A Feynman-Kac approach for logarithmic Sobolev inequalities*

Clément Steiner†

Abstract

This note presents a method based on Feynman-Kac semigroups for logarithmic Sobolev inequalities. It follows the recent work of Bonnefont and Joulin on intertwining relations for diffusion operators, formerly used for spectral gap inequalities, and related to perturbation techniques. In particular, it goes beyond the Bakry-Émery criterion and allows to investigate high-dimensional effects on the optimal logarithmic Sobolev constant. The method is illustrated on particular examples (namely Subbotin distributions and double-well potentials), for which explicit dimension-free bounds on the latter constant are provided. We eventually discuss a brief comparison with the Holley-Stroock approach.

1 Introduction

Since their introduction by Gross in 1975, the Logarithmic Sobolev Inequalities (LSI) became a widely used tool in infinite dimensional analysis. Initially studied in relation to the hypercontractivity property for Markov semigroups, they turned out to be prominent in many various domains, at the interface of analysis, probability theory and geometry (one of the best example of such prominence being their use in Perelmann's proof of Poincaré's conjecture in [23]).

For μ a probability measure on the Euclidean space $(\mathbf{R}^d, |\cdot|)$, this inequality provides a control on the entropy of any smooth function f in term of its gradient:

$$\mathrm{Ent}_\mu(f^2) \leq c \int_{\mathbf{R}^d} |\nabla f|^2 d\mu,$$

for some $c > 0$, where $\mathrm{Ent}_\mu(f^2) = \int_{\mathbf{R}^d} f^2 \log(f^2)\, d\mu - \left(\int_{\mathbf{R}^d} f^2\, d\mu\right) \log\left(\int_{\mathbf{R}^d} f^2\, d\mu\right)$. The optimal constant for the latter inequality to hold, often called the *logarithmic Sobolev constant* and denoted $c_{LSI}(\mu)$, is of primary importance in the study of the measure μ, since it encodes many of its properties. For instance, among many results in this area, Otto and Villani established in [22] a connection between LSI and some transportation inequalities (see also the related work by Bobkov and Götze in [9]), and Herbst provided a powerful argument that links LSI to Gaussian concentration inequalities (see the lecture notes by Ledoux [20] for more details and his reference monograph [21] about concentration of measure).

The case where μ is the invariant measure of some Markov process is also of great interest. For example, apart from Gross' initial results on hypercontractivity in [17], $c_{LSI}(\mu)$ encodes the decay in entropy of the related semigroup, and is linked to the Fisher information (defined for a positive function f as $\int_{\mathbf{R}^d} |\nabla \sqrt{f}|^2 d\mu$) through de Bruijn's identity. Significant advances in this setting were due to Bakry and Émery in [4], who stated their eponymous criterion, also known as "curvature-dimension criterion", that connects the logarithmic Sobolev inequality (and many functional inequalities) to geometric properties of μ. We refer to [5] for a comprehensive overview of this theory.

Although the value of $c_{LSI}(\mu)$ is key in the study of μ, its exact value is hardly ever known explicitly. Bakry-Émery theory provides sharp estimates on this constant for some log-concave measures, assumption that might be weakened according to some perturbation arguments. More precisely, although the Bakry-Émery criterion is defined in a more general situation, it can be reformulated conveniently in the Euclidean setting as follows.

Theorem (Bakry-Émery, [4]). Assume that $\mu(dx) \propto e^{-V(x)} dx$, for some smooth potential V. If there exists some $\rho > 0$ such that $\nabla^2 V(x) \geq \rho I_d$ for any $x \in \mathbf{R}^d$ (the Hessian matrix of V is uniformly bounded from below as a symmetric matrix), then μ satisfies a LSI with constant $2/\rho$.

We refer to [5] §5.7 for the general curvature-dimension criterion. We shall stick from know on to the assumption that $\mu(dx) \propto e^{-V(x)} dx$, for some smooth potential V. In particular, this bound is sharp for the standard Gaussian distribution γ, providing $c_{LSI}(\gamma) = 2$ (whatever the dimension of the underlying space is). Unfortunately, this criterion fails as soon as V is not uniformly convex. Yet, if this "lack of convexity" can be balanced by a bounded transformation, one may use perturbation techniques, such as the well-known Holley-Stroock method.

Theorem (Holley-Stroock, [19]). Assume that $d\mu \propto e^\Phi d\nu$ where ν is a probability measure that satisfies a LSI and Φ is continuous and bounded. Then μ satisfies a LSI with $c_{LSI}(\mu) \leq e^{2(\sup(\Phi) - \inf(\Phi))} c_{LSI}(\nu)$.

Note that perturbation by unbounded functions (under for example growth assumptions) has been studied, see for example [6]. However, authors in the latter explain that their method weakens the inequality as soon as the perturbation is not bounded. Nevertheless, the LSI can be preserved by unbounded perturbation in some specific cases, as will be developed in this article.

Apart from perturbation, stability of LSI by tensorization is also a key property of such inequalities, since it exhibits dimension-free behaviours for product measures, but fails in general to provide efficient bounds beyond this case. In particular, one may wish to keep track of the geometry of μ (dimension of the space, log-concavity, curvature, etc.) through $c_{LSI}(\mu)$, which can be difficult in many settings (as will be discussed in Section 4). For further reading, we refer to [5] §5.7 for detailed results and to the remarkably synthetic monograph [1] for a broader introduction.

In this note, we provide a probabilistic approach based on the study of some Feynman-

Kac semigroups to derive new estimates on the logarithmic Sobolev constant. It follows the recent work of Bonnefont and Joulin involving intertwinings and functional inequalities of spectral flavour [10, 11] and extends their approach to the latter. A somewhat similar approach can also be found in the recent work of Sturm and his collaborators on metric measure spaces [14]. Let us give an overview of our main results. They will be properly stated and proved in Section 3.

We first show a representation theorem for Feynman-Kac semigroups acting on gradient fields. Namely, for a perturbation function a satisfying some regularity and growth assumptions, the following result holds.

Theorem. There exist a stochastic process $(X_{t,a})_{t\geq 0}$, a martingale $(R_{t,a})_{t\geq 0}$ and a matrix-valued process $(J_t^{X_a})_{t\geq 0}$ such that for any smooth function f, one has

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f) = \mathbf{E}[R_{t,a} J_t^{X_a} \nabla f(X_{t,a})],$$

where $(\mathcal{P}_t^{\nabla^2 V})_{t\geq 0}$ is the Feynman-Kac semigroup of interest.

The invariant measure of the above process is known and closely related to μ. The martingale is given by Girsanov's theorem, while the matrix-valued process can be seen as the Jacobian matrix of $(X_{t,a})_{t\geq 0}$ (with respect to the initial condition).

Note that this formula can be related to other forms of derivatives of heat semigroups, including for example the well-known Bismut formula (as presented for instance in [16]). Originally derived using Malliavin calculus (see [8]), Elworthy and Li emphasize in [16] a more geometric approach and our proof relates to the differentiation of the flow of some stochastic differential equation (as presented for example in [24] §V.7). Yet, the above expression as a Feynman-Kac semigroup acting on a gradient field is particularly suitable when one aims to infer a logarithmic Sobolev inequality. In particular, this probabilistic representation allows to obtain Grönwall-type estimates on the semigroup, that lead to a new criterion for LSI. Namely, for a perturbation function a satisfying some growth assumptions, we can define a curvature $\kappa_a \in \mathbf{R}$ (depending on a and $\nabla^2 V$) which provides a Bakry-Émery-like condition.

Theorem. If $\kappa_a > 0$, then μ satisfies a LSI with $c_{LSI}(\mu) \leq C_a/\kappa_a$, for some $C_a > 0$.

This result indeed encompasses the Bakry-Émery criterion (taking $a \equiv 1$). Note that we derive, in the specific case of monotonic functions, a very similar result, yet allowing the function a to be unbounded.

The choice of a (provided that technical assumptions are satisfied) in the latter theorem is rather free, so that one expects to take it such that C_a/κ_a is minimal (to get the sharpest bound on $c_{LSI}(\mu)$). The precise value of C_a and κ_a and their behaviour with respect to a are discussed in more details around two examples.

The first one is the quadric potential, that is $V(x) = |x|^4/4$; the second is the double-well: $V(x) = |x|^4/4 - \beta|x|^2/2$ (for $\beta \in (0, 1/2)$). Bakry-Émery criterion fails in both cases, yet our main theorem applies and we manage to infer the following behaviour of $c_{LSI}(\mu)$.

