A PATHWISE APPROACH OF SOME CLASSICAL INEQUALITIES.

PATRICK CATTIAUX

ECOLE POLYTECHNIQUE AND UNIVERSITÉ PARIS X

ABSTRACT. The aim of this pedagogical paper is to show how some renowned inequalities may be obtained via a simple argument : entropy projection from the path space onto finite dimensional coordinates spaces. Some applications are given : ergodic behaviour, perturbation.

1. Introduction and Framework

In recent years, the study of the ergodic behaviour of symmetric (and sometimes non symmetric) semi-groups deserved a formidable growing interest, in connection with the deeper and deeper study of functional inequalities, like Poincaré, Log-Sobolev, Sobolev, Liggett or Nash inequalities in various forms. The picture is now rather complete, though many problems are still unsolved. We refer to the survey reviews by Bakry ([3]), Ledoux ([14]), Gross ([11]) and Guionnet-Zegarlinski ([12]) or Royer ([20]) for the theory and its application to spins systems.

The common feature of most of the papers on the topic is that they involve very clever arguments in semi-group theory and functional analysis, while the underlying stochastic process is absent. The question we shall ask is thus very natural: does a direct study of the stochastic process furnish interesting (or new) indications ? Here we identify the process with its law on the path space. Roughly speaking, semi-groups are mostly connected to time marginal laws of the process, hence the full law contains more information. It is thus natural to expect that the answer to our question is positive.

In this paper we shall focus on one possible use of the process, namely relative entropy on the path space, and we shall only consider the diffusion case (i.e. continuous processes with a "carré du champ".) Extensions to jump processes are possible. Precisely, we shall interpret the Dirichlet form associated to the semi-group as some relative entropy on the path space. This interpretation immediately yields basic inequalities that are very close to the above mentioned ones. In order to describe the contents of the paper we have first to describe the framework.

Framework.

For a probability measure μ on some measurable space E, let us first consider a μ stationary diffusion process $(\mathbb{P}_x)_{x\in E}$ and its associated semi-group $(P_t)_{t\geq 0}$ with generator A. Here by a

Date: April 28, 2002.

diffusion process we mean a strong Markov family of probability measures $(\mathbb{P}_x)_{x\in E}$ defined on the space of continuous paths $\mathcal{C}^0(\mathbb{R}^+, E)$ for some, say Polish, state space E, such that there exists some algebra $\mathbb{D}$ of uniformly continuous and bounded functions (containing constant functions) which is a core for the extended domain $D_e(A)$ of the generator (see [6]).
One can then show that there exists a countable family (C^n) of local martingales and a countable family (∇^n) of operators s.t. for all $f \in D_e(A)$

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Af(X_s)\, ds = \sum_n \int_0^t \nabla^n f(X_s)\, dC_s^n\,, \tag{1.1}$$

in $\mathbb{M}^2_{loc}(\mathbb{P}_\eta)$ (local martingales) for all probability measure η on E.
One can thus define the "carré du champ" Γ by

$$\Gamma(f,g) = \sum_n \nabla^n f\, \nabla^n g \stackrel{\text{def}}{=} (\nabla f)^2\,,$$

so that the martingale bracket is given by

$$< M^f >_t = \int_0^t \Gamma(f,f)(X_s)\, ds\,.$$

In terms of Dirichlet forms, all this, in the symmetric case, is roughly equivalent to the fact that the local pre-Dirichlet form

$$\mathcal{E}(f,g) = \int \Gamma(f,g)\, d\mu \quad f,g \in \mathbb{D}$$

is closable, and has a regular (or quasi-regular) closure $(\mathcal{E}, D(\mathcal{E}))$, to which the semi group P_t is associated. Notice that with our definitions, for $f \in \mathbb{D}$

$$\mathcal{E}(f,f) = \int \Gamma(f,f)\, d\mu = -2 \int f\, Af\, d\mu \; = \; -\frac{d}{dt} \parallel P_t f \parallel^2_{\mathbb{L}^2(\mu)} |_{t=0}\,. \tag{1.2}$$

It is then easy to check that

$$\Gamma(f,g) = A\,(fg) - f\, Ag - g\, Af\,,$$

and, that for f_i in $\mathbb{D}$, the following composition formula holds

$$A\Phi(f_1,\ldots,f_n) = \sum_{i=1}^n \frac{\partial \Phi}{\partial x_i}(f_1,\ldots,f_n)\, Af_i + \frac{1}{2} \sum_{i\,,j=1}^n \frac{\partial^2 \Phi}{\partial x_i\, \partial x_j}(f_1,\ldots,f_n)\, \Gamma(f_i,f_j)\,,$$

for Φ smooth enough.

Now pick some $f \in \mathbb{D}$ s.t.

$$0 < \inf_{x\in E} f(x) \leq \sup_{x\in E} f(x) < +\infty$$

and normalize it in order to have $\parallel f \parallel \stackrel{\text{def}}{=} \parallel f \parallel_{\mathbb{L}^2(\mu)} = 1$. We denote by $N^f \stackrel{\text{def}}{=} M^{\log f}$, which satisfies

$$< N^f >_t = \int_0^t \frac{\Gamma(f,f)}{f^2}(X_s)\, ds\,. \tag{1.3}$$

In many cases $\Gamma(f,f)$ is bounded for $f \in \mathbb{D}$, so that we may suppress the word "local" for the martingale N^f. In full generality (1.3) is defined up to an explosion time T_∞. Nevertheless, if we define

$$(1.4) \qquad \begin{cases} G_t^f & = exp\,(N_t^f - \frac{1}{2} < N^f >_t)\,,\, t < T_\infty \quad ; \quad 0 \text{ if } t \geq T_\infty \\ \frac{d\mathbb{Q}^f}{d\mathbb{P}_\mu}|_{\mathcal{F}_t} & = f^2(X_0)\, G_t^f\,, \end{cases}$$

it can be shown that, either if $\Gamma(f,f)$ is bounded, or if μ is symmetric and provided $\mathcal{E}(f,f) < +\infty$, $\mathbb{Q}^f$ is a probability measure (conservativeness). We shall study some properties of $\mathbb{Q}^f$, and because it is simpler, begin with the symmetric case.

Contents.

It is well known (see Remark 2.11 below) that any μ symmetric diffusion semi-group satisfies a robust version of Poincaré inequality, namely

$$\parallel g \parallel^2 \leq \; t\,\mathcal{E}(g,g) + \; \parallel P_t g \parallel^2 \; .$$

In section 2 we show, by using the relationship between the entropy on the path space and the Dirichlet form, that a similar statement holds for the Log-Sobolev inequality (Proposition 2.7). Stronger robust versions of Poincaré can then be deduced by linearization. These robust inequalities immediately indicate why contractivity properties are naturally linked with the usual Log-Sobolev or weak Poincaré inequalities. We also show that these inequalities are linked to convexity properties of the semi-group as a function of the time.

In section 3 we discuss the entropy minimization problem and show how it is linked with the superadditivity of Fisher's information. This property is the basic tool used by E.Carlen (see [4]) for proving Blachman-Stamm inequalities.

Section 4 is devoted to some calculations and results in the stationary (non reversible case). Here again the convexity properties of the semi-group (that can be obtained thanks to the spectral decomposition in the symmetric case) are obtained.

These properties are used in section 5 for studying ergodic properties. Almost all the results of this section are well known, may be not all. In particular we discuss weak Poincaré inequalities in the spirit of [19] and their links with Log-Sobolev inequality in the spirit of [18]. Section 6 contains a small remark on the use of martingales in order to study the ergodic behaviour.

The last two sections are devoted to perturbation theory. In section 7 we show that the perturbation result of Aida and Shigekawa (see [2]) can be obtained via the ideas of section 2 and how their hypotheses are linked to the integrability of the natural Girsanov density. In section 8 we study the transmission of a Log-Sobolev inequality to a perturbed semi-group in the spirit of the work by Kavian, Kerkyacharian and Roynette on ultracontractivity (see [13]). These two sections are a first application of the methodology of section 2. In order to keep the paper into a reasonable size, we do not include explicit examples. This will be done in another work. Very interesting results are already contained in [13].

Though it is one of the most interesting point we mainly will not discuss explicit expressions for the constants. The proofs can be used to get such expressions, and at our level of generality it does not seem really efficient to get such general expressions. Of course in many applications, exact bounds have to be obtained. We also did not try to optimize the results. In many cases a more accurate study allows to improve them. As we said before we intend this paper to be a pedagogical one, i.e. we have tried to understand how the process can be used to see classical or less classical results. How to go further will be the aim of future works.

Finally I would like to acknowledge F.Y. Wang who explained to me his joint paper with M.Röckner during a fabulous stay in Wuhan. I also benefited of two very nice discussions with M. Ledoux in Toulouse and P. Mathieu in Marseille. Last but not least, my friend C. Léonard wasted some time to hear about these ideas. He is gratefully acknowledged.

2. **Symmetric diffusion processes and relative entropy : first functional inequalities.**

We assume in this section that μ is symmetric for the process.
In this case, one easily checks that, provided $f \in D(\mathcal{E})$ is bounded (not to introduce problems with domains)

$$\int (Ah + \frac{\Gamma(h,f)}{f^2})\, f^2\, d\mu = 0\,,$$

for all $h \in \mathbb{D}$, i.e. $f^2\,\mu$ is an invariant measure for the perturbed generator. One can then show that,

$$\text{if } \mu \text{ is symmetric, } \mathbb{Q}^f \text{ is Markov and } f^2\mu \text{ reversible (hence stationary).} \tag{2.1}$$

For all this we refer to [10] section 6.3 or [6] in the Feller case. Of course when $\Gamma(f,f)$ is bounded, these results are immediate consequences of Girsanov theory of drift transformation. For (2.1) to hold, symmetry of μ is required.

The key observation is that the Dirichlet form is intimately connected with the relative entropy $H(\mathbb{Q}^f, \mathbb{P}_{f^2\,\mu})$. Indeed recall that

$$H_t(\mathbb{Q}^f, \mathbb{P}_{f^2\,\mu}) = \sup_{h \in \mathbf{B}_b(\mathcal{C}^0([0,t],E))} \left(\int h\, d\mathbb{Q}^f - log \int e^h\, d\mathbb{P}_{f^2\,\mu} \right), \tag{2.2}$$

so that as it is well known

$$H_t(\mathbb{Q}^f, \mathbb{P}_{f^2\,\mu}) = \int log(G_t^f)\, d\mathbb{Q}^f\,,$$

and using Girsanov theory

$$H_t(\mathbb{Q}^f, \mathbb{P}_{f^2\,\mu}) = \frac{1}{2}\, \mathbb{E}^{\mathbb{Q}^f}\, [\int_0^t \frac{\Gamma(f,f)}{f^2}(X_s)\, ds]\,.$$

In particular we get thanks to stationarity

$$H_t(\mathbb{Q}^f, \mathbb{P}_{f^2\,\mu}) = \frac{1}{2}\, t\, \mathcal{E}(f,f). \tag{2.3}$$

Looking at (2.2) and (2.3), we see that we can play the game of choosing various h in order to get a family of inequalities. In order to be able to calculate $\int e^h \, d\mathbb{P}_{f^2\mu}$ in terms of the semi group, it is natural to use functions h of the form

$$h(X) = \sum_j log\, h_j(X_{t_j}) \, ,$$

for some sequence $0 \leq t_1 \leq \cdots \leq t_k \leq t$. Here as in the previous section we assume that f is nonnegative, bounded above and bounded from below by a positive constant. Using the Markov property and stationarity again, this yields

$$\int (\sum_j log\, h_j)\, f^2\, d\mu \leq \frac{1}{2}\, t\, \mathcal{E}(f,f) + log \int f^2\, P_{t_1}(h_1 P_{t_2-t_1}(h_2 P_{t_3-t_2} \ldots))\, d\mu \, . \tag{2.4}$$

For (2.4) to turn a functional inequality, we are led to choose $h_j = f^{\alpha_j}$ which furnishes

$$\frac{1}{2}\, (\sum_j \alpha_j) \int f^2\, log f^2\, d\mu \leq \frac{1}{2}\, t\, \mathcal{E}(f,f) + log \int f^2 P_{t_1}(f^{\alpha_1} P_{t_2-t_1}(f^{\alpha_2} P_{t_3-t_2} \ldots))\, d\mu \, . \tag{2.5}$$

The simplest choice is to consider only two marginals, i.e. $k = 2$, $t_1 = 0$, $t_2 = t$. Actually, for (2.5) to be useful, we can take three marginals, but this does not furnish better results. We also choose $\alpha_2 = 1$ and $\alpha_1 = \alpha - 1$. This yields

$$\frac{\alpha}{2} \int f^2\, log f^2\, d\mu \leq \frac{1}{2}\, t\, \mathcal{E}(f,f) + log \int f^{1+\alpha}\, P_t f\, d\mu \, . \tag{2.6}$$

The first interesting situation is obtained for α between 0 and 1. Indeed, applying Hölder inequality and the normalization of f we get

$$\int f^2\, log f^2\, d\mu \leq \frac{t}{\alpha}\, \mathcal{E}(f,f) + \frac{1-\alpha}{\alpha}\, log \int (P_t f)^{\frac{2}{1-\alpha}}\, d\mu \, .$$

Now using standard cut-off, and the Markov property of the form, we have obtained

Proposition 2.7. *If μ is a symmetric probability measure for the diffusion semi-group P_t, then for all t, all $\alpha \in]0,1[$ and all non negative $f \in D(\mathcal{E})$,*

$$\int f^2\, log f^2\, d\mu \leq \frac{t}{\alpha}\, \mathcal{E}(f,f) + \frac{1-\alpha}{\alpha}\, \parallel f \parallel^2\, log \frac{\int (P_t f)^{\frac{2}{1-\alpha}}\, d\mu}{\parallel f \parallel^2} \, .$$

The case $\alpha = 1$ can be obtained similarly, i.e.

$$\int f^2\, log f^2\, d\mu \leq t\, \mathcal{E}(f,f) + 2\, \parallel f \parallel^2\, log\, \parallel P_t f \parallel_\infty \, .$$

Proposition 2.7 immediately shows why a Log-Sobolev inequality has something to do with hypercontractivity. Indeed as an immediate consequence, we get the easy half of Gross theorem

Corollary 2.8. *If in addition P_t is continuous from $\mathbb{L}^2$ in $\mathbb{L}^p$ for some $p > 2$, with norm c, then*

$$\int f^2\, log \frac{f^2}{\parallel f \parallel^2}\, d\mu \leq a\, \mathcal{E}(f,f) + b\, \parallel f \parallel^2 ,$$

where $c = exp\left(b\,(\frac{1}{2} - \frac{1}{p})\right)$ and $\frac{t}{a} = \frac{p-2}{p}$. In particular $b = 0$ if $c = 1$ i.e. if the semi-group is hypercontractive.

