
ETUDE DE MAXIMA DE FAMILLES GAUSSIENNES

par

Kevin Tanguy, sous la direction de M.Ledoux

Résumé. —

Le sujet de ce mémoire étudie plusieurs aspects dans l'étude de maxima de familles gaussiennes. Nous nous appuyons sur la monographie de S. Chatterjee "Superconcentration and related topics".

Nous parlerons essentiellement du phénomène de superconcentration. Nous rassemblerons, dans un premier temps, tout un ensemble de résultats analytiques qui nous permettra de développer les outils nécessaires à l'étude du phénomène de superconcentration. Dans un second temps, nous illustrerons, par différents exemples, cette propriété de familles de variables aléatoires gaussiennes.

Table des matières

Partie I. Notations	4
Partie II. Introduction	6
Partie III. Outils analytiques	8
1. Semi-groupe de Markov.....	8
2. Inégalités de Poincaré.....	13
3. Inégalité fonctionnelle et hypercontractivité.....	21
4. Méthode de Talagrand.....	27
5. Superconcentration et chaos.....	31
Partie IV. Modèles	36
6. Verres de spin.....	36
7. NK landscape.....	55
8. Modèle de percolation.....	57
9. Polymères dirigés en milieu aléatoire.....	62
Partie V. Annexe	71
10. Résultats sur les variables aléatoires gaussiennes.....	71
11. Théorème d'hypercontractivité.....	79
12. Inégalité de Hoeffding.....	80
Références.....	83

PARTIE I NOTATIONS

Nous présentons les notations que nous avons adopté et que nous avons essayé de maintenir cohérentes tout au long de ce mémoire.

En général, C désigne une constante positive qui peut souvent varier d'une ligne à une autre.

$|\cdot|$ désignera la norme Euclidienne sur \mathbb{R}^n . La semi-norme de Lipschitz $\|f\|_{Lip}$ d'une fonction de \mathbb{R}^n à valeurs réelles est donnée par :

$$\|f\|_{Lip} := \sup \left\{ \frac{|f(s) - f(t)|}{|s - t|} ; s \neq t, s, t \in \mathbb{R}^n \right\}.$$

Sauf mention contraire, $\langle \cdot \rangle$ désignera le produit scalaire Euclidien usuel. Cependant, nous utiliserons parfois le symbole \cdot pour ne pas alourdir les notations.

La mesure standard gaussienne sur \mathbb{R}^n sera notée par γ ; il s'agit de la mesure de probabilité sur \mathbb{R}^n de densité :

$$\frac{1}{(2\pi)^{n/2}} \exp(-|x|^2/2).$$

Nous noterons par X, Y, Z, \dots les variables aléatoires gaussiennes, que nous supposerons toujours centrées, et par B_t le mouvement Brownien standard de \mathbb{R}^n .

Γ désignera la matrice de covariance d'un vecteur gaussien. Nous supposerons toujours que Γ est semi-définie positive, nous noterons par M la matrice telle que ${}^t M M = \Gamma$.

Etant donné un vecteur gaussien $\mathbf{X} = (X_1, \dots, X_n)$, nous noterons par $\mathbf{M} := \max_{i=1, \dots, n} X_i$.

Pour simplifier les notations, nous posons $c_t := e^{-t}$ et $d_t := \sqrt{1 - e^{-2t}}$. Nous préciserons par moment par rapport à quelle mesure nous intégrons, par exemple $\mathbb{E}_\gamma[f]$. Lorsque que le cadre sera clair, nous omettrons cet indice pour alléger les notations.

H_k fera référence au k -ième polynôme normalisé d'Hermite.

P_t désignera un semi-groupe, L le générateur infinitésimal qui lui est associé. Lorsque nous parlerons du semi-groupe d'Ornstein-Uhlenbeck (OU), nous noterons le domaine de L par \mathcal{A} , où

$$\mathcal{A} := \left\{ f \in C^2(\mathbb{R}^n, \mathbb{R}), \text{ à croissance au plus polynômiale} \right\}.$$

\mathcal{E} fera référence à l'énergie de Dirichlet.

Dans le réseau \mathbb{Z}^d , $\mathcal{P}_{x,y}$ désignera un chemin reliant les points $x \in \mathbb{Z}^d$ et $y \in \mathbb{Z}^d$.

PARTIE II

INTRODUCTION

Les inégalités de concentration fournissent des bornes efficaces à horizon fini qui permettent de quantifier les phénomènes asymptotiques. Elles sont le plus souvent robustes, et ont démontré leur efficacité dans l'examen de nombreux modèles, en analyse géométrique et fonctionnelle, en physique statistique, en probabilités ou encore en statistiques.

Les inégalités de concentration traditionnelles concernent toutefois des classes générales de fonctionnelles, sans nécessairement prendre en compte les spécificités de certains exemples. Plusieurs modèles ont en effet récemment mis en évidence des comportements plus précis de la variance qui relèvent d'un phénomène de superconcentration précisant la concentration usuelle.

Dans ce mémoire nous nous attacherons à l'étude de ce phénomène. Le plus simple, pour fixer les idées sur la signification du mot superconcentration, est de l'illustrer par un exemple facile.

Considérons un « baby » modèle. Soit X une variable aléatoire réelle gaussienne standard, prenons alors X_1, \dots, X_n n copies indépendantes. Nous voulons nous intéresser au maximum de cette famille, posons alors $\mathbf{M} := \max_{i=1, \dots, n} X_i$.

Il est bien connu que $\mathbb{E}(\mathbf{M})$ se comporte en $\sqrt{\log n}$. De plus, l'inégalité de Poincaré, vérifiée par la mesure gaussienne sur \mathbb{R}^n , nous fournit l'inégalité suivante :

$$(1) \quad \text{Var}(\mathbf{M}) \leq 1.$$

Bien entendu, nous rappellerons ultérieurement ce qu'est une inégalité de Poincaré, de quelle manière on peut l'obtenir et enfin, comment nous permet-elle d'avoir l'inégalité (1).

Néanmoins, il faut retenir qu'une étude plus fine de la variance du maximum de nos gaussiennes nous assure que :

$$\text{Var}(\mathbf{M}) \leq \frac{C}{\log n}$$

où $C > 0$.

Ainsi, la théorie classique de la concentration nous donne un résultat sous-optimal. Notre variable aléatoire se concentre autour de sa moyenne bien plus rapidement que l'on pouvait le croire. Bien qu'étant un outil majeur de la théorie de la concentration, l'inégalité de Poincaré n'est pas assez précise pour nous donner le véritable comportement de la variance dans cet exemple. On assiste là à un phénomène que l'on peut qualifier de superconcentration. Nous verrons par la suite une définition précise de cette notion et nous développerons des outils permettant d'établir ce phénomène. Enfin, nous étudierons différents modèles, autrement plus compliqués que celui que nous venons d'énoncer, pour montrer qu'ils sont à même de vérifier un phénomène de superconcentration.

Nous allons aborder l'inégalité de Poincaré de deux points de vue différents pour obtenir de la superconcentration. Le premier consiste à utiliser l'outil de l'hypercontractivité pour améliorer l'inégalité de Poincaré : il s'agit de l'inégalité de Talagrand. Néanmoins cette approche ne fonctionnera pas pour tous les modèles considérés. C'est pourquoi, dans un deuxième temps, nous étudierons l'inégalité de Poincaré de façon spectrale. Plus précisément, nous exploiterons le fait que le générateur infinitésimal du semi-groupe d'Ornstein-Uhlenbeck (nous utiliserons parfois l'abréviation OU) admet pour fonctions propres une base hilbertienne de $L^2(\gamma^n)$: les polynômes d'Hermite.

Enfin, nous constaterons que nous pouvons améliorer l'un de nos résultats en combinant une méthode d'interpolation avec un Théorème de Bernstein sur les fonctions complètement monotones.

Nous illustrerons ces deux approches, via la décomposition spectrale et l'inégalité de Talagrand, en étudiant différents modèles : celui de Sherrington et Kirkpatrick sur les verres de spin, le modèle « NK landscape » de génétique et enfin un modèle de percolation ainsi qu'un modèle de polymères dirigés en milieu aléatoire.

PARTIE III OUTILS ANALYTIQUES

1. Semi-groupe de Markov

Dans cette section, nous allons faire des rappels sur les semi-groupes de Markov et nous montrerons de quelles manières ces derniers nous permettent d'établir l'inégalité de Poincaré pour la mesure gaussienne. Nous parlerons également du lien existant entre la décomposition spectrale d'un générateur infinitésimal et de l'inégalité de Poincaré. Nous aborderons ensuite les inégalités de Sobolev logarithmiques ainsi que la notion d'hypercontractivité. L'hypercontractivité est une des clés qui nous permettra d'obtenir, par la suite, des résultats de superconcentration.

1.1. Définition et propriétés. — Nous rappelons, par soucis de consistance, la définition d'un processus de Markov. Pour plus de détails voir le livre de D.Revuz et M.Yor (18).

Définition 1. — Soit (Ω, \mathcal{F}) un espace mesurable. Un noyau N sur Ω est une application de $\Omega \times \mathcal{F}$ dans $\mathbb{R}_+ \cup \{+\infty\}$ telle que

1. pour tout $\omega \in \Omega$, l'application $A \mapsto N(\omega, A)$ est une mesure positive sur \mathcal{F} ;
2. pour tout $A \in \mathcal{F}$, l'application $x \mapsto N(\omega, A)$ est \mathcal{F} -mesurable.

Remarque 1. — Un noyau P est appelé probabilité de transition si $P(\omega, \Omega) = 1$ pour tout $\omega \in \Omega$.

Définition 2. — Une fonction de transition sur (Ω, \mathcal{F}) est une famille $P_{s,t}$, $0 \leq s < t$ de probabilités de transition sur (Ω, \mathcal{F}) telle que pour n'importe quel triplet de nombres réels $s < t < v$, nous avons :

$$\int P_{s,t}(\omega, dy) P_{t,v}(y, A) = P_{s,v}(\omega, A),$$

pour tout $\omega \in \Omega$ et $A \in \mathcal{F}$.

Définition 3. — Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ un espace de probabilité filtré. Un processus adapté X est un processus de Markov par rapport à la filtration \mathcal{F}_t , de fonction de transition $P_{s,t}$ si pour toute fonction f mesurable positive et n'importe quel couple (s, t) tel que $s < t$,

$$\mathbb{E}[f(X_t) | \mathcal{F}_s] = P_{s,t}(X_s) \quad \mathbb{P} - p.s.$$

Introduisons quelques notions de bases sur les semi-groupes de Markov.

On suppose que $(X_t)_{t \geq 0}$ est un processus de Markov à valeurs dans un espace d'état abstrait. Ce processus de Markov permet de définir naturellement un semi-groupe d'opérateurs $(P_t)_{t \geq 0}$, à savoir $P_s \circ P_t = P_{s+t}$ et $P_0 = Id$, agissant sur un certain ensemble de fonctions :

$$P_t f(x) = \mathbb{E}(f(X_t) | X_0 = x),$$

en supposant que l'expression ci-dessus soit bien définie. Par la suite nous utiliserons un certain nombre de résultats essentiels sur un semi-groupe bien spécifique : celui d'Ornstein-Uhlenbeck.

Bien que nous allons énoncer des résultats généraux sur des processus de Markov vérifiant certaines hypothèses, l'exemple qui nous servira de fil conducteur est celui du processus d'Ornstein-Uhlenbeck. C'est pourquoi nous faisons l'hypothèse supplémentaire : les fonctions f que nous considérerons par la suite seront de classe C^2 à croissance au plus polynomiale. Ceci rendra licite les calculs que nous feront. Nous noterons par \mathcal{A} cet ensemble de fonctions.

La propriété de Markov de notre processus nous assure que $(P_t)_{t \geq 0}$ est bien un semi-groupe.

Pour une classe de fonction convenablement choisie, de telle sorte que la limite suivante ait un sens, on définit le générateur infinitésimal associé à notre semi-groupe par :

$$Lf := \lim_{t \rightarrow 0} \frac{P_t f - f}{t} = \partial_t P_t |_{t=0}.$$

Celui-ci permet de décrire le comportement local de notre processus. L'équation de la chaleur associée à notre semi-groupe P_t est :

$$\partial_t P_t = LP_t.$$

On obtient facilement en étudiant cette équation que :

$$P_t = e^{tL} = \sum_{k=0}^{\infty} \frac{t^k}{k!} L^k.$$

On dit qu'un processus de Markov X_t possède une mesure de probabilité invariante μ : si pour toute fonctions f sur lesquelles le semi-groupe $(P_t)_t$ peut agir, on a $\mu(P_t f) = \mu(f)$. Cette mesure permet de définir un espace L^2 et un produit scalaire :

$$\langle f, g \rangle_{\mu} := \int f g d\mu.$$

Cela permet également de définir une forme bilinéaire \mathcal{E} très importante que nous retrouverons par la suite. \mathcal{E} est appelée énergie de Dirichlet du semi-groupe de Markov P_t .

$$\mathcal{E}(f, g) := - \langle f, Lg \rangle = - \int f Lg d\mu.$$

Nous ferons l'hypothèse que notre processus de Markov X_t admet une telle mesure invariante μ . On supposera, enfin, que notre processus de Markov est réversible, ainsi L est un opérateur auto-adjoint et \mathcal{E} est symétrique. Autrement dit, pour toute fonction f et g on a :

$$\langle P_t f, g \rangle_\mu = \langle f, P_t g \rangle_\mu.$$

Remarque 2. — *Il convient de noter que les hypothèses que nous imposons ne sont pas gratuites. En général, il est difficile d'exhiber une mesure invariante. Le générateur infinitésimal peut ne pas être défini et si jamais il existait, il ne serait pas forcément autoadjoint.*

Nous allons établir un Lemme important qui permet de relier la forme bilinéaire de covariance à celle de Dirichlet.

Lemme 1 (Lemme de covariance). — *Pour tout $f, g \in L^2(\mu)$,*

$$Cov_\mu(f, g) = \int_0^\infty \mathcal{E}(f, P_t g) dt,$$

pourvu que l'on puisse dériver en t sous le signe intégrale et que l'équation de la chaleur $\partial_t P_t g = L P_t g$ soit vérifiée.

Démonstration. — On rappelle que $P_t g$ converge vers $\mathbb{E}_\mu(g)$ dans $L^2(\mu)$ lorsque $t \rightarrow \infty$ (il suffit d'utiliser le Théorème de convergence dominé). Ainsi en utilisant nos hypothèses on obtient :

$$\begin{aligned} Cov_\mu(f, g) &= \langle f, g \rangle - \lim_{t \rightarrow \infty} \langle f, P_t g \rangle \\ &= - \int_0^\infty \partial_t \langle f, P_t g \rangle dt \\ &= - \int_0^\infty \langle f, \partial_t P_t g \rangle dt \\ &= - \int_0^\infty \langle f, L P_t g \rangle dt = \int_0^\infty \mathcal{E}(f, P_t g) dt. \end{aligned}$$

□

1.2. Semi-groupe d’Ornstein-Uhlenbeck. — Voici à présent le semi-groupe que nous rencontrerons tout au long de ce mémoire. Il est d’un intérêt tout particulier car il admet la mesure gaussienne γ pour mesure invariante.

Par souci de consistance nous rappelons quelques résultats sur le semi-groupe de Ornstein Uhlenbeck. Nous nous inspirerons grandement du livre (1). Nous ne donnerons que la définition et les propriétés de ce semi-groupe sur la droite réelle. Bien entendu tout ceci se généralise à \mathbb{R}^n .

On munit également la droite réelle de la mesure Gaussienne standard : $d\gamma(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx$.

Définition 4. — Nous définissons le processus d’Ornstein-Uhlenbeck X_t comme solution de l’équation différentielle stochastique suivante :

$$dX_t = -X_t dt + \sqrt{2}dB_t,$$

où B_t est un mouvement Brownien standard.

Remarque 3. — Nous pouvons également voir le semi-groupe d’Ornstein-Uhlenbeck sur \mathbb{R} comme la famille d’opérateurs agissant sur les fonctions f de \mathcal{A} par :

$$P_t(f)(x) := \int_{\mathbb{R}} f(e^{-t}x + \sqrt{1 - e^{-2t}}y) d\gamma(y).$$

Ceci nous fournit une représentation intégrale de l’action du semi-groupe d’Ornstein-Uhlenbeck sur une fonction. On peut reformuler ceci plus simplement en adoptant les notations suivantes : soit X une variable gaussienne centrée réduite $\mathcal{N}(0, 1)$, on pose $c_t := e^{-t}$ et $d_t := \sqrt{1 - e^{-2t}}$.

On peut alors réécrire la définition précédente :

$$\begin{aligned} P_t(f)(x) &= \mathbb{E}_{\gamma}(f(e^{-t}x + \sqrt{1 - e^{-2t}}X)) \\ &= \mathbb{E}_{\gamma}(f(c_t x + d_t X)). \end{aligned}$$

Nous admettrons certains des résultats classiques suivants.

Proposition 1. — 1. Nous appellerons par la suite le générateur infinitésimal de la famille $(P_t)_{t \geq 0}$ l’opérateur L défini sur \mathcal{A} par :

$$L(f)(x) = f''(x) - xf'(x).$$

2. Propriété de semi-groupe : $P_t \circ P_s = P_{t+s}$ et $P_t(Id) = Id$.

3. *Propriété d'invariance et de symétrie* : soient X, Y deux variables aléatoires i.i.d. $\mathcal{N}(0, 1)$, alors pour tout $f, g \in \mathcal{A}$ et tout $t \geq 0$ on a :

$$\mathbb{E}_\gamma(fP_tg) = \mathbb{E}_\gamma(gP_tf).$$

En particulier, $\mathbb{E}_\gamma(P_tf) = \mathbb{E}_\gamma(f)$.

4. *Propriété de dérivation* : $\partial_t P_t f = (P_t \circ L)f = (L \circ P_t)f$.

5. *Propriété de contraction* : l'opérateur P_t est une contraction de $L^\infty(\gamma)$ et de $L^p(\gamma)$ quelque soit $p \geq 1$. Plus précisément, pour tout $p \in [1, +\infty)$ et tout $f \in \mathcal{A}$ on a :

$$\|P_t f\|_p \leq \|f\|_p.$$

Démonstration. — Démontrons la première assertion. Rappelons que le processus d'Ornstein-Uhlenbeck est solution de l'EDS suivante :

$$(2) \quad \begin{cases} dX_t = \sqrt{2}B_t - X_t dt, \\ X(0) = x \end{cases}$$

Nous noterons par la suite X_t^x le processus à l'instant t issu de x . La formule d'Itô (cf. (18)) appliquée à une fonction $f \in \mathcal{A}$ nous fournit :

$$f(X_t^x) = f(x) + \int_0^t [f''(X_s^x) - X_s^x f'(X_s^x)] ds + \sqrt{2} \int_0^t f'(X_s^x) dB_s.$$

Or, $f' \in L^2(\gamma)$. Ainsi le dernier terme dans l'égalité précédente est une martingale. Si on prend l'espérance dans cette dernière égalité nous obtenons

$$\lim_{t \rightarrow 0} \frac{\mathbb{E}[f(X_t^x) - f(x)]}{t} = \lim_{t \rightarrow 0} \frac{1}{t} \int_0^t \mathbb{E}[Af(X_s^x)] ds = Af(x),$$

où $Af := f'' - xf'$.

Prouvons la propriété d'invariance et de symétrie. Soit X et Y deux gaussiennes standards indépendantes. Par définition, pour toutes fonctions $f, g \in \mathcal{A}$ et pour tous $t \geq 0$, on a :

$$\mathbb{E}_\gamma[fP_tg] = \mathbb{E}_\gamma[f(X)P_tg(X)] = \mathbb{E}_\gamma[f(X)g(c_tX + d_tY)].$$

Or, les couples de variables $(X, c_tX + d_tY)$ et $(c_tX + d_tY, X)$ ont même loi. On peut donc inverser les rôles de f et g pour obtenir :

$$\mathbb{E}_\gamma(fP_tg) = \mathbb{E}_\gamma(gP_tf).$$

En appliquant ceci à $g = 1$, on trouve que $P_t 1 = 1$. Nous observons donc la propriété d'invariance de P_t sous γ . C'est-à-dire, pour toute fonction $f \in \mathcal{A}$ et tout $t \geq 0$,

$$\mathbb{E}_\gamma[P_t f] = \mathbb{E}_\gamma[f].$$

La propriété de dérivation découle de l'utilisation du Théorème de dérivation sous le signe intégral et d'une intégration par partie gaussienne (nous utiliserons fréquemment l'intégration par partie gaussienne dans nos calculs, cf. Annexe 12). \square

On peut généraliser le processus de OU en dimension n . On considère n processus d'Ornstein-Uhlenbeck de dimension un, tous indépendants les uns des autres, le processus d'OU en dimension n sera le vecteur de \mathbb{R}^n ayant pour coordonnées ces n processus indépendants. Ainsi notre semi-groupe et son générateur infinitésimal se réécrivent de la manière suivante :

- $P_t f(x) = \mathbb{E}(f(c_t x + d_t Z))$ où Z est un vecteur gaussien standard
- $Lf(x) = \Delta f(x) - x \cdot \nabla f(x)$

Une simple utilisation de l'intégration par partie gaussienne (cf. Annexe 12) nous permet d'établir un lien entre le gradient d'une fonction et son énergie de Dirichlet :

$$\mathcal{E}(f, g) = \mathbb{E}_{\gamma^n}(\nabla f \cdot \nabla g),$$

où γ est la mesure standard gaussienne sur \mathbb{R}^n .

2. Inégalités de Poincaré

2.1. Inégalité de Poincaré et semi-groupes. — Etablissons une inégalité fonctionnelle, pour le semi-groupe de OU, qui nous sera utile par la suite.

Définition 5. — Soient P_t un semi-groupe de Markov avec son énergie de Dirichlet associée \mathcal{E} et une mesure invariante μ . Ce semi-groupe vérifie une inégalité de Poincaré de constante $C > 0$ si pour toute fonction $f \in L^2(\mu)$,

$$\text{Var}_{\mu}(f) \leq C\mathcal{E}(f, f),$$

lorsque le second membre a un sens.

Proposition 2. — Le semi-groupe de OU de dimension n vérifie une inégalité de Poincaré avec une constante optimale $C = 1$.

Démonstration. — Remarquons tout d'abord que l'on a :

$$\nabla P_t f = e^{-t} P_t \nabla f,$$

où le vecteur $P_t \nabla f$ a pour i -ème composante $P_t \partial_i f$. Utilisons à présent le Lemme de covariance 1,

$$\begin{aligned} \text{Var}_\gamma(f) &= \int_0^\infty \mathcal{E}(f, P_t f) dt \\ &= \int_0^\infty \mathbb{E}_\gamma(\nabla f \cdot \nabla P_t f) dt \\ &= \int_0^\infty e^{-t} \mathbb{E}_\gamma(\nabla f \cdot P_t \nabla f) dt. \end{aligned}$$

Puis en appliquant l'inégalité de Cauchy-Schwarz on obtient les deux inégalités suivantes :

$$- \nabla f \cdot P_t \nabla f \leq |\nabla f| |P_t \nabla f| \text{ où } |\cdot| \text{ désigne la norme euclidienne.}$$

$$- \mathbb{E}_\gamma(|\nabla f| |P_t \nabla f|) \leq \left[\mathbb{E}_\gamma(|\nabla f|^2) \mathbb{E}_\gamma(|P_t \nabla f|^2) \right]^{1/2}.$$

De plus, par l'inégalité de Jensen et la propriété d'invariance :

$$\mathbb{E}_\gamma((P_t h)^2) \leq \mathbb{E}_\gamma(h^2),$$

pour n'importe quelle fonction réelle h . En combinant ces derniers résultats on obtient finalement que :

$$\text{Var}_\gamma(f) \leq \mathbb{E}_\gamma(|\nabla f|^2) = \mathcal{E}(f, f).$$

□

Remarque 4. — Nous avons choisit de présenter une démonstration du type noyau de la chaleur. Une autre méthode, pour prouver ce résultat, consiste à démontrer une inégalité de Poincaré pour la loi de Bernoulli. Enfin la propriété de tensorisation de la variance combinée avec le Théorème de la limite centrale nous permet d'obtenir l'inégalité de Poincaré pour la mesure gaussienne. Comme nous le verrons par la suite nous pouvons aussi démontrer ce résultat en utilisant un résultat de décomposition spectrale.

Donnons une application de cette inégalité, qui nous permettra de démontrer l'affirmation de l'introduction : $\text{Var}_\gamma(\max(X_1, \dots, X_n)) \leq 1$ où X_1, \dots, X_n sont des variables aléatoires gaussiennes standards indépendantes.