Theorem. • Assume that $V(x) = |x|^4/4$. Then μ satisfies a LSI and $c_{LSI}(\mu)$ does not depends on the dimension.
• Assume that $V(x) = |x|^4/4 - \beta|x|^2/2$, $\beta \in (0, 1/2)$. Then μ satisfies a LSI and $c_{LSI}(\mu)$ only depends on β.

We briefly compare both results with the Holley-Stroock method, and provide explicit constants.

The article is organised as follows. We introduce in Section 2 the framework of the paper, along with some results about intertwinings and Feynman-Kac semigroups. In Section 3, we properly state and prove our main results and discuss a comparison with the Holley-Stroock approach. Finally, Section 4 is devoted to examples, where explicit constants and detailed computations are provided.

2 Basic framework

In this first section, we recall the framework of our analysis, basic results and definitions about intertwinings and Feynman-Kac semigroups (as introduced in [10, 2]).

2.1 Setting

The whole analysis shall be performed on the d-dimensional Euclidean space $(\mathbf{R}^d, |\cdot|)$, for $d \in \mathbf{N}^\star$. We let $\mathcal{C}^\infty(\mathbf{R}^d, \mathbf{R})$ and $\mathcal{C}^\infty(\mathbf{R}^d, \mathbf{R}^d)$ be respectively the set of infinitely differentiable functions and vector fields on $\mathbf{R}^d$, and let $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ and $\mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ denote respectively the set of compactly supported and positive $\mathcal{C}^\infty$ functions on $\mathbf{R}^d$. We endow those spaces with the supremum norm $\|\cdot\|_\infty$. We consider throughout this article a probability measure μ on $\mathbf{R}^d$ whose density with respect to the Lebesgue measure is proportional to e^{-V}, for some potential V at least twice differentiable. To this measure, one can associate a Markov diffusion operator defined as

$$\mathbf{L} = \Delta - \nabla V \cdot \nabla,$$

where Δ and ∇ respectively stand for the usual Laplace operator and gradient on $\mathbf{R}^d$. The flow of the equation $\partial_t u = \mathbf{L}u$ over $\mathbf{R}_+$ defines a Markov semigroup $(\mathbf{P}_t)_{t\geq 0}$, invariant with respect to μ, which is, under standard assumptions on V, ergodic in $L^2(\mu)$. Such assumptions include for example that $\mathbf{L}$ vanishes only for constant functions and the latter are stable by $(\mathbf{P}_t)_{t\geq 0}$. See [5] §3.1.9 for a general result. Moreover, this semigroup describes the dynamics of a diffusion process $(X_t^x)_{t\geq 0}$ that solves the following Stochastic Differential Equation (SDE):

$$dX_t^x = \sqrt{2}\, dB_t - \nabla V(X_t^x)dt, \quad X_0^x = x \in \mathbf{R}^d \text{ a.s.,} \qquad (E)$$

where $(B_t)_{t\geq 0}$ denotes the standard d-dimensional Brownian motion. All stochastic processes are defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$, and we let $(\mathcal{F}_t)_{t\geq 0}$ denote the natural (completed) filtration associated to $(B_t)_{t\geq 0}$. Under mild assumptions on V, this process is non-explosive and converges in distribution towards μ, its invariant distribution. Moreover, regularity of V ensures that $x \mapsto X_t^x$ is (at least) differentiable over $\mathbf{R}^d$, for any $t \geq 0$. See Remarks 2.3 and 2.4 at the end of this section for more informations and references about non-explosion and regularity w.r.t. the initial condition.

In addition, $\mathbf{L}$ is symmetric on $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ with respect to μ, and the integration by parts formula rewrites as follows: for any $f, g \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$,

$$\int_{\mathbf{R}^d} f\mathbf{L}g \, d\mu = \int_{\mathbf{R}^d} g\mathbf{L}f \, d\mu = -\int_{\mathbf{R}^d} \nabla f \cdot \nabla g \, d\mu.$$

In particular, $\mathbf{L}$ is non-positive on $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$. Hence by completeness, this operator admits a unique self-adjoint extension (which shall still be denoted $\mathbf{L}$) on some domain $\mathcal{D}(\mathbf{L}) \subset L^2(\mu)$ for which $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ is a core, *i.e.* is dense for the norm induced by $\mathbf{L}$ (see [5] §3.1.8 and thereafter for more precise informations).

Finally, let us recall the definition of the logarithmic Sobolev inequality we will refer to.

Definition 2.1. *The measure μ is said to satisfy a* Logarithmic Sobolev Inequality *(in short LSI) with constant $c > 0$ if for any $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ one has*

$$\mathrm{Ent}_\mu(f^2) \leq c \int_{\mathbf{R}^d} |\nabla f|^2 d\mu.$$

We let $c_{LSI}(\mu)$ denote the optimal constant in the latter inequality, which we may as well refer as the logarithmic Sobolev constant.

Remark 2.2 (Embeddings and integrability)**.** The integration by parts formula entails that the bilinear form $(f, g) \mapsto -\int_{\mathbf{R}^d} f\mathbf{L}g\, d\mu$ extends likewise on some domain in which $\mathcal{D}(\mathbf{L})$ is dense for the $H^1(\mu)$ norm (see [5] §3.3.2 for more detailed statements). Hence, in some way, the LSI may be seen as a continuous embedding of $H^1(\mu)$ into some Orlicz space (see [25]), in quite a similar way as Sobolev inequalities provide a continuous (and compact, with Rellich-Kondrachov theorem) embedding of $H^1(\mu)$ into some $L^p(\mu)$ spaces (see for example [5] §6.4). Similarly, through Herbst's argument, LSI implies that the square of 1-Lipschitz functions is exponentially integrable (see [5] §5.4), and thus Gaussian concentration for μ, whereas a Sobolev inequality implies that such functions are actually bounded in $H^1(\mu)$ (see [5] §6.6). One may as well compare both Sobolev and logarithmic Sobolev inequalities to the (weaker) Poincaré inequality, we refer the interested reader to [5] §4.4 for further information.

We end this setting section with some details and references about diffusion processes.

Remark 2.3 (Diffusion processes: non-explosion)**.** The explosion time of the process $(X_t)_{t \geq 0}$ is defined as $\tau_e = \inf\{t \geq 0 : \limsup_{s \to t} |X_s| = +\infty\}$ (the definition is quite similar to the classical ODE one, except that τ_e is here a stopping time w.r.t. $(\mathcal{F}_t)_{t \geq 0}$). The process is said to be non-explosive (in finite time) as soon as τ_e is almost surely infinite. This is actually equivalent to the *mass preservation* for $(\mathbf{P}_t)_{t \geq 0}$, that is, $\mathbf{P}_t\mathbf{1} = \mathbf{1}$ for any $t \geq 0$, with $\mathbf{1}$ the constant function equal to 1 (understood as the increasing limit of a sequence of compactly supported $\mathcal{C}^\infty$ functions). Indeed, for any $t \geq 0$, $\mathbf{P}_t\mathbf{1} = \mathbf{P}(t \leq \tau_e)$, so that $\mathbf{P}_t\mathbf{1} = \mathbf{1}$ for any $t \geq 0$ if and only if $\tau_e = +\infty$ almost surely. This property is somewhat easier to handle, and Bakry inferred in [3] the following criterion: $(X_t)_{t \geq 0}$ is non-explosive as soon as there is $\rho \in \mathbf{R}$ such that $\nabla^2 V(x) \geq \rho I_d$ for any $x \in \mathbf{R}^d$. Note that ρ is not required to be positive, making it a very general condition. Roughly speaking, it states that V should not be "too concave".

From now on, we will assume that the latter is satisfied. This is for example true for $V(x) \propto |x|^\alpha$ for $\alpha > 1$, as shall be made clear in Section 4.

Remark 2.4 (Diffusion processes: initial condition)**.** As quickly mentioned above, as long as the process $(X_t)_{t \geq 0}$ is non-explosive and ∇V is smooth enough, the function $x \mapsto X_t^x(\omega)$ (for any fixed $t \geq 0$ and almost any fixed $\omega \in \Omega$) is differentiable on $\mathbf{R}^d$. Actually, as mentioned in [5] §B.4, this application is as smooth as ∇V is (up to the explosion time). In the following, we may only focus on the first order derivative (also know as tangent process or tangent flow), but general results for any order of differentiation can be found in [24] §V.7 (Theorems 39 and 40).

2.2 Intertwinings

We now focus on intertwinings (for a comprehensive introduction, see [10, 2]). Basically, we are interested in commutation relations between gradients and Markov generators, which give rise to the so-called Feynman-Kac semigroups. In the following proposition, we introduce some notations related to tensor operators and recall a chain rule commutation formula.

