}\mathbf{P}_{t-s}f}{\mathbf{P}_{t-s}f}\right).$$

At this step, we need a commutation relation between $\sqrt{\mathbf{\Gamma}}$ and $\mathbf{P}_t$. Consider the function β defined on $[0,t]$ by:

$$\beta(s) := \mathbf{P}_s\left(\sqrt{\mathbf{\Gamma}\mathbf{P}_{t-s}f}\right).$$

Then, with the notation $g = \mathbf{P}_{t-s}f$, its derivative is equal to:

$$\begin{aligned}
\beta'(s) &= \mathbf{P}_s\left(\mathbf{L}\left(\sqrt{\mathbf{\Gamma} g}\right) - \frac{1}{\sqrt{\mathbf{\Gamma} g}}\mathbf{\Gamma}(g, \mathbf{L}g)\right) \\
&= \mathbf{P}_s\left(\frac{1}{2\sqrt{\mathbf{\Gamma} g}}\left[2\sqrt{\mathbf{\Gamma} g}\,\mathbf{L}\left(\sqrt{\mathbf{\Gamma} g}\right) - 2\mathbf{\Gamma}(g, \mathbf{L}g)\right]\right) \\
&= \mathbf{P}_s\left(\frac{1}{\sqrt{\mathbf{\Gamma} g}}\left[\mathbf{\Gamma_2} g - \mathbf{\Gamma}\left(\sqrt{\mathbf{\Gamma} g}\right)\right]\right) \geqslant 0.
\end{aligned}$$

Therefore, the function β increases

$$\sqrt{\mathbf{\Gamma}\mathbf{P}_t f} \leqslant \mathbf{P}_t\left(\sqrt{\mathbf{\Gamma} f}\right).$$

This commutation relation provides:

$$\alpha'(s) \leqslant 2\,\mathbf{P}_s\left(\frac{\left(\mathbf{P}_{t-s}(\sqrt{\mathbf{\Gamma} f})\right)^2}{\mathbf{P}_{t-s}f}\right).$$

At last, the Cauchy-Schwarz inequality

$$(\mathbf{P}_{t-s}X)^2 \leqslant \mathbf{P}_{t-s}(X^2/Y)\mathbf{P}_{t-s}(Y)$$

leads to

$$\alpha'(s) \leqslant 2\,\mathbf{P}_t\left(\frac{\mathbf{\Gamma} f}{f}\right),$$

which achieves the proof. $\square$

Remark 1. Notice that it is perhaps more easy to derive the $\mathbf{\Gamma_2}$ bounds separately:

$$\mathbf{\Gamma_2}^{(1)} \geqslant \mathbf{\Gamma}^{(1)}\Big(\sqrt{\mathbf{\Gamma}^{(1)}}\Big) \quad \text{and} \quad \mathbf{\Gamma_2}^{(2)} \geqslant \mathbf{\Gamma}^{(2)}\Big(\sqrt{\mathbf{\Gamma}^{(2)}}\Big).$$

This yields to the following local inequalities for any $t > 0$ and any smooth $f : \mathbb{R}^d \to \mathbb{R}$:

$$\mathbf{Ent}(f(W_t)) \leqslant t\,\mathbb{E}\bigg(\frac{\mathbf{\Gamma}^{(1)} f}{f}(W_t)\bigg),$$

and

$$\mathbf{Ent}(f(Y_t)) \leqslant 2t\,\mathbb{E}\bigg(\frac{\mathbf{\Gamma}^{(2)} f}{f}(Y_t)\bigg),$$

where $\mathbf{Ent}(F) := \mathbb{E}(F\log F) - \mathbb{E}(F)\log\mathbb{E}(F)$. The first one, which concerns the Brownian motion $W_t := \sqrt{2\sigma}\,B_t + tb$, is well known and can be viewed as a direct consequence of the logarithmic Sobolev inequality for the standard Gaussian measure. Since $(X_t)_{t\geqslant 0}$ is an independent sum of the Brownian part $(W_t)_{t\geqslant 0}$ and the Poisson part $(Y_t)_{t\geqslant 0}$, the couple of inequalities above gives by tensorisation

$$\mathbf{Ent}(f(W_t, Y_t)) \leqslant t\,\mathbb{E}\bigg(\frac{\mathbf{\Gamma}^{(1)} f}{f}(W_t, Y_t) + 2\,\frac{\mathbf{\Gamma}^{(2)} f}{f}(W_t, Y_t)\bigg).$$

Finally, the latter yields (2) by taking $f(x, y)$ of the form $f(x + y)$.