It suffices to apply 2.7 to f^+ and f^-. Note that, unfortunately, the constants a and b are not the best ones (see [8] Theorem 6.1.14). One could also try to optimize by making α depend on t, as in the usual proof of Gross theorem.

As for Log-Sobolev inequalities, 2.7 will furnish some kind of Poincaré inequality. As usual if g is a bounded function in $\mathbb{D}$ satisfying $\int g\, d\mu = 0$, we will apply 2.7, or better (2.6), with $f = 1 + \varepsilon\, g$ and ε small enough. Thanks to stationarity, the leading term when ε goes to 0 is of order ε^2, and it furnishes after some simple manipulations

Proposition 2.9. *If μ is a symmetric probability measure for the diffusion semi-group P_t, then for all t, all $\alpha \in]0,1[$ and all $g \in D(\mathcal{E})$ such that $\int g\, d\mu = 0$,*

$$\| g \|^2 \leq \frac{t}{1+2\alpha}\, \mathcal{E}(g,g) + \frac{1+\alpha}{(1-\alpha)(1+2\alpha)} \ \| P_t g \|^2 \ .$$

This inequality extends to $\alpha = 0$ by continuity.
For all α, the following also holds

$$(\alpha + 1 - \frac{\alpha^2}{2}) \ \| g \|^2 \leq t\, \mathcal{E}(g,g) + (1+\alpha) \ \| P_t g \|^2 \ .$$

Notice that both inequalities agree for $\alpha = 0$.

One can use 2.9 in the spirit of the general weak Poincaré inequalities in [19]. Indeed an immediate consequence of 2.9 with $\alpha = 0$ is the following

Corollary 2.10. *If in addition $\int (P_t g)^2\, d\mu \leq \xi(t)\, C(g)$ for some functional C defined on a dense subset D_C of $\mathbb{L}^2(\mu)$ and some continuous ξ going to 0 at $+\infty$, then the following weak Poincaré inequality holds for all $r > 0$,*

$$\| g \|^2 \leq \beta(r)\, \mathcal{E}(g,g) + r\, C(g) \,,$$

for all $g \in D_C$ such that $\int g\, d\mu = 0$, where $\beta(r) = \inf\{t \geq 0, \ \text{such that } \xi(t) \leq r\}$.

Note that β is non increasing and goes to 0 when r goes to $+\infty$. It is reasonable to assume that $C(\lambda\, g) = \lambda^2 C(g)$ for the inequalities to be homogeneous. 2.10 is similar to Theorem 2.3 in [19] with a completely different coefficient in front of $\mathcal{E}(g,g)$. Notice that some converse of 2.10 is immediate as shown in [19] Theorem 2.1, provided $C(P_t g) \leq C(g)$. We shall come back to this point in a future section, and also improve the result.

Remark 2.11. For $\alpha = 0$, inequality 2.9 also is a simple consequence of the following two basic facts, for g in $\mathbb{D}$

$$\| g \|^2 - \| P_t g \|^2 = \int_0^t \mathcal{E}(P_s g, P_s g)\, ds \,, \tag{2.12}$$

and thanks to symmetry

$$\frac{d}{dt}\, \mathcal{E}(P_t g, P_t g) = -4 \int (A P_t(g))^2\, d\mu \,,$$

so that $t \mapsto \mathcal{E}(P_t g, P_t g)$ is non increasing (one can also use spectral decomposition to show this last property). This result already appeared in the proof of Liggett's Theorem 2.2 in [15], which deals with a weak form of Poincaré inequality. The constants α in 2.9 will only modify the constants in Liggett's result.

Here is another proof using martingales. Using Ito's formula in both time directions, for $g \in D(\mathcal{E})$, one has the sometimes called Lyons-Zheng decomposition formula

$$2(g(X_t) - g(X_0)) = M_t^g - (M_t^g) \circ R_t \,,$$

where R_t denotes the time reversal at time t. It follows that

$$\mathbb{E}^{\mathbb{P}_\mu}[|g(X_t) - g(X_0)|^2] = 2 \int g\,(g - P_t g)\, d\mu \leq \frac{t}{2}\, \mathcal{E}(g, g)\,.$$

The result follows (for $\frac{t}{2}$) just using the symmetry.

The previous (2.12) can be formulated in terms of convexity of $t \mapsto \parallel P_t g \parallel^2$. As remarked in [19] lemma 2.2,

$$t \to \, log \, \parallel P_t g \parallel \; \text{ is convex.}$$

This is an immediate consequence of the spectral decomposition. It is interesting to see that it also follows from the general inequalities we have obtained.
Indeed, go back to (2.6) and choose $\alpha = 0$. One thus has (recall that $\parallel f \parallel = 1$)

$$0 \leq \frac{1}{2}\, t\, \mathcal{E}(f, f) + \, log \int f\, P_t f\, d\mu \,,$$

and thanks to symmetry

$$0 \leq t\, \mathcal{E}(f, f) + \, log \, \parallel P_t f \parallel^2 \;. \tag{2.13}$$

But, (2.13) is an equality at time $t = 0$. Hence the time derivative of the right hand side is non negative at time $t = 0$. Actually it is equal to 0, hence the second time derivative is non negative, i.e. we get convexity at the origin.
Notice that this second derivative is given by

$$4 \int (Af)^2\, d\mu \, - \, \Big(\mathcal{E}(f, f)\Big)^2$$

which is unchanged when adding constants. So the result, which was true for nonnegative f, bounded away from 0, extends to all bounded $f \in D(A)$. But replacing f by $P_s f$, we get the desired convexity result by using the semi group property and density. Of course, we get convexity of $t \mapsto \parallel P_t g \parallel^\gamma$ for all nonnegative γ as a byproduct.

Remark 2.14. It is easy to see that (for a fixed t) the best possible α in 2.9 is

$$\alpha = 1 - \frac{\parallel P_t g \parallel}{\parallel g \parallel}\,,$$

so that we obtain

$$3 \parallel g \parallel^2 - 4 \parallel P_t g \parallel \parallel g \parallel + \parallel P_t g \parallel^2 \leq t\, \mathcal{E}(g, g)\,.$$

Remark that

$$\parallel g \parallel^2 - \parallel P_t g \parallel^2 \leq 3 \parallel g \parallel^2 - 4 \parallel P_t g \parallel \parallel g \parallel + \parallel P_t g \parallel^2\,,$$

so that,

$$\int_0^t \left(\mathcal{E}(P_s g, P_s g) - \mathcal{E}(g, g)\right) ds \; \leq \; 0\,.$$

Since the first derivative at time $t=0$ is equal to 0, we get that the second derivative is less or equal to 0, hence

$$\frac{d}{dt}\,\mathcal{E}(P_t g, P_t g)\,|_{t=0} \leq\, 0\,.$$

Since this inequality holds for all g, it holds for g replaced by $P_s g$, and thus, applying the semi group property, we recover the fact that $t \to \mathcal{E}(P_t g, P_t g)$ is non increasing.

One can also use the other inequality in Proposition 2.9. Again for a fixed t one can show that the best possible α is

$$\alpha = 1 - \frac{\parallel P_t g \parallel^2}{\parallel g \parallel^2}$$

which yields

$$\frac{3}{2}\, \parallel g \parallel^2 + \frac{1}{2}\,\frac{\parallel P_t g \parallel^4}{\parallel g \parallel^2} - 2 \parallel P_t g \parallel^2 \leq\, t\,\mathcal{E}(g,g)\,.$$

Differentiating with respect to t, we thus get

$$\int_0^t \left((2 - \frac{\parallel P_s g \parallel^2}{\parallel g \parallel^2})\mathcal{E}(P_s g, P_s g) \,-\, \mathcal{E}(g,g)\right) ds \,\leq\, 0\,.$$

Again, the second derivative at time $t=0$ has to be less than 0. This yields for $\parallel g \parallel = 1$,

$$\frac{d}{dt}\,\mathcal{E}(P_t g, P_t g)|_{t=0} + (\mathcal{E}(g,g))^2 \,\leq\, 0\,.$$

But this last quantity is equal to $-\,\frac{d^2}{dt^2}\,log \parallel P_t g \parallel^2 |_{t=0}$. So we again get the convexity result for the log.

Remark 2.15. Some of these results extend to the case of an unbounded nonnegative symmetric measure μ. Indeed the properties of $\mathbb{Q}^f$ do not require that f is bounded below, nor strictly positive (see [10] Theorem 6.3.3). So we may choose

$$h_1 = \left((f \vee \frac{1}{K}) \wedge L\right)^{\alpha-1} \;\text{ and }\; h_2 = f \wedge M$$

and take limits successively in L, K and M (the first one requires bounded convergence theorem in the right hand side, then use monotone convergence or inequalities). Thus all results up to 2.8 are still true without any change. Of course since 1 is not integrable, we cannot deduce the Poincaré like inequality, unless we have some additional properties.

3. Symmetric diffusion processes and relative entropy : others functional inequalities.

We still assume that μ is symmetric, but non necessarily bounded. In the beginning of this section, we furthermore assume that the family $\{\mathbb{P}_x\}$ is Feller, more precisely that the semi group maps continuous bounded functions into continuous bounded functions. The reason is we want to use some large deviations results.

In the preceding section, we built $\mathbb{Q}^f$ which is $f^2\,\mu$ symmetric (actually we only used stationarity) and has finite relative entropy with respect to $\mathbb{P}_{f^2\,\mu}$. But $\mathbb{Q}^f$ owes another interesting property : among all $f^2\,\mu$ stationary $\mathbb{Q}$, $\mathbb{Q}^f$ is the one that minimizes $H_t(\mathbb{Q}, \mathbb{P}_{f^2\,\mu})$. This can be interpreted in terms of mass transportation : among all possible measures on the paths space transporting $f^2\,\mu$ onto $f^2\,\mu$ at each time, the optimal one for relative entropy is $\mathbb{Q}^f$.

Remark 3.1. We only used time marginal laws at time 0 and time t. So, one can formulate the same problem of minimizing relative entropy for $\mathbb{Q}$ transporting $f^2\,\mu$ onto $f^2\,\mu$ at time t. This problem appears in the literature as the problem of construction of Schrödinger bridges (see e.g. [9] p.161-164 for the Brownian case, and [5] section 6). Of course $H_t(\mathbb{Q}^f, \mathbb{P}_{f^2\,\mu})$ will be greater than the minimal transportation cost for such a bridge.
The minimal transportation cost is known since $\mathbb{P}_x$ is Feller. When μ is a probability measure, it is given by

$$J(f) = \sup_{h \in \mathbf{B}_b^+(E)} \int log\,(\frac{h}{P_t h})\, f^2\, d\mu\,,$$

where B_b^+ denotes the set of nonnegative bounded h so that $1/h$ is also bounded. Hence we obtain

$$\int log(h_1\, h_2)\, f^2\, d\mu \leq \frac{t}{2}\mathcal{E}(f,f) + \int f^2\, log(h_1\, P_t h_2)\, d\mu\,, \tag{3.2}$$

for all positive and bounded h_1 and h_2.
If we take $h_1 = f^{\alpha-1}$ and $h_2 = f$, we get a better inequality than (2.6) (just use the concavity of the logarithm in the last integral). However, we do not see how to really use this improvement. Just notice that it yields another inequality

$$\frac{3}{2} \parallel g \parallel^2 \leq t\,\mathcal{E}(g,g) + 2\int (P_t g)^2\, d\mu - \frac{1}{2}\int (P_{2t} g)^2\, d\mu\,.$$

This minimality has a nice counterpart. Indeed, as shown in [6] Corollary 4.3, and using stationarity

$$H_t(\mathbb{Q}^f, \mathbb{P}_{f^2\,\mu}) = \tag{3.3}$$

$$= \sup_{h \in \mathcal{C}_b([0,t]\times E))} \left(\int \int_0^t h(s,x)ds\, f^2\, d\mu\; - \int log\, \mathbb{E}^{\mathbb{P}_x}[exp \int_0^t h(s,X_s)\, ds]\, f^2(x)\, d\mu \right).$$

As a first consequence, we see that we do not loose any information by taking h as we did in (2.4). So, the inequalities we obtained in the previous section, though not optimal, are not far to be. The second consequence is concerned with superadditivity of Fisher's information.

Definition 3.4. *For nonnegative ρ such that $\int \rho\, d\mu = 1$, the Fisher's information $I(\rho)$ is defined as*

$$I(\rho) = \int \frac{\Gamma(\rho,\rho)}{\rho}\, d\mu\; = 4\,\mathcal{E}(\sqrt{\rho}, \sqrt{\rho})\,,$$

when this quantity is well defined and finite, $+\infty$ otherwise.

Note that if $f = \sqrt{\rho}$, then $I(\rho) = 8\, H_1(\mathbb{Q}^f, \mathbb{P}_{f^2\,\mu})$, when the first quantity is finite. According to (3.3), we thus get a variational formulation for Fisher's information. Furthermore, as shown in [6] Theorem 4.4, the finiteness of the supremum in (3.3) implies that $I(\rho)$ is also finite. So we get a complete characterization.

We shall now consider two processes $\{\mathbb{P}^i_{x_i}\}_{x_i \in E_i}$ as in the previous section, and the corresponding $\mathbb{D}_i$, μ_i. The process $X = (X^1, X^2)$ associated to $\mathbb{P}^1_{x_1} \otimes \mathbb{P}^2_{x_2}$ share the same properties. The corresponding algebra is $\mathbb{D} =$ span $\mathbb{D}_1 \otimes \mathbb{D}_2$, so that it is easily seen that the corresponding $\mathcal{E}$ is given by

$$\mathcal{E}(f,f) = \int ((\nabla_1 f)^2 + (\nabla_2 f)^2)(x_1,x_2)\,\mu_1(dx_1)\,\mu_2(dx_2)\,.$$

It is also easy to see that for $f \in D(\mathcal{E})$,

$$\nabla_1 \int f(x_1,x_2)\,\mu_2(dx_2) = \int \nabla_1 f(x_1,x_2)\,\mu(dx_2) \;\; \mu_1 \text{ a.s.}$$

For a given density of probability ρ, we introduce the marginal densities ρ_i, and the conditional densities $\rho_{|i}(x_i,.)$ i.e.

$$\rho_1(x_1) = \int \rho(x_1,x_2)\,\mu(dx_2)\,, \quad \text{and} \quad \rho_{|1}(x_1,y) = \frac{\rho(x_1,y)}{\rho_1(x_1)}\,.$$

Superadditivity of Fisher's information is a statement like

$$I(\rho) \ge I_1(\rho_1) + I_2(\rho_2)$$

where I, I_i are Fisher's information respectively on $E_1 \times E_2$ and each E_i. For the usual gradient and Lebesgue measure,such a result was proved by Carlen ([4] Theorem 3). His proof lies on the stronger result (see Theorem 2 in [4] in the case $p = 2$)

$$I_1(\rho_1) \le \int I_1(\rho_{|2}(.,x_2))\,\rho_2(x_2)\,d\mu_2$$

we have rewritten in an appropriate form. This result is stronger since the sum of the right hand sides (for $i = 1, 2$) is equal to $I(\rho)$.

