Chapter 10
Remarks on Superconcentration and Gamma Calculus: Applications to Spin Glasses

Kevin Tanguy

Abstract This note is concerned with the so-called superconcentration phenomenon. It shows that the Bakry-Émery's Gamma calculus can provide relevant bound on the variance of function satisfying a inverse, integrated, curvature criterion. As an illustration, we present some variance bounds for the Free Energy in different models from Spin Glasses Theory.

10.1 Introduction

Superconcentration phenomenon has been introduced by Chatterjee in [7] and has given birth to a lot a work (cf. [15] for a survey). Each of these works, used various ad-hoc methods to improve upon sub-optimal bounds given by classical concentration of measure (cf. [4, 10]). In this note, we want to show that the celebrated Gamma calculus from Bakry and Émery's Theory is relevant to such improvements. To this task, we introduce an inverse, integrated, Γ_2 criterion which provides a useful bound on the variance of a particular function. As far as we know, this criterion seems to be new. We give below a sample of our modest achievement.

Denote by γ_n the standard Gaussian measure on $\mathbb{R}^n$ and by $(P_t)_{t\geq 0}$ the standard Ornstein–Uhlenbeck semigroup. Γ will stand for the so-called "carré du champ" operator, associated to the infinitesimal generator $L = \Delta - x \cdot \nabla$ of $(P_t)_{t\geq 0}$, and Γ_2 its iterated operator. We refer to Sect. 10.2 for more details about this topic.

Theorem 10.1.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a regular function and assume that there exists $\psi : \mathbb{R}_+ \to \mathbb{R}$ such that*

(1) for any $t \geq 0$,

$$\int_{\mathbb{R}^n} \Gamma_2(P_t f) d\gamma_n \leq \int_{\mathbb{R}^n} \Gamma(P_t f) d\gamma_n + \psi(t), \tag{10.1}$$

K. Tanguy (✉)
University of Angers, Angers, France
e-mail: kevin.tanguy@univ-angers.fr

© Springer Nature Switzerland AG 2019

N. Gozlan et al. (eds.), *High Dimensional Probability VIII*,
Progress in Probability 74, https://doi.org/10.1007/978-3-030-26391-1_10

(2)

$$\int_0^\infty e^{-2t} \int_t^\infty e^{2s} \psi(s) ds dt < \infty.$$

Then the following holds

$$\mathrm{Var}_{\gamma_n}(f) \leq \left| \int_{\mathbb{R}^n} \nabla f d\gamma_n \right|^2 + 4 \int_0^\infty e^{-2t} \int_t^\infty e^{2s} \psi(s) ds dt.$$

with $|\cdot|$ *the standard Euclidean norm.*

Remark Equation (10.14) can be seen as an inverse, integrated, curvature inequality for the function f.

As an application of Theorem 10.1.1, we show that some results due to Chatterjee can be expressed in terms of such criterion. From our point of view, this expression seems to ease the original scheme of proof and could possibly lead to various extensions. It also permits to easily recover some known variance bounds in Spin Glass Theory (cf. [5, 6, 11, 12]). Therefore let us present a short introduction to this theory.

Most of the time, in Spin Glasses Theory, it is customary to consider a centered Gaussian field $\left(H_n(\sigma)\right)_{\sigma \in \{-1,1\}^n}$ on the discrete cube $\{-1, 1\}^n$ (the map $\sigma \mapsto H_n(\sigma)$ is called the Hamiltonian of the system) and to focus on $\max_{\sigma \in \{-1,1\}^n} H_n(\sigma)$ (or $\min_{\sigma \in \{-1,1\}^n} H_n(\sigma)$). In general, this quantity is rather complex and presents a lack of regularity. Therefore, one focusses on a smooth approximation of the maximum (or the minimum) called the Free Energy $F_{n,\beta}$. This function is defined as follow

$$F_{n,\beta} = \pm \frac{1}{\beta} \log \left(\sum_{\sigma \in \{-1,1\}^n} e^{\pm \beta H_n(\sigma)} \right)$$

where $\beta > 0$ corresponds to (the inverse of) the temperature and its sign depends on whether you want to study the maximum or the minimum of H_n over the discrete cube.

For instance, for the Random Energy Model (REM in short), we have

$$H_n(\sigma) = \sqrt{n} X_\sigma, \quad \sigma \in \{-1, 1\}^n$$

where $(X_\sigma)_{\sigma \in \{-1,1\}^n}$ is a sequence of i.i.d. standard Gaussian random variables.

For the Sherrington and Kirkpatrick's model (SK model in short), the Hamiltonian is more complex,

$$H_n(\sigma) = -\frac{1}{\sqrt{n}} \sum_{i,j=1}^{n} X_{ij}\sigma_i\sigma_j, \quad \sigma \in \{-1,1\}^n$$

with $(X_{ij})_{1\le i,j\le n}$ a sequence of i.i.d. standard Gaussian random variables.

As an application of our methodology (cf. Sect. 10.4), we prove the following two Propositions.

Proposition 10.1.1 *The following holds for the SK model. Let* $0 < \beta < \frac{1}{2}$*, then*

$$\mathrm{Var}(F_{n,\beta}) \le C_\beta, \quad n \ge 1 \tag{10.2}$$

where $C_\beta > 0$ *is a constant depending only on* β*.*

Remark Talagrand obtained (cf. [11, 12]) such upper bound on the variance, for $0 < \beta < 1$, as a consequence of precise (and much harder to prove than our variance bounds) concentration inequalities for the Free Energy together with second moment method. As far as we know, it is the first time that such bound is obtained through semigroups arguments.

The methodology can also be used for the Random Energy Model (REM in short) (cf. Sect. 10.4 for more details) and provides the following bounds.