Remark 2. One can ask if the bound (6) holds when b is not constant, for example when $(X_t)_{t\geqslant 0}$ is the Markov process generated by $\mathbf{\Delta} - \nabla U \cdot \nabla + \lambda\,(K - I)$. In this case, $(X_t)_{t\geqslant 0}$ is not the independent sum of the diffusive part and the jump part. The additional term in $\mathbf{\Gamma_2}$ is

$$\langle \nabla f(x), \mathrm{Hess}(U)(x)\nabla f(x)\rangle + \lambda \int_{\mathbb{R}^d} \nabla f(x+y)(\nabla U(x+y) - \nabla U(x))(f(x+y) - f(x))\,\nu(dy).$$

Unfortunately, even in the case where $U(x) = |x|^2/2$ and $\nu = \delta_1$, due to the presence of the crossing term, the associated $\mathbf{\Gamma_2}$ operator can be negative. Notice that

$$\varepsilon^{-1} \int_{\mathbb{R}^d} (f(x+\varepsilon y) - f(x))\,\nu(dy) \underset{\varepsilon\to 0^+}{\longrightarrow} \nabla f(x) \cdot \int_{\mathbb{R}^d} y\,\nu(dy),$$

and then, the jump part acts asymptotically like a constant drift in the generator. Hence, one can see the jump part in general like a drift perturbation of the diffusive part of the generator. The behavior of the local inequalities under bounded perturbations of a convex drift is still an unsolved problem. On the other hand, one can ask if $\mathbf{\Gamma_2}$ remains bounded below when the jump kernel $K(x)$ is not of the form $\delta_x * \nu$. In this case, we have by denoting ν_x the probability measure associated to $K(x)$:

$$\mathbf{L}^{(2)} f(x) = \lambda \int_{\mathbb{R}^d} f(y)\,\nu_x(dy) - f(x),$$

and

$$\mathbf{\Gamma}^{(2)} f(x) = \frac{\lambda}{2} \int_{\mathbb{R}^d} [f(y) - f(x)]^2\,\nu_x(dy),$$

and

$$\mathbf{\Gamma_2}^{(2)} = \frac{\lambda^2}{4} \iiint [f(z) - f(y) - f(u) + f(x)]^2 \, \nu_y(dz) \, \nu_x(dy) \, \nu_x(du)$$
$$- 2 \iiint f(z)[f(y) - f(u)] \, \nu_y(dz) \, \nu_x(dy) \, \nu_x(du).$$

As we can see, $\mathbf{\Gamma_2}^{(2)}$ is not positive in general. Nevertheless, when $\nu_x = \delta_{\alpha(x)}$, the bound $\mathbf{\Gamma_2}^{(2)} \geqslant \mathbf{\Gamma}^{(1)}\Big(\sqrt{\mathbf{\Gamma}^{(1)}}\Big)$ still holds, for any function $x \mapsto \alpha(x)$.

2 Proof of Theorem 2

For simplicity, we give the proof for $n = 2$ and $(t_1, t_2) = (s, t)$. By denoting $\Phi(u) := u \log u$, we have that $\mathbb{E}(\Phi(F)) - \Phi(\mathbb{E}(F))$ is equal to

$$\mathbb{E}[\mathbb{E}(\Phi(F) \,|\, X_s) - \Phi(\mathbb{E}(F \,|\, X_s))] + \mathbb{E}(\Phi(\mathbb{E}(F \,|\, X_s))) - \Phi(\mathbb{E}(\mathbb{E}(F \,|\, X_s))). \tag{8}$$

In other words, $\mathbf{Ent}(X) = \mathbb{E}(\mathbf{Ent}(X|Y)) + \mathbf{Ent}(\mathbb{E}(X|Y))$. If $G_x(y) := F(x, x+y)$, we have by virtue of (2),

$$\mathbb{E}(\Phi(G_x(X_{t-s}))) - \Phi(\mathbb{E}(G_x(X_{t-s}))) \leqslant 2(t-s)\,\mathbb{E}\left(\frac{(\mathbf{\Gamma} G_x)(X_{t-s})}{G_x(X_{t-s})}\right).$$