Proposition 3. — Soit $\mathbf{X} = (X_1, \dots, X_n)$ un vecteur gaussien de matrice de covariance Γ . Soit $f \in L^2(\gamma)$ une fonction absolument continue. L'inégalité suivante est toujours vérifiée :

$$\text{Var}(f(\mathbf{X})) \leq \mathbb{E}_{\gamma^n} [\langle M \nabla f, M \nabla f \rangle] = \mathbb{E}_{\gamma^n} [\langle \Gamma \nabla f, \nabla f \rangle].$$

Démonstration. — Sans perdre de généralité on peut supposer que Γ , qui est toujours positive, soit également définie positive. Ainsi, il existe $M \in M_n(\mathbb{R})$ telle que $\Gamma = {}^t M M$. Notons alors que, si Z suit une loi $\mathcal{N}(0, I_d)$ alors, MZ suit

une loi $\mathcal{N}(0, \Gamma)$. En appliquant l'inégalité de Poincaré, vérifiée par la mesure gaussienne de \mathbb{R}^n , à la fonction $g(x) := f(Mx)$ nous obtenons que :

$$\text{Var}(f(\mathbf{X})) \leq \mathbb{E}_{\gamma^n} [\langle M\nabla f, M\nabla f \rangle] = \mathbb{E}_{\gamma^n} [\langle \Gamma\nabla f, \nabla f \rangle].$$

Si f n'est pas suffisamment régulière, on procède par convolution avec un noyau gaussien $\frac{1}{(2\pi\epsilon)^{n/2}} e^{-|x|^2/2\epsilon}$. On conclut par convergence dominée, en faisant tendre ϵ vers 0. \square

Nous en déduisons le résultat suivant :

Corollaire 1. — Soit $\mathbf{X} = (X_1, \dots, X_n)$ un vecteur gaussien de matrice de covariance Γ . L'inégalité suivante est toujours vérifiée :

$$\text{Var}(\max_{i=1\dots n} X_i) \leq \max_{i=1\dots n} \text{Var}(X_i).$$

Démonstration. — En remarquant que la fonction suivante :

$$f(x) := \frac{1}{\beta} \log\left(\sum_{i=1}^n e^{\beta x_i}\right),$$

où $x = (x_1, \dots, x_n)$ converge vers $\max_{i=1\dots n} x_i$ lorsque $\beta \rightarrow \infty$ (il s'agit d'une approximation de Gibbs). Il nous suffit de montrer la proposition pour la fonction f et conclure par convergence dominée.

En conservant les notations de la démonstration de la proposition précédente et en appliquant cette dernière à la fonction f définie ci-dessus, nous avons :

$$\text{Var}_{\gamma}(f(\mathbf{X})) \leq \mathbb{E}_{\gamma} [\langle M\nabla f, M\nabla f \rangle] = \mathbb{E}_{\gamma} [\langle \Gamma\nabla f, \nabla f \rangle],$$

où $\nabla f_i = \left(\frac{e^{\beta x_i}}{\sum_{i=1}^n e^{\beta x_i}} \right)$. On pose alors pour alléger les notations :

$$Z_n := \sum_{i=1}^n e^{\beta x_i}.$$

Nous obtenons donc :

$$\begin{aligned} \text{Var}_{\gamma}(f(\mathbf{X})) &\leq \mathbb{E}_{\gamma} [\langle \Gamma\nabla f, \nabla f \rangle] \\ &= \mathbb{E}_{\gamma} \left[\sum_{i,j=1}^n \Gamma_{i,j} \frac{e^{\beta X_i}}{Z_n} \frac{e^{\beta X_j}}{Z_n} \right] \\ &\leq \max_{i,j} \Gamma_{i,j} \times \mathbb{E}_{\gamma} \left[\sum_{i,j=1}^n \frac{e^{\beta X_i}}{Z_n} \frac{e^{\beta X_j}}{Z_n} \right] = \max_{i,j} \Gamma_{i,j} \\ &\leq \max_i \Gamma_{i,i} = \max_i \text{Var}_{\gamma}(X_i). \end{aligned}$$

En effet, pour tout i, j par Cauchy-Schwarz,

$$\text{Cov}(X_i, X_j) \leq \mathbb{E}_\gamma(X_i^2)^{1/2} \mathbb{E}_\gamma(X_j^2)^{1/2} \leq \max_i \mathbb{E}_\gamma(X_i^2).$$

Et comme $f(\mathbf{X}) \leq \log n \sum_{i=1}^n X_i \in L^2$ on peut conclure par convergence dominée. \square

Autrement dit, ceci signifie que l'ordre de fluctuation du maximum est toujours inférieur à l'ordre de fluctuation de la coordonnée la plus fluctuante. Nous noterons bien que ce résultat ne prend plus en compte la structure de corrélation de notre vecteur.

2.2. Inégalité de Poincaré et décomposition spectrale. — On considère de nouveau $(X_t)_{t \geq 0}$ un processus de Markov réversible à valeurs dans un espace d'état abstrait, de semi-groupe P_t , admettant une mesure d'équilibre μ , un générateur infinitésimal L et une forme de Dirichlet \mathcal{E} .

Dans cette section nous allons nous intéresser au spectre de l'opérateur L . L'étude du spectre combiné avec l'identité de Plancherel nous fournit un nouveau point de vue sur l'inégalité de Poincaré vérifiée par la mesure gaussienne. En particulier cette nouvelle approche nous permettra par la suite d'obtenir un outil pour montrer un phénomène de superconcentration, via le semi-groupe de OU et des polynômes d'Hermite qui sont associés au générateur infinitésimal de ce semi-groupe.

Puisque le processus que nous considérons est réversible cela implique que L est un opérateur autoadjoint. Plus encore, celui-ci est même défini négativement. En effet l'inégalité de Cauchy-Schwarz et celle de Jensen nous assurent que pour toute fonction $f \in L^2(\mu)$ nous avons :

$$\int f P_t f d\mu \leq \left(\int f^2 d\mu \right)^{1/2} \left(\int (P_t f)^2 d\mu \right)^{1/2} \leq \int f^2 d\mu.$$

C'est pourquoi l'utilisation du Lemme de Fatou nous fournit l'inégalité suivante :

$$\langle f, Lf \rangle = \int \lim_{t \rightarrow 0} \frac{f(P_t f - f)}{t} d\mu \leq \liminf_{t \rightarrow 0} \int \frac{f(P_t f - f)}{t} d\mu \leq 0.$$

On remarque alors que 0 est toujours une valeur propre de L puisque L , appliqué à n'importe quelle constante vaut toujours 0. Souvent, comme cela sera le cas du processus d'Ornstein-Uhlenbeck, les valeurs propres de l'opérateur $-L$ peuvent être ordonnées en une suite dénombrable croissante : $0 = \lambda_0 \leq \lambda_1 \leq \dots$ auxquelles correspondent, respectivement, une suite de fonctions propres $u_0 := 1, u_1, \dots$

Cette suite de fonctions propres forme une base hilbertienne de $L^2(\mu)$. Si $Lf = 0$ alors $P_t f = e^{tL} f = f$ pour toutes fonctions f . C'est pourquoi $f = \lim_{t \rightarrow \infty} P_t f = \int f d\mu$. En particulier, l'identité de Plancherel nous donne

$$\|f\|_{L^2(\mu)}^2 = \sum_{k=0}^{\infty} \langle u_k, f \rangle^2.$$

En traduisant ceci de manière probabiliste nous avons, $\langle u_0, f \rangle = \mathbb{E}_\mu(f)$ et que :

$$\text{Var}_\mu(f) = \sum_{k=1}^{\infty} \langle u_k, f \rangle^2.$$

Nous pouvons obtenir une relation entre l'énergie de Dirichlet et le spectre de L . Puisque

$$f = \sum_{k=0}^{\infty} \langle u_k, f \rangle u_k,$$

nous en déduisons que :

$$\mathcal{E}(f, f) = -(f, Lf) = \sum_{k=0}^{\infty} \lambda_k \langle u_k, f \rangle^2 = \sum_{k=1}^{\infty} \lambda_k \langle u_k, f \rangle^2.$$

Etablissons à présent un Lemme reliant l'inégalité de Poincaré du semi-groupe avec la décomposition spectrale.

Lemme 2. — *Un semi-groupe de Markov satisfait une inégalité de Poincaré si et seulement si $\lambda_1 > 0$, et dans ce cas, la constante optimale est $1/\lambda_1$.*

Démonstration. — Supposons $\lambda_1 > 0$. Alors pour n'importe quelle fonction $f \in L^2(\mu)$,

$$\text{Var}_\mu(f) = \sum_{k \geq 1} \langle u_k, f \rangle^2 \leq \frac{1}{\lambda_1} \sum_{k \geq 1} \lambda_k \langle u_k, f \rangle^2 = \frac{1}{\lambda_1} \mathcal{E}(f, f).$$

Pour montrer que $1/\lambda_1$ est la constante optimale, il suffit d'observer que l'égalité est atteinte lorsque $f = u_1$.

Si maintenant $\lambda_1 = 0$, alors u_1 doit être non constant et donc de variance non nulle. En effet u_1 est orthogonal aux fonctions constantes. D'un autre côté $\mathcal{E}(u_1, u_1) = 0$. Ceci prouve donc que le semi-groupe de Markov ne peut satisfaire une inégalité de Poincaré. \square

Remarque 5. — *Finalement tout ceci revient uniquement à comparer f et f' en norme L^2 après les avoir décomposées suivant les polynômes d'Hermite.*

Voici un résultat important qui exploite la version spectrale de l'inégalité de Poincaré. Nous l'utiliserons par la suite pour démontrer un premier résultat de superconcentration pour le modèle des verres de spin de Sherrington et Kirkpatrick.

Proposition 4. — Soit $m \geq 1$ et supposons que la m -ième valeur propre λ_m de l'opérateur L soit strictement positive. Alors pour n'importe quelle fonction $f \in L^2(\mu)$,

$$\text{Var}_\mu(f) \leq \sum_{k=1}^{m-1} \langle u_k, f \rangle^2 + \frac{\mathcal{E}(f, f)}{\lambda_m}.$$

Démonstration. — Grâce à la formule de Plancherel nous pouvons réécrire la variation de f sous forme de série. Nous utiliserons aussi le fait que nos valeurs propres sont rangées par ordre croissant.

$$\begin{aligned} \text{Var}_\mu(f) &= \sum_{k=1}^{m-1} \langle u_k, f \rangle^2 + \sum_{k=m}^{\infty} \langle u_k, f \rangle^2 \\ &\leq \sum_{k=1}^{m-1} \langle u_k, f \rangle^2 + \frac{1}{\lambda_m} \sum_{k=m}^{\infty} \lambda_k \langle u_k, f \rangle^2 \\ &\leq \sum_{k=1}^{m-1} \langle u_k, f \rangle^2 + \frac{1}{\lambda_m} \sum_{k=1}^{\infty} \lambda_k \langle u_k, f \rangle^2 = \sum_{k=1}^{m-1} \langle u_k, f \rangle^2 + \frac{\mathcal{E}(f, f)}{\lambda_m}. \end{aligned}$$

Ceci termine la preuve de la Proposition. \square

2.3. Décomposition spectrale du semi-groupe d'Ornstein-Uhlenbeck.

— Nous commençons par de brefs rappels sur les polynômes d'Hermite. Ceux-ci sont d'un intérêt tout particulier car, comme nous allons le voir, ils sont étroitement liés à la mesure gaussienne et au générateur infinitésimal du semi-groupe OU.

Définition 6. — Pour tout $k \geq 0$ on définit le k -ième polynômes d'Hermite H_k par

$$H_k(x) := (-1)^k e^{x^2/2} \frac{d^k}{dx^k} e^{-x^2/2}.$$

Exemple 1. — Par exemple si l'on calcule les premiers polynômes nous obtenons

$$- H_0 = 1$$

- $H_1(x) = x$
- $H_2(x) = x^2 - 1$

Voici deux résultats fondamentaux sur ces polynômes.

Proposition 5. — Les polynômes d’Hermite $(H_k)_k$ forment une base hilbertienne de $L^2(\gamma)$ et nous avons la relation suivante :

$$\int_{\mathbb{R}} H_n(x)H_m(x)d\gamma(x) = n!\delta_{n,m}.$$

En particulier $\|H_n\|_{L^2(\gamma)}^2 = n!$.

- De plus, si l’on note $Lf(x) = f''(x) - xf'(x)$ le générateur infinitésimal du semi-groupe d’Ornstein-Uhlenbeck. Alors H_k est la fonction propre associée à la valeur propre k de l’opérateur $-L$.

Remarque 6. — Une façon naturelle d’introduire ces polynômes est la suivante : les polynômes forment un sous-espace dense de $L^2(\gamma)$.

En effet, soit $f \in L^2(\gamma)$ orthogonale à tous les polynômes. La transformée de Laplace de la mesure $\nu = f\gamma$ est finie sur \mathbb{R} tout entier : pour $\lambda \in \mathbb{R}$

$$\int e^{\lambda x}d\nu(x) \leq \left(\int f^2d\gamma \right) \left(\int e^{2\lambda x}d\gamma(x) \right) < \infty.$$

La transformée de Laplace de ν est donc, en particulier, analytique au voisinage de 0. Puisque f est orthogonale à tous les polynômes, toutes les dérivées de cette transformée sont nulles en 0 et par suite la transformée de Laplace de f est également nulle. Ceci implique la nullité de f et la densité des polynômes.

Nous pouvons donc trouver une base hilbertienne polynômiale de $L^2(\gamma)$: il s’agit de la famille des polynômes d’Hermite. Une grande particularité de cette famille et qu’il s’agit des fonctions propres du générateur infinitésimal du semi-groupe d’Ornstein-Uhlenbeck.

Soulignons cependant le fait qu’étant donné un semi-groupe muni d’une mesure invariante μ et d’un générateur infinitésimal L , il n’est pas automatique que la base hilbertienne de $L^2(\mu)$ soit également les fonctions propres de l’opérateur L .

Nous pouvons étendre toutes ces définitions au cas multidimensionnel. Notons de nouveau par L , le générateur infinitésimal du semi-groupe d’Ornstein-Uhlenbeck de dimension n ,

$$i.e. \quad Lf(x) = \Delta f(x) - x \cdot \nabla f(x).$$

Les valeurs propres et les fonctions propres de l'opérateur L sont maintenant indexées par \mathbb{Z}_+^n .

Définition 7. — Pour chaque $k = (k_1, \dots, k_n) \in \mathbb{Z}_+^n$ nous définissons le k -ième polynôme d'Hermite H_k par :

$$H_k(x) := \prod_{i=1}^n H_{k_i}(x_i),$$

où $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Nous retrouvons alors des propriétés analogues à la dimension 1.

Proposition 6. — — Le k -ième polynôme d'Hermite H_k est une fonction propre de l'opérateur $-L$ correspondant à la valeur propre $k_1 + \dots + k_n$.

– Nous avons la formule suivante : $\|H_k\|_{L^2(\gamma)}^2 = k! := k_1! \cdots k_n!$.

– En utilisant la formule de Plancherel, nous obtenons, pour n'importe quelle fonction $f \in L^2(\gamma)$, la relation suivante :

$$(3) \quad \|f\|_{L^2(\gamma)}^2 = \sum_{k \in \mathbb{Z}_+^n} \frac{\langle f, H_k \rangle^2}{k!}.$$

Soit $f \in L^2(\gamma)$, supposons de plus que f soit assez régulière pour autoriser les opérations qui vont suivre ($f \in \mathcal{A}$ par exemple). Nous allons voir que l'intégration par partie gaussienne nous fournit une façon agréable de calculer les produits scalaires (f, H_k) .

Proposition 7. — Pour tout $k \in \mathbb{Z}_+^n$ nous avons la relation de récurrence suivante :

$$\langle f, H_k \rangle = \mathbb{E}_\gamma \left(\frac{\partial^{k_1 + \dots + k_n} f}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}} \right),$$

Démonstration. — Considérons la dimension 1 pour mieux comprendre d'où provient cette formule. En utilisant l'intégration par partie gaussienne nous trouvons la relation suivante :

$$\langle f, H_k \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) (-1)^k \frac{d^k}{dx^k} e^{-x^2/2} dx = \langle f', H_{k-1} \rangle.$$

Par une récurrence immédiate nous en déduisons que :

$$\langle f, H_k \rangle = \langle f^{(k)}, H_0 \rangle = \mathbb{E}_\gamma(f^{(k)}),$$

où $f^{(k)}$ désigne la dérivée k -ième de f .

Cette formule se généralise directement dans \mathbb{R}^n : pour $k = (k_1, \dots, k_n)$ et f une fonction régulière dont les dérivées partielles sont toutes dans $L^2(\gamma)$,

$$\langle f, H_k \rangle = \mathbb{E}_\gamma \left(\frac{\partial^{k_1 + \dots + k_n} f}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} \right).$$

□

La dernière Proposition nous permet de réécrire la formule (3) de manière probabiliste :

$$\|f\|_{L^2(\gamma)} = \sum_{k \in \mathbb{Z}_+^n} \frac{1}{k!} \left(\mathbb{E}_\gamma \left(\frac{\partial^{k_1 + \dots + k_n} f}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} \right) \right)^2.$$

Cette même somme peut se réorganiser de la façon suivante :

$$\begin{aligned} \|f\|_{L^2(\gamma)} &= \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{k \in \mathbb{Z}_+^n, k_1 + \dots + k_n = m} \frac{m!}{k!} \left(\mathbb{E}_\gamma \left(\frac{\partial^{k_1 + \dots + k_n} f}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} \right) \right)^2 \\ &= \sum_{m=0}^{\infty} \frac{\theta_m(f)}{m!}, \end{aligned}$$

où

$$(4) \quad \theta_m(f) := \sum_{1 \leq i_1, \dots, i_m \leq n} \left(\mathbb{E}_\gamma \left(\frac{\partial^{k_1 + \dots + k_n} f}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} \right) \right)^2.$$

Remarque 7. — Nous pouvons aussi écrire ceci de manière probabiliste :

$$\text{Var}_\gamma(f) = \sum_{m=1}^{\infty} \frac{\theta_m(f)}{m!}.$$

3. Inégalité fonctionnelle et hypercontractivité

3.1. Inégalités de Sobolev logarithmiques. — Nous allons faire un bref rappel sur les inégalités de Sobolev logarithmiques avant de parler du phénomène d'hypercontractivité qui sera un des outils majeurs de notre étude.

Définition 8. — On dit qu'une mesure μ satisfait une inégalité de Sobolev logarithmique sur une certaine classe de fonctions \mathcal{C}_{LS} , lorsqu'il existe une constante $c > 0$ telle que, pour toute fonction $f \in \mathcal{C}_{LS}$, on ait :

$$\text{Ent}_\mu(f^2) \leq c\mathcal{E}(f).$$

Où $\text{Ent}_\mu(f) := \mathbb{E}_\mu(f \log f) - \mathbb{E}_\mu(f) \log \mathbb{E}_\mu(f)$ désigne l'entropie de la fonction mesurable positive f .

Et $\mathcal{E}_\mu(f) = \mathbb{E}_\mu(f'^2)$ désigne l'énergie d'une fonction localement intégrable dont la dérivée est dans L^2 .

Remarque 8. — Remarquons qu'il existe une hiérarchie entre les inégalités de Sobolev logarithmiques et les inégalités de Poincaré. C'est à dire : si μ est une mesure vérifiant une inégalité de Sobolev logarithmique de constante c_{LS} pour une classe de fonctions \mathcal{C}_{LS} alors μ vérifie une inégalité de Poincaré sur au moins la même classe de fonctions avec une constante c_P et l'on a la relation suivante :

$$c_P \leq \frac{1}{2}c_{LS}.$$

En effet, soit $f \in \mathcal{C}_{LS}$. Un développement limité de la fonction

$$h \mapsto (1+h)\log(1+h)$$

en 0 nous donne :

$$Ent_\mu((1+\epsilon f)^2) = 2\epsilon^2 Var_\mu(f) + O(\epsilon^3).$$

D'autre part, on a

$$\mathcal{E}_\mu(1+\epsilon f) = \epsilon^2 \mathcal{E}_\mu(f).$$

En faisant tendre ϵ vers 0^+ pour obtenir l'inégalité de Poincaré pour f , de constante $\frac{1}{2}c_{LS}$.

Proposition 8. — La mesure gaussienne γ vérifie une inégalité de Sobolev logarithmique avec une constante $c_\gamma = 2$.i.e.

$$Ent_\gamma(f^2) \leq 2\mathcal{E}_\gamma(f).$$

Démonstration. — Soit f une fonction suffisamment régulière et positive. Nous allons démontrer que la mesure gaussienne vérifie une inégalité de Sobolev logarithmique. Pour cela nous allons utiliser le semi-groupe d'Ornstein-Uhlenbeck et les propriétés qu'il vérifie.

Rappelons que $P_0 f = f$ et que $\lim_{t \rightarrow \infty} P_t f = \int f d\gamma$. Ainsi nous pouvons réécrire l'entropie de la façon suivante :

$$Ent_\gamma(f) = - \int_0^\infty \frac{d}{dt} \left(\int P_t f \log P_t f d\gamma \right) dt$$

Puisque nous avons supposé notre fonction f suffisamment régulière nous pouvons dériver sous le signe intégrale. Nous utiliserons par la suite les résultats suivants :

$$- \frac{d}{dt} P_t = L P_t$$

$$- \int f(-Lg) d\gamma = \int \nabla f(-\nabla g) d\gamma$$

Ainsi

$$\begin{aligned}
 Ent_\gamma(f) &= - \int_0^\infty \int \frac{d}{dt} (P_t f \log P_t f) d\gamma dt \\
 &= - \int_0^\infty \int L P_t f \log P_t f d\gamma dt - \int_0^\infty \int L P_t f d\gamma dt \\
 &= \int_0^\infty \int \nabla P_t f \nabla \log P_t f d\gamma dt - \int_0^\infty \int P_t f L 1 d\gamma dt \\
 &= \int_0^\infty \int \frac{(\nabla P_t f)^2}{P_t f} d\gamma dt
 \end{aligned}$$

Nous avons utilisé dans la dernière égalité le fait que L est un opérateur autoadjoint et que $L1 = 0$. De plus $\nabla P_t f = e^{-t} P_t(\nabla f)$ donc $|\nabla P_t f| \leq e^{-t} P_t(|\nabla f|)$. L'utilisation de l'inégalité de Cauchy-Schwarz nous assure également que :

$$\left(P_t(|\nabla f|) \right)^2 \leq P_t \left(\frac{|\nabla f|^2}{f} \right) P_t(f).$$

En combinant ces dernières inégalités nous obtenons alors :

$$\begin{aligned}
 Ent_\gamma(f) &\leq \int_0^\infty e^{-2t} \left(\int P_t \left(\frac{|\nabla f|^2}{f} \right) d\gamma \right) dt \\
 &= \frac{1}{2} \int \frac{|\nabla f|^2}{f} d\gamma.
 \end{aligned}$$

Il suffit ensuite de remplacer f par f^2 pour terminer la preuve. \square

Une propriété admirable de ce type d'inégalité est que celle-ci se préserve par tensorisation : à partir de l'inégalité de Sobolev logarithmique pour la loi gaussienne réelle on peut obtenir le même résultat avec la même constante pour la loi gaussienne sur \mathbb{R}^n . Cette propriété permet de montrer le Théorème d'hypercontractivité qui atteste que l'opérateur d'Ornstein-Uhlenbeck est hypercontractif.

3.2. Hypercontractivité. —

Définition 9. — *Un semi-groupe $(P_t)_{t \geq 0}$, agissant sur un espace de fonctions \mathcal{B} de \mathbb{R}^n . Soit μ une mesure invariante pour cette famille d'opérateurs, i.e. pour tout $f \in \mathcal{B}$ $\int P_t f d\mu = \int f d\mu$.*

Ce semi-groupe est dit hypercontractif si pour tout $t > 0$ il existe des nombres $q > p > 1$ (pouvant dépendre de t) tels que pour tout $f \in L^q(\mu)$:

$$\|P_t f\|_q \leq \|f\|_p.$$

Rappelons le, la définition du semi-groupe d'Ornstein-Uhlenbeck se généralise naturellement à \mathbb{R}^n en conservant la même définition avec cette fois X un vecteur gaussien standard de \mathbb{R}^n .

Théorème 1 (Hypercontractivité). — Soient $1 < p < q < \infty$.

- Si $q - 1 \leq e^{2t}(p - 1)$, alors l'opérateur P_t d'Ornstein-Uhlenbeck est hypercontractif de $L^p(\gamma)$ dans $L^q(\gamma)$. C'est à dire que pour toute fonction $f \in \mathcal{A}$, on a :

$$\|P_t f\|_q \leq \|f\|_p.$$

- Si $q - 1 > e^{2t}(p - 1)$, alors l'opérateur P_t n'est pas continu de $L^p(\gamma)$ dans $L^q(\gamma)$.

Démonstration. — Nous ferons la démonstration du premier point dans l'annexe (11). \square

A noter que ce résultat ne dépend absolument pas de la dimension n . Ceci fait partie d'une des remarquables propriétés qui fait de l'hypercontractivité un outil puissant.

Nous attirons aussi l'attention du lecteur sur le fait qu'il existe une forte connection entre les inégalités de Sobolev logarithmiques et l'hypercontractivité d'un semi-groupe. En fait, si l'on considère $(P_t)_{t \geq 0}$ un semi-groupe de Markov de mesure invariante μ . Si, de plus, μ est réversible nous avons l'équivalence entre « μ vérifie une inégalité de Sobolev logarithmique » et « $(P_t)_{t \geq 0}$ est hypercontractif ». Voir pour plus de détails le Théorème de Gross dans (1) ou dans l'article de Gross lui même (??).

En suivant une idée de Talagrand, nous allons voir que cet outil d'hypercontractivité permet d'établir des résultats sur la superconcentration pour certains des modèles étudiés.

Rappelons brièvement le cadre considéré : nous avons un vecteur gaussien centré $\mathbf{X} = (X_i)_{1 \leq i \leq n}$ dont les composantes ne sont pas nécessairement indépendantes et $\Gamma = \Gamma(i, j)_{1 \leq i, j \leq n}$ qui désigne la matrice de covariance.

On peut toujours définir un semi-groupe $(\check{P}_t)_{t \geq 0}$ par :

$$\check{P}_t f(x) := \mathbb{E}(f(c_t x + d_t \mathbf{X})).$$

On remarque que ce semi-groupe conserve la propriété d'hypercontractivité de celui d'Ornstein Uhlenbeck avec exactement les mêmes relations entre p et q .