Proposition 10.1.2 *The following holds in the REM.*

High temperature regime: for $0 < \beta < \frac{1}{\sqrt{2n}}$*, we have*

$$\mathrm{Var}_{\gamma_{2^n}}(F_{n,\beta}) \le \frac{n}{2^n}\left(\frac{1-n\beta^2}{1-2n\beta^2}\right), \quad n \ge 1$$

with $C > 0$ *a universal constant.*

Remark

(1) The preceding bound has to be compared with the results exposed in [6, 7] (be careful with the different renormalization, in [6] the free energy is $\frac{F_{n,\beta}}{n}$). In [6], it is shown that

$$\mathrm{Var}_{\gamma_{2^n}}(F_{n,\beta}) \sim \frac{1}{2^n} \times \frac{e^{n\beta^2}}{\beta^2}, \quad \beta < \sqrt{\frac{\log 2}{2}}.$$

The dependance (in n and β) is clearly not optimal in this regime but, as presented in Proposition 10.1.1, the scheme of proof of our method is robust enough to treat more complicated models. It seems natural that it can fail to capture precise behaviour such as the one obtained in [6]. Notice also that in [6], the authors obtained various (according to the temperature β) asymptotic convergence results for the (renormalized) Free Energy. Therefore, their results only indicate the correct order of the variance of this functional. However, to our best knowledge, this is the first time that such non-asymptotic bounds on the variance of the Free energy is obtained for the high temperature regime temperature.

(2) In [6] the low temperature regime was also investigated. Non-asymptotic variance bound, in accordance with the convergence results from Bovier et al., was already obtained in [7] and is presented and commented in Sect. 10.4 (Proposition 10.4.4) for the sake of completeness.

(3) As we will see latter in this note, it is easier to do the proof (of the preceding result) with the standard Gaussian measure on $\mathbb{R}^n$ and then to perform the following substitutions

$$n \longleftrightarrow 2^n \quad \text{and} \quad \beta \longleftrightarrow \sqrt{n}\beta$$

to fit the framework of [6].

This note is organized as follows. In Sect. 10.2, we recall some facts about superconcentration and Gamma calculus. In Sect. 10.3, we will prove our main results. Finally, in Sect. 10.4, we will give some applications in Spin Glass Theory.

10.2 Framework and Tools

In this section, we briefly recall some notions about superconcentration, Gamma calculus and interpolation methods by semigroups. General references about these topics could be, respectively, [1, 7].

10.2.1 Superconcentration

It is well known (cf. [4, 10]), that concentration of measure of phenomenon is useful in various mathematical contexts. Such phenomenon can be obtained through functional inequalities. For instance, the standard Gaussian measure, on $\mathbb{R}^n$, γ_n satisfies a Poincaré's inequality:

Proposition 10.2.1 *For any function $f : \mathbb{R}^n \to \mathbb{R}$ smooth enough, the following holds*

$$\mathrm{Var}_{\gamma_n}(f) \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n \tag{10.3}$$

where $|\cdot|$ stands for the Euclidean norm.

Although this inequality holds for a large class of function, it could lead to sub-optimal bounds. A classical example is the function $f(x) = \max_{i=1,\ldots,n} x_i$. For such function, Poincaré's inequality implies that

$$\mathrm{Var}_{\gamma_n}(f) \leq 1$$

but it is known that $\mathrm{Var}_{\gamma_n}(f) \sim \frac{C}{\log n}$ for some constant $C > 0$. In Chatterjee's terminology, in this Gaussian framework, a function f is said to be superconcentrated when Poincaré's inequality (10.3) is sub-optimal.

As we have said in the introduction, this phenomenon has been studied in various manner: semigroup interpolation [14], Renyi's representation of order statistics [3], Optimal Transport [15], Ehrard's inequality [17],...(cf. the Thesis [16] for a recent survey about superconcentration). In this note, we want to show that some differential inequalities between the operator Γ and Γ_2 from Bakry and Émery's Theory could provide superconcentration.

10.2.2 Semigroups Interpolation and Gamma Calculus

For more details about semigroups interpolation and Γ calculus, we refer to [1, 9]. Although our work can easily be extended to a more general framework, we will focus on a Gaussian setting.

The Ornstein–Uhlenbeck process $(X_t)_{t\geq 0}$ is defined as follow:

$$X_t = e^{-t}X + \sqrt{1 - e^{-2t}}\,Y, \quad t \geq 0,$$

with X and Y i.i.d. standard Gaussian vectors in $\mathbb{R}^n$. The semigroup $(P_t)_{t\geq 0}$, associated to this process, acts on a class of smooth function $\mathcal{A}$ (due to the integrability of Gaussian densities, one can choose here for $\mathcal{A}$ the class of C^∞ functions whose derivatives are rapidly decreasing) and admits an explicit representation formula:

$$P_t f(x) = \int_{\mathbb{R}^n} f\big(xe^{-t} + \sqrt{1 - e^{-2t}}\,y\big) d\gamma_n(y), \quad x \in \mathbb{R}^n,\ t \geq 0$$

Its infinitesimal generator is given by

$$L = \Delta - x \cdot \nabla$$

Furthermore, γ_n is the invariant and reversible measure of $(P_t)_{t\geq 0}$. That is to say, for any function f and g belonging to $\mathcal{A}$,

$$\int_{\mathbb{R}^n} P_t f d\gamma_n = \int_{\mathbb{R}^n} f d\gamma_n \quad \text{and} \quad \int_{\mathbb{R}^n} f P_t g d\gamma_n = \int_{\mathbb{R}^n} g P_t f d\gamma_n.$$

Now, let us recall some properties satisfied by $(P_t)_{t\geq 0}$ which will be useful in the sequel.