The proof of superadditivity is an easy consequence of (3.3). Indeed, if $f = \sqrt{\rho}$, we may define $\mathbb{Q}^f$ on the product space (recall that we only need the finiteness of the Dirichlet form for defining $\mathbb{Q}^f$) , so that

$$\frac{1}{8}\,I(\rho) = H_1(\mathbb{Q}^f, \mathbb{P}_{\rho\,\mu_1\otimes\mu_2}) =$$

$$= \sup_{h\in\mathcal{C}_b([0,1]\times(E_1\times E_2))} \Bigg(\int\int_0^1 h(s,x_1,x_2)ds\,\rho(x_1,x_2)\,\mu_1(dx_1)\mu_2(dx_2) -$$

$$- \int log\,\mathbb{E}^{\mathbb{P}^1_{x_1}\otimes\mathbb{P}^2_{x_2}}[exp\int_0^1 h(s,X^1_s,X^2_s)\,ds]\,\rho(x_1,x_2)\,\mu_1(dx_1)\mu_2(dx_2)\Bigg)\,.$$

Of course this supremum is greater than the one obtained on functions

$$h(s,x_1,x_2) = h_1(s,x_1) + h_2(s,x_2)$$

which is easily seen to be equal to $I_1(\rho_1) + I_2(\rho_2)$, by using a similar argument. Note that finiteness of $I_i(\rho_i)$ is a byproduct of the proof. We have thus shown

Proposition 3.5. *For Feller processes, if ρ is a density of probability on the product space with marginal densities ρ_1 and ρ_2,*

$$I(\rho) \ge I_1(\rho_1) + I_2(\rho_2)\,.$$

Furthermore, when equality holds, it is easy to see that

$$\frac{\nabla\rho}{\rho} = \frac{\nabla_1\rho_1}{\rho_1} \oplus \frac{\nabla_2\rho_2}{\rho_2}$$

$\rho\mu$ a.s., where by convention a ratio is 0 whenever the denominator vanishes. It follows that

$$\rho = (\rho_1 \otimes \rho_2)\, e^h$$

where h is an invariant function, and finally $\rho = \rho_1 \otimes \rho_2$ if the μ_i are ergodic.

If we add some conditions, we can get the stronger statement of [4].

Assume from now on that $\int \rho\, log\rho\, d\mu < +\infty$. This is automatically satisfied if $I(\rho)$ is finite when a Log-Sobolev inequality holds (as for the euclidean case). It can also be assumed when the μ_i are probability measures. Indeed in this case, we can first assume that ρ is bounded below and above and then take limits.
Then for μ_2 almost all x_2

$$H(\rho_{|2}(.,x_2)\,\mu_1,\mu_1) = \int \rho_{|2}(x_1,x_2)\, log\rho_{|2}(x_1,x_2)\,\mu_1(dx_1) < +\infty\,.$$

Denote by $f_{2,1}$ the square root of this conditional density. Then we have the entropy decomposition

$$\frac{t}{8}\, I_1(\rho_{|2}(.,x_2)) \;=\; H_t(\mathbb{Q}^{f_{2,1}},\mathbb{P}^1_{\rho_{|2}(.,x_2)\,\mu_1}) = H_t(\mathbb{Q}^{f_{2,1}},\mathbb{P}^1_{\mu_1}) - H(\rho_{|2}(.,x_2)\,\mu_1,\mu_1)\,. \tag{3.6}$$

We want to integrate both hand sides with respect to $\rho_2\, d\mu_2$. To this end we need the integrability of the second term in the difference, thus we need

$$\int \rho_2\, log\rho_2\, d\mu_2 < +\infty\,.$$

This is automatically satisfied when the μ_i are probability measures (just use convexity) or when $I_i(\rho_i)$ is finite and a Log-Sobolev inequality holds. In particular, it will hold when $I(\rho)$ is finite when the processes are Feller and satisfy a Log-Sobolev inequality.
Using convexity of relative entropy, we get
(3.7)

$$\begin{aligned}\frac{t}{8}\int I_1(\rho_{|2}(.,x_2))\,\rho_2(x_2)\mu_2(dx_2) \;&\geq\; H_t(\mathbb{Q}^{f_1},\mathbb{P}^1_{\mu_1}) - \int H(\rho_{|2}(.,x_2)\,\mu_1,\mu_1)\rho_2(x_2)\mu_2(dx_2)\\ &\geq\; \frac{t}{8}\, I_1(\rho_1) + H(\rho_1\,\mu_1,\mu_1) - \int H(\rho_{|2}(.,x_2)\,\mu_1,\mu_1)\rho_2(x_2)\mu_2(dx_2)\,.\end{aligned}$$

Dividing by t and making t tend to infinity, we get the desired result. Remark that we do not need the Feller assumption since we did not use (3.3).
Furthermore, if equality holds, we must have

$$H(\rho_1\,\mu_1,\mu_1) - \int H(\rho_{|2}(.,x_2)\,\mu_1,\mu_1)\rho_2(x_2)\mu_2(dx_2)\,,$$

and since relative entropy is strictly convex, this implies $\rho_{|2}(.,x_2) = \rho_1$ for ρ_2 almost all x_2, i.e. $\rho = \rho_1 \otimes \rho_2$. Let us summarize the results we have obtained

Theorem 3.8. *Let μ_i be symmetric nonnegative measures for the processes $\{\mathbb{P}^i_{x_i}\}$. If one of the following conditions holds*

(1) *μ_i are probability measures,*

(2) *the processes are Feller continuous,*

then for all probability density ρ *with marginal laws* ρ_i *it holds*

$$I(\rho) \geq I_1(\rho_1) + I_2(\rho_2)$$

where I *denotes the Fisher's information.*
In addition in cases (1) and (2)+(the processes satisfy a Log-Sobolev inequality) the following stronger inequality holds

$$I_1(\rho_1) \leq \int I_1(\rho_{|2}(.,x_2))\, \rho_2(x_2)\, d\mu_2$$

where $\rho_{|i}(.,x_i)$ *denotes the conditional density knowing* x_i*.*
Finally in all cases, equality holds if and only if $\rho = \rho_1 \otimes \rho_2$*.*

Actually we proved a little bit more since Log-Sobolev is only required to ensure finiteness of some entropy. In the euclidean space in particular (where Log-Sobolev actually holds), one can use some approximations in order to get the result without using explicitly the Log-Sobolev condition. Also note that if there exist a pair (ψ_1, ψ_2) of probability densities such that $(\nabla_i log \psi_i)^2$ is $\rho\mu$ integrable, we may replace μ_i by $\psi_i \mu_i$ (which are probability measures) and ρ by $(\psi_1 \otimes \psi_2)^{-1}\rho$ in order to use the previous result with probability measures.

In the euclidean case, these inequalities are connected with Blachman-Stam inequality (see [4] Theorem 7) which furnishes another proof of Nelson's hypercontractivity theorem (to be useful one of course has to get a proof that does not use a priori Log-Sobolev). They also have some interesting applications in kinetic theory, we shall not discuss here. Thanks to Theorem 3.8, Theorem 7 in [4] immediately extends to more general Gaussian spaces.

4. **The general stationary case.**

We do no more assume that μ is symmetric. Of course the Dirichlet form only controls the symmetric part of the process, so we cannot expect to get similar results in the general stationary case. In some particular cases, one can rely the symmetric part to the whole process. For instance if A is normal (i.e. commutes with its adjoint) one can use the previous results for the symmetric process generated by $\frac{A+A^*}{2}$ and remark that

$$\| P_t g \|^2 = \| P_t^* g \|^2 = \| (P_{\frac{t}{2}} P_{\frac{t}{2}}^*) g \|^2$$

so that the results of the previous subsection starting from 2.9 are still true.

One can nevertheless ask about what happens if we play a similar game in the general stationary case assuming that $\Gamma(f,f)$ is bounded for $f \in \mathbb{D}$. First we assume that μ is a probability measure.

In this case (2.2) is still hold, so that we have an analogue of (2.4)

$$\int \sum_j log\, h_j\, f_{t_j}^2\, d\mu \leq \tag{4.1}$$

$$\leq \frac{1}{2}\, \mathbb{E}^{\mathbb{Q}^f} \left[\int_0^t \frac{\Gamma(f,f)}{f^2}(X_s)\, ds\right] + log \int f^2 P_{t_1}(h_1 P_{t_2-t_1}(h_2 P_{t_3-t_2} \ldots))\, d\mu\,,$$

where f_s^2 denotes the density of the $\mathbb{Q}^f$ law of X_s with respect to μ.

This time, we take $h_1 = f^\alpha$, $h_2 = f^\beta$, $t_1 = 0$ and $t_2 = t$. Since G_t^f is invariant when scaling f, we have, for all positive $f \in \mathbb{D}$, all α and β

$$\alpha \int f^2 log f \, d\mu + \beta \, \mathbb{E}^{\mathbb{P}_\mu} [f^2(X_0) \, log \, f(X_t) \, G_t^f] + 2 \parallel f \parallel^2 \; log \; \parallel f \parallel \leq \tag{4.2}$$

$$\leq \frac{1}{2} \, \mathbb{E}^{\mathbb{P}_\mu} [\, f^2(X_0) \, G_t^f \int_0^t \frac{\Gamma(f,f)}{f^2}(X_s) \, ds] + \parallel f \parallel^2 \; log \int f^{2+\alpha}(P_t f^\beta) \, d\mu \, .$$

Now as before, take $f = 1 + \varepsilon g$ for some bounded $g \in \mathbb{D}$ such that $\int g \, d\mu = 0$. We already saw that the leading term is in ε^2 in all terms except the second one in the left hand side, which is new (the first term in the right hand side is clearly of order ε^2).
In what follows we are using $\simeq$ to denote the coefficient of the ε^2 term, for each term in the previous inequality.

(i) $\alpha \int f^2 log f \, d\mu \simeq \frac{3}{2} \alpha \parallel g \parallel^2$,

(ii) $2 \parallel f \parallel^2 \; log \parallel f \parallel \simeq \parallel g \parallel^2$,

(iii) $\frac{1}{2} \, \mathbb{E}^{\mathbb{P}_\mu} [\, f^2(X_0) \, G_t^f \int_0^t \frac{\Gamma(f,f)}{f^2}(X_s) \, ds] \simeq \frac{t}{2} \, \mathcal{E}(g,g)$,

(iv) $\parallel f \parallel^2 log \int f^{2+\alpha}(P_t f^\beta) \, d\mu \simeq (2+\alpha)\beta \int g P_t g \, d\mu + \frac{1}{2} \Big((2+\alpha)(1+\alpha) + \beta(\beta - 1) \Big) \parallel g \parallel^2$,

and finally

$$\mathbb{E}^{\mathbb{P}_\mu} [f^2(X_0) \, log \, f(X_t) \, G_t^f] \approx \mathbb{E}^{\mathbb{P}_\mu} [(1 + 2\varepsilon \, g(X_0))(1 + \varepsilon \, M_t^g) \, (\varepsilon \, g(X_t) - \frac{\varepsilon^2}{2} \, g^2(X_t))] \, ,$$

$$\simeq -\frac{1}{2} \parallel g \parallel^2 + 2 \int g P_t g \, d\mu + \mathbb{E}^{\mathbb{P}_\mu} [g(X_t) \, M_t^g] \, .$$

The last term in the previous approximate equality can be calculated as follows (for g in the domain of A):

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}_\mu} [g(X_t) \, M_t^g] &= \mathbb{E}^{\mathbb{P}_\mu} [g(X_t) \Big(g(X_t) - g(X_0) - \int_0^t Ag(X_s) \, ds \Big)] \, , \\
&= \parallel g \parallel^2 - \int g P_t g \, d\mu - \int_0^t \int P_{t-s} g \, Ag \, d\mu \, ds \, , \\
&= \parallel g \parallel^2 - \int g P_t g \, d\mu + \int_0^t \int g \, A P_{t-s} g \, d\mu \, ds + \int_0^t \int \Gamma(g, P_{t-s} g) \, d\mu \, ds \, , \\
&= \int_0^t \int \Gamma(g, P_s g) \, d\mu \, ds \, ,
\end{aligned}$$

where we have used stationarity, the definition of Γ and of the generator A.
Putting all this together, we finally obtain for $g \in \mathbb{D}$ with mean 0 :

$$\frac{1}{2} \, (\beta^2 + \alpha^2) \parallel g \parallel^2 + \frac{t}{2} \, \mathcal{E}(g,g) + \alpha\beta \int g P_t g \, d\mu - \beta \int_0^t \int \Gamma(g, P_s g) \, d\mu \, ds \geq 0 \, ,$$

or after a simple manipulation

$$\text{(4.3)}\quad \Big(\beta - \frac{1}{2}\,(\beta^2+\alpha^2)\Big)\,\parallel g \parallel^2 \leq \frac{t}{2}\,\mathcal{E}(g,g) + \beta(\alpha+1)\int gP_tg\,d\mu + \beta\int_0^t \int Ag\,P_sg\,d\mu\,ds\,.$$

One should try to get conditions for this inequality to become more exciting. In particular when the last term is less than 0, or when one can use an integration by parts to control it. We shall come back later to some examples.