En effet, en supposant que \mathbf{X} ne soit pas le vecteur nul, on sait qu'il existe $d \leq n$, tel qu'il existe une matrice de rang plein $M \in \mathcal{M}_{n \times d}(\mathbb{R})$ vérifiant ${}^tMM = \Gamma$. On peut supposer l'existence d'un vecteur gaussien standard d dimensionnel \mathbf{Z} tel que $\mathbf{X} = M\mathbf{Z}$. Soit \mathbf{Z}' une copie indépendante de \mathbf{Z} et posons $\mathbf{X}' = B\mathbf{Z}'$. Etant donné une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$, on définit l'application $g : \mathbb{R}^d \rightarrow \mathbb{R}$ par :

$$g(x) := f(Mx).$$

Le vecteur aléatoire \mathbf{X} est porté par l'image de \mathbb{R}^d sous l'application B . Sur ce sous-espace, M est inversible, on a alors :

$$\begin{aligned} \check{P}_t f(x) &= \mathbb{E}(f(c_t x + d_t \mathbf{X})) \\ &= \mathbb{E}(f(c_t M M^{-1} x + d_t M \mathbf{Z})) \\ &= \mathbb{E}(g(c_t M^{-1} x + d_t \mathbf{Z})) \\ &= P_t g(M^{-1} x). \end{aligned}$$

Ainsi, nous obtenons pour tout $t \geq 0$,

$$\check{P}_t f(\mathbf{X}) = P_t g(B^{-1} \mathbf{X}) = P_t(\mathbf{Z}).$$

Finalement le Théorème de Nelson (1) nous fournit que :

$$\|\check{P}_t f(\mathbf{X})\|_q = \|P_t(\mathbf{Z})\|_q \leq \|g(\mathbf{Z})\|_p = \|f(\mathbf{X})\|_p.$$

La propriété d'hypercontractivité est donc conservée avec les mêmes p et q . Le lemme suivant est une conséquence directe de ceci et sera d'une importance cruciale par la suite. Nous rappelons que $\mathbf{X}^t := c_t \mathbf{X} + d_t \mathbf{X}'$ où \mathbf{X}' est une copie indépendante de \mathbf{X} .

Lemme 3. — *Pour tout fonction mesurable $f : \mathbb{R}^n \rightarrow [0, 1]$ et tout $t \geq 0$, nous avons l'estimation suivante :*

$$\mathbb{E}(f(\mathbf{X})f(\mathbf{X}')) \leq \mathbb{E}(f(\mathbf{X}))^{1+\tanh(t/2)}.$$

Démonstration. — Soient $p > 1$ et $q = 1 + e^{2t}(p - 1)$. On pose $q' = q/(q - 1)$, puisque f est à valeurs dans $[0, 1]$ nous avons, en utilisant l'inégalité de Hölder et la propriété d'hypercontractivité, que :

$$\begin{aligned} \mathbb{E}(f(\mathbf{X})f(\mathbf{X}')) = \mathbb{E}(f(\mathbf{X})\check{P}_t f(\mathbf{X})) &\leq \|f(\mathbf{X})\|_{q'} \|\check{P}_t f(\mathbf{X})\|_q \\ &\leq \|f(\mathbf{X})\|_{q'} \|f(\mathbf{X})\|_p \\ &\leq \left(\mathbb{E}(f(\mathbf{X})) \right)^{1/q' + 1/p}. \end{aligned}$$

On optimise ensuite en p , ce qui nous fournit :

$$p = 1 + e^{-t}.$$

Et enfin un calcul facile nous donne que :

$$\begin{aligned} \frac{1}{q'} + \frac{1}{p} &= 1 - \frac{1}{1 + e^{2t}(p-1)} + \frac{1}{p} \\ &= 1 + \tanh(t/2). \end{aligned}$$

Ceci conclut la démonstration. \square

Ce résultat nous permet d'obtenir la Proposition (9) suivante. Comme nous le ferons remarquer par la suite, l'inégalité que nous allons obtenir est optimale, néanmoins elle n'est pas très utile car les hypothèses imposées sont très fortes. Toutefois, cette proposition nous permettra de démontrer un phénomène de superconcentration dans un cadre simple.

On rappelle les notations suivantes : $\mathbf{M} = \max_{i=1,\dots,n} X_i$ et $\Gamma(i, j) = \text{Cov}(X_i, X_j)$.

Proposition 9. — *Supposons que $\Gamma(i, j) \leq \rho$ pour tout $i \neq j$ et $\Gamma(i, i) = \sigma^2$ pour tout i . Alors,*

$$\mathbb{E}[\mathbf{M}] \leq \frac{C\sigma^2}{\log n} + C\rho,$$

où C est une constante universelle.

Comme nous l'avons signalé auparavant, cette borne est optimale si l'on considère le cas où $\Gamma(i, j) = \rho$ pour tout $i \neq j$.

Démonstration. — Pour tout vecteur $x \in \mathbb{R}^n$ dont les coordonnées sont toutes distinctes on définit :

$$f_i(x) = \mathbb{I}_{\{x_i = \max_j x_j\}}.$$

Sous nos hypothèses et en utilisant le Lemme précédent nous obtenons que :

$$\begin{aligned} \mathbb{E}(\Gamma(I^0, I^t)) &\leq \sigma^2 \mathbb{P}(I^0 = I^t) + \rho \mathbb{P}(I^0 \neq I^t) \\ &= \sigma^2 \sum_{i=1,\dots,n} \mathbb{E}(f_i(\mathbf{X}^0) f_i(\mathbf{X}^t)) + \rho \mathbb{P}(I^0 \neq I^t) \\ &\leq \sigma^2 \sum_{i=1,\dots,n} \left(\mathbb{E}(f_i(\mathbf{X}^0)) \right)^{1+\tanh(t/2)} + \rho \\ &\leq \sigma^2 \left(\max_{i=1,\dots,n} \mathbb{P}(I^0 = i) \right)^{\tanh(t/2)} \sum_{i=1,\dots,n} \mathbb{P}(I^0 = i) + \rho \\ &= \sigma^2 \left(\max_{i=1,\dots,n} \mathbb{P}(I^0 = i) \right)^{\tanh(t/2)} + \rho. \end{aligned}$$

\square

Retrouvons notre « baby » modèle dont nous rappelons l'énoncé. Soit X une variable aléatoire réelle gaussienne standard, prenons alors X_1, \dots, X_n n copies indépendantes. Nous voulons nous intéresser au maximum de cette famille : $M = \max_i X_i$.

L'utilisation de la proposition précédente nous fournit alors que :

$$\text{Var}(M) \leq \frac{C}{\log n}$$

où $C > 0$.

Il faut toutefois noter que la proposition précédente est un outil bien trop puissant pour obtenir les résultats que nous venons de mentionner. En effet une analyse fine permet de retrouver cette estimation de la variance du max à la main. Néanmoins il est important de remarquer que l'hypercontractivité nous a permis d'estimer la variance du maximum de notre famille gaussienne sans avoir de renseignement sur la taille de l'espérance de celui-ci.

4. Méthode de Talagrand

Nous allons définir, dans la section suivante, deux nouvelles notions : la superconcentration et le chaos. Nous aurons même un Théorème nous assurant l'équivalence entre celles-ci. Nous devons à présent trouver des méthodes pour montrer qu'un modèle vérifie l'un de ses deux phénomènes.

Pour cela, nous allons utiliser une méthode due à Talagrand. Cette méthode s'appuie sur l'hypercontractivité du semi-groupe d'Ornstein-Uhlenbeck démontrée par Nelson (cf. (1)).

Nous rappelons brièvement son énoncé. Pour n'importe quel $p > 1$ et $t > 0$ il existe une fonction $q(t, p) = 1 + (p - 1)e^{2t}$ telle que pour toute fonctions $f \in L^p(\gamma^n)$:

$$\|P_t f\|_{L^q(\gamma^n)} \leq \|f\|_{L^p(\gamma^n)}.$$

L'idée de Talagrand nous permet d'utiliser l'hypercontractivité pour améliorer une inégalité de Poincaré. Pour en savoir plus sur l'origine de cette inégalité nous renvoyons le lecteur vers l'article de D. Cordero-Erausquin et M. Ledoux (7).

Théorème 2 (Talagrand). — Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction absolument continue. Notons par $\partial_i f$ la dérivée partielle de f en la i -ème coordonnée.

Supposons que pour tout $i = 1, \dots, n$ il existe un nombre A_i tel que $\|\partial_i f\|_{L^2(\gamma)} \leq A_i$. Alors

$$\text{Var}_\gamma(f) \leq C \sum_{i=1}^n \frac{A_i^2}{1 + \log \left(A_i / \|\partial_i f\|_{L^1(\gamma)} \right)},$$

où C est une constante indépendante de f et de n .

Avant de démontrer ce théorème, faisons quelques remarques sur celui-ci pour mieux le comprendre.

Rappelons que l'inégalité de Poincaré pour la mesure gaussienne nous dit que :

$$\text{Var}_\gamma(f) \leq \sum_{i=1}^n \|\partial_i f\|_{L^2(\gamma)}.$$

Ainsi le facteur logarithmique au dénominateur est une amélioration de l'inégalité de Poincaré. Ce facteur nous donne une meilleure estimation ; d'autant que la norme L^2 des dérivées partielles est beaucoup plus grande que la norme L^1 des dérivées partielles.

Considérons deux exemples.

Si l'on choisit $f(x) = \sum_{i=1}^n x_i$, alors $\partial_i f = 1$ et les normes L^1 et L^2 coïncident. Dans ce cas là, le Théorème de Talagrand ne donne aucune amélioration. Ce qui est logique, puisque l'inégalité de Poincaré est optimale pour un tel choix de fonction f .

En revanche, si $f(x) = \max_{i=1, \dots, n} x_i$ alors

$$\partial_i f = \mathbb{I}_{\{x_i \geq x_j \forall j\}}.$$

Ainsi, si l'on considère X_1, \dots, X_n des variables aléatoires réelles gaussiennes standard i.i.d (notre « baby » modèle). Nous obtenons donc

$$\|\partial_i f\|_{L^2(\gamma)}^2 = \|\partial_i f\|_{L^1(\gamma)} = \mathbb{P}(X_i \geq X_j \forall j) = \frac{1}{n}.$$

Cette fois-ci, la méthode de Talagrand nous fournit l'estimation suivante :

$$\text{Var}(\max_{i=1, \dots, n} X_i) \leq \frac{C}{\log n},$$

qui est optimale (nous retrouvons donc le résultat annoncé dans l'introduction). Alors que l'inégalité de Poincaré nous donne

$$\text{Var}(\max_{i=1, \dots, n} X_i) \leq 1.$$

Attention, il ne faudrait pas croire que la méthode de Talagrand nous fournit toujours un résultat optimal. Cet outil reste donc une méthode pour obtenir des estimations mais ce ne sera pas forcément la meilleure façon de procéder. Il s'agit seulement d'un moyen pour estimer la variance d'une fonction sans connaître forcément de résultat sur l'espérance de celle-ci.

Tournons nous à présent vers la démonstration de ce résultat.

Lemme 4. — Soit $h \in L^2(\gamma)$ telle qu'il existe une constante $A > 0$ vérifiant $\|h\|_{L^2(\gamma)} \leq A$. On pose aussi $p = 1 + e^{-2t}$.

Alors

$$\mathbb{E}_\gamma(hP_t h) \leq A^{3-\frac{2}{p}} \|h\|_{L^1(\gamma)}^{\frac{2}{p}-1} = A^2 \left(\frac{\|h\|_{L^1(\gamma)}}{A} \right)^{\tanh(t)}.$$

Démonstration. — Nous allons utiliser l'inégalité de Cauchy-Schwarz, celle de Hölder et le Théorème d'hypercontractivité de Nelson avec $q = 2$.

$$\begin{aligned} \mathbb{E}_\gamma(hP_t h) &\leq \|h\|_{L^2(\gamma)} \|P_t h\|_{L^2(\gamma)} \\ &\leq \|h\|_{L^2(\gamma)} \|h\|_{L^p(\gamma)} \end{aligned}$$

De plus, comme $1 < p < 2$ et $|h|^p = |h|^{2p-2}|h|^{2-p}$ l'inégalité de Hölder, avec $p' = \frac{1}{p-1}$ et $q' = \frac{1}{2-p}$, nous assure que :

$$\|h\|_{L^p(\gamma)} \leq \|h\|_{L^2(\gamma)}^{2-\frac{2}{p}} \|h\|_{L^1(\gamma)}^{\frac{2}{p}-1}.$$

Comme par hypothèse $\|h\|_{L^2(\gamma^n)} \leq A$ et $t = -\log(\sqrt{p-1})$, en combinant les deux inégalités précédentes, on obtient que :

$$\begin{aligned} \mathbb{E}_\gamma(hP_t h) &\leq \|h\|_{L^2(\gamma)}^{3-\frac{2}{p}} \|h\|_{L^1(\gamma)}^{\frac{2}{p}-1} \\ &\leq A^{3-\frac{2}{p}} \|h\|_{L^1(\gamma)}^{\frac{2}{p}-1} = A^2 \left(\frac{\|h\|_{L^1(\gamma)}}{A} \right)^{\tanh(t)}. \end{aligned}$$

Armé du précédent Lemme nous pouvons donc démontrer le Théorème de Talagrand.

Soit P_t le semi-groupe d'Ornstein-Uhlenbeck de dimension n . En utilisant le Lemme de covariance et le fait que $\nabla P_t = e^{-t} \nabla P_t$, il s'ensuit que :

$$\text{Var}_\gamma(f) = \int_0^\infty e^{-t} \sum_{i=1}^n \mathbb{E}_\gamma[\partial_i f P_t \partial_i f] dt$$

Utilisons maintenant le Lemme précédent avec $h_i = \partial_i f$. Et si l'on pose

$$a_i = \frac{\|\partial_i f\|_{L^1(\gamma)}}{A_i},$$

plus le fait que $\tanh(t) \leq 1 - e^{-t}$ nous obtenons que

$$\begin{aligned} \text{Var}_\gamma(f) &\leq \sum_{i=1}^n A_i^2 \int_0^\infty e^{-t} a_i^{1-e^{-t}} dt \\ &= \sum_{i=1}^n A_i^2 \int_0^1 a_i^u du \leq C \sum_{i=1}^n \frac{A_i^2}{1 + \log(1/a_i)}, \end{aligned}$$

où C est une constante universelle. Ainsi le Théorème est démontré. \square

Nous allons voir à présent que la méthode de Talagrand peut se généraliser.

Théorème 3 (Version fonction monotone). — Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction absolument continue, monotone en chacune de ses coordonnées. On pose :

$$a := \frac{\text{Var}_\gamma(f)}{\mathbb{E}_\gamma |\nabla f|^2} \quad \text{et} \quad b := \frac{\sum_{i=1}^n \|\partial_i f\|_{L^1(\gamma)}^2}{\sum_{i=1}^n \|\partial_i f\|_{L^2(\gamma)}^2}.$$

Dans ce cas, a et b sont compris entre 0 et 1 et vérifient l'inégalité suivante :

$$b \leq a \leq \frac{C}{1 + \log(b^{-1/2})},$$

où C est une constante universelle. Ceci implique donc que a est petit si et seulement si b l'est.

Démonstration. — Pour chaque i , posons $u_i := \|\partial_i f\|_{L^1(\gamma)}^2$ et $v_i := \|\partial_i f\|_{L^2(\gamma)}^2$. On définit également

$$g(x) := \frac{1}{1 + \log(x^{-1/2})} = -\frac{2}{\log(x/e^2)}.$$

En calculant la dérivée seconde on montre facilement que $y \mapsto \frac{1}{\log(y)}$ est convexe sur l'intervalle $]0, e^{-2}[$. Ceci implique donc que la fonction $y \mapsto g(y)$ est concave sur $]0, 1[$. Ainsi le théorème de Talagrand nous fournit l'inégalité suivante :

$$\begin{aligned} \text{Var}_\gamma(f) &\leq C \sum_{i=1}^n \frac{\|\partial_i f\|_{L^2(\gamma)}^2}{1 + \log\left(\frac{\|\partial_i f\|_{L^2(\gamma)}}{\|\partial_i f\|_{L^1(\gamma)}}\right)} \\ &= C \sum_{i=1}^n g(u_i/v_i) v_i. \end{aligned}$$

Il faut noter que le second membre peut être égal à plus l'infini si nos dérivées partielles ne sont pas suffisamment intégrables. En utilisant l'inégalité de Jensen dans la relation précédente nous obtenons la majoration voulue.

$$\text{Var}_\gamma(f) \leq Cg \left(\frac{\sum_{i=1}^n (u_i/v_i)v_i}{\sum_{i=1}^n v_i} \right) \sum_{i=1}^n v_i = Cg(b) \mathbb{E}_\gamma[|\nabla f|^2].$$

Démontrons à présent la minoration. On pose pour tout i , $c_i := \mathbb{E}_\gamma[\partial_i f]$ et on définit la fonction suivante :

$$h(x) := f(x) - \sum_{i=1}^n c_i x_i.$$

On remarque alors que pour tout i ,

$$\text{Cov}_\gamma(h, x_i) = \text{Cov}_\gamma(f, x_i) - c_i,$$

puisque nos variables aléatoires sont des gaussiennes centrées réduites indépendantes de même loi et donc $\text{Cov}_\gamma(x_i, x_j) = \delta_{i,j}$.

De plus l'intégration par partie gaussienne nous fournit la relation suivante :

$$\text{Cov}_\gamma(f, x_i) = \mathbb{E}_\gamma[x_i f] = \mathbb{E}[\partial_i f] = c_i.$$

D'où, $\text{Cov}_\gamma(h, x_i) = 0$. C'est pourquoi nous obtenons, grâce à cette non corrélation, que :

$$\text{Var}_\gamma(f) = \text{Var}_\gamma \left(h + \sum_{i=1}^n c_i x_i \right) = \text{Var}_\gamma(h) + \sum_{i=1}^n c_i^2 \geq \sum_{i=1}^n c_i^2.$$

Finalement, en remarquant que, grâce à la monotonie de f on a l'égalité : $c_i^2 = \|\partial_i f\|_{L^1(\gamma)}^2$ pour tout i . Ceci nous fournit la minoration voulue. \square

5. Superconcentration et chaos

Il est temps de définir formellement la notion de superconcentration dont nous parlons depuis le début de ce mémoire. Nous constaterons, ci-dessous, que la notion de superconcentration est en fait équivalente à une notion de chaos. Ceci nous permet donc d'obtenir des théorèmes de superconcentration d'aux, ainsi que l'opportunité d'aborder un problème de deux façons différentes.

Dans cette section nous allons introduire deux nouvelles notions : celle de superconcentration et celle du chaos. Nous rappelons rapidement le cadre dans lequel nous nous plaçons.

On considère un processus de Markov réversible X_t et P_t le semi-groupe lui étant associé. On notera L le générateur infinitésimal de ce semi-groupe,

μ la mesure invariante et \mathcal{E} la forme de Dirichlet. On suppose également que le semi-groupe vérifie une inégalité de Poincaré de constante C_p .

Définition 10. — Une fonction f à valeurs réelle est dite ϵ -superconcentrée si

$$\text{Var}_\mu(f) \leq \epsilon C_p \mathcal{E}(f, f).$$

Il faut comprendre ceci de la manière suivante : ϵ est petit, par exemple $\epsilon = \epsilon_n \rightarrow 0$ lorsque $n \rightarrow \infty$. Ainsi f se concentre autour de sa moyenne.

Pourquoi parler de superconcentration ? Encore une fois, cette notion est là pour nous donner des résultats plus fins que la théorie de la concentration. Comme on le sait, l'inégalité très générale suivante :

$$\text{Var}\left(\max_{i=1,\dots,n} X_i\right) \leq \max_{i=1,\dots,n} \text{Var}(X_i),$$

où X_1, \dots, X_n sont des variables aléatoires gaussiennes standard que l'on ne suppose pas indépendantes. Cette inégalité est optimale et ne peut-être améliorée. Pour s'en convaincre, il suffit de considérer une seule variable aléatoire. Malheureusement, dans sa grande généralité, cette inégalité ne peut être très précise : le cas où les X_i sont indépendantes et identiquement distribuées suivant une gaussienne standard nous fournit la relation suivante :

$$\text{Var}\left(\max_{i=1,\dots,n} X_i\right) \leq 1.$$

Néanmoins en travaillant plus précisément, et si l'on note par $f_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ définit par $x \mapsto \max_{i=1,\dots,n} x_i$, on obtient que :

$$\text{Var}_{\gamma^n}(f_1) \leq \frac{C}{\log n}.$$

Dans un exemple aussi simple nous constatons que la théorie de la concentration nous fournit un résultat sous-optimale alors que nous devrions observer un phénomène de superconcentration avec $\epsilon = \frac{C}{\log n}$.

On supposera dans le reste de l'article que l'opérateur $-L$ admet une décomposition spectrale de valeurs propres $0 = \lambda_0 < \lambda_1 \leq \dots$ et de fonctions propres associées u_0, u_1, \dots . On peut alors réécrire la variance de f ainsi que son énergie de Dirichlet via cette décomposition spectrale.

$$\text{Var}_\mu(f) = \sum_{k \geq 1} \langle u_k, f \rangle^2, \quad \mathcal{E}(f, f) = \sum_{k \geq 1} \lambda_k \langle u_k, f \rangle^2.$$

On remarque, par ailleurs, que l'on a toujours la relation suivante :

$$\begin{aligned} \mathcal{E}(f, P_t f) = - \langle f, L e^{tL} f \rangle &= \sum_{k \geq 1} \lambda_k e^{-\lambda_k t} \langle u_k, f \rangle^2 \\ &\leq e^{-\lambda_1 t} \mathcal{E}(f, f). \end{aligned}$$

Ceci nous amène vers la définition de chaos. Lorsque l'inégalité ci-dessus est sous-optimale à partir d'un certain temps on dira que f est chaotique. Formellement on obtient la définition suivante.

Définition 11. — f est (ϵ, δ) -chaotique si pour tout $t \geq \delta$

$$\mathcal{E}(f, P_t f) \leq \epsilon e^{-\lambda_1 t} \mathcal{E}(f, f).$$

Il faut avoir à l'esprit, une fois de plus, que ϵ et δ sont petits. Typiquement ils dépendront de n et tendront vers 0 lorsque n tendra vers l'infini.

Nous allons donner une motivation quant à la définition de ces deux dernières notions. Nous avons parlé de l'hypercontractivité, ce phénomène peut-être un outil pour exhiber un chaos. Comme nous le verrons par la suite, il y a une équivalence entre la notion de chaos et de superconcentration. Ceci nous permettra donc d'aborder de deux manières différentes un modèle donné.

Heuristique 1. — Soient X , un vecteur gaussien centré de \mathbb{R}^n et Y , une copie indépendante de X . On pose $X^t = c_t X + d_t Y$. Soit A , un borelien de \mathbb{R}^n et notons par f la fonction indicatrice de A . Alors en utilisant l'inégalité de Cauchy-Schwarz et l'hypercontractivité avec $q = 2$ on obtient :

$$\begin{aligned} \mathbb{P}(X \in A \cap X^t \in A) = \mathbb{E}_\gamma[f P_t f] &\leq \|f\|_{L^2} \|P_t f\|_{L^2} \\ &\leq \|f\|_{L^2} \|f\|_{L^p} \\ &= \mathbb{P}(X \in A)^{1/2+1/p} \end{aligned}$$

On en déduit que :

$$\mathbb{P}(X^t \in A | X \in A) \leq \mathbb{P}(X \in A)^{1/2+1/p-1}.$$

Puisque $1/2 + 1/p - 1 > 0$ ceci implique donc que si $\mathbb{P}(X \in A)$ est petit alors $\mathbb{P}(X^t \in A | X \in A)$ sera également très petit.

Autrement dit la probabilité que X^t , qui n'est rien d'autre qu'une perturbation de X , se trouve au même endroit que X , sera très faible. Ceci illustre une idée de chaos comme l'on pourra le voir dans le modèle des polymères dirigés en milieu aléatoire : le chemin d'énergie maximum et le chemin d'énergie maximum que l'on obtient après perturbation sont quasiment disjoints.

Enonçons à présent le théorème d'équivalence entre superconcentration et chaos.

Théorème 4 (Equivalence). — *On conserve le cadre énoncé au début de la section ainsi que les notations établies ci-dessus.*

- Si f est ϵ -superconcentré alors pour tout $\delta > 0$ f est (ϵ', δ) -chaotique où $\epsilon' = \frac{\epsilon e^{\lambda_1 \delta}}{\lambda_1 \delta}$.
- Si f est (ϵ, δ) -chaotique alors f est ϵ' -superconcentré pour $\epsilon' = \epsilon + \lambda_1 \delta$.

Démonstration. — La démonstration du théorème repose sur le Lemme de covariance :

$$Var_\mu(f) = \int_0^\infty \mathcal{E}(f, P_t f) dt.$$

Voici en quelques mots les idées directrices de la démonstration.

1. En utilisant la décomposition spectrale on constate que l'application $t \mapsto \mathcal{E}(f, P_t f)$ est décroissante. Ainsi, puisque $Var_\mu(f)$ tend vers 0 (superconcentration) $\mathcal{E}(f, P_t f)$ doit faire de même.
2. Si l'on a un phénomène de chaos $\mathcal{E}(f, P_t f)$ tend vers 0 rapidement. Ceci implique donc que $Var_\mu(f)$ doit être également petite.

Démontrons à présent la première assertion du théorème.