Proposition 10.2.2 *The Ornstein–Uhlenbeck semigroup $(P_t)_{t\geq 0}$ satisfies the following properties*

- *$P_t(f)$ is a solution of the heat equation associated to L*

$$\text{i.e.}\quad \partial_t(P_t f) = P_t(Lf) = L(P_t f). \tag{10.4}$$

- *$(P_t)_{t\geq 0}$ is ergodic, that is to say, for $f \in \mathcal{A}$*

$$\lim_{t\to+\infty} P_t(f) = \int_{\mathbb{R}^n} f\, d\gamma_n = \mathbb{E}_{\gamma_n}[f] \tag{10.5}$$

- *$(P_t)_{t\geq 0}$ commutes with the gradient ∇. More precisely, for any function $f \in \mathcal{A}$,*

$$\nabla P_t(f) = e^{-t} P_t(\nabla f), \quad t \geq 0. \tag{10.6}$$

- *$(P_t)_{t\geq 0}$ is a contraction in $L^p(\gamma_n)$, for any function $f \in L^p(\gamma_n)$ and every $t \geq 0$,*

$$\|P_t(f)\|_p \leq \|f\|_p. \tag{10.7}$$

As it is exposed in [1], it is possible to give a dynamical representation of the variance of a function f along the semigroup $(P_t)_{t\geq 0}$:

$$\mathrm{Var}_{\gamma_n}(f) = 2\int_0^\infty \int_{\mathbb{R}^n} |\nabla P_s(f)|^2 d\gamma_n ds = 2\int_0^\infty e^{-2s} \int_{\mathbb{R}^n} |P_s(\nabla f)|^2 d\gamma_n ds \tag{10.8}$$

10.2.3 Gamma Calculus and Poincaré's Inequality

Let us introduce the fundamental operator Γ_2 and Γ from Bakry and Emery's Theory. Given an infinitesimal generator L set, for f and g, two smooth functions,

$$\Gamma(f,g) = \frac{1}{2}\big[L(fg) - fLg - Lfg\big] \quad \text{and} \quad \Gamma_2(f,g) = \frac{1}{2}\big[L\Gamma(f,g) - \Gamma(f,Lg) - \Gamma(Lf,g)\big]$$

In the case of the Ornstein–Uhlenbeck's infinitesimal generator $L = \Delta - x\cdot\nabla$, it is easily seen that

$$\Gamma(f) = |\nabla f|^2 \quad \text{and} \quad \Gamma_2(f) = \|\mathrm{Hess} f\|_2^2 + |\nabla f|^2 \tag{10.9}$$

where $\|\mathrm{Hess} f\|_2 = \left(\sum_{i,j=1}^n \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)^2\right)^{1/2}$ is the Hilbert–Schmidt norm of the tensor of the second derivatives of f.

Now, let us briefly recall how a relationship between Γ and Γ_2 can be used to give a elementary proof of Poincaré's inequality (10.3).

First, notice that the representation formula of the variance (10.8) can be expressed in terms of Γ:

$$\mathrm{Var}_{\gamma_n}(f) = 2\int_0^\infty \int_{\mathbb{R}^n} \Gamma(P_t f) d\gamma_n ds. \tag{10.10}$$

Then, observe that (10.9) implies the celebrated curvature-dimension criterion $CD(1, +\infty)$ (cf. [1])

$$\Gamma_2 \geq \Gamma. \tag{10.11}$$

Set $I(t) = \int_{\mathbb{R}^n} \Gamma(P_t f) d\gamma_n$. It is classical that

$$I'(t) = -2\int_{\mathbb{R}^n} \Gamma_2(P_t f) d\gamma_n, \quad t \geq 0$$

Thus, the inequality (10.11) leads to a differential inequality

$$\int_{\mathbb{R}^n} \Gamma_2(P_t f) d\gamma_n \geq \int_{\mathbb{R}^n} \Gamma(P_t f) d\gamma_n \Leftrightarrow 2I + I' \leq 0, \quad t \geq 0 \tag{10.12}$$

which can be easily integrated between s and t (with $0 \leq s \leq t$). That is

$$I(t)e^{2t} \leq I(s)e^{2s}.$$

It is now classical to let $s \to 0$ to easily recover Poincaré's inequality (10.3) for the measure γ_n. As we will see in the next section, we will show that a differential inequality of the form

$$I' \geq -2(I + \psi), \tag{10.13}$$

for some function ψ, can be used to obtain relevant bound (with respect to superconcentration phenomenon) on the variance of the function f (being fixed) by letting s fixed and $t \to +\infty$.

Remark Let us make few remarks.

(1) As it is proved in [1], the integrated curvature dimension inequality (10.12) is, in fact, equivalent to the Poincaré's inequality (10.3).
(2) As we will see in the next section, the inequality $I' \geq -2(I + \psi)$ is equivalent to an inverse, integrated, curvature dimension inequality which seems to be new. However, notice that the major difference between (10.12) and (10.13) is that

the first one holds for a large class of function whereas the second is only true for a particular function f (and ψ depends on f).

10.3 Inverse, Integrated, Curvature Inequality

In this section, we will use the methodology exposed in the preceding section to obtain variance bounds for a (fixed) function f satisfying an inverse, integrated, curvature inequality $IC_{\gamma_n}(1, \psi)$.

First, let us state a definition. We want to highlight the fact that this definition will be stated in a Gaussian framework $(\mathbb{R}^n, \Gamma, \gamma_n)$ with Γ associated to the infinitesimal generator $L = \Delta - x \cdot \nabla$ and the Ornstein–Uhlenbeck's semigroup $(P_t)_{t \geq 0}$. The next definition can be extended, mutatis mutandis, to fit the general framework of [1].

Definition 10.3.1 Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function. We say that f satisfy an inverse, integrated, curvature criterion with function $\psi : \mathbb{R}_+ \to \mathbb{R}$ if

$$\int_{\mathbb{R}^n} \Gamma_2(P_t f) d\gamma_n \leq \int_{\mathbb{R}^n} \Gamma(P_t f) d\gamma_n + \psi(t), \quad t \geq 0 \tag{10.14}$$

When the previous inequality is satisfied we denote it by $f \in IC_{\gamma_n}(1, \psi)$.

