



Functional inequalities for perturbed measures with applications to log-concave measures and to some Bayesian problems

Patrick Cattiaux, Arnaud Guillin

► To cite this version:

Patrick Cattiaux, Arnaud Guillin. Functional inequalities for perturbed measures with applications to log-concave measures and to some Bayesian problems. *Bernoulli*, 2022, 28 (4), 10.3150/21-BEJ1419 . hal-03120626

HAL Id: hal-03120626

<https://uca.hal.science/hal-03120626>

Submitted on 26 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FUNCTIONAL INEQUALITIES FOR PERTURBED MEASURES WITH APPLICATIONS TO LOG-CONCAVE MEASURES AND TO SOME BAYESIAN PROBLEMS.

PATRICK CATTIAUX \spadesuit AND ARNAUD GUILLIN \diamond

\spadesuit UNIVERSITÉ DE TOULOUSE

\diamond UNIVERSITÉ CLERMONT-AUVERGNE

ABSTRACT. We study functional inequalities (Poincaré, Cheeger, log-Sobolev) for probability measures obtained as perturbations. Several explicit results for general measures as well as log-concave distributions are given. The initial goal of this work was to obtain explicit bounds on the constants in view of statistical applications for instance. These results are then applied to the Langevin Monte-Carlo method used in statistics in order to compute Bayesian estimators.

Key words : logconcave measure, Poincaré inequality, Cheeger inequality, logarithmic Sobolev inequality, perturbation, bayesian statistic, sparse learning

MSC 2010 : 26D10, 39B62, 47D07, 60G10, 60J60.

1. INTRODUCTION.

Let $\mu(dx) = Z_V^{-1} e^{-V(x)} dx$ be a probability measure defined on \mathbb{R}^n . We a priori do not require regularity for V and allow it to take values in $\mathbb{R} \cup \{-\infty, +\infty\}$. We denote by $\mu(f)$ the integral of f w.r.t. μ .

We define the Poincaré constant $C_P(\mu)$ as the best constant C satisfying

$$\mathrm{Var}_\mu(f) := \mu(f^2) - \mu^2(f) \leq C \mu(|\nabla f|^2), \quad (1.1)$$

for all smooth f , for instance all bounded f with bounded derivatives. Similarly the logarithmic-Sobolev constant $C_{LS}(\mu)$ is defined as the best constant such that, for all smooth f as before,

$$\mathrm{Ent}_\mu(f^2) := \mu(f^2 \ln(f^2)) - \mu(f^2) \ln(\mu(f^2)) \leq C \mu(|\nabla f|^2), \quad (1.2)$$

As it is well known, the Poincaré and the log-Sobolev constants are linked to the exponential stabilization of some markovian dynamics, like the Langevin diffusion i.e. the diffusion semi-group with generator

$$A = \Delta - \nabla V \cdot \nabla.$$

We shall give more explanations later. For simplicity we will say that μ satisfies a Poincaré or a log-Sobolev inequality provided $C_P(\mu)$ or $C_{LS}(\mu)$ are finite.

Date: January 26, 2021.

One can also introduce the \mathbb{L}^1 Poincaré constant $C_C(\mu)$ of μ is the best constant such that, for all smooth f ,

$$\mu(|f - \mu(f)|) \leq C_C(\mu) \mu(|\nabla f|).$$

Replacing $\mu(f)$ by $m_\mu(f)$ any μ median of f defines another constant, the Cheeger constant $C'_C(\mu)$ which satisfies $\frac{1}{2} C_C \leq C'_C \leq C_C$.

It is well known [31] that \mathbb{L}^1 and \mathbb{L}^2 Poincaré constants are related by the following

$$C_P(\mu) \leq 4 (C'_C)^2(\mu) \leq 4 C_C^2(\mu). \quad (1.3)$$

The Cheeger constant is connected to the isoperimetric profile of μ , see Ledoux [30].

The initial goal of this work is to study the transference of these inequalities to perturbed measures. Namely, let

$$\mu_F = Z_F^{-1} e^{-F} \mu \quad (1.4)$$

be a new probability measure. The question is: what can be said for the Poincaré or the log-Sobolev constant of μ_F in terms of the one of μ and the properties of F ? The question includes explicit controls, not only the finiteness of the related constants.

This question is of course not new and some results have been obtained for a long time. We shall only recall results with explicit controls on the constants. The most famous is certainly the following general result of Holley and Stroock (see for example [5])

Theorem 1.1. *If F is bounded, then*

$$C_P(\mu_F) \leq e^{Osc F} C_P(\mu), \quad \text{and} \quad C_{LS}(\mu_F) \leq e^{Osc F} C_{LS}(\mu).$$

Other results, where the constants are not easy to trace, have been obtained in [1] (also see [14] section 7). The result reads as follows: if μ satisfies a log-Sobolev inequality, and $e^{|\nabla F|^2}$ belongs to all the $\mathbb{L}^p(\mu)$ for $p < +\infty$, then μ_F also satisfies a log-Sobolev inequality. In particular the result holds true if F is Lipschitz. We shall revisit this result in subsection 2.2. For the Poincaré inequality, some general results have been obtained in [25] (see e.g. [14] proposition 4.4 for a simplified formulation).

Most of the other known results assume some convexity property.

We shall say that μ is *log-concave* if V is a convex function defined on some convex subset U . Since V can be infinite, this definition contains in particular the uniform measure on a convex body.

If V is strongly convex, i.e. for all $x \in \mathbb{R}^n$, $\langle u, Hess_V(x) u \rangle \geq \rho |u|^2$, where $\langle \cdot, \cdot \rangle$ denotes the euclidean scalar product and $Hess_V(x)$ the Hessian of V computed at point x , a consequence of Brascamp-Lieb inequality ([13]) is the inequality

$$C_P(\mu) \leq 1/\rho. \quad (1.5)$$

This relation was extended to more general situations and is often called the Bakry-Emery criterion or the curvature-dimension criterion $CD(\rho, \infty)$ (see [5] for a complete description of curvature-dimension criteria). Some improvements for variable curvature bounds are contained in [16]. Recall that the Bakry-Emery approach allows to show that in the strongly convex situation

$$C_{LS}(\mu) \leq 2/\rho. \quad (1.6)$$

Combining these results with the Holley-Stroock perturbation result shows that one can relax the strong convexity assumption in a bounded subset (i.e. assume strong convexity at infinity only).

That $C_P(\mu) < +\infty$ for general log-concave measures was first shown in 1999 by S. Bobkov in [11]. Another proof was given in [3] using Lyapunov functions, as introduced in [4].

Once one knows that $C_P(\mu)$ is finite, a particularly important problem is to get some explicit estimates. A celebrated conjecture due to Kannan, Lovász and Simonovits (KLS for short) is that there exists a universal constant C such that

$$\sigma^2(\mu) \leq C_P(\mu) \leq C \sigma^2(\mu) \quad (1.7)$$

where $\sigma^2(\mu)$ denotes the largest eigenvalue of the Covariance matrix $Cov_{i,j}(\mu) = Cov_\mu(x_i, x_j)$ and μ is log-concave. The left hand side is immediate. Since, a lot of works have been devoted to this conjecture, satisfied in some special cases. We refer to the book [2] for references before 2015, and to [19] for more information on the Poincaré constant of log-concave measures. Up to very recently the best general known result was

$$C_P(\mu) \leq C n^{\frac{1}{2}} \sigma^2(\mu) \quad (1.8)$$

as a consequence of the results by Lee and Vempala ([32]). It has been very recently announced (see [22] Theorem 1) a drastically better bound namely the existence of an universal constant C such that

$$C_P(\mu) \leq e^{C\sqrt{\ln(n) \ln(1+\ln(n))}} \sigma^2(\mu). \quad (1.9)$$

The dimension dependence of such results is what is important. Recall that for $n = 1$ one knows that $C \leq 12$ according to Bobkov's result (see [11] corollary 4.3).

In this framework, complementary perturbation results have been shown

Theorem 1.2. (1) (Miclo, see lemma 2.1 in Bardet et al [7]) *If $HessV \geq \rho Id$ for some $\rho > 0$ and F is L -Lipschitz then*

$$C_P(\mu_F) \leq \frac{2}{\rho} e^{4\sqrt{2n/\pi} \frac{L^2}{\rho^2}}.$$

(2) (see [17] example (3) section 7.1) *With the same assumptions as in (1),*

$$C_P(\mu_F) \leq \frac{1}{2} \left(\frac{2L}{\rho} + \sqrt{\frac{8}{\rho}} \right)^2 e^{L^2/2\rho}.$$

(3) (Barthe-Milman [9] Theorem 2.8) *If μ_F is log-concave and $\mu_F(e^{-F}) > K \mu(e^{-F}) \leq \frac{1}{8}$ then*

$$C_P(\mu_F) \leq C^2(1 + \ln K)^2 C_P(\mu).$$

Here C is a universal constant.

The final result (3) is the most general one obtained by transference in the log-concave situation. [9] contains a lot of other results in this direction, [19] contains alternative, simpler but worse results. Notice that (3) is wrongly recalled in [19] where a square is missing. The remarkable property of (3) or (2) is that the bound is dimension free (but actually dimension is hidden either in the Lipschitz constant or in the choice of K in most of the generic examples). One can find various other perturbation results in [6] relying on growth conditions, and usually stronger inequalities to get weaker ones.

In [19] we have studied several properties of the Poincaré constant for log-concave measures in particular the transference of these inequalities using absolute continuity, distance, mollification. This study was based on the fact that weak forms of the Poincaré inequality imply the usual form. A similar result was first stated by E. Milman ([34]) and the section 9.2 in [19] is devoted to extensions of E. Milman's results. Actually, in subsection 9.3.1 of [19] devoted to the transference via absolute continuity we missed the point. As explained in [9] the right way to obtain good results is to use the concentration properties of the initial measure (see the proof of Theorem 2.7 in [9]). The main (only) default in [34, 9] is that these papers obtain results up to universal constants that are difficult to trace. Thanks to the weak Poincaré inequalities used in [19], it is possible to obtain explicit bounds for these universal constants, sometimes up to a small loss. Why are we so interested in numerical bounds ?

The motivation of this note came from a statistical question. Indeed, log-concave distributions recently deserve attention in Statistics, see e.g. the survey [37]. Our starting point was a question asked to us by S. Gadat on the work [23] by Dalalyan and Tsybakov on sparse regression learning, we shall recall in more details in section 5. The question should be formulated as follows. Let

$$\mu_F = Z_F^{-1} e^{-F} \mu \quad (1.10)$$

be a new probability measure. Assume that μ is log-concave and that F is convex (log-concave perturbation of a log-concave measure). Is it possible to control $C_P(\mu_F)$ by $C_P(\mu)$ at least up to an universal multiplicative constant ? Since in a sense μ_F is “more” log-concave than μ , such a statement seems plausible, at least when the “arg-infimum” of F coincides with the one of μ .

A first partial answer was obtained by F. Barthe and B. Klartag in [8] Theorem 1:

Theorem 1.3. *For $n \geq 2$, choose $V(x) = \sum_{i=1}^n |x_i|^p$ for some $1 \leq p \leq 2$, and assume that F is an even convex function then*

$$C_P(\mu_F) \leq C (\ln(n))^{\frac{2-p}{p}} C_P(\mu),$$

where C is some universal constant.

Of course here $C_P(\mu)$ does not depend on the dimension since μ is a product measure and

$$C_P(\mu_1 \otimes \dots \otimes \mu_k) \leq \max_{j=1, \dots, k} C_P(\mu_j).$$

Unfortunately this result does not apply to sparse regression as in [23], where F is not even. We will thus first study perturbation for log-concave measures in section 3 and then see how it can be applied to the aforementioned statistical question. For the latter explicit numerical bounds are of key interest.