The second fact we remarked in the symmetric case is that some convexity properties are consequences of (2.6). Let us see what happens when we are using (4.2). For the computations below to be justified, we need to assume that $\mathbb{D}$ is stable, i.e. that $Af \in \mathbb{D}$ if $f \in \mathbb{D}$. This hypothesis is not really surprising in the context of such inequalities (see [3]).
Let us look at (4.2) with $\alpha = -1$ and $\beta = 1$, and for $\parallel f \parallel = 1$ ($f \in \mathbb{D}$).
At time $t = 0$ we have an equality. It is not hard to see that the first derivatives with respect to t are equal at time $t = 0$. Hence the second derivatives are still satisfy an inequality, and we have to compute them.
The second derivative of the left hand side (at $t = 0$) is given by

$$\int \bigg(f^2\,A\Big(\frac{1}{f}\,Af + \frac{1}{2}\frac{1}{f^2}\,\Gamma(f,f)\Big) + f\,\Gamma(f,\frac{1}{f}\,Af) + \frac{1}{2}\,f\,\Gamma(f,\frac{1}{f^2}\,\Gamma(f,f))\bigg)\,d\mu$$

while the one of the right hand side is given by

$$\int \bigg(\frac{1}{2}\,f^2\,A\Big(\frac{1}{f^2}\,\Gamma(f,f)\Big) + \frac{1}{2}\,f\,\Gamma(f,\frac{1}{f^2}\,\Gamma(f,f)) + f\,AAf\bigg)\,d\mu\; - \;\Big(\int fAf\,d\mu\Big)^2.$$

Using the composition formula and the fact that ∇ is a derivation, we obtain after some calculations

$$\text{(4.4)}\qquad \frac{1}{4}\,\Big(\mathcal{E}(f,f)\Big)^2 \;\leq\; \int (Af)^2\,d\mu\,.$$

But if A^* has the same properties, we may replace A by A^* and take the average of both inequalities. Up to a factor 2, this average is the second derivative at time $t = 0$ of $t \mapsto log \parallel P_tf \parallel^2$. As for the symmetric case, one can extend the result to all $f \in \mathbb{D}$ by adding constants.
It remains to extend the result to all t, by using the semi group property, provided $\mathbb{D}$ is also stable for the semi group. Actually we only need this stability for functions in $\mathbb{D}$ up to a constant, i.e with 0 mean. Another possibility is to extend (4.4) to all $f \in D(A)$ provided one can approximate Af by Af_n for some sequence $f_n \in \mathbb{D}$.
Finally using density we get,

Proposition 4.5. *Let μ be a stationary probability measure for the diffusion semi group P_t. Assume that $\mathbb{D}$ is a stable algebra for A (in particular $\Gamma(f,f)$ is bounded for $f \in \mathbb{D}$) and for A^*. Assume in addition that either $\{f \in \mathbb{D}\,,\ \int f\,d\mu = 0\}$ is stable for the semi group, or $\mathbb{D}$ is dense in $D(A)$ equipped with the norm $\parallel . \parallel + \parallel \nabla . \parallel + \parallel A. \parallel$. Then*

$$t \;\mapsto\; log \parallel P_tf \parallel$$

is convex for all $f \in \mathbb{L}^2(\mu)$.

This result is quite stable by approximation of semi-groups. We shall see some examples later.

Finally, one can also partly extend the results in the previous section. Here is the easiest result in this direction.

Proposition 4.6. *Assume that μ_i are stationary probability measures for the Feller processes $\{\mathbb{P}^i_{x_i}\}$. If ρ is a probability density, with marginal densities ρ_i, then*

$$I(\rho) \,\geq\, I_1(\rho_1) + I_2(\rho_2)\,.$$

If in addition $\mathcal{E}(h,h) = 0$ implies that h is a constant, then equality holds if and only if $\rho = \rho_1 \otimes \rho_2$.

Proof. If $\rho \in \mathbb{D}$ is bounded away from 0, one can again use the results of [6] in order to prove (recall that $f = \sqrt{\rho}$)

$$\mathbb{E}^{\mathbb{P}^1_{\mu_1} \otimes \mathbb{P}^2_{\mu_2}} \left[\, f^2(X_0)\, G^f_t \int_0^t \frac{\Gamma(f,f)}{f^2}(X_s)\, ds\right] \,\geq \tag{4.7}$$

$$\mathbb{E}^{\mathbb{P}^1_{\mu_1}} \left[\, f_1^2(X^1_0)\, G^{f_1}_t \int_0^t \frac{\Gamma(f_1,f_1)}{f_1^2}(X^1_s)\, ds\right] + \mathbb{E}^{\mathbb{P}^2_{\mu_2}} \left[\, f_2^2(X^2_0)\, G^{f_2}_t \int_0^t \frac{\Gamma(f_2,f_2)}{f_2^2}(X^2_s)\, ds\right].$$

The result follows by taking the first derivative at time $t = 0$. Next it extends to any $\rho \in D(\mathcal{E})$ bounded away from 0 by density, and finally to any $\rho \in D(\mathcal{E})$ by considering $\rho_K = (\rho \wedge K) \vee \frac{1}{K}$ and using monotone convergence.
Equality can be treated similarly. The only difference is that in the symmetric case, $\mathcal{E}(h,h) = 0$ is equivalent to h invariant. □

We shall now study some consequences of the inequalities we proved (or recalled) in terms of the ergodic behaviour of the process.

5. **Ergodic behaviour of stationary diffusion processes.**

The framework is the same as in section 1. As we saw, some interesting inequalities rely on convexity properties. Let us introduce some notations.

Notation 5.1. We shall say that the Convexity Property (CP) holds for P_t, when $t \to \parallel P_t g \parallel^2$ is convex (if one prefers if $t \to \mathcal{E}(P_t g, P_t g)$ is non increasing); that the Strong Convexity Property (SCP) holds when $t \to log \parallel P_t g \parallel$ is convex (if one prefers if $t \to \frac{\mathcal{E}(P_t g, P_t g)}{\|P_t g\|^2}$ is non increasing). As we saw in the previous sections, both properties hold when μ is symmetric, when A is normal or in some nice stationary cases. When no confusion is possible we just say that CP holds, without mentioning P_t.

We also introduce usual definitions

Definition 5.2. *We denote by $\mathcal{I}$ the σ-field generated by the invariant functions, i.e. the $f \in \mathbb{L}^2(\mu)$ such that $P_t f = f$ for all t. We shall say that μ is ergodic if $\mathcal{I}$ is reduced to the constants.*

It immediately follows from (2.12) that

(5.3) If CP holds, f is invariant if and only if $f \in D(\mathcal{E})$ and $\mathcal{E}(f,f) = 0$.

In general for any invariant function the Dirichlet form vanishes. Since P_t^* is also a contraction, it is easy to see that the invariant functions of P_t and P_t^* are the same, so that P_t and $\mathbb{E}[./\mathcal{I}]$ commute.
The renowned mean ergodic theorem tells that the Cesaro means of the $P_s f$ converge towards $\mathbb{E}[f/\mathcal{I}]$. In the symmetric case, by using the spectral decomposition and Lebesgue theorem, one easily sees that $P_t f$ converges towards $\mathbb{E}[f/\mathcal{I}]$ strongly in $\mathbb{L}^2(\mu)$. In the general stationary case we have

Proposition 5.4. *Assume that μ is stationary for the diffusion semi-group P_t. Then for all $f \in \mathbb{L}^2(\mu)$, $P_t f$ is $\mathbb{L}^2(\mu)$ weakly convergent to $\mathbb{E}^\mu[f/\mathcal{I}]$ when t goes to $+\infty$.*
If in addition CP holds for both P_t and P_t^* (in particular if μ is symmetric, or A is normal), convergence holds in the strong sense.
In particular, if μ is ergodic, $P_t f$ goes to $\int f\, d\mu$.

Proof. Let $g = f - \mathbb{E}^\mu[f/\mathcal{I}]$. Then $\int g\, d\mu = 0$ and g is orthogonal to the set of invariant functions. $t \mapsto \| P_t g \|^2$ is non increasing and bounded, hence has a limit m when t goes to $+\infty$. It follows that from any subsequence of $P_t g$, one can extract a weakly convergent subsequence $P_{t_n} g$. Let h be a weak cluster point. We are going to show that h is invariant.
Indeed since h is the weak limit of some sequence, it is the strong limit of a sequence h_n of convex combinations of the $(P_{t_k} g)_{k \geq n}$, according to Mazur's theorem. But, according to the convexity of the Dirichlet form and (2.12), for all t,

$$\| h_n \|^2 - \| P_t h_n \|^2 = \int_0^t \mathcal{E}(P_s h_n, P_s h_n)\, ds \leq \sup_{u \geq t_n} \int_u^{u+t} \mathcal{E}(P_s g, P_s g)\, ds\,,$$

which goes to 0 when n goes to $+\infty$. It follows thanks to the continuity of P_t and the strong convergence of h_n that for all t, $\| h \|^2 = \| P_t h \|^2$, hence that h is invariant.
In addition,

$$\| h \|^2 = \lim_n \int P_{t_n} g\, h\, d\mu = \lim_n \int g\, P_{t_n}^* h\, d\mu = 0$$

since h is P_t^* invariant and since g is orthogonal to invariant functions. It follows that $P_t g$ weakly converges to 0, when t goes to $+\infty$. In particular

(5.5) $$\text{for all } s,\ \lim_{t \to +\infty} \int P_s g\, P_t g\, d\mu = 0\,.$$

Hence if μ is symmetric, $m = 0$. In the normal case, apply this result with the symmetric $P_t P_t^*$.
If CP holds for both P_t and P_t^* we shall follow the same route this time with $P_t^*\, P_t g$. Indeed

$$\| P_t^*\, P_t g \| \leq \| P_t g \|$$

so that again, starting with any sequence t_n we may find a sequence h_n of convex combinations of the $(P^*_{t_k}\, P_{t_k} g)_{k\geq n}$ which is strongly convergent towards some h. But

$$\begin{aligned} \| h_n \|^2 - \| P_t h_n \|^2 &= \int_0^t \mathcal{E}(P_s h_n, P_s h_n)\, ds \\ &\leq t\, \mathcal{E}(h_n, h_n) \\ &\leq t\, \mathcal{E}(P^*_{t_n}\, P_{t_n} g, P^*_{t_n}\, P_{t_n} g) \\ &\leq t\, \mathcal{E}(P_{t_n} g, P_{t_n} g)\,, \end{aligned}$$

where we have used successively (2.12), convexity of $\mathcal{E}$, CP for P_t and for P^*_t. Taking the limit in n we obtain as before first that h is invariant and next that

$$\| h \|^2 = \lim_n \int P^*_{t_n} P_{t_n} g\, h\, d\mu = \lim_n \int P_{t_n} g\, P_{t_n} h\, d\mu = \lim_n \int P_{t_n} g\, h\, d\mu = 0\,,$$

i.e. any weak cluster point is 0. Hence, as in (5.5)

$$\lim_{t\to +\infty} \int P^*_t P_t g\, g\, d\mu = 0\,.$$

But this limit is also equal to m. The proof is completed. □

Note that the usual ergodic theorem is a consequence of the first statement in proposition 5.4 according to the Banach-Saks theorem.

We shall now use the material of the previous section in order to better understand the ergodic behaviour of P_t. The first step is the following elementary lemma which explains the lack of uniformity.

Lemma 5.6. *Assume that CP holds for the diffusion semi-group P_t. Then the following statements are equivalent:*

1. *μ is ergodic.*
2. *For any sequence $\{g_n\} \in D(\mathcal{E})$ such that $\int g_n\, d\mu = 0$, $\| g_n \| \leq 1$ and $\mathcal{E}(g_n, g_n) \to 0$ as n goes to $+\infty$, we have $g_n \to 0$ weakly in $\mathbb{L}^2(\mu)$.*
   We shall call this last property the "weak weak spectral gap property" (WWSGP).
3. *We may replace the weak convergence in (2) by the strong convergence of the Cesaro means of the g_n.*

Proof. If (2) holds, then applying it with $g_n = g$, we get ergodicity thanks to (5.3). Conversely, take g_n as in WWSGP. Again, each subsequence contains some weakly convergent subsequence and then some sequence of convex combinations h_n which is strongly convergent to some h. We have to prove that $h = 0$, or thanks to ergodicity, that h is invariant (since $\int h\, d\mu = 0$).
Using convexity, we see that h_n satisfies all the hypotheses in WWSGP. According to 2.12 and CP,

$$\| h \|^2 = \lim_n \| h_n \|^2 \leq \lim_n \left(t\mathcal{E}(h_n, h_n) + \int (P_t h_n)^2\, d\mu \right) = \int (P_t h)^2\, d\mu\,,$$

for all t. Hence h is invariant.
Finally (2) implies (3) thanks to the Banach-Saks theorem, and (3) implies (1) using the same argument as for (2) implies (1). □

If we replace $\mathbb{L}^2$ weak convergence by convergence in probability in WWSGP, we obtain the "weak spectral gap property" (WSGP) introduced by Aida and Kusuoka. Of course WSGP implies WWSGP (apply Vitali's convergence theorem). If we replace weak by strong convergence in $\mathbb{L}^2$ we get the usual spectral gap property (SGP) which is clearly equivalent to Poincaré inequality.

In the symmetric case, WSGP and SGP are known to be connected with uniform ergodic properties. Similar statements are true (and mainly known) when CP holds, namely

(5.7) If CP holds, the following three properties are equivalent
(i) SGP holds,
(ii) $\lim_{t\to+\infty} \sup_{\|f\|\leq 1} \parallel P_t f - \int f\, d\mu \parallel = 0$,
(iii) there exists $\lambda > 0$ such that for all $f \in \mathbb{L}^2(\mu)$, $\parallel P_t f - \int f\, d\mu \parallel \leq e^{-\lambda t} \parallel f \parallel$;

(5.8) If CP holds, the following four properties are equivalent (see [19])
(i) WSGP holds,
(ii) $\lim_{t\to+\infty} \sup_{\|f\|\leq 1} \parallel P_t^* f - \int f\, d\mu \parallel_{\mathbb{L}^1(\mu)} = 0$,
(iii) $\lim_{t\to+\infty} \sup_{\|f\|_\infty\leq 1} \parallel P_t f - \int f\, d\mu \parallel = 0$,
(iv) the weak Poincaré inequality

$$\text{for all } r > 0, \parallel g \parallel^2 \leq \beta(r)\, \mathcal{E}(g,g) + r \parallel g \parallel_\infty^2 ,$$

holds for all bounded $g \in D(\mathcal{E})$ such that $\int g\, d\mu = 0$.

Actually the equivalence between (i) and (iv) does not require CP and is shown in [19] Proposition 1.2, as well as the direct (iv) implies (iii) (see [19] Theorem 2.1). We already saw that the converse (iii) implies (iv) only requires CP (since Corollary 2.10 only uses $\alpha = 0$). Also see [19] theorem 2.5.

Remark 5.9. In (5.8) (iii), if we denote by

$$\xi(t) = \sup_{\|f\|_\infty\leq 1} \parallel P_t f - \int f\, d\mu \parallel^2 ,$$

it is shown in [19] that, provided SCP holds, one may choose

$$\beta(r) = 2r \inf_{s>0} \frac{1}{s}\, \xi^{-1}(s\, \exp(1 - \frac{s}{r}))$$

in (iv), which is sharper than the $2\xi^{-1}(r)$ furnished by Corollary 2.10 (the factor 2 comes from the fact that $\parallel f \parallel_\infty \leq 1$ implies $\parallel f - \int f\, d\mu \parallel_\infty \leq 2$). In particular the bound of [19], contrary to the one of 2.15, allows to show that if $\xi(t) = e^{-dt}$ for some nonnegative d, then the ordinary Poincaré inequality (or SGP) holds.