Remark

(1) Notice, again, that the inequality (10.14) holds, a priori, only for the function f.
(2) More generally, as it will be needed in the sequel, if μ is a Gaussian measure we will say that $f \in IC_\mu(1, \psi)$ if Eq. (10.14) is satisfied with μ instead of γ_n and with the operators Γ and Γ_2 associated to the Markov Triple $(\mathbb{R}^n, L, \mu)$.

Now, let us prove our main result Theorem 10.1.1.

Proof (of Theorem 10.1.1) Assume that $f \in IC_{\gamma_n}(1, \psi)$ (cf. Eq. (10.14)) holds. This is equivalent to the following differential inequality:

$$I' \geq -2(I + \psi), \tag{10.15}$$

where $I(t) = \int_{\mathbb{R}^n} |\nabla P_t f|^2 d\gamma_n$, $t \geq 0$. Set $I(t) = K(t)e^{-2t}$, inequality (10.15) becomes

$$K'(t) \geq -2e^{2t}\psi(t), \quad t \geq 0 \tag{10.16}$$

Now, integrate inequality (10.16) between s and t. That is

$$K(t) - K(s) \geq -2 \int_s^t e^{2u}\psi(u) du, \quad \text{for all} \quad 0 \leq s \leq t.$$

Then, let $t \to \infty$, this yields

$$K(s) \le \big[\lim_{t\to\infty} K(t)\big] + 2\int_s^{\infty} e^{2u}\psi(u)du, \quad s \ge 0,$$

To conclude, observe that

$$K(t) = I(t)e^{2t} \to_{t\to\infty} \left|\int_{\mathbb{R}^n} \nabla f d\gamma_n\right|^2$$

by ergodicity of $(P_t)_{t\ge 0}$. Finally, we have, for every $t \ge 0$,

$$I(t) = \int_{\mathbb{R}_n} \Gamma(P_t f)d\gamma_n \le e^{-2t}\left(\left|\int_{\mathbb{R}^n} \nabla f d\gamma_n\right|^2 + 2\int_t^{\infty} e^{2s}\psi(s)ds\right). \qquad (10.17)$$

It suffices to use the dynamical representation of the variance (10.8) with elementary calculus to end the proof. □

Remark This method of interpolation, between t and $+\infty$, has also been used in [13] in order to obtain Talagrand's inequality of higher order.

10.3.1 Another Variance Bound

As we will see in the last section, it is sometimes useful to restrict an $IC_\mu(1, \psi)$, for some probability measure μ, up to a time T in order to improve the dependance with respect to some parameter.

In other words, the setting is the following: assume that an $IC_\mu(1, \psi)$ holds and that we are able to produce some $T > 0$ such that the bound of $I(T)$ (given by Eq. (10.17)) is particularly nice (with respect to some parameter). Now, we have to bound the variance in a different manner in order to use the information on $I(T)$. To this task, we will prove the next proposition.

Proposition 10.3.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function smooth enough. Then, for any $T > 0$*

$$\mathrm{Var}_{\gamma_n}(f) \le \frac{2T I(0)}{1 - e^{-2T}}\left[\frac{1}{\log a} - \frac{1}{a\log a}\right]$$

with $a = \frac{I(0)}{I(T)}$ and $I(t) = \int_{\mathbb{R}^n} \Gamma(P_t f)d\gamma_n$.

Remark This proposition will be used to show that the Free Energy is superconcentrated for some Spin Glasses models. Although we stated the preceding Proposition 10.3.1 for the standard Gaussian measure γ_n, it will also hold (up to obvious renormalization) for μ the law of a centered Gaussian vector with covariance matrix M.

To prove the preceding theorem, we will need two further arguments.

First, we present an inequality due to Cordero-Erausquin and Ledoux [8]. The proof of this inequality rests on the fact that the Poincaré's inequality satisfied by γ_n implies an exponential decay of the variance along the semigroup $(P_t)_{t\geq 0}$.

Lemma 10.3.1 (Cordero-Erausquin–Ledoux) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function smooth enough. Then, for any $T > 0$, the following holds*

$$\mathrm{Var}_{\gamma_n}(f) \leq \frac{2}{1-e^{-2T}} \int_0^T I(t)dt \tag{10.18}$$

with $I(t) = \int_{\mathbb{R}^n} \Gamma(P_t f)d\gamma_n$.

Proof For the sake of completeness we give the proof of the preceding Lemma.

$$\begin{aligned}
\mathrm{Var}_{\gamma_n}(f) &= \mathbb{E}_{\gamma_n}[f^2] - \mathbb{E}_{\gamma_n}[(P_T f)^2] + \mathbb{E}_{\gamma_n}[(P_T f)^2] - \mathbb{E}_{\gamma_n}[P_T f]^2 \\
&= -\int_0^T \frac{d}{ds}\mathbb{E}_{\gamma_n}[(P_s f)^2]ds + \mathrm{Var}_{\gamma_n}(P_T f) \\
&\leq 2\int_0^T I(s)ds + e^{-2T}\mathrm{Var}_{\gamma_n}(f).
\end{aligned}$$

□

Secondly, we will use the fact that the infinitesimal generator $(-L)$ of the Ornstein–Uhlenbeck process $(X_t)_{t\geq 0}$ admits a (discrete) spectral decomposition. Then, denote by dE_λ the spectral resolution of $(-L)$. According to [1], this leads to a different representation of $t \mapsto I(t)$. With $f : \mathbb{R}^n \to \mathbb{R}$ being fixed, we have:

$$I(t) = \int_{\mathbb{R}^n} |\nabla P_t f|^2 d\gamma_n = \int_0^\infty \lambda e^{-2\lambda t} dE_\lambda(f), \quad t \geq 0$$

As it is proven in [2] (cf. Corollary 5.6), $t \mapsto I(t)$ satisfies, with the preceding representation, an Hölder-type inequality. That is to say, for every $T > 0$,

Lemma 10.3.2 (Baudoin–Wang)

$$I(s) \leq I(0)^{1-s/T} I(T)^{s/T}, \quad 0 \leq s \leq T \tag{10.19}$$

Now, we can prove Proposition 10.3.1 with the help of preceding Lemma.