We will now describe the contents of the present paper.

In the next section we describe general perturbation results both for the Poincaré and the log-Sobolev constants. The naive method we are using does not seem to have been explored with the exception of some results for log-Sobolev contained in [1, 25, 15]. The results can be summarized as follows: for a not too big Lipschitz perturbation F , μ_F one may explicitly

compare $C_P(\mu_F)$ and $C_P(\mu)$, the same for the log-Sobolev constants. If F is C^2 one can replace the Lipschitz norm by a bound for $AF - \frac{1}{2}|\nabla F|^2$.

In subsection 2.3 we give a surprisingly simple application to mollified measures, extending part of the results in [7].

In section 3 we show how to improve these results when μ_F is log-concave. We first study how to get explicit controls in the results obtained by Barthe and Milman [9] using concentration, by plugging our explicit results of [19]. We then transpose the results of the first section. For short, if in general one needs to control the uniform norm of ∇F , in the log-concave case it is enough to control its \mathbb{L}^2 norm.

As an immediate consequence, we obtain in subsection 4.1, explicit controls for the pre-constants in results by Bobkov [11]. The idea of proof is to consider $\mu(dx) = e^{-V(x)}dx$ as a perturbation ν_F of $\nu(dx) = e^{-V(x) - \frac{1}{2}\rho|x|^2}dx$ that satisfies Bakry-Emery criterion. Actually the constants we are obtaining here are worse than the known ones, but the methodology will be used in the sequel.

In subsection 4.2 we replace the gaussian perturbation $|x|^2$ by $\sum_i |x_i|^p$ for $p > 2$. Using another result obtained in [8] for unconditional measures, we recover the control $C \ln^2(n)$ of the Poincaré constant of unconditional log-concave measures first obtained by Klartag ([28]) we already recovered in [19]. The advantage of our method is that it furnishes an explicit bound for the constant C .

The final section is devoted to the application to two statistical models: linear regression, following [23] and identification as in [24]. In the linear regression case we give explicit controls for the rate of convergence of the Langevin Monte-Carlo algorithm proposed in [23], that are much better than the ones suggested therein.

2. SMOOTH PERTURBATIONS IN GENERAL.

2.1. Poincaré inequalities.

We shall follow a direct perturbation approach, as the one of [15] section 4 used for the log-Sobolev constant. Since we shall use a similar but slightly different approach in the next sections we first isolate the starting point of the proof.

For a smooth f and a constant a we may write

$$\text{Var}_{\mu_F}(f) \leq \mu_F((f - a)^2) = \mu^{-1}(e^{-F}) \mu \left(((f - a)e^{-\frac{1}{2}F})^2 \right).$$

We choose

$$a = \frac{\mu \left(f e^{-\frac{1}{2}F} \right)}{\mu \left(e^{-\frac{1}{2}F} \right)}$$

so that the function $(f - a)e^{-\frac{1}{2}F}$ is μ centered. One can thus use the Poincaré inequality for μ in order to get, for all a and all $\varepsilon > 0$,

$$\mu_F((f - a)^2) \leq C_P(\mu) \int |\nabla f - \frac{1}{2}(f - a)\nabla F|^2 d\mu_F. \quad (2.1)$$

One deduces, for all $\varepsilon > 0$,

$$\mu_F((f - a)^2) \leq C_P(\mu) \left((1 + \varepsilon^{-1}) \mu_F(|\nabla f|^2) + \frac{1 + \varepsilon}{4} \mu_F((f - a)^2 |\nabla F|^2) \right). \quad (2.2)$$

We thus may state

Theorem 2.1. *If F is L -Lipschitz on the support of μ and if there exists $\epsilon > 0$ such that*

$$s := \frac{1}{4}(1 + \epsilon) C_P(\mu) L^2 < 1,$$

then

$$C_P(\mu_F) \leq \frac{(1 + \epsilon^{-1}) C_P(\mu)}{1 - s}.$$

Proof. Let a be as before. If F is L -Lipschitz we deduce from (2.2),

$$\mu_F((f - a)^2) \leq (1 + \varepsilon^{-1}) C_P(\mu) \mu_F(|\nabla f|^2) + \frac{1 + \varepsilon}{4} C_P(\mu) L^2 \mu_F((f - a)^2),$$

so that

$$\text{Var}_{\mu_F}(f) \leq \mu_F((f - a)^2) \leq \frac{(1 + \varepsilon^{-1}) C_P(\mu)}{1 - \frac{1}{4}(1 + \varepsilon) C_P(\mu) L^2} \mu_F(|\nabla f|^2)$$

as soon as $\exists \varepsilon > 0$ such that

$$\frac{1}{4}(1 + \varepsilon) C_P(\mu) L^2 < 1.$$

□

This theorem is quite sharp. Indeed recall [5, Prop. 4.4.2]: for every 1-Lipshitz function f and $s < \sqrt{4/C_P(\mu)}$, one has $\mu(e^{sf}) < \infty$. So in a sense our Lipschitz perturbation is nearly the largest one preserving that μ_F is well defined.

Similarly for the Cheeger constant we may state

Proposition 2.2. *If F is L -Lipschitz on the support of μ and $C_C(\mu) L < 1$,*

$$C'_C(\mu_F) \leq \frac{C_C(\mu)}{1 - C_C(\mu) L}.$$

Proof. First, for all a ,

$$\mu_F(|f - m_{\mu_F}(f)|) \leq \mu_F(|f - a|) = \mu^{-1}(e^{-F}) \mu(|(f - a)e^{-F}|)$$

so that choosing this time

$$a = \frac{\mu(f e^{-F})}{\mu(e^{-F})}$$

we get the result for the Cheeger constant by using

$$\begin{aligned} \mu_F(|f - a|) &\leq C_C(\mu) (\mu_F(|\nabla f|) + \mu_F(|f - a| |\nabla F|)) \\ &\leq C_C(\mu) \mu_F(|\nabla f|) + C_C(\mu) L \mu_F(|f - a|). \end{aligned}$$

□

The proof of these two statement is so simple that we cannot believe that the result is not known. In fact, the only comparable result we found, using seemingly more intricate techniques, is in [1, Th. 2.7] but with roughly a factor 16 in our favor and an explicit Poincaré constant.

What is remarkable is that we may even allow more general condition than F to be L -Lipschitz however assuming more regularity for F . It requires to be careful when μ is compactly supported.

Theorem 2.3. *a) If V is of C^1 class and F is of C^2 class and satisfies for some $\varepsilon > 0$*

$$C_P(\mu) \sup_x \left(AF - \frac{1}{2} |\nabla F|^2 \right)_+ (x) \leq 2(1 - \varepsilon),$$

where $u_+ = \max(u, 0)$, then

$$C_P(\mu_F) \leq \frac{C_P(\mu)}{\varepsilon}.$$

b) Assume that $U := \{V < +\infty\}$ is an open subset with a smooth boundary ∂U and that V is of C^1 class in U . Let F be of C^2 class and such that $\partial_n F \geq 0$ on ∂U where ∂_n denotes the normal derivative pointing outward. If F satisfies for some $\varepsilon > 0$

$$C_P(\mu) \sup_{x \in U} \left(AF - \frac{1}{2} |\nabla F|^2 \right)_+ (x) \leq 2(1 - \varepsilon),$$

where $u_+ = \max(u, 0)$, then

$$C_P(\mu_F) \leq \frac{C_P(\mu)}{\varepsilon}.$$

Proof. a) If one allows F to be more regular, one can replace (2.2) by another inequality. Indeed starting with (2.1) we have

$$\begin{aligned} \mu_F((f - a)^2) &\leq C_P(\mu) \int |\nabla f - \frac{1}{2}(f - a)\nabla F|^2 d\mu_F \\ &\leq C_P(\mu) \left(\mu_F(|\nabla f|^2) - \frac{1}{2} \mu_F(\nabla(f - a)^2 \cdot \nabla F) + \frac{1}{4} \mu_F((f - a)^2 |\nabla F|^2) \right) \\ &\leq C_P(\mu) \mu_F(|\nabla f|^2) + \frac{1}{2} C_P(\mu) \mu_F \left((f - a)^2 \left[A_F F + \frac{1}{2} |\nabla F|^2 \right] \right) \end{aligned} \quad (2.3)$$

where $A_F = A - \nabla F \cdot \nabla = \Delta - \nabla V \cdot \nabla - \nabla F \cdot \nabla$. Finally

$$\mu_F((f - a)^2) \leq C_P(\mu) \mu_F(|\nabla f|^2) + \frac{1}{2} C_P(\mu) \mu_F \left((f - a)^2 \left[A F - \frac{1}{2} |\nabla F|^2 \right] \right). \quad (2.4)$$

b) We have to start again with (2.1) where integration holds in U . This yields

$$\begin{aligned} \mu_F(f - a)^2 &\leq C_P(\mu) \int_U |\nabla f - \frac{1}{2}(f - a)\nabla F|^2 d\mu_F \\ &\leq C_P(\mu) \left(\mu_F(|\nabla f|^2) - \frac{1}{2} \mu_F(\langle \nabla(f - a)^2, \nabla F \rangle) + \frac{1}{4} \mu_F((f - a)^2 |\nabla F|^2) \right). \end{aligned} \quad (2.5)$$

To control the second term we have to use Green's formula to integrate by parts

$$\mu_F(\langle \nabla(f - a)^2, \nabla F \rangle) = -\mu_F((f - a)^2 A F) + \mu_F^\partial((f - a)^2 \partial_n F) \quad (2.6)$$

where $A = \Delta - \nabla V \cdot \nabla$, μ_F^∂ denotes the surface measure on ∂U and ∂_n denotes the normal derivative pointing outward. The end of the proof is then similar. \square

Example 2.4. Let us describe a very simple case which illustrates the difference between Th.2.1 and Th.2.3. Let $\mu = \frac{1}{2}e^{-|x|}d\mu$ for which $C_P(\mu) = 4$, and consider $F(x) = \rho|x|$ which is $|\rho|$ -Lipschitz. An application of Th.2.1 shows that μ_F still satisfies a Poincaré inequality if $|\rho| < 1$ whereas Th.2.1 implies that μ_F satisfies a Poincaré inequality as soon as $\rho > -1$ which is optimal in this case. \diamond

Example 2.5. For $\rho \in \mathbb{R}^+$ consider

$$\mu^\rho(dx) = Z_\rho^{-1} e^{-V(x) - \frac{\rho}{2}|x|^2} dx,$$

i.e.

$$F(x) = \frac{\rho}{2}|x|^2$$

which is not Lipschitzian. We thus have

$$(AF - \frac{1}{2}|\nabla F|^2)(x) = \rho \left(n - x \cdot \nabla V(x) - \frac{1}{2}\rho|x|^2 \right).$$

Hence if $x \cdot \nabla V \geq -K - K'|x|^2$, $AF - \frac{1}{2}|\nabla F|^2 \leq \rho(n + K) + (K' - \rho/2)|x|^2$ so that $C_P(\mu^\rho) \leq \frac{C_P(\mu)}{1-\varepsilon}$ as soon as $\rho \leq \frac{2(1-\varepsilon)}{C_P(\mu)(n+K)}$ and $\rho > K'$.