But now,

$$\mathbf{\Gamma}(G_x)(y) = \sigma \left|\partial_2 F(x, x+y)\right|^2 + \frac{\lambda}{2} \int_{\mathbb{R}^d} |F(x, x+y+z) - F(x, x+y)|^2 \, \nu(dz).$$

Therefore, since $X_t = X_s + X_t - X_s$ where $X_t - X_s$ has the same law than X_{t-s} and is independent of X_s, the first term of (8) can be bounded above as

$$\mathbb{E}[\mathbb{E}(\Phi(F) \,|\, X_s) - \Phi(\mathbb{E}(F \,|\, X_s))] \leqslant 2(t-s)\,\mathbb{E}\left(\frac{\mathbf{\Gamma}^{2\cdots 2} F}{F}\right).$$

On the other hand, if we define H_{t-s} by

$$H_{t-s}(x) := \mathbb{E}(F(x, x + X_{t-s})),$$

we have by virtue of (2),

$$\mathbb{E}(\Phi(H_{t-s}(X_s))) - \Phi(\mathbb{E}(H_{t-s}(X_s))) \leqslant 2s\,\mathbb{E}\left(\frac{(\mathbf{\Gamma} H_{t-s})(X_s)}{H_{t-s}(X_s)}\right).$$

Now, the Cauchy-Schwarz inequality $\mathbb{E}(Z)^2 \leqslant \mathbb{E}(Z^2/Y)\,\mathbb{E}(Y)$ gives,

$$\begin{aligned}
(\mathbf{\Gamma} H_{t-s})(x) &= \sigma \left[\mathbb{E}(\partial_1 F(x, x+X_{t-s}) + \partial_2 F(x, x+X_{t-s}))\right]^2 \\
&\quad + \frac{\lambda}{2} \int_{\mathbb{R}^d} \left[\mathbb{E}(F(x+z, x+z+X_{t-s}) - F(x, x+X_{t-s}))\right]^2 \nu(dz) \\
&\leqslant \sigma\,\mathbb{E}\left(\frac{|\partial_1 F(x, x+X_{t-s}) + \partial_2 F(x, x+X_{t-s})|^2}{F(x, x+X_{t-s})}\right) H_{t-s}(x) \\
&\quad + \frac{\lambda}{2} \int_{\mathbb{R}^d} \mathbb{E}\left(\frac{|F(x+z, x+z+X_{t-s}) - F(x, x+X_{t-s})|^2}{F(x, x+X_{t-s})}\right) H_{t-s}(x)\,\nu(dz).
\end{aligned}$$

Therefore, the last term of (8) is bounded as

$$\mathbb{E}(\Phi(\mathbb{E}(F \,|\, X_s))) - \Phi(\mathbb{E}(\mathbb{E}(F \,|\, X_s))) \leqslant 2s\,\mathbb{E}\left(\frac{\mathbf{\Gamma}^{1\cdots 2} F}{F}\right).$$

This achieves the proof.

3 Proof of Corollary 1

For any smooth function $f : \mathbb{R}^d \to \mathbb{R}$, one have by the definition of $\mathbf{\Gamma}$:

$$\mathbf{\Gamma}(e^f)(x) = \sigma\, e^{2f(x)}|\nabla f|^2(x) + \frac{\lambda}{2}\, e^{2f(x)} \int_{\mathbb{R}^d} \big(e^{f(x+y)-f(x)} - 1\big)^2\, \nu(dy).$$

Hence, if $\|f\|_{\mathrm{Lip}} \leqslant \alpha$, we get

$$\left(\frac{\mathbf{\Gamma}(e^f)}{e^f}\right)(x) \leqslant \sigma\alpha^2\, e^{f(x)} + \frac{\lambda}{2} e^{f(x)} \int_{\mathbb{R}^d} \big(e^{f(x+y)-f(x)} - 1\big)^2\, \nu(dy).$$

Now, the inequality $(e^u - 1)^2 \leqslant u^2 e^{2|u|}$ yields,

$$\left(\frac{\mathbf{\Gamma}(e^f)}{e^f}\right)(x) \leqslant \left[\sigma\alpha^2 + \frac{\lambda}{2}\, K^2\alpha^2 e^{2\alpha K}\right] e^{f(x)} \leqslant \left(\sigma + \frac{\lambda K^2}{2}\right)\alpha^2 e^{2K\alpha} e^{f(x)}.$$