Conversely if (5.8) (iv) holds for some non increasing β,

$$\xi(t) \leq 2 \inf \{r > 0\, , \, -\beta(r)\, \log(r) \leq 2t\} .$$

Remark 5.10. As mentioned in [19], and already used in [15], the weak Poincaré inequality (5.8) (iv) is equivalent for $\beta(r) = c\, r^{1-p}$ to some Nash inequality

$$\parallel g \parallel^2 \leq c\, (\mathcal{E}(g,g))^{\frac{1}{p}}\, (\parallel g \parallel_\infty)^{\frac{1}{q}}$$

for all g such that $\int g\, d\mu = 0$ and $\frac{1}{p} + \frac{1}{q} = 1$. Some generalization of these inequalities (called generalized Poincaré inequality) has been extensively studied by Mathieu (we adopt his terminology) (see e.g [17] or [16]) who was also the first to prove the equivalence between (i) and (ii) in (5.8). Note that a generalized Poincaré inequality is stronger than WSGP.

Remark 5.11. Also notice that if we replace $\| \, . \, \|_\infty$ by $\| \, . \, \|_{\mathbb{L}^p(\mu)}$ for some $p \geq 2$, we still have (iii) implies (iv) in (5.8) thanks to Corollary 2.10. Actually, as we already said, there is an equivalence between both (thanks to [19] Theorem 2.1).
But the weak Poincaré inequality with the $\mathbb{L}^p$ norm implies the same one with the $\mathbb{L}^\infty$ norm, hence WSGP. It is then easy to see that the first half of the proof of Proposition 1.2 in [19] (i.e. WSGP implies weak Poincaré for the $\mathbb{L}^\infty$ norm), immediately extends to the $\mathbb{L}^p$ norm, provided $p > 2$ (just use Hölder inequality). Hence, WSGP is actually equivalent to the weak Poincaré inequality with any $\mathbb{L}^p$ norm. In other words, as for the relationship between Log-Sobolev and hypercontractivity, the uniformly ergodic behaviour of P_t (i.e. (5.8) (iii)) holds for all $p > 2$ as soon as it holds for one such p (but with different speeds β).

Remark 5.12. It is well known that a (tight) Log-Sobolev inequality (i.e. with $b = 0$ in 2.14) is equivalent to hypercontractivity, and implies both a Poincaré inequality and the $\mathbb{L}^2 \to \mathbb{L}^p$ continuity of P_u for all $p > 2$ and u large enough. Thanks to inequality 6.1.26 in [8], it is well known that the converse (with one p) also holds. It turns out that we may replace Poincaré (or SGP) by weak Poincaré (or WSGP) in this converse statement. The result below was first shown by Mathieu in the symmetric case ([18] with a slightly different version of WSGP denoted by P, see below; also see Proposition 7 in [17]) with a different proof.

Proposition 5.13. *Assume that CP and WSGP hold, and that for some $u > 0$ and $p > 2$, P_u is $\mathbb{L}^2 \to \mathbb{L}^p$ continuous. Then SGP holds, and accordingly the semi-group is hypercontractive.*

Proof. Let g such that $\int g \, d\mu = 0$. Recall that for all $s > 0$,

$$\| g \|^2 - \| P_s g \|^2 \leq s \, \mathcal{E}(g, g) \, .$$

According to the previous remark, WSGP implies uniform ergodicity for p, i.e.

$$\| P_s g \|^2 \leq \xi(s-u) \; \| P_u g \|_p^2 \leq \xi(s-u) \, c^2 \; \| g \|^2$$

for $s > u$ and c equal to the $\mathbb{L}^2 \to \mathbb{L}^p$ norm of P_u. Since ξ goes to 0 at ∞, we immediately get the usual Poincaré inequality for s large enough. □

Also note that we may replace $\| g \|$ by $\| g \|_\infty$ in the statement of WSGP (this is condition P in [18]) or WWSGP. It is immediate for WWSGP (truncating a function will make decay the Dirichlet form). It is a consequence of the proof of Proposition 1.2 in [19] for WSGP (with the notation therein, what is used is that $\int f_n d\mu = 0$, $\| f_n \|_\infty^2 \leq \frac{1}{r}$ and $\mathcal{E}(f_n, f_n) \to 0$ implies f_n goes to 0 in probability).

Remark 5.14. As for the usual Poincaré inequality, it is not hard to see that, if μ_1 and μ_2 both satisfy WSGP then so does $\mu_1 \otimes \mu_2$. However the corresponding β in the weak Poincaré inequality is not easy to describe.

Remark 5.15. We can also discuss non uniform results. Indeed if CP holds and $\int g \, d\mu = 0$,

$$1 \leq s \, \frac{\mathcal{E}(P_t g, P_t g)}{\| P_t g \|^2} + \frac{\| P_{t+s} g \|^2}{\| P_t g \|^2} \leq s \, \frac{\mathcal{E}(P_t g, P_t g)}{\| P_t g \|^2} + 1 \, , \tag{5.16}$$

for all s and t. So either

$$\liminf_{t\to+\infty} \frac{\mathcal{E}(P_t g, P_t g)}{\| P_t g \|^2} \geq \lambda^* > 0$$

in which case

$$\| P_t g \| \leq e^{-\lambda t} \; \| g \|$$

for any $\lambda < \lambda^*$ and t large enough, or the lim inf is equal to 0.
If the lim inf is a limit, for instance if SCP holds, and if this limit is equal to 0, then one get

$$\text{for all } s, \; \lim_{t\to+\infty} \frac{\| P_{t+s} g \|^2}{\| P_t g \|^2} = 1 \,.$$

In other words when SCP holds, either the decay is exponential or it is algebraic (for example one cannot have some decay like $e^{-\sqrt{t}}$). Examples of algebraic decay are given in [7] for the critical Ornstein-Uhlenbeck process in infinite dimension (see [7] Theorems 1.9).

6. **Ergodic behaviour through Martingales.**

Since we have definitely forgotten the underlying process in the previous section, let us see what can be said using the ideas of section 2. Here we assume that μ is a symmetric probability measure for the semi-group. This section only contains remarks.

Let $g \in D(\mathcal{E})$ be bounded. One can define

$$M^g_{\cdot} = \int_0^{\cdot} \nabla g(X_s).dC_s$$

(see (1.1)) the associated martingale as the $\mathbb{L}^2$ limit of the corresponding $M^{g_n}_{\cdot}$ for a sequence $g_n \in \mathbb{D}$ that converges to g for the Dirichlet norm. So we may define

$$Z^g_t = \exp\{M^g_t - < M^g >_t\} \,.$$

Remember that we know that

$$t \,\to\, e^{2g(X_0)} \, Z^g_t$$

is a $\mathbb{P}_\mu$ martingale thanks to [10] section 6.3 or [6] (remember that g is bounded), so that, since g is bounded, $Z^g_{\cdot}$ is a $\mathbb{P}_\mu$ martingale.

If we denote by $f^2 = \exp 2g$, remember that

$$H_t(\mathbb{Q}^f, \mathbb{P}_{f^2\mu}) = \mathbb{E}^{\mathbb{Q}^f}[log Z^g_t] = \frac{t}{2} \int \Gamma(g,g)\, e^{2g}\, d\mu \;\leq\; \frac{t}{2}\, e^{2\,\|g\|_\infty}\, \mathcal{E}(g,g) \,. \tag{6.1}$$

But relative entropy is non increasing under measurable transforms, hence thanks to symmetry

$$H(e^{2g}\mu, P_t(e^{2g})\mu) = H(\mathbb{Q}^f \circ (X_t)^{-1}, \mathbb{P}_{f^2\mu} \circ (X_t)^{-1}) \;\leq\; H_t(\mathbb{Q}^f, \mathbb{P}_{f^2\mu}) \,. \tag{6.2}$$

Finally applying Pinsker's inequality ($\| \nu - \mu \|_{TV} \leq \sqrt{2\, H(\nu,\mu)}$), we get

$$\int |P_t(e^{2g}) - e^{2g}|\, d\mu \;\leq\; \sqrt{t}\, e^{\|g\|_\infty} \sqrt{\mathcal{E}(g,g)} \,. \tag{6.3}$$

In particular, if we introduce as before

$$\xi(t) = \sup_{\|f\|_\infty \leq 1} \; \| P_t f - \int f \, d\mu \, \|^2 \,,$$

and using Cauchy-Schwarz inequality we get

$$\int |e^{2g} - \int e^{2g} d\mu|\, d\mu \leq \sqrt{t}\, e^{\|g\|_\infty} \sqrt{\mathcal{E}(g,g)} + e^{2\,\|g\|_\infty} \sqrt{\xi(t)}\,. \tag{6.4}$$

Of course what we achieved to do is just another derivation of an inequality similar to proposition 2.9, without using the preliminary log-Sobolev type inequality (2.6). These inequalities immediately show why the uniform ergodic behaviour (i.e. $\xi(t) \to 0$ when t goes to ∞) implies WSGP. Notice that for an uniformly bounded sequence g_n, convergence in probability and in $\mathbb{L}^2$ are equivalent. Also notice that one can recover lemma 5.6 by directly using (6.3) and convexity arguments.

7. **Perturbation theory.**

Since our approach in section 2 lies on some perturbation, it is natural to expect that it is well suited for perturbation theory. Let us explain what we mean. In all what follows μ is supposed to be symmetric for the semi-group.

In the framework of section 1, we shall say that μ satisfies a Log-Sobolev inequality LSI if for some universal constants a and b and all $f \in D(\mathcal{E})$,

$$\int f^2\, log\Big(\frac{f^2}{\| f \|^2_{\mathbb{L}^2(\mu)}}\Big)\, d\mu \leq a \int \Gamma(f,f)\, d\mu + b\, \| f \|^2_{\mathbb{L}^2(\mu)}\ .$$

When $b = 0$ we will say that the inequality is tight (TLSI).

What we want to study is the following problem: let μ and $\nu = e^{2F}\mu$, be two probability measures (the second one is the perturbation). What should be said on ν when μ for instance satisfies a Log-Sobolev inequality ? It is well known (see e.g. [20] proposition 3.1.18) that if F is bounded, and μ satisfies

$$\int f^2\, log\Big(\frac{f^2}{\| f \|^2_{\mathbb{L}^2(\mu)}}\Big)\, d\mu \leq c \int \Gamma(f,f)\, d\mu\,,$$

then ν satisfies

$$\int f^2\, log\Big(\frac{f^2}{\| f \|^2_{\mathbb{L}^2(\nu)}}\Big)\, d\nu \leq c\, e^{Osc(F)} \int \Gamma(f,f)\, d\nu\,,$$

where $Osc(F) = sup(F) - inf(F)$. This is due to Holley and Stroock.

If we remove the boundedness assumption on F, the situation becomes much more tricky. Such a study was first done by Aida and Shigekawa ([2]) (we will not study here the relationship with exponential integrability or concentration of measure phenomenon). We will try to understand the set of hypotheses made in [2], by using the approach of section 2.

As in the previous section we consider some $F \in D(\mathcal{E})$, such that $\nu = e^{2F}\mu$ is a probability measure. We can thus build the probability measure $\mathbb{Q}^{e^F}$ that solves the martingale problem associated with $\big(A^F = A + \Gamma(F,.), \mathbb{D}\big)$, and is reversible, provided we solve some problems related to domains. These problems are carefully studied in [2] (lemma 2.3, proposition 3.2, assumption (A.6) and so on) in the framework of Dirichlet forms.

Let us first look at what is needed for the construction.

Lemma 7.1. *Assume that*

(1) *for all* $f \in \mathbb{D}$, $\mathcal{E}_F(f,f) = \int \Gamma(f,f)\, e^{2F}\, d\mu < +\infty$,
(2) *for all* $f \in \mathbb{D}$, $Af \in \mathbb{L}^1(\nu)$,
(3) $\int \Gamma(F,F)\, e^{2F}\, d\mu < +\infty$.

Then $\mathbb{Q}^{e^F}$ *is a reversible probability measure and solves the martingale problem associated with* $(A^F, \mathbb{D})$.
Furthermore, when (1) and (3) are fulfilled, the pre-Dirichlet form $(\mathcal{E}_F, \mathbb{D})$ *is closable.*

Proof. Hypotheses (1), (2) and (3) ensure that $A^F f$ is well defined and belongs to $\mathbb{L}^1(\nu)$. Using the composition formula for e^{F_n} where F_n is a sequence of functions in $\mathbb{D}$ that goes to F for the Dirichlet norm (associated with $\mathcal{E}$) it is easy to see that $\int A^F f\, d\nu = 0$ for $f \in \mathbb{D}$. Hence we can perform the construction of $\mathbb{Q}^{e^F}$ as explained in section 2. This prove the first part.

Now consider $F_K = F \wedge K$. F is bounded from above, e^F is thus bounded, hence (2) is automatically satisfied for F_K. In addition

$$\int \Gamma(f,f)\, e^{2F_K}\, d\mu \;\leq\; \int \Gamma(f,f)\, e^{2F}\, d\mu < +\infty\,.$$

Hence any sequence $f_n \in \mathbb{D}$ which is a Cauchy sequence for $\mathcal{E}_F$, is still a Cauchy sequence for $\mathcal{E}_{F_K}$. If in addition f_n goes to 0 in $\mathbb{L}^2(\nu)$, it also goes to 0 in $\mathbb{L}^2(\nu_K)$. Denote by B the $\mathbb{L}^2$ limit of ∇f_n. This limit does not depend on K.
Now remember the Lyons-Zheng decomposition

$$2(f_n(X_t) - f_n(X_0)) = \int_0^t \nabla f_n(X_s).dC_s^K - (\int_0^t \nabla f_n(X_s).dC_s^K) \circ R_t$$

that holds for all t, $\mathbb{Q}^{e^{F_K}}$ almost surely (the meaning of the basic martingales C^K is clear). Taking limits in n we get that $\mathbb{Q}^{e^{F_K}}$ almost surely

$$\int_0^t B(X_s).dC_s^K - (\int_0^t B(X_s).dC_s^K) \circ R_t \;= 0\,.$$

Note that the first term is a true martingale (not only local). It follows, using the reversibility of $\mathbb{Q}^{e^{F_K}}$, that for all $g \in \mathbb{D}$,

$$\int g(X_t) \Big(\int_0^t B(X_s).dC_s^K \Big)\, d\mathbb{Q}^{e^{F_K}} \;= 0\,.$$

Using Ito's formula, dividing by t and letting t go to 0, we thus obtain

$$\int \nabla g\,.B\, d\nu_K = 0\,.$$

This shows that B is orthogonal to all the "gradients". But by construction B belongs to the $\mathbb{L}^2$ closure of the "gradients", hence $B = 0 \quad \nu_K$ almost surely, then ν almost surely thanks to the bounded convergence theorem. □

In many interesting cases $\Gamma(f,f)$ is bounded for $f \in \mathbb{D}$, so that (1) is automatically satisfied. We shall come back later to (1) in the general case.

Let us come back to the initial problem of perturbation. Not to introduce difficulties, we shall here first assume that F is bounded from above and then try to remove this assumption. Of course the estimates we want to obtain do not involve the $\mathbb{L}^\infty$ norm of F^+.