Thus a (very) small gaussian perturbation of a measure satisfying some Poincaré inequality is still satisfying some Poincaré inequality. Though natural, we do not know any other way to prove such a result. \diamond

This result can be compared with the one in [25] where μ is assumed to satisfy a log-Sobolev inequality. Actually, one can almost recover Gong-Wu result. Indeed, in (2.2) and (2.4), the final step requires to control a term in the form

$$\mu_F((f - a)^2 G).$$

According to the variational definition of relative entropy we have, for $\alpha > 0$,

$$\mu_F((f - a)^2 G) \leq \frac{1}{\alpha} \text{Ent}_{\mu_F}((f - a)^2) + \frac{1}{\alpha} \mu_F((f - a)^2) \ln \mu_F(e^{\alpha G}). \quad (2.7)$$

Replacing F by $F + \ln(\mu(e^{-F}))$ we may assume for simplicity that $\mu(e^{-F}) = 1$. Defining $g = e^{-F/2}(f - a)$ we have $\mu(g^2) = \mu_F((f - a)^2)$. Hence for all $\theta > 0$,

$$\begin{aligned} \text{Ent}_{\mu_F}((f - a)^2) &= \text{Ent}_\mu(g^2) + \mu(g^2 F) \\ &\leq C_{LS}(\mu) \mu(|\nabla g|^2) + \mu(g^2 F) \\ &\leq C_{LS}(\mu) (1 + \theta^{-1}) \mu_F(|\nabla f|^2) + C_{LS}(\mu) \frac{1 + \theta}{4} \mu_F((f - a)^2 |\nabla F|^2) \\ &\quad + \mu_F((f - a)^2 F), \end{aligned} \quad (2.8)$$

if we follow the proof of Theorem 2.1, or

$$\text{Ent}_{\mu_F}((f - a)^2) \leq C_{LS}(\mu) \mu_F(|\nabla f|^2) + C_{LS}(\mu) \mu_F \left((f - a)^2 [AF - \frac{1}{2}|\nabla F|^2] \right) + \mu_F((f - a)^2 F) \quad (2.9)$$

if we follow the proof of Theorem 2.3. Using again (2.7) we have obtained

(1) if for some positive s and t ,

$$\frac{1}{s} + \frac{1+\theta}{4t} C_{LS}(\mu) := D_1 \leq 1$$

then

$$\begin{aligned} \text{Ent}_{\mu_F}((f-a)^2) &\leq \frac{1}{1-D_1} \frac{C_{LS}(\mu)}{1+\theta} \mu_F(|\nabla f|^2) + \\ &+ \frac{1}{1-D_1} \mu_F((f-a)^2) \left[\frac{1}{s} \ln \mu_F(e^{sF}) + \frac{(1+\theta)C_{LS}(\mu)}{4t} \ln \mu_F(e^{t|\nabla F|^2}) \right]. \end{aligned}$$

(2) if for some positive s and t ,

$$\frac{1}{s} + \frac{1}{t} C_{LS}(\mu) := D_2 \leq 1$$

then

$$\begin{aligned} \text{Ent}_{\mu_F}((f-a)^2) &\leq \frac{1}{1-D_2} C_{LS}(\mu) \mu_F(|\nabla f|^2) + \\ &+ \frac{1}{1-D_2} \mu_F((f-a)^2) \left[\frac{1}{s} \ln \mu_F(e^{sF}) + \frac{C_{LS}(\mu)}{t} \ln \mu_F \left(e^{t[AF - \frac{1}{2}|\nabla F|^2]} \right) \right]. \end{aligned}$$

Finally we have

Proposition 2.6. *We suppose here that μ satisfies a logarithmic Sobolev inequality and thus that $C_{LS}(\mu)$ is finite.*

(i) *Assume that for some positive s and t ,*

$$\frac{1}{s} + \frac{1+\theta}{4t} C_{LS}(\mu) := D_1 \leq 1 \quad (2.10)$$

Assume in addition that there exist $\alpha > 0$, $\varepsilon > 0$ and $\theta > 0$ such that

$$T'_1 := \frac{(1+\varepsilon)C_P(\mu)}{4\alpha} T_1 < 1$$

where

$$T_1 := \ln \mu_F(e^{\alpha|\nabla F|^2}) + \frac{1}{1-D_1} \left[\frac{1}{s} \ln \mu_F(e^{sF}) + \ln(\mu(e^{-F})) + \frac{(1+\theta)C_{LS}(\mu)}{4t} \ln \mu_F(e^{t|\nabla F|^2}) \right].$$

Then,

$$C_P(\mu_F) \leq \frac{1}{1-T'_1} C_P(\mu) \left((1+\varepsilon^{-1}) + \frac{(1+\theta^{-1})(1+\varepsilon)}{4\alpha} C_{LS}(\mu) \right).$$

(ii) *Assume ome positive s and t ,*

$$\frac{1}{s} + \frac{1}{t} C_{LS}(\mu) := D_2 \leq 1 \quad (2.11)$$

Assume in addition that there exists $\alpha > 0$ such that

$$T'_2 := \frac{C_P(\mu)}{2\alpha} T_2 < 1$$

where

$$T_2 := \ln \mu_F(e^{\alpha[AF - \frac{1}{2}|\nabla F|^2]}) + \frac{1}{1-D_2} \left[\frac{1}{s} \ln \mu_F(e^{sF}) + \ln(\mu(e^{-F})) + \frac{C_{LS}(\mu)}{t} \ln \mu_F \left(e^{t[AF - \frac{1}{2}|\nabla F|^2]} \right) \right].$$

Then

$$C_P(\mu_F) \leq \frac{1}{1 - T_2'} C_P(\mu) \left(1 + \frac{C_{LS}(\mu)}{\alpha} \right).$$

Of course such a result is difficult to apply, but the method will be useful to get a perturbation result for the log-Sobolev constant. Part of the result has been described in [15]. Note however that as μ satisfies a logarithmic Sobolev inequality then one has Gaussian integrability properties, so that at least for every $s < 1/C_{LS}(\mu)$

$$\int e^{s|x|^2} d\mu < \infty$$

and thus if for some positive a, b sufficiently small, one has $|F|, |\nabla F|^2 < a + b|x|^2$, T_1 is then finite and can be made explicit.

2.2. Log-Sobolev inequality.

We can now similarly look at the log-Sobolev constant

Theorem 2.7. *Assume that F is L -Lipschitz on the support of μ and that μ satisfies a log-Sobolev inequality with constant $C_{LS}(\mu)$. Also assume for simplicity that $\mu(e^{-F}) = 1$.*

(1) *If $\sup_{x \in \text{supp}(\mu)} F(x) = M$, then for all $\theta > 0$,*

$$C_{LS}(\mu_F) \leq (1 + \theta^{-1}) C_{LS}(\mu) + C_P(\mu_F) \left(\frac{1 + \theta}{4} L^2 C_{LS}(\mu) + M + 2 \right).$$

(2) *For all $\beta > 0$, for all $\theta > 0$,*

$$\begin{aligned} C_{LS}(\mu_F) &\leq \frac{(\beta + 1)(1 + \theta^{-1})}{\beta} C_{LS}(\mu) + C_P(\mu_F) (2 + \mu(F)) \\ &\quad + L^2 C_P(\mu_F) C_{LS}(\mu) \left(\frac{(1 + \theta)(1 + \beta)}{4\beta} + \frac{\beta^2}{2} \right). \end{aligned}$$

If in addition the condition in Theorem 2.1 is satisfied for some $\varepsilon > 0$ we may replace $C_P(\mu_F)$ by the bound obtained in Theorem 2.1.

Remark that the first stement does not enter the framework of Holley-Stroock's theorem as only a one sided bound is assumed on F .

Proof. Let f be smooth and such that $\mu_F(f^2) = 1$. Defining $g = e^{-F/2} f$ we have $\mu(g^2) = 1$. It follows for all $\theta > 0$,

$$\begin{aligned} \text{Ent}_{\mu_F}(f^2) &= \mu_F(f^2 \ln(f^2)) = \mu(g^2 \ln(g^2)) + \mu(g^2 F) \\ &\leq C_{LS}(\mu) \mu(|\nabla g|^2) + \mu(g^2 F) \\ &\leq C_{LS}(\mu) (1 + \theta^{-1}) \mu_F(|\nabla f|^2) + C_{LS}(\mu) \frac{1 + \theta}{4} \mu_F(f^2 |\nabla F|^2) + \mu_F(f^2 F). \end{aligned} \tag{2.12}$$

In the first case, we can bound the sum of the last two terms by

$$\left(C_{LS}(\mu) \frac{1 + \theta}{4} L^2 + M \right) \mu_F(f^2)$$

and apply the Poincaré inequality for μ_F provided $\mu_F(f) = 0$. To conclude it is then enough to recall Rothaus lemma (Lemma 5.14 in [5]),

$$\text{Ent}_\nu(f^2) \leq \text{Ent}_\nu((f - \nu(f))^2) + 2 \text{Var}_\nu(f).$$

If F is not bounded above we can use the variational definition of relative entropy as before:

$$\mu_F(f^2 F) \leq \frac{1}{\alpha} \text{Ent}_{\mu_F}(f^2) + \frac{1}{\alpha} \mu_F(f^2) \ln(\mu_F(e^{\alpha F})).$$

Gathering all the previous bounds we have obtained, provided $\alpha > 1$,

$$\text{Ent}_{\mu_F}(f^2) \leq \frac{\alpha}{\alpha - 1} (C_{LS}(\mu) (1 + \theta^{-1}) \mu_F(|\nabla f|^2) + C \mu_F(f^2)) \quad (2.13)$$

with

$$C = C_{LS}(\mu) \frac{1 + \theta}{4} L^2 + \frac{1}{\alpha} \ln(\mu_F(e^{\alpha F})).$$

We may then argue as before using Rothaus lemma again. The bound

$$\mu_F(e^{\alpha F}) = \mu(e^{(\alpha-1)F}) \leq e^{(\alpha-1)\mu(F) + (C_{LS}(\mu) L^2 (\alpha-1)^2 / 2)}$$

is known as the Herbst argument (see e.g. [5] Proposition 5.4.1). \square

As before we may replace the Lipschitz assumption by an integrability condition yielding the next result whose proof, similar to the previous one, is omitted

Theorem 2.8. *Suppose that μ satisfies a logarithmic Sobolev inequality with constant $C_{LS}(\mu)$ and that $\mu(e^{-F}) = 1$. Assume that there exist $\alpha > 1$ and $\beta, \theta > 0$ such that*

$$\mu_F(e^{\alpha F}) < \infty, \quad \mu_F(e^{\beta |\nabla F|^2}) < \infty$$

and

$$C_{LS}(\mu) \frac{1 + \theta}{4\beta} + \frac{1}{\alpha} := \delta < 1$$

then μ_F also satisfies a logarithmic Sobolev inequality with constant $C_{LS}(\mu_F)$ equal to

$$\frac{1}{1 - \delta} \left[C_{LS}(\mu) (1 + \theta^{-1}) + C_P(\mu_F) \left(2 + C_{LS}(\mu) \frac{1 + \theta}{4\beta} \log \mu_F(e^{\beta |\nabla F|^2}) + \frac{1}{\alpha} \log \mu_F(e^{\alpha F}) \right) \right].$$

The previous Theorem is a version of the one obtained in [1] as recalled in the introduction.