Supposons que f soit ϵ -superconcentré. Alors le Lemme de covariance et la caractérisation de la constante optimale dans l'inégalité de Poincaré en terme spectral nous assurent que :

$$\frac{\epsilon}{\lambda_1} \mathcal{E}(f, f) \geq Var_\mu(f) = \int_0^\infty \mathcal{E}(f, P_t f) dt.$$

On se souvient également de la relation suivante :

$$\mathcal{E}(f, f) = -(f, Lf) = \sum_{k \geq 1} \lambda_k (u_k, f)^2,$$

et que $P_t = e^{tL}$. Il s'ensuit donc que :

$$(5) \quad \mathcal{E}(f, P_t f) = \sum_{k \geq 1} \lambda_k e^{-\lambda_k t} (u_k, f)^2.$$

En particulier $t \mapsto \mathcal{E}(f, P_t f)$ est une fonction positive décroissante. Donc pour n'importe quel $\delta > 0$,

$$\frac{\epsilon}{\lambda_1} \mathcal{E}(f, f) \geq Var_\mu(f) \geq \int_0^\delta \mathcal{E}(f, P_t f) dt \geq \delta \mathcal{E}(f, P_\delta f).$$

Ce qui entraîne que :

$$e^{\lambda_1 \delta} \mathcal{E}(f, P_\delta f) \leq \frac{\epsilon e^{\lambda_1 \delta}}{\lambda_1 \delta} \mathcal{E}(f, f).$$

De plus la formule (5) montre que $t \mapsto e^{\lambda_1 t} \mathcal{E}(f, P_t f)$ est également une fonction décroissante. En combinant ceci avec l'inégalité ci-dessus nous obtenons que pour tout $t \geq \delta$,

$$e^{\lambda_1 t} \mathcal{E}(f, P_t f) \leq \frac{\epsilon e^{\lambda_1 \delta}}{\lambda_1 \delta} \mathcal{E}(f, f).$$

Ceci prouve la première assertion du théorème.

Pour la deuxième partie de celui-ci, nous allons simplement utiliser la décroissance de la fonction $t \mapsto \mathcal{E}(f, P_t f)$. En effet, si l'on suppose que f est (ϵ, δ) -chaotique,

$$\begin{aligned} \text{Var}_\mu(f) &= \int_0^\infty \mathcal{E}(f, P_t f) dt \\ &\leq \int_0^\delta \mathcal{E}(f, P_t f) dt + \int_\delta^\infty \epsilon e^{-\lambda_1 t} \mathcal{E}(f, f) dt \\ &\leq \delta \mathcal{E}(f, f) + \frac{\epsilon}{\lambda_1} \mathcal{E}(f, f) = (\delta \lambda_1 + \epsilon) \frac{\mathcal{E}(f, f)}{\lambda_1}. \end{aligned}$$

Et ceci prouve que f est $(\epsilon + \delta \lambda_1)$ superconcentré. \square

PARTIE IV MODÈLES

6. Verres de spin

6.1. Introduction. — Dans cette partie nous allons étudier un modèle de physique statistiques introduit par D. Sherrington et S. Kirkpatrick (cf.(19)) Plus précisément nous allons montrer que ce modèle vérifie un phénomène de superconcentration. Pour arriver à ce résultat nous allons montrer que le modèle vérifie un principe de chaos pour ensuite utiliser le théorème d'équivalence entre ces deux phénomènes. Parmi les outils que nous avons déjà présentés, nous montrerons que nous ne pouvons malheureusement pas utiliser la méthode de Talagrand pour ce modèle. Nous ne pouvons même pas utiliser la version monotone de la méthode de Talagrand (3), car l'énergie libre $F_n(\beta)$ n'est pas une fonction monotone en ses coordonnées.

Présentons tout d'abord notre nouveau modèle. Pour donner une heuristique concernant ce modèle nous allons largement nous inspirer de la formidable introduction du livre de M. Talagrand (20).

Considérons une grande famille d'individu numérotés de 1 à N . Supposons que chaque individus connaisse le reste de la population considérée. Nous mesurons les sentiments d'un individu i envers un autre j par un nombre X_{ij} qui peut négatif ou positif. Supposons également une certaine symétrie : $X_{ij} = X_{ji}$, ainsi seuls les nombres $(X_{ij})_{i < j}$ sont pertinents. Nous voulons modéliser une situation où ces sentiments seraient aléatoires. Par souci de simplicité nous ferons l'hypothèse suivante : les $(X_{ij})_{i < j}$ sont des variables aléatoires réelles indépendantes. Bien entendu tout ceci n'est pas très réaliste, cependant ce modèle qui semble si simple de prime abord est en fait difficile à étudier.

Une des propriétés importantes de ce modèle est le fait que même si $X_{ij} > 0$ et que $X_{jk} > 0$. Autrement dit, si les individus i et j sont amis, tout comme les individus j et k . Et bien les individus i et k ont tout autant de chance d'être amis qu'ennemis. On parle dans ce cas de frustrations. Les interactions (X_{ij}) décrivent une situation sociale très complexe.

Pensons maintenant à une réalisation typique des variables aléatoires X_{ij} , c'est à dire un événement de probabilité proche de 1 (Qui se produit lorsque N est grand). Par exemple l'événement où la moitié des X_{ij} sont positives et l'autre moitié négatives est typique. Tandis que l'événement

où toutes ces variables aléatoires sont positives n'est certainement pas typique.

On se propose dès lors l'objectif suivant : séparer du mieux possible cette population en deux groupes de sorte que les amis restent ensemble et soient séparés de leurs ennemis. Ceci étant un moyen de soulager les tensions créées par les frustrations. Il apparaît rapidement que cette partition risque d'être effectuée de façon imparfaite : des amis seront séparés et des ennemis devront cohabiter. Pour introduire une façon quantitative de mesurer la qualité de notre partition il est pratique d'assigner à chaque individu i un nombre $\sigma_i \in \{-1, 1\}$, et ainsi définir deux classes différentes. Une des façons les plus simple de mesurer si nos deux classes rassemblent convenablement les amis tout en séparant les ennemis et la quantité suivante :

$$\sum_{i < j} X_{ij} \sigma_i \sigma_j$$

Essayer de rendre cette quantité grande nous entraîne à rendre les $X_{ij} \sigma_i \sigma_j$ positifs, et donc à prendre σ_i et σ_j de même signe lorsque $X_{ij} > 0$ sont de signe opposés sinon.

Malgré la simplicité de $\sum_{i < j} X_{ij} \sigma_i \sigma_j$, le problème d'optimisation revenant à déterminer le maximum de cette fonction pour une réalisation typique des X_{ij} est extrêmement difficile. De façon équivalente on peut également s'intéresser à la fonction suivante :

$$-\sum_{i < j} X_{ij} \sigma_i \sigma_j.$$

Trouver le minimum d'une telle fonction de configurations est appelé en physique un problème à température zéro, car à la température zéro un système est toujours dans sa configuration d'énergie minimale.

D'un point de vue physique les verres de spin modélisent le phénomène suivant : on considère des matériaux non magnétiques, comme le cuivre, dans lequel on introduit aléatoirement des impuretés magnétiques, du manganèse par exemple. On retrouve alors les deux propriétés essentielles des verres de spin :

- Le désordre dû à la répartition aléatoires des impuretés.
 - Les frustrations engendrées par les interactions des impuretés entre elles.
- Décrivons notre problème de façon plus formelle.

On considère n particules, chacune d'entre elles portant un spin valant $+1$ ou -1 . Une configuration de spin $\sigma = (\sigma_1, \dots, \sigma_n)$ est un élément de de

$\Sigma = \{-1; 1\}^n$. Soient $(X_{ij})_{1 \leq i < j \leq n}$ une collection de variables aléatoires gaussiennes standards indépendantes. Cette famille de *v.a.r* modélise le désordre de notre système.

Etant donné un tel désordre, nous pouvons définir l'énergie d'une configuration de spin σ par :

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

ainsi qu'une mesure de Gibbs sur Σ en associant à chaque configuration le poids suivant : $\exp(-\beta H_n(\sigma))$. Où β est l'inverse de la température, ainsi lorsque $\beta \rightarrow \infty$ cela correspond à la température du zéro absolu. Le poids que nous associons à chaque configuration se justifie physiquement par le fait qu'il permet de favoriser les configurations ayant une faible énergie.

Connaisant le désordre, la probabilité d'observer une configuration de spin $\sigma \in \Sigma$ est donc

$$\frac{\exp(-\beta H_n(\sigma))}{Z_n(\beta)}.$$

Où $Z_n(\beta) := \sum_{\sigma \in \Sigma} e^{-\beta H_n(\sigma)}$ est la fonction de partition associée au système (attention, malgré son nom, il s'agit simplement d'une constante de renormalisation pour définir une mesure de probabilité). Tout ceci définit le modèle des verres de spin de Sherrington et Kirkpatrick (SK).

En ayant défini l'énergie, nous introduisons l'objet suivant : l'énergie libre à « température inverse » $\beta \geq 0$

$$F_n(\beta) = -\frac{1}{\beta} \log(Z_n(\beta)).$$

Une autre quantité importante associée à ce modèle est le chevauchement (*overlap* en anglais) : ayant défini la mesure de Gibbs sur Σ , soient σ^1 et σ^2 deux configurations tirées de façon indépendante suivant la mesure de Gibbs. Le chevauchement entre σ^1 et σ^2 est défini par :

$$R_{1,2} := \frac{1}{n} \sum_{i=1}^n \sigma_i^1 \sigma_i^2.$$

Nous nous posons la question suivante : est ce que la mesure de Gibbs est chaotique sous de petites perturbations ? Autrement dit, si l'on considère un désordre perturbé $(X_{ij}^t)_{1 \leq i < j \leq n}$ (par ceci nous voulons signifier que, pour tout couple (i, j) , (X_{ij}, X_{ij}^t) est une variable aléatoire centrée gaussienne de covariance égale à e^{-t}) et que l'on définit une nouvelle mesure de Gibbs à partir des X_{ij}^t . Et que, à $\beta \geq 0$ fixé, l'on tire une configuration σ^1 suivant la

mesure de Gibbs initiale et σ^2 suivant la mesure de Gibbs perturbée, est ce que la fonction $F_n(\beta)$ est chaotique ?

6.2. Superconcentration via les polynômes d’Hermite. — Pour ce modèle nous allons mettre en oeuvre l’un des outils qui se trouve à notre disposition : la décomposition spectrale. L’idée que nous allons exploiter est la suivante : nous allons couper notre somme, obtenue par la décomposition spectrale, en deux, de manière astucieuse. Plus précisément nous allons utiliser la Proposition 4 et choisir judicieusement la valeur de m . Dans le cadre du semi-groupe d’Ornstein-Uhlenbeck, la Proposition 4 se reformule de la manière suivante :

$$\text{Var}_\gamma(f) \leq \sum_{k=1}^{m-1} \frac{\theta_k(f)}{k!} + \frac{\mathbb{E}_\gamma |\nabla f|^2}{m},$$

où $\theta_k(f)$ est défini par (4). Cette inégalité nous permettra de démontrer le Théorème suivant de superconcentration pour l’énergie libre $F_n(\beta)$ dans le modèle SK des verres de spin.

Théorème 5. — *Soit $F_n(\beta)$ l’énergie libre à température inverse β dans le modèle SK des verres de spin. Alors,*

$$\text{Var}(F_n(\beta)) \leq \frac{C_\beta n \log \log n}{\log n},$$

où C_β est constante ne dépendant que du paramètre β .

Démonstration. — L’idée principale de la démonstration est très simple : nous allons chercher à majorer les dérivées partielles de $F_n(\beta)$. Ceci nous permettra de majorer la fonction $\theta_k(F_n(\beta))$. Puis, en choisissant une valeur de m appropriée la Proposition 4 nous permettra de conclure.

Notons par $\langle \cdot \rangle$ l’espérance sous la mesure de Gibbs à température inverse β . Nous noterons par \mathbb{E} l’espérance totale : *i.e.* l’espérance sous la mesure de Gibbs et sous le désordre gaussien.

Soient $\sigma^1, \dots, \sigma^m$, m configurations *i.i.d.* de spin tirées sous la mesure de Gibbs. On constate que les σ^i sont indépendantes si l’on connaît le désordre gaussien, sinon ces variables aléatoires sont dépendantes. De plus, la symétrie du problème nous assure que la loi non conditionnelle de chaque σ^i est la loi uniforme sur l’hypercube $\Sigma = \{-1, 1\}^n$.

Lemme 5. — Fixons β et notons par $F_n := F_n(\beta)$ pour alléger les notations. Alors, pour tout $i_1, j_1, \dots, i_k, j_k$ nous avons :

$$(6) \quad \frac{\partial^k F_n}{\partial X_{i_1 j_1} \cdots \partial X_{i_k j_k}} = \frac{\beta^{k-1}}{n^{k/2}} \sum_{1 \leq l_1, \dots, l_k \leq k} c_k(l_1, \dots, l_k) \langle \sigma_{i_1}^{l_1} \sigma_{j_1}^{l_1} \cdots \sigma_{i_k}^{l_k} \sigma_{j_k}^{l_k} \rangle,$$

où les $c_k(l_1, \dots, l_k)$ sont des constantes vérifiant l'inégalité suivante, pour tout $1 \leq l_1, \dots, l_k \leq k$,

$$|c_k(l_1, \dots, l_k)| \leq (k-1)!.$$

Démonstration. — Nous allons démontrer ce Lemme par récurrence sur k . Pour $k = 1$ nous avons :

$$\frac{\partial F_n}{\partial X_{ij}} = \frac{-1 - \beta \sum_{\sigma \in \Sigma} \sigma_i \sigma_j e^{-\beta H_n(\sigma)}}{\beta \sqrt{n} Z_n},$$

où, rappelons le, $Z_n = \sum_{\sigma \in \Sigma} e^{-\beta H_n(\sigma)}$. Ceci peut se réécrire de la manière suivante :

$$\frac{\partial F_n}{\partial X_{ij}} = \frac{-1}{\sqrt{n}} \langle \sigma_i \sigma_j \rangle.$$

Par ailleurs, pour tout i, j , $|\langle \sigma_i, \sigma_j \rangle| \leq 1$, cela implique donc que :

$$|\nabla F|^2 = \sum_{i,j} \frac{1}{n} \langle \sigma_i, \sigma_j \rangle^2 \leq n.$$

Supposons que notre hypothèse soit vraie pour un certain k . Démontrons le au rang $k+1$. Il suffit alors de calculer

$$\frac{\partial}{\partial X_{i_{k+1} j_{k+1}}} \langle \sigma_{i_1}^{l_1} \sigma_{j_1}^{l_1} \cdots \sigma_{i_k}^{l_k} \sigma_{j_k}^{l_k} \rangle.$$

Après quelques calculs pénibles nous obtenons :

$$\frac{\partial}{\partial X_{i_{k+1} j_{k+1}}} \langle \sigma_{i_1}^{l_1} \sigma_{j_1}^{l_1} \cdots \sigma_{i_k}^{l_k} \sigma_{j_k}^{l_k} \rangle = \frac{\beta}{\sqrt{n}} \sum_{l_{k+1}=1}^{k+1} d_{k+1}(l_1, \dots, l_{k+1}) \langle \sigma_{i_1}^{l_1} \sigma_{j_1}^{l_1} \cdots \sigma_{i_{k+1}}^{l_{k+1}} \sigma_{j_{k+1}}^{l_{k+1}} \rangle,$$

où $|d_{k+1}(l_1, \dots, l_k)| \leq 1$ pour tout $1 \leq l \leq k$ et $|d_{k+1}(l_1, \dots, l_{k+1})| \leq k$. Ceci nous permet de conclure facilement notre démonstration par récurrence. \square

Terminons à présent la preuve du Théorème de superconcentration pour l'énergie libre dans le modèle des verres de spin de SK.

Soit $\tilde{X} = (\tilde{X}_{i,j})_{1 \leq i,j \leq n}$ un désordre gaussien, c'est à dire une nouvelle famille de variables aléatoires gaussiennes standards indépendantes. On suppose que ce nouveau désordre \tilde{X} est indépendant de la famille $X = (X_{i,j})_{1 \leq i,j \leq n}$ initiale. On tire alors k configurations de spin *i.i.d.* $\tilde{\sigma}^1, \dots, \tilde{\sigma}^k$ suivant la mesure de Gibbs induite par notre nouveau désordre \tilde{X} .

Les hypothèses d'indépendance nous permettent d'établir l'identité suivante :

$$(7) \quad \left(\mathbb{E}(\sigma_{i_1}^{l_1} \sigma_{j_1}^{l_1} \cdots \sigma_{i_k}^{l_k}) \right)^2 = \mathbb{E}(\sigma_{i_1}^{l_1} \sigma_{j_1}^{l_1} \cdots \sigma_{i_k}^{l_k}) \mathbb{E}(\tilde{\sigma}_{i_1}^{l_1} \tilde{\sigma}_{j_1}^{l_1} \cdots \tilde{\sigma}_{i_k}^{l_k})$$

$$(8) \quad = \mathbb{E}(\sigma_{i_1}^{l_1} \tilde{\sigma}_{i_1}^{l_1} \sigma_{j_1}^{l_1} \tilde{\sigma}_{j_1}^{l_1} \cdots \sigma_{i_k}^{l_k} \tilde{\sigma}_{i_k}^{l_k} \sigma_{j_k}^{l_k} \tilde{\sigma}_{j_k}^{l_k})$$

Avant d'utiliser le Lemme 6 pour conclure nous allons effectuer un calcul lorsque $k = 1$. Cela permettra de mieux comprendre les inégalités qui vont suivre sans se perdre dans la difficulté des notations.

Soient $X = (X_{ij})_{1 \leq i, j \leq n}$ un désordre gaussien et une configuration $\sigma = (\sigma_1, \dots, \sigma_n)$ tirée suivant la mesure de Gibbs induite par ce désordre. Soient à présent $\tilde{X} = (\tilde{X}_{ij})_{1 \leq i, j \leq n}$ une copie indépendante du désordre gaussien initial et $\tilde{\sigma} = (\tilde{\sigma}_1, \dots, \tilde{\sigma}_n)$ une nouvelle configuration tirée suivant ce nouveau désordre.

Observons que :

$$\frac{\partial F_n}{\partial X_{ij}} = -\frac{1}{\sqrt{n}} \langle \sigma_i \sigma_j \rangle = -\frac{1}{\sqrt{n}} \mathbb{E}[\sigma_i \sigma_j | X_{ij}].$$

Notamment, l'indépendance entre X et \tilde{X} nous assure que :

$$(9) \quad \mathbb{E}[\sigma_i \sigma_j | X_{ij}] = \mathbb{E}[\sigma_i \sigma_j | X_{ij}, \tilde{X}_{ij}].$$

Alors

$$\begin{aligned} \sum_{i, j=1}^n \left(\mathbb{E} \left[\frac{\partial F_n}{\partial X_{ij}} \right] \right)^2 &= \frac{1}{n} \sum_{i, j=1}^n \left(\mathbb{E} \left[\langle \sigma_i \sigma_j \rangle \right] \right)^2 \\ &= \frac{1}{n} \sum_{i, j=1}^n \left(\mathbb{E} \left[\mathbb{E}(\sigma_i \sigma_j | X_{i, j}) \right] \right)^2 \end{aligned}$$

Utilisons à présent le fait que :

$$\mathbb{E} \left[\mathbb{E}(\sigma_i \sigma_j | X_{i, j}) \right] = \mathbb{E} \left[\mathbb{E}(\tilde{\sigma}_i \tilde{\sigma}_j | \tilde{X}) \right].$$

En combinant tout ceci avec la remarque (9) nous trouvons que :

$$\begin{aligned}
\sum_{i,j=1}^n \left(\mathbb{E} \left[\frac{\partial F_n}{\partial X_{ij}} \right] \right)^2 &= \frac{1}{n} \sum_{i,j=1}^n \mathbb{E} \left[\mathbb{E}(\sigma_i \sigma_j | X) \right] \times \mathbb{E} \left[\mathbb{E}(\tilde{\sigma}_i \tilde{\sigma}_j | \tilde{X}) \right] \\
&= \frac{1}{n} \mathbb{E} \left[\mathbb{E} \left(\sum_{i=1}^n \sigma_i \tilde{\sigma}_j \right)^2 \middle| X, \tilde{X} \right] \\
&= \frac{1}{n} \mathbb{E} \left[\mathbb{E}(\sigma \cdot \tilde{\sigma})^2 \middle| X, \tilde{X} \right] \\
&= \frac{1}{n} \mathbb{E} \left[(\sigma \cdot \tilde{\sigma})^2 \right]
\end{aligned}$$

Dans les calculs qui vont suivre tout se déroule de la même manière, seuls les indices de notations compliquent légèrement l'écriture. Utilisons à présent le Lemme 6 et la majoration des $c_k(l_1, \dots, l_k)$.

$$\begin{aligned}
&\sum_{1 \leq i_1, j_1, \dots, i_k, j_k \leq n} \left(\mathbb{E} \frac{\partial^k F_n}{\partial g_{i_1 j_1} \cdots \partial g_{i_k j_k}} \right)^2 \\
&\leq \frac{\beta^{2k-2} [(k-1)!]^2}{n^k} \sum_{1 \leq l_1, \dots, l_k \leq k} \mathbb{E}((\sigma^{l_1} \cdot \tilde{\sigma}^{l_1})^2 \cdots (\sigma^{l_k} \cdot \tilde{\sigma}^{l_k})^2) \\
&\leq \frac{\beta^{2k-2} [(k-1)!]^2 k^k}{n^k} \mathbb{E}((\sigma^1 \cdot \tilde{\sigma}^1)^{2k}).
\end{aligned}$$

La dernière étape découle de l'inégalité de Hölder généralisée avec $p_1 = k, \dots, p_k = k$ et de la symétrie de notre modèle.

Puisque σ^1 et $\tilde{\sigma}^1$ sont indépendantes et uniformément distribuées sur l'hypercube, $\sigma^1 \cdot \tilde{\sigma}^1$ n'est rien d'autre qu'une somme de n variables aléatoires *i.i.d.* de Bernoulli. Ceci nous permet d'utiliser l'inégalité de Hoeffding (13) (c.f. Annexe), nous obtenons ainsi que :

$$\mathbb{E}((\sigma^1 \cdot \tilde{\sigma}^1)^{2k}) \leq k^k n^k,$$

Ainsi, pour tout $k \geq 1$,

$$\theta_k(F_n) \leq \beta^{2k-2} k^{4k}.$$

Ce résultat et le fait que $|\nabla F_n|^2 \leq n$ nous permet donc, grâce à la Proposition 4, de terminer la preuve en choisissant $m = c \log n / \log \log n$ avec c une constante suffisamment petite. □

Comme nous allons le voir, nous pouvons obtenir une inégalité plus précise et gagner un facteur « $\log \log n$ ». Pour cela nous allons avoir besoin de nouveaux outils et d'utiliser une méthode d'interpolation. Par la suite nous allons démontrer les deux résultats suivants :

Théorème 6. — Soit $R_{1,2}(t)$ défini comme ci dessous. Alors pour tout entier $k \geq 1$,

$$\mathbb{E}(R_{1,2}^{2k}(t)) \leq (C_1 k)^k n^{-C_2 k \min(1,t)},$$

où C_1 et C_2 sont des constantes positives ne dépendant que de β .

$$R_{1,2}(t) = \frac{1}{n} \sum_{i=1}^n \sigma_i^1 \sigma_i^2.$$

De plus, ceci nous permettra d'établir que l'énergie libre vérifie un phénomène de superconcentration.

Démonstration. — La démonstration de ce résultat est immédiate si l'on utilise le Théorème 9 que nous allons prouver grâce à une méthode d'interpolation. \square

Théorème 7. — Dans le modèle SK,

$$\text{Var}(F_n(\beta)) = o(n)$$

si et seulement si il existe $t_n \rightarrow 0$ lorsque $n \rightarrow \infty$ tel que

$$\mathbb{E}(R_{1,2}^2(t_n)) = o(1).$$

Démonstration. — Nous ne donnons que certains arguments de la démonstration, les détails sont laissés au lecteur en exercice.

Pour démontrer ce résultat nous allons utiliser le Théorème 4. Nous allons montrer que

$$- \text{Var}(F_n(\beta)) = o(n) \Leftrightarrow F_n \text{ est } \epsilon_n \text{ superconcentré, avec } \epsilon_{n \rightarrow \infty} \rightarrow 0.$$

$$- F_n \text{ est } (\epsilon_n, \delta_n) \text{ chaotique } \Leftrightarrow \text{il existe } t_n =_{n \rightarrow \infty} o(1) \text{ telle que } \mathbb{E}[R_{1,2}^2(t_n)] = o(1)$$

Supposons que F_n soit ϵ_n superconcentré :

$$\text{Var}(F_n) \leq \epsilon_n C \mathcal{E}(F_n, F_n) = \epsilon_n C \mathbb{E}[|\nabla f|^2] \leq \epsilon_n C n,$$

car nous avons déjà montré que $|\nabla F|^2 \leq n$.

Réciproquement, supposons que $\text{Var}(F_n) = o(n)$. Cela signifie qu'il existe une suite ϵ_n , où $\epsilon_n \rightarrow 0$ lorsque $n \rightarrow \infty$, telle que

$$\text{Var}(F_n) \leq \epsilon_n n.$$

Montrons le deuxième point.

De plus le choix de $t_n = \frac{1}{\log n}$, dans Théorème 6 avec $k = 1$, nous assure que $\mathbb{E}(R_{1,2}^2(t_n)) = o(1)$. D'après les équivalences précédentes, la fonction F_n est donc superconcentrée et

$$\text{Var}(F_n) \leq \frac{n}{\log n}.$$

□

6.3. Méthode d'interpolation. — Parmi les techniques que nous avons présentées auparavant, nous sommes en droit de nous demander si la méthode de Talagrand peut nous être d'une quelconque utilité. Malheureusement nous démontrerons que nous ne pouvons pas l'appliquer au modèle des verres de spin SK.

Pour obtenir les Théorèmes 6 et 7 nous allons procéder de façon différente en utilisant une méthode d'interpolation. Ceci nous permettra de prouver ce phénomène de chaos pour le modèle de verres de spin SK et de gagner le facteur « $\log \log n$ » dans le résultat de superconcentration. En fait, le Théorème est également valable pour des modèles plus généraux comme celui des p -spins ainsi que d'autres. Nous allons donc le démontrer dans toute sa généralité pour enfin montrer de quelle manière il s'applique à notre modèle des verres de spin.