Proposition-definition 2.5. *In the following, we let $\mathcal{L}$ denote the tensorized operator $\mathbf{L}^{\otimes d}$ and $(\mathcal{P}_t)_{t \geq 0}$ be the associated Markov semigroup, that both act on vector fields. For $F = (F_1, \ldots, F_d) \in \mathcal{C}^\infty(\mathbf{R}^d, \mathbf{R}^d)$, they write as*

$$\mathcal{L}F = (\mathbf{L}F_1, \ldots, \mathbf{L}F_d) \ \text{ and } \ \mathcal{P}_t F = (\mathbf{P}_t F_1, \ldots, \mathbf{P}_t F_d).$$

For $f \in \mathcal{C}^\infty(\mathbf{R}^d, \mathbf{R})$, we recall the intertwining relation:

$$\nabla \mathbf{L}f = (\mathcal{L} - \nabla^2 V)(\nabla f),$$

where $\nabla^2 V \nabla f$ is the standard matrix-vector product. Similarly, the Feynman-Kac semi-group $(\mathcal{P}_t^{\nabla^2 V})_{t \geq 0}$ associated to $\mathcal{L} - \nabla^2 V$ satisfies the following identity:

$$\nabla \mathbf{P}_t f = \mathcal{P}_t^{\nabla^2 V}(\nabla f), \quad t \geq 0,$$

provided that f has compact support.

This idea takes roots in various works in differential geometry and operators analysis, and relates (in some more general setting) to the Bochner-Lichnerowicz-Weitzenböck formula (see [13] for an enthusiast introduction). See also the works around Witten Laplacians arising in statistical mecanics, for which we refer to Helffer's monograph [18].

Remark 2.6. We can still define the Feynman-Kac semigroup associated to $\mathcal{L}$ and a general smooth map $M : \mathbf{R}^d \to \mathcal{M}_d(\mathbf{R})$ as the flow of the following PDE system:

$$\begin{cases} \partial_t u &= (\mathcal{L} - M)u \\ u(0, \cdot) &= u_0, \end{cases}$$

denoted by $(\mathcal{P}_t^M)_{t \geq 0}$, provided that solutions to this system exist at any time. Such an extension will be implicitly used later.

Remark 2.7. The original Feynman-Kac formula, that arises in quantum mechanics, is stated for scalar-valued functions f and m and writes as follows (see [5] §1.15.6):

$$\mathcal{P}_t^m f = \mathbf{E}\left[f(X_t) e^{-\int_0^t m(X_s) ds} \right].$$

We call $(\mathcal{P}_t^{\nabla^2 V})_{t \geq 0}$ a Feynman-Kac semigroup by analogy with this case (which shall clearly appear in the following), yet the representation of $(\mathcal{P}_t^{\nabla^2 V})_{t \geq 0}$ does not write as simply as the above. This is the object of the next section.

3 Main results

In this section, we state and prove our main results in two steps: we first provide a representation theorem, related to Feynman-Kac semigroups, then apply it to estimates on the logarithmic Sobolev constant.

In [28], Wang developed a somehow similar approach in the framework of manifolds with boundaries, based on the Girsanov's theorem for reflected processes. Yet in our case, we take advantages of some properties of the semigroup, namely invariance and ergodicity.

3.1 Representation of Feynman–Kac semigroups

This first part is devoted to the main representation theorem we shall make use of. It is presented for Feynman-Kac semigroups acting on gradients, but still holds for more general vector fields (in which case the proof relies on a classical martingale argument).

The perturbation technique that will be set up in the next section strongly relies on a Girsanov representation of the semigroup $(\mathbf{P}_t)_{t \geq 0}$. To this end, we introduce a smooth perturbation function in V and study the relation between $(X_t)_{t \geq 0}$ and the process obtained from this new potential.

Definition 3.1. *Let $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$. We let $(X_{t,a})_{t \geq 0}$ denote the solution of the SDE*

$$dX_{t,a} = \sqrt{2} dB_t - \nabla V_a(X_{t,a}) \, dt,$$

where $V_a = V + \log(a^2)$.

Straightforward computations show that the generator of this process writes down

$$\mathbf{L}_a = \mathbf{L} - 2\frac{\nabla a}{a} \cdot \nabla,$$

and we let $(\mathbf{P}_{t,a})_{t\geq 0}$ denote the associated Markov semigroup (in particular, for any $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, $\mathbf{P}_{t,a}f = \mathbf{E}[f(X_{t,a})]$). Moreover, if μ_a is defined such that $d\mu_a/d\mu = 1/a^2$, then $(\mathbf{P}_{t,a})_{t\geq 0}$ is μ_a-invariant and $\mathbf{L}_a$ is (essentially) self-adjoint in $L^2(\mu_a)$ (see as well [5] §3.1.8). Note that μ_a is not a probability (or even finite) measure a priori.

Provided that everything is well-defined, the intertwining relation of Proposition-definition 2.5 for semigroups is also available for $\mathbf{P}_{t,a}f$, and writes as follows:

$$\nabla \mathbf{P}_{t,a}f = \mathcal{P}_{t,a}^{\nabla^2 V_a}(\nabla f).$$

Before we state the main theorem of this section, let us define a condition on the perturbation function that naturally arises in the computations involving Girsanov's theorem.

Definition 3.2. *A function $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ is said to satisfy the* (G) condition *whenever $|\nabla a|/a$ is bounded.*

We can now turn to the representation result (the first theorem stated in the introduction). To fix the ideas, we may write down the initial condition in the following statements, and omit it in the proofs.

Theorem 3.3. *Let $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ and $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ satisfying (G). Then for any $t \geq 0$, $x \in \mathbf{R}^d$,*

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f)(x) = \mathbf{E}[R_{t,a}^x J_t^{X_a^x} \nabla f(X_{t,a}^x)],$$

where $(R_{t,a}^x)_{t\geq 0}$ is a martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$ defined as

$$R_{t,a}^x = \frac{a(X_{t,a}^x)}{a(x)} \exp\left(-\int_0^t \frac{\mathbf{L}_a a(x)}{a(x)}(X_{s,a}^x)\, ds\right), \quad t \geq 0,$$

and $(J_t^{X_a^x})_{t\geq 0}$ is a matrix-valued process that solves

$$\begin{cases} dJ_t^{X_a^x} &= -J_t^{X_a^x} \nabla^2 V(X_{t,a}^x)dt, \quad t > 0 \\ J_0^{X_a^x} &= I_d. \end{cases}$$

As mentioned before, this result is based on Girsanov's theorem (see [24] §III.8 Theorem 46 for a proper statement). Hence, before we turn to its proof, we need the following lemma, that establishes a relation between the Markov semigroups $(\mathbf{P}_t)_{t\geq 0}$ and $(\mathbf{P}_{t,a})_{t\geq 0}$.

Lemma 3.4. *Let $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ satisfying the (G) condition. Then for any function $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, any $t \geq 0$, $x \in \mathbf{R}^d$:*

$$\mathbf{P}_t f(x) = \mathbf{E}[f(X_t^x)] = \mathbf{E}\left[R_{t,a}^x f(X_{t,a}^x)\right],$$

where $(R_{t,a}^x)_{t\geq 0}$ is the $(\mathcal{F}_t)_{t\geq 0}$-martingale defined above.

Proof. We first set up a suitable exponential martingale before we identify the involved probability distributions with Girsanov's theorem. For the sake of legibility, the initial condition shall be omitted in the following.

We first apply Itō's formula to $\log(a(X_{t,a}))$:

$$\log a(X_{t,a}) = \log a + \sqrt{2}\int_0^t \nabla(\log a(X_{s,a})) \cdot dB_s + \int_0^t \mathbf{L}_a(\log a)(X_{s,a})\, ds.$$

Note that this decomposition is fairly general and is related to the martingale problem of Stroock and Varadhan, see [26] Chap. 6.

Expanding the right-hand side and taking exponential lead to the following expression for $R_{t,a}$:

$$R_{t,a} = \exp\left(\sqrt{2}\int_0^t \frac{\nabla a}{a}(X_{s,a}) \cdot dB_s - \int_0^t \left|\frac{\nabla a}{a}\right|^2(X_{s,a})\,ds\right).$$

The (G) condition ensures through standard arguments that the right-hand side is a true $(\mathcal{F}_t)_{t\geq 0}$-martingale, thus so is $(R_{t,a})_{t\geq 0}$. From now on, we set $Y_{t,a} = \sqrt{2}\frac{\nabla a}{a}(X_{t,a})$.