Finally, if F is more regular we may replace (2.12) by the following

$$\text{Ent}_{\mu_F}(f^2) \leq C_{LS}(\mu) \mu(|\nabla f|^2) + \frac{1}{2} C_{LS}(\mu) \mu_F \left(f^2 [AF - \frac{1}{2} |\nabla F|^2] \right) + \mu_F(f^2 F).$$

Arguing as before we thus obtain

Theorem 2.9. *Suppose that μ satisfies a logarithmic Sobolev inequality with constant $C_{LS}(\mu)$ and that V is C^1 . Assume that there exist $\alpha > 1$ and $\beta > 0$ such that*

$$\mu_F(e^{\alpha F}) < \infty, \quad \mu_F(e^{\beta [AF - \frac{1}{2} |\nabla F|^2]}) < \infty$$

and

$$C_{LS}(\mu) \frac{1}{2\beta} + \frac{1}{\alpha} := \delta < 1$$

then μ_F also satisfies a logarithmic Sobolev inequality with constant $C_{LS}(\mu_F)$ equal to

$$\frac{1}{1-\delta} \left[C_{LS}(\mu) + C_P(\mu_F) \left(2 + C_{LS}(\mu) \frac{1}{2\beta} \log \mu_F(e^{\beta[AF - \frac{1}{2}|\nabla F|^2]}) + \frac{1}{\alpha} \log \mu_F(e^{\alpha F}) \right) \right].$$

If in addition the condition in Theorem 2.3 is satisfied for some $\varepsilon > 0$ we may replace $C_P(\mu_F)$ by the bound obtained in Theorem 2.3.

Remark 2.10. Notice that if $AF - \frac{1}{2}|\nabla F|^2$ is non-positive at infinity (which is often the case in concrete examples), and F is C^2 , $e^{\beta[AF - \frac{1}{2}|\nabla F|^2]}$ is bounded for all $\beta > 0$, so that the condition in the previous theorem reduces to the integrability of $e^{\alpha F}$ for some $\alpha > 1$. The most stringent condition is thus the one in theorem 2.3 ensuring the finiteness of the Poincaré constant. \diamond

2.3. Application to mollified measures.

Let ν be a given probability measure (non necessarily absolutely continuous) and define ν^σ as the convolution $\nu^\sigma = \nu * \gamma_\sigma$ where γ_σ denotes the centered gaussian distribution with covariance matrix $\sigma^2 Id$. In other words ν^σ is the probability distribution of $X + \sigma G$ where X is a random variable with distribution ν and G is a standard gaussian variable. A natural question is to know when ν^σ satisfies a Poincaré or a log-Sobolev inequality and to get some controls on the corresponding constants. Notice that ν is not assumed to satisfy itself such an inequality.

When ν has compact support, included in the euclidean ball $B(0, R)$, this question has been partly studied in [39], and the results therein extended in [7]. [39] is using the Lyapunov function method of [20], while [7] is partly using the Bakry-Emery criterion. Indeed since for $\sigma > 0$, $\nu^\sigma(dx) = e^{-V^\sigma}(x) dx$ for some smooth V^σ , with

$$V^\sigma(x) = - \ln \left(\int e^{-\frac{|x-y|^2}{2\sigma^2}} (2\pi\sigma^2)^{-n/2} \nu(dy) \right).$$

A simple calculation (see [7] p.438) shows that

$$Hess V^\sigma \geq \left(\frac{1}{\sigma^2} - \frac{R^2}{\sigma^4} \right) Id, \quad (2.14)$$

so that $C_{LS}(\nu^\sigma) \leq \frac{2\sigma^4}{\sigma^2 - R^2}$ as soon as $\sigma > R$ (it seems that the factor 2 is lacking in [7]). The small variance case is more delicate and impose to use other arguments. Nevertheless it is not difficult using a variance decomposition to prove that the following is always true:

$$C_P(\nu^\sigma) \leq \sigma^2 e^{4R^2/\sigma^2}.$$

If V^σ is not necessarily strongly convex, the Hessian remains bounded from below. Using deep results by E. Milman ([35]) in the spirit of the ones we will recall in the next section, it is shown in Theorem 4.3 of [7] that ν^σ is still satisfying a log-Sobolev inequality that does not depend on the dimension n provided $\sigma > R/\sqrt{2}$, but this time the log-Sobolev constant is not explicit.

We will improve the latter result and furnish an explicit constant by directly using our perturbation results. To this end we simply write

$$\nu^\sigma(dx) = Z^{-1} e^{-F(x)} \gamma_\sigma(dx) \quad \text{with} \quad F(x) = V^\sigma(x) - \frac{|x|^2}{2\sigma^2}. \quad (2.15)$$

We have

$$\nabla F(x) = \int \frac{x-y}{\sigma^2} h(x,y) \nu(dy) - \frac{x}{\sigma^2} = - \int \frac{y}{\sigma^2} h(x,y) \nu(dy) \quad (2.16)$$

where

$$h(x,y) = \frac{e^{-\frac{|x-y|^2}{2\sigma^2}}}{\int e^{-\frac{|x-z|^2}{2\sigma^2}} \nu(dz)}.$$

Hence

$$|\nabla F(x)| \leq \frac{1}{\sigma^2} \int |y| h(x,y) \nu(dy) \leq \frac{R}{\sigma^2}. \quad (2.17)$$

It remains to apply Theorem 2.1, part (2) of Theorem 2.7 and the bounds $C_P(\gamma_\sigma) \leq \sigma^2$ and $C_{LS}(\gamma_\sigma) \leq 2\sigma^2$ in order to get

Theorem 2.11. *Let ν be any probability measure whose support is included in $B(0, R)$. Define $\nu^\sigma = \nu * \gamma_\sigma$ where γ_σ denotes the centered gaussian distribution with covariance matrix $\sigma^2 \text{Id}$. Then if*

$$s := \frac{1+\varepsilon}{4} \frac{R^2}{\sigma^2} < 1$$

it holds

$$C_P(\nu^\sigma) \leq \frac{1+\varepsilon^{-1}}{1-s} \sigma^2.$$

Similarly for all θ and β positive,

$$\begin{aligned} C_{LS}(\nu^\sigma) &\leq \left(\frac{2(\beta+1)(1+\theta^{-1})}{\beta} + 5 \frac{1+\varepsilon^{-1}}{1-s} \right) \sigma^2 \\ &\quad + 2 \frac{1+\varepsilon^{-1}}{1-s} \left(\frac{(1+\theta)(1+\beta)}{4\beta} + \frac{\beta^2}{2} \right) R^2. \end{aligned}$$

Notice that this result covers the range $\sigma > R/2$ which is larger than the one in [7]. Here we have used $\mu(F) \leq F(0) + L\mu(|x|)$ in order to simplify the (already intricate) bound for the log-Sobolev constant. Using the elementary general $C_P(\mu * \nu) \leq C_P(\mu) + C_P(\nu)$, one has $C_P(\nu^\sigma) \leq C_P(\nu^{\sigma_0}) + (\sigma - \sigma_0)^2$, yielding the correct asymptotic behaviour.

The previous proof can easily be extended to more general situations replacing γ_σ by some more general $\mu(dx) = e^{-H(x)} dx$, provided $\text{Hess}H$ is bounded, yielding

Theorem 2.12. *Let ν be any probability measure whose support is included in $B(0, R)$. Define $\nu^H = \nu * \mu$ where $\mu(dx) = e^{-H(x)} dx$ is a probability measure such that*

$$\sup_x |\text{Hess}H(x)| = K < +\infty.$$

Then if

$$s := \frac{1+\varepsilon}{4} K^2 R^2 C_P(\mu) < 1$$

it holds

$$C_P(\nu^H) \leq \frac{1 + \varepsilon^{-1}}{1 - s} C_P(\mu).$$

Proof. Following the notations of the previous proof we have

$$\nabla F(x) = \int \nabla H(x - y) h(x, y) \nu(dy) - \nabla H(x) = \int (\nabla H(x - y) - \nabla H(x)) h(x, y) \nu(dy)$$

with

$$h(x, y) = \frac{e^{-H(x-y)}}{\int e^{-H(x-z)} \nu(dz)}.$$

It remains to use

$$|\nabla H(x - y) - \nabla H(x)| \leq K |y|,$$

and to use Theorem 2.1. \square

Corollary 2.13. *Let X be a random variable supported by $B(0, R)$ and Y a random variable with distribution $\mu(dx) = e^{-H(x)} dx$ such that $\sup_x |\text{Hess} H(x)| = K < +\infty$. For $\sigma \in \mathbb{R}^+$ define $X^\sigma = X + \sigma Y$ and denote by $\nu^{\sigma H}$ the distribution of X^σ . Then if*

$$s := \left(\frac{1 + \varepsilon}{4} K^2 R^2 C_P(\mu) \right) / \sigma^2 < 1$$

it holds

$$C_P(\nu^H) \leq \frac{1 + \varepsilon^{-1}}{1 - s} C_P(\mu) \sigma^2.$$

Proof. It is enough to remark that the probability density of σZ is proportional to $e^{-H(x/\sigma)}$ so that $|\nabla F| \leq \frac{KR}{\sigma^2}$ and to remember that $C_P(\sigma Y) = \sigma^2 C_P(Y)$. \square

Remark 2.14. One can ask about what happens when ∇H is bounded, for instance if $\mu(dx) = Z^{-1} e^{-\sigma|x|} dx$ in \mathbb{R} . The proofs above furnish $|\nabla F| \leq 2/\sigma$ for all R and all σ so that the condition on s reads $s := (1 + \varepsilon) C_P(\mu) < 1$ which is impossible since $C_P(\mu) = 4$. This is another argument showing that our perturbation result is close to be optimal. \diamond

3. PERTURBATION WITH LOG-CONCAVITY.

We will now give some new results relating the Poincaré constant of both measures μ and μ_F when at least one of them is log-concave.

For log-concave distributions it is often better to use the Cheeger constant instead of the Poincaré constant. Recall the following

Proposition 3.1. *Recall that in all cases $C_P(\mu) \leq 4(C'_C)^2(\mu) \leq 4C_C^2(\mu)$. If in addition μ is log-concave the following converse inequality is satisfied :*

$$C'_C(\mu) \leq C_C(\mu) \leq \frac{16}{\pi} \sqrt{C_P(\mu)}.$$

The first inequality is contained in [11], while the second one is shown in [19] proposition 9.2.11. With the slightly worse constant 6 instead of $\frac{16}{\pi}$ the result is due to Ledoux in [30].

Finally a remarkable property of log-concave measures, we shall intensively use in the sequel, is that a very weak form of the Poincaré (or Cheeger) inequality is enough to imply the true one. For simplicity we recall here the two main results we obtained in [19] (Theorem 9.2.7 and Theorem 9.2.14), improving on the beautiful seminal result by E. Milman ([34])

Theorem 3.2. *Let ν be a log-concave probability measure.*

- (1) *Assume that there exists some $0 \leq s < 1/2$ and some $\beta(s)$ such that for any Lipschitz function f it holds*

$$\nu(|f - m_\nu(f)|) \leq \beta(s) \|\nabla f\|_\infty + s \text{Osc}(f).$$

Then

$$C'_C(\nu) \leq \frac{4\beta(s)}{\pi(\frac{1}{2} - s)^2}.$$

- (2) *Assume that there exists some $0 \leq s < 1/6$ and some $\beta(s)$ such that for any Lipschitz function f it holds*

$$\text{Var}_\nu(f) \leq \beta(s) \nu(|\nabla f|^2) + s \text{Osc}^2(f).$$

Then

$$C'_C(\nu) \leq \frac{4\sqrt{\beta(s) \ln 2}}{1 - 6s}.$$

In both cases recall that $C_P(\nu) \leq 4(C'_C(\nu))^2$.