Soit $\mathbf{X} = (X_1, \dots, X_n)$ un vecteur gaussien centré de matrice de covariance R . Soit Y une copie indépendante de X , et posons $X^t := c_t X + d_t Y$ comme précédemment. Notons par $[n]$ l'ensemble $\{1, \dots, n\}$.

On fixe un nombre $\beta \geq 0$. Et pour chaque $t, s \geq 0$ on définit une mesure de probabilité $G_{t,s}$ sur $[n] \times [n]$ qui assigne la masse

$$\frac{e^{\beta X_i^t + \beta X_j^s}}{\sum_{k,l} e^{\beta X_k^t + \beta X_l^s}}$$

au point (i, j) . L'espérance d'une fonction $h : [n] \times [n] \rightarrow \mathbb{R}$ sous la mesure $G_{t,s}$ sera notée par $\langle h \rangle_{t,s}$ défini par :

$$\langle h \rangle_{t,s} := \frac{\sum_{i,j} h(i,j) e^{\beta X_i^t + \beta X_j^s}}{\sum_{k,l} e^{\beta X_k^t + \beta X_l^s}}.$$

Nous considérons la matrice de covariance R comme une fonction $[n] \times [n]$ aussi bien qu'en tant que matrice carrée.

Dans un premier lieu, nous énonçons un Théorème qui nous servira à prouver notre résultat général.

Théorème 8. — *Supposons que $R(i, j) \geq 0$ pour tout i, j . Pour chaque i on pose :*

$$\nu_i := \mathbb{E} \left(\frac{e^{\beta X_i}}{\sum_j e^{\beta X_j}} \right).$$

Soit aussi $\phi(x) = \sum_k c_k x^k$ une série convergente à termes positifs. Alors pour tout $t \geq 0$,

$$\mathbb{E} \langle \phi \circ R \rangle_{0,t} \leq \inf_{s \geq t} \left(\mathbb{E} \langle \phi \circ R \rangle_{0,0} \right)^{1-t/s} \left(\sum_{i,j} \phi(R(i, j)) e^{2\beta^2 e^{-s} R(i, j)} \nu_i \nu_j \right)^{t/s}.$$

De plus, $t \mapsto \mathbb{E} \langle \phi \circ R \rangle_{0,t}$ est une fonction décroissante.

La démonstration de ce Théorème nécessite trois Lemmes. Le premier d'entre eux est simplement un résultat calculatoire. Le second met en relief l'un des ingrédients essentiels de la démonstration : nous constaterons que la fonction que nous cherchons à estimer est complètement monotone. Ceci nous permet via un Théorème de Bernstein de l'écrire comme une transformée de Laplace d'une mesure de probabilité. Cette dernière nous permettra alors de faire des majorations de type Hölder. Enfin le dernier Lemme met en oeuvre la méthode dite d'interpolation à l'aide des deux Lemmes précédents pour aboutir au résultat que l'on voulait démontrer.

Dans la suite nous noterons par $C_b^\infty(\mathbb{R}^n)$ les fonctions infiniment dérivables ayant toutes leurs dérivées bornées. Nous prolongeons également la définition de X^t au t négatifs : soit Z une copie indépendante de X et de Y on pose alors pour tout $t \leq 0$

$$X^{-t} := c_t X + d_t Z.$$

Lemme 6. — *Pour n'importe quelle fonction $f \in C_b^\infty(\mathbb{R}^n)$, on a :*

$$\frac{d}{dt} \mathbb{E}(f(X^{-t})f(X^t)) = -2c_t^2 \sum_{i,j} R(i, j) \mathbb{E}(\partial_i f(X^{-t}) \partial_j f(X^t))$$

.

Démonstration. — On pose $h^t := d_t X - c_t Y$ et $h^{-t} := d_t X - c_t Z$.

Il faut noter que la régularité de notre fonction f nous permet de dériver sous le signe intégrale. De plus nous avons des vecteurs gaussiens vérifiant les propriétés suivantes en loi : $h^t = h^{-t}$ et $X^t = X^{-t}$ et X^{-t} est indépendant du

couple (X^{-t}, Y) . Un simple calcul nous fournit alors les égalités suivantes :

$$\begin{aligned} \frac{d}{dt} \mathbb{E}(f(X^{-t})f(X^t)) &= -\frac{c_t}{dt} \mathbb{E} \left((h^{-t} \cdot \nabla f(X^{-t}))f(X^t) + (h^t \cdot \nabla f(X^t))f(X^{-t}) \right) \\ &= -\frac{2c_t}{dt} \mathbb{E} \left((h^{-t} \cdot \nabla f(X^{-t}))f(X^t) \right). \end{aligned}$$

De plus,

$$X^t = e^{-2t}X^{-t} + e^{-t}\sqrt{1-e^{-2t}}h^{-t} + \sqrt{1-e^{-2t}}Y.$$

Ainsi, pour tout i , l'intégration par partie gaussienne nous permet d'obtenir que

$$\mathbb{E}(h_i^{-t} \partial_i f(X^{-t})f(X^t)) = c_t dt \sum_j R(i, j) \mathbb{E}(\partial_i f(X^{-t}) \partial_j f(X^t)).$$

Et en combinant les deux dernières étapes nous démontrons la véracité de ce Lemme. \square

Démontrons à présent le deuxième Lemme. Nous rappelons qu'une démonstration du Théorème de Bernstein concernant les fonctions complètement monotones se trouve dans le livre de Feller (9).

Lemme 7. — Soit \mathcal{F} l'ensemble des fonctions u de $[0, \infty)$ qui peuvent être exprimées de la manière suivante :

$$u(t) = \sum_{i=1}^m e^{-c_i t} \mathbb{E}(f_i(X^{-t})f_i(X^t))$$

pour un entier $m \geq 1$ et des nombres positifs c_1, c_2, \dots, c_m ainsi que des fonctions $f_1, \dots, f_m \in C_b^\infty(\mathbb{R})$

Pour n'importe quelle fonction $u \in \mathcal{F}$ il existe une mesure de probabilité μ sur $[0, \infty)$ telle que pour tout $t \geq 0$,

$$u(t) = u(0) \int_{[0, \infty)} e^{-xt} d\mu(x).$$

En particulier, pour tout $0 < t \leq s$,

$$u(t) \leq u(0)^{1-t/s} u(s)^{t/s}.$$

Démonstration. — Il faut tout d'abord remarquer que n'importe quelle fonction $u \in \mathcal{F}$ doit être nécessairement positive. En effet X^{-t} et X^t sont *i.i.d.* conditionnellement à X , ce qui nous fournit

$$\mathbb{E}(f(X^{-t})f(X^t)) = \mathbb{E} \left(\mathbb{E}(f(X^t)|X)^2 \right).$$

Supposons que $u(0) = 0$, alors $u(t) = 0$ pour tout t et il n'y a rien à démontrer. Supposons à présent que $u(0) > 0$.

Puisque R est une matrice semi-définie positive, il existe une matrice carrée C , telle que $R = {}^t C C$. Ainsi, étant donné une fonction f , si l'on définit

$$w_i := \sum_j C_{ij} \partial_j f,$$

le Lemme précédent nous assure que :

$$\frac{d}{dt} \mathbb{E}(f(X^{-t})f(X^t)) = -2c_t^2 \sum_i \mathbb{E}(w_i(X^{-t})w_i(X^t)).$$

Ceci nous montre, au vu de la définition de \mathcal{F} , que si $u \in \mathcal{F}$ alors $-u' \in \mathcal{F}$. En procédant à une récurrence évidente nous observons que pour tout $k \geq 0$, $(-1)^k u^{(k)}$ est fonction positive. De telles fonctions sur $[0, \infty)$ sont dites complètement monotones (c.f. (9)). Comme nous l'avons déjà signalé, la propriété la plus importante de telles fonctions u est qu'elles peuvent être représentées en tant que transformée de Laplace d'une mesure Borélienne ν sur $[0, \infty)$. Autrement dit :

$$u(t) = \int_{[0, \infty)} e^{-xt} d\nu(x).$$

De plus, $u(0) = \nu(\mathbb{R})$. En posant $\mu(dx) = u(0)^{-1} \nu(dx)$, nous prouvons la première assertion du Lemme. Quant à la seconde, il suffit de remarquer qu'avec l'inégalité de Hölder, pour $p = s/t$ et $q = 1 - 1/p$, nous obtenons que pour tout $0 < t \leq s$

$$\int_{\mathbb{R}} e^{-xt} d\mu(x) \leq \left(\int_{\mathbb{R}} e^{-xs} d\mu(x) \right)^{t/s} = (u(s)/u(0))^{t/s}.$$

Et ceci conclut notre démonstration. \square

Le Lemme suivant, qui nous permettra enfin de démontrer notre Théorème, est obtenu comme variante d'une méthode d'interpolation Gaussienne pour analyser le modèle des verres de spin à haute température. L'idée est similaire à la preuve non publiée de R. Latala sur la solution des répliques similaires dans le modèle des verres de spin SK.

Lemme 8. — Soient ϕ et μ_i comme dans le Théorème 8. Alors pour tout $t \geq 0$,

$$\mathbb{E} \langle \phi \circ R \rangle_{0,t} \leq \sum_{i,j} \phi(R(i,j)) e^{2\beta^2 e^{-2t} R(i,j)} \mu_i \mu_j.$$

Démonstration. — Pour tout i , on définit la fonction suivante $p_i : \mathbb{R}^n \rightarrow \mathbb{R}$ par

$$p_i(x) := \frac{e^{\beta x_i}}{\sum_j e^{\beta x_j}}.$$

On remarque alors que :

$$\partial_j p_i = \beta(p_i \delta_{i,j} - p_i p_j),$$

Puisque p_i est bornée ceci prouve qu'en particulier $p_i \in C_b^\infty(\mathbb{R}^n)$.

Soient r un entier positif. Comme R est une matrice semi-définie positive, $R^{(r)} = (R(i, j)^r)_{1 \leq i, j \leq n}$ l'est également. C'est pourquoi il existe une matrice $C^{(r)} = (C(i, j)^r)_{1 \leq i, j \leq n}$ telle que $R^{(r)} = {}^t(C^{(r)})C^{(r)}$.

On pose également :

$$w_i := \sum_j C_{i,j}^{(r)} p_j, \quad i \in \{1, \dots, n\}.$$

Nous adopterons par la suite la convention suivante qui allégera les notations : pour tout $s \in \mathbb{R}$ $p_i^s := p_i(X^s)$ et $h_i^s := w_i(X^s)$. Soit alors

$$f_r(t) := \mathbb{E} \left(\sum_{i,j} R(i, j)^r p_i^{-t} p_j^t \right) = \mathbb{E} \left(\sum_i w_i^{-t} w_i^t \right).$$

Le Lemme 6 nous permet alors d'obtenir la relation suivante :

$$\begin{aligned} f_r'(t) &= -2c_t^2 \sum_i \sum_{k,l} R(k, l) \mathbb{E}(\partial_k w_i^{-t} \partial_l w_i^t) \\ &= -2\beta^2 c_t^2 \sum_{i,k,l} R(k, l) \mathbb{E} \left[\left(C_{ik}^{(r)} p_k^{-t} - \sum_j C_{ij}^{(r)} p_j^{-t} p_k^{-t} \right) \right. \\ &\quad \times \left. \left(C_{il}^{(r)} p_l^{-t} - \sum_j C_{ij}^{(r)} p_j^{-t} p_l^{-t} \right) \right] \\ &= -2\beta^2 c_t^2 \mathbb{E} \left(\sum_{k,l} R(k, l)^{r+1} p_k^{-t} p_l^t - \sum_{j,k,l} R(k, l) R(j, l)^r p_j^{-t} p_k^{-t} p_l^t \right. \\ &\quad \left. - \sum_{j,k,l} R(k, l) R(k, j)^r p_k^{-t} p_j^t p_l^t + \sum_{j,k,l,m} R(k, l) R(j, m)^r p_j^{-t} p_k^{-t} p_m^t p_l^t \right). \end{aligned}$$

Notre objectif est de minorer $f_r'(t)$. Pour cela nous pouvons déjà supprimer les deux termes du milieu dans l'expression ci-dessus car ils ont tous les deux une contribution positive.

Pour le quatrième terme, nous allons utiliser l'inégalité de Hölder avec

$p = r + 1$ et $q = \frac{r+1}{r}$. En effet :

$$\begin{aligned} & \left(\sum_{k,l} R(k,l)^r p_k^{-t} p_l^t \right) \left(\sum_{j,m} R(j,m)^r p_j^{-t} p_m^t \right) \\ & \leq \left(\sum_{k,l} R(k,l)^{r+1} p_k^{-t} p_l^t \right)^{\frac{1}{r+1}} \left(\sum_{j,m} R(j,m)^{r+1} p_j^{-t} p_m^t \right)^{\frac{r}{r+1}} \\ & = \left(\sum_{k,l} R(k,l)^{r+1} p_k^{-t} p_l^t \right) \end{aligned}$$

Ainsi la majoration précédente et le Lemme 7 nous fournit l'inégalité suivante :

$$(10) \quad 0 \leq f'_r(t) \leq -4\beta^2 c_t^2 \mathbb{E} \left(\sum_{k,l} R(k,l)^{r+1} p_k^{-t} p_l^t \right)$$

$$(11) \quad = -4\beta^2 c_t^2 f_{r+1}(t).$$

Posons à présent : $u_r(x) := f_r(-\log \sqrt{x})$ pour $0 < x < 1$. Ceci nous permet d'exprimer u'_r en fonction de f'_r , i.e. :

$$u'_r(x) = -\frac{f'_r(-\log \sqrt{x})}{2x}.$$

L'inégalité 10 devient alors

$$(12) \quad 0 \leq u'_r(x) \leq 2\beta^2 u_{r+1}(x).$$

Fixons $0 < x < 1$, $r \geq 1$. Définissons T_m , pour tout $m \geq 1$, par :

$$T_m := \int_0^x \int_0^{x_1} \cdots \int_0^{x_{m-1}} (2\beta^2)^{m-1} u'_{r+m-1}(x_m) dx_m dx_{m-1} \cdots dx_1.$$

En utilisant la relation 12 et le Théorème fondamental de l'analyse, nous obtenons :

$$\begin{aligned} 0 \leq T_m & \leq \int_0^x \int_0^{x_1} \cdots \int_0^{x_{m-1}} (2\beta^2)^m u'_{r+m}(x_m) dx_m dx_{m-1} \cdots dx_1. \\ & = \int_0^x \cdots \int_0^{x_{m-1}} (2\beta^2)^m \left(u_{r+m}(0) + \int_0^{x_m} u'_{r+m}(x_{m+1}) dx_{m+1} \right) dx_m \cdots dx_1. \\ & = \frac{(2\beta^2)^m u_{r+m}(0) x^m}{m!} + T_{m+1}. \end{aligned}$$

En procédant par récurrence, ceci implique que, pour n'importe quel $m \geq 1$,

$$u_r(x) = u_r(0) + T_1 \leq \sum_{l=0}^m \frac{u_{r+l}(0) (2\beta^2 x)^l}{l!} + T_m + 1.$$

Montrons que nous pouvons faire tendre m vers l'infini dans l'inégalité précédente. De nouveau, pour $m \geq 2$, nous obtenons

$$\begin{aligned} 0 \leq T_m &\leq \int_0^x \int_0^{x_1} \cdots \int_0^{x_{m-1}} (2\beta^2)^m u'_{r+m}(x_m) dx_m dx_{m-1} \cdots dx_1 \\ &\leq \frac{M^{r+m} (2\beta^2 x)^m}{m!}, \end{aligned}$$

où $M := \max_{i,j} R(i,j)$. Rappelons le,

$$u_r(x) := f_r(-\log \sqrt{x}) = \mathbb{E} \left(\sum_{i,j} R(i,j)^r p_i^{\log \sqrt{x}} p_j^{-\log \sqrt{x}} \right).$$

Ainsi, $\lim_{n \rightarrow \infty} T_m = 0$. De plus, $p_i^\infty = X$ est indépendant de $p_i^{-\infty} = Z$. Ce qui implique alors, pour tout m ,

$$u_m(0) = f_m(\infty) = \sum_{i,j} R(i,j)^m \nu_i \nu_j.$$

Si nous rassemblons tous les résultats obtenus, nous concluons donc que :

$$u_r(x) \leq \sum_{l=0}^{\infty} \frac{u_{r+l}(0) (2\beta^2 x)^l}{l!} = \sum_{i,j} R(i,j)^r e^{2\beta^2 x R(i,j)} \nu_i \nu_j.$$

Il suffit alors de choisir $x = e^{-2t}$. Puis, de sommer sur r en utilisant le fait que les coefficients de ϕ sont positifs et conservent le sens de l'inégalité précédente. Ces calculs terminés, nous obtenons le résultat annoncé par le Lemme. \square

Maintenant que ces Lemmes ont été prouvés, nous pouvons aborder la démonstration du Théorème 8.

Théorème 8. — Soit p_i^t défini comme dans le Lemme 8. Comme nous l'avons fait remarquer précédemment la matrice $(R(i,j)^r)_{i,j}$ est semi-définie positive pour tout entier $r \in \mathbb{N}$. Etant donné que ϕ ne possède que des coefficients positifs dans son développement en série, il s'ensuit que la matrice

$$\Phi := \left(\phi(R(i,j)) \right)_{i,j}$$

est également semi-définie positive. Il existe donc une matrice $C = (C_{i,j})_{i,j}$ telle que $\Phi = {}^t C C$. C'est pourquoi :

$$\langle \phi \circ R \rangle_{-s,s} = \sum_{i,j} \phi(R(i,j)) p_i^{-s} p_j^s = \sum_j \left(\sum_i C_{ij} p_j^{-s} \right) \left(\sum_j C_{ij} p_j^s \right).$$

Ainsi la fonction

$$u(t) := \mathbb{E} \langle \phi \circ R \rangle_{-t,t}$$

appartient à la classe \mathcal{F} définie dans le Lemme 7. Comme nous avons rassemblé les difficultés techniques dans les Lemmes précédents, la démonstration du Théorème est quasiment achevée. Il suffit de combiner le Lemme 7 avec le Lemme 8 et l'observation suivante, découlant du fait que $(X^{-t/2}, X^{t/2}) = (X^0, X^t)$ en loi :

$$u(t/2) = \mathbb{E} \langle \phi \circ R \rangle_{0,t}$$

pour obtenir la première assertion du Théorème :

$$\mathbb{E} \langle \phi \circ R \rangle_{0,t} \leq \inf_{s \geq t} \left(\mathbb{E} \langle \phi \circ R \rangle_{0,0} \right)^{1-t/s} \left(\sum_{i,j} \phi(R(i,j)) e^{2\beta^2 e^{-s} R(i,j)} \nu_i \nu_j \right)^{t/s}.$$

L'affirmation que $t \mapsto u(t)$ est une fonction décroissante provient simplement du fait que $u' \leq 0$. \square

Nous allons enfin pouvoir démontrer un nouveau Théorème de superconcentration pour le modèle des verres des spin SK. Comme nous l'avons déjà signifié, ce Théorème est valable pour une plus grande classe de modèle. Néanmoins, pour ne pas obscurcir la démonstration, nous allons uniquement traiter le cas du modèle des verres de spin en spécifiant toutefois, à la fin de démonstration, comment faire pour obtenir le Théorème dans toute sa généralité.

Théorème 9. — *Considérons le modèle de verres de spin SK. Nous avons l'inégalité suivante :*

$$\mathbb{E} \left\langle \psi \left(\frac{\sigma \cdot \sigma'}{n} \right)^k \right\rangle_{0,t} \leq (Ck)^k n^{-k \min\{1, t/C \log(1+C\beta)\}},$$

où $\psi(x) = x^2$.

Démonstration. — Rappelons que $H_n(\sigma) = \sum_{i,j=1}^n X_{ij} \sigma_i \sigma_j$ avec $\sigma \in \{-1, 1\}^n$ où g_{ij} sont des variables aléatoires gaussiennes standard *i.i.d.*. On remarque en premier lieu, que $(H_n(\sigma))_{\sigma \in \{-1, 1\}^n}$ est un vecteur gaussien centré et de fonction de covariance :

$$R(\sigma, \sigma') = \frac{1}{n} \sum_{i,j=1}^n \sigma_i \sigma_j \sigma'_i \sigma'_j = \frac{1}{n} (\sigma \cdot \sigma')^2 = n \psi \left(\frac{\sigma \cdot \sigma'}{n} \right)$$

Par symétrie du modèle on remarque que, pour tout σ ,

$$(13) \quad \nu_\sigma = \mathbb{E} \left(\frac{e^{-\beta H_n(\sigma)}}{\sum_{\sigma'} e^{-\beta H_n(\sigma')}} \right) = 2^{-n}.$$

Des arguments provenant de la théorie de la concentration, que nous détaillerons dans un Lemme, nous permettent de montrer l'existence de

constantes positives C et γ , indépendantes de n , telles que la relation suivante soit vérifiée :

$$(14) \quad 2^{-2n} \sum_{\sigma, \sigma'} \frac{(\sigma \cdot \sigma')^{2k}}{n^k} \exp\left(\frac{\gamma(\sigma \cdot \sigma')^2}{n}\right) \leq (Ck)^k,$$

ceci pour tout $k \in \mathbb{N}_*$ et pour tout $n \in \mathbb{N}$. Pour démontrer notre résultat nous allons utiliser le Théorème 8. Pour cela nous choisissons la fonction $\phi(x) = \frac{x^k}{n^k}$. Il s'agit bien d'une fonction dont les coefficients du développement en série sont positifs : ici $c_i = \frac{1}{n^k} \delta_{ik}$.

Le Théorème 8 nous fournit l'inégalité suivante :

$$\begin{aligned} \mathbb{E}[\langle \phi \circ R \rangle_{0,t}] &= \mathbb{E}\left[\left\langle \psi\left(\frac{\sigma \cdot \sigma'}{n}\right)^k \right\rangle_{0,t}\right] \\ &\leq \left(\mathbb{E}\left[\left\langle \psi\left(\frac{\sigma \cdot \sigma'}{n}\right)^k \right\rangle_{0,t}\right]\right)^{1-t/s} \\ &\quad \times \left(\sum_{\sigma, \sigma'} \psi\left(\frac{\sigma \cdot \sigma'}{n}\right)^k e^{2\beta^2 e^{-s} \psi\left(\frac{\sigma \cdot \sigma'}{n}\right)} \nu_i \nu_j\right)^{t/s} \end{aligned}$$

Autrement dit

$$\begin{aligned} \mathbb{E}\left[\left\langle \psi\left(\frac{\sigma \cdot \sigma'}{n}\right)^k \right\rangle_{0,t}\right] &\leq \left(\mathbb{E}\left[\left\langle \psi\left(\frac{\sigma \cdot \sigma'}{n}\right)^k \right\rangle_{0,t}\right]\right)^{1-t/s} \\ &\quad \times \left(\sum_{\sigma, \sigma'} \left(\frac{\sigma \cdot \sigma'}{n}\right)^{2k} e^{2\beta^2 e^{-s} \left(\frac{\sigma \cdot \sigma'}{n}\right)^2} \nu_\sigma \nu_{\sigma'}\right)^{t/s} \end{aligned}$$

Nous allons à présent, estimer séparément chacun des deux termes du membre de droite.

Choisissons s tel que $2\beta^2 e^{-s} = \min(\gamma, 2\beta^2)$. Ainsi

$$\exp\left(2\beta^2 e^{-s} \frac{(\sigma \cdot \sigma')^2}{n}\right) \leq \exp\left(\gamma \frac{(\sigma \cdot \sigma')^2}{n}\right).$$

En outre, pour tout $\sigma \in \{-1, 1\}^n$, $\nu_\sigma = 2^{-n}$ donc $\nu_\sigma \nu_{\sigma'} = 2^{-2n}$. En conclusion :

$$\left(\sum_{\sigma, \sigma'} \left(\frac{\sigma \cdot \sigma'}{n}\right)^{2k} e^{2\beta^2 e^{-s} \left(\frac{\sigma \cdot \sigma'}{n}\right)^2} \nu_\sigma \nu_{\sigma'}\right)^{t/s} \leq \left[2^{-2n} \sum_{\sigma, \sigma'} \frac{(\sigma \cdot \sigma')^{2k}}{n^{2k}} e^{\gamma \frac{(\sigma \cdot \sigma')^2}{n}}\right]^{t/s}$$

En utilisant maintenant la relation 14. Celle ci nous fournit le résultat suivant :

$$\left[2^{-2n} \sum_{\sigma, \sigma'} \frac{(\sigma \cdot \sigma')^{2k}}{n^{2k}} e^{\gamma \frac{(\sigma \cdot \sigma')^2}{n}}\right]^{t/s} \leq [(Ck)^k n^{-k}]^{t/s}.$$

Traisons l'autre terme. Par définition

$$\left\langle \frac{(\sigma \cdot \sigma')^{2k}}{n^{2k}} \right\rangle_{0,0} = \frac{\sum_{\sigma, \sigma'} \frac{(\sigma \cdot \sigma')^{2k}}{n^{2k}} \exp(-\beta H_n(\sigma) - \beta H_n(\sigma'))}{\sum_{\sigma, \sigma'} \exp(-\beta H_n(\sigma) - \beta H_n(\sigma'))}.$$

De plus, $\sigma \cdot \sigma' \leq n$. Ceci nous assure que :

$$\mathbb{E} \left[\left\langle \frac{(\sigma \cdot \sigma')^{2k}}{n^{2k}} \right\rangle_{0,0} \right] \leq 1.$$

En conclusion, si $0 \leq t \leq s$. Nous obtenons que :

$$\mathbb{E} \left[\left\langle \psi \left(\frac{\sigma \cdot \sigma'}{n} \right)^k \right\rangle_{0,t} \right] \leq (Ck)kn^{-kt/s}.$$

Il reste alors à considérer le cas où $t \geq s$. Le fin de la démonstration est laissée en exercice. \square

Lemme 9 (Démonstration du Lemme). — *En conservant le cadre et les notations précédentes, il existe des constantes γ et C positives, indépendantes de n , telles que :*

$$2^{-2n} \sum_{\sigma, \sigma'} \frac{(\sigma \cdot \sigma')^{2k}}{n^k} \exp\left(\frac{\gamma(\sigma \cdot \sigma')^2}{n}\right) \leq (Ck)^k,$$

pour tout n et tout $k \geq 0$.