We let $\mathbf{Q}_a$ be the probability measure defined as

$$\left.\frac{d\mathbf{Q}_a}{d\mathbf{P}}\right|_{\mathcal{F}_t} = R_{t,a}.$$

According to Girsanov's theorem, the process $(\tilde{B}_t)_{t\geq 0}$ defined as

$$\tilde{B}_t = B_t - \int_0^t Y_{s,a}\,ds,$$

is a $\mathbf{Q}_a$-Brownian motion. Furthermore, the process $(X_{t,a})_{t\geq 0}$ solves the SDE

$$dX_{t,a} = \sqrt{2}d\tilde{B}_t - \nabla V(X_{t,a})\,dt,$$

hence the law of $X_{t,a}$ under $\mathbf{Q}_a$ coincides with the one of X_t under $\mathbf{P}$. In particular, for any $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$,

$$\mathbf{P}_t f = \mathbf{E}[f(X_t)] = \mathbf{E}[R_{t,a}f(X_{t,a})],$$

and the proof is complete. □

We can now prove Theorem 3.3.

Proof. Recall that under the aforementioned non-explosion assumptions, the diffusion process defined by Equation (E) is differentiable with respect to its initial condition (see Remarks 2.3 and 2.4), so that for any $t \geq 0$:

$$\begin{aligned}
\mathcal{P}_t^{\nabla^2 V}(\nabla f) &= \nabla \mathbf{P}_t f \\
&= \mathbf{E}[\nabla(f(X_t))] \\
&= \mathbf{E}[J_t^X \nabla f(X_t)],
\end{aligned}$$

where $(J_t^X)_{t\geq 0}$ denotes the (matrix-valued) tangent process to $(X_t)_{t\geq 0}$ (that is, the Jacobian matrix of X_t with respect to the initial condition). Differentiating with respect to the initial condition in the SDE (E) provides the following formula for J_t^X:

$$J_t^X = I_d - \int_0^t J_s^X \nabla^2 V(X_s)\,ds.$$

One can replace X_s by $X_{s,a}$ in the previous expression, to define as well

$$J_t^{X_a} = I_d - \int_0^t J_s^{X_a} \nabla^2 V(X_{s,a})\,ds.$$

Note that the potential V is unchanged in the equation. Lemma 3.4 implies then, since $R_{t,a}$ is scalar-valued,

$$\mathbf{E}[J_t^X \nabla f(X_t)] = \mathbf{E}[R_{t,a}J_t^{X_a}\nabla f(X_{t,a})],$$

and the proof is complete. □

Remark 3.5. Note that the statement of Lemma 3.4 generalises to functionals of the trajectory $X_{[0,t]} := (X_s, 0 \leq s \leq t)$. More precisely, for $t > 0$ and $F : \mathcal{C}^0([0,t], \mathbf{R}) \to \mathbf{R}$ a continuous map, Girsanov's theorem yields as well:

$$\mathbf{E}[F(X_{[0,t]})] = \mathbf{E}[R_{t,a}F(X_{[0,t],a})].$$

This extension is implicitly used in the previous proof, taking $F(X_{[0,t]}) = J_t^X \nabla f(X_t)$.

Remark 3.6. In dimension $d = 1$, since gradients and functions are both 1-dimensional objects, Theorem 3.3 rewrites in a more standard way:

$$(\mathbf{P}_t f)' = \mathbf{P}_t^{V''}(f') = \mathbf{E}\left[R_{t,a}f'(X_{t,a}) \exp\left(-\int_0^t V''(X_{s,a})\, ds\right)\right].$$

This writing shall be useful when dealing with monotonic functions in dimension 1, as briefly discussed at the end of the next section.

3.2 Logarithmic Sobolev inequalities

In this section, we provide a Feynman-Kac-based proof of the logarithmic Sobolev inequality, stated for a scalar perturbation. The method can easily be refined to improve the bound on $c_{LSI}(\mu)$, for example when finer informations on the spectrum of the the generator are available or for a restricted set of test functions. For instance, we adapt the proof to derive estimates in restriction to monotonic (positive) functions.

3.2.1 General case

Notation. The proof of the following theorem requires some matrix analysis. Henceforward, if A is a symmetric matrix, we let $\rho_-(A)$ denote its smallest eigenvalue. We may also use M^T and u^T to denote the usual transpose of a matrix M or a vector u.

The following result states the Bakry-Émery-like criterion mentioned in the introduction.

Theorem 3.7. *Let $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$. Define*

$$\kappa_a = \inf_{x \in \mathbf{R}^d}\left\{2\rho_-(\nabla^2 V(x)) - a\mathbf{L}(a^{-1})(x)\right\}.$$

If a, a^{-1} and $|\nabla a|$ are bounded and $\kappa_a > 0$, then for any $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$,

$$\mathrm{Ent}_\mu(f^2) \leq \frac{4\|a\|_\infty \|a^{-1}\|_\infty}{\kappa_a} \int_{\mathbf{R}^d} |\nabla f|^2\, d\mu.$$

Proof. Let $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ be a non-negative function. Ergodicity and μ-invariance of $(\mathbf{P}_t)_{t \geq 0}$ give:

$$\mathrm{Ent}_\mu(f) = -\int_{\mathbf{R}^d}\int_0^{+\infty} \partial_t\left(\mathbf{P}_t f \log \mathbf{P}_t f\right) dt\, d\mu = -\int_{\mathbf{R}^d}\int_0^{+\infty} \mathbf{L}[\mathbf{P}_t f]\log \mathbf{P}_t f\, dt\, d\mu.$$

The integration by parts formula and the intertwining relation yield then:

$$\mathrm{Ent}_\mu(f) = \int_{\mathbf{R}^d}\int_0^{+\infty} \frac{|\nabla \mathbf{P}_t f|^2}{\mathbf{P}_t f}\, d\mu\, dt = \int_{\mathbf{R}^d}\int_0^{+\infty} \frac{\left|\mathcal{P}_t^{\nabla^2 V}(\nabla f)\right|^2}{\mathbf{P}_t f}\, dt\, d\mu.$$

We focus on the numerator of the right-hand side. More precisely, we aim to cancel out $\mathbf{P}_t f$ at the denominator, which is made possible by Girsanov's theorem. Indeed, the assumptions on a ensure that it satisfies the (G) condition, and Theorem 3.3 leads to

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f) = \mathbf{E}[R_{t,a} J_t^{X_a} \nabla f(X_{t,a})],$$

which rewrites

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f) = 2\mathbf{E}\left[R_{t,a}^{1/2} J_t^{X_a} \nabla \sqrt{f}(X_{t,a}) R_{t,a}^{1/2} \sqrt{f}(X_{t,a})\right].$$

Cauchy-Schwarz' inequality with Lemma 3.4 finally entail

$$
\begin{aligned}
\left|\mathcal{P}_t^{\nabla^2 V}(\nabla f)\right|^2 &\leq 4\mathbf{E}\left[\left|R_{t,a}^{1/2} J_t^{X_a} \nabla \sqrt{f}(X_{t,a})\right|^2\right] \mathbf{E}\left[R_{t,a} f(X_{t,a})\right] \\
&= 4\mathbf{E}\left[\nabla \sqrt{f}(X_{t,a})^T J_t^{X_a} R_{t,a} (J_t^{X_a})^T \nabla \sqrt{f}(X_{t,a})\right] \mathbf{P}_t f.
\end{aligned}
$$

This implies then for the entropy:

$$\mathrm{Ent}_\mu(f) \leq 4 \int_{\mathbf{R}^d} \int_0^{+\infty} \mathbf{E}\left[\nabla \sqrt{f}(X_{t,a})^T J_t^{X_a} R_{t,a} (J_t^{X_a})^T \nabla \sqrt{f}(X_{t,a})\right] \, dt \, d\mu.$$

In order to recover the energy term in the LSI, one should provide some spectral estimates for $J_t^{X_a} R_{t,a} (J_t^{X_a})^T$. Define then

$$J_t^a = J_t^{X_a} \exp\left(-\frac{1}{2} \int_0^t \frac{\mathbf{L}_a a}{a}(X_{s,a}) \, ds\right),$$

which solves the following equation:

$$dJ_t^a = -J_t^a \left(\nabla^2 V(X_{t,a}) - \frac{1}{2} a \mathbf{L}(a^{-1})(X_{t,a}) I_d\right) dt.$$

Indeed, we have on the one hand:

$$dJ_t^{X_a} = -J_t^{X_a} \nabla^2 V(X_{t,a}) \, dt,$$

and on the other hand:

$$d\left[\exp\left(-\frac{1}{2} \int_0^t \frac{\mathbf{L}_a a}{a}(X_{s,a}) \, ds\right)\right] = -\frac{1}{2} \frac{\mathbf{L}_a a}{a}(X_{t,a}) \exp\left(-\frac{1}{2} \int_0^t \frac{\mathbf{L}_a a}{a}(X_{s,a}) \, ds\right) dt.$$

Moreover, $\mathbf{L}_a(a)/a = -a\mathbf{L}(a^{-1})$, so that both previous points and a chain rule give the expected formula. Since $J_t^{X_a} R_{t,a} (J_t^{X_a})^T = \frac{a(X_{t,a})}{a(x)} J_t^a (J_t^a)^T$, one should focus on spectral estimates for the latter term.