3.1. From Holley-Stroock to Barthe-Milman.

We start by mimiking the proof of Holley-Stroock perturbation result. Let f be a bounded Lipschitz function. Of course in the definition of μ_F we may always replace e^{-F} by $e^{-(F - \min F)}$ provided F is bounded from below. Hence, for simplicity we may first assume that $F \geq 0$ so that $e^{-F} \leq 1$ is in all the $\mathbb{L}^p(\mu)$. Then :

$$\begin{aligned} \mu_F(|f - \mu_F(f)|) &\leq 2 \mu_F(|f - m_{\mu_F}(f)|) \leq 2 \mu_F(|f - \mu(f)|) \\ &\leq 2 \frac{\mu(|f - \mu(f)| e^{-F})}{\mu(e^{-F})} \\ &\leq \frac{2}{\mu(e^{-F})} \mu^{1/2}(|f - \mu(f)|^2) \mu^{1/2}(e^{-2F}) \\ &\leq \frac{2 \mu^{1/2}(e^{-2F})}{\mu(e^{-F})} C_P^{1/2}(\mu) \mu^{1/2}(|\nabla f|^2) \leq \frac{2 \mu^{1/2}(e^{-2F})}{\mu(e^{-F})} C_P^{1/2}(\mu) \|\nabla f\|_\infty. \end{aligned}$$

If F is not bounded from below, just using a cut-off we obtain the same result, with a possibly infinite right hand side. Using Theorem 3.2 (1), we can thus conclude

Proposition 3.3. *If μ_F is log-concave then*

$$C'_C(\mu_F) \leq \frac{32}{\pi} \frac{\mu^{1/2}(e^{-2F})}{\mu(e^{-F})} C_P^{1/2}(\mu)$$

so that

$$C_P(\mu_F) \leq \frac{4 \times 32^2}{\pi^2} \frac{\mu(e^{-2F})}{\mu^2(e^{-F})} C_P(\mu).$$

If interesting in comparison with Holley-Stroock the previous result is far from optimal. As shown in Theorem 2.7 of [9], there exists an universal constant c such that

$$C_P(\mu_F) \leq c \left(1 + \ln \left(\frac{\mu^{\frac{1}{2}}(e^{-2F})}{\mu(e^{-F})} \right) \right)^2 C_P(\mu).$$

We shall recover this result and furnish a numerical bound for c . To this end first recall

Definition 3.4. The concentration profile of a probability measure ν denoted by α_ν , is defined as

$$\alpha_\nu(r) := \sup \left\{ 1 - \nu(A + B(0, r)); \nu(A) \geq \frac{1}{2} \right\}, \quad r > 0,$$

where $B(y, r)$ denotes the euclidean ball centered at y with radius r .

The following is shown in [19] Corollary 9.2.10

Proposition 3.5. For any log-concave probability measure ν ,

$$C'_C(\nu) \leq \inf_{0 < s < \frac{1}{4}} \frac{16 \alpha_\nu^{-1}(s)}{\pi(1-4s)^2} \quad \text{and} \quad C_P(\nu) \leq \inf_{0 < s < \frac{1}{4}} \left(\frac{32 \alpha_\nu^{-1}(s)}{\pi(1-4s)^2} \right)^2.$$

Actually a better result, namely

$$C'_C(\nu) \leq \frac{\alpha_\nu^{-1}(s)}{1-2s}$$

that holds for all $s < \frac{1}{2}$ was shown by E. Milman in Theorem 2.1 of [36], when ν is the uniform measure on a convex body. The results extends presumably to any log-concave measure, but the proof of this result lies on deep geometric results (like the Heintze-Karcher theorem) while ours is elementary.

Now recall the statement of Proposition 2.2 in [9], in a simplified form: if $M = \frac{\mu^{\frac{1}{2}}(e^{-2F})}{\mu(e^{-F})}$, then

$$\alpha_{\mu_F} \leq 2M \alpha_\mu^{1/2}(r/2).$$

We may use this result together with proposition 3.5 to deduce corollary 9.3.2 in [19]

Corollary 3.6. If μ_F is log-concave, denoting $M = \frac{\mu^{\frac{1}{2}}(e^{-2F})}{\mu(e^{-F})}$,

$$C'_C(\mu_F) \leq \inf_{0 < s < \frac{1}{4}} \frac{32 \alpha_\mu^{-1}((s/2M)^2)}{\pi(1-4s)^2}.$$

Finally, as shown by Gromov and V. Milman, the concentration profile of a measure with a finite Poincaré constant is exponentially decaying. One more time it is not as easy to find an explicit version of Gromov-Milman's result. We found two of them in the literature: the first one in [38] Théorème 25 (in french) and Proposition 11

$$\alpha_\mu(r) \leq 16 e^{-\frac{r}{\sqrt{2} C_P(\mu)}},$$

the second one in [10] Theorem 2:

$$\alpha_\mu(r) \leq e^{-\frac{r}{3\sqrt{C_P(\mu)}}}.$$

We shall use the first one due to the lower constant and obtain

Theorem 3.7. *If μ_F is log-concave, denoting $M = \frac{\mu^{\frac{1}{2}}(e^{-2F})}{\mu(e^{-F})}$, for all $0 < s < \frac{1}{4}$,*

$$C_P(\mu_F) \leq \frac{C}{(1-4s)^4} (3\ln 2 + \ln(1/s) + \ln M)^2 C_P(\mu)$$

where the constant C satisfies $C \leq \left(\frac{64\sqrt{2}}{\pi}\right)^2$.

Remark 3.8. Of course $M \geq 1$. In order to get a presumably more tractable bound, first remark that adding a constant to F does not change M so that we may always assume that $\min F = 0$. It thus follows $\mu(e^{-2F}) \leq \mu(e^{-F})$. According to Jensen's inequality we also have $\mu(e^{-F}) \geq e^{-\mu(F)}$ so that finally

$$M \leq e^{\frac{1}{2}\mu(F)}.$$

Finally

$$C_P(\mu_F) \leq (C_1 + C_2 \mu^2(F)) C_P(\mu)$$

for some explicit constants C_1 and C_2 . ◇

Example 3.9. Gaussian perturbation.

For $\rho \in \mathbb{R}^+$ consider

$$\mu^\rho(dx) = Z_\rho^{-1} e^{-V(x) - \frac{1}{2}\rho|x|^2} dx,$$

i.e.

$$F(x) = \frac{\rho}{2} |x|^2.$$

If V is convex (hence μ^ρ log-concave), we deduce from the previous theorem and Bakry-Emery criterion that,

$$C_P(\mu^\rho) \leq \min \left(\frac{1}{\rho}; \left(C_1 + C_2 \ln^2 \left(\frac{\mu^{\frac{1}{2}}(e^{-\rho|x|^2})}{\mu(e^{-\rho|x|^2/2})} \right) \right) C_P(\mu) \right),$$

for some explicit universal constants C_1 and C_2 .

This indicates that we can find a bound for the Poincaré constant of μ^ρ that does not depend on ρ . We shall try to get some explicit result.

First according to remark 3.8, the \ln^2 can be bounded up to some universal constant by $\mu^2(\rho|x|^2)$ so that if μ is isotropic i.e. is centered with a covariance matrix equal to identity, we obtain a bound in $\rho^2 n^2$. Optimizing in ρ , we see that the worst case is for $\rho \sim n^{-\frac{2}{3}} C_P^{-\frac{1}{3}}(\mu)$ yielding

Proposition 3.10. *If $\mu(dx) = Z^{-1} e^{-V(x)} dx$ is log-concave and isotropic, then for all $\rho \geq 0$, and $\mu^\rho(dx) = Z_\rho^{-1} e^{-V(x) - \frac{1}{2}\rho|x|^2} dx$,*

$$C_P(\mu^\rho) \leq C n^{2/3} C_P^{\frac{1}{3}}(\mu),$$

for some universal constant C .

This result looks disappointing. A direct approach via the KLS inequality obtained by Chen will presumably give a better dimensional bound provided we are able to get a good bound for the covariance matrix of the perturbed measure. When V is even we can directly estimate the covariance matrix of μ^ρ as in Theorem 18 in [8]. \diamond

3.2. Smooth perturbations.

Using the specific properties of log-concave measures we can state a first result

Theorem 3.11. *If μ_F is log-concave we have for $\varepsilon > 0$,*

$$C_P(\mu_F) \leq \frac{64 \ln(2) (1 + \varepsilon^{-1}) C_P(\mu)}{(1 - 6s)^2} \quad \text{provided} \quad \frac{1 + \varepsilon}{4} C_P(\mu) \mu_F(|\nabla F|^2) := s < \frac{1}{6}.$$

In particular this bound is available as soon as F is L -Lipschitz with $L^2 < 2/(3(1 + \varepsilon) C_P(\mu))$, but this condition is worse than the one in theorem 2.1.

Proof. We deduce from (2.2)

$$\text{Var}_{\mu_F}(f) \leq (1 + \varepsilon^{-1}) C_P(\mu) \mu_F(|\nabla f|^2) + \frac{1 + \varepsilon}{4} C_P(\mu) \text{Osc}^2(f) \mu_F(|\nabla F|^2). \quad (3.1)$$

So that, if μ_F is log-concave, using Theorem 3.2 (2) we get the result. \square

We may similarly modify the proof of proposition 2.2 to similarly get a Cheeger inequality.

Theorem 3.12. *If μ_F is log-concave we have*

$$C'_C(\mu_F) \leq \frac{16 C_C(\mu)}{\pi(1 - 2s)^2} \quad \text{provided} \quad C_C(\mu) \mu_F(|\nabla F|) := s < \frac{1}{2}.$$

In particular if μ is also log-concave we have

$$C_P(\mu_F) \leq \frac{256 \times 64}{\pi^4} \frac{C_P(\mu)}{(1 - 2s)^4},$$

for s as before.

At the level of Cheeger inequality, we found no other comparable perturbation result despite, once again, the very simple argument involved here.

Starting with (2.4) we also have

$$\text{Var}_{\mu_F}(f) \leq \mu_F((f - a)^2) \leq C_P(\mu) \mu_F(|\nabla f|^2) + \frac{1}{2} C_P(\mu) \text{Osc}^2(f) \mu_F([AF - \frac{1}{2} |\nabla F|^2]_+)$$

so that we obtain an improvement of Theorem 2.3

Theorem 3.13. *If V is C^1 , μ_F is log-concave, F is of C^2 class and satisfies*

$$C_P(\mu) \mu_F([AF - \frac{1}{2} |\nabla F|^2]_+) := s < \frac{1}{3}$$

then

$$C_P(\mu_F) \leq \frac{64 \ln(2)}{(1 - 3s)^2} C_P(\mu).$$

Remark 3.14. As in the previous section it is interesting to extend the result to compactly supported log-concave measures. We will thus assume that the set $U = \{V < +\infty\}$ is convex with a smooth boundary and that V is C^1 in U . Then the conclusion of the previous Theorem is still available provided in addition $\partial_n F \geq 0$ on ∂U . \diamond

Remark 3.15. Assume that V is C^2 on \mathbb{R}^n and satisfies $\langle u, \text{Hess}_V(x)u \rangle \geq \rho|u|^2$ for all u and x in \mathbb{R}^n . Let U be an open convex subset given by $U = \{W < 1\}$ where W is a smooth (say C^2) convex function. Consider

$$\mu(dx) = Z^{-1} e^{-V(x)} \mathbf{1}_{W(x) \leq 1} dx.$$

It turns out that once again

$$C_P(\mu) \leq 1/\rho.$$

To prove this result one can use the Γ_2 theory of Bakry-Emery, but one has to carefully define the algebra \mathcal{A} (see section 1.16 in [5]). The devil is in this definition when looking at reflected semi-groups. We prefer to give an elementary proof of what we claimed.