Proof (of Proposition 10.3.1) First use Lemma 10.3.1 to get

$$\mathrm{Var}_{\gamma_n}(f) \leq \frac{2}{1-e^{-2T}} \int_0^T I(t)dt.$$

Then, use Lemma 10.3.2. This yields

$$\begin{aligned}\mathrm{Var}_{\gamma_n}(f) &\le \frac{2}{1-e^{-2T}}\int_0^T I(0)^{1-t/T} I(T)^{t/T}dt\\ &= \frac{2I(0)}{1-e^{-2T}}\int_0^T e^{-\frac{t}{T}\log a}dt\end{aligned}$$

where $a = \frac{I(0)}{I(T)} \ge 1$ and $I(t) = \int_{\mathbb{R}^n} \Gamma(P_t f)d\gamma_n$. Finally, elementary calculus ends the proof. □

10.4 Application in Spin Glasses's Theory

In the remaining of this section, we will show how Theorem 10.1.1 can be used to provide relevant bounds on the variance of $F_{n,\beta}$. We will focus on the REM and the SK Model. For the remaining of this note we will denote by f_β, for $\beta > 0$, the following function

$$f_\beta(x) = \frac{1}{\beta}\log\Big(\sum_{i=1}^n e^{\beta x_i}\Big), \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

10.4.1 Random Energy Model

In this section we will show how Theorem 10.1.1 is useful to obtain relevant bound on the variance of the Free Energy $F_{n,\beta}$ (with β close to 0) for the REM.

Proposition 10.4.1 *For any $\beta > 0$, $f_\beta \in IC_{\gamma_n}(1, \psi)$ with*

$$\psi(t) = 2\beta^2 e^{-2t} I(t)$$

where, let us recall it, $I(t) = \int_{\mathbb{R}^n} \Gamma(P_t f_\beta)d\gamma_n$ and Γ is the standard "carré du champ" operator.

We will need the following Lemma to prove the preceding Proposition.

Lemma 10.4.1 *Let $(u_i)_{i=1,\dots,n}$ be a family of functions, with $u_i : \mathbb{R}^n \to \mathbb{R}$ for any $i = 1, \dots, n$, satisfying the following condition*

$$\sum_{i=1}^n u_i^2(x) \le 1 \quad \textit{for all} \quad x \in \mathbb{R}^n$$

Then, for any function $v : \mathbb{R}^n \to \mathbb{R}_+$ and any probability measure μ, we have

$$\sum_{i=1}^{n}\left(\int_{\mathbb{R}^n} u_i(x)v(x)d\mu(x)\right)^2 \leq \left(\int_{\mathbb{R}^n} v d\mu\right)^2$$

Proof Consider the vector $U = (u_1 v, \ldots, u_n v) \in \mathbb{R}^n$ and recall that $|\cdot|$ stands for the Euclidean norm. Then, it holds

$$\left[\sum_{i=1}^{n}\left(\int_{\mathbb{R}^n} u_i(x)v(x)d\mu\right)^2\right]^{1/2} = \left|\int_{\mathbb{R}^n} U d\mu\right| \leq \int_{\mathbb{R}^n} |U| d\mu = \int_{\mathbb{R}^n}\left[\sum_{i=1}^{n} u_i^2(x)\right]^{1/2} v(x)d\mu$$
$$\leq \int_{\mathbb{R}^n} v(x)d\mu$$

where the first upper bound comes from Jensen's inequality. □

Now we turn to the proof of Proposition 10.4.1.

Proof (of Proposition 10.4.1) First, observe that the condition $IC_{\gamma_n}(1, \psi)$ is equivalent to

$$\int_{\mathbb{R}^n} \Gamma_2\big(P_t(f_\beta)\big)d\gamma_n \leq (1 + 2\beta^2 e^{-2t})\int_{\mathbb{R}^n} \Gamma\Big(P_t(f_\beta)\Big)d\gamma_n, \quad t \geq 0.$$

That is (since $\Gamma_2(f) = \|\mathrm{Hess} f\|_2^2 + |\nabla f|^2$ and $\Gamma(f) = |\nabla f|^2$)

$$\int_{\mathbb{R}_n} \|\mathrm{Hess} P_t(f_\beta)\|_2^2 d\gamma_n \leq 2\beta^2 e^{-2t}\int_{\mathbb{R}^n} |\nabla P_t(f_\beta)|^2 d\gamma_n, \quad t \geq 0. \tag{10.20}$$

Now, observe that, pointwise, Eq.(10.20) is equivalent to (thanks to the commutation property between ∇ and $(P_t)_{t\geq 0}$)

$$\sum_{i,j=1}^{n}[P_t(\partial_{ij}^2 f_\beta)]^2 \leq 2\beta^2\sum_{i=1}^{n}[P_t(\partial_i f_\beta)]^2, \quad \forall t \geq 0$$

Elementary calculus yields, for every $i = 1, \ldots, n$, and every $\beta > 0$,

$$\partial_i f_\beta = \frac{e^{\beta x_i}}{\sum_{k=1}^{n} e^{\beta x_k}}$$

and, for every $j = 1, \ldots, n$,

$$\partial_j\partial_i f_\beta = \beta(\partial_i f_\beta \delta_{ij} - \partial_i f_\beta \partial_j f_\beta).$$

Thus, for every $t \geq 0$,

$$\sum_{i,j=1}^{n} [P_t(\partial_{ij}^2 f_\beta)]^2 = \beta^2 \sum_{i=1}^{n} \left[P_t(\partial_i f_\beta)\right]^2 - 2\beta \sum_{i=1}^{n} P_t(\partial_i f_\beta) P_t\left[(\partial_i f_\beta)^2\right] + \beta^2 \sum_{i,j=1}^{n} \left[P_t(\partial_i f_\beta \partial_j f_\beta)\right]^2.$$

First ignore the crossed terms (which are always non positive), then apply Lemma 10.4.1 to the third term.