Démonstration. — Reformulons tout ceci. En premier lieu, notons que

$$\sigma_i \cdot \sigma'_i = \eta_i \quad \text{en loi,}$$

où η_i suit une loi de Bernoulli symétrique sur $\{-1, 1\}$. Il nous suffit alors de montrer que

$$\mathbb{E} \left[\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i \right)^{2k} e^{\gamma \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i \right)^2} \right] \leq (Ck)^k,$$

pour démontrer le Lemme. Intuitivement ce résultat peut s'obtenir en comparant notre somme, renormalisée, de variables aléatoires de Bernoulli avec des gaussiennes standard. Démontrons le.

On considère X , une variable aléatoire gaussienne centrée réduite. On pose

$$I_{2k} := \mathbb{E} \left[X^{2k} e^{\gamma X^2} \right].$$

Remarquons que l'on peut réécrire ceci de la manière suivante :

$$I_{2k} = \int_0^\infty \left(t^{2k-1} e^{\gamma t^2} + \gamma t^{2k+1} e^{\gamma t^2} \right) \mathbb{P}(X \geq t) dt.$$

Rappelons que la mesure gaussienne standard vérifie une inégalité de concentration :

$$\mathbb{P}(X \geq t) \leq e^{-t^2/2}.$$

Ainsi, nous obtenons que :

$$I_{2k} \leq 2k \int_0^\infty t^{2k-2} t e^{-\frac{t^2}{2}(1-2\gamma)} dt + \gamma \int_0^\infty t^{2k} t e^{-\frac{t^2}{2}(1-2\gamma)} dt.$$

Notons par $J_{2k} := \int_0^\infty t^{2k} t e^{-\frac{t^2}{2}(1-2\gamma)} dt$. Une intégration par partie gaussienne nous fournit la relation de récurrence suivante :

$$J_{2k} = \frac{2k}{1-2\gamma} J_{2k-2}.$$

Ce qui nous permet de trouver que :

$$J_{2k} = \frac{(2k) \times \dots \times 2}{(1-2\gamma)^k} J_0 = \frac{2^k k!}{(1-2\gamma)^{k+1}}.$$

En choisissant, par exemple, $\gamma = 1/4$, nous obtenons alors :

$$J_{2k} \leq (Ck)^k,$$

où C est une constante positive universelle. Nous obtenons donc le résultat pour les variables aléatoires gaussiennes : $I_{2k} \leq (Ck)^k$.

Il se trouve que nos variables aléatoires de Bernoulli vérifient également une inégalité de concentration. Adoptons la notation suivante : $S_n := \frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i$. L'inégalité de Hoeffding (13) nous fournit :

$$\mathbb{P}(S_n \geq t) \leq e^{-t^2/2}.$$

Nous pouvons alors reprendre la démonstration que nous avons faites pour les variables aléatoires gaussiennes mot pour mot. Et ceci nous permet de conclure la démonstration de ce Lemme. \square

Remarque 9. — *Expliquons comment faire pour obtenir ce même Théorème pour des modèles autre que celui des verres de spin de Sherrington et Kirkpatrick.*

Dans un cadre plus général, nous devons considérer un vecteur gaussien centré $(H_n(\sigma))_{\sigma \in \{-1,1\}^n}$ de fonction de covariance :

$$\text{Cov}(H_n(\sigma), H_n(\sigma')) = n\Psi\left(\frac{\sigma \cdot \sigma'}{n}\right).$$

Où Ψ est une fonction sur $[-1, 1]$ indépendante de n . Dans la démonstration précédente $\Psi(x) = x^2$. L'important est qu'il existe une constante $c > 0$ telle que $\Psi(x) \leq cx^2$.

Soit H'_n une copie indépendante de H_n , pour tout $t \geq 0$ on pose :

$$H_n^t := c_t H_n + d_t H'_n.$$

Etant donnée une fonction h sur $\{-1, 1\}^n \times \{-1, 1\}^n$, nous pouvons définir $\langle h(\sigma, \sigma') \rangle_{t,s}$ comme la moyenne de h par rapport à la mesure de Gibbs définie par H_n^t et H_n^s . C'est à dire :

$$\langle h(\sigma, \sigma') \rangle_{t,s} := \frac{\sum_{\sigma, \sigma'} h(\sigma, \sigma') e^{-\beta H_n^t(\sigma) - \beta H_n^s(\sigma')}}{\sum_{\sigma, \sigma'} e^{-\beta H_n^t(\sigma) - \beta H_n^s(\sigma')}}.$$

La démonstration reste la même, en modifiant seulement la définition de s de la façon suivante : s est choisi de sorte à ce que $2\beta^2 c e^{-s} = \min(\gamma, 2\beta^2 c)$.

En particulier cette version générale du Théorème précédent couvre le modèle des p - spins avec $\Psi(x) = x^p$.

7. NK landscape

Nous allons utiliser la méthode de Talagrand pour obtenir une inégalité sur le modèle NK - *landscape*. Il s'agit d'un modèle de génétique introduit par S. A. Kauffman et S. A. Levin (cf. (14)).

Tout d'abord, il convient de garder à l'esprit que ce modèle présente de grandes similitudes avec celui des verres de spin que nous avons abordé plus tôt (cf. 6). Bien que la méthode de Talagrand ne s'applique pas au cas du modèle des verres de spin.

Pour avoir une vision d'ensemble du modèle NK - *landscape* décrivons le en quelques mots. On s'intéresse à un génome, lui-même composés d'un certain nombre de gènes, eux mêmes composés d'un allèle. Par souci de simplicité nous supposons qu'il n'y a que deux choix possible d'allèle différent.

Les différents gènes qui composent notre génome interagissent entre eux de manière aléatoire. On cherche alors à mesurer la compatibilité d'un gène donné avec ses gènes voisins. En supposant que l'on puisse additionner ces mesures de compatibilité on veut définir une fonction permettant de mesurer la compatibilité globale d'un génome. Nous nous intéresserons ensuite à l'espérance et à la variance du maximum de cette fonction.

Retranscrivons tout ceci de manière formelle.

On note par $S = \{0, 1\}^N$ l'espace des génomes. Un élément $\sigma = (\sigma_1, \dots, \sigma_N) \in$

S est donc composé de N gènes, eux mêmes composé d'un allèle : 0 ou 1. Le paramètre K détermine le nombre d'interaction entre un gène et ses voisins : par exemple, à K fixé, σ_1 interagit avec $\sigma_2, \dots, \sigma_{K+1}$. Les autres interactions sont aussi simples que possible : *i.i.d.*.

Comme nous l'avons spécifié quelques lignes plus haut, ces interactions sont modélisées par des variables aléatoires gaussiennes indépendantes de même loi : $\mathcal{N}(0, 1)$.

Formellement, on se donne une famille de gaussiennes standard *i.i.d.* $Y(i, \eta)$ où $0 \leq i \leq N - 1$ et $\eta \in \{0, 1\}^{K+1}$.

On définit alors la fonction qui mesure la compatibilité d'un génome donné par

$$F(\sigma) = \sum_{i=0}^{N-1} Y(i, (\sigma_i, \dots, \sigma_{i+K})).$$

Une façon naturelle de mesurer l'écart entre deux génomes, $\sigma, \sigma' \in S$, est d'utiliser la distance de Hamming :

$$P_{N,K}(\sigma, \sigma') = |\{i : (\sigma_i, \dots, \sigma_{i+K}) = (\sigma'_i, \dots, \sigma'_{i+K})\}|.$$

A noter que $P_{N,K}(\sigma, \sigma') = Cov(F(\sigma), F(\sigma'))$.

Bien que cela ne donne pas plus d'information sur la variance du maximum de $F(\sigma)$, on peut néanmoins calculer simplement un encadrement de l'espérance de ce maximum.

Lemme 10. — *Indépendamment de la valeur de K , nous avons :*

$$\frac{N}{\sqrt{\pi}} \leq \mathbb{E}(\max_{\sigma \in S} F(\sigma)) \leq N\sqrt{2 \log 2}.$$

Remarque 10. — *En fait, ces bornes sont optimales. La borne inférieure est atteinte lorsque $K = 0$ et la borne supérieure est atteinte (asymptotiquement) lorsque $K = N - 1$. De plus comme nous allons le constater dans la preuve, l'espérance est une fonction décroissante de K .*

Démonstration. — L'obtention de la borne supérieure est directe en utilisant le Lemme de Sudakov (Cf. Annexe 13). Pour la borne inférieure, on observe tout d'abord que si $G(\sigma)$ est une autre fonction mesurant la compatibilité d'un génome (pour un modèle NK avec $K = 0$, alors pour $\sigma, \sigma' \in S$ nous avons :

$$Cov(G(\sigma), G(\sigma')) \geq Cov(F(\sigma), F(\sigma')).$$

Ceci découle du fait que $K \mapsto P_{N,K}(\sigma, \sigma')$ est une fonction décroissante.

De plus $\text{Var}(F(\sigma)) = \text{Var}(G(\sigma)) = N$. On peut donc appliquer le Lemme de Slepian qui nous dit que :

$$\mathbb{E}(\max_{\sigma} F(\sigma)) \geq \mathbb{E}(\max_{\sigma} G(\sigma)).$$

Maintenant, si $(Z(i, \eta))_{0 \leq i \leq N-1, \eta \in \{0,1\}}$ sont les variables aléatoires gaussiennes *i.i.d.* utilisées pour définir la fonction de compatibilité G on a :

$$\max_{\sigma} G(\sigma) = \sum_{i=0}^{N-1} \max\{Z(i, 0), Z(i, 1)\}.$$

Il est maintenant aisé de calculer l'espérance du maximum de deux variables aléatoires gaussiennes standard pour conclure. \square

Nous avons un résultat de superconcentration pour le maximum de la fonction de compatibilité.

Théorème 10. — *En conservant les notations précédentes, nous avons :*

$$\text{Var}(\max_{\sigma} F(\sigma)) \leq \frac{CN}{K}, \quad C > 0.$$

Démonstration. — Soit $M := \max_{\sigma} F(\sigma)$. Notons par $\tilde{\sigma}$ la configuration pour laquelle ce maximum est atteint. Avec des notations évidentes, nous avons :

$$\frac{\partial M}{\partial Y(i, \eta)} = \mathbb{I}_{\{(\tilde{\sigma}_i, \dots, \tilde{\sigma}_{i+K}) = (\eta_1, \dots, \eta_{K+1})\}}.$$

Et par symmétrie,

$$\left\| \frac{\partial M}{\partial Y(i, \eta)} \right\|_{L^2}^2 = \left\| \frac{\partial M}{\partial Y(i, \eta)} \right\|_{L^1} = \frac{1}{2^{K+1}}.$$

Nous obtenons alors, grâce à l'inégalité de Talagrand :

$$\text{Var}(\max_{\sigma} F(\sigma)) \leq \frac{CN}{K}.$$

\square

8. Modèle de percolation

Dans un premier temps nous allons définir ce nouveau modèle, ensuite nous énoncerons le résultat de superconcentration le concernant. Pour établir ce Théorème nous utiliserons la méthode de Talagrand 2. Cependant, comme nous allons le voir, cet outil est insuffisant pour faire la démonstration complète. Nous aurons besoin d'utiliser « l'astuce » de Benjamini-Kalai-Schramm (cf. (3)) pour contourner les difficultés techniques.

On considère le réseau \mathbb{Z}^d et nous noterons par $E(\mathbb{Z}^d)$ l'ensemble des arêtes

de ce réseau. Soit alors $(\omega_e)_{e \in \mathbb{Z}^d}$ une famille de variables aléatoires positives *i.i.d.*, celles-ci correspondent au temps de passage ou au poids de chacune de ses arêtes. Ces poids peuvent être vus comme modélisant le temps nécessaire à un fluide pour passer une arête (que l'on pourrait assimiler à un tuyau).

Pour un chemin p reliant deux sommets du réseau \mathbb{Z}^d on définit le temps de passage du chemin p comme la somme des temps de passage de chaque arête du chemin. Le premier temps de passage $T(x, y)$ d'un sommet x à un sommet y est défini comme le minimum des temps de passage parmi tous les chemins reliant x à y .

$$i.e., \quad T(x, y) := \inf_{\mathcal{P}_{x,y}} \sum_{p \in \mathcal{P}_{x,y}} \omega_p.$$

Ce modèle de premier temps de passage en percolation à été introduit par J. M. Hammersley et D. J. A. Welsh (cf. (12)).

Etant donné un $x \in \mathbb{R}^d$ et un entier n , notons par $T_n(x)$ le premier temps de passage $T(0, [nx])$, où 0 représente l'origine de \mathbb{Z}^d et $[nx]$ le point du réseau le plus proche de nx . il existe toutes sortes de résultats concernant $T_n(x)$, par exemple :

$$\lim_{n \rightarrow \infty} \frac{T_n(x)}{n}$$

existe et est une fonction déterministe de x . Dans ce mémoire nous nous intéresserons uniquement au comportement de $Var(T_n(x))$.

Par simplicité de notation posons $x = e_1 = (1, 0, \dots, 0)$ et notons $T_n(x)$ par T_n . Grâce à des arguments de martingales, Kesten a prouvé que $Var(T_n) \leq Cn$. Bien que cela n'atteigne pas encore l'ordre de fluctuations attendu par les physiciens ($n^{2/3}$). Benjamini-Kalai-Schramm (BKS) ont amélioré l'estimation de Kesten en démontrant un phénomène de superconcentration.

Théorème 11. — *Considérons le premier temps de passage en percolation sur \mathbb{Z}^d , $d \geq 2$. Si les poids des arêtes peuvent être obtenues comme la loi d'une fonction Lipschitzienne de variables aléatoires gaussiennes uniformément borné, alors pour tout n ,*

$$Var(T_n) \leq \frac{Cn}{\log(n)},$$

où C est une constante qui ne dépend que de la loi des poids et de la dimension.

Tout d'abord, nous allons constater pourquoi le Théorème de Talagrand monotone ne s'applique pas directement. On suppose que la loi du poids de nos arêtes est une fonction Lipschitzienne de gaussiennes, et que cette fonction est bornée. Autrement dit, si ω_e est le poids d'une arête, on suppose

qu'il existe une gaussienne standard X_e et une fonction absolument continue $F : \mathbb{R} \rightarrow \mathbb{R}$ telle qu'il existe une constante $K > 0$ vérifiant $|F'| \leq K$. Ainsi que des constantes $0 < a < b < \infty$ telles que $a \leq F(x) \leq b$ pour tout x . Pour exemple, la loi uniforme sur l'intervalle $[a, b]$ vérifie ce type de conditions.

Le fait que les poids des arêtes soient uniformément bornés loin de zéro et de l'infini implique que la longueur du chemin optimal entre deux sommets ne peut dépasser un multiple de leur distance euclidienne.

En effet, par hypothèses, $T(0, x) \leq b|x|$ pour $x \in \mathbb{Z}^d$. C'est pourquoi, $|\mathcal{P}_{0,x}| \leq (b/a)|x|$, où $|\mathcal{P}_{0,x}|$ désigne le nombre de sommet appartenant au chemin $\mathcal{P}_{0,x}$ réalisant l'infimum. Cette hypothèse est cruciale pour utiliser l'astuce BKS.

Appelons \hat{p}_n le chemin optimal entre l'origine et ne_1 . Avec des notations évidentes, ceci nous donne que :

$$\frac{\partial T_n}{\partial X_e} = F'(X_e)1_{\{e \in \hat{p}_n\}}.$$

En conséquence,

$$\left\| \frac{\partial T_n}{\partial X_e} \right\|_{L^1} \leq K p_e,$$

où $p_e := \mathbb{P}(e \in \hat{p}_n)$. C'est pourquoi nous devons certainement démontrer que p_e est petit pour quasiment tout les sommets. Malheureusement ceci est compliqué à démontrer directement. L'astuce de Benjamini-Kalai-Schramm consiste à contourner ces difficultés.

Démonstration. — Expliquons les grandes lignes de la preuve avant d'aborder la démonstration du Théorème. L'idée est de considérer une boîte de côté k (pour un k que nous choisirons judicieusement plus tard). Dans cette boîte nous pouvons définir un temps de passage moyen, que nous noterons \tilde{T} . L'astuce BKS réside dans le fait que l'on peut contrôler l'écart entre T_n et \tilde{T} facilement et que la méthode de Talagrand s'applique sans trop de difficultés à \tilde{T} . Ces deux estimations permettront d'obtenir notre théorème.

Soit F une fonction Lipschitzienne telle que $|F'| \leq K$, et le poids d'une arête e soit $F(X_e)$, où les X_e sont des variables aléatoires gaussiennes standards *i.i.d.*.

Soit n quelconque et fixons k que nous choisirons ultérieurement. On pose $T = T_n$. Et considérons la boîte B de côté de longueur k centrée à l'origine. Notons par $|B|$ le nombre de points du réseau appartenant à B . Pour chacun

d'entre eux, $x \in B$, notons par T_x le premier temps de passage de x à $x + ne_1$. On pose alors,

$$\tilde{T} := \frac{1}{|B|} \sum_{x \in B} T_x.$$

Puisque les poids des arêtes sont uniformément bornés loin de zéro et de l'infini, il est facile de voir que $|T_x - T| \leq Ck$ où C est une constante dépendant uniquement de la loi des poids et de la dimension.

En effet, le chemin d'énergie minimale entre 0 et ne_1 étant fixé, nous pourrions construire un chemin particulier $\tilde{\gamma}$ issu de x et se terminant en $x + ne_1$.

Procédons de la manière suivante : lors des C_1k premiers pas nous faisons en sorte que $\tilde{\gamma}$ rejoigne \hat{p}_n . Ensuite, notre chemin $\tilde{\gamma}$ va suivre la même route que \hat{p}_n . Une fois arrivé en ne_1 , il ne faudrait que C_2k pas pour rejoindre le point $x + ne_1$. La remarque que nous avons faite sur la taille d'un chemin optimal entre deux points va nous permettre de conclure. Par construction, nous avons obtenu ceci :

$$\sum_{e \in \tilde{\gamma}} F(X_e) - \sum_{e \in \hat{p}_n} F(X_e) \leq C_1k + 0 + C_2k,$$

où les constantes C_1 et C_2 correspondent à l'écart entre nos deux chemins lorsque $\tilde{\gamma} \neq \hat{p}_n$. Autrement dit,

$$\sum_{e \in \tilde{\gamma}} F(X_e) \leq Ck + T$$

Il suffit alors de prendre l'infimum sur les chemins reliant x à $x + ne_1$, puis de suivre le même argument en fixant \hat{p}_n^x (le chemin optimal entre x et $x + ne_1$) et en construisant un chemin entre 0 et ne_1 .

Nous obtenons donc bien :

$$|T_x - T| \leq Ck.$$

Ainsi

$$|\tilde{T} - T| \leq \frac{1}{|B|} \sum_{x \in B} |T_x - T| \leq Ck.$$

De plus, en utilisant l'inégalité suivante : pour $a, b \in \mathbb{R}$ $(a + b)^2 \leq 2(a^2 + b^2)$ nous obtenons que :

$$(15) \quad \text{Var}(T) \leq 2\text{Var}(\tilde{T}) + Ck^2.$$

Rappelons que \hat{p}_n le chemin optimal entre l'origine et ne_1 . Pour chaque $x \in B$, soit \hat{p}_x le chemin optimal entre x et $x + ne_1$. Alors pour n'importe quelle arête

e ,

$$\frac{\partial \tilde{T}}{\partial X_e} = \frac{F'(X_e)}{|B|} \sum_{x \in B} 1_{\{e \in \hat{p}_x\}},$$

et donc

$$\left\| \frac{\partial \tilde{T}}{\partial X_e} \right\|_{L^1} \leq \frac{K}{|B|} \sum_{x \in B} \mathbb{P}(e \in \hat{p}_x).$$

Si l'on adopte la convention suivante : $e - x$ désigne l'arête e translatée de x . L'invariance par translation de notre problème nous donne alors que

$$\mathbb{P}(e \in \hat{p}_x) = \mathbb{P}(e - x \in \hat{p}_n).$$

De même si $e - B$ désigne la boîte translatée de e , c'est à dire l'ensemble des arêtes $e - x$ lorsque x parcourt les sommets de B , les deux inégalités précédentes impliquent que :

$$(16) \quad \left\| \frac{\partial \tilde{T}}{\partial X_e} \right\|_{L^1} \leq \frac{K}{|B|} \mathbb{E}|(e - B) \cap \hat{p}_n| =: A'_e.$$

Or, $|B|$ ne peut excéder l'ordre k^d (pensez à un carré puis à un cube etc.), alors que le nombre d'arêtes dans $e - B$ qui appartiennent à \hat{p} ne peuvent dépasser un multiple de k . En effet $(e - B) \cap \hat{p}_n$ est sous ensemble du chemin de poids minimum entre le premier point d'entrée de \hat{p} dans $e - B$ et le dernier point de sorti, et ces deux points sont au moins à une distance k l'un de l'autre. C'est pourquoi :

$$(17) \quad A'_e \leq Ck^{1-d}.$$

D'un autre côté, la majoration grossière d'une indicatrice par 1 nous fournit l'estimation suivante :

$$\begin{aligned} \left\| \frac{\partial \tilde{T}}{\partial X_e} \right\|_{L^2}^2 &\leq \mathbb{E} \left(\frac{K}{|B|} \sum_{x \in B} 1_{\{e \in \hat{p}_{n-x}\}} \right)^2 \\ &\leq \mathbb{E} \left(\frac{K^2}{|B|} \sum_{x \in B} 1_{\{e \in \hat{p}_x\}}^2 \right) = \frac{K^2}{|B|} \sum_{x \in B} \mathbb{P}(e \in \hat{p}_x) \\ &= \frac{K^2}{|B|} \sum_{x \in B} \mathbb{P}(e - x \in \hat{p}_n) = \frac{K^2}{|B|} \mathbb{E}|(e - B) \cap \hat{p}_n| := A_e^2. \end{aligned}$$

Remarquons que $A_e^2 = KA'_e$, c'est pourquoi en combinant 16 et 17, nous obtenons :

$$\left\| \frac{\partial \tilde{T}}{\partial X_e} \right\|_{L^1} \leq A'_e = \sqrt{A'_e} \frac{A_e}{\sqrt{K}} \leq Ck^{-\frac{d-1}{2}} A_e.$$

Nous pouvons dès lors utiliser la méthode de Talagrand qui nous assure que, après simplification, que :

$$\begin{aligned} \text{Var}(\tilde{T}) &\leq \frac{C}{\log(k)} \sum_e A_e^2 \\ &\leq \frac{C}{\log(k)} \frac{1}{|B|} \sum_{x \in B} \sum_e \mathbb{P}(e - x \in \hat{p}_n) \\ &= \frac{C}{\log(k)} \mathbb{E}|\hat{p}_n| \leq \frac{Cn}{\log(k)}. \end{aligned}$$

En se souvenant de l'inégalité (15) nous en déduisons que :

$$\text{Var}(T) \leq \frac{Cn}{\log(k)} + Ck^2.$$

En choisissant $k = n^\alpha$ avec $\alpha < 1/2$ nous terminons la preuve car dans ce cas $n^\alpha = O(n/\log(n))$. \square

9. Polymères dirigés en milieu aléatoire

L'astuce BKS que nous venons de voir pour le modèle de percolation fonctionne dans notre nouveau cadre. Les difficultés techniques sont seulement plus importantes pour ce cas de polymères. Présentons tout d'abord notre modèle, qui a été introduit pour la première fois par J. Z. Imbrie et T. Spencer (cf. (13)).

Soit $n \in \mathbb{N}$, nous considérons les marches aléatoires issues de 0 de longueur n . Plus précisément, chacun de nos chemins est une suite $\{(0, a_0), (1, a_1), \dots, (n, a_n)\}$ où $a_0 = 0$ et $|a_{i+1} - a_i| = 1$ pour tout i . Ainsi il y a 2^n chemins envisageables et nous appellerons chacun d'eux un polymère dirigé de dimension $(1 + 1)$ (le $(1 + 1)$ représentant une dimension de temps et une dimension en espace).

il est important de saisir la différence entre ce modèle et le modèle de percolation précédent. En effet, dans le modèle de percolation le « fluide » qui se déplaçait le long des arêtes était libre de ses mouvements. Alors que nos polymères, comme le souligne le mot « dirigé », ont un déplacement imposé. Expliquons ceci plus en détail : nos polymères sont obligés d'avancer, quitte à changer les axes. On peut voir notre modèle comme une marche aléatoires « up-crossing ». C'est à dire que notre polymère ne peut aller que sur la droite ou vers le haut et ne peut pas revenir sur ses pas.

Soit à présent $(X_v)_{v \in \mathbb{Z}^2}$ une famille de gaussiennes standard *i.i.d.* que l'on

appellera « le milieu » ou bien « l'environnement ». Étant donné une marche aléatoire p de longueur n , on définit l'énergie de ce chemin par :

$$H_n(p) := - \sum_{v \in p} X_v.$$

On parle alors de polymères gaussiens aléatoires de dimension $(1 + 1)$.