Define $H(x) = (W(x) - 1)^4 \mathbf{1}_{W(x) \geq 1}$. H is smooth and

$$\partial_{ij}^2 H(x) = 4(W(x) - 1)^2 ((W - 1) \partial_{ij}^2 W + 3 \partial_i W \partial_j W) \mathbf{1}_{W(x) \geq 1}$$

so that

$$\langle u, \text{Hess}_H(x)u \rangle = 4(W(x) - 1)^2 ((W - 1) \langle u, \text{Hess}_W(x)u \rangle + 3 \langle u, \nabla W \rangle^2) \mathbf{1}_{W(x) \geq 1} \geq 0.$$

If we consider

$$\mu^\varepsilon(dx) = Z_\varepsilon^{-1} e^{-V(x) - \frac{1}{\varepsilon} H(x)} dx,$$

μ^ε satisfies the Bakry-Emery criterion so that $C_P(\mu^\varepsilon) \leq 1/\rho$. It remains to let ε go to 0 and to pass to the limit in (1.1) by using Lebesgue convergence theorem. \diamond

Let us finish by proving the same type of result at the level of Brascamp-Lieb inequality. Let us consider $d\mu = e^{-V} dx$ with $\text{Hess}(V) > 0$ in the sense of positive definite matrix, the celebrated Brascamp-Lieb inequality is then

$$\text{Var}_\mu(f) \leq \int (\nabla f)^t \text{Hess}(V)^{-1} \nabla f d\mu.$$

We will see that we can easily have some perturbation result, leading to some modified Brascamp-Lieb inequality.

Theorem 3.16. *Let us consider $d\mu = e^{-V} dx$ with $\text{Hess}(V) > 0$ in the sense of positive definite matrix and suppose that there exists ϵ such that*

$$\frac{1}{4}(1 + \epsilon) \|(\nabla F)^t \text{Hess}(V)^{-1} \nabla F\|_\infty < 1$$

then

$$\text{Var}_{\mu_F}(f) \leq \frac{(1 + \epsilon^{-1})}{1 - \frac{1}{4}(1 + \epsilon) \|(\nabla F)^t \text{Hess}(V)^{-1} \nabla F\|_\infty} \int \nabla f^t \text{Hess}(V)^{-1} \nabla f d\mu.$$

In particular if μ_F is log-concave then

$$C_P(\mu_F) \leq \frac{64(1 + \epsilon^{-1}) \ln(2)}{1 - \frac{1}{4}(1 + \epsilon) \|(\nabla F)^t \text{Hess}(V)^{-1} \nabla F\|_\infty} \int \| \text{Hess}(V)^{-1} \|_{HS} d\mu$$

where $\| \cdot \|_{HS}$ denotes the Hilbert-Schmidt norm.

Proof. We follow the idea already developed for the Poincaré inequality.

$$\mathrm{Var}_{\mu_F}(f) \leq \mu_F((f-a)^2) = \mu^{-1}(e^{-F}) \mu \left(((f-a)e^{-\frac{1}{2}F})^2 \right)$$

for which we choose

$$a = \frac{\mu \left(f e^{-\frac{1}{2}F} \right)}{\mu \left(e^{-\frac{1}{2}F} \right)}$$

so that we may apply Brascamp-Lieb inequality for μ .

$$\begin{aligned} \mu_F((f-a)^2) &\leq \int (\nabla f - \frac{1}{2}(f-a)\nabla F)^t \mathrm{Hess}(V)^{-1} (\nabla f - \frac{1}{2}(f-a)\nabla F) d\mu_F \\ &\leq (1+\epsilon^{-1}) \int (\nabla f)^t \mathrm{Hess}(V)^{-1} \nabla f d\mu_F \\ &\quad + \frac{1}{4}(1+\epsilon) \int (f-a)^2 (\nabla F)^t \mathrm{Hess}(V)^{-1} \nabla F d\mu_F. \end{aligned}$$

We then use our growth condition to conclude.

If μ_F is log-concave, we have

$$\mathrm{Var}_{\mu_F}(f) \leq \frac{(1+\epsilon^{-1})}{1 - \frac{1}{4}(1+\epsilon) \|(\nabla F)^t \mathrm{Hess}(V)^{-1} \nabla F\|_\infty} \| |\nabla f| \|_\infty^2 \int \| \mathrm{Hess}(V)^{-1} \|_{HS} d\mu,$$

and we can conclude by using Theorem 3.2 (2). \square

Of course it is illusory to expect a Brascamp-Lieb inequality for μ_F as $\mathrm{Hess}(V+F)$ is not necessarily positive. However it may be useful for concentration inequalities, indeed, reproducing the proof of the exponential integrability for Poincaré inequality due to Bobkov-Ledoux, see [5], under the assumptions of the previous theorem, if f is such that

$$\|(\nabla f)^t \mathrm{Hess}(V)^{-1} \nabla f\|_\infty \leq 1$$

then

$$\forall s < \sqrt{\frac{4(1 - \frac{1}{4}(1+\epsilon) \|(\nabla F)^t \mathrm{Hess}(V)^{-1} \nabla F\|_\infty)}{1+\epsilon^{-1}}}, \quad \int e^{sf} d\mu_F < \infty.$$

This thus implies exponential concentration for μ_F for some particular class of functions.

4. COMING BACK: FROM THE PERTURBED MEASURE TO THE INITIAL ONE.

Any probability measure $\mu(dx) = e^{-V(x)}dx$ can be seen as a perturbation of a perturbed measure, namely $\mu(dx) = Z^{-1} e^{F(x)} \mu_F(dx)$. In some cases the measure μ_F is simpler to study, so that one can expect some results for the initial one using our perturbation method.

4.1. Some consequences using gaussian perturbation.

As an immediate application, for $\rho \in \mathbb{R}^+$ consider

$$\nu(dx) = Z_\rho^{-1} e^{-V(x) - \frac{1}{2}\rho|x|^2} dx.$$

If we denote

$$F(x) = -\frac{1}{2}\rho|x|^2$$

we have with the previous notations

$$\mu(dx) = \nu_F(dx).$$

If V is convex, ν satisfies the Bakry-Emery criterion and accordingly $C_P(\nu) \leq 1/\rho$. We shall use the results in the previous subsection, starting with Theorem 3.11 with $\varepsilon = 1$ for simplicity.

Hence

$$\frac{1}{2} C_P(\nu) \nu_F(|\nabla F|^2) \leq \frac{1}{2\rho} \rho^2 \mu(|x|^2).$$

In order to apply Theorem 3.11 we thus need $s = \frac{1}{2}\rho\mu(|x|^2) \leq 1/6$. We have thus obtained, choosing ρ small enough

$$C_P(\mu) \leq 128 \ln(2) \frac{1}{2s(1-6s)^2} \mu(|x|^2).$$

The optimal choice of s is $1/18$ and of course we may always center μ without changing the Poincaré constant.

Corollary 4.1. *Let μ be a log-concave measure. Then*

$$C_P(\mu) \leq 32 \times 81 \ln(2) \mu(|x - \mu(x)|^2).$$

This result is well known and according to [2] p.11 is contained in [27] with a much better pre-constant 4 (also see [11] (1.8) with a non explicit constant).

Applying Theorem 3.2 it is easily seen (see (9.2.13) in [19]) that

$$C'_C(\mu) \leq \frac{16}{\pi} \mu(|x - m_\mu(x)|) \leq \frac{16}{\pi} \mu(|x - \mu(x)|). \quad (4.1)$$

If we replace Theorem 3.11 by Theorem 3.12 we obtain using our perturbation method the worse

$$C'_C(\mu) \leq \frac{100\sqrt{10}}{\pi^2} \mu(|x - \mu(x)|).$$

Remark 4.2. In what precedes we may replace $F(x) = -\frac{1}{2}\rho|x|^2$ by $F(x) = -\rho H(x)$ with $\text{Hess}H(x) \geq Id$, without changing anything. Hence for example Corollary 4.1 can be generalized in:

Corollary 4.3. *If μ is log-concave, for all C^2 function H satisfying $\text{Hess}H(x) \geq Id$, it holds*

$$C_P(\mu) \leq 32 \times 81 \ln(2) \mu(|\nabla H|^2). \quad (4.2)$$

◇

Remark 4.4. What happens if instead we try to use Theorem 3.13. With the notations of the previous subsection it holds

$$A = \Delta - \nabla V \cdot \nabla - \rho x \cdot \nabla$$

so that

$$AF - \frac{1}{2} |\nabla F|^2 = -\rho n + \rho x \cdot \nabla V(x) + \frac{1}{2} \rho^2 |x|^2, \quad (4.3)$$

and

$$C_P(\mu) \leq \frac{64 \ln(2)}{(1-3s)^2 \rho}$$

as soon as

$$\mu \left(\left[-\rho n + \rho x \cdot \nabla V(x) + \frac{1}{2} \rho^2 |x|^2 \right]_+ \right) \leq \frac{\rho}{3}.$$

Though the result looks stronger than the previous ones we did not succeed in really exploring some interesting consequences, in terms of dimensional controls of the Poincaré inequality.

◇

4.2. Using product Subbotin (exponential power) perturbations.

Here we shall use the idea of the previous subsection replacing the gaussian measure by the tensor product of Subbotin distributions and the Bakry-Emery criterion by results of Barthe-Klartag.

For $p \geq 1$ and $\lambda > 0$ consider

$$\nu(dx) = Z^{-1} e^{-V(x) - \lambda^p \sum_{i=1}^n |x_i|^p} dx,$$

where we assume that V is a convex function. Let X be a random variable with distribution ν , then λX has distribution

$$\nu(\lambda, dx) = Z_\lambda^{-1} e^{-V(x/\lambda) - \sum_{i=1}^n |x_i|^p} dx,$$

and the dilation property for the Poincaré constant gives

$$C_P(\nu) = \lambda^{-2} C_P(\nu(\lambda, \cdot)).$$

As before, if we denote

$$F(x) = -\lambda^p \sum_{i=1}^n |x_i|^p$$

we have

$$\mu(dx) = \nu_F(dx).$$

We thus have

$$|\nabla F|^2(x) \leq \lambda^{2p} p^2 \sum_{i=1}^n |x_i|^{2(p-1)}, \quad (4.4)$$

so that

$$\mu(|\nabla F|^2) = \nu_F(|\nabla F|^2) \leq \lambda^{2p} p^2 \mu \left(\sum_{i=1}^n |x_i|^{2(p-1)} \right). \quad (4.5)$$

Choosing for simplicity $\varepsilon = 1$ and $s = 1/12$ in Theorem 3.11 we thus have to choose (if this choice is possible)

$$\lambda^{2(p-1)} p^2 C_P(\nu(\lambda, .)) \mu \left(\sum_{i=1}^n |x_i|^{2(p-1)} \right) = \frac{1}{6}. \quad (4.6)$$

Notice that we may always use an upper bound for $C_P(\nu(\lambda, .))$ furnishing a lower bound for λ and an upper bound for $C_P(\mu)$.

For which p 's do we obtain interesting results ?

One cannot expect that (4.6) can be satisfied for $p = 1$, since the left hand side is of size n .

If $p \neq 1$ we obtain, provided (4.6) is satisfied,

$$C_P(\mu) \leq C 6^{\frac{1}{p-1}} p^{\frac{2}{p-1}} \mu^{\frac{1}{p-1}} \left(\sum_{i=1}^n |x_i|^{2(p-1)} \right) C_P^{\frac{p}{p-1}}(\nu(\lambda, .)), \quad (4.7)$$

with $C = 512 \ln(2)$.

In particular if V is even, we may apply Theorem 1.3 (since $C_P(\nu(\lambda, .))$ can be bounded independently of λ) and get for $1 < p < 2$,

$$C_P(\mu) \leq C 6^{\frac{1}{p-1}} p^{\frac{2}{p-1}} \mu^{\frac{1}{p-1}} \left(\sum_{i=1}^n |x_i|^{2(p-1)} \right) (\ln(n))^{\frac{2-p}{p-1}}, \quad (4.8)$$

for some universal C .

Compared with (4.1), this result for $p = \frac{3}{2}$ is however bad w.r.t. the dimension.