Indeed, let $i \in \{1, \ldots, n\}$ be fixed and set $u_j = \partial_j f_\beta$ and $v = \partial_i f_\beta$. Thus, Lemma 10.4.1 implies

$$\sum_{j=1}^{n} \left[P_t(\partial_i f_\beta \partial_j f_\beta)\right]^2 \leq P_t^2(\partial_i f_\beta).$$

This inequality finally yields,

$$\sum_{i,j=1}^{n} [P_t(\partial_{ij}^2 f_\beta)]^2 \leq \beta^2 \sum_{i=1}^{n} \left[P_t(\partial_i f_\beta)\right]^2 + \beta^2 \sum_{i,j=1}^{n} \left[P_t(\partial_i f_\beta \partial_j f_\beta)\right]^2 \leq 2\beta^2 \sum_{i=1}^{n} \left[P_t(\partial_i f_\beta)\right]^2.$$

□

Now, the criterion $IC_{\gamma_n}(1, \psi)$ can be used gives to provide relevant bound on the variance of $F_{n,\beta}$ as stated in Proposition 10.1.2.

Proof (of Proposition 10.1.2) As mentioned earlier, the proof will be done for the standard Gaussian measure on $\mathbb{R}^n$ and then it will be enough to perform a change of variable. As it will be useful in the sequel, observe that (by symmetry) the following holds

$$\int_{\mathbb{R}^n} \partial_i f_\beta d\gamma_n = \frac{1}{n}, \quad \forall i = 1, \ldots, n.$$

Now, let $\beta > 0$ and use Theorem 10.1.1 which implies that

$$\mathrm{Var}_{\gamma_n}(F_{n,\beta}) \leq \frac{1}{n} + 4\beta^2 \int_0^\infty e^{-2s}(1 - e^{-2s}) \sum_{i=1}^{n} \int_{\mathbb{R}^n} P_s^2(\partial_i f_\beta) d\gamma_n ds \tag{10.21}$$

where we used Fubini's Theorem and the commutation property between ∇ and P_s.

For the first bound, when $\beta \in \left(0, \frac{\sqrt{2}}{2}\right)$, it is possible to rewrite (thanks to the dynamical representation of the variance (10.21)) the integral in the right hand side as

$$2\beta^2 \mathrm{Var}_{\gamma_n}(F_{n,\beta}) - 4\beta^2 \int_0^\infty e^{-4s} \sum_{i=1}^{n} \int_{\mathbb{R}^n} P_s^2(\partial_i f_\beta) d\gamma_n ds \tag{10.22}$$

Furthermore, by Jensen's inequality and the invariance of $(P_t)_{t\geq 0}$ with respect to γ_n, we have

$$\int_{\mathbb{R}^n} P_s^2(\partial_i f_\beta)d\gamma_n \geq \left(\int_{\mathbb{R}^n} P_s(\partial_i f_\beta)d\gamma_n\right)^2 = \frac{1}{n^2}, \quad \forall i = 1, \ldots, n, \quad \forall s > 0$$

Thus, $\mathrm{Var}_{\gamma_n}(F_{n,\beta}) \leq \left(\frac{1-\beta^2}{1-2\beta^2}\right)\frac{1}{n}$.

To conclude, as announced, it is enough to substitute n by 2^n and β by $\sqrt{n}\beta$ to get the result. □

Remark Incidentally, the preceding proof can be used to get a lower bound on the variance of the Free energy. More precisely, it is possible to deduce from (10.21) and (10.22) the following lower bound

$$\mathrm{Var}_{\gamma_{2^n}}(F_{n,\beta}) \geq \frac{n}{2^n}\frac{(1-n\beta^2)}{(1-2\beta^2 n)}, \quad \text{for} \quad \beta > \frac{1}{\sqrt{2n}}$$

10.4.2 SK Model

In this section we show how some work of Chatterjee (from [7]) can be rewritten in term of an inverse, integrated, curvature criterion. Then, it allows us to easily recover a bound, obtained by Talagrand (cf. [11, 12]), on the variance of the Free Energy for the SK model at high temperature.

First, we need to express the Γ and Γ_2 operator when γ_n is replaced by μ the law of a centered Gaussian vector, in $\mathbb{R}^n$, with covariance matrix M.

Let X be a random Gaussian vector with $\mathcal{L}(X) = \mu$ and consider Y an independent copy of X. It is then possible to define the generalized Ornstein–Uhlenbeck process, which we will still denote by $(X_t)_{t\geq 0}$, as follow

$$X_t = e^{-t}X + \sqrt{1-e^{-2t}}Y, \quad t \geq 0$$

Similarly, we also denote by $(P_t)_{t\geq 0}$ the associated semigroup. Then, it is known (cf. [7, 14, 16]) that, for any smooth function $f : \mathbb{R}^n \to \mathbb{R}$,

$$I(t) = \int_{\mathbb{R}^n} \Gamma(P_t f)d\mu = 2\int_{\mathbb{R}^n} e^{-2t}\sum_{i,j} M_{ij}(\partial_i f)P_t(\partial_j f)d\mu, \quad t \geq 0$$

As we will see latter, it will be more convenient to work with

$$I_r(t) = 2\int_{\mathbb{R}^n} e^{-2t}\sum_{i,j}(M_{ij})^r(\partial_i f)P_t(\partial_j f)d\mu, \quad t \geq 0$$

where r is a positive integer. In the rest of this section, we choose $f = f_\beta$.

Proposition 10.4.2 (Chatterjee) *Assume that $M_{ij} \geq 0$ for all $(i, j) \in \{1, \ldots, n\}^2$. Then, for any $t \geq 0$, the following holds*

$$I_r'(t) \geq -2\big[I_r(t) + 2\beta^2 e^{-2t} J_{r+1}(t)\big] \tag{10.23}$$

with $J_r(t) = e^{2t} I_r(t)$.