Therefore, by virtue of (2), we have obtained that for any smooth Lipschitz function f such that $\|f\|_{\mathrm{Lip}} \leqslant \alpha$,

$$\mathbb{E}\big(f(X_t)e^{f(X_t)}\big) - \mathbb{E}(f(X_t))\mathbb{E}\big(e^{f(X_t)}\big) \leqslant B(\alpha)\, \mathbb{E}\big(e^{f(X_t)}\big),$$

where

$$B(\alpha) := 2t\left(\sigma + \frac{\lambda K^2}{2}\right)\alpha^2 e^{2K\alpha} =: A\alpha^2 e^{B\alpha}.$$

This inequality is known to give the desired concentration bound by the classical Herbst's argument, cf. [Led99, Cor. 2.12]. Namely, for any 1-Lipschitz function g, let

$$H(\beta) := \beta^{-1}\, \mathbb{E}(e^{\beta g(X_t)}).$$

Now, the previous inequality can be simply rewritten as $H'(\beta) \leqslant B(\beta)/\beta^2$. Since $H(0) = \mathbb{E}(g(X_t))$, we get

$$\begin{aligned}\mathbb{E}(e^{\beta g}) &\leqslant \exp\left(\beta\, \mathbb{E}(g(X_t)) + \beta \int_0^\beta \frac{B(u)}{u^2}\, du\right)\\ &= \exp\big(\beta\, \mathbb{E}(g(X_t)) + \beta A B^{-1}(e^{B\beta} - 1)\big).\end{aligned}$$

This yields by Chebychev's inequality,

$$\mathbb{P}(g(X_t) - \mathbb{E}(g(X_t)) \geqslant r) \leqslant \exp\big(-\beta r + \beta A B^{-1}(e^{B\beta} - 1)\big).$$

The desired result follows by choosing $\beta = r(4A)^{-1}$ when $r \leqslant 4AB^{-1}$, which gives

$$\exp\big(-\beta r + \beta A B^{-1}(e^{B\beta} - 1)\big) \leqslant \exp\big(-\beta r + 2A\beta^2\big) \leqslant \exp\left(-\frac{r^2}{8A}\right),$$

and $\beta = B^{-1}\log(Br/(2A))$ when $r \geqslant 4AB^{-1}$, for which

$$\exp\big(-\beta r + \beta A B^{-1}(e^{B\beta} - 1)\big) \leqslant \exp\left(-\frac{r}{2B}\log\frac{Br}{2A}\right).$$

Remark 3. Actually, the compact support hypothesis for ν can be relaxed, but the concentration inequality is then different. More precisely, one can take

$$B(u) := 2t\, u^2\left(\sigma + \frac{\lambda}{2}\int_{\mathbb{R}^d} |y|^2 e^{2u|y|}\, \nu(dy)\right).$$

This gives various concentration inequalities [Led99], depending on the behavior of ν at infinity. Notice that

$$\int_0^\alpha \frac{B(s)}{s^2}\, ds = 2\sigma\alpha t + \frac{\lambda t}{2}\int_{\mathbb{R}^d} |y|\big(e^{2\alpha|y|} - 1\big)\, \nu(dy).$$

4 Application to the Monte-Carlo method

As we said, the process $(X_t)_{t\geqslant 0}$ can be viewed as the independent sum

$$X_t = \sqrt{2\sigma}\, B_t + tb + Y_t,$$

where $(B_t)_{t\geqslant 0}$ is a standard Brownian motion and $(Y_t)_{t\geqslant 0}$ is a compound Poisson process with jump kernel K and jumping rate λ. This gives a way to simulate X_t for any time t, but in general, the law of X_t is not known explicitly. Therefore, by virtue of the law of large numbers, $\mathbf{E}(f(X_t))$ can be approximated by the empirical mean

$$\frac{1}{n}\left(f(X_t^{(1)})(\omega) + \cdots + f(X_t^{(n)})(\omega)\right),$$

where $X_t^{(1)}, \ldots, X_t^{(n)}$ are n i.i.d. copies of X_t. The problem is then to control the error in terms of all the parameters λ, σ, ν, t and n. The usual methods used in practice are asymptotic (CLT or LDP) or badly depend on n and on the variance of f (Chebychev's or Berry-Essen's bounds). In contrast, when f is Lipschitz, (2) gives *exact* confidence intervals for this approximation method, as stated in the following theorem. Moreover, n can be clearly chosen in a sharp way.