Looking at (4.8) it seems interesting to get an analogue of Theorem 1.3 i.e. a bound for $C_P(\nu(\lambda, .))$ that does not depend on λ but for $p > 2$.

Indeed, if V is even, $\mu(x) = 0$ and for $2 < p$,

$$\mu(|x_i|^{2(p-1)}) \leq \frac{\Gamma(2p+1)}{2^{p-1}} \mu^{p-1}(|x_i|^2) \leq (p-1)^{2(p-1)} \mu^{p-1}(|x_i|^2) \quad (4.9)$$

according to [26] corollary 5.7 and remark 5.8 (also see in [29] the discussion after definition 2). In particular we obtain, provided (4.6) i.e.

$$\lambda^2 = \left(\frac{1}{6p^2 C_P(\nu(\lambda, .)) n} \right)^{\frac{1}{p-1}} \frac{1}{(p-1)^2 \sigma^2(\mu)} \quad (4.10)$$

is satisfied, for $p > 2$,

$$C_P(\mu) \leq C 6^{\frac{1}{p-1}} p^{\frac{2}{p-1}} (p-1)^2 n^{\frac{1}{p-1}} C_P^{\frac{p}{p-1}}(\nu(\lambda, .)) \sigma^2(\mu), \quad (4.11)$$

with $C = 512 \ln(2)$. This time if p is of order $\ln(n)$ we will get an interesting result, provided $C_P(\nu(\lambda, .))$ is controlled by some not too bad constant (possibly depending on n).

Unfortunately, [8] contains an example (see subsection 3.4) where $C_P^{\frac{p}{p-1}}(\nu(\lambda, .))$ behaves like $n^{\frac{p-2}{p-1}}$, but for a measure μ which is highly non isotropic.

Nevertheless if we assume in addition that μ is unconditional (i.e $V(x_1, \dots, x_n) = V(|x_1|, \dots, |x_n|)$ for all x so that $V(./\lambda)$ is also unconditional), it follows from Theorem 17 in [8] that

$$C_P(\nu(\lambda, dx)) \leq C_P(S_p(dx_1))$$

for all λ , where

$$S_p(dx_1) = \frac{1}{z_p} e^{-|x_1|^p}.$$

According to Bobkov's one dimensional result ([11] Corollary 4.3),

$$C_P(S_p) \leq 12 \operatorname{Var}_{S_p}(x) = 12 \frac{\Gamma(3/p)}{\Gamma(1/p)}. \quad (4.12)$$

A better result is obtained by combining [12] Theorem 2.1 and the dilation property of Poincaré constants yielding

$$C_P(S_p) \leq \frac{p^{1-2/p}}{2(1+p)^{1-2/p}}.$$

Hence in the unconditional case,

$$\begin{aligned} C_P(\mu) &\leq C 6^{\frac{1}{p-1}} \left(\frac{p^2}{2}\right)^{\frac{1}{p-1}} (p-1)^2 n^{\frac{1}{p-1}} \frac{p^{\frac{p-2}{p-1}}}{(1+p)^{\frac{p-2}{p-1}}} \sigma^2(\mu) \\ &\leq 4C 3^{\frac{1}{p-1}} (p-1)^2 n^{\frac{1}{p-1}} \sigma^2(\mu), \end{aligned} \quad (4.13)$$

with $C = 512 \ln(2)$. Here we used $p^{2/(p-1)} \leq 4$ for $p \geq 2$.

It remains to optimize in p , the optimal value being $p-1 = \ln(3n)/2$ (which is larger than 1 for $n \geq 2$).

Proposition 4.5. *For $n \geq 2$, any unconditional log-concave probability measure μ satisfies*

$$C_P(\mu) \leq C \ln^2(3n) \sigma^2(\mu)$$

with $C = 512 e^2 \ln(2)$.

This result is not new, and is due to Klartag in [28] with a non explicit constant. Another proof (still with a difficult to trace constant) is contained in [19]. The constant here is explicit (but certainly far to be sharp), but the most interesting fact is that this result can be obtained via Subbotin perturbation.

Remark 4.6. Notice that the choice $p-1 = \ln(3n)/2$ gives $\lambda \sim C/\ln(n)$ for n large enough, according to (4.10).

In second place $C_p(S_p)$ is uniformly bounded in p (S_p weakly converges to the uniform distribution on $[-1, 1]$ as p goes to infinity). One can thus be tempted to use another product measure based on a one dimensional log concave family such that the Poincaré constant goes to 0 as p goes to infinity. We did not succeed in following this direction. \diamond

Remark 4.7. We have only introduced $\nu(\lambda, dx)$ in order to directly connect the previous proof to Barthe and Klartag results. We should of course directly estimate the Poincaré constant of ν . \diamond

5. APPLICATION TO SOME PROBLEMS IN BAYESIAN STATISTICS.

5.1. Sparse linear regression.

In [23] the authors proposed a Bayesian strategy for the liner regression model

$$Y_i = \langle X_i, \lambda^* \rangle + \xi_i \quad (5.1)$$

where λ^* and each X_i belong to \mathbb{R}^M , ξ_i are i.i.d. scalar noises and $i = 1, \dots, n$. n is thus the size of a sample while M is the dimension of the predictor. Given a collection of design points X_i the exponentially weighted aggregate estimator of λ^* is given by

$$\hat{\lambda}_n(X) = \int \lambda \hat{\pi}_{n,\beta}(d\lambda) \quad (5.2)$$

where

$$\hat{\pi}_{n,\beta}(d\lambda) = C \exp \left(-(1/\beta) \sum_{i=1}^n |Y_i - \langle X_i, \lambda \rangle|^2 \right) \pi(d\lambda) \quad (5.3)$$

is the posterior probability distribution associated to the prior $\pi(d\lambda) = e^{-W(\lambda)} d\lambda$ and the temperature β . Here and in all what follows C is a normalizing constant that can change from line to line.

In their Theorem 2 they obtain in particular an explicit bound for the \mathbb{L}^2 error, when the prior is (almost) chosen as

$$\pi(d\lambda) = C \prod_{j=1}^M \frac{e^{-\alpha|\lambda_j|}}{(\tau^2 + \lambda_j^2)^2} \mathbf{1}_{\sum_{j=1}^M |\lambda_j| \leq R} d\lambda = e^{-W(\lambda)} d\lambda, \quad (5.4)$$

for some positive α and R . This choice is motivated by dimensional reasons when $M \gg n$ and λ^* is sparse.

In order to compute $\hat{\lambda}_n$, they propose to use the ergodic theorem applied to the Langevin diffusion process

$$dL_t = \sqrt{2} dB_t - \nabla W(L_t) dt - \frac{2}{\beta} (\langle X_i, L_t \rangle - Y_i) dt \quad (5.5)$$

where B is a standard \mathbb{R}^M valued Brownian motion, i.e. the \mathbb{L}^1 convergence of $\frac{1}{t} \int_0^t L_s ds$ to the desired $\hat{\lambda}_n$ as $t \rightarrow +\infty$. The $\mathbf{1}_{\sum_{j=1}^M |\lambda_j| \leq R}$ is no more considered here, and thus denote

$$\nu_{n,\beta}(d\lambda) = C \exp \left(-(1/\beta) \sum_{i=1}^n |Y_i - \langle X_i, \lambda \rangle|^2 - \sum_{i=1}^n \log(\tau^2 + \lambda_i^2) \right) \prod_{j=1}^M e^{-\alpha|\lambda_j|} d\lambda. \quad (5.6)$$

Remark that $\nu_{n,\beta}$ is not logconcave. To justify some rate of convergence they call upon the Meyn-Tweedie theory (see e.g. [33]) of Foster-Lyapunov functions. Notice that using instead $\frac{1}{t} \int_t^{2t} L_s ds$ the bound (8) in [23] becomes $C \theta^t$ for some $\theta < 1$.

If the arguments give an exponential rate θ^t for some $\theta < 1$ of convergence in their proposition 1, they are far to provide us with a bound for $-\ln(\theta)$. The reason is that constants are very difficult to trace with this theory. In several papers, [3, 21, 18], we have established the links between the Poincaré inequality and the existence of Lyapunov functions, including their relationship with hitting times. One can find explicit bounds in these papers. A quick look

to the explicit bounds in [21] show that it furnishes a rate of convergence that heavily depend on M , assumed to be big, on β assumed to be big too, and the observations.

Since $\nu_{n,\beta}$ is reversible for the Langevin diffusion, and satisfies a Poincaré inequality, L_t itself converges to $\hat{\lambda}_n$ in $\mathbb{L}^2(\nu_{n,\beta})$ at an exponential rate given by $e^{-t/C_P(\nu_{n,\beta})}$. There is no need of Cesaro average.

Our goal is thus to get some interesting bounds for $C_P(\nu_{n,\beta})$. To this end we may use two methods. In what follows Z is a normalizing constant that may change from line to line.

First, we may write $\nu_{n,\beta} = \mu_F$ with

$$\mu(d\lambda) = Z^{-1} \exp \left(-(1/\beta) \sum_{i=1}^n \langle X_i, \lambda \rangle^2 \right) \prod_{j=1}^M e^{-\alpha |\lambda_j|} d\lambda$$

and

$$F(\lambda) = -\frac{2}{\beta} \sum_{i=1}^n Y_i \langle X_i, \lambda \rangle - \sum_{i=1}^n \log(\tau^2 + \lambda_i^2) = -\frac{2}{\beta} \sum_{j=1}^M \lambda_j \langle Y, \tilde{X}^j \rangle - \sum_{i=1}^n \log(\tau^2 + \lambda_i^2)$$

where $\tilde{X} = (X_1^j, \dots, X_n^j)$. But according to Barthe and Klartag result Theorem 1.3, the tensorisation property and the quadratic behaviour of the Poincaré with respect to dilation we have

$$C_P(\mu) \leq C \ln^2(M) \frac{1}{\alpha^2},$$

for some universal constant. Since F is L -Lipschitz with

$$L = \frac{2}{\beta} \sup_{j=1, \dots, M} |\langle Y, \tilde{X}^j \rangle| + \frac{1}{\tau}$$

we may apply Theorem 2.1

Theorem 5.1. *There exist two universal positive constants c and C such that, provided*

$$\frac{\ln(M)}{\beta \alpha} \left(\sup_{j=1, \dots, M} |\langle Y, \tilde{X}^j \rangle| + \beta \tau \right) \leq c$$

then

$$C_P(\nu_{n,\beta}) \leq C \frac{\beta}{\sup_{j=1, \dots, M} |\langle Y, \tilde{X}^j \rangle| + \beta \tau}.$$

Depending on whether the logarithmic factor $\ln(M)$ is necessary in Theorem 1.3 or not will yield a result that does not depend on the dimension.

As said before the situation considered here is $M \gg n$. Another approach to compute $C_P(\mu)$ will thus be to use Theorem 3.3 instead of Theorem 1.3. The discussion will thus be very similar to the one in Example 3.9. Of course because the quadratic form $\sum_{i=1}^n \langle X_i, \lambda \rangle^2$ is very degenerate on \mathbb{R}^M it is impossible to use Bakry-Emery criterion. Since our goal is not to rewrite [23] but to see how one can control the rate of convergence of the Langevin dynamics, we will assume for simplicity that the X_i 's are an orthogonal family. So, after an orthogonal transform, we may write

$$\mu(d\lambda) = Z^{-1} \exp \left(-(1/\beta) \sum_{i=1}^n |X_i|^2 \lambda_i^2 \right) e^{-V(\lambda)} d\lambda$$

where $e^{-V(\lambda)} d\lambda$ an orthogonal change of $\prod_{j=1}^M e^{-\alpha|\lambda_j|} d\lambda$, hence shares the same Poincaré constant. Since the Poincaré constant of the symmetric exponential isotropic distribution in dimension 1, hence the one of the tensor product of such distributions, is equal to 4, the dilation scaling yields $C_P(e^{-V(\lambda)} d\lambda) = 4/\alpha^2$.