Remark

(1) In [7], Chatterjee proved that $J_r'(t) \geq -4\beta^2 e^{-2t} J_{r+1}(t)$ for any $r \in \mathbb{N}^*$. The proof is similar the proof of Lemma 10.4.1 with the additional use of Hölder's inequality.
(2) In particular, when $r = 1$, Chatterjee's proposition amounts of saying that

$$f_\beta \in IC_\mu(1, \psi)$$

with $\psi(t) = 2\beta^2 e^{-2t} J_2(t)$. Unfortunately, it remains hard to upper bound this quantity by something relevant.

As observed in the preceding remark, the inverse, integrated, curvature criterion can not be used in the present form. However, it is possible to recycle the arguments of Sect. 10.3. That is, use l times, with $l \in \mathbb{N}$, the fundamental Theorem of analysis (on $t \mapsto I_r(t)$) together with the inequality (10.23) and let $l \to +\infty$. This leads to a useful bound on the function $t \mapsto I_r(t)$ for any $r \in \mathbb{N}^*$.

Theorem 10.4.1 (Chatterjee) *Assume that $M_{ij} \geq 0$ for all $(i, j) \in \{1, \ldots, n\}^2$. Then, for any $t \geq 0$, the following holds*

$$I_r(t) \leq e^{-2t} \sum_{i,j=1}^{n} (M_{ij})^r e^{2\beta^2 e^{-2t} M_{ij}} \nu_i \nu_j, \quad \forall r \geq 1 \tag{10.24}$$

where $\nu_i = \int_{\mathbb{R}^n} \partial_i f_\beta d\mu$ for all $i = 1, \ldots, n$.

Remark When $r = 1$, the main step of Chatterjee's proof is equivalent to show that $f_\beta \in IC_\mu(1, \psi)$ with $\psi(t) = 2\beta^2 e^{-2t} \sum_{i,j=1} M_{ij} e^{2\beta^2 e^{-2t} M_{ij}} \nu_i \nu_j$. The proof of this result can be found in [7, pp. 108–110] .

Unfortunately, the repeated use of the differential inequality (10.23) degrades the upper bound on $t \mapsto I_r(t)$. As we will briefly see in the next subsection, Chatterjee used Eq. (10.24) only for a fixed $T > 0$ (large enough). We show, in the next Proposition, that this bound (for $r = 1$) is still relevant to recover some work of Talagrand on the variance of $F_{n,\beta}$, with small β, for the SK model (cf. [11, 12]).

Now, let us prove Proposition 10.1.1.

Proof (of Proposition 10.1.1) First we show that inequality (10.24) leads to a general upper bound on the variance of $F_{n,\beta}$ which might be of independent interest. Then, we choose M to be the covariance structure of the SK model and proved inequality (10.2).

When $r = 1$, Eq. (10.24) combined with Eq. (10.10) implies that, for any $\beta > 0$,

$$\begin{aligned}\mathrm{Var}_\mu(F_{n,\beta}) &\leq 2\int_0^\infty e^{-2t}\sum_{i,j=1}^n M_{ij}e^{2\beta^2 e^{-2t}M_{ij}}\nu_i\nu_j dt\\ &\leq \frac{1}{2\beta^2}\sum_{i,j=1}^n e^{2\beta^2 M_{ij}}\nu_i\nu_j\end{aligned}$$

Following Chatterjee (cf. [7]), choose M to be the covariance structure of the SK model. That is,

$$M_{\sigma\sigma'} = \left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \sigma_i\sigma_i'\right)^2, \quad \forall \sigma, \sigma' \in \{-1, 1\}^n.$$

Besides, observe (by symmetry) that, for each $\sigma \in \{-1, 1\}^n$,

$$\nu_\sigma = \mathbb{E}_\mu\big[\partial_\sigma F_{n,\beta}\big] = \frac{1}{2^n}.$$

Thus,

$$\mathrm{Var}_\mu(F_{n,\beta}) \leq \frac{1}{2\beta^2}\mathbb{E}_{\sigma,\sigma'}\Big[e^{2\beta^2\left(\frac{1}{\sqrt{n}}\sigma_i\sigma_i'\right)^2}\Big]$$

where $\mathbb{E}_{\sigma'\sigma}$ stands for the expectation under the product measure induced by the Rademacher random variables $\sigma_i, \sigma_i', \ i = 1, \ldots, n$.

Finally, if $\beta \in \left(0, \frac{1}{2}\right)$ we have $\mathbb{E}_{\sigma,\sigma'}\Big[e^{2\beta^2\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n\sigma_i\sigma_i'\right)^2}\Big] = C(\beta)$. Indeed, observe first that $\sum_{i=1}^n \sigma_i\sigma_i'$ has the same distribution as $\sum_{i=1}^n \sigma_i$. Then, it is enough to use Hoeffding's inequality (cf.[4]), which gives the following deviation inequality

$$\mathbb{P}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \sigma_i > t\right) \leq e^{-t^2/2} \quad t \geq 0,$$

to conclude. □

10.4.3 Improvements of Variance Bounds with Respect to the Parameter β

Let us collect some results of Chatterjee and briefly explain how Proposition 10.3.1 can be used to improve the dependence of the variance bounds with respect to β. However, the dependance in n will be worse.

Chatterjee used, in [7], a Theorem of Bernstein about completely monotone function. As far as we are concerned, the spectral framework exposed in Sect. 10.3 seems to be more natural to work with and provides equivalent results.

The arguments, in order to improve the dependance in β, can be summarize as follow: choose T such that $I(T)$ can be bounded by a relevant quantity and apply Proposition 10.3.1.