If F is bounded from above we may apply lemma 7.1, since (1),(2) and (3) are satisfied.
Now for any $f \in \mathbb{D}$ we can play the same game as in section 2, i.e. build a perturbed probability measure $\mathbb{U}^f$ as

$$\frac{d\mathbb{U}^f}{d\mathbb{Q}^{e^F}}|_{\mathcal{F}_t} = f^2(X_0)\,(GF)^f_t\,,$$

where (GF) is similar to G (see (1.4)) replacing the $\mathbb{P}_\mu$ local martingale M^f, by its $\mathbb{Q}^{e^F}$ analogue $(MF)^f$ obtained by replacing A by A^F.
We thus obtain the analogue of (2.4) (restricting ourselves to two times $t_1 = 0$, $t_2 = t$)

$$\int (\sum_{j=1,2} log\, h_j)\, f^2\, e^{2F} d\mu \leq \frac{1}{2}\, t\, \mathcal{E}_F(f,f) + \, log\, \mathbb{E}^{\mathbb{Q}^{e^F}}\left[f^2(X_0) h_1(X_0) h_2(X_t)\right]. \tag{7.2}$$

But $Z^F_. = G^{e^F}_.$ being a $\mathbb{P}_\mu$ martingale, it easily follows that for μ almost all $x \in E$, the $\mathbb{P}_x$ supermartingale $Z^F_.$ is a martingale. Since E is Polish, we then obtain that a ν regular desintegration $(\mathbb{Q}^{e^F}_x)_x$ of $\mathbb{Q}^{e^F}$, is given by

$$\frac{d\mathbb{Q}^{e^F}_x}{d\mathbb{P}_x}|_{\mathcal{F}_t} = Z^F_t\,.$$

Hence (7.1) can be rewritten
(7.3)

$$\int (\sum_{j=1,2} log\, h_j)\, f^2\, e^{2F} d\mu \leq \frac{1}{2}\, t\, \mathcal{E}_F(f,f) + \, log\, \int \, f^2(x) h_1(x)\, \mathbb{E}^{\mathbb{P}_x}[h_2(X_t) Z^F_t]\, e^{2F}(x)\, \mu(dx).$$

What has to be done in (7.3) is to come back to the semigroup P_t. One immediately see that it will require some integrability conditions for Z^F_t. The easiest way to proceed is to use Hölder's inequality, thus we have to get conditions for Z^F_t to belong to $\mathbb{L}^p$ spaces.
To our knowledge such conditions are not well known, except when the bracket $\int_0^t \Gamma(F,F)(X_s)\, ds$ is bounded. This is of course a too strong assumption. But here one can use the particular form of Z^F_t.
Indeed

$$(Z^F_t)^p = Z^{pF}_t \, \exp\,\Big(\frac{p\,(p-1)}{2}\, \int_0^t \, \Gamma(F,F)(X_s)\, ds\Big)\,,$$

hence using successively the convexity of the exponential map, the stationarity of $\mathbb{Q}^{e^{pF}}$ and Hölder's inequality one has

$$\mathbb{E}^{\mathbb{P}_\mu}[e^{2pF(X_0)}\, (Z^F_t)^p\,] = \mathbb{E}^{\mathbb{Q}^{e^{pF}}}[\exp\,\Big(\frac{p\,(p-1)}{2}\, \int_0^t \, \Gamma(F,F)(X_s)\, ds\Big)] \tag{7.4}$$

$$\leq \mathbb{E}^{\mathbb{Q}^{e^{pF}}}\Big[\frac{1}{t}\Big(\int_0^t \exp\big(\frac{p\,(p-1)t}{2}\,\Gamma(F,F)(X_s)\big)\,ds\Big)\Big]$$
$$\leq \int \exp\big(\frac{p\,(p-1)t}{2}\,\Gamma(F,F)\big)\,e^{2pF}\,d\mu$$
$$\leq \Big(\int e^{2rpF}\,d\mu\Big)^{\frac{1}{r}}\Big(\int \exp\big(\frac{qp\,(p-1)t}{2}\,\Gamma(F,F)\big)\,d\mu\Big)^{\frac{1}{q}}$$

where (q,r) is a pair of conjugate numbers. Of course in order (7.4) to be tractable, one sees that what has to be assumed is

(7.5.1) e^F belongs to $\mathbb{L}^{\infty\,-}(\mu) = \bigcap_{p>1}\mathbb{L}^p(\mu)$,
(7.5.2) $e^{\Gamma(F,F)}$ belongs to $\mathbb{L}^{\infty\,-}(\mu)$.

Of course (7.5 .1) seems to be meaningless here, since F is bounded from above. But as we said, we are trying to get estimates that does not depend on $\mathbb{L}^\infty$ bounds.
It turns out that these assumptions are exactly the ones in [2]. As a matter of fact these authors and Masuda ([1] Theorem 3.1) have shown that provided μ satisfies a Log-Sobolev inequality, (7.5.2) implies (7.5.1).

Assuming (7.5), we may use (7.4) to get some estimates from (7.3). Indeed, we may apply Hölder's inequality for μ almost all x, so that if u is the conjugate of p,

$$\int\big(\sum_{j=1,2} log h_j\big)\,f^2\,e^{2F}d\mu \leq \frac{1}{2}\,t\,\mathcal{E}_F(f,f)+ \tag{7.6}$$
$$+\,log\int f^2(x)h_1(x)\,(P_t(h_2^u))^{\frac{1}{u}}\,\Big(\mathbb{E}^{\mathbb{P}_x}[e^{2pF}(X_0)(Z_t^F)^p]\Big)^{\frac{1}{p}}\,\mu(dx).$$

The second term in the right hand of (7.6) is bounded by

$$\frac{1}{u}\,log\Big(\int f^{2u}h_1^uP_t(h_2^u)\,d\mu\Big)+\frac{1}{p}\,log\Big(\mathbb{E}^{\mathbb{P}_\mu}[e^{2pF(X_0)}\,(Z_t^F)^p\,]\Big)\,,$$

and finally

$$\int\big(\sum_{j=1,2} log\,h_j\big)\,f^2\,e^{2F}d\mu \leq \frac{1}{2}\,t\,\mathcal{E}_F(f,f)+\frac{1}{u}\,log\Big(\int f^{2u}h_1^uP_t(h_2^u)\,d\mu\Big)+ \tag{7.7}$$
$$+\frac{1}{rp}\,log\Big(\int e^{2rpF}\,d\mu\Big)+\frac{1}{pq}\,log\Big(\int \exp\big(\frac{qp\,(p-1)t}{2}\,\Gamma(F,F)\big)\,d\mu\Big)\,.$$

We have written all the details in order to show that one can if necessary compute all the constants in what follows, where all constants depending only on the $\mathbb{L}^m$ bounds in (7.5) will be denoted by $C(F)$.
We choose $h_2 = f^{\alpha_2}$, with $u\alpha_2 = 1-\varepsilon$, and $h_1 = f^{\alpha_1}$. Denoting by $\gamma = (2+\alpha_1)u$ we will choose $u>1$ and $-\alpha_2 < \alpha_1 < 0$ in such a way that $\gamma < 2$ (this is always possible if ε is small enough, for example if $u = 1+\varepsilon$, $\alpha_1 = -\frac{\alpha_2}{2}$ it is enough that $\varepsilon < \frac{1}{5}$). It thus holds

$$\frac{\alpha_1+\alpha_2}{2}\int f^2\,log\,f^2\,d\nu \,\leq\, \frac{t}{2}\,\mathcal{E}_F(f,f)+\frac{1}{u}\,log\Big(\int f^\gamma\,P_t(f^{1-\varepsilon})\,d\mu\Big)+C(t,F)\,.$$

But we have to face a final problem. Indeed remember that the ad-hoc normalization here is

$$\int f^2\, d\nu = 1\,.$$

Hence in order to control the logarithm in the right hand side we have to introduce $e^{-\gamma F}$ i.e. denoting by θ the conjugate of $\frac{2}{\gamma}$,

$$\int f^\gamma\, P_t(f)\, d\mu \,\leq\, \Big(\int e^{-\gamma\,\theta\, F}\,\big(P_t(f^{1-\varepsilon})\big)^\theta\, d\mu\Big)^{\frac{1}{\theta}}\,,$$

hence what we need is that e^{-F} belongs to $\mathbb{L}^{\infty\,-}(\mu)$ too. But actually, Theorem 3.1 in [1] shows that provided μ satisfies a Log-Sobolev inequality, (7.5.2) implies that $e^{|F|}$ belongs to $\mathbb{L}^{\infty\,-}(\mu)$.

Hence, using one more time Hölder's inequality, we get an inequality like

$$\frac{\alpha_1+\alpha_2}{2}\int f^2\, log\, f^2\, d\nu \,\leq\, \frac{t}{2}\,\mathcal{E}_F(f,f) + c\, log\Big(\int (P_t(f^{1-\varepsilon}))^a\, d\mu\Big) + C(t,F)\,, \tag{7.8}$$

for some universal constants c, $\alpha_1+\alpha_2>0$, and $a>2$. The only dependence in F is contained in the constant $C(t,F)$.

It remains to remove the boundedness assumption on F. If $f\in\mathbb{D}$, (7.8) immediately extends to the unbounded case, provided F satisfies (7.5). Indeed just use the cut-off F_K and limits in K. But we do not know whether (1) in lemma 7.1 i.e. $\mathcal{E}_F(f,f)<+\infty$ is satisfied for $f\in\mathbb{D}$ or not (if not the inequality is true but not interesting). Actually if (7.5) holds, the above result can be shown as in lemma 2.4 in [2] (the hypothesis $\Gamma(F,F)$ bounded therein can be improved just using (7.5) and Hölder in their proof, as the authors are saying in the proof of lemma 3.1).

We have thus recovered Lemma 3.1 in [2] i.e.

Proposition 7.9. *If μ satisfies a Log-Sobolev inequality and $F\in D(\mathcal{E})$ is such that $e^{\Gamma(F,F)}$ belongs to $\mathbb{L}^{\infty\,-}(\mu)$, then $\nu = e^{2F}\,\mu$ satisfies a Log-Sobolev inequality.*

Proof. Use the $\mathbb{L}^2(\mu)\to\mathbb{L}^a(\mu)$ continuity of P_t for t large enough, and as before the comparison between the $\mathbb{L}^2(\mu)$ norm of $f^{1-\varepsilon}$ and the $\mathbb{L}^2(\nu)$ norm of f. □

Remark. As we said from the beginning, our aim was to understand the meaning of the assumptions made in [2]. The previous scheme of proof is nor simpler, nor worse than the one used by these authors. A comparison of the constants here and there is tedious. Actually we only used (repeatedly) Hölder's inequality and one time convexity of the exponential map, instead of Young's inequality. The interest of the previous derivation is just to show that these hypotheses have a nice interpretation in terms of the Girsanov density. But at the same time we used rough arguments to get some control on this density.

Also remark that the cut-off of F introduced bounded measures (not probability measures), but this fact is irrelevant just dividing by the total mass.

The second question addressed in [2] is the spectral gap property. The authors are showing that when (7.5.2) is fulfilled, if μ satisfies a tight Log-Sobolev inequality (or equivalently if P_t is hypercontractive) then so does ν. In particular ν will then satisfy SGP (recall section

5). This part of their work is surprisingly difficult, and we shall try to simplify it, thanks to the results in section 5. First we state the result

Proposition 7.10. *If μ satisfies a tight Log-Sobolev inequality and $F \in D(\mathcal{E})$ is such that $e^{\Gamma(F,F)}$ belongs to $\mathbb{L}^{\infty -}(\mu)$, then $\nu = e^{2F}\,\mu$ satisfies a tight Log-Sobolev inequality.*

Proof. We already know that ν satisfies a Log-Sobolev inequality. Furthermore,

$$\mathcal{E}(e^F, e^F) < +\infty$$

thanks to (7.5.2). We can thus build a ν symmetric semi-group P_t^F with generator A^F thanks to lemma 7.1. According to Proposition 5.13, it is enough to show that this semi-group satisfies WSGP.
We already remarked it is enough to prove a weaker version i.e.

> For any sequence $\{g_n\} \in D(\mathcal{E}_F)$ such that $\int g_n \, d\nu = 0$, $\| g_n \|_\infty \leq 1$ and $\mathcal{E}_F(g_n, g_n) \to 0$ as n goes to $+\infty$, we have $g_n \to 0$ in ν probability.

Step 1. F is bounded from above.
Assume first that F is bounded from above, i.e. e^{2F} is bounded. For any nonnegative k, introduce a non decreasing smooth φ_k (defined on $\mathbb{R}$) such that

$$\varphi_k(x) = 0 \text{ , if } x < -k-1 \text{ , } \varphi_k(x) = 1 \text{ , if } x > -k \text{ , } |\varphi_k'(x)| \leq 2 \text{ for all } x \text{ .}$$

Then for any $f \in \mathbb{D}$, $\varphi_k(F)f \in D(\mathcal{E})$ and

$$\Gamma(\varphi_k(F)f, \varphi_k(F)f) = \big(\varphi_k(F)\big)^2 \Gamma(f,f) + 2\, f\, \varphi_k'(F)\Gamma(f,F) + (\varphi_k'(F))^2\, f^2\, \Gamma(F,F)\,.$$

Accordingly we get

$$\begin{aligned} (7.11) \qquad \mathcal{E}(\varphi_k(F)f, \varphi_k(F)f) \;\; \leq \;\; & e^{2k+2}\, \mathcal{E}_F(f,f) \\ & +4 \, \| f \|_\infty \; e^{k+1} \big(\mathcal{E}_F(f,f)\big)^{\frac{1}{2}} \big(\mathcal{E}(F,F)\big)^{\frac{1}{2}} \\ & +4 \; \| f \|_\infty^2 \; \Big(\int 1\!\!1_{-k-1<F<-k}\, \Gamma(F,F)\, d\mu \Big)\,. \end{aligned}$$

(7.11) extends to any bounded $f \in D(\mathcal{E}_F)$, in particular to the sequence g_n.
Now let pick a nonnegative ε. It holds

$$(7.12) \qquad \nu(|g_n| > \varepsilon) \leq \nu(|g_n\, \varphi_k(F)| > \varepsilon) + \nu(F < -k)\,.$$

In order to control the first term in the sum, we will use (7.11) and the spectral gap property for $\mathcal{E}$. But we have to be careful with the means. Denote by $m(k,n) = \int g_n\, \varphi_k(F)\, d\mu$. Then, since e^{2F} is bounded (say by K), one has

$$\| (g_n\, \varphi_k(F) - m(k,n)) \|_{\mathbb{L}^2(\nu)}^2 \leq K \; \| (g_n\, \varphi_k(F) - m(k,n)) \|_{\mathbb{L}^2(\mu)}^2 \leq K\, C\, \mathcal{E}(\varphi_k(F)g_n, \varphi_k(F)g_n)\,,$$

where C is the spectral gap constant for $\mathcal{E}$. Now we see what it remains to do.
Let δ be a nonnegative number, smaller than ε. First we choose k such that $\nu(F < -k) < \frac{\delta}{10}$ (it is possible since $\nu(F < -k) < e^{-2k}$) and

$$4 \Big(\int 1\!\!1_{-k-1<F<-k}\, \Gamma(F,F)\, d\mu \Big) < \frac{\delta^2}{200\, KC}\,,$$

(it is possible since we know that $\Gamma(F,F)$ is in all $\mathbb{L}^p(\mu)$). Then, k being fixed, we choose n large enough for

$$e^{2k+2}\,\mathcal{E}_F(g_n,g_n) + 4\,e^{k+1}\,\big(\mathcal{E}_F(g_n,g_n)\big)^{\frac{1}{2}}\,\big(\mathcal{E}(F,F)\big)^{\frac{1}{2}} < \frac{\delta^2}{200\,KC}\,.$$

It thus holds

$$\|\,(g_n\,\varphi_k(F) - m(k,n))\,\|_{\mathbb{L}^2(\nu)} \leq \frac{\delta}{10}\,.$$

But it thus follows that

$$-\,\frac{\delta}{10} \leq \int (g_n\,\varphi_k(F) - m(k,n))\,d\nu \leq \frac{\delta}{10}\,.$$

In addition

$$|\int g_n\,\varphi_k(F)\,d\nu| < \frac{\delta}{10}$$

since $\nu(F<-k) < \frac{\delta}{10}$, g_n is bounded by 1 and $\int g_n\,d\nu = 0$. It follows that

$$|m(k,n)| \leq \frac{\delta}{5}\,.$$

We get finally

$$\nu(|g_n\,\varphi_k(F)|>\varepsilon) \leq \nu(|g_n\,\varphi_k(F) - m(k,n)| > \varepsilon - \frac{\delta}{5}) \leq \frac{25\delta^2}{1600\,\varepsilon^2}\,,$$

thanks to Tchebytcheff inequality, since $\delta < \varepsilon$. Together with (7.12) and our choice of k, we thus have proved that g_n goes to 0 in ν probability.

Step 2. General case.

If F is no more bounded from above, consider as before $F_K = F \wedge K$. According to Step 1, WSGP holds for $\mathcal{E}_{F_K}$. But $\mathcal{E}_F(g,g) \geq \mathcal{E}_{F_K}(g,g)$ for all $g \in D(\mathcal{E}_F)$. Hence for the sequence g_n we have

$$\nu(|g_n|>\varepsilon) \leq \nu_K(|g_n|>\varepsilon) + \nu(F>K)$$

and

$$|\nu_K(g_n)| \leq \nu(F>K)\,.$$

We can thus proceed as we did at the end of the first step, choosing K and then n large enough in order to get the result. □

Looking at the previous proof we see that we never used the LSI property for $\mathcal{E}_F$. Actually we have shown the following result

Proposition 7.13. *Let $F \in D(\mathcal{E})$ such that $\nu = e^{2F}\mu$ is a probability measure. Assume that conditions (1) and (3) of lemma 7.1 are fulfilled. Assume in addition that $\Gamma(F,F)$ belongs to $\mathbb{L}^p(\mu)$ for some $p>1$. Then if μ satisfies SGP, ν satisfies WSGP.*

This statement seems to be new. However a very similar one, concerned with the transmission of WSGP, is contained in [19] (Theorem 6.1). Actually this result also gives precise estimates on the constants (while we do not), but does not seem to recover 7.13.

8. More on perturbation theory.

The hypotheses made in the previous section in order to obtain a general perturbation theory are certainly too restrictive. In particular in many interesting cases $e^{\Gamma(F,F)}$ will not be μ integrable. However, looking at what we have done, one sees that this assumption is necessary in order to get some $\mathbb{L}^p$ control on the Girsanov density (see (7.4)), that can only be obtained with ν as initial law. So one can expect that direct controls (for any initial law, in particular δ_x) should be useful. This idea already appeared in [13] that mainly deals with ultracontractivity (but not only) for finite dimensional diffusion processes. We shall here show that, provided F is a little bit more regular, one can improve the results in section 7.

Let $F \in D(A)$ be such that $\nu = e^{-2F}\mu$ is a probability measure (or is bounded). We have changed the sign in accordance with the usual notation. We assume that the hypotheses (1), (2) and (3) in Lemma 7.1 are fulfilled (for $-F$). We can thus follow the same route as we did in the previous section. The Girsanov martingale Z_t^F is then

$$
\begin{aligned}
Z_t^F &= \exp\{-\int_0^t \nabla F(X_s).dC_s - \frac{1}{2}\int_0^t \Gamma(F,F)(X_s)\,ds\} \\
&= \exp\{F(X_0) - F(X_t) + \int_0^t \big(AF(X_s) - \frac{1}{2}\Gamma(F,F)(X_s)\big)ds\}\,.
\end{aligned} \tag{8.1}
$$

We know that (thanks to our assumptions) $Z_.^F$ (given by the first expression) is a $\mathbb{P}_x$ martingale for ν, hence μ almost all x. To get the second form it is enough to apply Ito's formula. If P_t^F denotes the associated (ν symmetric) semi-group, it holds ν a.s.

$$
(P_t^F h)(x) = e^F(x)\,\mathbb{E}^{\mathbb{P}_x}\left[h(X_t)\,e^{-F(X_t)}\,M_t\right], \tag{8.2}
$$

with

$$
M_t = \exp\Big(\int_0^t \big(AF(X_s) - \frac{1}{2}\Gamma(F,F)(X_s)\big)ds\Big)\,.
$$

In Lemme 2.1 of [13], the authors give a fairly general condition for the perturbed heat semi group on $\mathbb{R}^n$ to be ultracontractive. Their proof is based on an ingenious use of the Markov property which immediately extends to the present general framework. We will here be concerned with LSI (or TLSI).

Not to introduce intricacies, we shall here assume that μ is a probability measure. Hence $e^F \in \mathbb{L}^2(\nu)$, and a necessary condition for ν to satisfy LSI is thus

$$
P_t^F(e^F) = e^F\,\mathbb{E}^{\mathbb{P}_x}[M_t] \in \mathbb{L}^p(\nu) \tag{8.3}
$$

for all (some) $p > 2$ and t large enough. This condition is not always tractable so that we introduce the stronger

$$
\int e^{pF}\,\mathbb{E}^{\mathbb{P}_x}[M_t^p]\,d\nu = i(p,t) \tag{8.4}
$$

and will sometimes assume that $i(p,t) < +\infty$ for some pair (p,t) such that $p > 2$ and $t > 0$. There is a big difference between both assumptions because the second one is loosing the martingale property. This will be clear when we shall try to get conditions for (8.3) or (8.4) to hold.

The main additional assumption made by the authors is then

$$M_t \text{ is bounded by an universal constant } C(t)\,. \tag{8.5}$$

Let us follow the proof of Lemme 2.1 in [13].
Using the Markov property and Hölder's inequality we can write for $q > 2$, $s > q$ and r the conjugate of s ,

$$\begin{aligned}
\int (P^F_{2t}(|f|))^q \, d\nu &\leq \int e^{(q-2)F} \, (C(t))^q \, \Big(E^{\mathbb{P}_x}[M_t \, (P_t(|f| \, e^{-F}))(X_t)]\Big)^q \, d\mu \\
&\leq \int e^{(q-2)F} \, (C(t))^q \, \Big(E^{\mathbb{P}_x}[M_t^r]\Big)^{\frac{q}{r}} \, \Big(E^{\mathbb{P}_x}[(P_t(|f| \, e^{-F}))^s(X_t)]\Big)^{\frac{q}{s}} \, d\mu \\
&\leq (C(t))^q \, \Big(\int P_t\big((P_t(|f| \, e^{-F}))^s\big) \, d\mu\Big)^{\frac{q}{s}} \, \Big(\int e^{\frac{s(q-2)}{s-q} \, F} \mathbb{E}^{\mathbb{P}_x}[M_t^{\frac{sq}{s-q}}] \, d\mu\Big)^{\frac{s-q}{s}} \, .
\end{aligned}$$

It immediately follows

Proposition 8.6. *Assume that conditions (1),(2) and (3) in Lemma 7.1 are fulfilled for some probability measure μ and some $-F \in D(A)$. Assume in addition that e^{-F} is bounded, (8.5) is in force and $i(p,t) < +\infty$ in (8.4) (for some $p > 2$ and all $t > 0$ large enough).*
Then if μ satisfies LSI, so does $\nu = e^{-2F} \mu$.
If μ satisfies TLSI and $\Gamma(F,F) \in \mathbb{L}^{p'}(\mu)$ for some $p' > 1$, then ν satisfies TLSI.

Proof. Since μ satisfies LSI, P_t is hypercontractive, thanks to Gross theorem, hence

$$\Big(\int P_t\big((P_t(|f| \, e^{-F}))^s\big) \, d\mu\Big)^{\frac{q}{s}} \leq \parallel f \parallel^{\frac{q}{s}}_{\mathbb{L}^2(\nu)} \, ,$$

(use first the contraction property on $\mathbb{L}^s(\mu)$). The other terms are bounded (we have to use the boundedness of e^{-F} for the third one). The first part thus follows from Proposition 2.7. The final part of the proposition comes from Proposition 7.13 and Proposition 5.13. □

Remark 8.7. Actually for the first part of the proposition we do not need that μ is a probability measure. The second part (concerned with TLSI) however is not really tractable without this assumption.
The main problem with (8.4) is that it is much more difficult to check than (8.3). It is thus interesting to see what happens if we only assume (8.3). It is immediately seen that all the problems are coming from Hölder, so that in order to improve proposition 8.6 we have to assume a stronger assumption for P_t namely ultracontractivity. Just looking at the previous proof we see that the following holds

Proposition 8.8. *Assume that conditions (1),(2) and (3) in Lemma 7.1 are fulfilled for some nonnegative measure μ and some $-F \in D(A)$. Assume in addition that e^{-F} is bounded, (8.5) is in force and (8.3) holds (for some $p > 2$ and all $t > 0$ large enough).*
Then if P_t is ultracontractive (for some t), $\nu = e^{-2F} \mu$ satisfies LSI. One can also here give a bound, namely

$$\parallel P^F_{2t} \parallel_{\mathbb{L}^2 \to \mathbb{L}^q} \leq C(t) \, K(t) \, \parallel e^F \, \mathbb{E}^{\mathbb{P}_x}[M_t] \parallel_{\mathbb{L}^q(\nu)} \, ,$$

where $K(t)$ is the constant of ultracontractivity.
If in addition $\Gamma(F,F) \in \mathbb{L}^{p'}(\mu)$ for some $p' > 1$ and μ is a probability measure, then ν satisfies TLSI.

Examples of $\mathbb{L}^\infty$ bounds (instead of $\mathbb{L}^p$) in (8.3) are given in [13], as well as some sufficient convexity assumptions on F in order (8.5) to hold. Before to look at such examples, let us see what happens when one uses the ideas of section 2, instead of the Markov property.

Let us recall (2.4) in the appropriate form (recall that $\int f^2\,d\nu = 1$ and f is nonnegative),

$$\int (\sum_j log\, h_j)\, f^2 d\nu \leq \frac{1}{2}\, t\, \mathcal{E}_F(f,f) + log \int f^2(x) h_1(x)\, e^F(x)\, M(t,x) \nu(dx)\,. \tag{8.9}$$

where

$$M(t,x) = \mathbb{E}^{\mathbb{P}_x}[\Pi_{j\geq 2}\, h_j(X_{t_j})\, e^{-F}(X_t)\, M_t]\,.$$

Choose first $j = 1,2$ and $t_2 = t$. Then, denoting by q the conjugate of p, it holds

$$\begin{aligned} \int f^2(x)h_1(x)\, e^{-F}(x)\, M(t,x)\mu(dx) \;&\leq\; \int f^2 h_1\, e^{-F}\, \left(P_t(h_2^q\, e^{-qF})\right)^{\frac{1}{q}} \left(\mathbb{E}^{\mathbb{P}_x}[M_t^p\,]\right)^{\frac{1}{p}} d\mu\,, \\ &\leq\; \int f^2 h_1\, \left(P_t(h_2^q\, e^{-qF})\right)^{\frac{1}{q}} \left(e^{pF}\, \mathbb{E}^{\mathbb{P}_x}[M_t^p\,]\right)^{\frac{1}{p}} d\nu\,, \\ &\leq\; \left(i(p,t)\right)^{\frac{1}{p}} \left(\int f^{2q} h_1^q \left(P_t(h_2^q\, e^{-qF})\right) d\nu\right)^{\frac{1}{q}}. \end{aligned} \tag{8.10}$$

It is thus natural to choose

$$h_2 = f^{\frac{1}{q}} \quad , \quad h_1 = f^{\alpha-1}$$

for some $\alpha < 1$, such that $q(1+\alpha) < 2$. Using one more time Hölder, we obtain

$$\int f^2(x) h_1(x)\, e^{-F}(x)\, M(t,x)\mu(dx) \;\leq\; \left(i(p,t)\right)^{\frac{1}{p}} \left(\int \left(P_t(f\, e^{-qF})\right)^{\frac{2}{2-q(1+\alpha)}} d\nu\right)^{\frac{2-q(1+\alpha)}{2q}}. \tag{8.11}$$

If we assume in addition (as before) that $F \geq -\log(K)$ (i.e. $e^{-F} \leq K$). We thus obtained

$$\begin{aligned} \frac{\beta}{2} \int f^2 log f^2\, d\nu \leq \frac{1}{2}\, t\, \mathcal{E}_F(f,f) &+ \frac{2-q(1+\alpha)}{2q}\, log\Big(\int \left(P_t(f\, e^{-F})\right)^{\frac{2}{2-q(1+\alpha)}} d\mu\Big) \\ &+ \frac{1}{p}\, log(i(p,t)) + c(q,\alpha,K)\,, \end{aligned} \tag{8.12}$$

with $\beta = \frac{1}{q} + (\alpha - 1)$. For (8.12) to be useful, we have to choose $\alpha > 0$ such that

$$q(1+\alpha) < 2 \quad \text{and} \quad \frac{1}{q} + (\alpha - 1) > 0\,,$$

hence

$$1 - \frac{1}{q} < \alpha < \frac{2}{q} - 1$$

that imposes $q < \frac{3}{2}$ and thus $p > 3$.