Nous nous intéresserons à l'énergie minimum d'un chemin de longueur n . Nous appellerons ceci l'état fondamental (« Ground state » en anglais) et nous le noterons par E_n . Nous étudierons également le chemin \hat{p}_n qui réalise cet état fondamental. Comme nous l'avons signifié plus tôt, l'astuce BKS s'applique pour ce nouveau modèle après quelques difficultés techniques. Toutefois nous pouvons obtenir le résultat suivant :

Théorème 12. — *Si E_n est l'état fondamental d'un polymère gaussien aléatoire de dimension $(1 + 1)$, alors*

$$\text{Var}(E_n) \leq \frac{Cn}{\log n},$$

où C est une constante indépendante de n .

Comme pour le modèle de percolation, les méthodes classiques nous donnent une estimation de la variance d'ordre n au lieu de n/\log). En effet, si l'on utilise l'abus de notation évident que

$$\frac{\partial E_n}{\partial X_v} = -1_{\{v \in \text{chemin optimal}\}},$$

qui nous fournit que

$$\begin{aligned} |\nabla E_n|^2 &= \sum_v \left(\frac{\partial E_n}{\partial X_v} \right)^2 \\ &= \sum_v 1_{\{v \in \text{chemin optimal}\}} = n + 1 \end{aligned}$$

Alors l'inégalité de Poincaré nous assure que :

$$\text{Var}(E_n) \leq n + 1.$$

Avant de prouver ce dernier théorème, nous allons faire une petite digression sur le phénomène de chaos. Comme nous l'avons démontré, la notion de superconcentration et de chaos sont équivalentes. Néanmoins ce phénomène de chaos est bien visible dans le modèle des polymères dirigés et permet de donner une justification de la définition que nous avons donnée.

Rappelons de quelle manière nous pouvons perturber l'environnement gaussien de nos polymères. On considère un polymère gaussien de dimension

(1 + 1) et pour chaque sommets $v \in \mathbb{Z}^2$ soit (X_v, X_v^t) , $t > 0$ un couple de variables aléatoires gaussiennes centrées de covariance e^{-t} et de variance égale à 1. Lorsque t est proche de zéro X_v et X_v^t sont égaux en loi avec une forte probabilité. Soit $\{(X_v, X_v^t) : v \in \mathbb{Z}^2\}$ une famille de couple *i.i.d.*. Alors si $(X_v)_{v \in \mathbb{Z}^2}$ est notre environnement original, nous pouvons appeler $(X_v^t)_{v \in \mathbb{Z}^2}$ notre environnement t -perturbé.

Comme nous allons le voir, une petite perturbation de l'environnement original nous fournit un nouveau chemin réalisant l'état fondamental quasiment disjoint du chemin d'énergie minimum de départ. Le théorème d'équivalence entre superconcentration et chaos nous fournit dans cet exemple-ci le résultat suivant.

Théorème 13. — *Considérons le modèle des polymères gaussiens de dimension (1 + 1). Si E_n désigne l'état fondamental, alors*

$$\text{Var}(E_n) = o(n)$$

si et seulement si il existe $t_n \rightarrow 0$, lorsque n tend vers l'infini, tel que

$$\mathbb{E}|\hat{p}_n \cap \hat{p}_n^{t_n}| = o(n).$$

Où $\hat{p}_n^{t_n}$ désigne l'état fondamental dans l'environnement t -perturbé et $|\cdot|$ correspond au nombre de sommets communs entre \hat{p}_n et $\hat{p}_n^{t_n}$.

Démonstration. — Considérons le modèle de polymères Gaussiens de dimension (1 + 1) introduit précédemment. Soit E_n l'état d'énergie fondamentale d'un polymère, g_v le poids du sommet v et $(X'_v)_{v \in \mathbb{Z}^2}$ des variables aléatoires Gaussiennes standard *i.i.d.* indépendantes de X_v . On définit l'environnement perturbé par :

$$X_v^t = c_t X_v + d_t X'_v.$$

Soit $V := \{-n, -n + 1, \dots, n - 1, n\}^2$ et $N := |V|$, on pose $\mathbf{X} := (X_v)_{v \in V}$ et $\mathbf{X}^t := (X_v^t)_{v \in V}$. Comme toujours P_t représente le semi-groupe de OU et γ^N la mesure gaussienne standard N -dimensionnelle. Soit \hat{p}_n^t le chemin optimal dans l'environnement \mathbf{X}^t et \hat{p}_n le chemin optimal dans l'environnement \mathbf{X} (*i.e.* $\hat{p}_n = \hat{p}_n^0$).

Puisque $\frac{\partial E_n}{\partial X_v} = -1_{\{v \in \text{chemin optimal}\}}$ et que $\nabla P_t = e^{-t} P_t \nabla$ nous obtenons

que,

$$\begin{aligned}
 \mathcal{E}(E_n, P_t E_n) &= \mathbb{E}_{\gamma^N}(\nabla E_n \cdot \nabla P_t E_n) \\
 &= e^{-t} \mathbb{E}_{\gamma^N}(\nabla E_n \cdot P_t \nabla E_n) \\
 &= e^{-t} \mathbb{E}_{\gamma^N}(\nabla E_n(g) \cdot \nabla E_n(g^t)) \\
 &= e^{-t} \sum_{v \in V} \mathbb{P}(v \in \hat{p}_n \text{ et } v \in \hat{p}^t)_n \\
 &= e^{-t} \mathbb{E}|\hat{p}_n \cap \hat{p}_n^t|.
 \end{aligned}$$

De plus pour le semi-groupe de OU $\lambda_1 = 1$. Ainsi, dire que E_n est (ϵ, δ) -chaotique (i.e $\mathcal{E}(E_n, P_t E_n) \leq \epsilon e^{-\lambda_1 t} \mathcal{E}(E_n, E_n)$ pour $t \geq \delta$) revient à dire que pour tout $t \geq \delta$

$$\mathbb{E}|\hat{p}_n \cap \hat{p}_n^t| \leq \epsilon(n+1).$$

C'est pourquoi, si ϵ et δ tendent vers zéro lorsque n tend vers l'infini, ceci signifie qu'une petite perturbation de l'environnement engendre un nouveau chemin optimal quasiment disjoint du précédent. Ce qui est bien l'idée que l'on se fait du chaos. Nous venons de démontrer que E_n est (ϵ_n, δ_n) -chaotique, pour ϵ_n et δ_n tendant vers 0, si et seulement si, il existe $t_n \rightarrow 0$ tel que $\mathbb{E}|\hat{p}_n \cap \hat{p}_n^{t_n}| = o(n)$.

Il ne reste plus qu'à démontrer que E_n est ϵ_n -superconcentré, pour $\epsilon_n \rightarrow 0$; si et seulement si $Var(E_n) = o(n)$. Dans ce cas le Théorème d'équivalence entre la superconcentration et le chaos nous permettra de conclure. Or, l'inégalité suivante

$$Var(E_n) \leq \mathcal{E}(E_n, E_n) = n + 1,$$

nous permet de montrer l'équivalence manquante. En effet, si E_n est ϵ_n -superconcentré cela signifie que

$$Var(E_n) \leq \epsilon_n C \mathcal{E}(E_n, E_n) = \epsilon_n C(n+1) = o(n).$$

L'autre sens se démontre aussi facilement. \square

Maintenant que nous avons démontré ce théorème d'équivalence pour le modèle de polymères gaussiens nous pouvons aborder la preuve de la superconcentration dans le théorème 12

Théorème 12. — L'idée est la même que pour le modèle de percolation. C'est à dire que l'on introduit une boîte B d'une longueur k que nous déterminerons ultérieurement, ainsi qu'une moyenne \tilde{E}_n des énergies des chemins de taille n issus d'un point x de cette boîte B . Comme nous pouvons utiliser l'inégalité suivante : $(a+b)^2 \leq 2(a^2 + b^2)$, il ne reste plus qu'à contrôler la variance de notre moyenne \tilde{E}_n ainsi que la variance de $E_n - \tilde{E}_n$ pour contrôler celle de E_n .

La difficulté réside dans l'estimation de $Var(E_n - \tilde{E}_n)$, pour y parvenir nous devons comprendre comment se comporte l'écart entre un chemin et son translaté (nous définirons tout ceci précisément plus loin). L'autre estimation, comme nous le verrons procède de la même manière que pour le modèle de percolation.

Après ce descriptif des grandes lignes de la preuve, introduisons des notations.

Fixons $n \in \mathbb{N}^*$, ainsi que $k = k(n)$ que nous choisirons ultérieurement. Soit B l'ensemble des points du réseau $\{0\} \times \mathbb{Z}$ qui sont à une distance inférieure ou égale à k de l'origine et dont les coordonnées sont paires. Notons par $|B|$ le nombre de points appartenant à B . Pour chaque point $x \in B$ on définit E_n^x comme étant l'énergie minimum parmi tout les polymères de taille n issus du point x . On pose alors

$$\tilde{E}_n = \frac{1}{|B|} \sum_{x \in B} E_n^x.$$

Soit \hat{p}_n le chemin optimal de taille n partant de l'origine. Notons par $(0, v_0), \dots, (n, v_n)$ les sommets de ce polymère \hat{p}_n , où $v_0 = 0$ et $|v_{j+1} - v_j| = 1$ pour tout j .

Si l'on rajoute le point $X_{n+1, v_{n+1}}$ où bien $X_{n+1, v_{n-1}}$ (qui sont les seules possibilités envisageables pour prolonger notre polymère) au chemin \hat{p}_n , nous obtenons un chemin de taille $n+1$. Donc par définition de E_{n+1} nous obtenons que

$$E_{n+1} \leq E_n - \max(X_{n+1, v_{n+1}}, X_{n+1, v_{n-1}}).$$

Ceci montre qu'il existe une constante $C > 0$ telle que $E_{n+1} - E_n \leq -C$. D'où, pour tout $0 \leq m \leq n < \infty$,

$$(18) \quad \mathbb{E}(E_n) - \mathbb{E}(E_m) \leq -C(n - m).$$

Soit τ le premier instant j tel que $v_{j+1} - v_j = 1$, c'est à dire le premier moment où notre polymère remonte. Si un tel j n'existe pas on pose $\tau = n$. Choisissons $m \leq n$, soit alors $E_{n,m}$ le chemin d'énergie minimum parmi tout les polymères de taille n issus de l'origine qui sont au point $-m$ à l'instant m . Ainsi les m premiers sommets de ces polymères sont déterministes : en effet celui-ci doit descendre jusqu'au point $-m$ durant les m premiers instants. Puisque l'espérance des m premiers poids vaut donc zéro, nous obtenons donc que $\mathbb{E}(E_{n,m}) = \mathbb{E}(E_{n-m})$. Ainsi d'après 18, $\mathbb{E}(E_n - E_{n,m}) \leq -Cm$.

En utilisant l'inégalité de concentration Gaussienne (12) avec

$$r = \mathbb{E}(E_n - E_{n,m}),$$

nous obtenons aisément que

$$\mathbb{P}(E_n = E_{n,m}) \leq \mathbb{P}(E_n - E_{n,m} \leq 0) \leq e^{-Cm^2/n}.$$

Remarquons que l'événement $\{\tau \geq m\}$ implique que $E_n = E_{n,m}$. En effet si dans le chemin optimal le premier instant où l'on remonte, après l'instant m , implique que l'on soit descendu au moins jusqu'au point $-m$. D'où

$$(19) \quad \mathbb{P}(\tau \geq m) \leq e^{-Cm^2/n},$$

nous connaissons donc le comportement de la queue de distribution de la variable aléatoire τ .

Choisissons le point $x \in B$ comme étant le point $(0, 2)$. Nous constatons à présent que la suite de points $(0, 2), (1, 2+v_1), (2, 2+v_2), \dots, (\tau+1, v_{\tau+1}), (\tau+2, v_{\tau+2}), \dots, (n, v_n)$ est un polymère de longueur n débutant du point x . En fait nous venons de recoller en l'instant τ notre polymère optimal \hat{p}_n issu de 0 avec son polymère translaté débutant en x . On remarque également que $v_j = -j$ pour $j \leq \tau$, par définition de τ .

En conséquence, par définition de E_n^x et grâce à la remarque précédente nous obtenons que

$$(20) \quad E_n^x \leq -\sum_{j=0}^{\tau} g_{(2,2-j)} - \sum_{j=\tau+1}^n g_{(j,v_j)}$$

$$(21) \quad = E_n - \sum_{j=0}^{\tau} X_{(2,2-j)} + \sum_{j=0}^{\tau} X_{(j,-j)}.$$

De plus, l'utilisation de l'estimation de la taille des moments Gaussiens donnée par le Lemme (16) (Cf. Annexe) nous fournit l'inégalité suivante :

$$(22) \quad \mathbb{E} \max_{0 \leq m \leq n} \left| \frac{\sum_{j=0}^m X_{(j,-j)}}{\sqrt{m+1}} \right|^2 \leq C \log n,$$

et la même borne reste valable si l'on remplace $X_{(j,-j)}$ par $X_{(j,2-j)}$. Notons par $A_k := \sum_{j=0}^k X_{(j,-j)}$ et par $Z_n := \max_{0 \leq k \leq n} \frac{A_k^2}{k+1}$. En utilisant les inégalités (19), (20) et (22) nous allons montrer que

$$\mathbb{E}(E_n^x - E_n)_+^2 \leq C\sqrt{n} \log n,$$

où $(\cdot)_+$ désigne la partie positive.

Tout d'abord $\mathbb{E}(E_n^x - E_n)_+^2 \leq \mathbb{E}(A_\tau^2) \leq \mathbb{E}(mZ_n 1_{\tau \leq m}) + \mathbb{E}(\tau Z_n 1_{\tau \geq m})$. En choisissant pour la suite $m = \sqrt{n}$ on constate que le premier terme de l'inégalité précédente est contrôlé par $\sqrt{n} \log n$. Pour le second, établissons un Lemme intermédiaire.

Lemme 11. — *Avec les notations précédentes nous avons :*

$$\mathbb{E}(\tau Z_n 1_{\tau \geq m}) \leq C\sqrt{n} \log(n).$$

Démonstration. — D'après l'inégalité de Cauchy-Schwarz nous obtenons que :

$$\mathbb{E}(\tau Z_n 1_{\tau \geq m}) \leq \mathbb{E}(Z_n^2)^{1/2} \mathbb{E}(\tau^2 1_{\tau \geq m})^{1/2}.$$

en utilisant, de nouveau, l'estimation des moments gaussiens du Lemme (16) (Cf. Annexe) on obtient que :

$$\mathbb{E}(Z_n^2)^{1/2} \leq C' \sqrt{\log n}.$$

De plus en utilisant nos connaissances sur la queue de distribution de τ ainsi que la comparaison série/intégrale nous avons que :

$$\begin{aligned} \mathbb{E}(\tau^2 1_{\tau \geq m}) &\leq \sum_{k=m}^n k^2 e^{Ck^2/n} \\ &\leq C_1 n \int_m^n x e^{-Cx^2/n} dx \\ &\leq C_1 \sqrt{nn} \int_1^{\sqrt{n}} u e^{-Cu^2} du \\ &\leq C\sqrt{nn} e^{-n} = O(n) \end{aligned}$$

D'où le résultat. □

En combinant les inégalités précédentes on obtient bien que :

$$\mathbb{E}(E_n^x - E_n)_+^2 \leq C\sqrt{n} \log n.$$

Par symétrie, la même borne reste valable pour la partie négative. Autrement dit, pour $x = (0, 2)$

$$\mathbb{E}(E_n^x - E_n)^2 \leq C\sqrt{n} \log n,$$

Nous allons maintenant utiliser ce moyen de comparaison pour estimer $\mathbb{E}(E_n^x - E_n)^2$.

Choisissons $x = (0, 2l)$ pour $l \in \mathbb{N}$, alors, en utilisant Cauchy-Schwarz,

l'inégalité précédente et la symétrie de translation, nous obtenons que

$$\begin{aligned} \mathbb{E}\left[(E_n^x - E_n)^2\right] &= \mathbb{E}\left[\left(\sum_{j=1}^l E_n^{(0,2j)} - E_n^{(0,2j-2)}\right)^2\right] \\ &\leq l \sum_{j=1}^l \mathbb{E}\left[(E_n^{(0,2j)} - E_n^{(0,2j-2)})^2\right] \\ &\leq Cl^2 \sqrt{n} \log n. \end{aligned}$$

Clairement, la même relation reste valable pour des l négatifs, en remplaçant dans le membre de droite l par $|l|$. Donc, pour n'importe quel $x \in B$,

$$\mathbb{E}(E_n^x - E_n)^2 \leq Ck^2 \sqrt{n} \log n.$$

C'est pourquoi

$$\text{Var}(E_n) \leq 2\text{Var}(\tilde{E}_n) + 2\mathbb{E}(\tilde{E}_n - E_n)^2 \leq 2\text{Var}(\tilde{E}_n) + Ck^2 \sqrt{n} \log n.$$

Comme nous allons l'expliquer ci-dessous, nous pouvons montrer que $\text{Var}(\tilde{E}_n) \leq Cn/\log k$. Il ne restera plus qu'à choisir $k = n^\alpha$ avec $\alpha < 1/2$ pour conclure cette démonstration.

Pour démontrer que la variance de \tilde{E}_n est bornée par $Cn/\log k$. Il suffit de reprendre la méthode utilisée dans la preuve du Théorème de superconcentration pour le modèle de percolation. En effet, il suffit de remarquer que :

$$\frac{\partial \tilde{E}_n}{\partial X_e} = \frac{1}{|B|} \sum_{x \in B} 1_{\{e \in \hat{p}_n^x\}},$$

où, nous le rappelons, \hat{p}_n^x est le polymère d'énergie minimum issu de x de taille n .

Ainsi nous obtenons que :

$$\begin{aligned} \left\| \frac{\partial \tilde{E}_n}{\partial g_e} \right\|_{L^1} &\leq \frac{1}{|B|} \sum_{x \in B} \mathbb{P}(e \in \hat{p}_n^x) \\ &= \frac{1}{|B|} \sum_{x \in B} \mathbb{P}(e - x \in \hat{p}_n) \\ &= \frac{1}{|B|} \mathbb{E}|(e - B) \cap \hat{p}_n| := A'_e \end{aligned}$$

Or, par définition, notre boîte B ne peut excéder l'ordre k c'est pourquoi nous obtenons que :

$$A'_e \leq \frac{C}{k} \mathbb{E}|(e - B) \cap \hat{p}_n|.$$

Puisque nous voulons appliquer la méthode de Talagrand il nous faut estimer

$$\left\| \frac{\partial \tilde{E}_n}{\partial X_e} \right\|_{L^2}^2.$$

En utilisant exactement le même type de majoration que dans le cas du modèle de percolation nous trouvons :

$$\begin{aligned} \left\| \frac{\partial \tilde{E}_n}{\partial X_e} \right\|_{L^2}^2 &= \mathbb{E} \left(\frac{1}{|B|} \sum_{x \in B} 1_{\{e \in \tilde{p}_x\}} \right)^2 \\ &\leq \frac{1}{|B|} \sum_{x \in B} \mathbb{P}(e \in \hat{p}_n^x) \\ &= \frac{1}{|B|} \mathbb{E} |(e - B) \cap \hat{p}_n| := A_e^2 = A'_e \end{aligned}$$

La méthode de Talagrand nous assure, après simplification, que :

$$\begin{aligned} \text{Var}(\tilde{E}_n) &\leq \frac{C}{\log n} \frac{1}{|B|} \sum_{x \in B} \sum_e \mathbb{P}(e - x \in \hat{p}_n) \\ &= \frac{C}{\log n} \mathbb{E} |\hat{p}_n| = \frac{Cn}{\log n}. \end{aligned}$$

Car $|\hat{p}_n|$ est une somme de n variables aléatoires gaussiennes standard indépendantes. Enfin, comme pour la démonstration de superconcentration pour le modèle de percolation, il ne reste plus qu'à choisir $k = n^\alpha$ avec $\alpha < 1/2$. \square

PARTIE V
ANNEXE

Dans cette annexe nous exposerons un certains nombres de résultats qui nous serviront d'outils dans l'étude de la superconcentration. Dans un premier temps nous rappellerons des résultats classiques concernant les variables aléatoires gaussiennes. Dans un second temps, nous démontrerons le Théorème d'hypercontractivité 1 à l'aide de l'inégalité de Sobolev Logarithmique 8 vérifiée par la mesure gaussienne.

10. Résultats sur les variables aléatoires gaussiennes

Nous allons énoncer un résultat que nous utiliserons implicitement tout au long du mémoire. La plupart du temps il sera combiné avec un résultat de concentration. Pour plus de détails à ce sujet, nous renvoyons le lecteur vers l'Appendice du livre de Talagrand (20).

Proposition 10 (Intégration par partie probabiliste)

Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, X une variable aléatoire positive et F une fonction croissante sur \mathbb{R}_+ , alors

$$\mathbb{E}[F(X)] = F(0) + \int_0^\infty \mathbb{P}(X > t) dF(t).$$

En particulier, lorsque $0 < p < \infty$,

$$\mathbb{E}[X^p] = p \int_0^\infty \mathbb{P}(X > t) t^{p-1} dt.$$

Par la suite nous considérerons un vecteur gaussien $\mathbf{X} = (X_1, \dots, X_n)$ de matrice de covariance Γ . Nous pouvons obtenir une inégalité de concentration pour $\mathbf{M} := \max_{i=1, \dots, n} X_i$. Nous noterons par $m := \mathbb{E}_\gamma[\mathbf{M}]$.

Proposition 11. — Pour tout $r \geq 0$,

$$\mathbb{P}(\mathbf{M} - m \geq r) \leq e^{-r^2/2\sigma^2}.$$

Cette même borne est également valable pour $\mathbb{P}(\mathbf{M} - m \leq -r)$.

Démonstration. — Pour obtenir ce résultat il suffit d'utiliser l'inégalité de Markov avec la fonction $x \mapsto \exp(x)$, puis de contrôler la transformée de Laplace de $\mathbf{M} - m$. En fait nous avons le même résultat en remplaçant le maximum \mathbf{M} par une fonction Lipschitzienne. Plus précisément, soit f une

fonction Lipschitzienne sur \mathbb{R} de norme de Lipschitz $\|f\|_{Lip} \leq 1$. Alors, pour tout $t \geq 0$,

$$\int_{\mathbb{R}} f d\gamma \leq e^{\int_{\mathbb{R}} f d\gamma - t^2/2}.$$

Ce qui nous permettra d'obtenir un résultat analogue à l'assertion de la proposition.

$$\gamma(f - \mathbb{E}_\gamma[f] \geq t) \leq e^{-t^2/2}.$$

Pour obtenir notre résultat il suffira de prouver que la fonction maximum est bien Lipschitzienne de constante de Lipschitz plus petite que un.

Nous allons utiliser essentiellement les propriétés du semi-groupe d'Ornstein-Uhlenbeck associé à la mesure gaussienne.

Avant cela, il est nécessaire de rappeler certaines définitions et propriétés. Nous désignerons par $L := \Delta - x \cdot \nabla$ le générateur infinitésimal du semi-groupe d'Ornstein-Uhlenbeck. A ce générateur, on peut lui associer la forme bilinéaire symétrique : l'opérateur « carré du champ » Γ .

Définition 12. — — Soient f, g deux fonctions appartenant au domaine du générateur infinitésimal L . On définit l'opérateur Γ de la manière suivante :

$$\Gamma(f, g) := \frac{1}{2}(L(fg) - fLg - gLf).$$

— On dit qu'un opérateur \mathbf{L} est une diffusion si, pour toute fonction Φ , C^∞ de \mathbb{R}^n dans \mathbb{R} et toute fonction $f = (f_1, \dots, f_n)$ appartenant au domaine de \mathbf{L} ,

$$\mathbf{L}(\Phi(f)) = \sum_i \Phi'_i(f) L(f_i) + \sum_{i,j} \Phi''_{i,j}(f) \Gamma(f_i, f_j).$$

On se convaincra sans peine, que l'opérateur $L = \Delta - x \cdot \nabla$ est une diffusion. Nous avons dans ce cas la relation suivante :

$$\Gamma(\Phi(f), g) = \Phi'(f) \Gamma(f, g).$$

Rappelons la représentation intégral du semi-groupe d'Ornstein-Uhlenbeck :

$$P_t(x) = \int_{\mathbb{R}^n} f(c_t x + d_t y) d\gamma(y) \text{ pour } x \in \mathbb{R}^n.$$

On vérifie sans peine que P_t est solution de l'équation de la chaleur :

$$\partial_t f = Lf.$$

Soit $u = u(t, x) = P_t f(x)$ pour $x \in \mathbb{R}^n$ et $t \geq 0$ une solution de l'équation de la chaleur : $\partial_t u = \frac{1}{2} \Delta_x u$ sur $\mathbb{R}^n \times]0, \infty[$.

On peut supposer, sans perdre de généralité que f est centrée sous γ . On remarque, par ailleurs, que si f est de classe C^2 alors

$$(23) \quad \Delta P_t f = e^{-t} P_t \Delta f.$$

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^2 de norme Lipschitz $\|f\|_{Lip} \leq 1$. Notons que la fonction f admet des moments exponentiels de tout ordre par rapport à la mesure γ car, pour tout $\lambda > 0$,

$$\mathbb{E}_\gamma[e^{\lambda f}] \leq e^{\lambda f(0)} \mathbb{E}_\gamma[e^{\lambda |x|^2}] < \infty.$$

Soit $\lambda \in \mathbb{R}$. On définit la fonction ϕ de la manière suivante :

$$\phi(t) = \mathbb{E}_\gamma[e^{\lambda P_t f}]$$

En particulier $\phi(0) = \mathbb{E}_\gamma[e^{\lambda f}]$. Nous allons chercher à comparer ϕ et ϕ' pour appliquer le Lemme de Gronwall.