Therefore, if we let $\varphi(t) = y^T J_t^a (J_t^a)^T y$, for some $y \in \mathbf{R}^d$, symmetry of $\nabla^2 V$ entails

$$
\begin{aligned}
d\varphi(t) &= y^T dJ_t^a (J_t^a)^T y + y^T J_t^a (dJ_t^a)^T y \\
&= -y^T J_t^a \left(\nabla^2 V(X_{t,a}) - \frac{1}{2} a\mathbf{L}(a^{-1})(X_{t,a}) I_d\right) (J_t^a)^T y \, dt \\
&\quad - y^T J_t^a \left(\nabla^2 V(X_{t,a}) - \frac{1}{2} a\mathbf{L}(a^{-1})(X_{t,a}) I_d\right)^T (J_t^a)^T y \, dt \\
&= -y^T J_t^a \left(2\nabla^2 V(X_{t,a}) - a\mathbf{L}(a^{-1})(X_{t,a}) I_d\right) (J_t^a)^T y \, dt \\
&\leq -\kappa_a y^T J_t^a (J_t^a)^T y \, dt = -\kappa_a \varphi(t) \, dt,
\end{aligned}
$$

by definition of κ_a. Hence, $\varphi(t) \leq e^{-\kappa_a t}\varphi(0)$ for any $t \geq 0$, which yields

$$y^T J_t^a (J_t^a)^T y \leq e^{-\kappa_a t}|y|^2.$$

We can apply the previous inequality to $y = \sqrt{\frac{a(X_{t,a})}{a(x)}} \nabla \sqrt{f}(X_{t,a})$ to get

$$\mathrm{Ent}_\mu(f) \leq 4 \int_{\mathbf{R}^d} \int_0^{+\infty} e^{-\kappa_a t} \mathbf{E}\left[\frac{a(X_{t,a})}{a(x)} |\nabla \sqrt{f}(X_{t,a})|^2\right] \, dt \, d\mu,$$

which rewrites

$$\mathrm{Ent}_\mu(f) \le 4 \int_0^{+\infty} e^{-\kappa_a t} \int_{\mathbf{R}^d} \frac{1}{a} \mathbf{P}_{t,a} \left(a|\nabla\sqrt{f}|^2 \right) \, d\mu \, dt.$$

Recall that $d\mu_a/d\mu = 1/a^2$. Then, since a is bounded,

$$\mathrm{Ent}_\mu(f) \le 4\|a\|_\infty \int_0^{+\infty} e^{-\kappa_a t} \int_{\mathbf{R}^d} \mathbf{P}_{t,a} \left(a|\nabla\sqrt{f}|^2 \right) \, d\mu_a \, dt.$$

One can use invariance of $\mathbf{P}_{t,a}$ with respect to μ_a, then assumption on κ_a to get

$$\mathrm{Ent}_\mu(f) \le \frac{4\|a\|_\infty}{\kappa_a} \int_{\mathbf{R}^d} a|\nabla\sqrt{f}|^2 \, d\mu_a.$$

Finally, boundedness of a^{-1} entails

$$\mathrm{Ent}_\mu(f) \le \frac{4\|a\|_\infty \|a^{-1}\|_\infty}{\kappa_a} \int_{\mathbf{R}^d} |\nabla\sqrt{f}|^2 \, d\mu,$$

and the proof is complete replacing f by f^2. □

Remark 3.8. In terms of perturbation matrices (as presented in [2] through weighted intertwinings) one has here $A = aI_d$. To take into account the geometry of $\nabla^2 V$, a natural extension to this result would be to consider non-homothetic perturbations, for instance of the form $A = \mathrm{diag}(a_1, \ldots, a_d)$, where $a_1, \ldots, a_d \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ are distinct functions. In spite of many attempts, the above proof does not transpose to this case, and more general spectral estimates are besides much harder to derive. Generalisation of the representation theorem and Grönwall-like estimates for such perturbations would then allow an interesting extension to this result.

Remark 3.9 (Holley-Stroock criterion)**.** One may wish to compare this technique to the well-known Holley-Stroock method (introduced in [19] for the Ising model). As a reminder, if ν is a probability measure that satisfies a LSI and there exists $\Phi : \mathbf{R}^d \to \mathbf{R}$ a bounded continuous function such that $d\mu \propto e^\Phi d\nu$, then μ satisfies a LSI and

$$c_{LSI}(\mu) \le e^{2osc(\Phi)} c_{LSI}(\nu),$$

where $osc(\Phi) = \sup(\Phi) - \inf(\Phi)$. Note that $osc(\Phi)$ can poorly depend on the dimension, for example if $\Phi(x) = \sum_{i=1}^d \varphi(x_i)$, in which case $osc(\Phi) = d \cdot osc(\varphi)$. To stick to our framework, one might choose $\Phi = \log(a^2)$ for some bounded perturbation function $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$. The above inequality becomes

$$c_{LSI}(\mu) \le \|a\|_\infty^4 \|a^{-1}\|_\infty^4 c_{LSI}(\mu_a),$$

so that Holley-Stroock method leads to show that μ_a satisfies a LSI. This is conveniently ensured as soon as μ_a satisfies the Bakry-Émery criterion, namely

$$\inf_{x\in\mathbf{R}^d} \{\rho_-(\nabla^2 V_a(x))\} > 0.$$

In terms of a and V, the above condition rewrites explicitly:

$$\inf_{\mathbf{R}^d} \left\{ \rho_- \left(\nabla^2 V + \frac{2}{a}\nabla^2 a - \frac{2}{a^2}\nabla a(\nabla a)^T \right) \right\} > 0,$$

which shall be compared to the spectral estimates involved in κ_a, that can be expressed as:

$$\inf_{\mathbf{R}^d} \left\{ \rho_-(\nabla^2 V) + \frac{\Delta a}{a} - \nabla V \cdot \nabla a - \frac{2}{a^2}|\nabla a|^2 \right\} > 0.$$

Both expressions do not compare to each other, yet the second one seems to be far more tractable, as it could be illustrated on various examples.

As mentioned before, the above proof can be adapted in some particular cases to improve the estimate on $c_{LSI}(\mu)$. In the following, we thus study the restriction of the latter to monotonic (positive) functions.

3.2.2 Monotonic functions

Definition 3.10. *A measurable function $f : \mathbf{R}^d \to \mathbf{R}$ is said to be* monotonic *(in each direction) if for any $i = 1, \ldots, d$, for any fixed $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) \in \mathbf{R}^{d-1}$, $f_i : x_i \mapsto f(x_1, \ldots, x_d)$ is monotonic.*

In particular, if f is differentiable, then f is monotonic if and only if $\partial_i f$ has a constant sign on $\mathbf{R}^d$ for any i.

Remark 3.11. In the following, we shall focus on smooth functions f such that all f_i are non-decreasing (resp. non-increasing). In such cases, f will be said to be itself non-decreasing (resp. non-increasing).

Definition 3.12 ((BM) condition)**.** *Given the potential V, a function $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ is said to satisfy the* Bakry-Michel condition *(in short (BM)) if:*

1. *for any $i, j \in [\![1, d]\!]$, $i \neq j$, $\partial_{ij}^2 V_a \leq 0$;*
2. *for any $i \in [\![1, d]\!]$, $\sum_{j=1}^d \partial_{ij}^2 V_a$ is upper bounded,*

The following proposition is one of the main arguments that allows to improve the estimate on $c_{LSI}(\mu)$ for monotonic functions.

Proposition 3.13. *Let $f \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ and $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ satisfying (BM). Assume furthermore that f and a are both non-decreasing. Then*

$$\mathbf{P}_{t,a} f \leq \mathbf{P}_t f, \quad t \geq 0.$$

This proposition is based on a lemma provided by Bakry and Michel in [7], used initially to infer some FKG inequalities in $\mathbf{R}^d$.