Theorem 3. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be an α-Lipschitz function and $(X_t)_{t\geqslant 0}$ be the Lévy process generated by* (1)*, where ν is compactly supported with* $\operatorname{supp}(\nu) \subset B(0, K)$*. Then, for every $t > 0$, $r > 0$, $n \in \mathbb{N}^*$,*

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} f(X_t^{(i)}) - \mathbb{E}(f(X_t))\right| \geqslant r\right) \;\leqslant\; 2\,\exp\left(-\frac{r\sqrt{n}}{8K\alpha}\log\left(1 + \frac{2Kr\sqrt{n}}{\alpha t(\sigma + \lambda K^2)}\right)\right).$$

where $X_t^{(1)}, \ldots, X_t^{(n)}$ are n i.i.d. copies of X_t.

Proof. The proof is based on the tensorisation of (2), which yields an inequality of the same type of (2) with the same constant for the probability measure $\mathcal{L}(X_t)^{\otimes n}$ on $(\mathbb{R}^d)^{\otimes n}$ with the sums of the $\boldsymbol{\Gamma}$ operators on the n coordinates. The desired result follows with an argument similar to the proof of Corollary 1. □

5 Simpler and stronger

Let $\mathbf{P}_t(f)\,(x) := \mathbf{E}(f(X_t)\,|\,X_0 = x)$ where $(X_t)_{t\geqslant 0}$ is the Markov process with infinitesimal generator $\mathbf{L}$ given by

$$(\mathbf{L}f)(x) := \lambda \int_{\mathbb{R}^d} \mathrm{D}_y f(x)\,\nu(dy),$$

where $\mathrm{D}_y g(x) := f(x+y) - f(x)$ and ν is a Lévy measure (i.e. $\int_{\mathbb{R}^d} (|y|^2 \wedge 1)\,\nu(dy) < +\infty$). We have for any $t > 0$ and any $f : \mathbb{R}^d \to \mathbb{R}$ with $f > 0$:

$$\mathbf{P}_t(f \log f) - \mathbf{P}_t(f)\log \mathbf{P}_t(f) = \alpha(t) - \alpha(0) = \int_0^t \alpha'(s)\,ds,$$

where $\alpha(s) := \mathbf{P}_s(\mathbf{P}_{t-s}(f)\log \mathbf{P}_{t-s}(f))$. But now $\alpha'(s) = \mathbf{P}_s(\mathbf{L}(g\log g) - (1 + \log g)\mathbf{L}g)$ where $g := \mathbf{P}_{t-s}(f)$. At this stage, we notice that

$$\mathbf{L}(g\log g) - (1 + \log g)\mathbf{L}g = \lambda \int_{\mathbb{R}^d} [\mathrm{D}_y\Phi(g) - \Phi'(g)\mathrm{D}_y g]\,\nu(dy),$$

where $\Phi(u) := u \log u$. But $\mathrm{D}_y \Phi(g) - \Phi'(g)\mathrm{D}_y g = \Psi(g, \mathrm{D}_y g)$, where

$$\begin{aligned}\Psi(u,v) : &= \Phi(u+v) - \Phi(u) - \Phi'(u)v \\ &= (u+v)\log(u+v) - u\log u - (1+\log u)v,\end{aligned}$$

for any $(u,v) \in \mathbb{R}^2$ with $u > 0$ and $u+v > 0$. Hence, by the Fubini Theorem,

$$\alpha'(s) = \lambda \int_{\mathbb{R}^d} \mathbf{P}_s(\Psi(g, \mathrm{D}_y g)) \; \nu(dy).$$

Now, since $g = \mathbf{P}_{t-s}(f)$ and since the process have independent increments, we have $\mathrm{D}_y g = \mathrm{D}_y \mathbf{P}_{t-s}(f) = \mathbf{P}_{t-s}(\mathrm{D}_y f)$. Then, by the Jensen inequality for the bivariate convex function Ψ and the probability measure $\mathbf{P}_{t-s}(\cdot)(x) = \mathcal{L}(X_{t-s} \,|\, X_0 = x)$:

$$\Psi(g, \mathrm{D}_y g) = \Psi(\mathbf{P}_{t-s}(f)\,, \mathbf{P}_{t-s}(\mathrm{D}_y f)) \leqslant \mathbf{P}_{t-s}(\Psi(f, \mathrm{D}_y f))\,,$$