Proposition 10.4.3 (Chatterjee) *In the SK model the following holds*

$$\mathrm{Var}_\mu(F_{n,\beta}) \leq \frac{C_1 n \log(2 + C_2\beta)}{\log n}, \quad \forall\beta > 0$$

with $C_1, C_2 > 0$ two numerical constants.

Remark Here $T > 0$ is chosen such that

$$\mathbb{E}_{\sigma,\sigma'}\left[M_{\sigma\sigma'} e^{2\beta^2 e^{-2T} M_{\sigma\sigma'}}\right] = C_\beta, \quad \forall\beta > 0$$

where $M_{\sigma\sigma'} = \left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \sigma_i\sigma_i'\right)^2$ and $C_\beta > 0$ is a constant that does not depend on n. That is $T = \frac{1}{2}\log\left(\frac{2\beta^2}{\gamma}\right)$ for some sufficiently small constant $\gamma > 0$ (cf. [7]).

Proposition 10.4.4 (Chatterjee) *In the REM, the following holds for $\beta > 2\sqrt{\log 2}$,*

$$\mathrm{Var}_\mu(F_{n,\beta}) \leq C_\beta$$

where $C_\beta > 0$ is a constant that does not depend on n.

Remark Here T is chosen as $T = \frac{1}{2}\log(2\beta^2)$ so that $I(T) \leq \frac{n}{2^n}e^{-2T}e^n$ and the upper bound is relevant in the low temperature regime (cf. [6, 7]). Again, notice the difference of renormalization with Proposition 10.1.2 (one has to replace the number of random variables n by 2^n and the i.i.d. standard Gaussian random variables $(X_i)_{i=1,\ldots,2^n}$ by $\sqrt{n}X_i$ in the Proposition). In [7], Chatterjee also proved that the upper bound is tight.

In fact, it also possible to use hypercontractive arguments instead of Theorem 10.4.1 to achieve the upper bound of Proposition 10.4.4. Indeed, one can use the inequality (10.21) together with hypercontractive estimates of $(P_t)_{t\geq 0}$ (cf. [7, 8, 15, 16]). More precisely, we have

$$\|P_s(\partial_i f_\beta)\|_2^2 \leq \|\partial_i f_\beta\|_{1+e^{-2s}}^2, \quad \forall i = 1,\ldots,n, \quad \forall s > 0$$

It is then standard, cf. Section 4 in [16] for instance, to prove that

$$\int_0^\infty e^{-2s}(1 - e^{-2s})\|\partial_i f_\beta\|_{1+e^{-2s}}^2 ds \leq \frac{C\|\partial_i f_\beta\|_2^2}{\left[1 + \log\frac{\|\partial_i f_\beta\|_2}{\|\partial_i f_\beta\|_1}\right]^2}$$

where $C > 0$ is a numerical constant. Then, it is elementary to conclude. Notice that such estimates are already implicit in the celebrated L^1/L^2 Talagrand's inequality (presented in [7, 8] for instance), which one can also be directly used to recover the content of Proposition 10.4.4.

Acknowledgements I thank M. Ledoux for fruitful discussions on this topic. I also warmly thank the referee for helpful comments in improving the exposition and the simplification of the proof of Lemma 10.4.1.

References

1. D. Bakry, I. Gentil, M. Ledoux, *Analysis and Geometry of Markov Diffusion Operators*. Grundlehren der Mathematischen Wissenschaften, vol. 348 (Springer, Berlin, 2014)
2. F. Baudoin, J. Wang, Curvature dimension inequalities and subelliptic heat kernel gradient bounds on contact manifolds. Potential Anal. **40**, 163–193 (2014)
3. S. Boucheron, M. Thomas, Concentration inequalities for order statistics. Electron. Commun. Probab. **17**, 1–12 (2012)
4. S. Boucheron, G. Lugosi, P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independance* (Oxford University Press, Oxford, 2013)
5. A. Bovier, *Gaussian Processes on Trees: From Spin Glasses to Branching Brownian Motion*. Cambridge Studies in Advanced Mathematics, vol. 163 (Cambridge University Press, Cambridge, 2016)
6. A. Bovier, I. Kurkova, M. Löwe, Fluctuations of the free energy in the REM and the p-spin SK models. Ann. Probab. **30**(2), 605–651 (2002)
7. S. Chatterjee, *Superconcentration and Related Topics* (Springer, Berlin, 2014)
8. D. Cordero-Erausquin, M. Ledoux, Hypercontractive measures, Talagrand's inequality, and influences, in *Geometric Aspects of Functional Analysis*. Lectures Notes in Mathematics, vol. 2050 (Springer, Berlin, 2012), pp.169–189
9. M. Ledoux, The geometry of Markov diffusions operators. Ann. Fac. Sci. Toulouse Math. 6 **9**(2), 305–366 (2000)
10. M. Ledoux, *The Concentration of Measure Phenomenon*. Mathematical Surveys and Monographs, vol. 89 (American Mathematical Society, Providence, 2001)
11. M. Talagrand, *Mean Field Models for Spin Glasses. Volume I*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics (Springer, Berlin, 2011)
12. M. Talagrand, *Mean Field Models for Spin Glasses. Volume II: Advanced Replica-Symmetry and Low Temperature*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, vol. 55 (Springer, Heidelberg, 2011)
13. K. Tanguy, Talagrand's inequality at higher order and application to Boolean analysis (2019). https://arxiv.org/pdf/1801.08931.pdf
14. K. Tanguy, Some superconcentration inequalities for extrema of stationary Gaussian processes. Stat. Probab. Lett. **106**, 239–246 (2015)
15. K. Tanguy, Non asymptotic variance bounds and deviation inequalities by optimal transport (2017). Preprint, http://arxiv.org/abs/1708.08620
16. K. Tanguy, *Quelques inégalités de concentration: théorie et applications (in French)*. Ph.D. thesis, Institute of Mathematics of Toulouse, 2017
17. P. Valettas, On the tightness of Gaussian concentration for convex functions. J. Anal. Math. (to appear)