Of course we may consider, when $M \geq n$,

$$\eta_n(d\lambda_1, \dots, d\lambda_n) = \left(\int e^{-V(\lambda)} d\lambda_{n+1} \dots d\lambda_M \right) d\lambda_1 \dots d\lambda_n$$

which is a new log-concave distribution according to Prekopa-Leindler theorem. It is still isotropic up to a dilation of scale α and

$$\int \exp \left(-(1/\beta) \sum_{i=1}^n |X_i|^2 \lambda_i^2 \right) e^{-V(\lambda)} d\lambda = \int \exp \left(-(1/\beta) \sum_{i=1}^n |X_i|^2 \lambda_i^2 \right) \eta_n(d\lambda_1, \dots, d\lambda_n).$$

In addition the Poincaré constant of η_n is less than the one of $e^{-V} d\lambda$. We may thus argue as in Example 3.9 when $\alpha = 1$, and then use the dilation property of the Poincaré inequality to conclude that for all β and X ,

$$C_P(\mu) \leq C \frac{n^{2/3}}{\alpha^2},$$

for some universal constant C . We can now follow what we did previously to get

Theorem 5.2. *There exist two universal positive constants c and C such that, provided*

$$\frac{n^{1/3}}{\beta \alpha} \left(\sup_{j=1, \dots, M} |\langle Y, \tilde{X}^j \rangle| + \beta \tau \right) \leq c$$

then

$$C_P(\nu_{n,\beta}) \leq C \frac{\beta}{\sup_{j=1, \dots, M} |\langle Y, \tilde{X}^j \rangle| + \beta \tau}.$$

5.2. Parameter identification.

We shall briefly indicate another problem, parameter identification via a Bayesian approach as studied in the recent [24]. Given a convex function $U : \mathbb{R}^q \times \mathbb{R}^d \mapsto \mathbb{R}$ one considers the family of probability densities

$$f_\theta(x) := f(x, \theta) = Z_\theta^{-1} e^{-U(x, \theta)}$$

on \mathbb{R}^q . Given the observation of a sample $X = (X_1, \dots, X_n)$ of i.i.d. random vectors with density f_{θ^*} , one wants to estimate the unknown parameter θ^* , here again using a Bayesian procedure.

For a prior distribution density $\pi_0(\theta)$, the posterior density is thus

$$\pi_n(\theta) = \pi_0(\theta) \prod_{i=1}^n f_\theta(X_i), \quad (5.7)$$

and the natural bayesian estimator of θ^* is once again given by

$$\hat{\theta}_n = \int \theta \pi_n(\theta) d\theta. \quad (5.8)$$

It is shown in [24] that under mild assumptions this estimator is consistent and a bound for the \mathbb{L}^p error is given, provided there exists a constant C_P^U such that

$$C_P(f_\theta(x) dx) \leq C_P^U \quad \text{for all } \theta \in \text{support}(\pi_0), \quad (5.9)$$

An important case where this assumption is satisfied is the location problem, i.e. when $U(x, \theta) = V(x - \theta)$ for some convex function V , and of course $q = d$.

Here again the authors propose to use a Langevin Monte Carlo procedure to compute $\hat{\theta}_n$, and as in the previous subsection the problem is now to estimate the Poincaré constant of the measure $\pi_n(\theta) d\theta$. The situation is of course much simpler here if one chooses π_0 as a strictly log-concave log concave distribution since we obtain a bound that only depends on the curvature of π_0 . The case of a general log-concave measure π_0 is studied in [24].

Acknowledgements

This work has been (partially) supported by the Project EFI ANR-17-CE40-0030 of the French National Research Agency.

REFERENCES

- [1] S. Aida and I. Shigekawa. Logarithmic Sobolev inequalities and spectral gaps: perturbation theory. *J. Funct. Anal.*, 126(2):448–475, 1994.
- [2] D. Alonso-Gutierrez and J. Bastero. *Approaching the Kannan-Lovasz-Simonovits and variance conjectures*, volume 2131 of *LNM*. Springer, 2015.
- [3] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measures. *Elec. Comm. in Prob.*, 13:60–66, 2008.
- [4] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes : Lyapunov versus Poincaré. *J. Funct. Anal.*, 254:727–759, 2008.
- [5] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov diffusion operators.*, volume 348 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 2014.
- [6] D. Bakry, M. Ledoux, and F-Y. Wang. Perturbations of functional inequalities using growth conditions. *J. Math. Pures Appl. (9)*, 87(4):394–407, 2007.
- [7] J.B Bardet, N. Gozlan, F. Malrieu, and P. A. Zitt. Functional inequalities for gaussian convolutions of compactly supported measures: explicit bounds and dimension dependence. *Bernoulli*, 24(1):333–353, 2018.
- [8] F. Barthe and B. Klartag. Spectral gaps, symmetries and log-concave perturbations. *Bull. Hellenic Math. Soc.*, 64:1–31, 2020.
- [9] F. Barthe and E. Milman. Transference principles for log-sobolev and spectral-gap with applications to conservative spin systems. *Comm. Math. Physics*, 323:575–625, 2013.
- [10] N. Berestycki and R. Nickl. Concentration of measure. Available at <http://www.statslab.cam.ac.uk/~beresty/teach/cm10.pdf>, 2009.
- [11] S. G. Bobkov. Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Probab.*, 27(4):1903–1921, 1999.
- [12] M. Bonnefont, A. Joulin, and Y. Ma. A note on spectral gap and weighted Poincaré inequalities for some one-dimensional diffusions. *ESAIM Probab. Stat.*, 20:18–29, 2016.
- [13] H.J. Brascamp and E.H. Lieb. On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Funct. Anal.*, 22:366–389, 1976.
- [14] P. Cattiaux. A pathwise approach of some classical inequalities. *Potential. Anal.*, 20:361–394, 2004.
- [15] P. Cattiaux. Hypercontractivity for perturbed diffusion semigroups. *Ann. Fac. Sciences Toulouse Math. (6)*, 14(4):609–628, 2005.
- [16] P. Cattiaux, M. Fathi, and A. Guillin. Self-improvement of the Bary-Emery criterion for Poincaré inequalities and Wasserstein contraction using variable curvature bounds. Available on Math. ArXiv 2002.09221 [math FA], 2020.

- [17] P. Cattiaux and A. Guillin. Semi log-concave Markov diffusions. In *Séminaire de Probabilités XLVI*, volume 2123 of *Lecture Notes in Math.*, pages 231–292. Springer, Cham, 2014.
- [18] P. Cattiaux and A. Guillin. Hitting times, functional inequalities, Lyapunov conditions and uniform ergodicity. *J. Funct. Anal.*, 272(6):2361–2391, 2017.
- [19] P. Cattiaux and A. Guillin. On the Poincaré constant of log-concave measures. In *Geometric aspects of functional analysis, Israel Seminar 2017-2019, Vol.1*, volume 2256 of *Lecture Notes in Math.*, pages 171–217. Springer, Berlin, 2020.
- [20] P. Cattiaux, A. Guillin, and L. Wu. A note on Talagrand’s transportation inequality and logarithmic Sobolev inequality. *Probab. Theory Related Fields*, 148(1-2):285–304, 2010.
- [21] P. Cattiaux, A. Guillin, and P. A. Zitt. Poincaré inequalities and hitting times. *Ann. Inst. Henri Poincaré. Prob. Stat.*, 49(1):95–118, 2013.
- [22] Y. Chen. An almost constant lower bound of the isoperimetric coefficient in the KLS conjecture. Available on Math. ArXiv 2011.13661 [math PR], 2021.
- [23] A. S. Dalalyan and A. B. Tsybakov. Sparse regression learning by aggregation and Langevin Monte-Carlo. *J. Comput. System Sci.*, 78(5):1423–1443, 2012.
- [24] S. Gadat, F. Panloup, and C. Pellegrini. On the cost of bayesian posterior mean strategy for log-concave models. Available on Math. ArXiv. 2010.06420 [math.PR]., 2020.
- [25] F. Gong and L. Wu. Spectral gap of positive operators and applications. *C. R. Acad. Sci. Paris Sér. I Math.*, 331(12):983–988, 2000.
- [26] O. Guédon, P. Nayar, and T. Tkocz. Concentration inequalities and geometry of convex bodies. In *Analytical and probabilistic methods in the geometry of convex bodies*, volume 2 of *IMPAN Lect. Notes*, pages 9–86. Polish Acad. Sci. Inst. Math., Warsaw, 2014.
- [27] R. Kannan, L. Lovasz, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.*, 13(3-4):541–559, 1995.
- [28] B. Klartag. A Berry-Esseen type inequality for convex bodies with an unconditional basis. *Probab. Theory Relat. Fields*, 145(1-2):1–33, 2009.
- [29] R. Latała. On some problems concerning log-concave random vectors. In *Convexity and concentration*, volume 161 of *IMA Vol. Math. Appl.*, pages 525–539. Springer, New York, 2017.
- [30] M. Ledoux. Spectral gap, logarithmic Sobolev constant, and geometric bounds. In *Surveys in differential geometry.*, volume IX, pages 219–240. Int. Press, Somerville MA, 2004.
- [31] M. Ledoux. From concentration to isoperimetry: semigroup proofs. In *Concentration, functional inequalities and isoperimetry*, volume 545 of *Contemp. Math.*, pages 155–166. Amer. Math. Soc., Providence, RI, 2011.
- [32] Y. T. Lee and S. S. Vempala. Eldan’s stochastic localization and the KLS hyperplane conjecture: An improved lower bound for expansion. In *Proc. of IEEE FOCS*, 2017., 2017.
- [33] S. P. Meyn and R. L. Tweedie. Stability of markovian processes III: Foster-Lyapunov criteria for continuous-time processes. *Adv. Appl. Proba.*, 25:518–548, 1993.
- [34] E. Milman. On the role of convexity in isoperimetry, spectral-gap and concentration. *Invent. math.*, 177:1–43, 2009.
- [35] E. Milman. Isoperimetric and concentration inequalities: Equivalence under curvature lower bound. *Duke Math. J.*, 154(2):207–239, 2010.
- [36] E. Milman. Isoperimetric bounds on convex manifolds. In *Concentration, Functional inequality and Isoperimetry*, volume 545 of *Contemporary Mathematics*, pages 195–208. Amer. Math. Soc., 2011.
- [37] A. Saumard and J. A. Wellner. Log-concavity and strong log-concavity: a review. *Statistics Surveys.*, 8:45–114, 2014.
- [38] M. Troyanov. Concentration et inégalité de Poincaré. Available at <https://infoscience.epfl.ch/record/118471/files/concentration2001.pdf>, 2001.
- [39] D. Zimmermann. Logarithmic Sobolev inequalities for mollified compactly supported measures. *J. Funct. Anal.*, 265:1064–1083, 2013.

Patrick CATTIAUX,, INSTITUT DE MATHÉMATIQUES DE TOULOUSE. CNRS UMR 5219., UNIVERSITÉ PAUL SABATIER,, 118 ROUTE DE NARBONNE, F-31062 TOULOUSE CEDEX 09.

Email address: `patrick.cattiaux@math.univ-toulouse.fr`

Arnaud GUILLIN,, UNIVERSITÉ CLERMONT AUVERGNE, CNRS, LMBP, F-63000 CLERMONT-FERRAND, FRANCE.

Email address: `arnaud.guillin@uca.fr`