Hence, we have:

$$\alpha'(s) \leqslant \lambda \int_{\mathbb{R}^d} \mathbf{P}_s(\mathbf{P}_{t-s}(\Psi(f, \mathrm{D}_y f))) \; \nu(dy) = \lambda \int_{\mathbb{R}^d} \mathbf{P}_t(\Psi(f, \mathrm{D}_y f)) \; \nu(dy).$$

Therefore, again by the Fubini Theorem:

$$\alpha'(s) \leqslant \lambda\, \mathbf{P}_t\left(\int_{\mathbb{R}^d} \Psi(f, \mathrm{D}_y f)\, \nu(dy) \right).$$

Finally, we have:

$$\mathbf{P}_t(f \log f) - \mathbf{P}_t(f) \log \mathbf{P}_t(f) \leqslant \lambda t\, \mathbf{P}_t\left(\int_{\mathbb{R}^d} \Psi(f, \mathrm{D}_y f)\, \nu(dy) \right).$$

This inequality gives two bounds in terms of $(\mathrm{D}_y f)^2/f$ and $\mathrm{D}_y f\, \mathrm{D}_y \log f$ since we have

$$\Psi(u,v) \leqslant \frac{v^2}{u} \quad \text{and} \quad \Psi(u,v) \leqslant v(\log(u+v) - \log u).$$

In other words:

$$\mathbf{P}_t(f \log f) - \mathbf{P}_t(f) \log \mathbf{P}_t(f) \leqslant \lambda t\, \mathbf{P}_t\left(\int_{\mathbb{R}^d} \min\left(\frac{(\mathrm{D}_y f)^2}{f}\;,\; \mathrm{D}_y f\, \mathrm{D}_y \log f \right) \nu(dy) \right).$$

Notice that $\mathbf{\Gamma} f = \frac{\lambda}{2} \int_{\mathbb{R}^d} (\mathrm{D}_y f)^2\, \nu(dy)$.

References

[AL00] C. Ané and M. Ledoux, *On logarithmic Sobolev inequalities for continous time random walks on graphs.*, Probab. Theor. Relat. Fields **116** (2000), no. 4, 573–602.

[Bak97] D. Bakry, *On Sobolev and logarithmic Sobolev inequalities for Markov semigroups*, New trends in stochastic analysis (Charingworth, 1994) (River Edge, NJ), Taniguchi symposium, World Sci. Publishing, 1997, pp. 43–75.

[BL98] S.G. Bobkov and M. Ledoux, *On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures.*, J. Funct. Anal. **156** (1998), no. 2, 347–365.

[Gro75] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083.

[Led99] M. Ledoux, *Concentration of measure and logarithmic Sobolev inequalities*, Séminaire de Probabilités, XXXIII, Lecture Notes in Math., Springer, Berlin, 1999, pp. 120–216.

[Nua95] David Nualart, *The Malliavin calculus and related topics*, Springer-Verlag, New York, 1995.

[Pri00] N. Privault, *On logarithmic Sobolev inequalities for normal martingales*, Ann. Fac. Sci. Toulouse Math. (6) **9** (2000), no. 3, 509–518.

[Wu00] Liming Wu, *A new modified logarithmic Sobolev inequality for Poisson point processes and several applications*, Probab. Theor. Relat. Fields **118** (2000), no. 3, 427–438.

E-mail: `chafai@math.ups-tlse.fr`
URL : `http://www.lsp.ups-tlse.fr/Chafai/`
E-mail: `malrieu@math.ups-tlse.fr`
URL : `http://www.lsp.ups-tlse.fr/Fp/Malrieu/`
Postal Adress: Laboratoire de Statistique et Probabilités, U.M.R. C.N.R.S. C5583, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 4, France.
Keywords: Log-Sobolev inequality, Concentration of measure, Lévy Processes.
Subj. Class. MSC-2000 : 46-99 Functional analysis, 26D10 Inequalities involving derivatives, differential and integral operators.