Lemma 3.14. *Let $M : \mathbf{R}^d \to \mathcal{M}_d(\mathbf{R})$ be a measurable map such that $M_{ij} \leq 0$ for any $i \neq j$ and $\sum_{j=1}^d M_{ij}$ is upper bounded for any i, and let F be a smooth vector field on $\mathbf{R}^d$. Then all components of $\mathcal{P}_t^M F$ are non-negative whenever all components of F are so.*

We refer the reader to [7] for the proof. We can now provide a proof of Proposition 3.13.

Proof. The proof relies on very classical techniques. Let $t \geq 0$ be fixed and take $f \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ a non-decreasing function. Define, for any $s \in [0, t]$,

$$\Psi(s) = \mathbf{P}_s(\mathbf{P}_{t-s,a} f).$$

Since $\Psi(0) = \mathbf{P}_{t,a} f$ and $\Psi(t) = \mathbf{P}_t f$, we aim to prove that Ψ is non-decreasing. One has, for any $s \in [0, t]$,

$$\Psi'(s) = \mathbf{P}_s[(\mathbf{L} - \mathbf{L}_a)\mathbf{P}_{t-s,a} f],$$

which rewrites accordingly

$$\Psi'(s) = \mathbf{P}_s \left[\frac{\nabla a}{a} \cdot \nabla \mathbf{P}_{t-s,a} f \right] = \mathbf{P}_s \left[\frac{\nabla a}{a} \cdot \mathcal{P}_{t-s,a}^{\nabla^2 V_a}(\nabla f) \right].$$

Since f is non-decreasing, all entries of ∇f are non-negative, and since a satisfies *(BM)*, Lemma 3.14 implies that all entries of $\mathcal{P}_{t-s,a}^{\nabla^2 V_a}(\nabla f)$ are non-negative. Moreover, a is positive and non decreasing, so that

$$\frac{\nabla a}{a} \cdot \mathcal{P}_{t-s,a}^{\nabla^2 V_a}(\nabla f) \geq 0.$$

Hence, since $\mathbf{P}_s$ preserves the positivity, $\Psi'(s) \geq 0$ and the proof is over. $\qquad\square$

Remark 3.15. In dimension $d = 1$, due to the particular form of the Feynman-Kac semigroup $(\mathcal{P}_{t,a}^{\nabla^2 V_a})_{t\geq 0}$, Proposition 3.13 still holds if one only assumes that a is positive and a and f are both non-decreasing.

Proposition 3.13 enables us to adapt the proof of Theorem 3.7 and improve the estimate on $c_{LSI}(\mu)$. Moreover, the proof allows to handle unbounded perturbation functions (as long as the (G) condition is satisfied).

Theorem 3.16. Let $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ be non-decreasing. Define

$$\tilde{\kappa}_a = \inf_{x \in \mathbf{R}^d} \left\{ \rho_-(\nabla^2 V(x)) - a\mathbf{L}(a^{-1})(x) \right\}.$$

If a satisfies (BM), (G) and $\tilde{\kappa}_a > 0$, then for any non-decreasing $f \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$,

$$\text{Ent}_\mu(f^2) \leq \frac{2}{\tilde{\kappa}_a} \int_{\mathbf{R}^d} |\nabla f|^2 \, d\mu.$$

Proof. Let $f \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ be non-decreasing. The beginning of the proof is very similar to the one of Theorem 3.7. Indeed, the entropy rewrites

$$\text{Ent}_\mu(f) = \int_{\mathbf{R}^d} \int_0^{+\infty} \frac{\left| \mathcal{P}_t^{\nabla^2 V}(\nabla f) \right|^2}{\mathbf{P}_t f} \, dt \, d\mu,$$

with the representation

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f) = 2\mathbf{E}\left[R_{t,a} J_t^{X_a} \nabla \sqrt{f}(X_{t,a}) \sqrt{f}(X_{t,a}) \right],$$

since a satisfies (G). Theorem 3.3 and Cauchy-Schwartz' inequality imply here

$$\left| \mathcal{P}_t^{\nabla^2 V}(\nabla f) \right|^2 \leq 4\mathbf{E}\left[R_{t,a}^2 |J_t^{X_a} \nabla \sqrt{f}(X_{t,a})|^2 \right] \overbrace{\mathbf{E}[f(X_{t,a})]}^{\mathbf{P}_{t,a}f}$$

$$\leq 4\mathbf{E}\left[R_{t,a}^2 |J_t^{X_a} \nabla \sqrt{f}(X_{t,a})|^2 \right] \mathbf{P}_t f,$$

using Proposition 3.13. Plugged into the entropy, this yields

$$\text{Ent}_\mu(f) \leq 4 \int_{\mathbf{R}^d} \int_0^{+\infty} \mathbf{E}\left[\nabla \sqrt{f}(X_{t,a})^T J_t^{X_a} R_{t,a}^2 (J_t^{X_a})^T \nabla \sqrt{f}(X_{t,a}) \right] \, dt \, d\mu.$$

Here we let

$$J_t^a = J_t^{X_a} \exp\left(-\int_0^t \frac{\mathbf{L}_a a}{a}(X_{s,a}) \, ds \right),$$

and the same reasoning as in the proof of Theorem 3.7 gives then

$$\text{Ent}_\mu(f) \leq 4 \int_{\mathbf{R}^d} \int_0^{+\infty} e^{-2\tilde{\kappa}_a t} \mathbf{E}\left[\frac{a(X_{t,a})^2}{a(x)^2} |\nabla \sqrt{f}(X_{t,a})|^2 \right] \, dt \, d\mu.$$

Hence, using μ_a-invariance of $(\mathbf{P}_{t,a})_{t\geq 0}$,

$$\text{Ent}_\mu(f) \leq 4 \int_0^{+\infty} e^{-2\tilde{\kappa}_a t} \int_{\mathbf{R}^d} \mathbf{P}_{t,a}\left(a^2 |\nabla \sqrt{f}|^2 \right) \, d\mu_a \, dt$$

$$= 4 \int_0^{+\infty} e^{-2\tilde{\kappa}_a t} \int_{\mathbf{R}^d} |\nabla \sqrt{f}|^2 \, d\mu \, dt = \frac{2}{\tilde{\kappa}_a} \int_{\mathbf{R}^d} |\nabla \sqrt{f}|^2 \, d\mu,$$

and the proof is achieved replacing f by f^2. $\qquad\square$

4 Examples

In this section, we illustrate the Feynman-Kac approach on some examples. Since the perturbation function we introduce is scalar-valued, the method will be particularly suitable for potentials whose Hessian matrix admits many symmetries, for instance radial potentials. The examples we focus on shall then pertain to this class of potentials, namely here Subbotin and double-well potentials. Let us mention that, using other techniques, similar results for compactly supported radial measures were recently derived by Cattiaux, Guillin and Wu in [15].

For the sake of concision, we restrain ourselves to the illustration of Theorem 3.7. We eventually briefly resume the comparison to Holley-Stroock method.

4.1 Subbotin potentials

The first example we focus on is the general Subbotin[1] distribution [27]. We take then $V(x) = |x|^\alpha/\alpha$ for $\alpha > 2$, to ensure that V is at least twice continuously differentiable, but Bakry-Émery criterion does not apply (see the following proof and remark thereafter).

Lemma 4.1. Let $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$. Then for any $x \in \mathbf{R}^d$,

$$\rho_-(2\nabla^2 V(x)) - a\mathbf{L}(a^{-1})(x) = 2|x|^{\alpha-2} - a\mathbf{L}(a^{-1})(x).$$

Proof. First, notice that for any fixed $x \in \mathbf{R}^d$,

$$\nabla^2 V(x) = (\alpha - 2)|x|^{\alpha-4}xx^T + |x|^{\alpha-2}I_d.$$

Hence, $T_x := 2\nabla^2 V(x) - a\mathbf{L}(a^{-1})(x)I_d$ (seen as an element of $\mathcal{L}(\mathbf{R}^d)$), can be written as the sum of a rank 1 operator (projection on $\mathbf{R}x$) and a full-rank operator (multiple of the identity). One can then write $\mathbf{R}^d = \mathbf{R}x \oplus (\mathbf{R}x)^\perp$. Let λ be a non-zero eigenvalue of T_x and y be an associated eigenvector. Then

- either $y \in \mathbf{R}x$, that is, $y = \beta x$ for some $\beta \in \mathbf{R}^*$, and one can write

$$\lambda y = T_x y = 2\beta(\alpha - 2)|x|^{\alpha-2}x + 2\beta|x|^{\alpha-2}x - \beta a\mathbf{L}(a^{-1})(x)x,$$

which leads to

$$\lambda = 2(\alpha - 1)|x|^{\alpha-2} - a\mathbf{L}(a^{-1})(x);$$

- either $y \in (\mathbf{R}x)^\perp$, in which case

$$\lambda y = T_x y = 2|x|^{\alpha-2}y - a\mathbf{L}(a^{-1})(x)y,$$

which entails

$$\lambda = 2|x|^{\alpha-2} - a\mathbf{L}(a^{-1})(x).$$

Hence for any $x \in \mathbf{R}^d$, since $\alpha > 2$,

$$\rho_-(2\nabla^2 V(x)) - a\mathbf{L}(a^{-1})(x) = \rho_-(T_x) = 2|x|^{\alpha-2} - a\mathbf{L}(a^{-1})(x). \qquad \square$$

In the following, we may focus on the $\alpha = 4$ (quadric) case. Indeed, computations turn out to be particularly difficult in full generality, as well as keeping track of dependency in both parameters d and α. Bakry-Émery criterion clearly does not apply to this particular potential, since $\rho_-(\nabla^2 V(x))$ vanishes at point $x = 0$.