En utilisant le fait que $P_t f$ est solution de l'équation de la chaleur, *i.e.* $\partial P_t f = L P_t f$, nous obtenons la relation suivante :

$$\phi'(t) = \mathbb{E}_\gamma \left[\lambda L P_t f e^{\lambda P_t f} \right].$$

On constate, par un argument de convergence dominée, l'ergodicité du semi-groupe d'OU. Plus précisément, lorsque t tend vers l'infini, $P_t f(x) \rightarrow \mathbb{E}_\gamma[f] = 0$. Autrement dit $\phi(t)_{t \rightarrow \infty} \rightarrow 1$. Le théorème fondamental de l'analyse nous assure donc que

$$\phi(t) = 1 - \int_t^\infty \phi'(s) ds.$$

En outre, le fait que γ soit une mesure invariante et symétrique, nous permet de réécrire ϕ' en terme de Γ

$$\begin{aligned} H'(s) &= \lambda \int e^{\lambda P_s f} L P_s f d\gamma \\ &= -\lambda \int \mathbf{\Gamma} \left(P_s f e^{\lambda P_s f} \right) d\gamma \end{aligned}$$

De plus le semi-groupe d'OU est une diffusion, en particulier

$$\mathbf{\Gamma}(f) := \mathbf{\Gamma}(f, f) = |\nabla f|^2.$$

Dans notre cas, grâce à (23), nous obtenons

$$\begin{aligned} \phi'(s) &= -\lambda^2 \int |\nabla P_s f|^2 e^{\lambda P_s f} d\gamma \\ &= -\lambda^2 \int e^{-2s} |P_s(\nabla f)|^2 e^{\lambda P_s f} d\gamma \end{aligned}$$

L'inégalité de Jensen nous fournit l'estimation suivante :

$$|P_s(\nabla f)|^2 \leq P_s(|\nabla f|^2).$$

On se souvient à présent que, par hypothèse, $\|f\|_{Lip} \leq 1$. Le Théorème de Rademacher nous assure alors, que $|\nabla f|_2^2 \leq 1$. Nous obtenons donc

$$\begin{aligned} \phi(t) &\leq 1 + \lambda^2 \int_t^\infty e^{-2s} \left(\int e^{\lambda P_s f} d\gamma \right) ds \\ &= 1 + \lambda^2 \int_t^\infty e^{-2s} H'(s) ds \end{aligned}$$

Le Lemme de Gronwall entraîne alors :

$$H(t) \leq \exp \left(\lambda^2 \int_t^\infty e^{-2s} ds \right).$$

En particulier,

$$H(0) = \mathbb{E}_\gamma[e^{\lambda f}] \leq e^{\lambda^2/2}.$$

En rassemblant tout ces résultats, nous avons démontré que, pour f une fonction Lipschitzienne, $\|f\|_{Lip} \leq 1$, centrée sous la mesure gaussienne γ et suffisamment régulière,

$$\gamma(f - \mathbb{E}_\gamma[f] \geq t) \leq e^{-\lambda t + \lambda^2/2}.$$

Un optimisation en λ nous permet de conclure la démonstration.

Si f est toujours Lipschitzienne mais n'est plus suffisamment régulière, il suffit de considérer la fonction $p_\epsilon f$, où l'on fait convoler f avec la densité gaussienne $\frac{1}{(2\pi\epsilon)^{n/2}} e^{-|x|^2/2\epsilon}$, $x \in \mathbb{R}^n$. On pourra appliquer l'argument précédent à cette fonction, pour finalement faire tendre ϵ vers 0, par convergence dominée, à la fin de la preuve.

Justifions brièvement que la fonction max est Lipschitzienne. Rappelons le cadre : $\mathbf{X} = (X_1, \dots, X_n)$ est un vecteur gaussien centré de matrice de covariance Γ . Sans perdre de généralité on peut supposer que Γ n'est pas dégénérée, il existe donc une matrice M telle que $\Gamma = M^t M$. Ainsi nous avons l'égalité en loi suivante : $\mathbf{X} = MG$ ou G est un vecteur gaussien standard de \mathbb{R}^n . Notons par $\sigma^2 := \max_{k=1 \dots n} \mathbb{E}[X_k^2]$.

On pose $f : \mathbb{R}^n \rightarrow \mathbb{R}$ où $f(x) = \max_{k=1, \dots, n} |(Mx)_k|$ Nous allons montrer que f est Lipschitz et $\|f\|_{Lip} \leq \sigma^2$.

En effet, soit $x, y \in \mathbb{R}^n$ et $k_0 \in \{1, \dots, n\}$ tel que $f(x) = |(Mx)_{k_0}|$.

$$\begin{aligned} f(x) - f(y) &= |(Mx)_{k_0}| - |(My)_{k_0}| \leq \left| (M(x-y))_{k_0} \right| = \left| \sum_{j=1}^n M_{jk_0}(x_j - y_j) \right| \\ &\leq \left(\sum_{j=1}^n M_{jk_0}^2 \right)^{1/2} \left(\sum_{j=1}^n (x_j - y_j)^2 \right)^{1/2} \\ &= \Gamma_{k_0 k_0}^{1/2} |x - y| = \mathbb{E}(X_{k_0}^2)^{1/2} |x - y| \\ &\leq \sigma |x - y| \end{aligned}$$

Nous avons utilisé, respectivement, l'inégalité triangulaire et l'inégalité de Cauchy-Schwarz. Il suffit, à présent, d'échanger les rôles de x et de y pour obtenir le résultat voulu. \square

Bien qu'il s'agisse d'un résultat profond et puissant, cela ne nous donne pas plus d'informations que la proposition précédente. En effet cette inégalité implique seulement que la fluctuation du maximum est, au plus, d'ordre σ^2 . En particulier, il s'agit d'une borne grossière du pire des cas qui n'utilise pas la structure de corrélation de \mathbf{X} . Néanmoins, on ne peut pas obtenir mieux en utilisant la théorie classique de la concentration de la mesure.

Nous allons à présent énoncer des inégalités concernant la queue de distribution d'une variable aléatoire gaussienne standard. On considère donc Z une variable aléatoire réelle gaussienne standard.

Lemme 12 (Bornes de Mills). — *Pour tout $x > 0$ on a l'encadrement suivant :*

$$\frac{x e^{-x^2/2}}{\sqrt{2\pi}(1+x^2)} \leq \mathbb{P}(Z > x) \leq \frac{e^{-x^2/2}}{x\sqrt{2\pi}}.$$

Démonstration. — Montrons d'abord l'inégalité de droite.

$$(24) \quad \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \leq \int_x^{+\infty} \frac{t}{x} \frac{1}{\sqrt{(2\pi)}} e^{-t^2/2} dt$$

$$(25) \quad = \frac{e^{-x^2/2}}{x\sqrt{2\pi}}.$$

Pour la minoration, nous allons introduire de force $\frac{1}{1+t^2}$ dans notre intégrale, puis séparer celle-ci en deux et à l'aide d'une intégration par partie bien choisie

nous obtiendrons le résultat escompté.

$$\begin{aligned} \int_x^{+\infty} \frac{1+t^2}{1+t^2} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt &= \int_x^{+\infty} \frac{1}{1+t^2} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt + \int_x^{+\infty} \frac{t}{1+t^2} \frac{te^{-t^2/2}}{\sqrt{2\pi}} dt \\ &= \frac{x}{1+x^2} \frac{e^{-x^2/2}}{\sqrt{2\pi}} + \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \frac{2}{(1+t^2)^2} e^{-t^2/2} dt \\ &\geq \frac{x}{1+x^2} \frac{e^{-x^2/2}}{\sqrt{2\pi}}. \end{aligned}$$

Car la dernière intégrale est strictement positive. \square

Nous utiliserons par la suite quelques résultats techniques concernant les variables aléatoires gaussiennes. Notamment le Lemme de minoration de Sudakov.

Lemme 13 (Sudakov). — *Supposons qu'il existe une constante $a > 0$ telle que $\mathbb{E}[(X_i - X_j)^2] \geq a$ pour tout $i \neq j \in [1, \dots, n]$. Alors*

$$\mathbb{E}[\max_{i=1, \dots, n} X_i] \geq Ca\sqrt{\log n},$$

où C est une constante universelle.

Démonstration. — Il suffit de comparer notre vecteur gaussien à un vecteur gaussien $\mathbf{Y} = (Y_1, \dots, Y_n)$, où les Y_i sont toutes *i.i.d* de variance bien choisie, via le Théorème de comparaison de Slepian. Le lecteur pourra en trouver la preuve dans le livre (16). \square

Remarque 11. — *Dans un certain sens, l'hypothèse $\mathbb{E}[(X_i - X_j)^2] \geq a$ signifie que nos variables sont très peu corrélées.*

Lemme 14. — *Soient X_1, \dots, X_n des variables aléatoires gaussiennes standards. On ne suppose pas que ces variables $(X_i)_{i=1 \dots n}$ soient indépendantes. Notons par $\mathbf{M} := \max(X_1, \dots, X_n)$. Alors*

$$\mathbb{E}[\mathbf{M}] \leq \sqrt{2 \log n}.$$

Démonstration. — Soit $\beta > 0$, nous avons les relations suivantes :

$$\begin{aligned} \mathbb{E}[\mathbf{M}] &= \frac{1}{\beta} \mathbb{E}[\log(e^{\beta \max(X_i)})] \leq \frac{1}{\beta} \mathbb{E}[\log(\sum_{i=1}^n e^{\beta X_i})] \\ &\leq \frac{1}{\beta} \log \sum_{i=1}^n \mathbb{E}[e^{\beta X_i}] = \frac{\beta}{2} + \frac{\log n}{\beta} \end{aligned}$$

Nous avons utilisé l'inégalité de Jensen ainsi que $\mathbb{E}[e^{\beta X_i}] = e^{\beta^2/2}$ pour tout $i = 1 \dots n$. Il suffit enfin d'optimiser en β pour obtenir le résultat annoncé par le Lemme. \square

Remarque 12. — *Ce résultat reste valable pour des variables aléatoires sous-gaussienne : i.e. leur transformée de Laplace est majorée par la transformée de Laplace d'une gaussienne $\mathcal{N}(0, 1)$.*

En fait, si l'on suppose, de plus, que nos variables aléatoires précédentes X_1, \dots, X_n sont indépendantes on peut obtenir une minoration de l'espérance du maximum du même ordre de grandeur.

Lemme 15. — *Soient X_1, \dots, X_n des variables aléatoires gaussiennes standard indépendantes. Notons par $M := \max(X_1, \dots, X_n)$. Alors*

$$\mathbb{E}[M] \geq C\sqrt{\log n},$$

où $C > 0$ est une constante universelle.

Démonstration. — Nous allons utiliser la formule d'intégration par partie 10. Pour cela nous devons nous ramener à une variable aléatoire positive. Nous utiliserons donc l'inégalité suivante :

$$\mathbb{E}\left[\max_{i=1, \dots, n} |X_i|\right] \leq \mathbb{E}[X_1] + 2\mathbb{E}\left[\max_{i=1, \dots, n} X_i\right].$$

Ainsi, une fois que nous aurons obtenus le résultat pour $\max_{i=1, \dots, n} |X_i|$ nous pourrons conclure la démonstration du Lemme.

Le fait que nos X_i soit *i.i.d.* nous permet d'obtenir les inégalités suivantes, pour tout $\delta > 0$,

$$\begin{aligned} \mathbb{E}\left[\max_{i=1, \dots, n} |X_i|\right] &\geq \int_0^\delta \left[1 - (1 - \mathbb{P}(|X_1| > t))^n\right] dt \\ &\geq \delta \left[1 - (1 - \mathbb{P}(|X_1| > \delta))^n\right]. \end{aligned}$$

En outre,

$$\mathbb{P}(|X_1| > \delta) = \sqrt{\frac{2}{\pi}} \int_\delta^\infty \exp(-t^2/2) dt \geq \sqrt{\frac{2}{\pi}} \exp(-(\delta + 1)^2/2).$$

Si l'on choisit, par exemple, $\delta = \sqrt{\log n}$ (pour n suffisamment grand) de tel sorte que $\mathbb{P}(|X_1| > \delta) \geq 1/N$. Nous obtenons donc :

$$\mathbb{E}\left[\max_{i=1, \dots, n} |X_i|\right] \geq \delta \left[1 - \left(1 - \frac{1}{n}\right)^n\right] \geq \delta \left(1 - \frac{1}{e}\right).$$

Et ceci nous permet de conclure. □

En fait, nous avons même un résultat plus général concernant les moments gaussiens.

Lemme 16 (Moments gaussiens). — Soient X_1, \dots, X_n des variables aléatoires gaussiennes standards i.i.d.. Si $n \geq 2$ lors pour tout $p \geq 1$

$$\mathbb{E}|\max_i X_i|^p \leq \mathbb{E}\max_i |X_i|^p \leq C(p)(\log n)^{p/2},$$

où $C(p)$ est une constante ne dépendant que de p .

Nous allons énoncer un résultat important qui nous a servi tout au long de ce mémoire : il s'agit de l'intégration par partie gaussienne.

Proposition 12 (Intégration par partie gaussienne)

Soit X une variable aléatoire gaussienne centrée de variance σ^2 .

– Soit F une fonction dérivable. Nous supposons que F vérifie une condition de décroissance à l'infini :

$$\lim_{|t| \rightarrow \infty} F(t) \exp(-t^2/2\sigma^2) = 0.$$

Alors nous avons la relation suivante :

$$\mathbb{E}[Xf(X)] = \sigma^2 \mathbb{E}[F'(X)].$$

Ceci peut se généraliser au cas multidimensionnel.

– Soient X, Z_1, \dots, Z_n des variables aléatoires gaussiennes centrées indépendantes. Soit également $F : \mathbb{R}^n \rightarrow \mathbb{R}$ différentiable, telle que, pour n'importe quel $a > 0$:

$$\lim_{|x| \rightarrow \infty} |F(x)| \exp(-a|x|^2) = 0.$$

Alors nous avons la formule suivante :

$$\mathbb{E}[XF(Z_1, \dots, Z_n)] = \sum_{l \leq n} \mathbb{E}[XZ_l] \mathbb{E}\left[\frac{\partial F}{\partial x_l}(Z_1, \dots, Z_n)\right].$$

Démonstration. — – Démontrons la première assertion. Il suffit simplement d'utiliser la formule d'intégration par partie usuelle et de tenir compte de la propriété de décroissance à l'infini de F .

$$\begin{aligned} \mathbb{E}[XF(X)] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} t \exp\left(-\frac{t^2}{2\sigma^2}\right) F(t) dt \\ &= \frac{\sigma^2}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} \exp\left(-\frac{t^2}{2\sigma^2}\right) F'(t) dt \\ &= \sigma^2 \mathbb{E}[F'(X)]. \end{aligned}$$

– Procédons à la démonstration du second point. Considérons les variables aléatoires gaussiennes

$$Z'_l = Z_l - G \frac{\mathbb{E}[Z_l G]}{\sigma^2}.$$

Celles-ci vérifient $\mathbb{E}[Z'_l G] = 0$. C'est pourquoi, G est indépendant de la famille (Z'_1, \dots, Z'_l) . Nous pouvons appliquer à chacune d'elles l'assertion précédente. Puisque $Z_l = Z'_l + g\mathbb{E}[GZ_l]/\sigma^2$, nous obtenons le résultat voulu. □

11. Théorème d'hypercontractivité

Dans cette section nous allons donner une preuve de l'inégalité d'hypercontractivité pour le semi-groupe d'Ornstein Uhlenbeck. Bien qu'il ne s'agisse que d'un outil dans l'étude de la Superconcentration, son rôle est tellement important que ne nous pouvions le laisser à l'état de « boîte noire ». Par soucis de consistance et pour satisfaire le lecteur curieux nous allons donc fournir une démonstration de ce résultat.

Nous allons procéder à une approche classique, faisant appel à l'inégalité de Sobolev logarithmique (8) vérifiée par la mesure gaussienne standard γ sur \mathbb{R}^n . Cette inégalité, rappelons le, atteste que pour toutes fonctions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ absolument continue, nous avons la relation suivante :

$$\int f^2 \log \frac{f^2}{\int f^2 d\gamma} d\gamma \leq 2 \int |\nabla f|^2 d\gamma,$$

lorsque les deux membres sont finis.

Démonstration. — Nous allons enfin donner la preuve du Théorème d'hypercontractivité 1, attestant que le semi-groupe d'Ornstein-Uhlenbeck est hypercontractif. Rappelons ce que cela signifie.

Pour tout $p > 1$, tout $t \geq 0$ et nimporte quelle fonction $f \in L^p(\gamma)$,

$$(26) \quad \|P_t f\|_{L^{q(t)}(\gamma)} \leq \|f\|_{L^p(\gamma)};$$

où,

$$q(t) := 1 + (p - 1)e^{2t}.$$

A noter que, $q(t) > p$ lorsque $t > 0$.

Nous allons obtenir cette inégalité d'hypercontractivité grâce à l'inégalité de Sobolev logarithmique vérifiée par la mesure gaussienne standard γ .

Soit f quelconque et posons $v := |f|$. Puisque $|P_t f| \leq P_t v$ et que $\|f\|_{L^p(\gamma)} = \|v\|_{L^p(\gamma)}$, il suffit de prouver notre inégalité pour des fonctions f positives. Soit $f \geq 0$ une telle fonction. Posons $f_t := P_t f$ et $r(t) := \int f_t^{q(t)} d\gamma$. Utilisons de nouveau l'équation de la chaleur $-\partial_t P_t f = LP_t f$. Puis le fait

que $-\int fLgd\gamma = \mathbb{E}_\gamma[\nabla f \cdot \nabla g]$, pour n'importe quelles fonctions f et g suffisamment intégrables. Nous en déduisons que :

$$\begin{aligned} r'(t) &= \int f_t^{q(t)} \frac{\partial}{\partial t} (q(t) \log f_t) d\gamma \\ &= q'(t) \int f_t^{q(t)} \log f_t d\gamma + q(t) \int f_t^{q(t)-1} \frac{\partial f_t}{\partial t} d\gamma \\ &= q'(t) \int f_t^{q(t)} \log f_t d\gamma + q(t) \int f_t^{q(t)-1} Lf_t d\gamma \\ &= \frac{q'(t)}{q(t)} \int f_t^{q(t)} \log f_t^{q(t)} d\gamma - q(t)(q(t)-1) \int f_t^{q(t)-2} |\nabla f_t|^2 d\gamma. \end{aligned}$$

En outre, remarquons que $q'(t) = 2(q(t)-1)$. C'est pourquoi, en remplaçant $r'(t)$ par sa valeur dans la deuxième égalité, nous obtenons

$$\begin{aligned} \frac{\partial}{\partial t} \log \|f_t\|_{L^{q(t)}(\gamma)} &= \frac{\partial \log r(t)}{\partial q(t)} = -\frac{q'(t) \log r(t)}{q(t)^2} + \frac{r'(t)}{q(t)r(t)} \\ &= \frac{q'(t)}{q(t)^2 r(t)} \int f_t^{q(t)} \log \frac{f_t^{q(t)}}{r(t)} d\gamma - \frac{q(t)-1}{r(t)} \int f_t^{q(t)-2} |\nabla f_t|^2 d\gamma \\ &= \frac{q'(t)}{q(t)^2 r(t)} \left(\int f_t^{q(t)} \log \frac{f_t^{q(t)}}{r(t)} d\gamma - \frac{q(t)^2}{2} \int f_t^{q(t)-2} |\nabla f_t|^2 d\gamma \right) \end{aligned}$$

Enfin l'inégalité de Sobolev logarithmique, appliquée à la fonction $f_t^{q(t)/2}$ nous fournit :

$$\int f_t^{q(t)} \log \frac{f_t^{q(t)}}{r(t)} d\gamma \leq \frac{q(t)^2}{2} \int f_t^{q(t)-2} |\nabla f_t|^2 d\gamma.$$

Donc, $\frac{\partial}{\partial t} \log \|f_t\|_{L^{q(t)}(\gamma)} \leq 0$ pour tout t . Et ceci prouve notre résultat. \square

12. Inégalité de Hoeffding

Proposition 13 (Inégalité de Hoeffding). — Soient, X_1, \dots, X_n des variables aléatoires i.i.d. de Bernoulli sur $\{-1, 1\}$, alors on a la relation suivante, pour tout $t > 0$:

$$\mathbb{P}\left(\sum_{i=1}^n X_i > t\right) \leq e^{-\frac{t^2}{2n}}.$$

Démontrons, tout d'abord, un Lemme qui nous sera utile pour la démonstration de l'inégalité d'Hoeffding.

Lemme 17. — Soit Y une variable aléatoire centrée, telle que $Y \in [a, b]$ presque sûrement. Posons $\psi_Y(\lambda) = \log \mathbb{E}[e^{\lambda Y}]$. Alors

- $\psi_Y''(\lambda) \leq (b-a)^2/4$.
- $\psi_Y(\lambda) \leq \lambda^2(b-a)^2/8$.

Démonstration. — Observons, en premier lieu, que

$$\left| Y - \frac{(b+a)}{2} \right| \leq \frac{b-a}{2}.$$

C'est pourquoi,

$$\text{Var}(Y) = \text{Var}(Y - (b+a)/2) \leq \frac{(b-a)^2}{2}.$$

Notons par P la loi de Y et par P_λ la loi de probabilité admettant pour densité :

$$x \mapsto e^{-\psi_Y(\lambda)} e^{\lambda x}.$$

Puisque P_λ est concentrée sur $[a, b]$, la variance d'une variable aléatoire Z suivant la loi P_λ est bornée par $(b-a)^2/4$. Ainsi, après un calcul élémentaire, nous obtenons :

$$\begin{aligned} \psi_Y''(\lambda) &= e^{-\psi_Y(\lambda)} \mathbb{E}[Y^2 e^{\lambda Y}] - e^{-2\psi_Y(\lambda)} (\mathbb{E}[Y e^{\lambda Y}])^2 \\ &= \text{Var}(Z) \leq \frac{(b-a)^2}{4}. \end{aligned}$$

La deuxième assertion découle simplement, en remarquant que $\psi_Y(0) = \psi_Y'(0) = 0$. Ainsi la formule de Taylor entraîne que, pour un $\theta \in [0, \lambda]$,

$$\psi_Y(\lambda) = \psi_Y(0) + \lambda \psi_Y'(0) + \frac{\lambda^2}{2} \psi_Y''(\theta) \leq \frac{\lambda^2(b-a)^2}{8}.$$

□

Démontrons à présent l'inégalité de Hoeffding

Inégalité de Hoeffding. — Nous pouvons nous placer dans un cadre légèrement plus général. Considérons des variables aléatoires X_1, \dots, X_n indépendantes, admettant un moment d'ordre un. Supposons de plus que, pour tout $i \in \{1, \dots, n\}$, il existe des constantes a_i, b_i telles que $X_i \in [a_i, b_i]$ presque sûrement.

Il suffira par la suite de choisir des variables aléatoires de Rademacher (Bernoulli symétrique sur $\{-1, 1\}$) et de remplacer, pour tout i , $a_i = -1$ et $b_i = 1$.

La démonstration repose sur le Lemme précédent et sur l'utilisation de l'inégalité de Chernoff. Nous avons l'inégalité suivante :

$$\mathbb{P}(Z \geq t) \leq e^{-\lambda t} \mathbb{E}[e^{\lambda Z}].$$

Il suffit alors de contrôler la transformée de Laplace de la variables aléatoire Z , pour enfin optimiser en λ .

Posons $S := \sum_{i=1}^n (X_i - \mathbb{E}[X_i])$, par indépendance nous avons, pour tout λ pour lesquels la transformée de Laplace de X_i est finie,

$$\psi_S(\lambda) = \sum_{i=1}^n \log \mathbb{E} \left[e^{\lambda(X_i - \mathbb{E}[X_i])} \right].$$

De plus, comme chaque $X_i \in [a_i, b_i]$ nous pouvons utiliser le Lemme précédent :

$$\psi_S(\lambda) \leq \frac{\lambda^2}{8} \sum_{i=1}^n (b_i - a_i)^2.$$

Ceci nous permet de contrôler la transformée de Laplace. Il ne reste plus qu'à optimiser en λ pour obtenir :

$$\mathbb{P}(S \geq t) \leq \exp \left(- \frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2} \right).$$

□

Références

- [1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, G. Scheffer *Sur les inégalités de Sobolev logarithmiques*
- [2] M. Benaïm, R. Rossignol *Exponential concentration for first passage percolation through modified Poincaré inequalities*
- [3] I. Benjamini, G. Kalai, O. Schramm *First passage percolation has sublinear distance variance*
- [4] S. Boucheron, G. Lugosi, P. Massart *Concentration Inequalities, a non asymptotic theory of independence*
- [5] S. Chatterjee *Chaos, concentration, and multiple valleys*
- [6] S. Chatterjee *Superconcentration and related topics*
- [7] D. Cordero-Erausquin, M. Ledoux *Hypercontractive measures, Talagrand's inequality, and influences*
- [8] S. Evans et D. Steinsaltz *Estimating some features of NK fitness landscape*
- [9] W. Feller *An introduction to probability theory Vol II*
- [10] B. Graham *Sublinear variance for directed last-passage percolation*
- [11] L. Gross *Logarithmic Sobolev inequalities*
- [12] J. M. Hammersley, D. J. A. Welsh *First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory*
- [13] J. Z. Imbrie, T. Spencer *Diffusion of directed polymers in a random environment*
- [14] S. A. Kauffman, S. A. Levin *Towards a general theory of adaptive walks on rugged landscapes*
- [15] M. Ledoux *The concentration of measure phenomenon*
- [16] M. Ledoux, M. Talagrand *Probability in Banach spaces*
- [17] E. Nelson *Probability theory and euclidian field theory*
- [18] D. Revuz, M. Yor *Continuous martingales and Brownian motion*
- [19] D. Sherrington, S. Kirkpatrick *Solvable model of a spin glass*
- [20] M. Talagrand *Mean field models for spin glasses*
- [21] K. Yoshida *Functional analysis*