We thus have obtained a different version of Proposition 8.6

Proposition 8.13. *The statements of Proposition 8.6 remain true without assuming (8.5), provided $i(p,t) < +\infty$ in (8.4) for some $p > 3$ and all $t > 0$ large enough.*

Of course as we did from the beginning we did not try to obtain sharp estimates, so that the value 3 for p is a little bit artificial. The key point in proposition 8.13 is that the boundedness assumption (8.5) is no more assumed. Also notice that we assumed that $i(p,t) < +\infty$ for all large enough t, in order to use the hypercontractivity property for P_t. Of course if this property holds for all t and all $\mathbb{L}^p$ norm (in particular if P_t is ultracontractive for all $t > 0$) it is enough to check (8.4) for one t. But again (8.4) is much stronger than (8.3). We shall try to better explore (8.9).

For convenience we write

$$G(t,x) = P_t^F(e^F)(x) = e^F(x)\,\mathbb{E}^{\mathbb{P}_x}[M_t]\,,$$

and will assume that e^{-F} is bounded by K. If we choose

$$h_1 = f^{-\alpha}\,e^{-\beta F} \quad , \quad h_2 = f^{\delta}\,,$$

(8.9) becomes

$$\text{(8.14)}\quad (\delta-\alpha)\int f^2\,log\,f d\nu \;\leq\; \frac{1}{2}\,t\,\mathcal{E}_F(f,f) + \beta\int F\,f^2\,d\nu + \\ + log\int f^{2-\alpha}(x)\,e^{(1-\beta)F}(x)\,\mathbb{E}^{\mathbb{P}_x}[f^{\delta}(X_t)\,e^{-F}(X_t)\,M_t]\nu(dx)\,.$$

But using the martingale property of $t \to e^{-F}(X_t)\,M_t$, the Markov property and the bound for e^{-F}, one obtains for all nonnegative u, and all conjugate pair (p,q)

$$\text{(8.15)}\qquad \mathbb{E}^{\mathbb{P}_x}[f^{\delta}(X_t)\,e^{-F}(X_t)\,M_t] \;\leq\; K\,\mathbb{E}^{\mathbb{P}_x}[f^{\delta}(X_t)\,e^{-F}(X_t)\,M_t\,G(u,X_t)]$$

$$\leq\; K\,e^{-F}(x)\,\mathbb{E}^{\mathbb{P}_x^F}[f^{\delta}(X_t)\,G(u,X_t)]$$
$$\leq\; K\,e^{-F}(x)\left(\mathbb{E}^{\mathbb{P}_x^F}[f^{q\,\delta}(X_t)\,e^{-q\theta\,F}(X_t)]\right)^{\frac{1}{q}}\left(\mathbb{E}^{\mathbb{P}_x^F}[e^{p\theta\,F}(X_t)\,G^p(u,X_t)]\right)^{\frac{1}{p}},$$

so that

$$\int f^{2-\alpha}(x)\,e^{(1-\beta)F}(x)\,\mathbb{E}^{\mathbb{P}_x}[f^{\delta}(X_t)\,e^{-F}(X_t)\,M_t]\nu(dx) \;\leq$$
$$\leq\; K\Big(\int f^{(2-\alpha)q}\,e^{-q\beta F}\mathbb{E}^{\mathbb{P}_x^F}[f^{q\,\delta}(X_t)\,e^{-q\theta F}(X_t)]\,\nu(dx)\Big)^{\frac{1}{q}}\Big(\int \mathbb{E}^{\mathbb{P}_x^F}[e^{p\theta F}(X_t)\,G^p(u,X_t)]\,\nu(dx)\Big)^{\frac{1}{p}}$$
$$\leq\; K\Big(\int f^{(2-\alpha)q}\,e^{-(q\beta+1)F}\mathbb{E}^{\mathbb{P}_x}[f^{q\,\delta}(X_t)\,e^{-(q\theta+1)F}(X_t)M_t]\,\mu(dx)\Big)^{\frac{1}{q}}\Big(\int e^{p\theta F}\,G^p(u,x)\,\nu(dx)\Big)^{\frac{1}{p}}.$$

In the last inequality we have used the fact that ν is P_t^F stationary. Now we see that it is natural to impose (8.5) and (8.3). This time we will be accurate with the constants.
First, provided $p\,\theta < 2$ (i.e $e^{p\theta F}$ has some ν moment) we have

$$\text{(8.16)}\qquad \int e^{p\theta F}\,G^p(u,x)\,\nu(dx) \leq \Big(\int G^{p(\frac{2}{2-p\theta})}(u,x)\,\nu(dx)\Big)^{\frac{2-p\theta}{2}}.$$

Here we are using the fact that μ is a probability measure.

Next if we choose $q\delta = q\theta + 1 < 2$ and $(2-\alpha)q = q\beta + 1 < 2$, we have

$$\text{(8.17)}\qquad \int f^{(2-\alpha)q}\,e^{-(q\beta+1)F}\,\mathbb{E}^{\mathbb{P}_x}[f^{q\,\delta}(X_t)\,e^{-(q\theta+1)F}(X_t)M_t]\,\mu(dx) \leq$$

$$\begin{aligned} &\leq \; C(t) \int (fe^{-F})^{(2-\alpha)q} \, P_t\big((fe^{-F})^{q\delta}\big) \, d\mu \\ &\leq \; C(t) \left(\int \left(P_t\big((fe^{-F})^{q\delta}\big)\right)^{\frac{2}{2-(2-\alpha)q}} \, d\mu \right)^{\frac{2-(2-\alpha)q}{2}} , \end{aligned}$$

where $C(t)$ is defined by (8.5).

(8.16) and (8.17) are showing that the third term in the sum in (8.14) will be controlled thanks to (8.3) and to the hypercontractivity of P_t.

But it remains to control the second term. To this end just use Young's inequality

$$fF \leq f \, log(f) - f + e^F \, .$$

It yields

$$\int f^2 \, F \, d\nu \leq \int f^2 \, log(f) \, d\nu - 1 + \int f \, e^F \, d\nu \leq \int f^2 \, log(f) \, d\nu \, , \tag{8.18}$$

thanks to Cauchy-Schwarz (here again we are using the fact that μ is a probability measure).

Now plug (8.16) , (8.17) and (8.18) into (8.14). We get

$$\begin{aligned} (\delta - \alpha - \beta) \int f^2 \, log \, f d\nu \; &\leq \; \frac{1}{2} \, t \, \mathcal{E}_F(f,f) + logK + \frac{1}{q} \, log(C(t)) \\ &+ \; \frac{2 - p\theta}{2p} \, log\Big(\int G^{p(\frac{2}{2-p\theta})}(u,x) \, \nu(dx) \Big) \\ &+ \; \frac{2-(2-\alpha)q}{2q} \, log\Big(\int \Big(P_t\big((fe^{-F})^{q\delta}\big)\Big)^{\frac{2}{2-(2-\alpha)q}} \, d\mu \Big) \, . \end{aligned} \tag{8.19}$$

We shall now see that we can find constants such that (8.19) is interesting. Recall that we have to choose positive α, β, δ and θ such that

$$\begin{aligned} &p \, \theta < 2 \, , \\ &q\delta = q\theta + 1 < 2 \, , \\ &(2-\alpha)q = q\beta + 1 < 2 \, , \\ &\delta - \alpha - \beta > 0 \, , \end{aligned}$$

where (p,q) is a pair of conjugate real numbers.

Since $\alpha + \beta = 2 - \frac{1}{q}$ it follows $\frac{2}{q} > \delta > 2 - \frac{1}{q}$, hence $q < \frac{3}{2}$, i.e. $p > 3$ as before. We will choose

$$q = \frac{3}{2(1+\varepsilon)} \quad \text{hence } p = \frac{3}{1-2\varepsilon} \, , \quad \text{and} \quad \delta = \frac{4}{3} \, (1-\varepsilon) \, .$$

We then have

$$\alpha + \beta = \frac{4}{3} - 2\varepsilon \quad , \quad \theta = \frac{2}{3} \, (1 - 3\varepsilon) \, ,$$

so that $q\delta = 2 \, \frac{1-\varepsilon}{1+\varepsilon} \; < 2$, $p \, \theta = 2 \, \frac{1-3\varepsilon}{1-2\varepsilon} \; < 2$ and $\delta - \alpha - \beta = \frac{2}{3} \, \varepsilon$. Hence we may choose any $\beta < \frac{2(1+\varepsilon)}{3}$ and $\varepsilon < \frac{1}{3}$.

Even if the calculations are tedious let us see what is happening with (8.19).

First $(fe^{-F})^{q\delta} \in \mathbb{L}^{\frac{1+\varepsilon}{1-\varepsilon}}(\mu)$ and its norm in this space is equal to 1 (again we use that μ is a probability measure). So we get

$$\frac{\varepsilon}{3}\int f^2\, log(f^2)\, d\nu \leq \frac{1}{2}\, t\, \mathcal{E}_F(f,f) + logK + \frac{2(1+\varepsilon)}{3}\, log(C(t)) + \tag{8.20}$$
$$+ \frac{\varepsilon}{3}\, log\Big(\int G^{\frac{3}{\varepsilon}}(u,x)\, \nu(dx)\Big) + \frac{2(1+\varepsilon)}{3}\, log\Big(\parallel P_t \parallel_{\mathbb{L}^{\frac{1+\varepsilon}{1-\varepsilon}} \to \mathbb{L}^{\frac{2(1+\varepsilon)}{\varepsilon}}}\Big).$$

We thus have proved

Theorem 8.21. *Assume that conditions (1),(2) and (3) in Lemma 7.1 are fulfilled for some probability measure μ and some $-F \in D(A)$. Assume in addition that e^{-F} is bounded, (8.5) is in force and (8.3) holds for some $p > 9$ and some $t > 0$.*
Then if μ satisfies LSI, then so does $\nu = e^{-2F}\mu$.
If in addition μ satisfies TLSI and $\Gamma(F,F) \in \mathbb{L}^{p'}(\mu)$ for some $p' > 1$, then ν satisfies TLSI.
Conversely if ν satisfies LSI, (8.3) holds for all p and t large enough.

Remarks. In Proposition 3.7 of [13], the authors give some conditions for immediate hypercontractivity. The proof of proposition 8.6 is closed to their proof. However, (8.5) is crucial there, so that proposition 8.13 and its proof seems to be an improvement. Theorem 8.21 is however the good extension (for hypercontractivity) of the results in [13].

If μ is not bounded (say the Lebesgue measure), one can of course modify F (say replace F by $F + cx^2$) and check that the modified F is such that $\Gamma(F,F)$ belongs to some $\mathbb{L}^p$ for the modified measure (say some gaussian measure). In particular on $\mathbb{R}^d$ equipped with Lebesgue measure, as soon as $|\nabla F|$ is controlled by some polynomial, LSI and TLSI will be equivalent for $e^{-2F}\mu$. The situation for $F = |x|^\alpha$ is for instance completely known : for $\alpha \geq 2$, TLSI holds; for $0 < \alpha < 2$ LSI does not hold but for $1 \leq \alpha$ SGP holds (see [19] example 1.4.c). One can see (left to the reader) that the first part of Theorem 8.21 can be extended to a general nonnegative measure μ , provided e^{-F} belongs to some $\mathbb{L}^s(\mu)$ for some $s > 0$ (possibly less than 1) small enough. The constants have to be modified by introducing the related norm. We mentioned during the proof each time we have used the boundedness assumption for μ.

As we said in the introduction, we will not study here explicit examples. The reader will find in sections 3 and 4 of [13] very beautiful ideas for checking that (8.3) (in the stronger bounded form) is satisfied. Examples will be developed elsewhere.

References

[1] S. Aida, T. Masuda, and I. Shigekawa. Logarithmic Sobolev inequalities and exponential integrability. *J. Func. Anal.*, 126:83–101, 1994.

[2] S. Aida and I. Shigekawa. Logarithmic Sobolev inequalities and spectral gaps: perturbation theory. *J. Func. Anal.*, 126:448–475, 1994.

[3] D. Bakry. L'hypercontractivité et son utilisation en théorie des semi groupes, Ecole d'été de probabilités de Saint-Flour. *Lect. Notes Math.*, 1581:1–114, 1994.

[4] E. Carlen. Superadditivity of Fisher's information and Logarithmic Sobolev inequalities. *J. Func. Anal.*, 101:194–211, 1991.

[5] P. Cattiaux and F. Gamboa. Large deviations and variational theorems for marginal problems. *Bernoulli*, 5:81–108, 1999.

[6] P. Cattiaux and C. Léonard. Minimization of the Kullback information for general Markov processes. Séminaire de Probas XXX. *Lect. Notes Math.*, 1626:283–311, 1996.
[7] J. D. Deuschel. Algebraic $\mathbb{L}^2$ decay of attractive processes. *Ann. Prob.*, 22:264–283, 1994.
[8] J. D. Deuschel and D. W. Stroock. *Large deviations.* Academic Press, 1989.
[9] H. Föllmer. Random fields and diffusion processes, Ecole d'été de probabilités de Saint-Flour. *Lect. Notes Math.*, 1362:101–204, 1988.
[10] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet Forms and Symmetric Markov Processes.* Number 19 in Studies in Mathematics. Walter de Gruyter, Berlin New York, 1994.
[11] L. Gross. Logarithmic Sobolev inequalities and contractivity properties of semi-groups in Dirichlet forms. *Lect. Notes Math.*, 1563:54–88, 1993.
[12] A. Guionnet and B. Zegarlinski. Cours I.H.P. , Paris. 1999.
[13] O. Kavian, G. Kerkyacharian, and B. Roynette. Quelques remarques sur l'ultracontractivité. *J. Func. Anal.*, 111:155–196, 1993.
[14] M. Ledoux. Concentration of measure and logarithmic sobolev inequalities. Séminaire de Probas XXXIII. *Lect. Notes Math.*, 1709:120–216, 1999.
[15] T. M. Liggett. $\mathbb{L}^2$ rates of convergence for attractive reversible nearest particle systems. *Ann. Prob.*, 19:935–959, 1991.
[16] P. Mathieu. Convergence to equilibrium for spin glasses. To appear in Comm. Math. Phys.
[17] P. Mathieu. Hitting times and spectral gap inequalities. *Ann. Inst. Henri Poincaré. Prob. Stat.*, 33:437–465, 1997.
[18] P. Mathieu. Quand l'inégalité Log-Sobolev implique l'inégalité de trou spectral. Séminaire de Probas XXXII. *Lect. Notes Math.*, 1686:30–35, 1998.
[19] M. Röckner and F. Y. Wang. Weak Poincaré inequalities and L^2 convergence rates of Markov semigroups. To appear in J. Func. Anal.
[20] G. Royer. *Une initiation aux inégalités de Sobolev logarithmiques.* S.M.F., Paris, 1999.

Patrick CATTIAUX,, Ecole Polytechnique, CMAP, F- 91128 Palaiseau cedex, CNRS 756, and Université Paris X Nanterre, équipe MODAL'X, UFR SEGMI, 200 avenue de la République, F-92001 Nanterre, Cedex.

E-mail address: `cattiaux@cmapx.polytechnique.fr`