---

[1]after Mikhail Fedorovich Subbotin, 1893–1966, Soviet mathematician

Theorem 4.2. *Let $\mu(dx) \propto \exp(-|x|^4/4)dx$. There exists a universal explicit constant $c > 0$ such that for any $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, one has*

$$\mathrm{Ent}_\mu(f^2) \leq c \int_{\mathbf{R}^d} |\nabla f|^2 d\mu.$$

In particular, c does not depend on the dimension.

From the proof, one infers that $c = 3e\sqrt{3}$ is a suitable constant, yet highly dependent on the way computations are handled.

Proof. The first concern about making use of Theorem 3.7 stands in the choice of the perturbation function a. In practice, a should correct a lack of convexity of V where it occurs (namely where $\nabla^2 V(x) \leq 0$, here at $x = 0$). One of the first choices turns out to be the function

$$a(x) = \exp\left(\frac{\varepsilon}{2} \arctan(|x|^2)\right), \quad x \in \mathbf{R}^d.$$

Indeed, the arctangent function behaves like the identity near zero (where lies the lack of convexity of V) and like a constant at infinity (ensuring that a is bounded above and below). Furthermore, the square function is uniformly convex on $\mathbf{R}^d$, so that the Hessian matrix of the above is positive definite near the origin. Finally, taking exponential, a is indeed positive and computations are easier. Note that this choice is motivated by some results on the spectral gap, in which case the choice of a perturbation function that is close to non-integrability can provide relevant estimates on the Poincaré constant (see for example [10, 2]).

The next step in the method consists in the explicit computation of κ_a. With this definition of a, one has

$$-a\mathbf{L}(a^{-1})(x) = \varepsilon \frac{d + |x|^4(d-4)}{(1 + |x|^4)^2} - \varepsilon \frac{|x|^4}{1 + |x|^4} - \varepsilon^2 \frac{|x|^2}{(1 + |x|^4)^2}, \quad x \in \mathbf{R}^d,$$

and shall then minimize in $x \in \mathbf{R}^d$:

$$2|x|^2 + \varepsilon \frac{d + |x|^4(d-4)}{(1 + |x|^4)^2} - \varepsilon \frac{|x|^4}{1 + |x|^4} - \varepsilon^2 \frac{|x|^2}{(1 + |x|^4)^2},$$

which rewrites, setting $t = |x|^2$,

$$\kappa_a = \inf_{t \geq 0} \left(2t + \varepsilon \frac{d + t^2(d-4)}{(1 + t^2)^2} - \varepsilon \frac{t^2}{1 + t^2} - \varepsilon^2 \frac{t}{(1 + t^2)^2}\right).$$

Optimization of polynomials is hardly explicit in most cases, especially when one must keep track of all parameters (namely ε and d). We shall then focus here on the case where the infimum is reached for $t = 0$, that is, for any $t \geq 0$,

$$2t^4 - \varepsilon(d+1)t^3 + 4t^2 - \varepsilon(d+5)t + 2 - \varepsilon^2 \geq 0.$$

Let us denote by g the above polynomial function. Clearly, $\varepsilon \leq \sqrt{2}$ is a necessary, yet not sufficient condition for g to be non negative. In order to make computations more tractable, let us assume that g'' is positive. This is true as soon as

$$\varepsilon < \frac{8}{\sqrt{3}(d+1)}.$$

Consider then $\varepsilon \leq \dfrac{8}{3\sqrt{3}(d+1)}$. With this choice of ε, given that $d \geq 1$, one has for any $t \geq 0$

$$g(t) \geq 2t^4 - \frac{8t^3}{3\sqrt{3}} + 4t^2 - \frac{8t}{\sqrt{3}} + 2 - \frac{16}{27}.$$

It is easy to see that the above right-hand side is non-negative, so that g is non-negative either over $\mathbf{R}_+$. We can then take $\kappa_a = \varepsilon d$, and Theorem 3.7 entails the following estimate:

$$c_{LSI}(\mu) \leq \frac{4e^{\varepsilon\pi/4}}{\varepsilon d},$$

with $\varepsilon \leq \dfrac{8}{3\sqrt{3}(d+1)}$ (which implies that $\varepsilon \leq \sqrt{2}$). We finally minimize this bound with respect to $\varepsilon \in \left(0, \frac{8}{3\sqrt{3}(d+1)}\right]$ to get

$$c_{LSI}(\mu) \leq \frac{3\sqrt{3}(d+1)}{2d}e^{2\pi/3\sqrt{3}(d+1)}.$$

The above is uniformly bounded with respect to $d \in \mathbf{N}^\star$, and one can take $c = 3e\sqrt{3}$ as the universal constant mentioned in the theorem. $\qquad\square$

Remark 4.3. This proof points out the main concerns about Theorem 3.7. Indeed, the choice of the function (or family of functions) a is a key point. Nevertheless, the most important, yet technical, part of the proof is the explicit computation of κ_a, given that track should be kept of all parameters.

Note that, up to some numerical constant, the bound on ε in the previous proof is optimal (with this optimization method). Recall that the problem reduces to the prove that the function g defined on $\mathbf{R}_+$ as

$$g(t) = 2t^4 - \varepsilon(d+1)t^3 + 4t^2 - \varepsilon(d+5)t + 2 - \varepsilon^2, \quad t \geq 0,$$

is non-negative. If we assume that ε is of order $(d+1)^{-r}$ for some $r \in (0,1)$, then when d is large, for any fixed positive t,

$$g(t) \sim 2t^4 - d^{1-r}t^3 + 4t^2 - d^{1-r}t + 2 - d^{-2r},$$

and taking $t = 3/d^{1-r}$ leads to

$$g(3/(d+1)^r) \sim \frac{162}{d^{4(1-r)}} + \frac{9}{d^{2(1-r)}} - \frac{1}{2^{2r}} - 1 < 0$$

when d increases, which prevents the infimum of $t \mapsto \varepsilon d + tg(t)$ to be reached at $t = 0$.

We do not know if the constant we inferred is optimal (in terms of the dimension). Yet, one can note that, for example from [12], since the spectral gap for the quadric Subbotin distribution is of order $\sqrt{d}$, it is reasonable to expect $c_{LSI}(\mu)$ to be of order $1/\sqrt{d}$ (since μ satisfies a Poincaré inequality with constant c (which is the inverse of the spectral gap) as soon as it satisfies a LSI with constant $2c$, see [5] §5.1.2). It is then unclear that we can reach optimality with this very optimization procedure. More reliable optimization techniques would be then a good improvement regarding explicit estimates using this result.

Remark 4.4. The Holley-Stroock method as developed in Remark 3.9 leads, in the present case and after tedious computations, to a conclusion somewhat comparable to ours. Nevertheless, the involved constants are not fully explicit and leave less room for improvement than our above approach.

4.2 Double-well potentials

The following example is a perturbation of the previous one, known as the double-well potential. Consider $V(x) = |x|^4/4 - \beta|x|^2/2$, where $\beta > 0$ controls the size of the concave region. Although V is convex at infinity, its Hessian matrix is negative definite near the

origin, and Bakry-Émery criterion does not apply. Still, one can expect to recover the behaviour inferred in Theorem 4.2 when β is small.

Similarly to the Subbotin case, one can explicitly compute the Hessian matrix of V to get the following lemma.

Lemma 4.5. *Let $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$. Then for any $x \in \mathbf{R}^d$,*

$$\rho_-\left(2\nabla^2 V(x) - a\mathbf{L}(a^{-1})(x)I_d\right) = 2|x|^2 - 2\beta - a\mathbf{L}(a^{-1})(x).$$

Proof. The proof is identical to the one of Lemma 4.1. $\qquad\square$

Theorem 4.6. *Let $\mu(dx) \propto \exp(-|x|^4/4 + \beta|x|^2/2)dx$, $\beta \in (0, 1/2)$. There exists $c_\beta > 0$ a universal constant such that, for any function $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, one has*

$$\mathrm{Ent}_\mu(f^2) \le c_\beta \int_{\mathbf{R}^d} |\nabla f|^2 d\mu.$$

Again, c_β does not depend on the dimension.

A suitable constant is here $c_\beta = \frac{4e}{1-2\beta}$, the blow-up when $\beta \to 1/2$ is a computation artefact and has, to our knowledge, no qualitative significance.

Proof. This proof is very similar to the previous one. In particular, we set for any $x \in \mathbf{R}^d$

$$a(x) = \exp\left(\frac{\varepsilon}{2}\arctan(|x|^2)\right),$$

so that, for $t = |x|^2$,

$$\kappa_a = \inf_{t \ge 0}\left(2t - 2\beta + \varepsilon\frac{d + t^2(d-4)}{(1+t^2)^2} - \varepsilon(t-\beta)\frac{t}{1+t^2} - \varepsilon^2\frac{t}{(1+t^2)^2}\right).$$

Again, we aim to show that this infimum is equal to $\varepsilon d - 2\beta$, reached for $t = 0$, which amounts to prove that, for any $t \ge 0$,

$$g(t) := 2t^4 - \varepsilon(d+1)t^3 + (4+\beta)t^2 - \varepsilon(d+5)t + 2 - \varepsilon^2 + \beta \ge 0,$$

along with, to ensure positivity of κ_a, $\varepsilon > 2\beta/d$.

The first necessary condition that arises is $\varepsilon \le \sqrt{\beta+2}$. Moreover, in light of both previous proof and remark, ε should be of order $\frac{1}{d+1}$. To make computations easier, we take $\varepsilon = \frac{2}{d+1}$. Plugging this into both conditions $\varepsilon > 2\beta/d$ and $\varepsilon \le \sqrt{\beta+2}$ imply that β should not exceed $d/d+1$ for any d, which equates to $\beta < 1/2$. To summarize, we have

$$\varepsilon = \frac{2}{d+1} \quad \text{and} \quad 0 \le \beta < \frac{1}{2}.$$

Under those assumptions, g can be bounded from below as follows

$$g(t) \ge 2t^4 - 2t^3 + 4t^2 - 2t + 1 + \beta, \quad t \ge 0.$$

The right-hand term is positive on $\mathbf{R}_+$, so that with this choice of ε, one has

$$\kappa_a = \frac{2d}{d+1} - 2\beta.$$

This amounts, using Theorem 3.7, to

$$c_{LSI}(\mu) \le \frac{4(d+1)}{2d(1-\beta) - 2\beta}e^{\frac{\pi}{2(d+1)}}.$$

The above is uniformly bounded with respect to $d \in \mathbf{N}^\star$, and one can take $c_\beta = \dfrac{4e}{1-2\beta}$ as the aforementioned universal constant. $\qquad\square$

Remark 4.7. Note that the restriction on β is a computation artefact, and one has more $c_\beta \xrightarrow[\beta \to \frac{1}{2}^-]{} +\infty$. Nevertheless, the behaviour in term of the dimension is similar to what was derived for the Subbotin distribution in Theorem 4.2.

Remark 4.8. As for the quadric distribution, the Holley-Stroock method provides somewhat similar results, yet computations are far more tedious in this case, particularly in keeping track of the dependency in β.

References

[1] Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C. and Scheffer, G.: Sur les inégalités de Sobolev logarithmiques. *Panoramas et Synthèse*, Société Mathématique de France, 2000. MR-1845806

[2] Arnaudon, M., Bonnefont, M. and Joulin, A.: Intertwinings and Generalized Brascamp-Lieb Inequalities, *Rev. Mat. Iberoam.*, **34**, (2018), 1021–1054. MR-3850277

[3] Bakry, D.: Un critère de non-explosion pour certaines diffusions sur une variété riemannienne complète. *C. R. Acad. Sc. Paris*, **303**, (1986), 23–26. MR-0849620

[4] Bakry, D. and Émery, M.: Diffusions hypercontractives. *Séminaire de probabilités (Strasbourg)*, XIX, **1123**, (1985), 177–206. MR-0889476

[5] Bakry, D., Gentil, I. and Ledoux, M.: Analysis and Geometry of Markov Diffusion Operators. *Grundlehren der mathematischen Wissenschaften*, Springer, 2014. MR-3155209

[6] Bakry, D., Ledoux, M. and Wang, F.-Y.: Perturbations of functional inequalities using growth conditions. *J. Math. Pures Appl.*, **87**, (2007), 394–407. MR-2317340

[7] Bakry, D. and Michel, D.: Sur les inégalités FKG. *Séminaire de probabilités (Strasbourg)*, XXVI, **1526**, (1992), 170–188. MR-1231994

[8] Bismut, J.-M.: Large deviations and the Malliavin calculus. *Progress in Mathematics*, Birkhäuser, 1984. MR-0755001

[9] Bobkov, S. and Götze, F.: Exponential integrabiblity and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, **163**, (1999), 1–28. MR-1682772

[10] Bonnefont, M. and Joulin, A.: Intertwining Relations for One-Dimensional Diffusions and Application to Functional Inequalities. *Potential Anal.*, **41**, (2014), 1005–1031. MR-3269712

[11] Bonnefont, M. and Joulin, A.: Intertwinings, Second-Order Brascamp-Lieb Inequalities and Spectral Estimates. To appear in *Stud. Math.*, (2021).

[12] Bonnefont, M., Joulin, A. and Ma, Y.: Spectral gap for spherically symmetric log-concave probability measures, and beyond. *J. Funct. Anal.*, **270**, (2016), 2456–2482. MR-3464047

[13] Bourguignon, J.-P.: The "magic" of Weitzenböck formulas, in Variational Methods. *Progress in Nonlinear Differential Equations*, Birkhäuser, 1990. MR-1205158

[14] Braun M., Habermann, K. and Sturm, K.-T.: Optimal transport, gradient estimates, and pathwise Brownian coupling on spaces with variable Ricci bounds. *J. Math. Pures Appl.*, **147**, (2021), 60–97. MR-4213679

[15] Cattiaux, P., Guillin, A. and Wu, L.: Poincaré and logarithmic Sobolev inequalities for nearly radial measures. arXiv:1912.10825.

[16] Elworthy, K. D. and Li, X.-M.: Formulae for the derivatives of heat semigroups. *J. Funct. Anal.*, **125**, (1994), 252–286. MR-1297021

[17] Gross, L.: Logarithmic Sobolev Inequalities. *Am. J. Math*, **97**, (1975), 1061–1083. MR-0420249

[18] Helffer, B.: Semiclassical Analysis, Witten Laplacians, and Statistical Mechanics. *Progress in Mathematics*, Birkhäuser, 2002. MR-1936110

[19] Holley, R. and Stroock, D.: Logarithmic Sobolev Inequalities and Stochastic Ising Models. *J. Stat. Phys.*, **46**, (1987), 1159–1194. MR-0893137

[20] Ledoux, M.: Concentration of measure and logarithmic Sobolev inequalities. *Lecture Notes in Math.*, **1709**, (1999), 120–216. MR-1767995

[21] Ledoux, M.: The concentration of measure phenomenon. *Mathematical Surveys and Monographs*, AMS, 2001. MR-1849347

[22] Otto F. and Villani, C.: Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, **173**, (2000), 361–400. MR-1760620

[23] Perelman, G.: The entropy formula for the Ricci flow and its geometric applications. arXiv:math/0211159.

[24] Protter, P.E.: Stochastic Integration and Differential Equations. *Stochastic Modelling and Applied Probability*, Springer, 2004. MR-2273672

[25] Roberto, C. and Zegarlinski, B.: Orlicz-Sobolev inequalities for sub-Gaussian measures and ergodicity of Markov semi-groups. *J. Funct. Anal.*, **243**, (2007), 28–66. MR-2289793

[26] Stroock, D.W. and Varadhan, S.R.S.: Multidimensional Diffusion Processes. *Grundlehren der mathematischen Wissenschaften*, Springer, 2006. MR-0532498

[27] Subbotin, M.: On the Law of Frequency of Error. *Mat. Sb.*, **31**, (1923), 296–301.

[28] Wang, F.-Y.: Modified Curvatures on Manifolds with Boundary and Applications. *Potential Anal.*, **41**, (2014), 699–714. MR-3264816

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm