THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *29 juin 2017* **par :**

Kevin Tanguy

Quelques inégalités de superconcentration : théorie et applications

JURY

Michel Ledoux	Professeur des universités	Directeur
Stéphane Boucheron	Professeur des universités	Rapporteur
Nathaël Gozlan	Professeur des universités	Rapporteur
Mireille Capitaine	Chargée de recherche	Examinatrice
Ivan Gentil	Professeur des universités	Examinateur
Ivan Nourdin	Professeur des universités	Examinateur

École doctorale et spécialité :
MITT : Domaine Mathématiques : Mathématiques appliquées
Unité de Recherche :
Institut de mathématiques de Toulouse (UMR 5219)
Directeur de Thèse :
Michel Ledoux
Rapporteurs :
Stéphane Boucheron et *Nathaël Gozlan*

À mes parents

Le silence éternel de ces espaces infinis m'effraie - B. Pascal

Remerciements

La fameuse page des remerciements : une des rares parties entièrement lue par l'ensemble des personnes assistant à la soutenance. J'ai beau y avoir songé durant ma thèse, je trouve qu'il est assez difficile de résumer ces quatres années sans oublier trop de choses, trop de personnes ou encore trop d'événements. Je vais pourtant m'essayer à cet exercice dans les lignes qui vont suivre et profiter du côté informel que l'on peut s'accorder dans cette partie de la thèse.

Bien évidemment, mes premiers remerciements vont à mon directeur de thèse Michel Ledoux. Peu importe les mots qui vont suivre, je pense qu'ils ne pourront jamais rendre compte de tout ce que j'ai pu apprendre sous sa supervision. J'estime qu'il a su diriger ma thèse d'une main de maître et m'a habilement formé au métier de chercheur. Grâce à lui, j'ai pu découvrir une manière d'explorer le monde de la recherche mathématique. Pour le citer : « *Les meilleures questions sont les vôtres, celles qui vous intéressent* ». En suivant ses recommandations, j'ai pu donc mieux appréhender le quotidien d'un mathématicien en me posant des questions, souvent naïves, sur les domaines qui me plaisaient : « c'est possible de faire ça avec cet outil ? Et ce résultat, pourquoi s'y prennent-ils comme ça ? On pourrait pas faire autrement ? De manière plus simple, plus élégante ? ». Grâce à ses conseils avisés, je parvenais à clarifier mes interrogations, en dégageant une problématique concise et dépourvue d'atours superflus, pour enfin essayer d'y répondre en résolvant des sortes de petits exercices qui se compliquaient au fur et à mesure que ma compréhension du problème augmentait. Finalement un chercheur prend alternativement le rôle de l'enseignant (en élaborant des exercices servant à mieux comprendre une problématique donnée) puis de l'étudiant (qui se doit de résoudre l'exercice à l'aide de ses connaissances en mathématiques). Au delà de cette formation, il a pu m'inculquer sa philosophie des mathématiques : la recherche constante de solutions élégantes aux problèmes posés. J'ai également pu profiter de sa vaste connaissance de certaines mathématiques, lesquelles ont sans cesse attisé ma curiosité d'en savoir plus. Pour toutes ces choses et les différentes discussions que l'on a pu avoir sur la pédagogie, l'enseignement, le système universitaire et l'univers de la recherche, je lui serai éternellement reconnaissant. Pour l'anecdote, lors d'une discussion que nous avons pu avoir concernant mon futur professionnel et l'éventuallité d'une carrière académique, je lui ai demandé ce qui l'avait amené à devenir professeur à l'université. Il fut amusant qu'un probabiliste tel que lui réponde, involontairement, avec modestie : « *le hasard* ».

Je tiens également à remercier mes rapporteurs, Stéphane Boucheron et Nathaël Gozlan pour avoir relu ma thèse avec attention. Les différents commentaires, remarques et conseils de Nathaël ont été extrêmement précieux pour améliorer et corriger mon manuscrit. Je ne le remercierai jamais assez de s'être intéressé à mes travaux et de m'avoir permis de les exposer au Mexique devant une prestigieuse assemblée de mathématiciens. Bien évidemment, je suis très honoré d'avoir Mireille Capitaine, Ivan Gentil et Ivan Nourdin dans mon jury. Mireille a toujours su s'armer d'une patience infinie lorsque je passais lui poser des questions sur les matrices aléatoires, Ivan N.

(alias « Mister Carte Metrolor », personne activement surveillée par les services de police de Washington) qui m'a encouragé à maintes reprises durant ma thèse et fut mon premier contact dans le monde de la recherche (hors Toulouse) et Ivan G. qui a agréablement répondu par mail à mes diverses interrogations concernant le transport optimal et l'inégalité de Poincaré. Merci à vous tous de faire partie de mon jury de thèse, j'en suis très touché.

Bien que les enseignants qui vont suivre ne sont par intervenus durant ma thèse, c'est pourtant un peu grâce à eux, ainsi qu'aux cours qu'ils ont pu me dispenser, que j'en suis là aujourd'hui et je tiens à les remercier : Jean-Jacques Horrent, Friedrich Wagemann (avec qui j'ai pu partager une scène de concert avec grand plaisir), François Nicoleau, Gilles Carron, Philippe Carmona qui m'a recommandé l'université de Toulouse pour mon master 2 et Laurent Thomann qui a encadré mon premier travail d'étude et de recherche à Nantes.

Mes pensées vont ensuite aux membres de ma famille qui ont su m'encourager durant cette thèse, qui ont fait brillamment semblant de m'écouter lorsque je parlais de mes problèmes mathématiques auxquels ils ne saisissaient pas grand chose. Une mention spéciale revient à mon père qui a relu plusieurs fois mon manuscrit dans son intégralité à la recherche de fautes d'orthographe, de syntaxes tordues, etc.... Pour ton aide précieuse, merci Papa. Je remercie également, bien entendu, ma mère qui a su apprécier la beauté mystérieuse de ces petits caractères inconnus. Tout ceci doit te sembler si éloigné du théorème de Pythagore, merci pour ton soutien Maman. Merci également à Arthur d'avoir partagé avec moi des moments d'évasion musicale lors de nos retrouvailles à Nantes ainsi qu'à Lola (la grande saucisse) pour m'appeler affectueusement « gros thon ». Evidemment, je suis très heureux que certains membres de ma famille aient pu venir voir ma soutenance à Toulouse : certains appâtés par le buffet, d'autres par curiosité, d'autres un peu contraints d'effectuer un si long voyage depuis Nantes alors qu'il fallait nourrir les chats. Merci à vous d'être venus de si loin : Nathalie, Rémy, Pierre, Mamie, Papy, Agnès et Jean-Luc. Je remercie également ceux qui n'ont pas pu venir pour diverses raisons : Marylène, Denis, William, Igor (il faudra que tu m'indiques à nouveau l'erreur que tu avais relevé dans mon article), Célia, Daphné, Salomé, le petit Vadim (fraîchement arrivé), ma cousinette chérie Sophie, la famille du sud (trop longue à énumérer), les autres Tanguy (Olivier, Véro, Margaux, Louis), Alp, Floriane, Sylvine, Antoine et Prime.

Enfin, le plus important à mes yeux, je ne me remercierai jamais assez Mélanie pour son soutien constant et sa présence quotidienne à mes côtés. J'ai hâte de partir à l'aventure avec toi dans une nouvelle ville, du fond du coeur, merci pour tout ce que tu m'apportes.

Quatres années de thèse cela peut sembler une longue période mais elle s'est écoulée si vite au contact de certains. Ces années passées à Toulouse m'ont permis de nouer de nouvelles amitiés, que ce soit à l'université ou en dehors. Je pense que le doctorat aurait perdu en saveur si je n'avais pas côtoyé Stéphane-Pépito-José qui a su illuminer certaines journées par ses capacités de juke-box (bien que le style musical soit assez ciblé, son déhanché combiné à un claquement de doigts qui ferait danser tout Cancùn est inoubliable), sans oublier les différents contes et légendes entourant sa mamie ainsi que sa phrase gimmick « *rahh vous me cassez les gonades, je me casse à Subway* » ou encore le fameux « *vous en faites plus une !* ». Je ne pourrais jamais oublier la rencontre avec Sofiane-Momo-l'endive (ou le navet ? Je ne sais plus trop, tout ceci semble confus), son amour pour les frites et la viande, sa passion du cinéma qu'il a pu me transmettre, nos discussions interminables sur nos futurs achats vidéoludiques, nos running-gags sur les oeufs et les tabourets qui ne faisaient rire que nous, nos idées pour obtenir un poste de maitre de conf' à coup sûr, notre ping-pong humoristique à l'INSA, sa dernière compétition d'aqua-gym, sa blessure (à l'origine d'un fameux krach boursier) lors du dernier chameau-racing des 24h d'Oman... Bref, comme dirait une certaine *Doudounette*, Sofiane c'est ma *Besta*. Merci pour tout mec ! A ce duo phénoménal vient s'ajouter trois personnes importantes. Tout d'abord, Raphaël-Boubou-Inspecteur Lababouche dont la verve n'a d'égal que sa démarche, mon grand frère de thèse qui a relu avec intêret l'un des chapitres du manuscrit et a su nous faire rire lors de certaines après-midis footbalistiques, par ses interprétations de Brassens à la guitare, par certaines de ses pensées philosophiques (pour n'en citer qu'une : « *raaah, longue marche* ») et qui fut l'un de mes partenaires récurrent de coinche. Ensuite, Yuriy-Yurko-Mme Nemish pour sa passion du club montagne, son intêret pour l'escalade, sa participation à l'installation d'une cloche à l'entrée de son bureau (qui semble avoir fortement surpris Michel Ledoux la première fois), son organisation sans failles (un véritable coucou Suisse) et sa générosité infinie. Enfin, Fanny-Gigi-Louis XIV, qui a su apporter une touche féminine à notre équipe de glands, et a aussi emmené un vent d'insouciance et de naïveté (heureusement wikipédia a pu apporter des réponses à certaines de ses interrogations). Grâce à elle nous avons été dotés de multiples dates d'anniversaire (si ma mémoire est bonne, Sofiane prend trois ans de plus au cours de chaque année civile). Elle a accepté et s'est essayée fougueusement à l'escalade lors d'une conférence aux Houches et fut une interlocutrice de très grande qualité lorsqu'il s'agissait de parler de recherche (d'ailleurs, j'ai l'impression de ne pas avoir été suffisamment à la hauteur concernant ses différentes questions sur la concentration de la mesure). Merci à vous tous pour ces années passées à vos côtés.

Il est évident que je ne pouvais pas faire de remerciements sans mentionner les membres de mon bureau. Tout d'abord, Claire Delplansky que je côtoie depuis le master 2, avec qui j'ai pu beaucoup rire et partager des moments d'entraide et de soutien moral. Je m'étonne encore de sa ferveur quasi religieuse pour Nicolas Cage et de son amour pour le raisin fermenté. Le bureau a également vu débarquer, après 14h, l'accent chantant de Ioana (« *Mais il est où Kevin ?* ») avec qui j'ai pu passer de bons moments comme des batailles de catapultes, des sorties escalade à Saint-Flour, le projet de construire un bonhomme de neige (« *Hai afară la Zăpadă ?* ») et l'apprentissage de la belote roumaine. Enfin, depuis l'année dernière Pierre a débarqué dans le bureau. Nous avons longtemps cru qu'il était empaillé avant de nous rendre compte qu'il prenait un malin plaisir à faire du bruit avec ses crayons sur son bureau. Malheureusement, la 4ième année à l'INSA de Toulouse ne m'a pas permis de passer plus de temps avec lui alors que nous aurions pu beaucoup échanger sur le cinéma. C'est étrange et un peu triste de se dire qu'une page se tourne et que l'on risque de ne pas se revoir de sitôt alors que vous avez fait partie de mon quotidien durant toute la thèse. En tout cas, merci à vous trois pour ces agréables années !

Bien entendu, je remercie vivement les autres doctorants, que j'ai pu fréquenter, pour leur agréable compagnie : Anäis-Tripoux-Pépita et son fauve Cannelle (bon elle ne fait pas partie des doctorants mais tant pis), Malika et son k-way rose bonbon, Mélanie-English-Blazère (aussi connue sous le prénom de Myriam), Daniel qui doit être la personne qui m'a le plus de fois dit « bonjour, ça va ? Merci Kevin » durant la thèse (alors que nous ne nous croisions qu'occasionnellement dans notre bureau à l'INSA), Magalie, Pierre Monmarché et notre retour à pied d'un pub de Bath, Hélène, Hugo mon partenaire de coinche et de chambrée à Saint-Flour (équipe invaincue, on est tellement des beaux gosses), Tatiana-Tatoche et son amour de la cuisine. Mais aussi ceux que j'ai rencontré trop tardivement : Valentin, Antoine B., Maylis,... J'en profite aussi pour remercier Benjamin (Monsieur Benj') pour les débats animés que nous avons pu avoir à l'INSA avec Sofiane (politique, son amour des animaux, sa brosse à cheveux, le personnel administratif, le partage de certains de ses rêves...), mais aussi Mélisande (passion mots-fléchés) ainsi que Loïc (le roi du ruban, le Philippe Candeloro de la GRS).

Il est important de ne pas oublier les amis de toujours : Joss (le gros Jojo, qui a hérité de ce sobriquet affectueux suite à une remarque de la prof d'anglais ≪ *Il est pas là, votre copain, le gros qui parle fort ?* ≫) qui m'a transmis sa passion de l'escalade, qui prend un malin plaisir à m'envoyer un sms en milieu de semaine concernant la parution (ou non !) du dernier chapitre de One Piece et qui arrive (encore et toujours) à finir vainqueur d'un grand chelem à MKDD. En parlant de ces compétitions sur GC, je suis obligé de citer Guillaume qui n'était jamais loin lorsque ce genre de rencontre se produisait et avec qui j'ai pu passer des réunions hebdomadaires très plaisantes (parfois fatigantes), au cours desquelles j'ai pu découvrir de nombreux jeux de société, résoudre des petits problèmes d'électricité et faire de belles parties d'Heroes 3 ! Durant ces différentes réunions ou compétitions de Mario Kart, il y avait toujours un troisième larron : Ivan (Princesse) ! Ivan fait partie de ma famille mais c'est aussi un ami très proche, qui fut mon compagnon de cordée depuis le master 2 ! Ce fut assez stimulant d'essayer de progresser aussi vite que l'autre en escalade, on a aussi beaucoup rigolé durant diverses soirées. Etonnamment, bien que cela arrive fréquemment, je n'arrive pas à me souvenir d'une remarque naïve d'Ivan qui ne m'ait pas fait rire (peut-être certaines confusions sur le nom des légumes ?). Sans m'étendre plus je suis obligé de mentionner, tellement ces moments étaient agréables, toutes les soirées que nous avons pu passer devant Heroes ou Civilization avec Arthur et lui chez nos grands-parents. En parlant de soirées jeux-vidéo, je pense forcément à Mathieu (notre doyen) avec qui j'ai l'impression d'avoir passé plus de temps devant l'ordinateur que je n'en ai passé dans une salle d'escalade ; en tout cas, chaque moment à ses côtés est toujours un plaisir (peut-être au détriment de la pauvre Mathilde qui sait qu'elle risque de récupérer un gros sac en fin de soirée, merci de me le prêter de temps en temps. Promis, j'en prends soin !) . Certaines amitiés datant du lycée, j'ai une tendresse toute particulière pour ces personnes et il me semble impossible d'énoncer (peut-être dans le volume 2 de ma thèse ?) tout ce qui me vient à l'esprit en pensant à eux : tout d'abord, Louis pour sa fraîcheur constante (d'un point de vue humoristique et intestinal), qui restera pour toujours mon meilleur ami, Fabien qui a fait le choix, contestable, de faire de la physique et que je ne vois pas autant que je le voudrais, Caro et sa bonne humeur quotidienne (qu'il est loin le temps ou tu m'accueillais dans ta chambre étudiante pour boire des gin-fizz ou lorsque tu m'apprenais à faire du poney chez ta mère ! !). D'ailleurs, je pense que tu devrais déménager pour que je puisse te voir plus souvent (surtout que tu fais de l'escalade maintenant !).

Bien que je les ai rencontrés plus tardivement, ces personnes me sont également très chères : Olivier (sa passion pour la seconde guerre mondiale, les figurines, les cochons de jade, les jeux de société), Marion (son rire communicatif, les gros yeux qu'elle peut faire à Louis, lorsqu'il s'emballe dans une discussion. Pour info, je réclame toujours la barboteuse Totoro pour Mathieu ! !), Delphine (sa constante discrétion, sa moue discrète lors de certains jeux de société), Maguie (sa bonne humeur, les petits apéritifs saucisson/fromage que l'on a pu partager, son amour de la cuisine et son GROS chat qu'elle nous prête généreusement de temps à autre), Anne-Laure (pour sa jovialité et son enthousiasme constant), Eric (pour égayer certaines séances de bloc), Jessica (the Kraaken et son chat, noir comme le fondement de Satan, la vieille Cocoon), Plouf (le hobbit brasseur de bière), Etienne (expert gustatif en houblon et dans la fermentation de fruits en tout genre). Merci à vous tous.

J'espère n'avoir oublié personne. Cependant, si c'est le cas j'en suis fort désolé mais comme pourrait dire Freud, dans son ouvrage *Zur Psychopathologie des Alltagslebens. Über Vergessen, Versprechen, Vergreifen, Aberglaube und Irrtum* : ≪ *Ein Mann erzählt von irgendwelchen Vorgängen, die er beanstandet, und setzt fort : Dann aber sind Tatsachen zum, Vorschwein' gekommen. ([. . .] Auf Anfrage bestätigt er, dass er diese Vorgänge als, Schweinereien' bezeichnen* *wollte.), Vorschein und Schweinerei' haben zusammen das sonderbare, Vorschwein' entstehen lassen* ≫. En français : ≪ il y a sûrement une raison inconsciente (ou pas) à ces oublis ! Posez-vous des questions, y aurait-il une raison ? ≫(Traduction approximative de l'Allemand à l'aide de mes souvenirs issus de certains cours prodigués par ma marraine la bonne fée).

Ces remerciements (peut-être trop longs, mais tant pis ! Je ne les ferai qu'une fois dans ma vie alors j'en profite !) s'achèvent et il est temps de passer aux choses sérieuses ! Bien que la plupart d'entre vous ne m'écoutera pas, je vous invite à lire la suite. J'ai essayé de rendre ce manuscrit le plus digeste possible. Pour vous mettre un peu l'eau à la bouche : Gauss va-t-il retrouver l'assassin de sa fille ? Emmy Noether va-t-elle rencontrer l'amour ? Je promets, dans le Chapitre 4, un cliffangher renversant dont la résolution n'apparaîtra que dans les dernières lignes de cette thèse ! J'espère que vous prendrez autant de plaisir en lisant mon manuscrit que j'ai pu en avoir à l'écrire et à réaliser cette thèse. Bonne lecture.

Table des matières

Introduction

La théorie de la concentration de la mesure s'est imposée comme un outil essentiel dans différents domaines mathématiques (analyse fonctionnelle, théorie des probabilités, méchanique statistique,...). Ainsi, l'une des forces de la théorie de la concentration de la mesure est sa généralité, ce qui lui permet d'être utilisée de manière efficace dans des situations diverses et variées. Cependant, une telle généralité manque de précision dans certains cas particuliers. Cette sous-optimalité a été mise en avant par Chatterjee dans son ouvrage ≪ Superconcentration and Related Topics ≫ dans lequel il liste un certains nombre d'exemples exhibant un tel phénomène et propose des méthodes souples permettant de palier ces problèmes d'imprécisions. Notre sujet de thèse s'inscrit dans la continuité des travaux de Chatterjee, nous avons donc cherché et obtenu de nouveaux résultats (certains améliorant ceux de Chatterjee) permettant une meilleure compréhension du phénomène dit de superconcentration. Nous proposons, ci-dessous, de décrire l'organisation du manuscrit de thèse.

Le premier chapitre propose un survol de la théorie de la concentration de la mesure et les différentes manières de l'aborder (isopérimétrie, inégalités fonctionnelles, transport optimal). Cela permet de souligner la polyvalence d'une telle théorie et donc les enjeux de notre sujet de thèse. Nous nous concentrons principalement sur le cas gaussien. Nous en profitons également pour introduire les outils, provenant de la théorie des semi-groupes, sur lesquels la majeure partie de nos travaux de recherche repose (notamment la propriété d'hypercontractivité).

Le chapitre deux présente le phénomène de superconcentration ainsi que l'état de l'art sur la question. Nous profitons de cette exploration pour établir les questions essentielles qui ont occupé cette thèse, à savoir : par quels moyens peut-on améliorer les bornes connues sur la variance et (le cas échéant) obtenir une inégalité de concentration reflétant ce gain ainsi que les asymptotiques d'une éventuelle convergence en loi. Ensuite, nous avons listé de manière exhaustive les différents modèles connus qui satisfont de la superconcentration. Comme le lecteur pourra le constater les domaines abordés sont très variés (matrices aléatoires, percolation, mécanique statistique, statistique d'ordre,...), rendant difficile la mise en place d'une méthodologie générale.

Le chapitre trois permet de rappeller des résultats généraux sur la théorie des extrêmes, comme nous le verrons ce domaine est étroitement lié au phénomène du superconcentration. Nous découvrirons qu'une des principales difficultés de cette thèse est d'obtenir des bornes non asymptotiques sur la variance ou des inégalités de déviations non-asymptotiques en adéquation avec la théorie des extrêmes. Ce domaine des mathématiques nous servira de source d'exemples pour illustrer certains théorèmes abstraits que nous avons obtenu durant cette thèse.

Le chapitre quatre est le premier chapitre à partir duquel nous exposons des résultats obtenu au cours de cette thèse. Nous y proposons un théorème abstrait permettant d'obtenir des inégalités de superconcentration exponentielles, celui-ci correspond au renforcement de certains travaux de Chatterjee. Nous illustrons ce nouveau théorème avec la théorie des extrêmes, en portant une attention particulière aux processus gaussiens stationnaires. Une illustration en statistique complète l'exposition. Une grande partie de ce chapitre s'appuit sur l'article ≪ Some Superconcentration inequalities for extrema of stationnary Gaussian processes ≫ que nous avons publié dans la revue ≪ Statistics and Probability letters ≫. Cet article est incorporé dans le manuscrit à la suite de la bibliographie.

Le chapitre cinq introduit la notion d'extremalité. Cette notion fut developpée dans le livre de Chatterjee comme une autre manière d'obtenir de la superconcentration. Nous étendons les travaux de Chatterjee au niveau exponentiel et constatons que nos résultats sont meilleurs, dans un certains régime, que ceux de la littérature existante.

Le chapitre six s'inspire d'une approche de Chatterjee de la superconcentration pour des modèles de verres de spins. Nous développons des méthodes d'interpolation permettant d'obtenir des inégalités de Talagrand d'ordre supérieur. Nous mettons également en avant le fait qu'il est possible d'obtenir de la superconcentration à partir de nouvelles inégalités fonctionnelles. En accord avec la littérature existante, nous avons choisis de les nommer ≪ inégalités de courbures dimension, intégrées, inversées ≫. Les méthodes d'interpolations mises en place dans ce chapitre sont suffisamment robustes pour être appliquées au cube discret. Nous proposons une application de ceci en analyse booléenne sur l'influence de coordonnées, cette application nous a également amené à définir la notion d'influence ≪ double ≫ qui semblait absente de la littérature. Les résultats de ce chapitre font parti d'un article en cours de rédaction.

Le chapitre sept aborde la superconcentration via la théorie du transport optimal. Nous utilisons des arguments similaires à ceux proposés par Gozlan pour des inégalités de Poincaré à poids pour produire de la superconcentration. Ces méthodes nous permettent de retrouver toute une série de bornes sur la variance de fonctionnelles. Nous parvenons également à obtenir des inégalités de déviations exponentielles. La majeur partie des ces résultats s'inscrit dans le cadre de la théorie des extrêmes et permettent d'illustrer de manière non-asymptotique les théorèmes présentés dans le chapitre 3. Notre approche permet également de redémontrer, avec des conditions plus faibles, une partie des travaux de Boucheron et Thomas sur les statistiques d'ordre. Ce chapitre fera également l'objet d'une note en vue d'une éventuelle publication.

Enfin, le chapitre huit fait office de conclusion et liste une série de questions ayant germées au cours de cette thèse. Ces questions sont de difficultés variées, certaines semblent pouvoir être résolues à l'aide des outils développés dans ce manuscript, d'autres, plus délicates, demanderons certainement des ingrédients supplémentaires.

Les articles rédigés durant cette thèse seront placés en annexe.

Notations

Nous regroupons, ci-dessous, les différentes notations que nous rencontrerons au cours de ce manuscrit. Nous souhaitons également signaler que nous avons adopté la convention suivante : un vecteur gaussien de $\mathbb{R}^n$, $X = (X_1, \ldots, X_n)$, sera toujours supposé, sauf mention du contraire, centré et réduit. C'est à dire que, pour tout $i = 1, \ldots, n$, $\mathbb{E}[X_i] = 0$ et $\mathrm{Var}(X_i) = 1$. Toutes les variables aléatoires de ce recueil sont définies sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$.

μ, ν, μ^n, ν^n	Mesures de probabilités et leurs produits tensoriels.
γ_n	Mesure gaussienne standard sur $\mathbb{R}^n$.
σ_n	Mesure uniforme sur la sphère $\mathbb{S}^n \subset \mathbb{R}^{n+1}$.
μ_α	Mesure gamma de paramètre $\alpha \geq 0$ sur $\mathbb{R}$.
card	Cardinal d'un ensemble fini.
$d(\cdot, \cdot)$	Distance sur un espace métrique.
$\langle \cdot, \cdot \rangle$	Produit scalaire euclidien.
$\\|x\\|_p^p = \sum_{i=1}^n \|x_i\|^p$	Norme l_p, $p \geq 1$ d'un vecteur de $\mathbb{R}^n$.
$L^p(E, \mu)$, $p \geq 1$	Espace de Lebesgue.
$\\|f\\|_p^p = \int_E \|f\|^p d\mu$	Norme $L^p(\mu)$ d'une fonction $f : E \to \mathbb{R}$.

$X, Y, Z, \ldots$	Variables aléatoires.
$\lambda_{\max}$	Plus grande valeur propre d'une matrice aléatoire.
$X_{(1)} > \ldots > X_{(n)}$	Statistiques d'ordre d'un vecteur aléatoire $(X_1, \ldots, X_n)$.
$(X_t)_{t \in T}$	Processus stochastique indéxé par un ensemble T.
$(B_t)_{t \geq 0}$	Mouvement brownien.
$(P_t)_{\geq 0}$	Semi-groupe d'Ornstein-Uhlenbeck.
$(Q_t)_{t \geq 0}$	Semi-groupe de Bonami-Beckner.
$M_n = \max_{i=1,\ldots,n} X_i$	Maximum d'un vecteur aléatoire $(X_1, \ldots, X_n)$.
$M_T = \sup_{t \in [0,T]} X_t$	Supremum d'un processus sur l'intervalle $[0, T]$.
I	Argmax de M_n, autrement dit $M_n = X_I$.
Γ, M	Matrice de covariance et sa racine carrée (*i.e.* $M^t M = \Gamma$).
$(H_k)_{k \in \mathbb{N}^n}$	Polynômes d'Hermite.
$(W_S)_{S \subset \{1,\ldots,n\}}$	Polynômes de Walsh.
$C_n = \{-1, 1\}^n$	Le cube discrèt.
$F_\beta, F_{n,\beta}$	Fonction d'énergie libre.
$V : \mathbb{R}^n \to \mathbb{R}$	Potentiel convexe.
$G, H, g, h \ldots$	Fonctions de répartition, respectivement densité de probabilité d'une variable aléatoire réelle.
$\kappa_\mu(x) = \frac{f(x)}{1-F(x)}$	Fonction de risque associé a une variable aléatoire réelle X de densité f par rapport à la mesure de Lebesque et de fonction de répartition F.
G^{-1}	Inverse généralisée d'une fonction de répartition.
Φ	Fonction de repartition d'une gaussienne standard.
$\Lambda_0, \Xi_\alpha, \Psi_\alpha$	Fonctions de répartitions des lois de Gumbel, de Weibull et Fréchet.
TW	Loi de Tracy et Widom.

$\mathbb{E}, \mathbb{E}_\mu$	Espérance.
$\mathrm{Var}, \mathrm{Var}_\mu$	Variance.
$\mathrm{Ent}, \mathrm{Ent}_\mu$	Entropie.
L	Opérateur de diffusion, générateur infinitésimal d'un processus de Markov.
$\mathcal{A}, \mathcal{D}(L)$	Domaine d'un générateur infinitisémal.
$\mathcal{E}$	Energie de Dirichlet associée à une diffusion.
Γ_n, $n \geq 0$	Opérateurs carré du champ de la théorie de Bakry et Émery.
$CD(\rho, \infty)$, $\rho \geq 0$	Critère de courbure dimension infinie.
∇, ∂_i	Gradient, dérivée partielle.
D, D_i	Gradient et dérivée partielle discrete.
Δ,	Laplacien.
$\mathcal{C}(r)$, $r \geq 0$	« Recouvrement » de $\{1, \ldots, n\}$.
$\rho(r)$	Supremum sur les ensembles $D \in \mathcal{C}(r)$ de la probabilité que $I \in D$.
$\tau(r)$	Nombre moyen d'éléments D de $\mathcal{C}(r)$ contenant I.
$(A_i)_{i=1,\ldots,n}$	Partition de l'espace ambiant (principalement $\mathbb{R}^n$).

Chapitre 1

Théorie de la concentration

Considérons (E, d, μ) un espace métrique mesuré. Si l'on suppose que μ est une mesure de probabilité sur les boréliens de E et que l'on note, pour tout $A \subset E$ et $r > 0$, par $A_r = \{x \in E,\ d(x, A) < r\}$ le r-élargissement de A ; il est évident que $\mu(A_r) \to 1$ lorsque $r \to \infty$. En revanche, il est plus surprenant que l'on soit capable d'exprimer, sur un grand nombre d'exemples non-triviaux, la vitesse à laquelle ceci s'accomplit. Plus précisément, si l'on définit la fonction, dite de concentration, associée à (E, d, μ) par

$$\alpha_{(E,d,\mu)}(r) = \sup \{1 - \mu(A_r) \,;\, A \subset E,\, \mu(A) \geq \frac{1}{2}\}, \quad r > 0$$

il est possible de contrôler cette fonction par un terme qui tend rapidement vers 0 lorsque $r \to \infty$. Si jamais il existe des constantes $C, c > 0$ telles que

$$\alpha_{(E,d,\mu)}(r) \leq Ce^{-cr^2}, \quad r > 0, \tag{1.1}$$

on parlera de concentration gaussienne. Tandis que s'il existe des constantes $C, c > 0$ telles que

$$\alpha_{(E,d,\mu)}(r) \leq Ce^{-cr}, \quad r > 0$$

on parlera de concentration exponentielle. Il est à noter que d'autres comportements peuvent-être obtenus et que la constante c peut dépendre d'une dimension n sous-jacente à l'espace E. La terminologie de concentration de la mesure provient de l'observation suivante : si l'on considère le cas (1.1) et un ensemble A de mesure $1/2$, alors quasiment tout les points (au sens de la mesure) de E seront à une distance $r = 10$ (par exemple) de l'ensemble A.

La théorie de la concentration de la mesure s'est developpée durant le siècle dernier et s'est imposée comme un outil incontournable dans différents domaines des mathématiques : théorie des probabilités, statistiques, géométrie, analyse fonctionnelle ... [112]. Historiquement, cette théorie est née au début des années soixante-dix avec une nouvelle preuve, reposant sur l'inégalité de Lévy (de nature isopérimétrique), de Milman, du Théorème de Dvoretzky [126] sur les sections sphériques de domaines convexes en grande dimension. Durant les années 1980 et 1990, cette théorie fut révolutionnée par Talagrand et ses remarquables travaux sur le phénomène de la concentration de la mesure dans les espaces produits [151, 145, 146]. Il proposa un regard novateur sur la notion d'indépendance de variables aléatoires. Il suggéra de considérer une variable aléatoire dépendant (de manière régulière) d'une grande famille de variables aléatoires comme étant quasiment constante. A titre illustratif considérons l'exemple suivant : soient $X_1, X_2, \ldots$

une suite de variables aléatoires indépendantes, prenant les valeurs ± 1 avec la même probabilité et posons, pour $n \geq 1$, $S_n = X_1 + \cdots + X_n$. On peut considérer S_n comme une fonction des variables aléatoires X_i. L'inégalité classique suivante, exprimant une version non-asymptotique du théorème de la loi forte des grands nombres, satisfait par les variables X_i,

$$\mathbb{P}\left(\frac{|S_n|}{n} \geq t\right) \leq 2e^{-nt^2/2}, \quad t \geq 0$$

nous assure que S_n/n est essentiellement nul. Cette approche lui permet d'obtenir des résultats majeurs dans cette théorie pouvant facilement s'appliquer à la résolution de problèmes complexes. Par exemple, son outil de l'inégalité de distance convexe lui a permis de développer des applications en probabilité combinatoire, de mécaniques statistiques et de processus empiriques ainsi que l'étude du supremum de processus [148, 149, 156]. Parallèlement, une remarque de Herbst sur les inégalités de Sobolev logarithmiques, permit une approche supplémentaire, via les inégalités fonctionnelles, pour étudier les propriétés de la concentration de la mesure. Il est à noter que cette approche est plus souple que l'approche isopérimétrique historique. Enfin, en 1997 une remarque de Marton [124, 123], permit une troisième approche, complémentaire à celle des inégalités fonctionnelles, par le biais de la théorie du transport optimal.

Ce chapitre rassemble, de manière non-exhaustive, des résultats théoriques essentiels de la concentration de la mesure. Nous renvoyons le lecteur vers les ouvrages [112, 38] pour plus de détails. Nous présenterons, dans un premier temps, des résultats d'isopérimétrie et indiquerons de quelle manière ils ont été source d'inspiration dans la théorie de la concentration de la mesure. Nous nous focaliserons sur le problème isopérimétrique gaussien et ses conséquences. Nous aborderons ensuite la concentration de la mesure via l'approche fonctionnelle, notamment par le prisme des inégalités de Poincaré et de Sobolev logarithmiques. Nous fournirons des démonstrations de ces inégalités d'un point de vue dynamique en suivant l'exposition faite dans [16, 111]. à l'aide de semi-groupes markoviens. Bien que sous jacente, nous n'entrerons pas dans les détails de la théorie de Bakry-Emery et la géométrie des opérateurs de diffusions. Les méthodes d'interpolation et de représentation de la variance (ou de l'entropie) fourniront les outils principaux qui seront utilisés pour démontrer les résultats essentiels de cette thèse. Enfin, nous développerons une dernière approche du phénomène de concentration de la mesure au travers de la théorie du transport optimal [163, 162].

1.1 Isopérimétrie

Le problème isopérimétrique dans le plan euclidien est un problème concret et vieux de plusieurs millénaires. Ce problème s'énonce comme suit : quels sont les ensembles minimisant le périmètre lorsque l'aire est fixée ? La légende raconte que sa solution, les disques euclidiens, avait déjà été démontrée de manière pragmatique par la reine Didon de Carthage. Ce résultat se généralise à l'espace euclidien, $\mathbb{R}^n$, pour lequel les ensembles minimisant la mesure de bord à volume fixé sont les boules euclidiennes. Bien que cette solution semble naturelle, la résolution rigoureuse de ce problème isopérimétrique n'est survenue qu'à la fin du XIXe siècle par Brunn, Minkowski puis Lusternik [120].

D'un point de vue plus abstrait, un problème isopérimétrique peut s'énoncer dans des espaces très généraux. En effet, il suffit d'une mesure pour définir le « volume » d'un ensemble et d'une distance pour définir le bord d'un ensemble. Par souci de simplicité et de concision, nous présentons une version « intégrée » du problème isopérimétrique pour des exemples spécifiques. Nous renvoyons le lecteur vers [29, 112, 33] pour plus de détails.

Considérons le problème isopérimétrique sur la sphère $\mathbb{S}^n \subset \mathbb{R}^{n+1}, n \geq 1$ équipée de sa mesure de volume (renormalisée) riemannienne σ_n. Ce résultat a été obtenu par Lévy au début des années 20 [119] puis, par Schmidt [140], de manière rigoureuse dans les années 40, par des techniques délicates de symétrisation. Essentiellement, ce théorème s'énonce comme suit :

Théorème 1.1.1. *Soient $A \subset \mathbb{S}^n$ et C une calotte sphérique satisfaisant $\sigma_n(C) = \sigma_n(A)$, alors*

$$\forall r > 0, \sigma_n(A) = \sigma_n(C) \Rightarrow \sigma_n(A_r) \geq \sigma_n(C_r),$$

où $A_r = \{x \in E, d(x, A) \leq r\}$ avec $d(x, A) = \inf\{d(x, y), y \in A\}$ et d qui désigne la distance canonique de la sphère.

Ce résultat propose un exemple non trivial du phénomène de concentration de la mesure en grande dimension. En effet, ce théorème implique que l'élargissement d'ordre $1/\sqrt{n}$ d'un ensemble de demi-mesure recouvre presque toute la sphère. En effet, si $\sigma_n(A) = 1/2$,

$$\sigma_n(A_r) \geq \sigma_n(C_r) \geq 1 - 2e^{-nr^2/2}.$$

De manière (équivalente) fonctionnelle, ceci implique de la concentration sous-gaussienne pour des fonctions lipschitziennes sur la sphère via la proposition suivante, démontrée par Lévy.

Proposition 1.1.2. *Soit $f : \mathbb{S}^n \to \mathbb{R}$ une fonction L-lipschitzienne et $M = \mathrm{Med}(f)$ sa médiane. Alors,*

$$\sigma_n(|f - M| \geq t) \leq 2e^{-\frac{nt^2}{2L^2}}, \quad t \geq 0$$

Cette observation est à l'origine du phénomène du concentration de la mesure et du fait que certaines fonctionnelles aléatoires sont quasiment constantes en grande dimension. Outre la naissance de la concentration de la mesure, cette proposition à également des conséquences sur d'autres problèmes isopérimétriques. Elle permet notamment de résoudre le problème isopérimétrique dans l'espace gaussien $(\mathbb{R}^n, \|\cdot\|_2, \gamma_n)$. Une observation attribuée, par erreur semble-t-il, à Poincaré stipule que l'espace gaussien peut-être « vu » comme limite de sphère N dimensionnelle de rayon $\sqrt{N}$. Au milieu des années 70, en combinant cette observation avec le résultat de Lévy, les mathématiciens Sudakov et Tsirelson [144], ainsi que Borell, [36] de manière indépendante, résolvent le problème isopérimétrique dans l'espace gaussien $(\mathbb{R}^n, \|\cdot\|_2, \gamma_n)$ et démontrent que les ensembles extrémaux sont des demi-espaces H. On rappelle que de tels ensembles sont définis par

$$H = \{x \in \mathbb{R}^n \,;\, \langle x, u \rangle \leq a\}$$

avec $u \in \mathbb{R}^n$ un vecteur unitaire et $a \in \mathbb{R}$.

Théorème 1.1.3. *Soit $A \subset \mathbb{R}^n$ et H un demi-espace tel que $\gamma_n(A) = \gamma_n(H)$. Alors*

$$\gamma_n(A) = \gamma_n(H) \Rightarrow \gamma_n(A_r) \geq \gamma_n(H_r), \quad r \geq 0$$

De même que sur les sphères, le résultat précédent fournit un énoncé fonctionnel incontournable du phénomène du concentration de la mesure. Bien que non présenté ici, signalons que l'on peut obtenir ce résultat par des méthodes d'interpolation et de monotonie le long du flot de la chaleur [109].

Théorème 1.1.4. *Soient $f : \mathbb{R}^n \to \mathbb{R}$ une fonction L-lipschitzienne et $M = \mathrm{Med}(f)$ sa médiane, alors*

$$\gamma_n(|f - M| \geq t) \leq 2e^{-\frac{t^2}{2L^2}}, \quad t \geq 0 \tag{1.2}$$

Une propriété remarquable et inhérente aux mesures gaussiennes est que ce résultat ne dépend pas de la dimension, contrairement à celui des sphères euclidiennes. Ce résultat permet de fournir des démonstrations alternatives du théorème de Dvoretzky en géometrie ou bien du lemme de Johnson-Lindenstrauss [75, 1]. Nous considérerons plus tardivement dans ce manuscrit l'étude de deux autres problèmes isopérimétriques, l'un étudié par Talagrand pour la mesure exponentielle symétrique [145] et l'autre par Bobkov pour la mesure exponentielle [24], et nous démontrerons leur pertinence vis à vis de la problématique de cette thèse.

1.2 Inégalités fonctionnelles

Bien que le phénomène de la concentration de la mesure soit d'inspiration isopérimétrique, il peut être abordé de plusieurs manières. L'une d'entre elles, beaucoup plus souple que l'approche isopérimétrique, repose sur des inégalités fonctionnelles. Celles-ci ont été abordées de différentes manières et dans différents contextes. Nous faisons le choix de présenter des démonstrations de ces inégalités du point de vue des semi-groupes de Markov et des méthodes d'interpolation introduites par Bakry et Émery dans le cadre des opérateurs du ≪ carré du champ ≫. Pour plus de détails, nous renvoyons le lecteur vers l'ouvrage très complet [16], ou bien pour une exposition plus succincte vers [111] dont nous nous inspirons. Ce choix de présentation est motivé par la raison suivante : la majeure partie des résultats obtenus durant cette thèse s'appuie sur des méthodes d'interpolation par semi-groupes. Nous présenterons, essentiellement, les inégalités fonctionnelles satisfaites par la mesure gaussienne ainsi que leurs conséquences vis à vis de la concentration de la mesure. Puisqu'il existe une certaine hiérarchie parmi ces inégalités fonctionnelles, nous commencerons par présenter la plus faible d'entre elles : l'inégalité de Poincaré, nous aborderons ensuite l'inégalité de Sobolev logarithmique. Dans un second temps, nous discuterons des méthodes d'interpolation provenant de la théorie de Bakry-Émery pour présenter des preuves élémentaires de ces inégalités fonctionnelles.

1.2.1 Inégalités de Poincaré

Dans la suite, nous considérons (E, d, μ) un espace métrique mesuré. Pour $f : E \to \mathbb{R}$, une fonction localement lipschitzienne sur (E, d), on peut généraliser la notion de longueur de gradient de f en un point $x \in E$ de la manière suivante :

$$|\nabla f|(x) = \limsup_{y \to x} \frac{|f(x) - f(y)|}{d(x, y)}. \tag{1.3}$$

Sur $\mathbb{R}^n$, muni de la distance euclidienne, cette définition coïncide avec la norme euclidienne d'un gradient.

Définition 1.2.1. On dit que μ vérifie une inégalité de Poincaré de constante $C_P > 0$ pour une famille de fonction $\mathcal{A}_P$, si pour toute fonction $f \in \mathcal{A}_P$,

$$\mathrm{Var}_\mu(f) \leq C_P \int_E |\nabla f|^2 d\mu,$$

lorsque le second membre a un sens et où $\mathrm{Var}_\mu(f) = \int_E f^2 d\mu - \left(\int_E f d\mu\right)^2$. Dans l'équation précédente, la quantité $|\nabla f|$ est définie par le gradient métrique introduit en (1.3).

Considérons $\mathbb{R}$ muni de sa métrique euclidienne ainsi que la mesure gaussienne standard sur $\mathbb{R}$, admettant la densité de probabilité suivante par rapport à la mesure de Lebesgue

$$d\gamma_1(x) = e^{-x^2/2} \frac{dx}{\sqrt{2\pi}}.$$

Proposition 1.2.2. *La mesure γ_1 vérifie une inégalité de Poincaré avec une constante optimale $C_P = 1$ et $\mathcal{A}_P$ consiste en l'ensemble des fonctions $f : \mathbb{R} \to \mathbb{R}$ de l'espace de Sobolev $H^1(\mathbb{R}, \gamma_1)$.*

Une propriété remarquable des inégalités de Poincaré est qu'elles se tensorisent indépendamment de la dimension, plus précisément nous avons le résultat suivant [112].

Proposition 1.2.3. *Soient (E_i, d_i, μ_i), $\quad i = 1, \ldots, n$, une famille d'espace métrique mesuré, alors la mesure produit μ^n satisfait l'inégalité suivante, pour $f \in L^2(\mu^n)$,*

$$\mathrm{Var}_{\mu^n}(f) \leq \sum_{i=1}^{n} \int_E \mathrm{Var}_{\mu_i}(f) d\mu^n,$$

où Var_{μ_i} signifie que l'on intègre uniquement par rapport à la i-ème coordonnée. En particulier, si les mesures μ_i, $\quad i = 1, \ldots, n$ vérifient une inégalité de Poincaré de constante $C_i > 0$, alors la mesure produit μ^n satisfait une inégalité de Poincaré de constante $C = \max_{i=1,\ldots,n} C_i$.

Cette proposition entraine que la mesure (produit) gaussienne γ_n sur $\mathbb{R}^n$ vérifie une égalité de Poincaré de constante $C_P = 1$, c'est à dire, pour toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulière,

$$\mathrm{Var}_{\gamma_n}(f) \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n \tag{1.4}$$

Décomposition spectrale et polynômes orthogonaux

Il se trouve qu'une inégalité de Poincaré permet d'obtenir de la concentration à partir de l'étude du spectre d'un opérateur de diffusion sur une variété riemannienne [112]. Nous rappelons que L est un opérateur de diffusion sur $\mathbb{R}^n$ s'il s'écrit

$$L = \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x) \frac{\partial}{\partial x_i},$$

avec a_{ij} et b_i, $i, j = 1, \ldots, n$, des fonctions suffisamment régulières. Nous noterons par $\mathcal{D}(L)$ l'ensemble des fonctions sur lesquelles L peut agir. Une mesure μ sera dite invariante (cf. [8] page 29), par rapport à un opérateur de diffusion L, lorsque pour toutes fonctions $f \in \mathcal{D}(L)$ nous avons la relation suivante :

$$\int_E Lf d\mu = 0.$$

Comme affirmé précédemment, les inégalités de Poincaré sont spectrales et correspondent à une formulation variationnelle de la première valeur propre non nulle d'un opérateur de diffusion L sur une variété riemannienne. Nous allons développer cet aspect dans cette section.

Sur $\mathbb{R}$, si une mesure μ admet un moment exponentiel (c'est à dire qu'il existe $\alpha > 0$ tel que $\int_{\mathbb{R}} e^{\alpha|x|} \mathrm{d}\mu(\mathrm{x}) < \infty$) il est possible de construire de manière canonique une base hilbertienne de $L^2(\mu)$. En effet, une telle base est obtenue en orthogonalisant, par rapport au produit scalaire de $L^2(\mu)$, la suite des polynômes $(x^k)_{k \geq 0}$ qui est alors un ensemble dense dans $L^2(\mu)$. A la normalisation et au signe près, une telle famille de polynômes orthogonaux est unique. Nous prendrons d'ailleurs pour convention le choix de normaliser (dans $L^2(\mu)$) ces polynômes et d'imposer au coefficient du terme dominant d'être strictement positif.

Dans un nombre assez restreint de cas, ces polynômes sont aussi les fonctions propres d'un opérateur de diffusion. En fait, en dimension un, il n'y a que les familles des polynômes d'Hermite, de Laguerre et de Jacobi qui vérifient une telle propriété. Nous nous focaliseront uniquement sur les polynômes d'Hermite et de Laguerre dans cette section. N'étant pas exactement de même nature, nous choisissons d'aborder le cas du cube discret $C_n = \{-1,1\}^n$ et des polynômes de Walsh dans un chapitre ultérieur. Il est à noter que les exemples des polynômes d'Hermite, Laguerre et Jacobi font parti des rares cas pour lesquels la description des suites de fonctions propres et de valeurs propres d'un opérateur de diffusion, dit de Sturm-Liouville, est complète (cf. [16]). Nous présentons ci-dessous deux exemples d'opérateur de diffusion et leur décomposition spectrale associée, à savoir l'exemple d'Ornstein-Uhlenbeck et de l'opérateur de Laguerre.

Nous verrons ensuite qu'une inégalité de Poincaré peut facilement être obtenue via la décomposition spectrale de l'opérateur de diffusion. Etant donné un opérateur de diffusion L, nous désignerons par $\mathcal{E}(f,g) = \int_E f(-Lg)d\mu$, pour f,g suffisamment régulière, l'énergie de Dirichlet associée. Notons, que dans la plupart des exemples que nous considérerons par la suite, une intégration par partie permet de réexprimer le second membre des inégalités de Poincaré, $\int_E |\nabla f|^2 d\mu$, en fonction de l'énergie de Dirichlet. Par exemple, dans le cas gaussien, pour f suffisamment régulière, on a

$$\int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n = \mathcal{E}(f,f).$$

En dimension un, la famille des polynômes orthogonaux associée à l'opérateur de diffusion d'Ornstein-Uhlenbeck L et à la mesure gaussienne γ_1 est celle des polynômes d'Hermite que nous noterons $(H_k)_{k\geq 0}$. Ils sont obtenus, dans leur version normalisée, via la formule suivante

$$H_n(x) = \frac{(-1)^n}{\sqrt{n!}} e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}, \quad x \in \mathbb{R}.$$

Rappelons que L agit sur un ensemble de fonctions f suffisament régulières par la formule suivante :

$$Lf = f'' - xf'$$

et que l'énergie de Dirichlet s'exprime simplement par $\mathcal{E}(f) = \int_{\mathbb{R}} f'^2 d\gamma_1$, où $\mathcal{E}(f)$ est un abus de notation pour désigner $\mathcal{E}(f,f)$. Comme annoncé plus haut, les polynômes d'Hermite sont des fonctions propres de l'opérateur L, plus précisément nous avons la relation suivante :

$$-LH_k = kH_k, \quad k \in \mathbb{N},$$

Ainsi le spectre $(\lambda_k)_{k\geq 0}$ de l'opérateur L est discret et $\lambda_k = k, \quad \forall k \in \mathbb{N}$. En utilisant la décomposition d'une fonction de $L^2(\gamma_1)$ sur la base des polynômes d'Hermite on obtient la formule suivante :

$$f = \sum_{k=0}^{\infty} a_k H_k,$$

avec $a_k = \int_{\mathbb{R}} f H_k \mathrm{d}\gamma_1$. Ceci entraine facilement les relations suivantes :

1. $$\|f\|_2^2 = \sum_{k\geq 0} a_k^2 \tag{1.5}$$

2. $P_t f = \sum_{k\geq 0} e^{-kt} a_k H_k, \quad t \geq 0$

3. $-Lf = \sum_{k\geq 0} ka_k H_k$.

Puisque l'énergie de Dirichlet s'exprime en fonction de L par $\mathcal{E}(f) = \int_{\mathbb{R}} f(-Lf)\mathrm{d}\gamma_1$, les formules de décomposition précédentes entrainent que, pour tout $f \in L^2(\gamma_1)$,

$$\mathcal{E}(f) = \sum_{k\geq 0} ka_k^2. \tag{1.6}$$

Il est alors facile, à partir de (1.5) et (1.6), de démontrer de manière spectrale l'inégalité de Poincaré 2.1 pour une fonction $f \in L^2(\gamma_1)$ de moyenne nulle. L'utilisation de la terminologie de trou spectral servant, parfois, à désigner une inégalité de Poincaré devient alors évidente. En effet, comme mentionné plus tôt, la meilleure constante C_P obtenue telle que (1.2.1) soit vérifiée, correspond à la première valeur propre non nulle de l'opérateur L.

Nous pouvons étendre toutes ces définitions au cas multidimensionnel. Notons de nouveau par L l'opérateur d'Ornstein-Uhlenbeck dans $\mathbb{R}^n$,

$$\text{i.e.} \quad Lf = \Delta f - x \cdot \nabla f,$$

avec $f \in \mathcal{D}(L)$. Les valeurs propres et les fonctions propres de l'opérateur L sont maintenant indexées par $\mathbb{N}_+^n$.

Définition 1.2.4. Pour chaque $k = (k_1, \dots, k_n) \in \mathbb{N}_+^n$ 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, auxquelles nous ajoutons des formules d'intégrations utiles pour certains calculs.

Proposition 1.2.5. *1. Le k-ième polynôme d'Hermite H_k est une fonction propre de l'opérateur $-L$ correspondant à la valeur propre $k_1 + \cdots + k_n$.*

2. Puisque $L(x_i) = -x_i$ pour tout $1 \leq i \leq n$, pour n'importe quelle fonction $f \in \mathcal{D}(L)$,

$$\int_{\mathbb{R}^n} x_i f d\gamma_n = -\int_{\mathbb{R}^n} L(x_i) f d\gamma_n = \int_{\mathbb{R}^n} \partial_i f d\gamma_n,$$

ce genre de résultat s'obtient facilement en faisant directement une intégration par partie avec la densité gaussienne.

Plus généralement, pour f suffisament régulière de sorte que la formule suivante fasse sens, on a

Proposition 1.2.6. *Pour tout $k \in \mathbb{N}^n$, nous avons la relation de récurrence suivante :*

$$\int_{\mathbb{R}^n} f H_k d\gamma_n = \frac{1}{\sqrt{k!}} \int_{\mathbb{R}^n} \frac{\partial^{k_1+\cdots+k_n} f}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}} d\gamma_n,$$

avec $k! = k_1! \dots k_n!$.

Cette dernière proposition nous permet de réécrire la formule (1.5) d'une autre manière :

$$\|f\|_2^2 = \sum_{k\in\mathbb{N}^n} a_k^2 = \sum_{k\in\mathbb{N}^n} \frac{1}{k!}\left(\int_{\mathbb{R}^n} \frac{\partial^{k_1+\cdots+k_n} f}{\partial x_1^{k_1}\cdots\partial x_n^{k_n}} d\gamma_n\right)^2,$$

où $a_k = \int_{\mathbb{R}^n} fH_k d\gamma_n, \quad \forall k = (k_1,\ldots,k_n) \in \mathbb{N}^n$ et $k! = k_1!\ldots k_n!$. Cette même somme peut se réorganiser de la façon suivante :

$$\|f\|_2^2 \quad = \quad \sum_{m=0}^{\infty} \theta_m(f),$$

où

$$\begin{aligned}
\theta_m(f) \quad &= \quad \sum_{|k|=m,\, k\in\mathbb{N}^n} \frac{1}{k!}\left(\int_{\mathbb{R}^n} \frac{\partial^{k_1+\cdots+k_n} f}{\partial x_1^{k_1}\cdots\partial x_m^{k_n}} d\gamma_n\right)^2 &(1.7)\\
&= \quad \sum_{|k|=m,\, k\in\mathbb{N}^n} \left(\int_{\mathbb{R}^n} fH_k d\gamma_n\right)^2, &(1.8)
\end{aligned}$$

correspondant aux carrés des normes des projections de f sur le chaos gaussien d'ordre m. Comme indiqué par Chatterjee dans son ouvrage [48], nous verrons plus tard que cette égalité fournit des conditions nécessaires sur f pour avoir de la superconcentration. Il existe dans la littérature différentes décompositions de la norme $L^2(\gamma_n)$ d'une fonction f, autrement dit sa variance, du même type. Nous renvoyons le lecteur vers les articles suivants : [49, 92, 93, 94, 127, 110].

Présentons maintenant un dernier exemple d'opérateur de diffusion, qui peut-être diagonalisé à l'aide de polynômes orthogonaux, l'opérateur de Laguerre. Celui-ci présente beaucoup de similarités avec le cas d'Ornstein-Uhlenbeck, nous omettrons donc certains détails.

En dimension un, l'opérateur de Laguerre L_α, avec $\alpha > 0$, sur $E = \mathbb{R}_+$ agit sur des fonctions régulières f sur $\mathbb{R}_+$ de la manière suivante :

$$L_\alpha f = xf'' + (\alpha - x)f'.$$

Cet opérateur est naturellement associé à la mesure gamma μ_α [16], sur $\mathbb{R}_+$ dont la densité par rapport à la mesure de Lebesgue est donnée par

$$d\mu_\alpha(x) = \gamma_a^{-1} x^{\alpha-1} e^{-x} dx,$$

où γ_a^{-1} désigne une constante de renormalisation. Son énergie de Dirichlet est donnée par, pour f suffisament régulière, $\mathcal{E}(f) = \int_{\mathbb{R}_+} xf'^2 d\mu_\alpha$. Comme pour l'exemple d'Ornstein-Uhlenbeck, L_α peut-être diagonalisé dans une base orthogonale de polynôme $(L_k)_{k\geq 0}$ (avec les mêmes conventions que précédemment concernant la normalisation et le signe du coefficient du terme dominant) de $L^2(\mu_\alpha)$. Cette famille porte le nom de polynômes de Laguerre et ils vérifient la formule suivante :

$$-L_\alpha L_k = kL_k,$$

montrant que L_α possède le même spectre que l'opérateur L d'Ornstein-Uhlenbeck.

Il est instructif de considérer le cas particulier (avec $\alpha = 1$) de la mesure exponentielle $d\mu_1(x) = e^{-x}dx$ sur $\mathbb{R}_+$. Similairement au cas d'Ornstein-Uhlenbeck, nous pouvons démontrer de manière spectrale que μ satisfait une inégalité de Poincaré. Autrement dit, pour toute fonction régulière $f : \mathbb{R}_+ \to \mathbb{R}$ dans le domaine de l'opérateur de Laguerre,

$$\mathrm{Var}_{\mu_1}(f) \leq \int_{\mathbb{R}_+} x f'^2 d\mu_1. \tag{1.9}$$

Notons que le second membre correspond bien à l'énergie de Dirichlet associée à L_1. Le spectre de L_1 est discret et l'inégalité de Poincaré représente bien l'écart entre 0 et la première valeur propre non nulle 1. De manière assez surprenante, il est également possible (c.f. [16]) d'obtenir une autre inégalité de Poincaré pour la mesure μ par rapport à l'énergie de Dirichlet $\mathcal{E}(f) = \int_{\mathbb{R}_+} f'^2 d\mu_1$.

Proposition 1.2.7. *Soit $d\mu_1(x) = e^{-x}dx$ la mesure exponentielle sur $\mathbb{R}_+$. Alors pour toute fonction $f : \mathbb{R}_+ \to \mathbb{R}$ suffisamment régulière,*

$$\mathrm{Var}_{\mu_1}(f) \leq 4 \int_{\mathbb{R}_+} f'^2 d\mu_1 \tag{1.10}$$

Démonstration. Nous reprenons la démonstration exposée dans [16]. Il est suffisant de considérer des fonctions régulières à support compact $f : \mathbb{R}_+ \to \mathbb{R}$. Remarquons tout d'abord que

$$\mathrm{Var}_{\mu_1}(f) \leq \int_{\mathbb{R}_+} [f - f(0)]^2 d\mu_1$$

puisque la variance minimise la distance aux constantes dans $L^2(\mu_1)$. Puisque $\int_{\mathbb{R}_+} f'^2 \mathrm{d}\mu_1$ n'est pas affectée par l'ajout d'une constante à f, il sera suffisant d'établir l'inégalité souhaitée pour des fonctions f satisfaisant $f(0) = 0$. Alors, d'après le Théorème de Fubini,

$$\begin{aligned} \int_{\mathbb{R}_+} f^2 d\mu_1 &= \int_0^\infty f(x)^2 e^{-x} dx \\ &= \int_0^\infty \left(2 \int_0^x f'(t) f(t) \mathrm{d}t \right) e^{-x} dx \\ &= 2 \int_0^\infty f'(t) f(t) e^{-t} \mathrm{d}t = 2 \int_{\mathbb{R}_+} f f' d\mu_1. \end{aligned}$$

La conclusion s'ensuit en appliquant l'inégalité de Cauchy-Schwarz à la dernière intégrale. □

Comme expliqué dans [16], ce même type d'argument s'applique également à la mesure exponentielle symétrique sur $\mathbb{R}$ $d\mu(x) = \frac{1}{2} e^{-|x|}$. En fait, il est possible d'étendre considérablement ce genre de résultats via des arguments (que nous ne présentons pas ici) reposant sur des fonctions de Lyapounov [16]. C'est le contenu du théorème suivant du à Kannan, Lovász, Simonovits et Bobkov [102] d'après [16].

Théorème 1.2.8. *(Kannan, Lovász, Simonovits, Bobkov) Soit μ une mesure log-concave sur les boréliens de $\mathbb{R}^n$, $d\mu = e^{-V}dx$ avec $V : \mathbb{R}^n \to \mathbb{R}$ une fonction régulière convexe, alors μ satisfait une inégalité de Poincaré par rapport à l'énergie de Dirichlet usuelle $\mathcal{E}(f) = \int_{\mathbb{R}^n} |\nabla f|^2 d\mu$.*

Remarque. La raison principale de ne pas explorer plus avant ce type de démonstration avec les fonctions de Lyapounov, est qu'elle ne permet pas d'obtenir les constantes optimales.

Nous verrons plus tard dans le manuscrit, lorsque nous aborderons les inégalités de type Talagrand L^1/L^2, en quoi le fait que les mesures gamma vérifient deux inégalités de Poincaré par rapport à deux énergies de Dirichlet différentes est important.

Conséquence de l'inégalité de Poincaré et concentration de la mesure

L'inégalité de Poincaré satisfaite par la mesure gaussienne est pleine de conséquences pour celle-ci. Nous débutons cette section par un résultat sur la variance d'un vecteur gaussien admettant une matrice de covariance Γ. Nous présenterons ensuite une proposition attestant qu'une mesure satisfaisant une inégalité de Poincaré admet de la concentration au moins exponentielle.

Pour commencer, nous déduisons de l'inégalité de Poincaré pour la mesure gaussienne standard une inégalité de Poincaré pour la mesure induite par un vecteur gaussien dans $\mathbb{R}^n$ de matrice de covariance Γ.

Proposition 1.2.9. *Soit $X = (X_1, \ldots, X_n)$ un vecteur gaussien de matrice de covariance Γ. L'inégalité suivante est toujours vérifiée :*

$$\mathrm{Var}\big(f(X)\big) \leq \mathbb{E}\big[\langle M\nabla f, M\nabla f\rangle\big] = \mathbb{E}\big[\langle \Gamma\nabla f, \nabla f\rangle\big],$$

avec ${}^tMM = \Gamma$.

Démonstration. Nous pouvons supposer, sans perdre de généralité, que Γ, qui est toujours positive, soit également définie positive. Ainsi, il existe $M \in M_n(\mathbb{R})$ telle que $\Gamma = {}^tMM$. 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 :

$$\mathrm{Var}\big(f(X)\big) \leq \mathbb{E}\big[\langle M\nabla f, M\nabla f\rangle\big] = \mathbb{E}\big[\langle \Gamma\nabla f, \nabla f\rangle\big].$$

□

Nous en déduisons le résultat suivant, également présent dans l'article de Houdré [92],

Corollaire 1.2.10. *Soit $X = (X_1, \ldots, X_n)$ un vecteur gaussien de matrice de covariance Γ. L'inégalité suivante est toujours vérifiée :*

$$\mathrm{Var}\big(\max_{i=1,\ldots,n} X_i\big) \leq \max_{i=1,\ldots,n} \mathrm{Var}(X_i).$$

Démonstration. Posons $M_n = \max_{i=1,\ldots,n} X_i$ le maximum des coordonnées d'un vecteur gaussien standard. On remarque alors que $f(X_1, \ldots, X_n) = M_n = \sum_{i=1}^n X_i 1_{A_i}$, avec, pour tout i=1,. . . ,n, $A_i = \{X_i = \max_{j=1,\ldots,n} X_j\}$. Ainsi, $\partial_i f = 1_{A_i}$ pour tout $i = 1, \ldots, n$. Puisque $(A_i)_{i=1,\ldots,n}$ est une partition de $\mathbb{R}^n$, l'inégalité est immédiate en appliquant l'inégalité de Poincaré, satisfaite par la loi de $(X_1, \ldots, X_n)$, à la fonction $f(x) = \max_{i=1,\ldots,n} x_i$. Cependant, pour être rigoureux, il aurait fallut faire la démonstration avec une fonction plus régulière que $\max_{i=1,\ldots,n}$. On aurait pu utiliser la fonction suivante

$$F_\beta(x) = \frac{1}{\beta}\log\Big(\sum_{i=1}^n e^{\beta x_i}\Big),\ \beta > 0$$

où $x = (x_1, \ldots, x_n)$, qui converge vers $\max_{i=1,\ldots,n} x_i$ lorsque $\beta \to \infty$ (il s'agit d'une approximation de Gibbs) et conclure par convergence dominée. □

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 pas en compte la structure de corrélation de notre vecteur. Cependant, on peut se convaincre facilement, en considérant une seule variable aléatoire ($n = 1$), ou au contraire toutes les mêmes ($X = (X_1, \dots, X_1)$), que cette inégalité est optimale.

Intégrabilité exponentielle

Comme il est exposé dans [112, 16] une inégalité de Poincaré de constante $C > 0$ entraine que les fonctions lipschitziennes sont exponentiellement intégrables. D'un point de vue plus historique sur les liens entre la concentration de la mesure et les inégalités de Poincaré nous renvoyons le lecteur curieux vers les articles [90, 141, 5]. Plus précisement nous avons la proposition suivante [112].

Proposition 1.2.11. *Si μ vérifie une inégalité de Poincaré de constante $C > 0$, alors pour toute fonction f 1-lipschitzienne et pour tout $s < \sqrt{\frac{4}{C}}$*

$$\int_E e^{sf} d\mu \leq e^{s\int_E f d\mu} \prod_{l=0}^{\infty} \left(1 - \frac{Cs^2}{4^{l+1}}\right)^{-2^l} \tag{1.11}$$

Remarque. Il est à noter que la démonstration précédente repose sur le fait suivant : pour une fonction f fixée on applique l'inégalité de Poincaré à la fonction $g = e^{sf/2}$, $s \in \mathbb{R}$, ceci permettant d'obtenir une inégalité entre la fonction $Z(s) = \int_E e^{sf} \mathrm{d}\mu$ et $Z(s/2)$. Cette inégalité servira à montrer que $Z(s)$ est bornée par 3 si $s < \sqrt{\frac{4}{C}}$. Il suffit alors d'appliquer l'inégalité de Chernoff pour conclure. Cette idée de preuve sera réutilisée ultérieurement en faisant dépendre C de la dimension.

Une conséquence primordiale de l'intégrabilité exponentielle des fonctions lipschitziennes sous la condition d'une inégalité de Poincaré est le phénomène de concentration de la mesure suivant, obtenu par l'inégalité de Chernoff,

Proposition 1.2.12. *Soit μ une mesure vérifiant une inégalité de Poincaré de constante $C > 0$. Alors, pour toute fonction f lipschitzienne et tout $t \geq 0$,*

$$\mu\left(f \geq \int_E f d\mu + t\right) \leq 3e^{-3t/\sqrt{C}\|f\|_{Lip}}$$

Remarque. L'inégalité précédente correspond plutôt une inégalité de déviation (de la moyenne). Cependant le même argument appliqué à la fonction $-f$ fournit

$$\mu\left(\left|f - \int_E f d\mu\right| \geq t\right) \leq 6e^{-3t/\sqrt{C}\|f\|_{Lip}}, \quad t \geq 0,$$

qui exprime bien une inégalité de concentration (autour de la moyenne).

1.2.2 Inégalités de Sobolev logarithmiques

Les inégalités de Sobolev logarithmiques font partie des inégalités fonctionnelles qui ont été les plus étudiées durant ces trentes dernières années. Elles sont nées avec les travaux de Nelson en théorie quantique des champs au début des années 70 et prennent le nom qu'elles portent aujourd'hui depuis l'article fondateur de Gross en 1975.

Définition 1.2.13. On dit qu'une mesure μ statisfait une inégalité de Sobolev logarithmique pour une certaine classe de fonctions $\mathcal{A}_{LS}$, lorsqu'il existe une constante $C_{LS} > 0$ telle que, pour toute fonction $f \in \mathcal{A}_{LS}$, on ait :

$$\mathrm{Ent}_\mu(f^2) \leq C_{LS} \int_E |\nabla f|^2 d\mu.$$

Où $\mathrm{Ent}_\mu(f) = \int_E (f \log f) d\mu - \left(\int_E f d\mu \right) \log \left(\int_E f d\mu \right)$ désigne l'entropie d'une fonction mesurable positive f.

Remarque. Nous rappelons le fait qu'une inégalité de Sobolev logarithmique entraine une inégalité de Poincaré et que nous avons la relation suivante entre les constantes C_{LS} et C_P. :

$$C_P \leq \frac{1}{2} C_{LS}.$$

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

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

en 0 nous donne :

$$\mathrm{Ent}_\mu\big((1+\epsilon f)^2\big) = \frac{\epsilon^2}{2}\mathrm{Var}_\mu(f) + o(\epsilon^2).$$

D'autre part, nous avons

$$\int_E [(1+\epsilon f)']^2 d\mu = \epsilon^2 \int_E f'^2 d\mu$$

En faisant tendre ϵ vers 0^+, nous obtenons une inégalité de Poincaré, de constante $\frac{1}{2}C_{LS}$.

Notons que l'ensemble $\mathcal{A}_{LS}$ contient strictement $\mathcal{A}_P$, une inégalité de Sobolev logarithmique permet donc de considérer une classe beaucoup plus grande de fonctions.

Proposition 1.2.14. *La mesure gaussienne γ_1 sur $\mathbb{R}$ vérifie une inégalité de Sobolev logarithmique avec une constante $C_{LS} = 2$, autrement dit*

$$\mathrm{Ent}_\gamma(f^2) \leq 2 \int_{\mathbb{R}} f'^2 d\gamma_1.$$

Tout comme les inégalités de Poincaré, les inégalités de Sobolev logarithmique jouissent d'une propriété de tensorisation indépendante de la dimension. Plus précisément, nous avons le résultat suivant [112]

Proposition 1.2.15. *Soient $(E, d_i, \mu_i), \quad i = 1, \dots, n$ une famille d'espaces mesurés alors, pour toute fonction f suffisamment régulière, la mesure produit μ^n vérifie l'inégalité suivante*

$$\mathrm{Ent}_{\mu^n}(f) \leq \sum_{i=1}^n \int_E \mathrm{Ent}_{\mu_i}(f) d\mu^n.$$

En particulier, si les mesures $\mu_i, \quad i = 1, \dots, n$ vérifient une inégalité de Sobolev logarithmique de constante $C_i > 0$ alors μ^n en vérifie également une, de constante $C = \max_{i=1,\dots,n} C_i$.

Ce résultat entraine que la mesure gaussienne standard γ_n sur $\mathbb{R}^n$ vérifie une inégalité de Sobolev logarithmique, c'est à dire, pour f suffisamment régulière,

$$\mathrm{Ent}_{\gamma_n}(f^2) \leq 2 \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n. \tag{1.12}$$

Nous allons voir dans la section suivante les conséquences d'une telle inégalité vis à vis de la concentration de la mesure.

Conséquence de l'inégalité de Sobolev logarithmique et concentration de la mesure

L'inégalité de Sobolev logarithmique étant plus forte que celle de Poincaré, il est naturel de s'attendre à obtenir des propriétés de concentration plus fortes. Ce fut l'observation de Herbst qui a remarqué qu'une telle inégalité entrainait un contrôle beaucoup plus précis de la transformée de Laplace. Plus précisement nous avons le résultat suivant :

Proposition 1.2.16. *Si μ satisfait une inégalité de Sobolev logarithmique de constante $C = C_{LS}$, alors pour toute fonction f 1-lipschitzienne et tout $\sigma^2 < \frac{1}{C}$ nous avons*

$$\int_E e^{\sigma^2 f^2/2} d\mu < \infty.$$

Plus précisément, toute fonction f 1-lipschitzienne est intégrable et pour tout $\lambda \in \mathbb{R}$

$$\int_E e^{\lambda f} \mathrm{d}\mu \leq e^{s \int_E f d\mu + C\lambda^2/2}.$$

Comme dans le cas de l'inégalité de Poincaré, ceci se retranscrit via l'argument de Chernoff en inégalité de déviation autour de la moyenne.

Proposition 1.2.17. *Si μ satisfait une inégalité de Sobolev logarithmique de constante $C = C_{LS}$, alors pour toutse fonctions f lipschitziennes et tout $t \geq 0$*

$$\mu\left(f \geq \int_E f d\mu + t\right) \leq e^{-t^2/2C\|f\|^2_{Lip}}.$$

Une inégalité de concentration autour de la moyenne est facilement obtenue en remplaçant f par $-f$ dans ce qui précède :

$$\mu\left(\left|f - \int_E f d\mu\right| \geq t\right) \leq 2e^{-t^2/2C\|f\|^2_{Lip}}, \quad t \geq 0$$

Une application fondamentale de ceci est un résultat de concentration pour le supremum de processus gaussien [82, 83, 117].

Proposition 1.2.18. *Soient (T, d) un espace métrique et $(X_s)_{s\in T}$ un processus gaussien, centré, vérifiant $\sup_{s\in T} X_s < \infty$ alors*

$$\mathbb{P}(\left(\left|\sup_{s\in T} X_s - \mathbb{E}[\sup_{s\in T} X_s]\right| \geq t\right) \leq 2e^{-t^2/2}, \quad r \geq 0 \tag{1.13}$$

Le plus remarquable dans cette inégalité est le fait qu'elle soit indépendante de la dimension. Cela permet notamment d'obtenir des résultats pour la mesure de Wiener en dimension infinie. Ce genre d'inégalité a été au coeur (sous une autre formulation) de l'étude menée par Talagrand sur le supremum de processus via des méthodes de chainage générique [153, 117].

1.3 Outils : Semi-groupes et méthodes d'interpolation

1.3.1 Notations

Cette section a pour objet de proposer des démonstrations élémentaires des inégalités fonctionnelles présentées précédemment. Nous nous inspirons de la présentation faite dans [111, 16], auxquels nous renvoyons le lecteur pour plus de détails. Dans un premier temps, nous introduirons quelques notations provenant de la théorie de Bakry-Émery afin de mettre en place les méthodes d'interpolation via l'utilisation de semi-groupes.

Considérons un espace mesurable $(E, \mathcal{F})$ equipé d'une mesure σ-finie μ. Lorsque μ est finie, nous la renormaliserons afin d'avoir une mesure de probabilité. Notons par $L^p = L^p(\mu)$, $1 \leq p \leq \infty$, les espaces de Lebesgue par rapport à la mesure μ, et nous poserons $\|\cdot\|_p$ pour désigner la norme L^p.

L'objet d'intérêt fondamental est la famille $(P_t)_{t\geq 0}$ d'opérateurs positifs agissant sur l'ensemble des fonctions bornées f sur E par

$$P_t f(x) = \int_E f(y) p_t(x, dy), \quad x \in E,$$

et satisfaisant la propriété dite de semi-groupe :

$$P_s \circ P_t = P_{s+t},\ s, t \geq 0,\ P_o = Id.$$

Les noyaux positifs $p_t(x, dy)$ sont appelés noyaux de transition. Nous supposerons toujours que les opérateurs P_t sont bornés et continus dans $L^2(\mu)$ au sens où, pour tout f dans $L^2(\mu)$, $\|P_t f\|_2 \leq C(t)\|f\|_2$ pour tout $t \geq 0$ et $\|P_t f - f\|_2 \to 0$ lorsque $t \to 0$.

Nous dirons que $(P_t)_{t\geq 0}$ est markovien si $P_t 1 = 1$ pour tout $t \geq 0$. Ces semi-groupes markoviens sont naturellement associés à des processus de Markov $(X_t)_{t\geq 0}$, à valeurs dans E, par la relation

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

Le mouvement brownien $(B_t)_{t\geq 0}$ à valeurs dans $\mathbb{R}^n$, issu de l'origine avec pour noyau (de la chaleur)

$$p_t(x, dy) = \frac{1}{(2\pi t)^{n/2}} e^{-|x-y|^2/2t} dy, \quad t \geq 0, \quad x \in \mathbb{R}^n,$$

est un exemple classique de tels semi-groupes. Néanmoins, nous travaillerons la plupart du temps (pour une raison qui deviendra évidente par la suite) avec le processus d'Ornstein-Uhlenbeck $(X_t)_{t\geq 0}$ qui peut être représenté comme le processus gaussien suivant

$$X_t = \sqrt{2} e^{-t} \int_0^t e^s dB_s,\ t \geq 0.$$

Le semi-groupe associé peut être représenter par la formule, dite de Mehler, suivante :

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

Un objet essentiel lié aux semi-groupes est leur générateur infinitésimal L. Notons par $\mathcal{D}(L)$ le domaine dans $L^2(\mu)$ du générateur infinitésimal du semi-groupe $(P_t)_{t\geq 0}$, celui-ci étant défini comme l'ensemble des fonctions f de $L^2(\mu)$ pour lesquelles la limite suivante existe

$$Lf = \lim_{t \to 0} \frac{1}{t}(P_t f - f).$$

Le semi-groupe $(P_t)_{t\geq 0}$ laisse le domaine $\mathcal{D}(L)$ stable, et pour toute fonction $f \in \mathcal{D}(L)$, $P_t f$ satisfait l'équation (dite de la chaleur) associée à L suivante ,

$$\partial_t P_t f = P_t L f = L P_t f. \tag{1.14}$$

Réciproquement, le générateur L et son domaine $\mathcal{D}(L)$ détermine complètement $(P_t)_{t\geq 0}$: il existe un unique semi-groupe $(P_t)_{t\geq 0}$ d'opérateurs bornés sur $L^2(\mu)$ satisfaisant (1.14) pour toutes les fonctions du domaine $\mathcal{D}(L)$ [164]. Par exemple, le semi-groupe du mouvement brownien a pour générateur un demi du laplacien Δ sur $\mathbb{R}^n$ et dans ce cas particulier (1.14) est l'équation de la chaleur classique. Dans notre contexte, avec le semi-groupe d'Ornstein-Uhlenbeck, nous avons le générateur suivant $L = \Delta - x \cdot \nabla$ qui peut être vu comme le laplacien de l'espace gaussien $(\mathbb{R}^n, \gamma)$ (on parle dans ce cas de Witten-Laplacien). L'équation de la chaleur associée à $L = \Delta - x \cdot \nabla$ correspond à une formulation probabiliste de l'équation, plus usuelle en théorie des dérivées partielles, de Fokker-Planck linéaire.

Certaines mesures μ auront un rôle particulier vis à vis du semi-groupe $(P_t)_{t\geq 0}$ lorsque celles-ci vérifient des propriétés d'invariance et de reversibilité. Plus précisément une mesure μ est dite réversible (en temps) par rapport à $(P_t)_{t\geq 0}$ si pour tout $f, g \in L^2(\mu)$,

$$\int_E f P_t g d\mu = \int_E g P_t f d\mu.$$

μ est invariante par rapport à $(P_t)_{t\geq 0}$ si pour toute fonction $f \in L^1(\mu)$,

$$\int_E P_t f d\mu = \int_E f d\mu.$$

Lorsque μ est finie, on peut choisir $g = 1$ pour constater que les mesures reversibles sont des mesures invariantes. Les mesures reversibles, respectivement invariantes, μ sont décrites de manière équivalentes comme celles pour lesquelles $\int fLgd\mu = \int gLfd\mu$, respectivement $\int Lfd\mu = 0$, pour tout $f, g \in \mathcal{D}(L)$. Concernant le semi-groupe d'Ornstein-Uhlenbeck, sa mesure invariante et réversible est la mesure gaussienne standard γ_n sur $\mathbb{R}^n$ (il s'agit de la raison principale qui permet de considérer ce semi-groupe dans un cadre gaussien). Etant donné un générateur infinitésimal L, on peut définir $\mathcal{E}$, l'énergie de Dirichlet (qui coïncide avec la définition donnée préalablement) du semi-groupe de Markov P_t, par

$$\mathcal{E}(f, g) = \int_E f(-Lg) d\mu.$$

Comme nous pouvons le vérifier dans le cas du semi-groupe d'Ornstein-Uhlenbeck, après une intégration par partie en espace, nous obtenons l'expression suivante :

$$\mathcal{E}(f, g) = \int_{\mathbb{R}^n} \nabla f \cdot \nabla g d\gamma_n.$$

Lorsque μ est une mesure de probabilité, nous dirons que $(P_t)_{t\geq 0}$ est ergodique si $P_t f \to \int_E f d\mu$ μ-presque sûrement lorsque $t \to \infty$.

1.3.2 Inégalités fonctionnelles par interpolation

Une des forces de la théorie de Bakry-Émery et des méthodes d'interpolation, est qu'elle permet de démontrer de manière élémentaire et synthétique les inégalités de Poincaré et de Sobolev logarithmique. Nous ne rentrerons pas dans le degré de généralité présenté dans l'ouvrage [16], notamment sur les triplets markoviens, les inégalités de courbures-dimensions ainsi que les inégalités de Poincaré ou Sobolev logarithmique dites locales. Néanmoins, il se peut que nous y fassions référence par endroits au cours de ce manuscrit de thèse, à nouveau nous inviterons le lecteur à se référer aux livres suivants [16, 17, 111]. Nous énonçons ci-dessous une proposition, regroupant les propriétés fondamentales satisfaites par le semi-groupe d'Ornstein-Uhlenbeck, celles-ci sont essentielles pour les démonstrations. Nous démontrerons ensuite l'inégalité de Poincaré, puis celle de Sobolev logarithmique pour la mesure gaussienne γ_n sur $\mathbb{R}^n$.

Proposition 1.3.1. *Le semi-groupe d'Ornstein-Uhlenbeck $(P_t)_{t\geq 0}$ admet pour mesure invariante la mesure gaussienne γ_n sur $\mathbb{R}^n$ et vérifie les assertions suivantes :*

1. *$(P_t)_{t\geq 0}$ est ergodique.*

2. *Les inégalités ci-dessous, sont valables pour tout $x \in \mathbb{R}^n$ que nous omettrons pour alléger les notations. $(P_t)_{t\geq 0}$ satisfait une relation de commutation exacte avec l'opérateur ∇, c'est à dire, pour toute fonction f suffisamment régulière,*

$$\nabla P_t f = e^{-t} P_t \nabla f, \quad t \geq 0. \tag{1.15}$$

En particulier, ceci fournit l'inégalité suivante, valable pour tout $t \geq 0$ et toute fonction f suffisamment régulière,

$$|\nabla P_t f|^2 = e^{-2t} |P_t(\nabla f)|^2 \leq e^{-2t} P_t(|\nabla f|^2). \tag{1.16}$$

L'inégalité de Cauchy-Schwarz implique également, pour tout $t \geq 0$ et $f : \mathbb{R}^n \to \mathbb{R}_+$ suffisamment régulière,

$$|\nabla P_t f|^2 \leq P_t\left(\frac{|\nabla f|^2}{f}\right) P_t(f) \tag{1.17}$$

3. *En considérant $(P_t)_{t\geq 0}$ comme une famille d'opérateurs de $L^p(\gamma_n)$, $p \geq 1$, dans lui-même. L'inégalité de Jensen nous affirme que ces opérateurs sont des contractions de L^p, autrement dit pour toute fonction $f \in L^p(\gamma_n)$ et tout $t \geq 0$,*

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

Inégalité de Poincaré

Nous débutons par le cas le plus simple : l'inégalité de Poincaré.

Proposition 1.3.2. *La mesure gaussienne standard γ_n sur $\mathbb{R}^n$ satisfait l'inégalité suivante, pour f suffisamment régulière,*

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

Démonstration. Par ergodicité du semi-groupe, puis avec le théorème fondamental de l'analyse on obtient le développement de la variance le long du semi-groupe suivant :

$$
\begin{aligned}
\mathrm{Var}_{\gamma_n}(f) &= \lim_{s\to\infty} \int_{\mathbb{R}^n} (P_0 f)^2 d\gamma_n - \left(\int_{\mathbb{R}^n} P_s f d\gamma_n \right)^2 && (1.18)\\
&= -\int_0^\infty \frac{d}{ds} \int_{\mathbb{R}^n} (P_s f)^2 d\gamma_n ds && (1.19)\\
&= -2 \int_0^\infty \int_{\mathbb{R}^n} P_s f L(P_s f) d\gamma_n ds && (1.20)\\
&= 2 \int_0^\infty \int_{\mathbb{R}^n} |\nabla P_s f|^2 d\gamma_n ds && (1.21)\\
&= 2 \int_0^\infty e^{-2s} \int_{\mathbb{R}^n} |P_s(\nabla f)|^2 d\gamma_n ds && (1.22)
\end{aligned}
$$

La dernière ligne est obtenue après une intégration par partie en espace, puis, en utilisant la relation (1.15) . Pour conclure, il suffit d'utiliser (1.16), le fait que le semi-groupe est une contraction des espaces $L^p(\gamma_n)$ et enfin en utilisant l'invariance de γ_n par rapport à $(P_t)_{t\geq 0}$.

$$
\begin{aligned}
2 \int_0^\infty \int_{\mathbb{R}^n} |\nabla P_s f|^2 d\gamma_n ds &= 2 \int_0^\infty e^{-2s} \int_{\mathbb{R}^n} |P_s(\nabla f)|^2 d\gamma_n ds\\
&\leq 2 \int_0^\infty e^{-2s} \int_{\mathbb{R}^n} P_s(|\nabla f|^2) d\gamma_n ds\\
&= \left(2 \int_0^\infty e^{-2s} ds \right) \left(\int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n \right)
\end{aligned}
$$

□

Remarque. Notons que la démonstration ci-dessus, fournit une représentation dynamique de la variance, le long du semi-groupe, qui sera au centre de certains des résultats obtenus durant cette thèse

$$\mathrm{Var}_{\gamma_n}(f) = 2 \int_0^\infty e^{-2s} \int_{\mathbb{R}^n} |P_s(\nabla f)|^2 d\gamma_n ds \tag{1.23}$$

Mentionnons également que, dans le cadre de la théorie de Bakry-Émery, l'inégalité de Poincaré peut s'exprimer, de manière équivalente par un critère de « courbure dimension intégré ». Pour la mesure gaussienne γ_n, il s'agit de l'équivalence suivante, pour $f \in \mathcal{D}(L)$,

$$\mathbb{E}_{\gamma_n}[\Gamma_2(f)] \geq \mathbb{E}_{\gamma_n}[\Gamma_1(f)] \quad \Longleftrightarrow \mathrm{Var}_{\gamma_n}(f) \leq \mathbb{E}_{\gamma_n}[\Gamma_1(f)], \tag{1.24}$$

avec $\Gamma_1(f) = |\nabla f|^2$ et $\Gamma_2(f) = \|\mathrm{Hess} f\|_2^2 + |\nabla f|^2$.

Dans un certain sens, la constante $C_P = 1$ de l'inégalité de Poincaré gaussienne dépend de la « pire » fonction de l'ensemble $\mathcal{D}(L)$. On peut se demander s'il est possible d'exhiber un ensemble plus petit $\mathcal{A} \subset \mathcal{D}(L)$ sur lequel une inégalité de Poincaré serait satisfaite avec une constante $\rho_{\mathcal{A}} < 1$. Il sera évident par la suite, via l'inégalité de Talagrand, que si $\mathcal{A}$ est constitué des statistiques d'ordres d'un échantillon gaussien de taille $n \geq 2$, on peut choisir $\rho_{\mathcal{A}} = C/\log n$.

Inégalité de Sobolev logarithmique

Comme nous allons le voir, ce type d'argument permet, de manière analogue, de démontrer une inégalité de Sobolev logarithmique.

Proposition 1.3.3. *Rappelons que γ_n satisfait une inégalité de Sobolev logarithmique, c'est-à-dire, pour toute fonction f suffisament régulière,*

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

Démonstration. En procédent de manière analogue à la démonstration de l'inégalité de Poincaré, on obtient le développement suivant de l'entropie d'une fonction f le long du semi-groupe d'Ornstein-Uhlenbeck :

$$\mathrm{Ent}_{\gamma_n}(f) = \int_0^\infty \int_{\mathbb{R}^n} \frac{|\nabla P_t f|^2}{P_t f} d\gamma_n dt$$

Il suffit ensuite d'utiliser l'item (1.17) et l'invariance du semi-groupe par rapport à la mesure gaussienne pour obtenir

$$\begin{aligned}\mathrm{Ent}_{\gamma_n}(f) &\leq \int_0^\infty \int_{\mathbb{R}^n} e^{-2t} P_t \left(\frac{|\nabla f|^2}{f} \right) d\gamma_n dt \\ &= \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} d\gamma_n.\end{aligned}$$

Le résultat s'ensuit en remplaçant f par f^2.

□

1.3.3 Hypercontractivité

Comme signifié auparavant, il est intéressant de considérer un semi-groupe $(P_t)_{t\geq 0}$ comme une famille d'opérateurs de $L^p(\mu), p \geq 1$, dans lui-même, où μ est la mesure invariante du semi-groupe. L'inégalité de Jensen nous affirme que ces opérateurs sont des contractions de $L^p(\mu)$, autrement dit pour toute fonction $f \in L^p(\mu)$ et tout $t \geq 0$,

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

Parfois, le semi-groupe peut vérifier une propriété plus forte dite d'hypercontractivité. C'est le contenu de la définition ci-dessous.

Définition 1.3.4. Considérons un semi-groupe $(P_t)_{t\geq 0}$ et sa mesure invariante μ. On dit que le semi-groupe $(P_t)_{t\geq 0}$ est hypercontractif si, pour tout $t \geq 0$, il existe des nombres $q > p > 1$ (pouvant dépendre de t) tels que pour tout $f \in \mathrm{L}^q(\mu)$:

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

Lorsque la mesure invariante μ satisfait une inégalité de Sobolev logarithmique, cela induit une propriété plus forte sur le semi-groupe sous-jacent. Il s'agit du théorème de Gross [91] qui démontre l'équivalence entre l'hypercontractivité du semi-groupe et le fait que la mesure invariante satisfasse une inégalité de Sobolev logarithmique. Cette observation est survenue la première fois dans la théorie quantique des champs. Nous énonçons ce résultat dans le cas de la mesure gaussienne γ_n sur $\mathbb{R}^n$ et du semi-groupe d'Ornstein-Uhlenbeck associé.

Théorème 1.3.5 (Nelson)**.** *Soient* $1 < p < q < \infty$.

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

$$\|P_t f\|_q \leq \|f\|_p. \tag{1.25}$$

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

A noter que ce résultat (1.25) 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 donnerons dans la section suivante une application de l'hypercontractivité pour obtenir une inégalité de Talagrand [147]. Cette inégalité, qui améliore l'inégalité de Poincaré, sera l'un des outils majeurs dans l'étude de la superconcentration.

Inégalité de Talagrand, un renforcement de l'inégalité de Poincaré

L'inégalité de Talagrand permet d'obtenir une amélioration de l'inégalité de Poincaré. De plus, sa mise en application est relativement simple puisqu'elle se résume à calculer des normes $L^1(\gamma_n)$ et $L^2(\gamma_n)$ de dérivées partielles. Voici son énoncé :

Théorème 1.3.6 (Talagrand)**.** *Soit* $f : \mathbb{R}^n \to \mathbb{R}$ *suffisament régulière. On désigne par* $\partial_i f$ *la dérivée partielle de* f *en sa* i*-ème coordonnée, alors*

$$\mathrm{Var}_{\gamma_n}(f) \leq C \sum_{i=1}^{n} \frac{\|\partial_i f\|_2^2}{1 + \log\left(\|\partial_i f\|_2 / \|\partial_i f\|_1\right)},$$

avec $C > 0$ *une constante numérique, où* $\|\cdot\|_p$ *désigne la norme de l'espace* $L^p(\gamma_n)$.

Avant de démontrer ce résultat, quelques commentaires s'imposent. L'inégalité de Poincaré pour la mesure gaussienne γ_n nous assure que

$$\mathrm{Var}_{\gamma_n}(f) \leq \sum_{i=1}^{n} \|\partial_i f\|_2^2.$$

Ainsi, nous avons bien un gain logarithmique au niveau du dénominateur. Pour que ce gain soit effectif il est important que la norme $L^2(\gamma_n)$ de $\partial_i f$, $i = 1, \dots, n$ soit beaucoup plus grande que la norme $L^1(\gamma_n)$ de cette même dérivée partielle. Illustrons, à présent, cette inégalité sur deux exemples. Soit $f(x) = \sum_{i=1}^n x_i$, alors $\partial_i f = 1$ et les normes L^1 et L^2 des dérivées partielles sont identiques. Dans ce cas, l'inégalité de Talagrand ne dit rien de mieux que l'inégalité de Poincaré (en fait, il est bien connu que le cas d'égalité dans Poincaré est atteint pour ce types de fonctions). Si maintenant, on choisit $f(x) = \max_{i=1,\dots,n} x_i$ alors, pour tout $1 \leq i \leq n$,

$$\partial_i f = 1_{\{x_i \geq x_j \, j=1,\dots,n\}},$$

puisque $f(x) = \sum_{i=1}^n X_i 1_{A_i}$ avec $A_i = \{x_i \geq x_j \, j = 1, \dots, n\}$, $\quad 1 \leq i \leq n$. Pour être parfaitement rigoureux, il aurait fallu utiliser une approximation régulière (de Gibbs, par exemple) afin de calculer les dérivées partielles de la fonction $\max_{i=1,\dots,n} x_i$. En effet, la fonction $F_\beta(x) = \frac{1}{\beta}\log\left(\sum_{i=1}^n e^{\beta x_i}\right)$ converge, uniformément en $x \in \mathbb{R}^n$, vers $\max_{i=1,\dots,n} x_i$ lorsque β tend vers l'infini. Ainsi, il s'agit uniquement d'un problème technique et notre simplification n'altèrera en rien le résultat final.

Finalement, dans notre exemple nous obtenons par symétrie,

$$\|\partial_i f\|_2^2 = \|\partial_i f\|_1 = \mathbb{P}(X_i \geq X_j \,\forall j) = \frac{1}{n}.$$

D'où, d'après l'inégalité de Talagrand,

$$\mathrm{Var}(M_n) \leq \frac{C}{\log n},$$

où $M_n = \max_{i=1,\ldots,n} X_i$, au lieu de $\mathrm{Var}(M_n) \leq 1$ d'après l'inégalité de Poincaré. Cet exemple est un prototype du phénomène de superconcentration, nous y reviendrons en détail dans le chapitre 2. Esquissons brièvement la preuve précédente avec l'approximation de Gibbs. En utilisant des arguments de symétrie et le fait que $\sum_{i=1}^n \|\partial_i F_{\beta,n}\|_1 = 1$, nous trouvons les estimées suivantes

$$\|\partial_i F_{\beta_n}\| = \frac{1}{n}, \quad i \in \{1, \ldots, n\}$$

et

$$\|\partial_i F_{\beta,n}\|_2^2 \leq \frac{1}{n}, \quad i \in \{1, \ldots, n\}.$$

Il suffit alors d'appliquer l'inégalité de Talagrand et de conclure par convergence dominée.

Démonstration. (de l'inégalité de Talagrand) L'idée principale de la preuve repose sur des méthodes de semi-groupes et d'interpolation, présentées dans la section 1.3.2, le long du flot d'Ornstein-Uhlenbeck. Le début de la preuve se déroule de manière similaire à la démonstration de l'inégalité de Poincaré. La majeure différence entre ces deux preuves provient de l'application de l'hypercontractivité du semi-groupe au lieu de l'inégalité de Jensen et du fait que l'on décompose l'énergie de Dirichlet $\mathcal{E}(f) = \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n$ suivant les « directions » $(\partial_i f)^2$ (c'est à dire $|\nabla f|^2 = \sum_{i=1}^n (\partial_i f)^2$). Cette observation est à l'origine d'un article de Cordero-Erausquin et Ledoux et sera réutilisée plus tard dans la thèse.

Nous rappelons au lecteur la formule de Melher du semi-groupe d'Ornstein-Uhlenbeck : pour toute fonction f suffisament régulière P_t est définie par

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

Notons que la propriété d'hypercontractivité de ce semi-groupe entraine que, pour toute fonction f suffisamment régulière, nous avons :

$$\|P_t f\|_2 \leq \|f\|_p, \quad t \geq 0,$$

avec $p = p(t) = 1 + e^{-2t}$. Nous utiliserons cette estimée, au lieu de la propriété de contraction fournie par l'inégalité de Jensen, puis appliquerons l'inégalité d'Hölder permettant l'interpolation de l'espace $L^p(\gamma_n)$ par les espaces $L^1(\gamma_n)$ et $L^2(\gamma_n)$.

$$
\begin{aligned}
\mathrm{Var}_{\gamma_n}(f) &= -\int_0^\infty \frac{d}{dt}\int_{\mathbb{R}^n}\big(P_t f\big)^2 d\gamma_n dt \\
&= 2\int_0^\infty e^{-2t}\int_{\mathbb{R}^n}\big|P_t(\nabla f)\big|^2 d\gamma_n dt \\
&= 2\sum_{i=1}^n \int_0^\infty e^{-2t}\|P_t(\partial_i f)\|_2^2 dt \\
&\le 2\sum_{i=1}^n \int_0^\infty e^{-2t}\|\partial_i f\|_{p(t)}^2 dt.
\end{aligned}
$$

Au vu de la propriété d'hypercontractivité satisfaite par $(P_t)_{t\ge 0}$, il est clair que $1 < p(t) < 2$. De plus, l'inégalité de Hölder entraine que

$$
\|\partial_i f\|_{p(t)} \le \|\partial_i f\|_1^\theta \|\partial_i f\|_2^{1-\theta},
$$

où $\theta = \theta(t) \in [0,1]$ vérifie $1/p(t) = \theta + (1-\theta)/2$.

Après un changement de variable, nous trouvons

$$
\mathrm{Var}_{\gamma_n}(f) \le \sum_{i=1}^n \|\partial_i f\|_2^2 \int_1^2 b^{2\theta(v)} dv,
$$

où $b = \|\partial_i f\|_1/\|\partial_i f\|_2 \le 1$. Enfin,

$$
\int_1^2 b^{2\theta(v)} dv \le \int_0^2 b^s ds \le \frac{2}{1+\log(1/b)},
$$

ce qui conclut la démonstration. □

Remarque. Le fait suivant sera utile plus loin dans la thèse : l'application $(x,y) \mapsto \frac{x^2}{1+\log(x/y)}$ est croissante en chacune des variables (tant que $x \ge y > 0$). C'est pourquoi, pour certains modèles il sera suffisant de majorer les normes $L^1(\gamma_n)$ et $L^2(\gamma_n)$ des dérivées partielles pour obtenir de la superconcentration. En fait, l'une des difficultés sera de majorer la probabilité que l'argmax soit atteint en la i-ième coordonnée. Le cas du vecteur gaussien standard est particulièrement facile puisque, par symétrie, l'argmax est uniformément réparti. Cependant nous verrons qu'il est possible d'obtenir des majorations similaires en imposant certaines restrictions sur la structure de corrélation du vecteur gaussien considéré.

L'inégalité de Talagrand est pratique car elle est simple d'utilisation, cependant elle est inefficace dans un grand nombre d'exemples (notamment les modèles de verres de spins). Néanmoins, Chatterjee a démontré dans [48] que lorsque la fonction considérée est monotone en ses coordonnées, l'inégalité de Talagrand fournissait le bon ordre de grandeur de la variance. Il s'agit du contenu du théorème suivant :

Théorème 1.3.7. *(Chatterjee)[Inégalité de Talagrand, version fonction monotone] Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction absolument continue, monotone en chacune de ses cordonnnées. On pose :*

$$
a = \frac{\mathrm{Var}_{\gamma_n}(f)}{\mathbb{E}_{\gamma_n}|\nabla f|^2}
$$

et

$$b = \frac{\sum_{i=1}^n \|\partial_i f\|_1^2}{\sum_{i=1}^n \|\partial_i f\|_2^2}.$$

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

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

où $C > 0$ est une constante numérique. Ceci implique donc que a est petit si et seulement si b l'est.

Extension de l'inégalité de Talagrand

L'argument d'hypercontractivité utilisé dans la démonstration de l'inégalité repose sur le fait suivant : la mesure invariante du semi-groupe est hypercontractive et que l'énergie de Dirichlet se décompose suivant des directions. De plus, une relation de commutation entre ces directions et le semi-groupe doit être satisfaite.

Plus précisément, comme observé par Cordero-Erausquin et Ledoux dans [62], qui considérent un cadre plus abstrait de triplet de Markov, nous supposons que la mesure invariante μ , du semi-groupe sous-jacent, est hypercontractive de constante $\rho > 0$ et que l'énergie de Dirichlet $\mathcal{E}$ peut-être décomposée le long de directions Γ_i pour lesquelles une relation de commutation avec le semi-groupe est satisfaite. C'est à dire, qu'il existe Γ_i, $i = 1, \ldots, n$ telles que, pour toute fonction régulière f,

$$\mathcal{E}(f) = \sum_{i=1}^n \int_E \Gamma_i^2(f) d\mu$$

et qu'il existe $\kappa \in \mathbb{R}$ tel que, pour tout $t \geq 0$,

$$P_t\big(\Gamma_i(f)\big) \le e^{\kappa t}\Gamma_i\big(P_t(f)\big), \quad i = 1, \ldots, n$$

Il est facile de voir que la mesure gaussienne γ_n rentre dans ce cadre avec le semi-groupe d'Ornstein-Uhlenbeck. On peut calculer $\mathcal{E}(f) = \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n$, avec $f \in \mathcal{D}(L)$ (où $\mathcal{D}(L)$ désigne le domaine de l'opérateur d'Ornstein-Uhlenbeck) et remarquer que $\mathcal{E}$ peut se décomposer comme suit :

$$\mathcal{E}(f) = \sum_{i=1}^n \int_{\mathbb{R}^n} (\partial_i f)^2 d\gamma_n = \sum_{i=1}^n \int_{\mathbb{R}^n} \Gamma_i(f)^2 d\gamma_n, \quad \Gamma_i(P_t f) \le e^{-t} P_t(\Gamma_i f).$$

Sous ces hypothèses, le principal résultat de [62] est le suivant

Théorème 1.3.8. *Pour $f : \mathbb{R}^n \to \mathbb{R}$ suffisament régulière, $n \geq 1$,*

$$\mathrm{Var}_\mu(f) \le C_{\rho,\kappa} \sum_{i=1}^n \frac{\|\Gamma_i(f)\|_2^2}{1+\log\left(\frac{\|\Gamma_i(f)\|_2}{\|\Gamma_i(f)\|_1}\right)}, \tag{1.26}$$

où $C_{\rho,\kappa} = 4e^{(1+(\kappa/\rho))_+}/\rho$.

Ce résultat sera utilisé dans le chapitre 4 de la thèse pour obtenir de la superconcentration pour le maximum des coordonnées d'un vecteur de loi uniforme sur la sphère $\mathbb{S}^n$ et pour montrer, de manière alternative, que la fonction maximum n'est pas superconcentrée sous la mesure gamma.

Remarque. En particulier, si μ est une mesure de probabilité sur $\mathbb{R}^n$, $n \geq 1$ dite strictement « log-concave ». C'est à dire admettant la densité suivante par rapport à la mesure de Lebesgue dx sur $\mathbb{R}^n$: $d\mu = e^{-V}dx$ avec $V : \mathbb{R}^n \to \mathbb{R}$ satisfaisant une condition de convexité, $\nabla\nabla V \geq \rho Id$ où $\rho > 0$. Alors l'inégalité de Cordero-Erausquin-Ledoux précédente est valable.

1.4 Transport optimal

Cette théorie est née à la fin du XVIIIème siècle sous l'impulsion du géomètre français Monge. En 1781, il publie l'un des travaux les plus célèbres « mémoire sur les théorie des déblais et des remblais ». Le problème considéré par Monge est le suivant : supposons que l'on dispose à un endroit E d'une quantité de sable que l'on souhaite déplacer sur un site de construction Y. Chacun des déplacements d'une partie du sable x de E à un emplacement y sur le site F coûte une certaine somme $c(x, y)$. En modélisant ceci (le tas de sable et le site de contruction) par des mesures de probabilités μ et ν, la théorie du transport optimal cherche à minimiser le coût (par la suite ce coût sera la distance d entre les points $x \in E$ et $y \in F$) d'un tel déplacement. Ce problème d'optimisation porte le nom de problème de Monge-Kantorovich, nous renvoyons le lecteur vers les livres [163, 162] pour plus détails à ce sujet.

Concernant notre thématique, l'approche via le transport optimal ([163, 162, 88, 113]) fournit une troisième description du phénomène de la concentration de la mesure. Nous nous inspirerons de l'exposition faite dans [113].

Débutons par l'inégalité classique de Pinsker-Csizsar-Kullback [134], celle-ci indique que pour n'importe quelle mesure de probabilité μ et ν la relation suivante est vérifiée

$$\|\mu - \nu\|_{TV} \leq \sqrt{\frac{1}{2}H(\nu|\mu)},$$

où $\|\cdot\|_{TV}$ désigne la distance en variation totale et $H(\nu, \mu)$ est l'entropie relative de ν par rapport à μ. Plus précisément,

$$H(\nu|\mu) = \int \log \frac{d\nu}{d\mu} d\nu,$$

lorsque ν est absolument continue par rapport à μ (et $\frac{d\nu}{d\mu}$ désigne alors la dérivée de Radon-Nykodym de ν par rapport à μ) et $+\infty$ sinon. De telles inégalités ont souvent été considérées en théorie de l'information.

Le fait qu'une telle inégalité soit liée à des propriétés de concentration a été remarqué par Marton [124, 123] et peut-être présentée de la façon suivante. Considérons un espace métrique (E, d) et des mesures de probabilités boreliennes sur E μ et ν, on définit alors la distance de Kantorovich-Wasserstein entre μ et ν par

$$\mathrm{W}_1(\mu, \nu) = \inf \int\int_{E\times E} d(x, y) d\pi(x, y),$$

où l'infimum parcourt toutes les mesures de probabilités μ sur l'espace produit $E \times E$ admettant μ et ν comme marginales (on dit que π est un couplage de μ et ν) et possédant un moment d'ordre un (par rapport à d) fini. Le théorème de Monge-Kantorovich (cf. [163, 162]) assure que la distance en variance totale correspond au choix de la métrique triviale sur E. Soit μ, une mesure de probabilité fixée, considérons l'inégalité suivante

$$\mathrm{W}_1(\mu,\nu) \le \sqrt{2CH(\nu|\mu)}, \tag{1.27}$$

avec $C > 0$ et ν quelconque. Soient A et B deux ensembles boréliens de mesure $\mu(A), \mu(B) > 0$ et considérons les deux probabilités conditionnelles $\mu_A = \mu(\cdot|A)$ et $\mu_B = \mu(\cdot|B)$. En appliquant l'inégalité triangulaire à la distance W_1 et (1.27), il s'ensuit que

$$\begin{aligned} W_1(\mu_A,\mu_B) &\le W_1(\mu,\mu_A) + W_1(\mu,\mu_B) \\ &\le \sqrt{2CH(\mu_A|\mu)} + \sqrt{2CH(\mu_B|\mu)} \\ &= \sqrt{2C\log\frac{1}{\mu_A}} + \sqrt{2C\log\frac{1}{\mu_B}} \end{aligned}$$

En outre, toutes les mesures ayant pour marginales μ_A et μ_B doivent avoir leur support dans $A \times B$, ceci entraine, par définition de W_1,

$$W_1(\mu_A,\mu_B) \ge d(A,B) = \inf\{d(x,y)\,;\, x \in A,\, y \in B\}.$$

En combinant ceci avec l'inégalité $W_1(\mu_A,\mu_B) \le \sqrt{2C\log\frac{1}{\mu_A}} + \sqrt{2C\log\frac{1}{\mu_B}}$, nous obtenons de la concentration pour la mesure μ. En effet, étant donné A et B dans E tels que $d(A,B) \ge t > 0$, nous obtenons

$$t \le \sqrt{2C\log\frac{1}{\mu(A)}} + \sqrt{2C\log\frac{1}{\mu(A_t)}}, \tag{1.28}$$

avec, rappelons-le, $A_t = \{x \in E\,;\, d(x,A) < t\}$. L'inégalité (1.28) apparait alors comme une forme de concentration. Si $\mu(A) \ge 1/2$, on a donc

$$t \le \sqrt{2C\log 2} + \sqrt{2C\log\frac{1}{\mu(A_t)}},$$

autrement dit, lorsque, par exemple, $t \ge 2\sqrt{2C\log 2}$,

$$1 - \mu(A_t) \le e^{-t^2/8C}.$$

De manière analogue aux transformées de Laplace, la distance de Kantorovich-Wasserstein W_1 se comporte mal vis-à-vis de la dimension par rapport aux produits euclidiens lorsque l'on souhaite tensoriser une inégalité telle que (1.27) (pour plus de détails voir [87, 112]). L'exemple des mesures gaussiennes suggère qu'il serait plus intéressant de considérer un coût quadratique pour atteindre des inégalités de concentration indépendantes de la dimension. Par simplicité, nous nous restreignons à l'espace euclidien $\mathbb{R}^n$ muni de sa norme $|\cdot|_2$. Etant donnée une mesure de probabilité μ sur les boréliens de $\mathbb{R}^n$, on dit que celle-ci satisfait une inégalité de transport de coût quadratique lorsqu'il existe une constante $C > 0$ telle que pour toute mesure de probabilité ν,

$$W_2(\mu,\nu) \le \sqrt{CH(\nu|\mu)}, \tag{1.29}$$

où W_2 désigne la distance de Kantorovich-Wasserstein avec le coût quadratique

$$W_2(\mu,\nu) = \inf \left(\int\int_{E\times E} \frac{1}{2}|x-y|^2 d\pi(x,y) \right)^{1/2}$$

où l'infimum parcourt toutes les mesures de probabilités μ sur l'espace produit $\mathbb{R}^n \times \mathbb{R}^n$ admettant μ et ν comme marginale (cette quantité est finie dès lors que μ et ν possèdent un moment d'ordre deux, ce que nous supposerons toujours dans la suite). Il est évident, par l'inégalité de Jensen, que l'inégalité de transport de coût quadratique est plus forte que l'inégalité (1.27) (avec le coût $d(x,y) = |x-y|$, $x,y :\in \mathbb{R}^n$).

Il a été démontré par Talagrand que la mesure gaussienne standard γ_n sur $\mathbb{R}^n$ satisfait l'inégalité (1.29) avec la consante $C = 1$ [150]. Il est instructif de voir la démonstration, qui repose sur des arguments de transport optimal, en dimension un. Cette méthodologie de transport de mesure nous sera utile lorsque nous aborderons la superconcentration via les inégalités de Poincaré à poids dans le chapitre sept.

Considérons donc le cas $n = 1$. Soit $f \geq 0$ telle que $\int f d\gamma_1 = 1$ et posons $d\nu = f d\mu$. Par simplicité, supposons que $f > 0$ sur tout $\mathbb{R}$. On définit l'application de transport monotone $T : \mathbb{R} \to \mathbb{R}$ par

$$\nu\big(] - \infty, T(x)]\big) = \gamma_1\big(] - \infty, x]\big), \quad x \in \mathbb{R},$$

de telle sorte à ce que ν soit la mesure image de γ par l'application T. La formule de changement de variable nous fournit l'équation suivante, dite de Monge-Ampère,

$$f\big(T(x)\big)T'(x)e^{-T(x)^2/2} = e^{-x^2/2}, \quad x \in \mathbb{R}.$$

Ainsi, en prenant le logarithme de l'équation précédente, nous obtenons pour tout $x \in \mathbb{R}$,

$$\log f\big(T(x)\big) + \log T'(x) - \frac{1}{2}T(x)^2 = -\frac{1}{2}x^2$$

En intégrant cette égalité par rapport à la mesure γ et en utilisant le fait que $\nu = T_{\#}\gamma_1$, nous trouvons

$$\int \log f d\nu = \frac{1}{2}\int [T(x)^2 - x^2] d\gamma_1 - \int \log T' d\gamma_1.$$

Une intégration par partie fournit que

$$\int x(T-x)d\gamma_1 = \int (T'-1)d\gamma_1$$

d'où

$$\begin{aligned} \int \log f d\nu &= \frac{1}{2}\int |x-T(x)|^2 d\gamma_1 + \int [T' - 1 - \log T'] d\gamma_1 \\ &\geq \frac{1}{2}\int |x-T(x)|^2 d\gamma_1 \end{aligned}$$

puisque $y \mapsto y - 1 - \log y \geq 0$ lorsque $y \geq 0$. Puisque ν est l'image de γ_1 par l'application de transport T, la mesure image π de γ_1 par l'application $x \mapsto \big(x, T(x)\big)$ a pour marginales γ_1 et ν respectivement. Ceci entraine, par définition de W_2,

$$\frac{1}{2}\int |x-T(x)|^2 d\gamma_1 = \int\int \frac{1}{2}|x-y|^2 d\pi \geq W_2^2(\gamma_1,\nu),$$

ce qui conlut la démonstration. Le cas n-dimensionnel s'ensuit facilement par tensorisation comme indiqué dans l'article de M. Talagrand. L'argument est formalisé dans la proposition suivante, (cf. [112]).

Proposition 1.4.1. *Soit $P = \mu_1 \otimes \ldots, \otimes \mu_n$ une mesure de probabilité sur les boréliens de $\mathbb{R}^n$. Supposons que pour tout $i = 1, \ldots, n$, μ_i satisfait une inégalité de transport quadratique*

$$W_2(\mu_i, \nu_i) \leq \sqrt{C_1 H(\mu_i | \nu_i)},$$

pour n'importe quelle mesure ν_i sur $\mathbb{R}$ et $C_1 > 0$. Alors,

$$W_2(P, R) \leq \sqrt{\max_{i=1,\ldots,n} C_i H(R|P)},$$

pour n'importe quelle mesure de probabilité R sur $\mathbb{R}^n$.

Le cas n-dimensionel peut également être obtenu de manière similaire au cas uni-dimensionel (sans utiliser de tensorisation). L'idée est d'utiliser l'application de transport monotone de Brenier-Mc Cann [61]. Soient μ et ν deux mesures de probabilité sur $\mathbb{R}^n$, une application $T : \mathbb{R}^n \to \mathbb{R}^n$ envoit μ sur ν (ou transporte μ sur ν) si ν est l'image de la mesure ν par T. Autrement dit, pour toute fonction borélienne positive bornée $f : \mathbb{R}^n \to \mathbb{R}$,

$$\int f(y) d\mu(y) = \int f\big(T(x)\big) d\nu(x).$$

Si μ et ν admettent un moment d'ordre deux, une application T poussant μ sur ν est dite optimale par rapport à la distance de Kantorovich-Wasserstein W_2 si

$$W_2(\mu, \nu)^2 = \frac{1}{2} \int |x - T(x)|^2 d\mu(x).$$

Un résultat fondamental de Brenier [46] et Mc Cann [125] (cf. [163, 162]) assure que lorsque μ est absolument continu par rapport à la mesure de Lebesgue, il existe une fonction convexe ϕ telle que $T = \nabla\phi$ transporte μ sur ν de manière optimale (au sens précédent).

Soit $\mu = \gamma_n$ la mesure gaussienne standard dans $\mathbb{R}^n$ et supposons que $d\nu = f d\gamma_n$ avec $f \geq 0$ et $\int f \mathrm{d}\gamma_{\mathrm{n}} = 1$. Lorsque celui-ci fait sens, la formule de changement de variables dans le transport de γ_n à ν fournit l'équation de Monge-Ampère suivante :

$$f\big(T(x)\big) \det\big(\mathrm{Hess}\phi(x)\big) e^{-|T(x)|^2/2} = e^{-|x|^2/2}, \quad x \in \mathbb{R}^n,$$

où Hessϕ désigne la hessienne de ϕ (nous ignorons les problèmes de régularité et le fait que $T = \nabla\phi$ ne pourrait exister que presque partout). En reproduisant la démonstration du cas uni-dimensionel et en utilisant le fait que

$$\log \det\big(\mathrm{Hess}\phi(x)\big) \leq \Delta\phi - n = \Delta\big(\phi - \frac{|x|^2}{2}\big),$$

on obtient que $W_2(\gamma_n, \nu) \leq \sqrt{H(\nu|\gamma_n)}$. Cette argument s'étend aisement aux mesures de probabilités $d\mu = e^{-V} dx$ avec un potentiel strictement convexe V.

Théorème 1.4.2. *Soit $d\mu = e^{-V} dx$ où $\mathrm{Hess}V(x) \geq cId$, $c > 0$ uniformément en $x \in \mathbb{R}^n$. Alors, pour toute mesure de probabilité ν sur $\mathbb{R}^n$,*

$$W_2(\nu, \mu) \leq \sqrt{\frac{1}{c} H(\nu|\mu)}$$

Remarque. Vis à vis de la hiérarchie des inégalités fonctionnelles, une telle inégalité s'intercale entre celle de Poincaré et celle de Sobolev logarithmique. C'est-à-dire qu'une inégalité de transport quadratique est plus forte que celle de Poincaré mais moins forte qu'une inégalité de Sobolev logarithmique. Pour plus de détails, nous renvoyons le lecteur vers [163]. Il également possible de démontrer une telle inégalité par des arguments de semi-groupe, c'est notamment l'approche employée par Bobkov, Gentil et Ledoux dans [31] via le semi-groupe d'Hamilton-Jacobi.

Plus généralement, la littérature concernant les liens entre les inégalités fonctionnelles, les méthodes de transport optimal et les méthodes de semi-groupe est très vaste (cf. [163, 162, 88]). Mentionnons, tout particulièrement, l'article de Bobkov, Gentil et Ledoux [31] concernant le semi-groupe Hamilton-Jacobi et l'inégalité de Sobolev logarithmique.

Chapitre 2

Phénomène de superconcentration

2.1 Introduction

Comme nous l'avons déjà mentionné dans la partie précédente, la théorie de la concentration de la mesure s'est imposée comme un outil incontournable dans différents domaines des mathématiques. Cependant une telle généralité et polyvalence risque de manquer de précision dans certains cas particuliers, lorsque cela arrive on parlera de superconcentration. Chatterjee fut le premier à essayer de formaliser cette notion et de collecter différents exemples pour lesquels la théorie de la concentration est sous-optimale [48]. La partie qui suit a donc pour objectif de présenter plus en détails le phénomène de superconcentration. Nous nous focaliserons, la plupart du temps, sur le cas de la mesure gaussienne γ_n sur $\mathbb{R}^n$. Après avoir expliqué les enjeux et la problématique de cette thèse nous énoncerons les différents outils existants pour exhiber de la superconcentration. Ce phénomène apparait dans différents contextes : verres de spin, percolation dirigé, champ libre discret gaussien, matrices aléaoires, statistiques d'ordres ... ; les techniques utilisées pour prouver qu'un modèle est superconcentré sont souvent *ad hoc* et inhérentes au modèle étudié. Nous tenterons de lister les travaux et articles récemment parus et non présents dans l'ouvrage [48] et d'expliquer briévement leurs approches.

Comme première manifestation de la concentration de la mesure gaussienne sur $\mathbb{R}^n$, nous avons montré dans le chapitre un qu'elle satisfaisait une inégalité de Poincaré par rapport au carré du champ usuel (il s'agit d'une appelation standard de l'énergie de Dirichlet dans ce cas particulier).

Théorème 2.1.1. *Pour $f : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulière,*

$$\mathrm{Var}_{\gamma_n}(f) \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n(x) \tag{2.1}$$

Nous avons également vu que le choix de la fonction $f(x) = \max_{i=1,\ldots,n} x_i$ entrainait le résultat suivant,

Proposition 2.1.2. *Soit X un vecteur gaussien dans $\mathbb{R}^n$ de matrice de covariance Γ, alors l'inégalité suivante est vérifiée*

$$\mathrm{Var}(M_n) \leq \max_{i=1,\ldots,n} \mathrm{Var}(X_i), \tag{2.2}$$

où $M_n = \max_{i=1,\ldots,n} X_i$.

Remarque. Nous avons déjà mentionné le fait que cette inégalité est optimale et ne peut-être améliorée. De plus, soulignons le, cette inégalité ne dépend pas de la matrice de covariance. Elle ne prend donc pas en compte la structure des corrélations du vecteur gaussien. Ainsi, il est possible que la généralité de l'inégalité (2.2) fournisse des résultats sous-optimaux pour des cas particuliers. Par exemple, appliquons l'inégalité (2.2) lorsque $\Gamma = I_d$ la matrice identité de $\mathbb{R}^n$. Nous trouvons que,

$$\operatorname{Var}(M_n) \leq 1.$$

Cependant, pour un tel choix de matrice de covariance, toutes les coordonnées du vecteur gaussien X sont indépendantes. Il est donc possible de faire directement le calcul de la variance de $M_n = \max_{i=1,\ldots,n} X_i$ et d'obtenir

$$\operatorname{Var}(M_n) \leq \frac{C}{\log n}, \tag{2.3}$$

avec $C > 0$ une constante numérique. Il apparait donc que l'inégalité (2.1), valable pour une grande classe de fonction, provenant de la théorie, dite ≪ classique ≫, de la concentration fournit une borne sous-optimale pour le choix particulier de la fonction $f(x) = \max_{i=1,\ldots,n} x_i$. Selon la terminologie introduite par Chatterjee dans [48] il s'agit du phénomène de superconcentration.

Cet exemple simple fournit la première question, abordée dans [48], que nous developperons dans cette thèse : ≪ si la borne fournie par l'inégalité de Poincaré (2.1) est sous-optimale, comment obtenir le bon ordre de grandeur ? ≫.

De manière plus formelle, la définition de la notion de superconcentration s'énonce comme suit

Définition 2.1.3. Soit μ une mesure de probabilité satisfaisant une inégalité de Poincaré (1.2.1) de constante C_p. Alors, une fonction $f : E \to \mathbb{R}$ est dite superconcentrée s'il existe $0 < \epsilon < 1$ tel que

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

Il faut comprendre ceci de la manière suivante : ϵ est petit et peut dépendre de la dimension de E, par exemple $\epsilon = \epsilon_n \to 0$ lorsque $n \to \infty$ (dans (2.3) $\epsilon = C/\log n$). Nous nous focaliserons principalement sur le cas de la fonction maximum, néanmoins nous présenterons aussi d'autres fonctionnelles, comme la médiane où les normes l_p, ou encore l'énergie libre F_β qui exhibent un phénomène de superconcentration. Néanmoins, nous souhaitons souligner le fait que le phénomène de concentration classique, induit par une inégalité de Poincaré par exemple, concerne une famille de fonctions tandis que le phénomène de superconcentration n'est valable, *a priori*, que pour une fonction particulière.

Le lecteur curieux pourrait se demander si des propriétés plus fortes de concentration satisfaites par la mesure μ peuvent améliorer la borne obtenue via (2.1) et fournir le bon ordre de grandeur de la variance. Par exemple, la mesure gaussienne sur $\mathbb{R}^n$ vérifie l'inégalité de Sobolev logarithmique (1.12). Rappelons que cette inégalité, via l'argument de Herbst, entraîne la propriété de concentration suivante

Théorème 2.1.4. *Soient $X_1, \ldots, X_n$ des variables aléatoires gaussiennes i.i.d. et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction lipschitzienne avec $\|f\|_{Lip} \leq 1$. Alors, pour tout $t \geq 0$,*

$$\mathbb{P}\big(|f(X_1, \ldots, X_n) - \mathbb{E}[f(X_1, \ldots, X_n)| \geq t\big) \leq 2e^{-t^2/2} \tag{2.4}$$

A nouveau, comme pour l'inégalité de Poincaré, constatons que cette inégalité est optimale dans le sens où le cas d'égalité est atteint lorsque f est une fonction linéaire. De manière similaire à l'inégalité de Poincaré, la grande généralité de (2.4) risque de fournir des résultats sous-optimaux pour des choix de fonctions f particuliers. Par exemple, au niveau de la variance de f, (2.4) implique, après intégration, que pour toute fonction lipschitzienne f avec $\|f\|_{Lip} \leq 1$,

$$\mathrm{Var}_{\gamma_n}(f) \leq 4.$$

En particulier, si l'on considère la fonction $f(x_1, \dots, x_n) = \max_{i=1,\dots,n} x_i$ nous obtenons que

$$\mathrm{Var}(M_n) \leq 4.$$

Comme nous l'avons déjà vu plus haut, cette borne est sous-optimale. En conclusion, l'inégalité de concentration (2.4) ne reflète pas le véritable comportement du maximum au niveau de la variance. Cependant, elle pourrait être pertinente vis à vis d'un théorème de convergence en loi. Ce qui entraine la considération suivante : si jamais la taille de la variance est correctement estimée, il est naturel de se demander si une convergence en loi est satisfaite. Il est aisé d'illustrer ce propos par l'exemple de la fonction maximum et la théorie des extrêmes [67, 108, 78]. Il est bien connu que M_n peut-être renormalisé afin d'obtenir une convergence en loi. Plus précisement, si l'on choisit $a_n = \sqrt{2\log n}$ et $b_n = a_n - \frac{\log\log n + \log(4\pi)}{2a_n}$ alors la convergence en loi suivante est satisfaite

$$a_n(M_n - b_n) \to \Lambda_0, \quad n \to \infty,$$

où Λ_0 désigne la loi de Gumbel ayant pour fonction de répartition $\mathbb{P}(\Lambda_0 \leq t) = e^{-e^{-t}}, \quad t \in \mathbb{R}$. Il serait donc intéressant d'obtenir une inégalité de concentration non-asymptotique reflétant cette convergence en loi (et donc l'ordre correct de la variance). Dans le cas de la convergence des extrêmes, auquel nous reviendrons régulièrement durant cette thèse, nous aimerions obtenir quelque chose du type

$$\mathbb{P}(a_n(M_n - b_n) \geq t) \leq g_1(t) \quad \text{et} \quad \mathbb{P}(a_n(M_n - b_n) \leq -t) \leq g_2(t) \quad t \geq 0,$$

avec $g_i : \mathbb{R}_+ \to \mathbb{R}$, $i = 1, 2$ telle que les asymptotiques de g reflètent les queues de distribution de la loi de Gumbel. C'est à dire,

$$g_1(t) \simeq e^{-t}, \, t \to +\infty \qquad g_2(t) \simeq e^{-e^t} \, t \to +\infty.$$

Il est naturel mais naïf de voir ce que l'on peut obtenir avec le théorème 2.1.4. Celui-ci entraine que, pour tout $t \geq 0$,

$$\mathbb{P}(a_n|M_n - b_n| \geq t) \leq 2e^{-t^2/2a_n^2}.$$

Cette inégalité se comporte très mal vis à vis de la dépendance en n, elle devient inutile lorsque $n \to \infty$ et elle ne reflète pas les asymptotiques de la loi de Gumbel. En revanche, elle reste pertinente, si l'on fait dépendre t de n et que l'on choisit $t = u\sqrt{a_n}$, $u \geq 1$.

Cet exemple introduit la deuxième problématique à laquelle nous avons tenté d'apporter des éléments de réponse durant cette thèse :

Supposons qu'on ait réussi à obtenir le bon ordre de grandeur de la variance et que M_n satisfasse une convergence en loi (après une éventuelle renormalisation). Est-il possible d'obtenir une inégalité de concentration reflétant à la fois le bon

ordre de grandeur de la variance et les asymptotiques des queues de distributions de la loi limite ?

Comme nous allons le voir, ceci n'est pas simple et il ne semble pas être possible de trouver une méthode générale permettant de traiter un modèle gaussien quelconque. La section suivante est dédiée à la présentation des différents travaux portant sur différents modèles gaussiens. Nous insistons sur le fait que chaque modèle présenté possède des structures de corrélations fondamentalement différentes les unes des autres et que des méthodes *ad hoc* sont utilisées pour chacun d'entre eux. Ceci amène le problème ouvert suivant, déjà présent dans [48] :

Est-il possible de trouver des conditions sur la structure de covariance qui serait équivalentes au phénomène de superconcentration ?

Après avoir effectué un survol de la littérature et des différents modèles (gaussiens, pour la plupart) qui présentent un phénomène de superconcentration, nous allons présenter une nouvelle inégalité, améliorant l'inégalité de Poincaré, reposant sur l'hypercontractivité du semi-groupe d'Ornstein-Uhlenbeck. Celle-ci est due à Chatterjee et est exposée dans [48]. L'autre inégalité, plus utilisée, de par sa souplesse pour obtenir de la superconcentration, est celle de Talagrand que l'on a présentée dans le chapitre précédent. Ces résultats sont les premiers outils suffisament souples pour être appliqués à des situations variées. Cependant, le défaut de ces inégalités est qu'elles ne permettent d'obtenir qu'un gain logarithmique (vis à vis de l'inégalité de Poincaré). Pour un nombre limité de modèles cela permet d'atteindre le bon ordre de grandeur de la variance tandis que pour la plupart des autres la nouvelle borne obtenue reste encore sous-optimale. Cette approche hypercontractive, utilisant la représentation de la variance le long du semi-groupe d'Ornstein-Uhlenbeck, est un des axes de recherche majeurs de cette thèse.

2.2 Modèles

Dans cette section, nous allons présenter différents modèles présentant de la superconcentration. Nous esquisserons, pour chaque thématique, les outils utilisés pour obtenir de la superconcentration. Il sera alors évident pour le lecteur, que chaque exemple nécessite des outils mathématiques spécifiques, inhérents au modèle et parfois très techniques, et qu'une méthodologie générique ne semble pas envisageable. Bien que ce panorama de la littérature existante a pour vocation d'être exhaustif, il n'est que le reflet de nos connaissances et pourrait, involontairement, omettre certains travaux mathématiques.

2.2.1 Matrices aléatoires

Dans cette section nous allons brièvement décrire des résultats concernant les matrices aléatoires. Une attention toute particulière sera donnée aux ensembles de matrices gaussiens Nous indiquerons par une remarque lorsque les résultats présentés restent valides dans un cadre plus général. Nous renvoyons le lecteur vers les ouvrages [7, 14] pour plus de détails.

Historiquement la théorie des matrices aléatoires est née durant les années 20 sous l'impulsion des mathématiciens Wigner et Wishart. Dès lors, ce domaine des mathématiques à connu un développement intense. L'un des résultats majeurs de cette théorie concerne l'étude de la plus grande valeur propre $\lambda_{\max}$ d'une matrice aléatoire. Sous certaines conditions de régularité cette valeur propre converge presque sûrement vers une valeur deterministe [7, 114]. Un résultat de fluctuations pour $\lambda_{\max}$ (sorte d'analogue du théorème de la limite centrale pour des matrices aléatoires) est, par la suite, obtenu pour différents modèles de matrices aléatoires.

Nous présentons ci dessous la définition de l'ensemble unitaire gaussien GUE (pour « Gaussian Unitary Ensemble ») et de l'ensemble orthogonale gaussien GOE (pour « Gaussian Orthogonal Ensemble »), nous fournirons également la définition d'une matrice, dite, de covariance et les ensembles gaussiens associés. Nous décrirons ensuite les résultats de fluctuations concernant la plus grande valeur propre d'une matrice appartenant à ces ensembles. Enfin, nous ferons un rapide historique de l'évolution des résultats non-asymptotiques, exhibant un phénomène de superconcentration, connu à ce jour.

Etude asymptotique de $\lambda_{\max}$

Considérons, pour tout entier $n \geq 1$, la matrice (de taille $n \times n$) auto-adjointe gaussienne centrée $X = X^n = (W_{ij}^n)_{1\leq i,j\leq n}$, de variance σ^2. Ceci revient à considérer une matrice hermitienne X de taille $n \times n$ dont les entrées, au dessus de la diagonale, sont des variables aléatoires gaussiennes complexes (respectivement réelles sur la diagonale) centrées et de variance σ^2 (autrement dit, la partie réelle et imaginaire sont des variables aléatoires gaussiennes centrées de variance $\sigma^2/2$). De manière équivalente, la matrice aléatoire X possède la densité de probabilité suivante

$$\mathbb{P}(dX) = \frac{1}{Z}\exp\big(-\mathrm{Tr}(X^2)/2\sigma^2\big)dX$$

sur l'espace des matrices hermitiennes $\mathcal{H}_n \cong \mathbb{R}^{n^2}$ de taille $n \times n$, avec

$$dX = \prod_{1\leq i\leq n} dX_{ii} \prod_{1\leq i<j\leq n} d\mathrm{Re}(X_{ij})d\mathrm{Im}(X_{ij})$$

la mesure de Lebesgue sur $\mathcal{H}_n$ et $Z = Z_n$ est une constante de renormalisation. Nous désignerons par $\lambda_{\max}$ la plus grande valeur propre d'une matrice du GUE.

Théorème 2.2.1. *(Tracy-Widom)*
Soit $X = X^n$, une matrice du GUE, avec $\sigma^2 = \frac{1}{4n}$, alors la suite $(\lambda_{max}^n)_{n\in\mathbb{N}}$, proprement renormalisée, converge en loi vers une distribution indépendante du modèle. Plus précisément,

$$n^{2/3}(\lambda_{max}^n - 1) \to TW, \quad n \to \infty$$

où TW désigne la loi de Tracy-Widom.

Remarque. La fonction de répartition F_{TW} de cette loi peut s'exprimer de manière plus explicite (cf. [114]) Ce résultat de convergence reste vrai si l'on considère des ensembles de matrices plus généraux, on parle alors de matrice de Wigner, en imposant des conditions sur les moments de telles matrices. Citons par exemples les travaux de Péché-Soshnikov [133], Ruzmaikina [136], Khorunzhiy[103] ou encore Lee-Yin [118].

Définition 2.2.2. Soit X une matrice aléatoire de taille $m \times n$ dont les coefficients sont indépendants, de moyenne nulle et de variance 1. On suppose également, dans le cas où les coefficients sont complexes, que leurs parties réelles et imaginaires sont indépendantes. La matrice $S_{m,n} = \frac{1}{m}X^*X$, de taille $n \times n$, est une matrice de covariance empirique.

Remarque. La matrice $S_{n,m}$ étant semi-définie positive, ses n valeurs propres sont réelles, positives ou nulles et sont notées $0 \leq \lambda_1 \leq \ldots, \leq \lambda_n = \lambda_{\max}$. Si les coefficients de la matrice X suivent une loi gaussienne, la matrice est dite appartenir à l'ensemble unitaire de Laguerre (LUE) ou à l'ensemble orthogonal de Laguerre (LOE) selon qu'elle est complexe ou réelle. La densité de probabilité d'une matrice aléatoire appartenant au LUE (ou au LOE) peut s'exprimer de manière similaire à celle d'une matrice du GUE (ou du GOE).

Le comportement de la plus grande valeur propre d'une matrice du LUE est semblable à celui de la plus grande valeur propre d'une matrice du GUE. Il s'agit d'un résultat de Johnstone [99] (le cas du LOE ayant été démontré par Johansson [97]) que nous présentons ci-dessous.

Théorème 2.2.3. *(Johnstone). Soit $S_{m,n}$ une matrice du LUE. On suppose que $\frac{m}{n} \to \rho \in [1, \infty[$. Alors*

$$n^{2/3} \frac{\lambda_{\max} - b_{m,n}}{b_{m,n}^{2/3} \left(\frac{m}{n}\right)^{-1/6}} \to TW,$$

où TW désigne la loi de Tracy-Widom.

Remarque. Tout comme pour les matrices du GUE, des résultats d'universalités sont obtenus par Soshnikov [142] et Péché [132].

Résultats non-asymptotiques

Cette partie présente les résultats non-asymptotiques, s'inscrivant dans le cadre du phénomène de superconcentration, connus à ce genre dans la théorie des matrices aléatoires. Nous renvoyons le lecteur vers la thèse de Dallaporta [64] ou les articles de survols [114, 161] pour plus de détails. Remarquons, tout d'abord, que la plus grande valeur propre peut s'exprimer comme un supremum de variables gaussiennes par le biais du théorème de Courant-Fisher. Dans le cas du GUE, l'inégalité de Poincaré satisfaite par la mesure gaussienne entraine que

$$\mathrm{Var}(\lambda_{\max}) \leq \frac{C}{n}$$

qui ne correspond pas à la renormalisation du théorème (2.2.1) de Tracy-Widom lequel suggère que $\mathrm{Var}(\lambda_{\max}) \leq \frac{C}{n^{4/3}}$. Les premières tentatives d'étude non-asymptotique d'inégalités de déviation pour la plus grande des valeurs propres du GUE sont obtenues par Aubrun dans [10], ce résultat peut être également obtenu à partir des travaux de Johansson [98], comme annoncé dans [114]. L'inégalité obtenue s'exprime de la manière suivante, pour tout $t \geq 0$

$$\mathbb{P}(\lambda_{\max} \geq 1 + t) \leq C e^{-n \max(t^2, t^{3/2})/C}.$$

Même si nous pouvons citer un article de Tao et Vu [158] qui fournit une inégalité de déviation, reflétant le taux de convergence en $n^{-2/3}$, pour la plus grande valeur propre ; celle-ci ne reflète pas les asymptotiques de la fonction de répartition de la loi de Tracy-Widom et n'est pas valable pour les petites déviations. Il faudra attendre les travaux de Ledoux et Rider [116] qui obtiennent, par des démonstrations unifiées, pour les β-ensembles (qui contiennent les ensembles GUE et GOE), des inégalités de concentration optimales. En effet, ils retrouvent le résultat de Johansson ainsi qu'une inégalité de déviation à gauche de la forme

$$\mathbb{P}(\lambda_{\max} \leq 1 - t) \leq C e^{-n^2 t^3 / C}, \quad t \geq 0.$$

De plus, leur article fournit également les mêmes inégalités de déviation optimales à droite et à gauche pour la plus grande valeur propre d'une matrice du LUE ou du LOE. En intégrant ces inégalités de déviation, ces résultats permettent d'obtenir le bon ordre de grandeur de la variance de la plus grande valeur propre d'une matrice du LOE, LUE, GUE et GOE. Il faudra attendre les travaux de Dallaporta pour obtenir des résultats sur la variance du $\lambda_{\max}$ pour des ensembles de matrices plus généraux. Dans [64, 63], elle obtient notamment des majorations, de la variance de $\lambda_{\max}$, du bon ordre de grandeur pour des matrices de Wigner ou de covariances satisfaisant certaines hypothèses. A notre connaissance, il n'existe que très peu de travaux proposant des inégalités de déviations non asymptotiques pour des matrices de Wigner ou de Wishart. Citons toutefois l'article de Feldheim et Sodin [81] dans lequel ils obtiennent des inégalités de déviations à droite optimales pour des petites déviations.

Gaz de Coulomb et ensemble de Ginibre complexe

D'autres résultats provenant de la théorie des matrices aléatoires fournissent des exemples où la théorie de la concentration est sous-optimale. Citons par exemple l'article de Chafaï et Péché [47] ainsi que celui de Rider [135] qui s'interesse à un gaz de particules chargées $\{z_1, \dots, z_n\}$ dans le plan complexe, où chaque particule est confinée par un champ externe Q ainsi qu'une interaction répulsive (avec les autres particules) coulombienne. La densité de probabilité d'un tel objet est proportionnelle à

$$(z_1, \dots, z_n) \mapsto \prod_{j=1}^{n} e^{-nQ(z_j)} \prod_{1 \le j < k \le n} |z_j - z_k|^{\beta},$$

avec $\beta > 0$ un paramètre fixé et Q une fonction régulière donnée. Dans la suite, nous supposerons que Q est radialement symétrique, au sens où

$$Q(z) = V(|z|),$$

avec $V : \mathbb{R}_+ \to \mathbb{R}$ est un potentiel régulier tel que $t \mapsto tV'(t)$ est croissante ou V est convexe. Les auteurs des articles mentionnés ci-dessus s'intéressent aux statistiques d'ordres associées à ces particules $z_1, \dots, z_n$, que l'on désignera par $|z_{(1)}| \ge \dots \ge |z_{(n)}|$, celles-ci ont une signification particulière vis à vis des matrices aléatoires (voir [47]).
Lorsque $\beta = 2$ (et $V(r) = r^2$), Rider à obtenu un théorème de représentation de la loi de ces statistiques d'ordre :

$$\big(|z_{(1)}|, \dots, |z_{(n)}|\big) = \big(R_{(1)}, \dots, R_{(n)}\big), \quad \text{en loi}$$

où les $R_{(1)} \ge \dots R_{(n)}$ sont les statistiques d'ordre de variables indépendantes $R_1, \dots, R_n$ avec R_k de densité proportionnelle à

$$t \mapsto t^{2k-1} e^{-nV(t)} 1_{[0,\infty)}(t).$$

Remarque. Notons que ce choix particulier de β et du potientiel V pour le gaz de Coulomb correspond au modèle de Ginibre complexe en matrice aléatoire.

Rider obtient alors un théorème de convergence pour la particule ayant le plus grand module, ce résultat est ensuite prolongé par Chafaï et Péché pour des potentiels $V : \mathbb{R}_+ \to \mathbb{R}$ plus géneraux. Plus précisément,

Théorème 2.2.4. *Avec les notations précédentes (avec $\beta = 2$), soit $|z_{(1)}| = \max_{k=1,\dots,n} |z_k|$. Supposons que $V(t) = t^{\alpha}$ pour tout $t \ge 0$ et $\alpha \ge 1$. Posons $c_n = \log n - 2\log\log n - \log 2\pi$ et*

$$a_n = 2\Big(\frac{\alpha}{2}\Big)^{1/\alpha + 1/2} \sqrt{nc_n} \quad \text{et} \quad b_n = \Big(\frac{2}{\alpha}\Big)^{1/\alpha}\left(1 + \frac{1}{2}\sqrt{\frac{2}{\alpha}\frac{c_n}{n}}\right).$$

Alors, $(a_n(|z_{(1)}| - b_n))_{n\geq 1}$ *converge en loi, lorsque* $n \to \infty$*, vers la loi de Gumbel, i.e., pour tout* $t \in \mathbb{R}$

$$\mathbb{P}(a_n(|z_{(1)}| - b_n) \leq t) \to e^{-e^{-t}}.$$

Remarque. Ce théorème reste valable en supposant certaines conditions de stricte convexité sur $t \mapsto V(t)$ (voir [47]). Il serait intéressant d'obtenir des inégalités (sur la variance et de concentration) non-asymptotiques reflétant le résultat de convergence ci-dessus. Nous proposerons dans le chapitre sept, une majoration non-asymptotique de la variance de $|z_{(1)}|$ ainsi qu'une inégalité de déviation à droite en adéquation avec le théoreme 2.2.4. Des travaux sont en cours pour obtenir les déviations à gauche ainsi que l'extension de ces résutats à certains modèles déterminantaux présentés dans l'article [96].

2.2.2 Théorème de Dvoretzky aléatoire

Le théorème classique de Dvoretzky [75], est un résultat important de la théorie de la concentration de la mesure (cf. [126]). Celui-ci concerne le problème géométrique des sections sphériques d'un domaine convexe.

Soit $\mathcal{K}$ un domaine symétrique convexe dans $\mathbb{R}^n$, c'est à dire, un sous ensemble convexe compact d'intérieur non vide de $\mathbb{R}^n$ symétrique par rapport à l'origine. On dit que $\mathcal{K}$ contient « presque » des sections euclidiennes de dimension k, si pour tout $\epsilon > 0$, il existe un sous espace H de dimension k et un ellipsoïde $\mathcal{E}$ dans H tel que

$$(1-\epsilon)\mathcal{E} \subset \mathcal{K} \cap H \subset (1+\epsilon)\mathcal{E}.$$

Cette description géométrique admet une formulation fonctionnelle équivalente. Un espace de Banach E (de norme $\|\cdot\|$) est dit contenir un sous espace $(1+\epsilon)$-isomorphe à l'espace euclidien $\mathbb{R}^k$ s'il existe des vecteurs linéairement indépendants $v_1, \dots, v_k$ dans E tels que pour tout $t = (t_1, \dots, t_k) \in \mathbb{R}^k$,

$$(1-\epsilon)|t| \leq \Big\| \sum_{i=1}^{k} t_i v_i \Big\| \leq (1+\epsilon)|t|,$$

où $|\cdot|$ désigne la norme euclidienne sur $\mathbb{R}^k$. Le théorème suivant est le célèbre résultat de Dvoretzky dans la théorie locale des espaces normés.

Théorème 2.2.5. *Pour tout* $\epsilon > 0$ *il existe* $\eta(\epsilon) > 0$ *tel que tout espace de Banach* E *de dimension* n *contient un sous-espace* $(1+\epsilon)$*-isomorphe à* $\mathbb{R}^k$ *avec* $\lfloor k = \eta(\epsilon) \log n \rfloor$*.*

Les auteurs Schechtman [139], Paouris *et al.* [131] s'intéressent à une version aléatoire de ce théorème dans les espaces l_p, $p \geq 1$. Comme indiqué par Schechtman dans son article, le théorème de Dvoretzky classique entraine qu'une vaste majorité de sous-espace de dimension $k \leq c(\epsilon) \log n$ sont $(1+\epsilon)$ isomorphe à l_2^k. Le théorème de Dvoretzky aléatoire consiste à étudier le même phénomène avec l_p^k, $p \geq 1$. Par exemple, lorsque $p = \infty$ Schechtman obtient le théorème suivant

Théorème 2.2.6 (Schechtman)**.** *Pour tout* $\epsilon > 0$ *et tout* $k \geq c\epsilon \log n$*, avec une probabilité strictement supérieure à* $1 - e^{ck}$ *la norme* l_∞^n *et un multiple de la norme* l_2^n *sont* $1+\epsilon$ *équivalentes sur un sous-espace de dimension* k*.*

Remarque. Nous avons retranscrit le théorème de Schechtman tel qu'il apparait dans l'article [139], nous renvoyons le lecteur vers celui-ci pour plus de détails. Signalons tout de même que la mesure de probabilité sous-jacente, dans le résultat précédent, est la loi d'un vecteur gaussien standard de $\mathbb{R}^n$.

Dans [131], les auteurs obtiennent des résultats très fins sur la variance des normes l_p, $p \geq 1$ d'un vecteur gaussien standard dans $\mathbb{R}^n$ ainsi que des inégalités de deviation. Ces raffinements font partis des ingrédients utilisés par les auteurs de [131] pour obtenir un théorème de Dvoretzky aléatoire.

Soit $p \geq 1$, nous noterons par $\|x\|_p^p = \sum_{i=1}^n |x_i|^p$ la norme l_p d'un vecteur $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Dans l'article Paouris *et al.* [131], il est remarqué que la variance de $\|X\|_p$ n'est pas estimée précisément par les outils classiques de la concentration de la mesure. Plus précisément, l'inégalité de Poincaré ou l'inégalité isopérimétrique vérifiée par la mesure gaussienne standard sur $\mathbb{R}^n$ fournissent la majoration suivante, pour $p \geq 1$,

$$\mathrm{Var}(\|X\|_p) \leq \max(n^{2/p-1}, 1).$$

Selon [131], cette borne n'est optimale que pour $1 \leq p \leq 2$. Les auteurs de [131] améliorent cette borne en utilisant des estimations précises et une application astucieuse de l'inégalité de Sobolev logarithmique satisfaite par γ_n. A notre connaissance, c'est la première fois que de la superconcentration est obtenue directement par l'inégalité de Sobolev logarithmique sans passer par sa formulation hypercontractive équivalente. L'inégalité de Talagrand 1.3.6 fait, tout de même, parti des outils fonctionnels utilisés lors des démonstrations lorsque p est grand vis à vis de $\log n$. Ils obtiennent également des inégalités de superconcentration reflétant ces bornes. Nous proposerons dans le chapitre sept de retrouver ceci par des méthodes de transport. Concernant la variance, voici leurs résultats

Théorème 2.2.7 (Paouris-Valettas-Zinn)**.** *Soit X un vecteur gaussien standard dans $\mathbb{R}^n$, alors les inégalités suivantes sont satisfaites, pour tout $n \geq 2$,*

$$\mathrm{Var}(\|X\|_p) \leq \begin{cases} C\frac{2^p}{p} n^{2/p-1}, & 2 < p \leq c\log n, \\ C/\log n, & p > c\log n, \end{cases}$$

avec $C, c > 0$ des constantes numériques indépendentes de n et de p.

Concernant les inégalités de déviations, ils obtiennent

Théorème 2.2.8 (Paouris-Valettas-Zinn)**.** *Soient $2 < p < \infty$ et X un vecteur gaussien standard dans $\mathbb{R}^n$. Alors, pour tout $n \geq 1$,*

$$ce^{-Cp\alpha_1(n,p,t)} \leq \mathbb{P}\Big(\big|\|X\|_p - \mathbb{E}[\|X\|_p]\big| \geq t\mathbb{E}[\|X\|_p]\Big) \leq Ce^{-c\alpha_2(n,p,t))},$$

pour tout $0 < t < 1$, avec $\alpha_i(n, p, \cdot)$ étant définie par :

$$\alpha_1(n,p,t) = \min\left\{C^{2p}t^2 n, (tn)^{2/p}\right\}$$

et

$$\alpha_2(n,p,t) = \min\left\{C^{-2p}t^2 n, (tn)^{2/p}\right\}$$

avec $C, c > 0$ des constantes numériques.

Remarque. La raison de restreindre t à l'intervalle $]0,1[$ est que les grandes déviations sont controlées par l'inégalité isopérimétrique gaussienne. La véritable difficulté est d'affiner les bornes pour des petites déviations, c'est le contenu du théorème ci-dessus. Notons que le cas $p=\infty$ et l'étude de théorème de Dvoretzky aléatoire, a été faite, via des calculs élémentaires mais précis, par Schechtman dans [139], plus précisément, il prouva les inégalités suivantes, valables pour tout $0<t<1$,

$$C_1 e^{-c_1 t M^2/2} \leq \mathbb{P}\big(\|X\|_\infty > (1+t)M\big) \geq C_2 e^{-tM^2},$$

et

$$e^{-c_3 e^{tM^2}} \leq \mathbb{P}\big(\|X\|_\infty < (1-t)M\big) \leq C_4 e^{-c_4 e^{3tM^2/4}}$$

où $M=\mathrm{Med}$ désigne la médiane de $\|X\|_\infty$ et les constantes $C_1, C_2, c_1 > 0$ sont explicitement données tandis que $c_3, c_4, C_4 > 0$ sont des constantes numériques. Nous verrons dans la suite de la thèse que des arguments hypercontractifs ou isopérimétriques permettent de retrouver ce genre de résultats. Citons également les articles [159, 121] qui abordent le phénomène de superconcentration et le théorème de Dvoretzky aléatoire.

Dans un autre article [130], Paouris et Valettas obtiennent une inégalité de déviation à gauche qui renforce, pour les petites déviations et une large classe de fonctions, la concentration fournie par l'inégalité isopérimétrique gaussienne. Leurs arguments reposent sur l'inégalité d'Ehrhard [76] et les outils utilisés dans [131].

Théorème 2.2.9. *[Paouris-Valettas] Soit X un vecteur gaussien dans $\mathbb{R}^n$ et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe, alors pour tout $t>1$,*

$$\mathbb{P}\left(f(X) - \mathbb{E}[f(X)] \leq -t\sqrt{\mathrm{Var} f(X)}\right) \leq e^{-ct^2},$$

avec $c>0$ une constante numérique.

Remarque. Ce résultat est inspiré d'un article de Kwapien [107] concernant des fonctions convexes d'un aléa gaussien. Ce genre de renforcement d'inégalité fonctionnelle lorsqu'une hypothèse de convexité est satisfaite, est déja apparu plusieurs fois dans la littérature comme, par exemple, dans les articles de Bobkov et Houdré pour l'inégalité de Poincaré [30, 28]. Concernant la superconcentration, si l'on applique ce résultat à fonction $f(x) = \max_{i=1,\dots,n} x_i$ (qui est convexe en tant que supremum de fonctions linéaires) avec le résultat de Eldan, Ding et Zhai [69], qui assure que pour un vecteur gaussien standard $\mathrm{Var}(M_n) \geq C/\log n$ (avec une constante numérique $C>0$), on obtient

$$\mathbb{P}\big(\sqrt{\log n}(M_n - \mathbb{E}[M_n]) \leq -t\big) \leq e^{-cr^2}, \quad t>1 \tag{2.5}$$

Ce résultat améliore bien la concentration habituelle, cependant il n'exprime pas l'asymptotique de la loi de Gumbel. Celle-ci admet une décroissance très rapide pour sa queue de distribution à gauche en $e^{-e^{-t}}$, $t\to-\infty$. Ce résultat a été obtenu dans les travaux de Schechtman [139]. Nous emploierons cette inégalité (2.5) dans le cadre des suites gaussiennes stationnaires.

2.2.3 Statistiques d'ordre

La théorie des extrêmes, notamment la convergence en loi du maximum des coordonnées d'un vecteur aléatoire, est un cadre naturellement lié à la théorie de la superconcentration. Les résultats classiques de convergence des extrêmes présentés dans [108, 67, 78] fournissent toute une zoologie d'exemples pour lesquels il serait intéressant de fournir des inégalités de superconcentration, non asymptotiques, reflétant ces résultats de convergence. Au niveau de la variance, Boucheron et Thomas proposent dans [37] une série d'inégalités pour les statistiques d'ordres d'un échantillon *i.i.d.*. Ils présentent également une inégalité de type Bernstein pour la norme infinie d'un vecteur gaussien standard.

Le cadre qu'ils considèrent est le suivant : soit $X_1, \dots X_n$ un échantillon i.i.d d'une loi μ et les statistiques d'ordre associées

$$X_{(1)} > \dots > X_{(n)}.$$

Les auteurs de [37] obtiennent, sous des hypothèses techniques sur la mesure μ qui seront levées dans le chapitre sept, le résultat suivant

$$\mathrm{Var}(X_{(k)}) \le \frac{2}{k}\mathbb{E}\left[\frac{1}{\kappa_\mu(X_{(k+1)})^2}\right], \quad k = 1, \dots, n,$$

où κ_μ désigne la fonction de risque associée à la loi μ (c.f. [37] ou chapitre sept). Leurs arguments reposent essentiellement sur la représentation de Renyi des statistiques d'ordre (qui fournit une expression de $X_{(k)}$, $k = 1, \dots, n$ en fonction d'une somme renormalisée de variables aléatoires exponentielles standards i.i.d. c.f. [67, 37]). Ils combinent cette représentation avec l'inégalité d'Efron-Stein (cf. [38]) et l'inégalité d'association négative de Harris (pour cela, ils sont obligé de supposer une certaine décroissance de la fonction dite de « risque » κ_μ afin d'obtenir une borne sur la variance de $X_{(k)}$) Ils obtiennent également, dans le cas gaussien, une inégalité de déviation à droite. Plus précisément, pour $|X_i|$ avec $\mathcal{L}(X_i) = \mathcal{N}(0, 1)$ et $U(s) = \Phi^{-1}(1 - 1/(2s))$, où Φ designe la fonction de répartition d'une variable aléatoire gaussienne standard.

$$\mathbb{P}\Big(X_{(1)} - \mathbb{E}_{\mu^n}[X_{(1)}] \le t/\big(3U(n)\big) + \sqrt{t}/U(n) + \delta_n\Big) \le e^{-t}, \quad t \ge 0$$

avec $\delta_n > 0$ et $[U(n)]^3\delta_n \to \frac{\pi^2}{12}$ lorsque $n \to \infty$.

Nous pourrons retrouver la majeure partie de leurs travaux via l'approche par le transport optimal. Nous clarifierons également la nécessité de certains arguments. En particulier, nous démontrerons que les bornes sur la variance peuvent être obtenues sans l'hypothèse de décroissance de la fonction de risque. Tandis que cette hypothèse semble nécessaire pour obtenir des inégalités de déviations.

2.2.4 Verres de spin

La littérature concernant les verres de spins est très vaste, nous renvoyons le lecteur vers les ouvrages de référence, dont nous nous inspirons, suivant [154, 155, 129, 39] pour plus de détails. Un modèle de verres de spin consiste en un système de spins (des « ± 1 ») soumis à des interactions aléatoires. L'espace d'état est, en général, le cube discret $C_n = \{-1, 1\}^n$, $n \ge 1$ (on considérera parfois la sphère $\mathbb{S}^n$). Sur cet espace, on définit un processus gaussien centré $H_n : C_n \to \mathbb{R}$ avec la fonction de covariance suivante

$$\mathbb{E}[H_n(\sigma)H_n(\sigma')] = ng(\sigma, \sigma'), \quad \sigma, \sigma' \in C_n,$$

et où $g : C_n \times C_n \to R$ est une forme quadratique définie positive. H_n est appelé hamiltonien du système et $H_n(\sigma)$, $\sigma \in C_n$ représente, en physique, l'énergie du système dans la configuration

σ. Dans la théorie des verres de spins de champ moyen, g est supposée ne dépendre uniquement que de la distance entre σ et σ'. Parmi les distances les plus utilisées, citons

1. La distance de Hamming,

$$d_{\mathrm{Ham}}(\sigma,\sigma') = \sum_{i=1}^{n} 1_{\{\sigma_i\neq\sigma'_i\}} = \frac{1}{2}\left(n - \sum_{i=1}^{n} \sigma_i\sigma'_i\right). \tag{2.6}$$

   Dans ce cas, on choisit g comme étant

$$g(\sigma,\sigma') = p\big(1 - 2d_{\mathrm{Ham}}(\sigma,\sigma')/n\big),$$

   où p désigne une fonction polynomiale ayant tout ses coefficients positifs. Le cas $p(x) = x^2$ correspond au modèle SK de Sherrington et Kirkpatrick.

2. La distance lexicographique,

$$d_{\mathrm{lex}}(\sigma,\sigma') = n + 1 - \min(i\,:\,\sigma_i \neq \sigma'_i). \tag{2.7}$$

   Dans ce cas,

$$g(\sigma,\sigma') = A\big(1 - n^{-1}d_{\mathrm{lex}}(\sigma,\sigma')\big),$$

   où $A\,:\,[0,1]\to[0,1]$ peut-être choisie comme une fonction croissante telle que $A(0) = 0$ et $A(1) = 1$. Les modèles obtenus dans ce cas de figure, sont ceux introduit par Gardner et Derrida et connu sous le nom de GREM (pour ≪ Generalized Random Energy Models ≫).

Dans les modèles de verres de spins, on cherche à comprendre le comportement de $\min_{\sigma\in C_N} H_N(\sigma)$ (la présence du minimum provient de la physique statistique dans laquelle on souhaite minimiser l'énergie plutôt que de la maximiser). Ce problème est, en général, très difficile à résoudre. Il suffit de considérer l'hamiltonien du modèle SK pour s'en convaincre. Celui-ci, aux renormalisations près, s'exprime de la manière suivante

$$H_n(\sigma) = \sum_{i,j=1}^{n} X_{ij}\sigma_i\sigma_j, \quad \sigma = (\sigma_1,\ldots,\sigma_n) \in C_n,$$

et où $(X_{ij})_{1\leq i\leq j\leq n}$ est une famille de gaussiennes standards indépendantes de même loi. Pour contourner ces difficultés, on considère une version plus régulière de l'hamiltonien H_n avec une température $\beta > 0$ (en fait, il s'agit de l'inverse de la température). Il s'agit de l'énergie libre, définie par

$$F_{n,\beta} = \frac{1}{\beta}\log\bigg(\sum_{\sigma\in C_n} e^{-\beta H_n(\sigma)}\bigg), \beta > 0.$$

Notons, que lorsque $\beta\to\infty$, $F_{n,\beta}$ converge vers $\min_{\sigma\in C_n} H_n(\sigma)$. L'énergie libre donne lieu à une mesure de Gibbs très étudiée dans cette théorie. Comme nous le verrons par la suite, les résultats obtenus dépendent souvent de la température et les modèles présentent des comportements différents suivant que cette température est supérieure à une valeur critique $\beta_c > 0$ ou non.

Nous allons présenter les différents résultats de superconcentration connus pour les modèles de verres de spins. A notre connaissance, il n'existe essentiellement que deux approches mathématiques pour étudier ces modèles : l'une utilise des méthodes d'interpolations tandis que l'autre utilise des méthodes de « second moment modifié ». Nous choisissons donc de partager notre présentation en deux suivant que l'on utilise l'une ou l'autre des deux méthodes évoquées ci-dessus.

Méthode d'interpolation

Dans cette partie nous allons présenter plus en détail le modèle de physique statistique introduit par Sherrington et Kirkpatrick. Pour donner une heuristique concernant ce modèle, et les raisons qui ont motivées son étude, nous allons largement nous inspirer de l'introduction du livre de Talagrand [152].

Considérons une grande famille d'individus numérotés de 1 à n. Supposons que chaque individu 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-être 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 ou 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 maintenenant 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 apparait 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 la plus grande possible nous entraine à rendre les $X_{ij}\sigma_i\sigma_j$ positifs, et donc à prendre σ_i et σ_j de même signe lorsque $X_{ij} > 0$, et de signes opposés sinon.

Malgré la simplicité apparente 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. On parle alors d'état fondamental.

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 du à la répartition aléatoire 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 $\sigma = (\sigma_1, \ldots, \sigma_n)$ est un élément de $C_n = \{-1;1\}^n$. Soient $(X_{ij})_{1\leq i<j\leq n}$ une collection de variables aléatoires gaussiennes standards indépendantes. Cette famille de variables aléatoires réelles 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 S_n en associant à chaque configuration le poids suivant : $\exp\big(-\beta H_n(\sigma)\big)$, avec $\beta > 0$, l'inverse de la température. 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.

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

$$\frac{\exp(-\beta H_n(\sigma))}{Z_{n,\beta}}.$$

Où $Z_{n,\beta} = \sum_{\sigma\in C_n} e^{-\beta H_n(\sigma)}$ 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).

L'objet principal d'intérêt sera la fonction d'energie libre, qui se réexprime comme suit

$$F_{n,\beta} = \frac{1}{\beta}\log\big(Z_{n,\beta}\big), \quad \beta > 0$$

A partir de la formule de représentation de la variance, le long du semi-groupe d'Ornstein-Uhlenbeck (généralisé), Chatterjee a obtenu dans [48] une série d'inégalités différentielles. Celles-ci lui ont permis d'obtenir une majoration de la variance de l'énergie libre pour le modèle SK (ainsi que pour d'autres modèles plus compliqués, dit de p-spin). Nous reviendrons sur cette démonstration dans le chapitre six du manuscrit, car une relecture de celle-ci a permis de clarifier certains arguments proposés par Chatterjee et d'obtenir une nouvelle inégalité sur la variance pour la mesure gaussienne γ_n. Son théorème dit essentiellement la chose suivante

Théorème 2.2.10. *(Chatterjee) Dans le modèle SK, la majoration suivante est vérifiée pour l'energie libre, pour tout* $n \geq 2$*,*

$$\mathrm{Var}(F_{n,\beta}) \leq \frac{C_\beta n}{\log n}, \quad \beta \geq 0.$$

Remarque. Sa méthodologie fournit aussi des bornes pertinentes pour des modèles plus simples de verres de spin, comme le REM (« Random Energy Model »). Nous verrons cependant, dans le chapitre cinq, qu'il est possible d'affiner ces majorations en fonction de la température. Dans son livre, Chatterjee propose également une approche spectrale de la superconcentration et obtient la borne suivante pour le modèle SK

$$\mathrm{Var}\big(F_{n,\beta}\big) \leq \frac{C_\beta n \log\log n}{\log n},$$

où C_β est constante ne dépendant que du paramètre β. Toutefois, aucune de ces deux bornes n'est optimale, il est attendu que $\mathrm{Var}(F_{n,\beta}) = O(1)$ pour toute température $\beta > 0$. Néanmoins, les travaux de Chatterjee sont les premiers à obtenir une borne sous linéaire de la variance de l'énergie libre améliorant la majoration fournie par l'inégalité de Poincaré.

Cette approche de la superconcentration dans les modèles de verres de spins a donnée naissance à toute une série d'article. Bien qu'il s'agisse essentiellement de la même chose, les auteurs parlent de phénomène de chaos en le désordre gaussien (« chaos disorder ») au lieu de superconcentration. L'équivalence entre ces deux notions est détaillée dans l'ouvrage de Chatterjee [48].

De nombreux travaux ont été fait pour répondre à la question suivante : dans un modèle de verres de spin donné, quelle est l'influence de la présence d'un champ externe sur le phénomène de superconcentration ? Plus précisément, cela revient à ajouter un vecteur $h \in \mathbb{R}^n$ à l'hamiltonien. L'énergie d'une configuration est alors mesurée par

$$-H_n(\sigma) + \sum_{i=1}^{n} h_i \sigma_i, \quad \sigma = (\sigma_1, \dots, \sigma_n) \in C_n$$

et où H_n est l'hamiltonien initial du système. Chen montre dans [50] que le phénomène de superconcentration, obtenu par Chatterjee dans le modèle SK, est toujours valable en présence d'un champ externe. Dans cet article, ainsi que les suivants, il s'appuie sur les travaux de Chatterjee et sur la formule de représentation de la variance pour étudier l'énergie libre $F_{n,\beta}$. Cependant, il utilise également des arguments propres à la théorie des verres de spins, comme les répliques symétriques de Guerra ou des résultats pointus comme la formule de Parisi pour le modèle considéré. Chen démontre alors la présence de chaos pour les modèles de p-spin pairs mixtes dans [51]. Chen, Dey et Panchenko étudient le modèle des p-spin mixtes avec présence d'un champ externe dans [52] et démontre que la présence du champ externe empèche le phénomène de superconcentration ; en d'autres termes, ils montrent que l'inégalité de Poincaré fournit le bon ordre de grandeur de la variance, pour ce modèle, en présence d'un champ externe. Citons également, les articles [55, 56, 57, 11] de Chen, avec différent co-auteurs, dans la même thématique. Chen, Hsieh, Hwang et Sheu montrent la présence de chaos, en le désordre, pour le modèle de champ moyen sphérique dans [54]. Dans [53], Chen, Handschy et Lerman montrent que l'état fondamental (le minimum de l'hamiltonien, que l'on désigne en anglais par « ground state ») du modèle des p-spins mixtes paires est superconcentré si et seulement si le champ externe est nul ; en présence d'un champ externe l'inégalité de Poincaré fournit le bon ordre de grandeur pour la variance de l'état fondamental et un théorème de convergence en loi vers une gaussienne est obtenu. Enfin, dans [15] Baik et Lee étudient l'énergie libre pour le modèle SK sphérique, pour lequel les spins $\sigma \in \mathbb{S}^n$ au lieu du cube discret C_n, et démontrent un phénomène de transition de phase. Ils obtiennent le résultat de convergence en loi suivant

Théorème 2.2.11. *(Baik-Lee) Si $0 < \beta < \frac{1}{2}$, alors,*

$$n\big(F_{n,\beta} - F_\beta\big) \to \mathcal{N}(f, \alpha), \quad n \to \infty$$

où $\mathcal{N}(f, \alpha)$ désigne une gaussienne de moyenne f et de variance α. Tandis que si $\beta > \frac{1}{2}$, la loi limite est différente

$$\frac{1}{\beta - \frac{1}{2}} n^{2/3}\big(F_{n,\beta} - F_\beta\big) \to TW, \quad n \to \infty,$$

où TW correspond à la loi de Tracy-Widom provenant de la théorie des matrices aléatoires.

L'expression détaillée des paramètres de renormalisation est fournie dans [15], les auteurs précisent également le lien entre l'état fondamental et la plus grande valeur propre d'une matrice du GOE.

Méthode de second moment

Parallèlement aux travaux qui font suite aux résultats de Chatterjee et à son approche, d'autres auteurs ont étudiés des modèles de verres dans lesquels une structure d'arbre est présente. Une grande partie de ces travaux sont antérieurs à ceux de Chatterjee et les questions abordées sont légèrement différentes. L'étude des extrêmes et de la convergence en loi de l'énergie libre font parties de leurs questionnements, cependant les auteurs cherchent à obtenir des résultats plus précis comme la convergence de processus ponctuels de Poisson associés aux extrêmes. Les résultats obtenus sont, pour la plupart, démontrés via la méthode de second moment (et ses variantes). Celle-ci repose sur l'inégalité de Paley-Zygmund et une étude fine, rendue possible grâce à la structure d'arbres sous-jacente, des trajectoires des particules. Nous renvoyons le lecteur, vers l'ouvrage de Bovier [39] ou l'excellent cours de Kistler [104] pour plus de détails. Comme nous le verrons dans la section suivante, cette méthodologie de second moment reposant sur la structure hiérarchique de l'arbre sera également un outil essentiel dans l'étude de champ libre gaussien discret sur $\mathbb{Z}^2$. Bien qu'efficace, ces techniques sont assez lourdes en calcul. Il sera intéresssant de retrouver les résultats obtenu dans ces articles par des méthodes plus simple d'emploi, nous verrons notamment dans le chapitre six que des arguments d'interpolation permettent de retrouver certains d'entre eux.

Dans [42], Bovier, Kurkova et Löwe obtiennent des résultats de fluctuations pour les modèles de verres de spin suivants : le REM (« Random Energy Model ») et le p-spin SK. Leurs résultats de fluctuations concernent des convergences en loi de la fonction de partition ou de convergence de processus ponctuel de Poisson. Nous allons présenter briévement le REM et les résultats obtenus le concernant. Nous verrons dans le chapitre six que ce modèle fait parties des exemples pour lesquels les méthodes de semi-groupes sont efficaces pour retrouver l'ordre de grandeur de la variance de la fonction de partition.

Le REM est sans doute le plus simple des modèles de verres de spins existant et est construit comme suit : on considère $(X_\sigma)_{\sigma \in S_n}$ une famille de gaussiennes réelles standards indépendantes et de même loi. On définit l'hamiltonien du système H_n par

$$H_n(\sigma) = \sqrt{n} X_\sigma, \quad \sigma \in C_n.$$

Les auteurs de [42], s'intéressent à la fonction de partition $Z_{n,\beta}$ définie par $Z_{n,\beta} = 2^{-n} \sum_{\sigma \in C_n} e^{\beta H_n(\sigma)}$, où $\beta > 0$ joue le rôle de l'inverse de la température. Parmi les résultats obtenus dans cet article mentionons ceux que nous pourrons redémontrer par d'autres méthodes.

Théorème 2.2.12. *(Bovier, Kurkova, Löwe) Avec les notations précédentes, les convergences en loi suivantes sont vérifiées*

1. *Si $\beta < \sqrt{\log 2/2}$, alors*

$$\exp \frac{n}{2}(\log 2 - \beta^2) \log \frac{Z_{n,\beta}}{\mathbb{E}[Z_{n,\beta}]} \to \mathcal{N}(0,1), \quad n \to \infty$$

2. *Si $\beta > \sqrt{2 \log 2}$, alors*

$$\log Z_{n,\beta} - \mathbb{E}[\log Z_{n,\beta}] \to \log \int_{\mathbb{R}} e^{\alpha z} \mathcal{P}(dz) - \mathbb{E} \log \int_{\mathbb{R}} e^{\alpha z} \mathcal{P}(dz), \quad n \to \infty$$

où $\mathcal{P}$ désigne le processus ponctuel de Poisson sur $\mathbb{R}$ de mesure d'intensité $e^{-x}dx$.

Remarque. L'article contient beaucoup d'autres résultats, complétant l'énoncé du théorème précédent. Notons également, qu'après une renormalisation et un changement d'échelle, on peut retrouver la fonction d'énergie libre $F_{n,\beta}$ introduite plus tôt à partir de $Z_{n,\beta}$.

Dans la continuité de cet article, Bovier et Kurkova publient deux articles sur le GREM [40, 41], contenant le même type de résultats de convergence et, implicitement, montre que la fonction de partition est superconcentrée dans certains cas. Toujours avec des méthodes de second moment (modifié), Kistler et Schmidt obtiennent un théorème de convergence pour des processus ponctuels de Poisson dans [105]. Ils insistent sur la présence d'une structure hiérarchique et de la présence d'un arbre dans leur modèle pour mettre en place leurs démonstrations. Ils parviennent notamment à traiter la convergence d'un processus ponctuel de Poisson N_N^α. Celui-ci permettant une interpolation entre le modèle du REM et de la marche aléatoire branchante (dont nous parlerons plus en détails dans la section suivante). Il serait intéressant de voir si une adaptation du lemme 3.1 de leur article est possible. Ce, afin d'obtenir une borne optimale sur la variance de la marche aléatoire branchante en combinant des méthodes d'interpolation et de second moment. Signalons également l'article de Subag et Zeitouni [143] qui obtiennent un résultat de convergence en loi de l'état fondamental (renormalisé) vers une loi de Gumbel (négative).

2.2.5 Champ libre gaussien discret

Pour présenter ce nouveau modèle, nous suivrons l'exposition faite dans le cours de Zeitouni [165]. Pour des raisons pédagogiques, il est plus intéressant de présenter le modèle de marche aléatoire branchante (« branching random walk ») en premier. Il fait parti des modèles gaussiens dit « log-corrélés » les plus simples. Ensuite, nous parlerons du champ libre gaussien discret sur $\mathbb{Z}^2$ et de l'évolution des résultats obtenus le concernant. Nous expliquerons aussi, brièvement, les méthodes mises en place pour l'étude de cet objet.

Fixons tout d'abord les notations. Soit $\mathcal{T}$ un arbre ayant pour racine un sommet o, nous noterons par V l'ensemble des sommets et par E l'ensemble des arêtes de cet arbre. Nous désignerons par $|v|$ la distance d'un sommet à la racine. Autrement dit, il s'agit de la longueur du chemin géodésique reliant v à o et l'on désignera par $o \leftrightarrow v$ l'ensemble des sommets composant cette géodésique (par abus de notation, nous noterons également par $o \leftrightarrow v$ l'ensemble des arêtes sur la géodésique reliant v à o). Similairement, pour $v, w \in V$, nous noterons par $|v \leftrightarrow w|$ la longueur de l'unique géodésique reliant v à w. Pour $n \geq 1$, la n-ième génération de l'arbre correspond l'ensemble suivant $\{v \in V\,;\, |v| = n\}$. Tandis que pour $m < n$ et $v \in D_m$ on désignera l'ensemble des descendants de v (dans D_n) par $D_n^v = \{w \in V\,;\, |v \leftrightarrow w| = n - m\}$. Enfin, pour tout sommet $v \in V$ on désignera par d_v son degré.

Soit $(X_e)_{e \in E}$ une collection de variables aléatoires μ attachées à chacunes des arêtes e de l'arbre $\mathcal{T}$. Pour $v \in V$, on pose $S_v = \sum_{e \in o \leftrightarrow v} X_e$. La marche aléatoire branchante n'est alors rien d'autre que $(S_v)_{v \in V}$ et l'objet principal d'intérêt sera la quantité suivante :

$$M_n = \max_{v \in D_n} S_v, \quad n \geq 1.$$

Les hypothèses suivantes sont faites lors de l'étude de ce modèle :

1. les variables aléatoires $(X_e)_{e \in E}$ sont supposées i.i.d. de loi commune μ.

2. On suppose que μ admet un moment exponentiel, c'est à dire $\mathbb{E}[e^{\lambda X_e}] < \infty$ pour un $\lambda > 0$.

3. L'arbre est k-aire avec $k \geq 2$ ($k = 2$ correspond à un arbre binaire), $d_0 = k$ et $d_v = k + 1$ pour $v \neq o$.

Comme présenté dans le cours de Zeitouni, à l'aide de méthodes dites de « second moment » (liées à l'inégalité de Paley et Sygmund) et de résultats du type « Ballot's theorem ». Une étape importante dans ce genre de méthode est de déterminer une barrière au delà de laquelle les configurations extrémales S_v, $v \in D_n$ ne peuvent aller. Il sera aussi important de suivre l'évolution des incréments S_v^k, $k = 1, \ldots n$, composant la « trajectoire » de S_v (autrement dit $S_v = \sum_{k=1}^n S_v^k$), qui ne peuvent pas dépasser un certain seuil. Suivant ces techniques, on peut montrer que

Théorème 2.2.13. *Sous les hypothèses précédentes on a l'estimation précise suivante, $n \geq 1$,*

$$\mathbb{E}[M_n] = C_1 n - C_2 \log n + O(1),$$

avec $C_1, C_2 > 0$ des constantes numériques dépendant uniquement de la loi μ.

Remarque. Le résultat précédent, via l'argument de Dekking-Host [165], permet de montrer que la suite $(M_n - \mathbb{E}[M_n])_{n \geq 1}$ est tendue. Bien que la preuve explicite semble être absente, à notre connaissance, de la littérature, le fait que $\mathrm{Var}(M_n) = O(1)$ fait dorénavant parti du folklore. Il est intéressant de mentionner que les arguments utilisés pour démontrer le théorème ci-dessus, bien qu'élémentaires, reposent sur des calculs assez techniques et précis. Tandis que la méthode hypercontractive de Talagrand permet d'atteindre de manière élementaire la borne suivante (sous-optimale) $\mathrm{Var}(M_n) \leq C \log n$ améliorant la borne $\mathrm{Var}(M_n) \leq n$ fournit par l'inégalité de Poincaré (cf. chapitre un). Il serait instructif de trouver une preuve moins technique à partir de la formule de représentation de la variance le long du semi-groupe d'Ornstein-Uhlenbeck. Il semblerait qu'une preuve reposant sur cette représentation et des calculs simples liés aux méthodes de second moment permettraient de s'approcher d'un tel résultat. Il suffirait d'obtenir une extension du lemme 3.1 de l'article [105] pour contrôler, pour tout $t \geq 0, \sigma \neq \tau \in D_n$, $\mathbb{E}[1_{A_\sigma} P_t 1_{A_\tau}]$ (avec $A_\sigma = \{S_\sigma = M_n\}$). Ceci permettrait d'attester que les contributions significatives dans le calcul de la variance ne proviennent que des termes diagonaux $\sigma = \tau$.

La marche aléatoire branchante est une version plus simple du champ libre gaussien discret. Nous définirons ce nouveau modèle ci-dessous, puis nous tenterons de résumer l'évolution des progrès obtenus durant son étude. Enfin, nous mentionnerons briévement les articles de recherche, s'inspirant des techniques utilisées pour le champ libre gaussien discret, portant sur des modèles gaussiens logarithmiquement corrélés plus généraux.

Considérons $G = (V, E)$ un graphe (fini) non-orienté et connexe. Choisissons un sommet particulier $v \in V$ et désignons le comme ≪ la racine ≫ du graphe, le champ libre gaussien discret correspond alors à la collection de variables aléatoires gaussiennes $(X_v)_{v\in V}$ avec $X_o = 0$ et la densité de probabilité, par rapport à la mesure de Lebesgue, des variables aléatoires restantes est donnée par

$$f\big((x_v)_{v\in V,\, v\neq o}\big) = \frac{1}{Z_G} e^{-\frac{1}{2}\sum_{(v,w)\in V}(x_v - x_w)^2}, \tag{2.8}$$

où Z_G est une constante de renormalisation.

Remarque. Il est important de faire quelques commentaires après cette définition. Tout d'abord, si G est un arbre, le champ libre gaussien discret consiste à assigner à chaque arête $e \in E$ une variable aléatoire gaussienne standard Y_e et de poser $X_v = \sum_{e\in o\leftrightarrow v} Y_e$, $v \in V$. En particulier, si G correspond à l'arbre binaire, alors les valeurs de champ libre gaussien discret (noté $(X_v^T)_{v\in D_n}$) suit le même comportement que M_n de la marche aléatoire branchante. On obtient donc, le résultat suivant

$$\mathbb{E}\big[\max_{v\in D_n} X_v^T\big] = \sqrt{2\log 2}n - \frac{3}{2\sqrt{2\log 2}}\log n + O(1)$$

Ensuite, le champ libre gaussien possède des liens avec certaines marches aléatoires. En effet, considérons la matrice de transition suivante

$$L_{v,w} = \begin{cases} 1, (v,w) \in E, \\ -d_v, v = w \neq o, \\ 0, \text{sinon}, \end{cases}$$

où d_v désigne le degré du sommet v. A partir de (2.8), on remarque que la structure de corrélation du champ libre gaussien discret $(X_v)_{v\in V,\, v\neq o}$ est obtenue grâce à la matrice $(\overline{L})^{-1}$, où $\overline{L}$ est construite à partir de la matrice L à laquelle on a supprimé la colonne et la ligne correspondant au sommet $v = o$. Ceci coincide avec la fonction de Green d'une marche aléatoire simple (à temps continu) $(S_t)_{t\geq 0}$ sur G qui serait tuée lorsqu'elle atteint o. Autrement dit, en notant τ le temps d'arrêt définit par $\tau = \min\{t \geq 0,\, S_t = o\}$, on obtient

$$\mathbb{E}[X_v X_w] = \mathbb{E}^v\left[\int_0^\tau 1_{\{S_t = w\}} dt\right],$$

avec $\mathbb{E}^v[\cdot]$ désignant l'espérance conditionnelle sachant $S_0 = v$. Dans le cas particulier où $d_v = d$ pour tout $v \neq o$, le même raisonnement montre que la fonction de covariance du champ libre gaussien discret correspond à la fonction de Green d'une marche aléatoire simple (à temps discret) tuée en o. Autrement dit,

$$\mathbb{E}[X_v X_w] = \mathbb{E}^v\left[\frac{1}{d}\sum_{n=1}^{\tau-1} 1_{\{S_n = w\}}\right].$$

En fait, il s'agit d'un champ libre gaussien avec des conditions de Dirichlet au bord. Le comportement de la fonction de covariance est donc déterminé par celui de la marche aléatoire simple sur le graphe. Il est bien connu que la marche aléatoire symétrique sur $\mathbb{Z}^d$ présente des comportements différents suivant la dimension. Lorsque $d = 2$ la fonction de Green associée à cette marche aléatoire admet un comportement logarithmique tandis qu'en dimension supérieur elle se comporte comme l'inverse d'une fonction puissance. Tandis que la majeure partie de la littérature se concentre sur le cas difficile de la dimension deux, les articles de [60, 59, 58] s'intéressent aux comportements du champ libre gaussien lorsque $d \geq 3$. En fait, il est possible de construire un champ libre gaussien discret sur un groupe infini $\mathcal{S}$ de type fini. Dans [60, 59, 58] les auteurs considèrent un tel objet sur $\mathbb{Z}^d$, $d \geq 3$. Comme nous le verrons, le comportement de la marche aléatoire symétrique sur $\mathcal{S}$ conditionne le comportement de la fonction de covariance.

On peut alors considèrer un champ libre gaussien discret $(X_s)_{s\in\mathcal{S}}$ indexé par $\mathcal{S}$, un groupe infini de type fini. En effet, on peut considérer une marche aléatoire symétrique $(S_n)_{n\geq 0}$ sur Σ une partie génératrice symétrique de $\mathcal{S}$. Sous certaines hypothèses (cf [138]) sur la croissance de volume de boule, la fonction de covariance de ce champ libre gaussien discret admet un comportement similaire à celui étudié par Cipriani *et al.* dans [60, 59, 58].

Nous proposons ci-dessous un bref survol historique de la littérature concernant les travaux portant sur le champ-libre gaussien discret et les modèles liés. Les mathématiciens Bramson, Bolthausen, Deuschel, Giacomin, Kumagaï et Zeitouni sont à la base d'une série d'articles étudiant le maximum du champ libre gaussien discret sur $\mathbb{Z}^2$, beaucoup d'effort sont fait pour prouver que la suite du maximum (recentré) est tendue [34, 35, 106, 45]. Dans les articles de Ding [68], puis Ding-Zeitouni [72] des inégaliés de déviations (à droite et à gauche) précises du maximum recentré sont obtenues. Outre les méthodes de second moment, les principaux arguments utilisés sont alors des outils de comparaison comme l'inégalité de Slepian et ses variantes ainsi que des résultats classiques propres à la concentration de famille de variables aléatoires gaussiennes. Ding et Zeitouni proposent des modèles hierarchiques plus manipulables, présentant une structure d'arbre sous jacente, auxquels une comparaison avec le champ libre gaussien discret est possible. Signalons que ce genre d'arguments est repris dans l'article d'Acosta [2] Enfin, Bramson, Ding et Zeitouni obtiennent un résultat de convergence en loi du maximum recentré vers une variable aléatoire non triviale, correspondant une loi de Gumbel translatée aléatoirement [44]. Ils étendent ensuite leurs travaux à des marches aléatoires branchantes dans [43]. Ces résultats font écho aux travaux de Aïdekon et Madaule, sur des modèles aléatoires dit « log-corrélé »[6, 122], notamment sur l'interprétation de la loi limite en tant que dérivée de martingale. Nous renvoyons le lecteur curieux, vers les articles de survols de Duplantier, Rhodes, Sheffield et Vargas [74], ainsi que celui d'Arguin [9] pour plus de détails à ce sujet. Récemment, Ding, Roy et Zeitouni ont aussi étendu leurs résultats de convergence, obtenus pour le champ libre gaussien discret sur $\mathbb{Z}^2$, à des modèles log-corrélés plus généraux [71].

Faisant suite à cet intense développement autour du champ libre gaussien discret sur $\mathbb{Z}^2$, Cipriani *et al.* proposent dans une série d'article [60, 59, 58] sur l'étude du champ libre gaussien discret sur $\mathbb{Z}^d$, $d \geq 3$. Le comportement de la marche aléatoire symétrique sous jacente est alors différent et rend la structure de covariance plus manipulable. En plus des outils de comparaison déjà présent dans la littérature, les auteurs de [60, 59, 58] utilisent la méthode de Stein-Chen pour obtenir des convergences de processus ponctuel de Poisson associés aux extrêmes du champ libre gaussien discret. Nous verrons dans le chapitre 3 de la thèse, que nous pouvons retrouver une partie de leurs travaux via des arguments hypercontractif. Nous proposerons aussi une démonstration différente de la convergence en loi du maximum du champ libre gaussien discret sur $\mathbb{Z}^d$, $d \geq 3$. Nous aimerions porter à l'attention du lecteur que les articles mentionnés précédemment, sur le champ libre gaussien discret en dimension deux, sont souvent longs et difficiles techniquement, il serait intéressant de proposer des démonstrations plus souples, éventuellement par des méthodes de semi-groupes. Il conviendrait alors de traiter le cas de la marche aléatoire branchante en premier.

2.2.6 Percolation

Les modèles de percolation sont parmi les plus célèbres exemples présentant de la superconcentration. Cette célébrité est notamment du au fait que certains modèles de percolation dirigée ont de profonds liens avec la théorie des matrices aléatoires (cf. [114]).

Dans un premier temps nous allons définir un modèle, plus simple, de percolation, ensuite nous énoncerons le résultat de superconcentration le concernant. Ce théorème a été établit grâce à l'inégalité de Talagrand 1.3.6. Cependant, comme nous le mentionnerons, cet outil est insuffisant pour faire la démonstration complète. Les auteurs Benjamini, Kalai, Schramm ont trouvé une astuce (cf. [22]) pour contourner ces difficultés techniques.

On considère le réseau $\mathbb{Z}^d$ et nous noterons par E 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 ces 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. Ce modèle de premier temps de passage en percolation à été introduit par Hammersley et Welsh.

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 :

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

existe et est une fonction déterministe de x.

Par simplicité de notation posons $x=e_1=(1,0,...,0)$ et notons $T_n(x)$ par T_n. Grâce à des arguments de martingales, Kesten a prouvé que $\mathrm{Var}(T_n)\leq Cn$. Bien que cela n'atteigne pas encore l'ordre de fluctuations attendu par les physiciens ($n^{2/3}$). Benjamini, Kalai et Schramm ont amélioré l'estimation de Kesten, lorsque les poids sont des variables aléatoires de Bernoulli, en démontrant un phénomène de superconcentration. Chatterjee à étendu ce résultat lorsque les poids sont des variables aléatoires gaussiennes standards [48].

Théorème 2.2.14 (Chatterjee)**.** *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, uniformément bornée, de variables aléatoires gaussiennes , alors pour tout $n\geq 2$,*

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

où $C>0$ 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 ne s'applique pas directement. On suppose que la loi du poids de nos arêtes est une fonction lipschizienne 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 $F : \mathbb{R} \to \mathbb{R}$ 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. Sous ces conditions les poids s'expriment de la manière suivante $\omega_e = F(X_e)$. 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. Cette hypothèse est cruciale pour utiliser l'astuce BKS (de Benjamini-Kalai-Schramm).

Appelons p_I le chemin optimal entre l'origine et ne_1. Avec des notations évidentes, ceci nous donne que :

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

En conséquence,

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

où $p_e = \mathbb{P}(e \in p_I)$. C'est pourquoi nous devrons 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, pour cela ils introduisent un autre temps de passage $\tilde{T_n}$. Celui-ci est construit de la manière suivante : on choisit une boite B de longueur k fixée autour de l'origine et l'on définit $T_n^x, x \in B$ comme étant le premier temps de passage issu de x, ensuite $\tilde{T_n}$ est obtenu comme étant la moyenne sur les points de depart x de T_n^x. Autrement dit,

$$\tilde{T_n} = \frac{1}{\mathrm{Card}(B)} \sum_{x \in B} T_n^x,$$

avec $\mathrm{Card}(B)$ désignant le cardinal des points du réseau $\mathbb{Z}^d$ contenu dans B. Ce procédé de moyenne, permet d'obtenir un objet plus régulier sur lequel l'inégalité de Talagrand peut être appliquée et l'écart entre T_n et $\tilde{T_n}$ peut-être facilement contrôlé.

La première tentative d'extension à de nouveaux poids est faite par Benaïm et Rossignol [21]. Ils sont les premiers à proposer une approche plus abstraite permettant de considérer des poids plus généraux. Ils se restreignent à une classe de mesures de probabilités dite « nearly gamma » dans le sens où ces lois de probabilités ont des propriétés de concentration proches des lois gammas. Leur article apporte des arguments novateurs en percolation : une idée provenant de la théorie du transport combinée avec une inégalité alternative à celle de Sobolev logarithmique ainsi qu'une inégalité de Poincaré exponentielle dépendant de la dimension. Plus précisement, ils considèrent leurs poids comme étant la mesure image de la mesure gaussienne γ_n par une application T, ensuite ils utilisent l'inégalité de Falik et Samorodnisky [80] pour enfin appliquer l'inégalité de Sobolev logarithmique satisfaite par la mesure gaussienne γ_n. L'autre point important est l'utilisation du lemme suivant, s'apparentant à une inégalité de Poincaré exponentielle,

Lemme 2.2.15. *Soit Z une variable aléatoire et $K > 0$. Supposons que pour tout $|\theta| \leq 2/\sqrt{K}$,*

$$\mathrm{Var}\left(e^{\theta Z/2}\right) \leq K \frac{\theta^2}{4} \mathbb{E}\left[e^{\theta Z}\right]. \tag{2.9}$$

Alors,

$$\mathbb{P}\left(|Z - \mathbb{E}[Z]| > t\right) \leq 6e^{-ct/\sqrt{K}},$$

pour tout $t \geq 0$, avec $c > 0$ une constante numérique.

ils utilisent ce résultat en faisant dépendre K de n. Grossièrement, après voir introduit le temps de passage moyen de Benjamini-Kalai et Schramm, ils appliquent le résultat précédent et montre que K_n est de l'ordre de $n/\sqrt{\log n}$. Ils obtiennent donc une inégalité de superconcentration pour le premier temps de passage qui renforce alors, au niveau exponentiel, les travaux de Chatterjee. Le principal défaut de leur approche est que la condition « nearly-gamma » est trop restrictive et exclue des lois usuelles de probabilités.

Néanmoins, leur approche plus abstraite a permit d'ouvrir la voie à Damron, Hanson et Sosoe qui, dans une série de plusieurs articles [66, 65], clarifient les arguments de Benaïm et Rossignol et étendent leurs résultats. Ils montrent, dans l'article [66], que les résultats de Chatterjee sur la variance sont toujours valides si l'on suppose uniquement que les poids sont dans l'espace d'Orlicz $L^2 \log L$ et clarifient les arguments de [21]. Dans leur deuxième article, ils obtiennent une inégalité de superconcentration exponentielle, identique à celle de Benaïm et Rossignol, sous la condition que les poids possèdent un moment exponentiel. Les outils qu'ils utilisent sont essentiellement les mêmes que ceux de [21], en revanche, au lieu « d'encoder » leurs poids comme provenant d'une mesure gaussienne, ils proposent un encodage à partir de variables aléatoires de Bernoulli qui satisfont également une inégalité de Sobolev logarithmique. Cette observation accorde beaucoup plus de souplesse dans le choix des poids et permet d'étendre considérablement les travaux de [21], en revanche cela induit beaucoup plus de technicité lors du calculs des dérivées discrètes intervenant dans l'inégalité de Sobolev logarithmique. Pour plus de détails sur ces modèles de percolation, nous renvoyons le lecteur vers l'article de survol [12].

2.2.7 Polymères dirigés en milieu aléatoire

Les polymères dirigés en milieu aléatoire sont un autre exemple où l'on peut exhiber de la superconcentration. Bien que les démonstration de tels résultats soient plus difficiles, notamment à cause de la complexité géométrique de ce modèle, les techniques employées sont similaires à celles présentées dans la section précédente et repose sur l'inégalité de Talagrand et l'astuce de Benjamini-Kalai-Schramm. Présentons à présent ce nouveau modèle, qui a été introduit pour la première fois par Imbrie et Spencer.

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 $\{(0, S_0), (1, S_1), \ldots, (n, S_n)\}$ où $S_0 = 0$ et $|S_{i+1} - S_i| = 1$ 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-right 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 $(X_e)_{e\in\mathbb{Z}^2}$ une famille de gausiennes standards indépendantes de même loi que l'on appellera ≪ le milieu ≫ ou bien ≪ l'environnement ≫. Étant donné un polymère, obtenu via une marche aléatoire, π de longueur n, on définit l'énergie de ce chemin par :

$$H_n(\pi) := -\sum_{e\in\pi} X_e.$$

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 ≫) et nous le noterons par E_n. Nous étudierons également le chemin p_I 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, du à Chatterjee [48] :

Théorème 2.2.16. *(Chatterjee) Si E_n est l'état fondamental d'un polymère gaussien aléatoire de dimension $(1+1)$, alors, pour tout $n \geq 2$,*

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

où $C > 0$ est une constante numérique.

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

$$\frac{\partial E_n}{\partial X_e} = -1_{\{e\in p_I\}},$$

qui nous fournit que

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

Alors l'inégalité de Poincaré nous assure que :

$$\mathrm{Var}(E_n) \leq n.$$

L'idée de la preuve est la même que pour le modèle de percolation. C'est à dire que l'on introduit une boite B d'une longueur fixée (à choisir convenablement), ainsi qu'une moyenne $\tilde{E}_n$ des énergies des chemins de taille n issus d'un point $x \in B$ sur laquelle l'inégalité de Talagrand s'applique plus facilement. Il reste ensuite à contrôler l'ecart entre E_n et $\tilde{E}_n$.

Ce résultat de Chatterjee a été étendu par Graham pour des polymères de dimension $d+1$

(avec $d \geq 1$) dans [89]. Les travaux de [21] ont également inspiré Alexander et Zygouras, qui ont adopté la même méthodologie pour étudier un modèle de polymères dirigés en milieu aléatoire avec une température. Ils obtiennent alors dans [13] une inégalité de superconcentration exponentielle pour la fonction de partition d'un tel modèle.

2.3 Approche hypercontractive et méthode d'interpolation par semi-groupe

Cette dernière approche de la superconcentration est celle que nous avons privilégiée durant cette thèse.

2.3.1 Inégalité de Chatterjee

Nous énonçons ci-dessous un résultat du à Chatterjee [48], que nous étendrons au niveau exponentiel pour obtenir des inégalités de superconcentration dans le chapitre suivant. Son théorème permet de borner la variance du maximum d'un vecteur gaussien sous certaines conditions sur la matrice de covariance. Il s'agit donc d'une extension de l'inégalité de Talagrand 1.3.6 pour le cas particulier de la fonction maximum. L'énoncé du théorème est légèrement différent de celui présenté dans l'ouvrage [48], cela permet plus de souplesse lors de son application. Nous adoptons les notations suivantes, qui seront conservées durant tout ce manuscrit, $I = \mathrm{argmax}_{i=1,\dots,n} X_i$ pour X un vecteur gaussien de matrice de covariance Γ et par analogie $I^t = \mathrm{argmax}_{i=1,\dots,n} X_{t,i}, \quad t \geq 0$ pour $(X_t)_{t\geq 0}$ le processus d'Ornstein-Uhlenbeck dans $\mathbb{R}^n$. Celui-ci est construit de la manière suivante :

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

avec X un vecteur gaussien standard de $\mathbb{R}^n$ et Y une copie indépendante de X. Il se trouve que l'on peut étendre la définition du processus d'Ornstein-Uhlenbeck en remplaçant X par un vecteur gaussien centré de matrice de covariance Γ et, à nouveau, Y une copie indépendante de X. Par abus de notations, nous continuerons à désigner par $(X_t)_{t\geq 0}$ cette généralisation du processus d'Ornstein-Uhlenbeck.

On peut facilement montrer (c.f. [48]) que les propriétés d'hypercontractivité du processus d'Ornstein-Uhlenbeck standard sont conservées par le processus d'Ornstein-Uhlenbeck généralisé $(X_t)_{t\geq 0}$ défini ci-dessus et son semi-groupe associé. En effet, soit $M \in M_n(\mathbb{R})$ telle que $M^tM = \Gamma$, que l'on peut supposer inversible sans perdre de généralité (dans le cas contraire il suffirait de se restreindre à un sous-espace de $\mathbb{R}^n$), et définissons le semi-groupe $(Q_t)_{t\geq 0}$ par

$$Q_t f(x) = \mathbb{E}\left[f\big(X_t\big)\big|\, X = x\right], \quad x \in \mathbb{R}^n$$

pour des fonctions $f : \mathbb{R}^n \to \mathbb{R}$ suffisament régulières. Soit maintenant Z un vecteur gaussien standard dans $\mathbb{R}^n$, on a donc $Y = MZ$ en loi. Alors, en définissant $g : \mathbb{R}^n \to \mathbb{R}$ par $g(x) = f(Mx)$, nous avons pour tout $x \in \mathbb{R}^n$, l'expression suivante du semigroupe $(Q_t)_{t\geq 0}$ associé au processus d'Ornstein-Uhlenbeck général

$$\begin{aligned}
Q_t f(x) &= \mathbb{E}\left[f\big(e^{-t}x + \sqrt{1-e^{-2t}}Y\big)\right] \\
&= \mathbb{E}\left[f\big(e^{-t}MM^{-1}x + \sqrt{1-e^{-2t}}MZ\big)\right] \\
&= \mathbb{E}\left[g\big(e^{-t}M^{-1}x + \sqrt{1-e^{-2t}}Z\big)\right] = P_t g(M^{-1}x)
\end{aligned}$$

Où, rappelons le, $(P_t)_{t\geq 0}$ désigne le semigroupe associé au processus d'Ornstein-Uhlenbeck standard. Ainsi, la propriété d'hypercontractivité de $(P_t)_{t\geq 0}$ est transmise à $(Q_t)_{t\geq 0}$ avec les mêmes exposants p et q. Par la suite, par abus de notation, nous noterons $(Q_t)_{t\geq 0}$ par $(P_t)_{t\geq 0}$. Le même type de formule de représentation de la variance le long de ce semi-groupe est également disponible, cette fois-ci avec la présence supplémentaire des termes de la matrice de covariance. Plus précisément (cf. [48] lemme 10.2),

Lemme 2.3.1. *Soit X un vecteur gaussien dans $\mathbb{R}^n$ de matrice de covariance Γ. Alors, l'égalité suivante est satisfaite :*

$$\frac{d}{dt}\mathbb{E}[\big(P_t f(X)\big)^2] = -2e^{-2t}\sum_{i,j=1}^{n}\Gamma_{ij}\mathbb{E}[\partial_i f(X) P_t \partial_j f(X)]$$

Ce lemme permet d'obtenir la représentation de la variance sus-mentionnée

Corollaire 2.3.2. *Soit X un vecteur gaussien dans $\mathbb{R}^n$ de matrice de covariance Γ, avec $\Gamma_{ij}\geq 0$ pour tout $i,j=1,\dots,n$, alors pour toute fonction $f:\mathbb{R}^n\to\mathbb{R}$ suffisamment régulière*

$$\mathrm{Var}\big(f(X)\big) = 2\int_0^{\infty} e^{-2t}\sum_{i,j=1}^{n}\Gamma_{ij}\mathbb{E}[\partial_i f(X) P_t(\partial_j f(X))]dt. \tag{2.10}$$

En particulier, si $f(x) = \max\limits_{i=1,\dots,n} x_i$,

$$\begin{aligned}\mathrm{Var}(M_n) &= 2\int_0^{\infty} e^{-2t}\sum_{i,j=1}^{n}\Gamma_{ij}\mathbb{E}[1_{I=i}P_t(1_{I=j})]dt\\ &= 2\int_0^{\infty} e^{-2t}\sum_{i,j=1}^{n}\Gamma_{ij}\mathbb{P}(I=i, I^t=j)dt\\ &= 2\int_0^{\infty} e^{-2t}\mathbb{E}[\Gamma_{II^t}]dt,\end{aligned}$$

où, rappelons le, $I = \mathrm{argmax}_{i=1,\dots,n}X_i$ et $I^t = \mathrm{argmax}_{i=1,\dots,n}X_i^t, \quad t\geq 0$.

Nous pouvons donc énoncer le théorème de Chatterjee.

Théorème 2.3.3. *(Chatterjee) Soit $X=(X_1,\dots,X_n)$ un vecteur gaussien de matrice de covariance Γ.*

Supposons qu'il existe $r_0\geq 0$ tel que pour tout $r\geq r_0$ il existe un ensemble $\mathcal{C}(r)\subset\mathcal{P}\big(\{1,\dots,n\}\big)$ satisfaisant la condition suivante : pour tout $i,j\in\{1,\dots,n\}$ vérifiant $\Gamma_{ij}\geq r$, il existe $D\in\mathcal{C}(r)$ tel que $i,j\in D$. Posons $\sigma^2=\max_{i=1,\dots,n}\mathrm{Var}(X_i)$ et définissons, pour tout $r\geq r_0$,

$$\rho(r) = \max_{D\in\mathcal{C}(r)}\mathbb{P}(I\in D),$$

$$\tau(r) = \sum_{D\in\mathcal{C}(r)}\mathbb{P}(I\in D) = \mathbb{E}\Big[\mathrm{Card}\big(\{D\in\mathcal{C}(r)\,:\,I\in D\}\big)\Big].$$

Alors,

$$\mathrm{Var}(M_n)\leq r_0 + \int_{r_0}^{\sigma^2}\frac{2\tau(r)\big(1-\rho(r)\big)}{\log(1/\rho(r))}dr. \tag{2.11}$$

Remarque. Faisons quelques commentaires sur les nouveaux objets introduits dans ce théorème

1. $\rho(r)$ désigne le sous-ensemble de $\mathcal{C}(r)$ qui a le plus de chance de contenir l'argmax I.

2. $\tau(r)$ désigne, en moyenne, le nombre de sous-ensembles de $\mathcal{C}(r)$ contenant I.

3. Il est simple de montrer que s'il existe un tel « recouvrement » $\mathcal{C}_{r_0}$ alors ce même recouvrement fonctionne pour tout $r_0 \leq r \leq \sigma^2$.

4. Il est implicite dans le livre de Chatterjee [48] qu'un argument de convergence monotone permet d'obtenir le même résultat avec un supremum (sur un sous ensemble A de $\mathbb{R}^n$) au lieu du maximum.

5. Comme nous le constaterons sur des exemples, il n'est pas nécessaire d'obtenir un véritable recouvrement de $\{1, \ldots, n\}$. L'important est de construire $\mathcal{C}(r_0)$ de telle sorte que si $\mathrm{Cov}(X_i, X_j)$, $\forall i, j \in \{1, \ldots, n\}$ est plus gros qu'une certaine quantité, alors i, j doivent nécessairement appartenir à un sous ensemble $\{1, \ldots, n\}$. Puis il faut s'assurer que ces sous-ensembles de $\{1, \ldots, n\}$ (contenu dans $\mathcal{C}(r_0)$) ne s'intersectent pas trop.

Démonstration. La preuve repose sur l'idée suivante, on souhaite regrouper par paquets de « même taille » (suivant leurs corrélations) les atomes $1_{I=i}1_{I^t=j}$ puis d'appliquer la propriété d'hypercontractivité du semi-groupe d'Ornstein-Uhlenbeck. Choisissons $p > 1$ et soit $q = 1 + e^{2t}(p-1)$. Posons $q' = q/(q-1)$. Alors pour toute fonction $f : \mathbb{R}^n \to [0,1]$ l'inégalité de Hölder et la propriété d'hypercontractivité de P_t entrainent que

$$\begin{aligned}
\mathbb{E}\big[f(X)f(X^t)\big] &= \mathbb{E}\big[f(X)P_tf(X)\big] \leq \|f(X)\|_{q'}\|P_tf(X)\|_q & (2.12)\\
&\leq \|f(X)\|_{q'}\|f(X)\|_p & (2.13)\\
&\leq \Big(\mathbb{E}\big[f(X)\big]\Big)^{\frac{1}{q'}+\frac{1}{p}} & (2.14)
\end{aligned}$$

où dans la dernière inégalité nous avons utilisé le fait que $f(X) \in [0,1]$ ce qui implique que $f(X)^k \leq f(X)$ pour tout $k \geq 0$. Le choix de $p = 1 + e^{-t}$ entraine, après un calcul élémentaire,

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

C'est pourquoi, pour tout $r \geq r_0$, l'hypothèse sur $\mathcal{C}(r)$ entraine

$$\begin{aligned}
\mathbb{P}\big(\Gamma_{II^t} \geq r\big) &\leq \mathbb{P}(\exists D \in \mathcal{C}(r) \,:\, I, I^t \in D)\\
&\leq \sum_{D \in \mathcal{C}(r)} \mathbb{P}(I, I^t \in D)\\
&\leq \sum_{D \in \mathcal{C}(r)} \mathbb{P}(I \in D)^{1+\tanh(t/2)}\\
&\leq \rho(r)^{\tanh(t/2)} \sum_{D \in \mathcal{C}(r)} \mathbb{P}(I \in D) = \tau(r)\rho(r)^{\tanh(t/2)}.
\end{aligned}$$

Grâce à l'estimation (2.12) pour le choix de $f(x) = 1_{x \in D}$ et les définitions de ρ et τ.

Il ne reste plus qu'à utiliser la représentation de la variance 2.10 pour conclure. On a l'identité suivante

$$\mathrm{Var}(M_n) \leq \int_0^\infty e^{-t} \mathbb{E}[\Gamma_{II^t}] dt$$

qui peut se réécrire

$$\mathrm{Var}(M_n) \leq \int_0^\infty \int_0^{\sigma^2} e^{-t} \mathbb{P}\big(\Gamma_{II^t} \geq r\big) dr dt.$$

Après avoir échangées les intégrales grâce au théorème de Fubini-Tonelli, les majorations précédemments obtenues fournissent

$$\mathrm{Var}(M_n) \leq \int_0^{r_0} \int_0^\infty e^{-t} \mathbb{P}\big(\Gamma_{II^t} \geq r\big) dt dr + \int_{r_0}^{\sigma^2} \int_0^\infty e^{-t} \tau(r) \rho(r)^{\tanh(t/2)} dt dr.$$

Soit $r_0 \leq r \leq \sigma^2$, on trouve

$$\begin{aligned}
\int_0^\infty e^{-t} \rho(r)^{\tanh(t/2)} dt &\leq \int_0^\infty e^{-t} \tau(r) \rho(r)^{1-e^{t/2}} dt \\
&\leq \int_0^\infty e^{-t/2} \tau(r) \rho(r)^{1-e^{t/2)}} dt \\
&= \int_0^1 2\tau(r) \rho(r)^u du = \frac{2\tau(r)\big(1-\rho(r)\big)}{\log(1/\rho(r))}.
\end{aligned}$$

après un simple changement de variables ($u = 1 - e^{-t/2}$) dans l'intégrale, ce qui conclut la démonstration. □

2.3.2 Exemples d'applications

Enonçons, brièvement, des exemples présentés par Chatterjee dans [48]. Nous proposerons dans le chapitre suivant, pour chacun d'entre eux, des inégalités de superconcentration exponentielles renforçant les bornes sur la variance obtenues par Chatterjee via son Théorème 2.3.3.

Vecteur gaussien « échangeable »

Voici une application simple de l'inégalité due à Chatterjee. L'idée principale est d'utiliser la taille des corrélations pour se localiser dans l'espace des indices. Plus précisément, comme nous le verrons dans le modèle ci-dessous, lorsque deux variables X_i et X_j, $i, j \in \{i, \ldots, n\}$, sont très corrélées alors nécessairement $i = j$.

Proposition 2.3.4. *Soit $(X_1, \ldots, X_n)$ un vecteur gaussien centré tel que $\mathrm{Var}(X_i) = 1$, $i = 1, \ldots, n$ et supposons qu'il existe $0 < \epsilon \leq \frac{1}{2}$ tel que*

$$\Gamma_{ij} = \mathbb{E}[X_i X_j] \leq \epsilon, \ \forall i, j \in \{1, \ldots, n\}, i \neq j.$$

Alors, pour tout $n \geq 2$,

$$\mathrm{Var}(M_n) \leq \epsilon + \frac{C}{\log n},$$

avec $C > 0$ une constante numérique.

La preuve est instructive puisqu'elle permet de mieux saisir comment construire le « recouvrement » apparaissant dans l'inégalité de Chatterjee.

Le principe de minoration de Sudakov (cf. [117, 17, 156, 153]) nous sera également utile par la suite, nous rappelons son énoncé ci-dessous

Proposition 2.3.5 (Sudakov)**.** *Soit X un vecteur gaussien centré de $\mathbb{R}^n$, supposons qu'il existe une constante $\delta > 0$ telle que $\mathbb{E}[(X_i - X_j)^2]^{1/2} \geq \delta$ pour tout $i \neq j \in \{1, \dots, n\}$ alors*

$$\mathbb{E}[\max_{i=1,\dots,n} X_i] \geq C\delta\sqrt{\log n}, \tag{2.15}$$

avec $C > 0$ une constante numérique.

Démonstration. Soit $r_0 = \epsilon$, on définit alors, pour tout $r > r_0$, l'ensemble $\mathcal{C}(r)$ par

$$\mathcal{C}(r) = \big\{\{1\}, \dots, \{n\}\big\}.$$

Ce recouvrement vérifie les hypothèses du théorème 2.3.3. En effet, si $\Gamma_{ij} = \mathbb{E}[X_i X_j] \geq r > \epsilon$, cela implique que $i = j$ puisque la corrélation entre deux variables est au plus ϵ. Donc $i, j \in \big\{\{1\}, \dots, \{n\}\big\}$.

Il reste à obtenir des bornes sur $\mu(r)$ et $\rho(r)$ pour $r > r_0$.

Par construction $\mathcal{C}(r)$ est une partition de $\mathbb{R}^n$ donc $\tau(r) = \sum_{i=1}^n \mathbb{P}(I = i) = 1$. Pour $\rho(r)$, considérons un indice $1 \leq i \leq n$ et $t \in \mathbb{R}$

$$\begin{aligned}
\mathbb{P}(I = i) &\leq \mathbb{P}(X_i \geq t) + \mathbb{P}(\max_{i=1,\dots,n} X_j \leq t) \\
&\leq e^{-t^2/2} + e^{-\frac{(t - \mathbb{E}[M_n])^2}{2}}
\end{aligned}$$

où nous avons utilisé les propriétés de concentration satisfaites pour les fonctions lipschitziennes de variables gaussiennes (1.1). A présent, optimisons en $t = \mathbb{E}[M_n]/2$. Ceci entraine

$$\mathbb{P}(I = i) \leq 2e^{-\frac{\mathbb{E}[M_n]^2}{2}}.$$

Finalement, puisque $\epsilon \leq \frac{1}{2}$ l'inégalité de Sudakov (2.15) implique que

$$\mathbb{E}\left[\max_{i=1,\dots,n} X_j\right] \geq c\sqrt{\log n},$$

avec $c > 0$ une constante numérique. Alors $\mathbb{P}(I = i) \leq n^{-c}$ et l'application de l'inégalité (2.11) permet de conclure. $\square$

Champ gaussien euclidien

Comme nous l'avons mentionné plus tôt l'inégalité de Chatterjee (2.11) s'étend, par convergence monotone, pour obtenir un résultat sur des supremums de processus gaussiens. Dans son livre, Chatterjee montre que cette extension est pertinente pour des champs gaussiens euclidiens stationnaires ; notamment vis à vis de la théorie des extrêmes, mais nous reviendrons plus tard sur ce sujet. Il devient nécessaire d'adopter de nouvelles notations avant de présenter le résultat de Chatterjee.

1. Nous noterons le nombre d'entropie $N(A,\epsilon)$ d'un ensemble $A \subset \mathbb{R}^d$ par $N(A)$ lorsque $\epsilon = 1$. Rappelons, à toute fin utile, que $N(A,\epsilon)$ correspond au nombre maximum de points, appartenant à l'ensemble A, tel que la distance entre chaque couple vaut, au moins, ϵ.
2. Nous désignerons par $X = (X_u)_{u\in\mathbb{R}^d}$ un champ gaussien centré sur $\mathbb{R}^d$ avec $\mathrm{Var}(X_u) = 1$ for tout $u \in \mathbb{R}^d$.
3. Pour tout $u \in \mathbb{R}^n$ et tout $r > 0$ on désignera la boule ouvert de rayon r, centrée en u, par $B(u,r)$.
4. Soit $A \subset \mathbb{R}^d$, on désignera le supremum de $(X_u)_{u\in\mathbb{R}^d}$ par

$$M_A = \sup_{u\in A} X_u$$

Si l'on suppose que le champ gaussien vérifie les hypothèses suivantes :

1. $L = \sup_{u\in\mathbb{R}^d} \mathbb{E}[M_{B(u,1)}] < \infty$.
2. Il existe $\phi : [0,\infty) \to \mathbb{R}$ une fonction **décroissante** telle que pour tout $u, v \in \mathbb{R}^d$

$$\mathrm{Cov}(X_u, X_v) \le \phi\big(|u-v|\big).$$

   De plus, supposons que

$$\lim_{s\to\infty} \phi(s) = 0.$$

   Remarquons que cette hypothèse n'est utilisée que pour des temps très grands puisque $\mathrm{Cov}(X_u, X_v) \le 1$ pour tous $u, v \in A$.

Sous ces hypothèses Chatterjee a obtenu le théorème suivant :

Théorème 2.3.6. *(Chatterjee) Pour tout borélien $A \subset \mathbb{R}^d$ tel que $\mathrm{diam}(A) > 1$, nous avons*

$$\mathrm{Var}(M_A) \le C_1\left(\phi\big(N(A)^{C_2}\big) + \frac{1}{\log N(A)}\right),$$

avec $C_1 = C_1(\phi,d)$ et $C_2 = C_2(\phi,d)$ des constantes numériques ne dépendant que de la fonction ϕ et de la dimension d.

Remarque. L'inégalité précédente montre donc qu'il suffit de comparer la vitesse de $s \mapsto \phi(s)$ avec $s \mapsto 1/\log s$ pour éventuellement obtenir de la superconcentration. La preuve est omise pour le moment, les principales idées de la preuve seront utilisées plus tard dans le manuscrit pour obtenir de la superconcentration pour des champs libres gaussien discrets sur des groupes infinis de type fini satisfaisant des hypothèses de croissance de volume ($\mathbb{Z}^d$, $d \ge 3$ par exemple).

Remarquons aussi que les bornes obtenues par ce résultat implique la quantité $N(A)$ pour un ensemble A dépendant, en pratique, d'un paramètre n ou T (par exemple $A = [0, T]$ ou bien $A = \{1, \ldots n\}$). Ces bornes ne sont pertinentes qu'à condition que $N(A_n) \to \infty$ lorsque $n \to \infty$. En effet, si $N(A_n)$ est constante, la borne obtenue est d'ordre 1 et correspond à celle fournie par l'inégalité de Poincaré.

Chapitre 3

Rappels : théorie des extrêmes

Dans cette section, nous allons faire de brefs rappels concernant la théorie des extrêmes. Nous nous focaliserons principalement sur la convergence en loi du maximum. Comme nous le verrons dans le chapitre suivant, nous avons obtenu, au cours de cette thèse, des bornes non asymptotiques sur la variance du maximum reflétant la renormalisation des théorèmes de convergence des extrêmes. Similairement, nous exhiberons des inégalités de (super)concentration, non-asymptotique, en adéquation avec la convergence des extrêmes et le comportement des queues de distributions des lois limites. Pour plus de détails sur la théorie des extrêmes, nous renvoyons vers les ouvrages de référence suivants : [108, 67, 78].

Etant donné une mesure de probabilité μ ainsi qu'un échantillon de variables aléatoires $X_1, \dots X_n$, indépendantes, telles que $\mathcal{L}(X_1) = \mu$; on désignera par $M_n = \max_{i=1,\dots,n} X_i$, le maximum des X_i, $i = 1, \dots, n$. C'est un fait classique, lorsque μ est absolument continue par rapport à la mesure de Lebesgue, que l'on puisse déterminer des constantes de renormalisation a_n et b_n de telle sorte que $a_n(M_n - b_n)$ converge en loi, vers une limite non-triviale, lorsque $n \to \infty$.

Les lois limites sont complétement caractérisées (cf. [108, 67, 78]) et sont restreintes au nombre de trois. Nous présentons, ci-dessous, leurs fonctions de répartitions.
La loi de Fréchet :

$$\Xi_\alpha(x) = \begin{cases} 0, \quad x \leq 0, \\ \exp(-x^{-\alpha}), \quad x > 0, \quad \alpha > 0 \end{cases}$$

La loi de Gumbel :

$$\Lambda_0(x) = \exp(-e^{-x}), \quad x \in \mathbb{R}.$$

Enfin, la loi de Weibull :

$$\Psi_\alpha(x) = \begin{cases} \exp\big(-(-x)^\alpha\big), \quad x < 0, \quad \alpha > 0 \\ 1, \quad x \geq 0. \end{cases}$$

Des conditions nécessaires et suffisantes existent pour déterminer si une loi μ appartient au domaine d'attraction d'une de ces trois lois limites. Par exemple, si G désigne la fonction de répartition de la loi μ, ces conditions peuvent s'exprimer en terme de la queue de distribution de $1 - G$ (*cf.* [108, 67, 78]) et les constantes de renormalisation a_n et b_n peuvent être déterminées à partir de cette même fonction. A titre illustratif, une loi μ de fonction de répartition G appartient au domaine d'attraction de la loi de Fréchet Ξ_α, $\alpha > 0$ si et seulement si $1 - G(x) = x^{-\alpha}L(x)$ avec L une fonction à variation lente. De plus, a_n peut être choisi comme valant $1/\inf\{x \in \mathbb{R} : \frac{1}{1-G(x)} \geq n\}$. Des conditions similaires peuvent être obtenues pour le domaine d'attraction de la loi de Weibull, tandis que le cas de la loi de Gumbel est un peu plus compliqué et ne s'énonce pas aussi facilement. Dans ce qui suit, nous allons présenter quelques résultats de convergence pour les extrêmes de lois de probabilités usuelles. Dans le reste du manuscrit nous obtiendrons des majorations de la variance de M_n, du même ordre de grandeur que $(a_n)^{-2}$, pour ces exemples spécifiques. Nous porterons une attention particulière sur les lois appartenant au domaine d'attraction de la loi de Gumbel. Nous parlerons également de la convergence des extrêmes dans le cas stationnaire lorsque l'aléa est gaussien, un résultat sur la sphère sera également obtenu. Dans les chapitres quatre et sept nous proposerons des inégalités de superconcentration non asymptotiques, pour le domaine d'attraction de la loi de Gumbel, reflétant ces théorèmes de convergence.

3.1 Echantillon de variables indépendantes et identiquement distribuées, mesures produits

Soit $n \geq 1$, dans cette section $M_n = \max_{i=1,\ldots,n} X_i$ avec $X_1, \ldots, X_n$ des variables aléatoires indépendantes identiquement distribuées de loi μ.

Théorème 3.1.1. *Voici quelques exemples de convergence des extrêmes pour des lois de probabilités usuelles (cf. [78] pages 153-157).*

1. *Domaine d'attraction de la loi de Fréchet. :*
   — *Si μ désigne la loi de Pareto de paramètre $\alpha > 0$ et $K > 0$, de densité*

$$f_{\alpha,K}(x) = \begin{cases} \frac{\alpha K^\alpha}{x^{\alpha+1}}, & x \geq K, \\ 0, & x < K \end{cases}$$

   par rapport à la mesure de Lebesgue, alors

$$a_n M_n \to \Xi_\alpha \tag{3.1}$$

   lorsque $n \to \infty$ avec $a_n = (Kn)^{-1/\alpha}$.

2. *Domaine d'attraction de la loi de Weibull :*
   — *Si μ désigne la loi uniforme sur $[0,1]$, alors*

$$a_n(M_n - 1) \to \Psi_1 \tag{3.2}$$

   lorsque $n \to \infty$ avec $a_n = n$.

   — *Plus géneralement, si μ désigne la loi beta de paramètres $a, b > 0$, de densité*

$$f_{a,b}(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \quad 0 < x < 1$$

   par rapport à la mesure de Lebesgue, alors

$$a_n(M_n - 1) \to \Psi_\alpha, \quad \alpha > 0 \tag{3.3}$$

lorsque $n \to \infty$ avec $a_n = \left(n\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b+1)}\right)^{-1/b}$ *(Γ désignant la fonction d'Euler).*

3. *Domaine d'attraction de la loi de Gumbel :*
   — *Si μ désigne la loi normale, alors*
   $$a_n(M_n - b_n) \to \Lambda_0, \tag{3.4}$$
   lorsque $n \to \infty$ avec $a_n = \sqrt{2\log n}$ et $b_n = a_n - \frac{\log(4\pi)+\log\log n}{2a_n}$.

   — *Si μ désigne la loi gamma de paramètres $\alpha, > 0$ et $\beta = 1$, alors*
   $$a_n(M_n - b_n) \to \Lambda_0 \tag{3.5}$$
   lorsque $n \to \infty$ avec $a_n = 1$ et $b_n = \log n + (\alpha - 1)\log\log n - \log\Gamma(\alpha)$. La loi exponentielle standard correspondant au cas $\alpha = 1$.

   — *Si μ désigne la loi de Gumbel, alors*
   $$a_n(M_n - b_n) \to \Lambda_0 \tag{3.6}$$
   lorsque $n \to \infty$ avec $a_n = 1$ et $b_n = \log n$.

N'ayant pas trouvé dans la littérature de démonstration concernant la convergence des extrêmes de lois dites log-concave, *i.e.* $d\mu(x) = e^{-V(x)}\frac{dx}{Z}$ avec $V : \mathbb{R} \to \mathbb{R}$ convexe. Nous proposons ci-dessous une démonstration lorsque V est de classe $C^2(\mathbb{R})$ et satisfait quelques conditions supplémentaires. Dans le cas particulier où $V(x) = |x|^\alpha$, $\alpha \geq 1$ nous fournirons également les constantes de renormalisation explicites.

Pour démontrer notre résultat, nous utiliserons la condition nécessaire suivante :

Proposition 3.1.2. *(Von Mises) Soit X une variable aléatoire de fonction de répartition $G \in C^2(\mathbb{R})$ et de loi μ. Si,*
$$\lim_{x\to\infty} \frac{(1 - G(x))G''(x)}{G'(x)^2} = -1 \tag{3.7}$$
alors μ appartient au domaine d'attraction de la loi de Gumbel et il existe a_n et b_n tels que $a_n(M_n - b_n) \to \Lambda_0$ lorsque $n \to \infty$.

Remarque. Il existe un résultat similaire, plus général, attestant que le résultat précédent reste vrai si l'on remplace ∞ par x_μ (désignant l'extrémité droite du support de la loi μ) dans la limite précédente. Par la suite, nous ne considérons que des lois log-concaves à support non borné.

Le cas gaussien étant traité grâce à l'inégalité de Mill (cf. annexe [48]), qui fournit des estimées de la queue de distribution gaussienne, nous allons utiliser un résultat anologue pour des lois log-concaves (cf. [8] chapitre 6).

Lemme 3.1.3. *Soit μ une mesure log-concave, dont le potentiel V satisfait les hypothèses suivantes*

1. *$V \in \mathcal{C}^2(\mathbb{R})$,*

2. *il existe $A > 0$ tel que $V'(x)$ ne s'annule pas lorsque $|x| \geq A$,*

3. $\frac{V''(x)}{[V'(x)]^2} \to 0$ *lorsque* $|x| \to +\infty$,

alors nous avons le résultat suivant

$$\int_x^\infty e^{-V(t)} \frac{dt}{Z} \sim_{+\infty} \frac{e^{-V(x)}}{V'(x)Z}, \tag{3.8}$$

pour x *suffisament grand.*

Ce lemme technique nous permet d'énoncer le théorème suivant

Théorème 3.1.4. *(Extrêmes log-concave) Sous les mêmes hypothèses, portant sur le potentiel* V*, que celles du lemme 3.8 la loi* μ *(log-concave) appartient au domaine d'attraction de la loi de Gumbel.*

Démonstration. La démonstration est immédiate. En effet, le lemme 3.8 permet d'utiliser le critère de Von Mises (3.7) et de conclure. □

Nous présentons quelques exemples de potentiels satisfaisant les conditions du théoreme 3.1.4.

Exemple 3.1.5. *Si l'on choisit* $V(x) = ax^2 + bx^4$ *les hypothèses sont satisfaites si* $a, b > 0$ *ou bien* $a = 0, b > 0$ *ou encore* $a > 0, b > 0$*. On peut également choisir* $V(x) = ax^2 + bx^{2p}$ *avec* $p \geq 1$ *ou* $V(x) = \frac{x^2}{2} + e^{-x}$*. Le potentiel, plus classique,* $V(x) = |x|^\alpha$ *avec* $\alpha \geq 1$ *fonctionne également.*

Il est important de déterminer les constantes de renormalisation a_n et b_n assurant la convergence de $a_n(M_n - b_n)$ vers la loi de Gumbel. Nous proposons, ci dessous, de traiter le cas du potentiel $V(x) = |x|^\alpha$ avec $\alpha \geq 1$.

Pour cela, nous allons suivre la démonstration du cas gaussien, qui correspond au potentiel quadratique avec $\alpha = 2$. Nous aurons besoin de l'asymptotique fournie par le lemme 3.8 dans nos calculs.

Proposition 3.1.6. *Dans le cas du potentiel* $V(x) = |x|^\alpha, \alpha \geq 1)$*, nous obtenons les constantes de renormalisation suivantes*

$$a_n = \sqrt{\alpha(\log n)^{2(\alpha-1)/\alpha}}$$

et

$$b_n = (\log n)^{1/\alpha} - \frac{\log(\alpha Z) + \frac{\alpha-1}{\alpha}\log\log n}{(\log n)^{(\alpha-1)/\alpha}},$$

avec $Z = \int_{\mathbb{R}} e^{-V(x)}dx$*, de telle sorte que* $a_n(M_n - b_n) \to \Lambda_0$ *lorsque* $n \to \infty$*.*

Remarque. Le cas $\alpha = 2$ correspond, aux constantes numériques près, au cas gaussien. Tandis que $\alpha = 1$ redonne le cas de la mesure exponentielle sur $\mathbb{R}_+$. Notons également que la même démonstration devrait fonctionner (pourvu que le même type d'asymptotiques soit satisfait) si l'on remplace $|x|^\alpha$ par un potentiel V abstrait. Toutefois, il nous a semblé plus pertinent de travailler avec un exemple plus explicite.

Démonstration. Nous suivons la démonstration de [108] du cas gaussien. Similairement, nous posons $u_n = \frac{x}{a_n} + b_n$, $n \geq 1$ et l'on désigne par G la fonction de répartition de la loi log-concave ν. La relation suivante est satisfaite pour tout $x \in \text{supp}(\mu)$,

$$\mathbb{P}\big(a_n(M_n - b_n) \leq x\big) = G(u_n)^n = \big[1 - \big(1 - G(u_n)\big)\big]^n$$

Il suffit donc de trouver a_n et b_n telles que $1-G(u_n) = \frac{e^{-x}}{n}$ pour conclure que $a_n(M_n - b_n) \to \Lambda_0$ lorsque $n \to \infty$. Puisque, $1 - G(u_n) \sim_{n\to+\infty} \frac{e^{-u_n^\alpha}}{Z\alpha u_n^{\alpha-1}}$, nous pouvons réécrire $1 - G(u_n) \sim \frac{e^{-x}}{n}$ de la manière suivante

$$-\log n - x + \log\alpha + (\alpha-1)\log(u_n) + \log Z + u_n^\alpha \to_{n\to+\infty} 0.$$

Ceci entraine que $\frac{u_n^\alpha}{\log n} \to_{n\to+\infty} 1$, autrement dit $\alpha \log u_n - \log\log n \to_{n\to+\infty} 0$. En reformulant ceci, on trouve que $\log u_n = \frac{1}{\alpha}\log\log n + o(1)$. Ainsi, on obtient

$$-\log n - x + \log(\alpha) + \frac{\alpha-1}{\alpha}\log\log n \log Z + u_n^\alpha = o(1), \quad n \to +\infty$$

et des calculs élémentaires fournissent que

$$u_n = (\log n)^{1/\alpha}\left[1 + \frac{x - \log(\alpha Z) - \frac{\alpha-1}{\alpha}\log\log n}{\alpha \log n} + o(\frac{1}{\log n})\right]$$

ce qui est exactement la conclusion désirée. □

Remarque. La démonstration précédente suggère que $\mathrm{Var}(M_n) \leq \frac{C}{(\log n)^{2(\alpha-1)/\alpha}}$. Il est bien connu que lorsque $\alpha > 2$ la mesure log-concave μ provenant du potentiel $V(x) = x^\alpha$ est hypercontractive ([16]), cependant la variance de M_n est plus petite que le cas gaussien ($\alpha = 2$). En particulier, l'approche hypercontractive et l'inégalité de Talagrand L^1/L^2 (1.3.6) améliore la concentration classique mais ne fournit pas le bon ordre de grandeur de la variance de M_n. Nous parviendrons à obtenir des bornes non-asymptotiques optimales, pour $\alpha > 2$, par des méthodes de transport. Lorsque $1 < \alpha < 2$ nous n'avons aucune estimation *a priori* de la variance de M_n, puisque la mesure μ n'est plus hypercontractive.

La démonstration de la proposition 3.1.6 permet d'obtenir aisément des résultats de convergence pour les extrêmes lorsque l'on perturbe le potentiel quadratique du cas gaussien.

Corollaire 3.1.7. *On considère $V(x) = \frac{x^2}{2} + v(x)$, avec $v \in \mathcal{C}^2$ une fonction convexe satisfaisant $v(x) = o(x^2)$ et $v'(x) = o(x)$ lorsque $x \to \infty$.*

Alors les constantes de renormalisation a_n et b_n, nécessaires pour avoir la convergence du maximum, sont les mêmes que dans le cas gaussien. En particulier, la variance de M_n doit être de l'ordre de $C/\log n$.

Remarque. Par exemple, on peut choisir $v(x) = x^\alpha$, avec $0 \leq \alpha < 2$ ou bien $v(x) = e^{-x}$. Une généralisation évidente montre que que l'on peut aussi considérer $V(x) = ax^\alpha + bx^\beta$ pour $\alpha > \beta \geq 1$. Auquel cas, a_n et b_n correspondront aux mêmes constantes que celles obtenues pour $V(x) = x^\alpha$ (au constantes numériques près).

3.2 Famille de variables aléatoires gaussiennes

La théorie des extrêmes ne se limite pas uniquement au cas de variables indépendantes de mêmes lois. En effet, il existe de nombreux résultats concernant les suites stationnaires (cf. [108, 67, 78]). Il se trouve que le cas de suite, ou de processus gaussiens stationnaires est plus facile à traiter grâce à des outils de comparaison (propres aux variables gaussiennes) comme l'inégalité de Slepian et ses variantes. Les hypothèses assurant la convergence en loi des extrêmes, pour ce cas particulier d'aléa, s'expriment facilement à l'aide de la fonction de covariance. Nous présentons ci-dessous un certain nombre de résultats dans cet esprit. L'illustration non-asymptotique de ces théorèmes fut l'un des points majeurs de l'article [157] publié durant cette thèse. Après avoir décrit le cas gaussien stationnaire, nous présenterons le cas de la mesure uniforme sur la sphère. Ceci fournira un autre exemple de mesure non-produit, pour lequel nous énoncerons un résultat de convergence des extrêmes.

3.2.1 Suites gaussiennes stationnaires

Soit $(X_i)_{i\geq 0}$ une suite gaussienne stationnaire centrée telle que $\mathbb{E}\left[X_i^2\right] = 1$, $i \geq 0$, avec pour fonction de covariance $\mathrm{Cov}(X_i, X_j) = \phi(|i-j|)$, $i, j \geq 0$, où $\phi : \mathbb{N} \to \mathbb{R}$. De nombreux résultats ont été obtenus dans ce cadre afin d'assurer la convergence du maximum $M_n = \max_{i=1,\dots,n} X_i$ (cf. Berman, Mittal, Pickands... Cf. e.g. [108, 67, 78]). Nous présentons ci-dessous les différents comportements asymptotiques de M_n, après renormalisation, en fonction de la vitesse de décroissance de ϕ à l'infini.

Théorème 3.2.1. *Soit $(X_i)_{i\geq 0}$ une suite gaussienne stationnaire avec, pour fonction de covariance, ϕ telle que $\phi(n)\log n \to 0$ lorsque $n \to \infty$. Alors*

$$a_n(M_n - b_n) \underset{n\to\infty}{\longrightarrow} \Lambda_0$$

en loi, où $a_n = (2\log n)^{1/2}$ et

$$b_n = (2\log n)^{1/2} - \frac{1}{2}(2\log n)^{-1/2}(\log\log n + \log 4\pi)$$

Si jamais $\phi(n)\log(n) \to \eta > 0$ lorsque $n \to \infty$, la loi limite est modifiée. C'est le contenu du théorème suivant (celui-ci est présenté, dans [108] pages 137-138, en terme de processus ponctuel ; nous préferons reformuler ceci en terme fonction de répartition).

Théorème 3.2.2. *Soit $(X_i)_{i\geq 0}$ une suite gaussienne stationnaire avec, pour fonction de covariance, ϕ telle que $\phi(n)\log n \to \eta > 0$ as $n \to \infty$. Alors, pour tout $x \in \mathbb{R}$,*

$$\mathbb{P}(a_n(M_n - b_n) \leq x) \underset{n\to\infty}{\longrightarrow} \mathbb{P}\big(N([0,1]) = 0\big)$$

où $a_n = (2\log n)^{1/2}$ et

$$b_n = (2\log n)^{1/2} - \frac{1}{2}(2\log n)^{-1/2}(\log\log n + \log 4\pi)$$

et où N désigne un processus de Cox (cf. [108]). Plus précisément, il s'agit d'un processus ponctuel de Poisson d'intensité $e^{-x-\gamma+\sqrt{2\gamma}X}$, avec X une gaussienne standard réelle, et

$$\mathbb{P}\big(N([0,1]) = 0\big) = \int_{\mathbb{R}} e^{-e^{-x-\eta+\sqrt{2\eta}z}} d\gamma_1(z), \tag{3.9}$$

où, rappelons le, $d\gamma_1$ désigne la mesure Gaussienne standard. On peut voir ce processus de Poisson comme la convolution d'une loi de Gumbel avec une gaussienne. Nous renvoyons le lecteur vers l'ouvrage [108] (page 135) pour plus de détails à propos du processus ponctuel de Cox.

Si jamais les corrélations sont trop importantes et que $\phi(n)\log(n) \to \infty$ lorsque $n \to \infty$, la loi limite devient gaussienne.

Théorème 3.2.3. *(Mittal-Ylvisaker) Soit $(X_i)_{i\geq 0}$ une suite gaussienne stationnaire avec, pour fonction de covariance, ϕ telle que $\phi(n)\log n \to \infty$ lorsque $n \to \infty$, ϕ est décroissante et $\phi(n) \to 0$ lorsque $n \to \infty$. Alors*

$$c_n\left(M_n - \left(1 - c_n^{-2}\right)^{1/2} b_n\right) \underset{n\to\infty}{\longrightarrow} X,$$

où X suit une loi gaussienne standard et $c_n = \sqrt{\phi(n)^{-1}}$ et $b_n = (2\log n)^{1/2} - \frac{1}{2}(2\log n)^{-1/2}(\log\log n + \log 4\pi)$.

Remarque. Ce résultat est techniquement difficile. En revanche, au prix d'hypothèses supplémentaires sur la fonction de covariance, il est possible d'obtenir une preuve plus simple de ce résultat. Comme présenté dans [108], on peut supposer que la suite $(\phi(n))_{n\geq 0}$ est convexe et que $\phi(n) = C/(\log n)^\eta$ avec $C > 0$, $0 < \eta < 1$ à partir d'un certain $n_0 \geq 2$ pour arriver au même résultat de convergence. Ceci suggère alors que la variance de M_n est de l'ordre de $C/(\log n)^\eta$, $0 < \eta < 1$.

3.2.2 Processus gaussiens stationnaires

Des théorèmes de convergence similaires, en temps continu, ont été obtenus dans le cadre des processus gaussiens stationnaires. Par exemple, si $(X_t)_{t\geq 0}$ désigne un processus centré gaussien stationnaire tel que $\mathbb{E}\left[X_t^2\right] = 1$, $t \geq 0$, de fonction de covariance $\mathrm{Cov}(X_s, X_t) = \phi\big(|t-s|\big)$, pour $t, s \geq 0$, où $\phi : [0, +\infty) \to \mathbb{R}$. On peut considérer différents types de comportements, lorsque $t \to 0$, de la fonction de covariance ϕ :

$$\phi(t) = 1 - \frac{\lambda_2 t^2}{2} + o(t^2) \tag{3.10}$$

$$\phi(t) = 1 - C|t|^\alpha + o\big(|t|^\alpha\big) \tag{3.11}$$

Le premier cas entraine que X_t est differentiable et $\lambda_2 = -\phi''(0)$ est un moment spectral, tandis que le deuxième cas concerne les processus non différentiables (mais néanmoins continus). Par exemple, le processus d'Ornstein-Uhlenbeck relève du second cas. Pour plus de détails à ce sujet, cf. [108]. Pour tout $T > 0$, on pose $M_T = \sup_{t\in[0,T]} X_t$. Dans les deux cas, on obtient le résultat de fluctuations suivant.

Théorème 3.2.4. *Soit $(X_t)_{t\geq 0}$ un processus gaussien stationnaire tel que $\phi(t)\log t \to 0$ lorsque $t \to \infty$. Alors*

$$a_T(M_T - b_T) \underset{T\to\infty}{\longrightarrow} \Lambda_0$$

en loi, où $a_T = (2\log T)^{1/2}$ et b_T depend de l'hypothèse (3.10) ou (3.11).

3.2.3 Mesure uniforme sur la sphère

Celle-ci semblant absente de la littérature, nous proposons ci-dessous, une démontration d'un résultat de convergence en loi des extrêmes pour une autre mesure non-produit. Considèrons U un vecteur aléatoire tel que $\mathcal{L}(U) = \mathrm{U_n}(\mathbb{S}^n)$ et l'on note par M_n le maximum des U_i, $i = 1, \ldots, n$. On désigne à nouveau par a_n et b_n les constantes de renormalisation du cas gaussien standard (3.4), qui assurent la convergence en loi suivante du maximum de n gaussiennes standards indépendantes de même loi

$$a_n\Big(\max_{i=1,\ldots,n} X_i - b_n\Big) \to \Lambda_0$$

lorsque $n \to \infty$, vers Λ_0 la loi de Gumbel.

Théorème 3.2.5. *Avec les notations précédentes, nous avons la convergence en loi suivante $n \to \infty$*

$$\sqrt{n}a_n\big(M_n - \frac{b_n}{\sqrt{n}}\big) \to \Lambda_0$$

Démonstration. Nous utilisons le fait que l'on peut représenter le vecteur U par des gaussiennes, c'est à dire : $(U_1, \dots, U_n) = (\frac{X_1}{|X|}, \dots, \frac{X_n}{|X|})$ en loi, avec $|X| = \sqrt{\sum_{i=1}^n X_i^2}$ où X_i désigne des variables aléatoires gaussiennes standards indépendantes.

C'est pourquoi nous obtenons l'égalité en loi suivante

$$\sqrt{n}a_n\big(\max_{i=1,\dots,n} U_i - \frac{b_n}{\sqrt{n}}\big) = \frac{\sqrt{n}}{|X|}a_n\big(\max_{i=1,\dots,n} X_i - b_n\big) + \frac{b_n}{\sqrt{n}}\sqrt{n}\left(\frac{\sqrt{n}}{|X|} - 1\right)$$

Pour conclure, il suffit de traiter chaque terme séparément avec le lemme de Slutsky (cf. [160]).

Lemme 3.2.6. *(Slutsky) Soient $(X_n)_{n\geq 1}$, $(Y_n)_{n\geq 1}$, X des variables aléatoires et $c \in \mathbb{R}$ telles que $X_n \to X$ en loi et $Y_n \to c$ en probabilité , lorsque $n \to \infty$. Alors*

$$(X_n, Y_n) \to (X, c)$$

en loi lorsque $n \to \infty$. En particulier, $X_n + Y_n \to X + c$ et $X_nY_n \to cX$ en loi lorsque $n \to \infty$.

Pour le premier terme, on exprime $\frac{\sqrt{n}}{|X|}$ comme $1/\sqrt{\frac{1}{n}\sum_i X_i^2}$. Le théorème de la loi forte des grands nombres, appliqué aux variables X_i^2, entraine que $\frac{\sqrt{n}}{|X|} \to 1$ presque sûrement lorsque $n \to \infty$. Puisque $a_n(\max X_{i=1,\dots,n} - b_n) \to \Lambda_0$ en loi, le lemme de Slutsky 3.2.6 implique que $\frac{\sqrt{n}}{|X|}a_n(\max_{i=1,\dots,n} X_i - b_n) \to \Lambda_0$ lorsque $n \to \infty$.

Pour le second terme, nous allons combiner le théorème de la limite centrale à la méthode delta (cf. [160]). Plus précisément, on applique le théorème de la limite centrale à la suite $(X_i^2)_i$, puis la méthode delta avec la fonction $f : x \mapsto \frac{1}{\sqrt{x}}$. Il est immédiat, par le théorème de la limite centrale que

$$\sqrt{n}\left(\frac{|X|^2}{n} - 1\right) \to \mathcal{N}(0, \sigma^2),$$

en loi, lorsque $n \to \infty$, avec $\sigma^2 = \mathrm{Var}[X_1^2]$. De plus, $f(x) = \frac{1}{\sqrt{x}}$ est de classe C^1 au point $x_0 = 1$. Ainsi, la méthode delta entraine que

$$\sqrt{n}\left(f\left(\frac{|X|^2}{n}\right) - f(1)\right) \to f'(1)\mathcal{N}(0, \sigma^2)$$

en loi, lorsque $n \to \infty$. On remarque ensuite, que $\frac{b_n}{\sqrt{n}} \to 0$ lorsque $n \to \infty$. Ce qui permet d'en conclure, d'après le lemme de Slutsky 3.2.6, que le second terme tend vers 0 en probabilité. Une dernière utilisation du lemme de Slutsky permet d'achever la démonstration. □

Remarque. On illustrera l'extension de Cordero-Erausquin/Ledoux de l'inégalité de Talagrand 4.4.1 sur cet exemple.

3.2.4 Champ libre gaussien discret

Dans [60, 59], les auteurs ont considéré le champ libre gaussien discret sur $\mathbb{Z}^d$, $d \geq 3$. Ils ont notamment montré que le maximum de ce champ gaussien converge, après renormalisation, vers une loi de Gumbel. Pour cela, ils ont utilisé des arguments reposant sur des processus ponctuels de Poisson et la méthode de Stein-Chen. Nous montrons ci-dessous que des arguments, plus simples, de comparaison découlant de l'inégalité de Slepian permettent d'arriver à la même conclusion et restent valables pour des groupes plus généraux que $\mathbb{Z}^d$, $d \geq 3$.

Le cadre est le suivant : on considère un groupe infini S de type fini. Pour simplifier la lecture, on peut se restreindre au groupe $\mathbb{Z}^d$, $d \geq 3$. Rappelons quelques définitions concernant le croissance du volume de boule dans un tel groupe, pour plus de détail cf. [138]. Considérons Σ une partie génératrice symétrique finie de $\mathcal{S}$ et désignons par e l'élement neutre du groupe. Dans le cas de $\mathbb{Z}^d$, on peut choisir Σ comme étant l'ensemble suivant $(\pm e_1, \ldots, \pm e_d)$ avec e_i, $i = 1, \ldots, d$ les vecteurs de la base canonique de $\mathbb{R}^d$. Pour n'importe quel élément $x \in \mathcal{S}$ nous définissons la quantité suivante $|x|_\Sigma = \min\{k : \exists s_1, \ldots, s_k \in \Sigma, x = s_1 \ldots s_k\}$. En d'autres termes, $|x|_\Sigma$ est le nombre minimal d'éléments $s_1, \ldots, s_k$ appartenant à Σ nécessaires pour écrire x sous forme de produit $x = s_1, \ldots, s_k$. On peut alors définir une distance $d = d_\sigma$ sur S par $d(x, y) = |x^{-1}y|_\Sigma = d(e, x^{-1}y)$ pour tout $x, y \in \mathcal{S}$.

Cette distance induit une topologie sur S et l'on peut considérer la croissance du volume $n \mapsto V(n)$ d'une boule, centrée en l'élément neutre de S et de rayon n, avec

$$V(n) = \mathrm{Card}\{x \in S : |x|_\Sigma \leq n\}$$

Beaucoup de travaux ont été réalisés concernant le comportement des croissances de volume V [138, 137]. On dira que la croissance est polynomiale s'il existe $C, D > 0$ telles que $V(n) \leq Cn^D$, $n \geq 1$; on dira que la croissance est sous-exponentielle s'il existe $C > 0$ telle que $V(n) \leq Ce^n$, $n \geq 1$.

Remarque. Lorsque S est fini, $D = 0$ et $V(n) \leq C$ pour une constante $C > 0$. Il ne semble pas pertinent d'étudier la convergence des extrêmes sur ce type de groupe puisque pour ensemble $A \subset \mathcal{S}$ de cardinal $\mathrm{Card}(A) =$ on ne peut pas faire tendre n vers l'infini.

Etant donné une famille génératrice symétrique de S $\Sigma = \{e, s_1^{\pm 1}, \ldots s_k^{\pm 1}\}$, il est alors possible de définir une marche aléatoire symétrique $(S_n)_n$ sur $\mathcal{S}$. En partant d'un point $S_0 = h \in \mathcal{S}$, on tire uniformément un point $g \in \Sigma$ pour se déplacer en $S_1 = gh$ et ainsi de suite. Remarquons que si $\mathcal{S} = \mathbb{Z}^d$, cette procédure correspond bien à la marche aléatoire symétrique usuelle. Comme mentionné dans le chapitre précédent, nous pouvons construire un champ libre gaussien discret associé à cette marche aléatoire symétrique.

Théorème 3.2.7. *Soit $(X_s)_{s \in \mathcal{S}}$ un champ gaussien discret, indexé par $\mathcal{S}$ un groupe infini de type fini. Supposons que $\mathcal{S}$ admet une croissance de volume polynomial au moins quadratique, i.e. il existe $C > 0$ et $D > 2$ tels que*

$$V(n) \sim Cn^D,$$

pour tout $n \geq 1$. Alors, pour tout sous-ensemble $A_n \subset \mathcal{S}$ avec $\mathrm{Card}(A_n) = n^q$, avec $q \geq 1$, nous avons la convergence en loi suivante

$$a_n(M_{A_n} - b_n) \to \Lambda_0, n \to \infty,$$

où $a_n = \sqrt{\log \mathrm{Card}(A_n)}$ et b_n correspondant aux mêmes constantes de renormalisation que dans le cas gaussien standard (voir théorème 3.4) où l'on aurait remplacé n par $\mathrm{Card}(A_n)$ et $M_{A_n} = \max_{i \in A_n} X_i$.

Remarque. D'après [138] (théorème 2.5), les hypothèses de croissance de volume entraine l'existence d'une fonction $\phi : \mathbb{N} \to \mathbb{R}_+$ décroissante telle que

$$|\mathrm{Cov}(X_i, X_j)| \leq \phi\big(d(i, j)\big), \quad i, j \in \mathcal{S}$$

où d désigne la distance induite par une partie génératrice Σ de $\mathcal{S}$. De plus, $\phi(n) = o(1/\log n)$ lorsque $n \to \infty$.

Comme nous le verrons durant la démonstration, la même conclusion reste vérifiée si l'on suppose que $Cn^D \leq V(n) \leq Ce^n$, avec $D > 3$ et $\phi(n) = o(1/n)$.

Nous présentons, ci-dessous, une adaptation d'un résultat de comparaison pour des vecteurs gaussiens (cf. [108] p.84).

Proposition 3.2.8. *Soient A un ensemble fini et $(X_i)_{i\in A}$ un vecteur gaussien de covariance $\Gamma_{ij} = \mathrm{Cov}(X_i, X_j)$, $i, j \in \{1, \dots, n\}$ tel que $\delta = \max_{i\neq j} |\Gamma_{ij}| < 1$. Alors, pour tout $u \in \mathbb{R}$ et tout s-uplet $(i_1, \dots, i_s) \in A^s$, $\quad 1 \leq s \leq \mathrm{Card}(A)$,*

$$|\mathbb{P}(X_{i_l} \leq u, \quad l = 1, \dots s) - \Phi(u)^s| \leq K \sum_{i\neq j} |\Gamma_{ij}| \exp\left(-\frac{u^2}{1 + |\Gamma_{ij}|}\right),$$

où K est une constante numérique ne dépendant que de δ.

Remarque. Il est important que $\mathrm{Card}(A) < \infty$. Il suffit alors de réindexer A pour retrouver le corollaire de [108]. Nous appliquerons la proposition 3.2.8 pour $s = \mathrm{Card}(A)$, puis utiliserons le fait que, sous ces hypothèses de croissance, la fonction de covariance est dominée par ϕ d'après [138].

Démonstration. (du théorème 3.2.7) Nous allons suivre la démonstration de [108] dans le cas stationnaire. Soit $x \in \mathbb{R}$ et posons $u_{A_n} = u_n = x/a_n + b_n$. On a le résultat suivant

$$|\mathbb{P}(M_{A_n} \leq u_n) - \mathbb{P}(\Lambda_0 \leq x)| \leq |\mathbb{P}(M_{A_n} \leq u_n) - \Phi(u_n)^{\mathrm{Card}(A_n)}| + |\Phi(u_n)^{\mathrm{Card}(A_n)} - \mathbb{P}(\Lambda_0 \leq x)|,$$

avec Φ la fonction de répartition d'une gaussienne standard. Ainsi, il suffit de montrer que $|\mathbb{P}(M_{A_n} \leq u_n) - \Phi(u_n)^{\mathrm{Card}(A_n)}| \to 0$ lorsque $n \to \infty$ puisque l'on sait déjà, d'après le cas gaussien 3.4, que le deuxième terme tend vers zéro lorsque $n \to \infty$. L'utilisation de la proposition 3.2.8 (cf. [108] page 84) entraine

$$|\mathbb{P}(M_{A_n} \leq u_n) - \Phi(u_n)^{\mathrm{Card}(A_n)}| \leq K\mathrm{Card}(A_n) \sum_{i\in A_n} \big|\phi(d(e,i))\big| \exp\left(-\frac{u_{A_n}^2}{1 + \big|\phi(d(e,i))\big|}\right).$$

On pose $\delta = \sup_{i\in A_n\,;\,d(e,i)>0} \phi\big(d(e,i)\big)$ et l'on choisit η tel que $0 < \eta < \frac{1-\delta}{Dq(1+\delta)}$, où D est l'exposant apparaissant dans la condition de croissance de volume et q l'exposant de $\mathrm{Card}(A_n) = n^q$. Nous utiliserons les asymptotiques suivantes dans la démonstration (cf. [108] page 84)

$$u_n \underset{n\to\infty}{\sim} \sqrt{\log \mathrm{Card}(A_n)}, \text{et} \quad e^{-u_n^2/2} \underset{n\to\infty}{\sim} K u_n/\mathrm{Card}(A_n),$$

avec $K > 0$.

Séparons la somme en deux parties, suivant que $d(e,i) \leq n^\eta$ ou non. D'une part

$$\begin{aligned}
\Sigma_1 &= K\mathrm{Card}(A_n) \sum_{i\in A_n; d(e,i)\le n^\eta} \big|\phi(d(e,i))\big| \exp\left(-\frac{u_n^2}{1+\big|\phi(d(e,i))\big|}\right)\\
&\le K\mathrm{Card}(A_n)\exp\left(-\frac{u_n^2}{1+\delta}\right) \sum_{i\in A_n, d(e,i)\le \mathrm{Card}(A_n)^\eta} 1\\
&\le K\mathrm{Card}(A_n)^{1+D\eta}\left(e^{-u_n^2/2}\right)^{2/1+\delta}.
\end{aligned}$$

Notons que la dernière majoration obtenue est équivalente à la quantité suivante, lorsque $n\to\infty$,

$$K\mathrm{Card}(A_n)^{1+D\eta}\mathrm{Card}(A_n)^{-2/1+\delta}(\log\mathrm{Card}(A_n))^{1/1+\delta}$$

et celle-ci tend vers zéro lorsque $n\to\infty$ d'après le choix de η. Pour cette majoration, nous avons uniquement utilisé la propriété de croissance de volume pour contrôler $\sum_{i\in A, d(e,i)\le \mathrm{Card}(A)^\eta} 1$ par $\mathrm{Card}(A_n)^{1+D\eta}$, ainsi que les asymptotiques de u_n décrits plus haut.

Pour la seconde partie de la somme, nous utiliserons le fait que ϕ est décroissante pour $n\to\infty$, $\mathrm{Card}(A_n)=n^q$, et les asymptotiques de u_n.

$$\begin{aligned}
\Sigma_2 = K \quad & \mathrm{Card}(A_n) \sum_{i\in A_n\,;\, d(e,i)\ge \mathrm{Card}(A_n)^\eta} \big|\phi(d(e,i))\big| \exp\left(-\frac{u_n^2}{1+\big|\phi(d(e,i))\big|}\right)\\
\le \quad & K\mathrm{Card}(A_n)|\phi(n^\eta)| \sum_{i\in A_n, d(e,i)\ge \mathrm{Card}(A_n)^\eta} e^{-u_n^2}\exp\left(\frac{u_n^2|\phi(d(e,i))|}{1+|\phi(d(e,i))|}\right)\\
\le \quad & K\mathrm{Card}(A_n)^2|\phi(\mathrm{Card}(A_n)^\eta)|e^{-u_n^2}e^{u_n^2|\phi(\mathrm{Card}(A_n)^\eta)|}.
\end{aligned}$$

Notons que la dernière majoration obtenue est équivalente à la quantité suivante, lorsque $n\to\infty$,

$$K\big|\phi\big(\mathrm{Card}(A_n)^\eta\big)\big|u_n^2 e^{u_n^2|\phi(\mathrm{Card}(A_n)^\eta)|}$$

cette quantité tend vers zéro lorsque $n\to\infty$ puisque

$$\phi(\mathrm{Card}(A_n)^\eta)u_n^2 \underset{n\to\infty}{\sim} K\phi(\mathrm{Card}(A_n)^\eta)\log(\mathrm{Card}(A_n)^\eta) = \phi(n^{q\eta})\log(n^{q\eta}).$$

$\square$

Remarque. Pour une condition de croissance de volume du type $Cn^D \le V(n) \le Ce^n$ avec $\phi(n)=o(1/n)$ il suffirait de partager la somme selon $\log n^\eta$ with $0<\eta<\frac{1-\delta}{1+\delta}$.

Voici une liste de groupe $\mathcal{S}$ satisfaisant les hypothèses du théorème 3.2.7 (cf. [138]).

1. Si $\mathcal{S}=\mathbb{Z}^d$, $d\ge 3$ alors $V(n)\sim Cn^D$ pour $D=d\ge 3$ et

$$\mathrm{Cov}(X_i,X_j)\le \frac{C}{d(i,j)^{D-2}},\quad i,j\in\mathcal{S}$$

donc $\phi(n)=C/n^{D-2}=o(1/\log n)$ lorsque $n\to\infty$.

2. Si $\mathcal{S} = \text{Heis} = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \middle| a, b, c \in \mathbb{R} \right\}$ le groupe d'Heiseinberg, nous avons les mêmes résultats que dans l'item précédent avec $D = 4$.

3. Si $\mathcal{S}$ désigne le groupe des matrices triangulaires supérieures ayant des 1 sur la diagonale et $a_{i,j} \in \mathbb{Z}$ pour $i < j$ ($a_{ij} = 0$ if $i > j$) au dessus. Pour ce groupe $V(n) \sim n^D$ avec D suffisament grand.

4. Si $\mathcal{S} = \mathbb{F}_2$ le groupe libre à deux éléments, alors $V(n) \sim 2^n$ et l'on peut montrer que $Cn^D \leq V(n) \leq Ce^n$ avec $D > 3$ afin d'utiliser un résultat de Saloff-Coste ([138]) et obtenir

$$\mathrm{Cov}(X_i, X_j) \leq \frac{C}{d(i,j)^{D-2}}, \qquad i, j \in \mathcal{S}$$

qui fournit $\phi(n) = C/n^{D-2} = o(1/n)$ puisque $D > 3$.

5. Dans un cadre plus abstrait, on pourrait également considérer un groupe de Lie simple, connexe et nilpotent ou bien un groupe nilpotent de type fini pour lequel on peut estimer l'exposant de croissance de volume D. En effet, il se trouve qu'il existe une formule explicite, mais compliquée, pour la condition de croissance de volume [138].

3.3 Théorème de Darling-Erdös

Nous présentons un dernier résultat provenant de la théorie des extrêmes. Celui-ci est intéressant puisqu'il fait intervenir le maximum d'une somme partielles de variables *i.i.d.* au lieu du simple maximum d'un échantillon. Nous énonçons le théorème de Darling-Erdös [79] dans le cas gaussien, bien que celui-ci soit valide pour des lois satisfaisants une condition d'intégrabilité (cf. [77, 23, 100]).

Théorème 3.3.1. *Soient $(X_i)_{i\geq 0}$ des variables gaussiennes standards indépendantes, de mêmes loi. On définit, pour tout $k \in \mathbb{N}$, $S_k = X_1 + \cdots + X_k$ et l'on pose $Z_i = \frac{S_i}{\sqrt{i}}$ pour $i = 1, \ldots, n$. $Z = (Z_1, \ldots, Z_n)$ est un vecteur gaussien centré dans $\mathbb{R}^n$, admettant la structure de covariance suivante*

$$\mathrm{Cov}(Z_i, Z_j) = \frac{\min(i,j)}{\sqrt{ij}}, \quad i, j = 1 \ldots, n$$

Alors, nous avons le résultat de convergence en loi

$$a_n(M_n - b_n) \to \Lambda_0,$$

lorsque $n \to \infty$, avec $M_n = \max_{i=1,\ldots,n} Z_i$, $a_n = \sqrt{2 \log\log n}$ et $b_n = a_n + \frac{1/2 \log\log\log n - \log 2\sqrt{\pi}}{a_n}$.

Remarque. Nous obtiendrons une inégalité de concentration exponentielle reflétant ce résultat de convergence dans le chapitre suivante.

En fait, ce résultat reste vrai si l'on remplace S_k par un mouvement brownien. C'est à dire

Théorème 3.3.2. *Soit* $(B_t)_{t\geq 0}$ *un mouvement brownien standard. Pour* $n > 1$ *on définit*

$$M_n = \sup_{t\in[1,n]} \frac{B_t}{\sqrt{t}}$$

Alors, lorsque $n \to \infty$

$$a_n(M_n - b_n) \to \Lambda_0,$$

en loi, avec les mêmes constantes de renormalisation que dans le théorème précédent.

Chapitre 4

Approche hypercontractive, renforcement des résultats de Chatterjee

Cette partie concerne les principaux résultats obtenus au cours de cette thèse, via l'approche hypercontractive, concernant le phénomène de superconcentration. L'inégalité de Talagrand est l'outil fréquemment utilisé en superconcentration comme nous l'avons vu dans le chapitre deux. Nous l'utiliserons à nouveau dans cette section pour traiter des variations du cas gaussien indépendant. Ces résultats nous permettrons de comprendre jusqu'à quel point l'on peut perturber la matrice de covariance d'un vecteur gaussien standard et conserver de la superconcentration. Ensuite, nous présenterons un théorème abstrait, correspondant à une extension au niveau exponentiel du théorème 2.3.3 de Chatterjee sur la variance du maximum. Ce résultat fut au coeur de notre article [157]. Nous illustrerons notre théorème sur une série d'exemples. Les suites et processus gaussiens stationnaires seront tout particulièrement étudiés. Une illustration statistique sera également fournie, justifiant à nouveau l'utilité des inégalités de superconcentration. Enfin, nous conclurons ce chapitre en abordant la superconcentration pour des mesures non gaussiennes.

4.1 Utilisation de l'inégalité de Talagrand

Rappelons que l'inégalité de Poincaré entraine que $\mathrm{Var}(M_n) \leq 1$, les inégalités de superconcentration sont donc pertinentes si la borne obtenue est plus petite que 1.

4.1.1 Vecteur gaussien échangeable

Que peut-on dire du maximum des coordonnées d'un vecteur gaussien centré échangeable $(Z_1, \ldots, Z_n)$? Il s'agit d'un vecteur gaussien centré tel que $\mathbb{E}[Z_i Z_j] = \rho, \forall i \neq j = 1, \ldots, n$, avec $\rho \in]-1, 1[$. Autrement dit, la matrice de covariance Γ_ρ du vecteur $(Z_1, \ldots, Z_n)$ s'exprime comme suit

$$\Gamma_\rho = \begin{pmatrix} 1 & \rho & \dots & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \rho \\ \rho & \dots & \dots & \rho & 1 \end{pmatrix} \tag{4.1}$$

Nous supposerons également que $\mathbb{E}[Z_i^2] = 1$ pour tout $i = 1, \dots, n$ et, à nouveau, le maximum des Z_i sera noté par $M_n = \max_{i=1,\dots,n} Z_i$.

Proposition 4.1.1. *Pour un tel vecteur gaussien, l'inégalité suivante est satisfaite, pour $n \geq 2$,*

$$|\rho| \leq \mathrm{Var}(M_n(\rho)) \leq \frac{C}{\log n} + |\rho|$$

Démonstration. La preuve repose sur le lemme de Berman's (cf. [108]). Notons par $M_n(\rho)$ le maximum des Z_i lorsque $\mathbb{E}[Z_i Z_j] = \rho, \forall i, j = 1, \dots, n$. Avec ces notations, $M_n(0)$ correspond au maximum de gaussiennes standards indépendantes. Le lemme de Berman peut s'énoncer comme suit :

Lemme 4.1.2. *(Berman)*

$$M_n(\rho) = (1-\rho)^{1/2} M_n(0) + \rho^{1/2} X, \tag{4.2}$$

en loi, avec X une gaussienne standard indépendante du reste.

On obtient donc $\mathrm{Var}(M_n(\rho) = (1-\rho)\mathrm{Var}(M_n(0)) + |\rho|$. Enfin, puisque $\mathrm{Var}(M_n(0)) \leq \frac{C}{\log n}$ d'après le théorème 1.3.6, la preuve est achevée. □

Remarque. Le phénomène de superconcentration a donc lieu si $|\rho| < 1$ puisque la borne précédente améliore celle de Poincaré. C'est notamment le cas si $\rho = \rho_n = o(1/\log n)$ lorsque $n \to \infty$, auquel cas $M_n(\rho)$ présente le même comportement que le cas gaussien indépendant.

4.1.2 Perturbation d'un vecteur gaussien standard

Que se passe-t-il si l'on adjoint à un vecteur gaussien standard dans $\mathbb{R}^{n-k}, k \geq 2$ un vecteur gaussien indépendant de $\mathbb{R}^k$ admettant des corrélations ? Plus précisément, soient Y un vecteur gaussien de $\mathbb{R}^k, k \geq 2$ de matrice de covariance Γ_k et X un vecteur gaussien standard dans $\mathbb{R}^{n-k}$ indépendant de Y. Alors $Z = (X, Y)$ est un vecteur gaussien standard dans $\mathbb{R}^n$. Que peut-t-on dire de $\mathrm{Var}(\max_{i=1,\dots,n} Z_i)$?

Proposition 4.1.3. *En conservant le cadre précédent (rappelons que $\Gamma_{i,i} = 1$ pour tout $i = 1, \dots, n$), nous avons l'inégalité suivante de satisfaite, pour tout $k \geq 2$ et $n > k + 1$,*

$$\mathrm{Var}\Big(\max_{i=1,\dots,n} Z_i\Big) \leq C\left(\frac{k}{n-k} + \frac{1}{\log(n-k)}\right),$$

avec $C > 0$ une constante numérique.

Démonstration. Soit $M = (A_{i,j})_{1 \leq i,j \leq k}$ telle que $M^t M = \Gamma_k$, la matrice de covariance Γ de Z est donc diagonale par bloc et s'exprime comme suit

$$\Gamma = \begin{pmatrix} Id_{n-k} & 0 \\ 0 & \Gamma_k \end{pmatrix} \tag{4.3}$$

Par abus de notation, on désignera également par $X = (X_1, \dots, X_n)$ un vecteur gaussien standard de $\mathbb{R}^n$ et par M la matrice satisfaisant $M^t M = \Gamma$. Il est bien connu que $Z = MX$ en loi, ce qui se retranscrit de manière équivalente par

$$\begin{cases} Z_i = X_i, \quad 1 \le i \le n-k \\ Z_i = \sum_{j=1}^k M_{i-(n-k),j} X_{n-k+j}, \quad n-k+1 \le i \le n \end{cases} \tag{4.4}$$

Comme mentionné précédemment, l'inégalité de Talagrand 1.3.6 satisfait des propriétés de monotonie. Il suffit donc de majorer les normes L^p, $p = 1, 2$ des dérivées partielles de $F(x) = \max_{i=1,\dots,n} x_i$ pour démontrer notre résultat.

Puisque $F(Z) = \max_{i=1,\dots,n} Z_i = \sum_{i=1}^{n-k} Z_i 1_{A_i} + \sum_{i=n-k+1}^{n} Z_i 1_{A_i}$ avec $A_i = \{Z_i = \max_{j=1,\dots,n} Z_j\}$. Pour $i \le n-k$, d'après (4.4), on trouve $\partial_i F = 1_{A_i}$. Ainsi,

$$\|\partial_i F\|_1 = \|\partial_i F\|_2^2 \le \frac{1}{n-k},$$

puisque

$$\begin{aligned} \mathbb{P}(A_i) = \mathbb{P}(Z_i \ge Z_j, \forall j) &\le \mathbb{P}(Z_i \ge Z_j \,;\, j \ne i,\, 1 \le j \le n-k) \\ &= \mathbb{P}(X_i \ge X_j \,;\, j \ne i,\, 1 \le j \le n-k). \end{aligned}$$

Ceci permet d'obtenir la majoration désirée, puisque $X = (X_1, \dots, X_{n-k})$ est un vecteur gaussien standard de $\mathbb{R}^{n-k}$ donc, par symétrie,

$$\mathbb{P}(X_i \ge X_j \,;\, j \ne i,\, 1 \le j \le n-k) = \frac{1}{n-k}.$$

Il reste donc à majorer $\partial_i F(Z) = \sum_{l=n-k+1}^{n} M_{l-(n-k), i-(n-k)} 1_{A_l}$ pour $n-k+1 \le i \le n$. D'une part, nous avons (après changement d'indice)

$$\begin{aligned} \|\partial_i F\|_1 &\le \sum_{j=1}^{k} |M_{j,i-(n-k)}| \mathbb{P}(A_{j+(n-k)}) \\ &\le \min\left(1, \frac{k}{n-k}\right). \end{aligned}$$

En effet, pour n'importe quel $m \in \{1, \dots, k\}$, $\sum_{l=1}^k M_{lm}^2 = \Gamma_{mm} = 1$. Donc, pour n'importe quel $m, l \in \{1, \dots, k\}$, $M_{lm}^2 \le 1$ et $|M_{lm}| \le 1$. De plus, $(A_l)_{l=1,\dots,n}$ est une partition de $\mathbb{R}^n$ ainsi, $\sum_{l=1}^n \mathbb{P}(A_l) = 1$. Ceci entraine que $\sum_{l=1}^m \mathbb{P}(A_l) \le 1$ pour tout $m \le n$.

D'autre part, en utilisant le même type d'arguments que pour les deux majorations précédentes,

$$\|\partial_i F\|_2^2 = \sum_{l=1}^{k} M_{l,i-(n-k)}^2 \mathbb{P}(A_{l+(n-k)}) \le \left(\sum_{l=1}^{k} M_{l,i-(n-k)}^2\right) \max_{l=1,\dots,k} \mathbb{P}(A_{l+(n-k)}) \le \frac{1}{n-k}$$

Finalement,

$$\begin{aligned}\mathrm{Var}(M_n) &\leq C\Big(\sum_{i=1}^{k}\frac{\|\partial_i F\|_2^2}{1+\log(\|\partial_i F\|_2/\|\partial_i F\|_1)}+\sum_{i=k+1}^{n}\frac{\|\partial_i F\|_2^2}{1+\log(\|\partial_i F\|_2/\|\partial_i F\|_1)}\Big)\\ &\leq C\Big(\frac{k}{n-k}+\frac{1}{\log(n-k)}\Big).\end{aligned}$$

□

Remarque. Cette inégalité montre jusqu'à quel point on peut perturber un vecteur gaussien standard et conserver de la superconcentration. En effet, notons que si $k = k_n = O(n)$ alors $\mathrm{Var}(M_n) \leq C$. Tandis que si $k_n = cn^{\eta}$, avec $0 < \eta < 1$ alors

$$\mathrm{Var}(M_n) \leq \frac{C}{n^{1-\eta}} + \frac{C}{\log n} \leq \frac{C}{\log n}.$$

4.1.3 Modèle de covariance par blocs

Que se passe-t-il si l'on construit un vecteur gaussien dans $\mathbb{R}^n$ à partir de vecteur gaussien multidimensionel ? Plus précisément, supposons que l'on puisse décomposer $\mathbb{R}^n = \oplus_{j=1}^{d}\mathbb{R}^{n_j}$. On considère ensuite, pour $1 \leq j \leq d$, d vecteurs gaussiens indépendants $Z_1, \ldots, Z_d$ tels que, pour tout $1 \leq j \leq d$, Z_j soit un vecteur gaussien de $\mathbb{R}^{n_j}$ de matrice de covariance $\Gamma_j \in \mathcal{M}_{n_j}(\mathbb{R})$. On construit alors le vecteur gaussien $Z = (Z_1, \ldots, Z_n)$ de matrice de covariance $\Gamma = \mathrm{diag}(\Gamma_{n_1}, \ldots, \Gamma_{n_d})$. Autrement dit, Γ est diagonale par bloc :

$$\Gamma = \begin{pmatrix} \Gamma_{n_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \Gamma_{n_d} \end{pmatrix} \tag{4.5}$$

Comme dans l'ensemble de ce manuscrit, sauf mention du contraire, les variables aléatoires gaussiennes sont suposées centrées et renormalisées de sorte que leurs variances soient égales à 1. Autrement dit, $\Gamma_{ii} = 1$ pour tout $i \in \{1, \ldots, n\}$. Soit M telle que $M^t M = \Gamma$, et par $(M^i)_{1\leq i\leq n}$ les colonnes de la matrice M. Comme auparavant X désigne un vecteur gaussien standard dans $\mathbb{R}^n$, on peut donc représenter (en loi) $Z = (Z_1, \ldots, Z_n)$ par $Z = \big(\langle M^1, X\rangle, \ldots, \langle M^n, X\rangle\big)$, où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire euclidien.

Proposition 4.1.4. *Soit $n \geq 2$, sous le cadre précédent, nous obtenons la majoration suivante de la variance de $M_n = \max_{i=1,\ldots,n} Z_i$, pour $d \geq 1$,*

$$\mathrm{Var}(M_n) \leq \frac{1}{d}\sum_{j=1}^{d}\sum_{i=1}^{n_j}\frac{1}{1+\log\left(\frac{1}{\sqrt{d}\min(1,n_j/d)}\right)}$$

Démonstration. Pour tout $j = 1, \ldots, d$ et tout $i = 1, \ldots, n_j$, nous avons $\partial_i F(Z) = \sum_{l=1}^{n_j} M_{li} 1_{A_l}$, avec $A_l = \{X_l = \max_{k=1,\ldots,n} X_k\}$. Ainsi,

$$\begin{aligned}\|\partial_i F\|_1 &\leq \sum_{l=1}^{n_j} |A_{li}|\mathbb{P}(A_l)\\ &\leq \min(1, n_j/d)\end{aligned}$$

En effet, pour tout $l, i = 1, \dots, n$ $|M_{li}| \le 1$ et $\sum_{l=1}^{n_j} \mathbb{P}(A_l) \le 1$ puisque $(A_l)_{l=1,\dots,n}$ est une partition de $\mathbb{R}^n$. Nous avons également, pour tout $l \in \{1, \dots, n\}$ $\mathbb{P}(A_l) \le 1/d$, que l'on obtient, en ne conservant qu'une coordonnée des vecteurs indépendants, Z_j de $\mathbb{R}^{n_j}$ pour $j = 1, \dots, d$ lors de la majoration.

De plus, puisque $\sum_{l=1}^{n_j} |M_{li}|^2 = \Gamma_{ii} = 1$ pour tout $j = 1, \dots, d$ et tout $i = 1, \dots, n$.

$$\|\partial_i F\|_2^2 = \sum_{l=1}^{n_j} |M_{li}|^2 \mathbb{P}(A_l) \le 1/d$$

La conclusion s'ensuit via l'inégalité de Talagrand 1.3.6. □

Remarque. Remarquons que si $d = n$ alors $n_i = 1$ pour tout $i = 1, \dots, n$ et $\mathrm{Var}(M_n) \le C/\log n$. Tandis que si l'on considère des vecteurs gaussiens bi-dimensionnels, c'est à dire $d = n/2$ et $n_i = 2$ pour tout $i = 1, \dots, n$ la proposition précédente entraine que

$$\mathrm{Var}(M_n) \le C/\log n.$$

Cependant, si $d = 2$ et $n_i = n/2$ pour $i = 1, 2$, la borne obtenue par la proposition 4.1.4 n'est pas pertinente. Finalement, pour ce type de vecteurs gaussiens, nous conservons de la superconcentration si n_i est petit comparé à d.

4.1.4 Approximation régulière de la fonction maximum et comparaison de variance

Cette section rassemble quelques résultats d'approximation pour la fonction maximum et permet d'étendre les résultats obtenus sur la variance du maximum à d'autres fonctionnelles comme l'énergie libre pour une certaine échelle de température.

Remarquons que pour tout $n \ge 1$ et tout $\beta > 0$ fixés nous avons

$$\|F_\beta - \max_{i=1,\dots n} x_i\|_\infty \le \frac{\log n}{\beta},$$

avec $F_\beta(x) = \frac{1}{\beta} \log\left(\sum_{i=1}^n e^{\beta x_i}\right)$ l'énergie libre. Ceci fournit le lemme de comparaison suivant :

Lemme 4.1.5. *Pour tout $n \ge 1$ et $\beta > 0$, nous avons*

$$\mathrm{Var}(F_{n,\beta}) \le 3\mathrm{Var}(M_n) + \frac{6(\log n)^2}{\beta^2}$$

avec $F_{n,\beta} = F_\beta(X_1, \dots, X_n)$ et $M_n = \max_{i=1,\dots,n} X_i$.

Démonstration. En effet, en utilisant l'inégalité de Cauchy-Schwarz, nous avons

$$\mathrm{Var}(F_{n,\beta}) \le 3\Big(\mathrm{Var}(M_n) + \mathbb{E}[(F_{n,\beta} - M_n)^2] + \big(\mathbb{E}[F_{n,\beta} - M_n]\big)^2\Big).$$

□

Remarque. Il est facile de voir que l'on peut obtenir les mêmes inégalités en échangeant le rôle de $F_{n,\beta}$ et de M_n. Alors, si l'on obtient des bornes sur la variance de l'énergie libre ou du maximum, cela entraine que, mis à part une erreur de $(\log n)^2/\beta^2$, la même borne est valable pour l'autre fonctionnelle.

Les estimées suivantes pourraient également être utiles,

Lemme 4.1.6. *Soit X un vecteur gaussien standard dans $\mathbb{R}^n, n \geq 1$. Alors, pour tout $p \geq 1$, les inégalités suivantes sont satisfaites*

1. $\mathrm{Var}\big(\|X\|_p\big) \leq 3\mathrm{Var}(\max_{i=1,\dots,n}|X_i|) + 6\big(e^{\frac{1}{p}\log n} - 1\big)^2 \log n$.

2. $\mathrm{Var}\big(\max_{i=1,\dots,n}|X_i|\big) \leq 3\mathrm{Var}(\|X\|_p) + 6\big(e^{\frac{1}{p}\log n} - 1\big)^2 \log n$.

Démonstration. Notons que,

$$\big|\|x\|_p - \max_{i=1,\dots,n}|x_i|\big| \leq \left(e^{\frac{1}{p}\log n} - 1\right)\max_{i=1,\dots,n}|x_i|, \quad \forall x \in \mathbb{R}^n.$$

La conclusion s'ensuit aisément (à l'aide de l'inégalité de Cauchy-Schwarz) puisque, par des estimées gaussiennes classiques (cf. annexe [48]), nous avons $\mathbb{E}\big[(\max_{i=1,\dots,n}|X_i|)^2\big] \leq 2\log n$. □

Remarque. Ce genre d'estimées permet d'étendre les travaux de Paouris *et al.* sur le théorème de Dvoretzky aléatoire [131]. En effet, lorsque X est un vecteur gaussien standard dans $\mathbb{R}^n$, ces auteurs ont prouvé que, pour $p > (\log n)^2$, $n \geq 2$,

$$\mathrm{Var}\big(\|X\|_p\big) \leq \frac{C}{\log n}.$$

Puisque nous savons déja que $\mathrm{Var}(\max_{i=1,\dots,n}|X_i|) \leq \frac{C}{\log n}$, en utilisant le lemme précédent, nous obtenons une inégalité similaire à celle de Paouris *et al.*. Il est également possible d'étendre leurs travaux en ajoutant des corrélations. Au lieu de choisir un vecteur gaussien standard de $\mathbb{R}^n$, on considère X un vecteur gaussien de matrice de covariance Γ satisfaisant les mêmes types d'hypothèses que celles requises pour appliquer le théorème de Chatterjee 2.3.3, pour la variance, ou bien celui de Tanguy dans [157] au niveau exponentiel.

En revanche, il semblerait que ces bornes soient trop grossières pour être d'une quelconque pertinence pour la variance de l'énergie libre à basse température (β grand).

4.2 Inégalités de superconcentration

Le but de cette section est de quantifier les résultats asymptotiques, énoncés dans le chapitre précédent, en inégalité de concentration reflétant pleinement les fluctuations du maximum (respectivement le supremum). De telles bornes sur les variances avec le bon ordre de grandeur ont d'abord été obtenus par Chatterjee dans [48] (section 9.6) dans le cadre du phénomène de superconcentration. Les résultats présentés ici renforcent les bornes sur la variance de Chatterjee en inégalité de concentration exponentielle. Il s'agit d'une partie majeure des travaux obtenus durant cette thèse. Un exemple d'application est le théorème suivant.

Théorème 4.2.1. *Soit $(X_i)_{i\geq 0}$ une suite gaussienne centrée stationnaire de fonction de covariance ϕ. Supposons que ϕ est décroissante et satisfait $\phi(1) < 1$. Alors, il existe $\alpha = \alpha(\phi) \in (0,1)$ et $c = c(\phi, \alpha) > 0$ tel que pour tout $n \geq 2$,*

$$\mathbb{P}\left(|M_n - \mathbb{E}[M_n]| > t\right) \leq 6e^{-ct/\sqrt{\max(\phi(n^\alpha),1/\log n)}}, \quad t \geq 0. \tag{4.6}$$

Remarque. Sous les hypothèses du théorème 3.2.1, $\max\left(\phi(n^\alpha), 1/\log n\right) = 1/\log n$ pour n assez grand, ce qui correspond exactement au taux de fluctuaction. Remarquons également qu'en intégrant (4.6), on retrouve les bornes sur la variance de [48]. En particulier, dans le cadre du théorème 3.2.1, la variance est d'ordre $\frac{1}{\log n}$ ce qui est optimal (cf. corollaire 1.9 dans [69]). Une dernière observation est que, sous les hypothèses du théorème 3.2.1, $\phi(n) \to 0$ lorsque $n \to \infty$. Ceci implique que $\sup_{n\geq 1} |\phi(n)| < 1$ (cf. [108] p.86), dans ce cas particulier l'hypothèse $\phi(1) < 1$ peut-être enlevée.

Il est important de comparer le théorème 4.2.1 avec la concentration gaussienne classique (cf. [112]) qui produit, typiquement,

$$\mathbb{P}\left(|M_n - \mathbb{E}\left[M_n\right]| \geq t\right) \leq 2e^{-t^2/2}, \quad t \geq 0. \tag{4.7}$$

Bien que la décroissance soit gaussienne, de telles bornes ne reflètent pas les fluctuations des extrêmes M_n du théorème 4.2.1. De plus, par rapport à cette borne gaussienne, le théorème 4.2.1 fournit le bon ordre de grandeur pour les déviations à droite de la moyenne M_n sous la forme d'une inégalité de superconcentration en accord avec les résultats de fluctuations et la loi de Gumbel limite (puisque $\mathbb{P}(\Lambda_0 > t) \sim e^{-t}$ lorsque $t \to +\infty$). Pour les queues de déviation à gauche, il faudra utiliser une méthode de transport pour atteindre le bon comportement puisque $\mathbb{P}(\Lambda_0 \leq -t) \sim e^{-e^{-t}}$ lorsque $t \to -\infty$. Soulignons également le fait que le théorème 4.2.1 couvre le cas classique de gaussiennes standards, lorsque toutes les variables X_i sont indépendantes, en choisissant $\phi = 0$ et fournissant l'inégalité de concentration suivante qui sera utile pour une application en statistique,

$$\mathbb{P}\left(|M_n - \mathbb{E}\left[M_n\right]| \geq t\right) \leq 6e^{-ct\sqrt{\log n}}, \quad t \geq 0. \tag{4.8}$$

Remarquons également que le théorème 4.2.1 exprime une propriété de concentration du maximum autour de sa moyenne tandis que, dans le régime de convergence des extrêmes du théorème 3.4, les termes de recentrage sont produits par des valeurs explicites b_n. En fait, à des constantes numériques près, les mêmes inégalités sont satisfaites avec b_n au lieu de la moyenne. A cet effet, il suffit de prouver que $\sup_n \mathbb{E}\left[|a_n(M_n - b_n)|\right] < \infty$. On pose $Z_n = a_n(M_n - b_n)$. Soit M'_n une copie indépendante de M_n et posons similairement $Z'_n = a_n(M'_n - b_n)$. Or, en intégrant (4.8), $\sup_n \mathbb{E}\left[a_n|M_n - \mathbb{E}[M_n]|\right] < \infty$. Ainsi, $\sup_n \mathbb{E}\left[|Z_n - Z'_n|\right] < \infty$ ce qui entraine facilement que $\sup_n \mathbb{E}\left[|Z_n|\right] < \infty$. Ces remarques étant faites (elles sont valables pour chaques inégalités de superconcentration obtenues au cours de cette thèse), nous allons énoncer le résultat de superconcentration abstrait que nous appliquerons ensuite à des exemples provenant de la théorie des extrêmes.

4.2.1 Version exponentielle du Théorème de Chatterjee

La démonstration du prochain théorème est basée sur l'approche hypercontractive de Chatterjee pour obtenir des bornes sur la variance. La tâche principale de la preuve sera d'adapter ses arguments pour atteindre des inégalités de concentration exponentielle reflétant le régime correct de fluctuation. Ceci sera obtenu par des bornes sur la variance, similaires à celle de [48], au niveau des transformées de Laplace.

Pour tout $n \geq 1$, désignons par $X = (X_1, \ldots, X_n)$ un vecteur gaussien centré, avec le maximum de ses coordonnées $M_n = \max_{i=1,\ldots,n} X_i$. Posons $I = \mathrm{argmax}_{i=1,\ldots,n} X_i$.

La partie technique principale de ce travail est fournie par l'extension suivante de l'approche de Chatterjee [48] adaptée aux bornes de concentration exponentielle. La base de l'argument employé est l'hypercontractivité du semi groupe d'Ornstein-Uhlenbeck combiné avec la version exponentielle de l'inégalité de Poincaré (cf. [17, 16, 48]).

Théorème 4.2.2. *Soit $X = (X_1, \ldots, X_n)$ un vecteur gaussien centré de matrice de covariance Γ. Supposons que pour un $r_0 \geq 0$, il existe un recouvrement non trivial $\mathcal{C}(r_0)$ de $\{1, \ldots, n\}$ satisfaisant les conditions suivantes :*

— *pour tout $i, j \in \{1, \ldots, n\}$ tels que $\Gamma_{ij} \geq r_0$, il existe $D \in \mathcal{C}(r_0)$ tel que $i, j \in D$;*

— *il existe $C \geq 1$ tel que, p.s., $\sum_{D \in \mathcal{C}(r_0)} 1_{\{I \in D\}} \leq C$.*

Soit $\rho(r_0) = \max_{D \in \mathcal{C}(r_0)} \mathbb{P}(I \in D)$.
Alors, pour tout $\theta \in \mathbb{R}$,

$$\operatorname{Var}\Big(e^{\theta M_n/2}\Big) \leq C\,\frac{\theta^2}{4}\Big(r_0 + \frac{1}{\log\left(1/\rho(r_0)\right)}\Big)\mathbb{E}\left[e^{\theta M_n}\right]. \tag{4.9}$$

En particulier,

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| \geq t) \leq 6e^{-ct/\sqrt{K_{r_0}}}, \quad t \geq 0,$$

avec $K_{r_0} = \max\left(r_0, \frac{1}{\log(1/\rho(r_0))}\right)$ et $c > 0$ une constante numérique.

Remarque. Par convergence monotone, nous pouvons obtenir le même résultat pour le supremum à la place du maximum. Fournissant une version exponentielle du théorème 2.3.6 de Chatterjee. Ceci sera utile pour obtenir une application du théorème 4.2.2 pour des champs gaussiens euclidiens.

Démonstration. Comme annoncé, le schéma de preuve suit [48]. Le point de départ est la formule de représentation

$$\operatorname{Var}(f) = \int_0^\infty e^{-t}\mathbb{E}\left[\nabla f \cdot P_t \nabla f\right] dt \tag{4.10}$$

pour la variance d'une fonction (régulière) $f : \mathbb{R}^n \to \mathbb{R}$ le long du semi-groupe d'Ornstein-Uhlenbeck $(P_t)_{t\geq 0}$. Rappelons les ouvrages de référence [17, 16] concernant le semi groupe d'Ornstein-Uhlenbeck et des semi groupes de Markov plus généraux, ainsi que des méthodes d'interpolations.

En suivant [48], étant donné un vecteur gaussien centré X de matrice de covariance $\Gamma = (\Gamma_{ij})_{1\leq i,j\leq n}$, on applique (4.10) à (une approximation régulière) de $f = e^{\theta M_n/2}$ où $\theta \in \mathbb{R}, F(x) = \max_{i=1,\ldots,n}(Mx)_i$ et $\Gamma = M\,{}^tM$. Cela entraine, que pour tout $\theta \in \mathbb{R}$,

$$\operatorname{Var}\left(e^{\theta M_n/2}\right) = \frac{\theta^2}{4}\int_0^\infty e^{-t}\,\mathbb{E}\Big[\sum_{i,j=1}^n \Gamma_{ij}\,1_{\{X\in A_i\}}e^{\theta M_n/2}1_{\{X^t\in A_j\}}e^{\theta M_n^t/2}\Big]dt.$$

Ici $\{X \in A_i\} = \{I = i\}$, pour $i = 1, \ldots, n$, et $X^t = e^{-t}X + \sqrt{1-e^{-2t}}\,Y$, où Y est une copie indépendante de X, avec son maximum correspondant M_n^t. Rappelons brièvement que le semi-groupe $(Q_t)_{t\geq 0}$ d'Ornstein-Uhlenbeck généralisé satisfait une propriété d'hypercontractivité avec les mêmes constantes que le semi-groupe d'Ornstein-Uhlenbeck standard $(P_t)_{t\geq 0}$:

$$\mathbb{E}\left[|Q_t f(X)|^q\right]^{1/q} = \|P_t g\|_q \leq \|g\|_p = \mathbb{E}\left[|f(X)|^p\right]^{1/p}, \tag{4.11}$$

où $\|\cdot\|_r$, pour tout $r \geq 1$, désigne la norme L^r par rapport à la mesure gaussienne standard sur $\mathbb{R}^n$.
De plus, pour tout $t \geq 0$, avec $I^t = \mathrm{argmax}_{i=1,\dots,n} X_i^t$,

$$\begin{aligned}
\mathcal{I} &= \mathbb{E}\Big[\sum_{i,j=1}^n \Gamma_{ij}\, 1_{\{X\in A_i\}} e^{\theta M_n/2} 1_{\{X^t\in A_j\}} e^{\theta M_n^t/2}\Big] \\
&= \mathbb{E}\Big[e^{\theta(M_n+M_n^t)/2} \sum_{i,j=1}^n \Gamma_{ij}\, 1_{\{I=i\}} 1_{\{I^t=j\}}\Big] \\
&= \mathbb{E}\left[e^{\theta\left(M_n+M_n^t\right)/2}\, \Gamma_{II^t}\right] \\
&= \sum_{k\geq 0} \mathbb{E}\left[e^{\theta\left(M_n+M_n^t\right)/2}\, \Gamma_{II^t} 1_{\{2^{-k-1}\leq \Gamma_{II^t}\leq 2^{-k}\}}\right].
\end{aligned}$$

Pour alléger les notations, on notera Γ_{II^t} par Γ et on posera $F^t = M_n + M_n^t$. Soit $k_0 = \min\{k \geq 0\,; r_0 \leq 2^{-k-1}\}$. En coupant la somme précédente en deux, on obtient

$$\begin{aligned}
\mathcal{I} &\leq \sum_{k=0}^{k_0} 2^{-k}\, \mathbb{E}\left[e^{\theta F^t/2} 1_{\{\Gamma \geq r_0\}}\right] + \sum_{k=k_0+1}^{\infty} 2^{-k}\, \mathbb{E}\left[e^{\theta F^t/2} 1_{\{2^{-k-1}\leq \Gamma\leq 2^{-k}\}}\right] \\
&\leq 2 \sum_{D\in\mathcal{C}(r_0)} \mathbb{E}\left[e^{\theta F^t/2} 1_{\{I,I^t\in D\}}\right] + \sum_{k\geq 0} r_0\, \mathbb{E}\left[e^{\theta F^t/2} 1_{\{2^{-k-1}\leq \Gamma\leq 2^{-k}\}}\right] \\
&= 2 \sum_{D\in\mathcal{C}(r_0)} \mathbb{E}\left[e^{\theta F^t/2} 1_{\{I,I^t\in D\}}\right] + r_0\, \mathbb{E}\left[e^{\theta F^t/2}\right].
\end{aligned}$$

Puis, d'après l'inégalité de Cauchy-Schwarz et l'invariance rotationnelle gaussienne

$$\mathbb{E}\left[e^{\theta F^t/2}\right] \leq \mathbb{E}\left[e^{\theta M_n}\right].$$

D'autre part, d'après l'inégalité d'Hölder avec $\frac{1}{p} + \frac{1}{q} = 1$,

$$\begin{aligned}
\mathbb{E}\left[e^{\theta F^t/2} 1_{\{I,I^t\in D\}}\right] &= \mathbb{E}\left[e^{\theta M_n/2} 1_{\{I\in D\}}\, e^{\theta M_n^t/2} 1_{\{I^t\in D\}}\right] \\
&\leq \mathbb{E}\left[e^{\theta p M_n/2} 1_{\{I\in D\}}\right]^{1/p}\, \mathbb{E}\left[e^{\theta q M_n^t/2} 1_{\{I^t\in D\}}\right]^{1/q}.
\end{aligned}$$

Ensuite, d'après la propriété d'hypercontractivité (4.11),

$$\mathbb{E}\left[e^{\theta q M_n^t/2} 1_{\{I^t\in D\}}\right]^{1/q} \leq \mathbb{E}\left[e^{\theta p M_n/2} 1_{\{I\in D\}}\right]^{1/p}$$

à condition que $e^{2t} = \frac{q-1}{p-1}$, c'est à dire $p = 1 + e^{-t} < 2$. En conséquence,

$$\mathbb{E}\left[e^{\theta F^t/2} 1_{\{I,I^t\in D\}}\right] \leq \mathbb{E}\left[e^{\theta p M_n/2} 1_{\{I\in D\}}\right]^{2/p}$$

et une dernière utilisation de l'inégalité d'Hölder sur le terme de droite de l'inégalité précédente,

$$\mathbb{E}\left[e^{\theta F^t/2} 1_{\{I, I^t \in D\}}\right] \leq \mathbb{P}(I \in D)^{\frac{2-p}{p}} \, \mathbb{E}\left[e^{\theta M_n} 1_{\{I \in D\}}\right].$$

En combinant ce qui précède avec les hypothèses du théorème, nous obtenons que

$$\mathcal{I} \leq \left(r_0 + C\rho(r_0)^{\frac{2-p}{p}}\right) \mathbb{E}\left[e^{\theta M_n}\right].$$

Finalement,

$$\mathrm{Var}\left(e^{\theta M_n/2}\right) \leq C\,\frac{\theta^2}{4}\left(r_0 + \int_0^\infty e^{-t} \rho(r_0)^{\tanh(t/2)} dt\right) \mathbb{E}\left[e^{\theta M_n}\right].$$

L'inégalité annoncée (4.9) s'ensuit, puisque

$$\int_0^\infty e^{-t} \left(\rho(r_0)\right)^{\tanh(t/2)} dt \leq \frac{1}{\log\left(1/\rho(r_0)\right)}.$$

Pour achever la démonstration du théorème 4.2.2 et obtenir des inégalités de concentration exponentielles, nous combinons (4.9) avec le lemme suivant (cf. [112] p.50).

Lemme 4.2.3. *Soit Z une variable aléatoire et $K > 0$. Supposons que pour tout $|\theta| \leq 2/\sqrt{K}$,*

$$\mathrm{Var}\left(e^{\theta Z/2}\right) \leq K\,\frac{\theta^2}{4}\,\mathbb{E}\left[e^{\theta Z}\right]. \tag{4.12}$$

Alors,

$$\mathbb{P}\left(|Z - \mathbb{E}\left[Z\right]| > t\right) \leq 6e^{-ct/\sqrt{K}},$$

pour tout $t \geq 0$, avec $c > 0$ une constante numérique.

Remarque. Un tel résultat à été utilisé dans [21, 66, 65], pour des modèles de percolations, afin d'obtenir des inégalités de concentration exponentielles pour le premier temps de passage.

□

4.2.2 Superconcentration et argument de Herbst

Au vu des arguments précédents, il est naturel de se demander si la même méthodologie peut s'appliquer, au niveau exponentiel, avec l'entropie au lieu de la variance. Pour, ensuite, utiliser l'argument de Herbst. Plus précisément, que se passe-t-il si on trouve une suite $\epsilon_n \to 0$ lorsque $n \to \infty$ telle que

$$\mathrm{Ent}(e^{\lambda f}) \leq \epsilon_n \lambda^2 \mathbb{E}[e^{\lambda f}],$$

pour tout $\lambda \in I$, où $I \subset \mathbb{R}$ est un intervalle. Par exemple, si $I = [0, C]$, $C > 0$, alors l'argument de Herbst entraine que $\mathbb{E}[e^{\lambda f}] \leq e^{\epsilon_n \lambda^2}$, $\lambda \in I$. Ainsi, l'inégalité de Chernoff fournit

$$\mathbb{P}(f - \mathbb{E}[f] \geq t) \leq e^{-\lambda t + \epsilon_n \lambda^2}, \quad \lambda \in I$$

On optimise ensuite en $\lambda = \frac{t}{2\epsilon_n}$ si $t \leq 2C\epsilon_n$. Le problème dans cette optimisation est que $\epsilon_n \to 0$, donc la taille de l'intervalle $[0, 2C\epsilon_n]$ sur lequel nous aurions une décroissance gaussienne reflétant la taille de la variance tend vers 0. Finalement, pour n assez grand, on obtiendrait une inégalité de concentration comparable à l'inégalité 1.13. Ce qui n'apporte rien de pertinent vis-à-vis de la superconcentration de la mesure.

4.2.3 Illustrations du théorème 4.2.2

Nous présentons différentes applications du théorème 4.2.2, nous signalerons également lorsque les nouvelles inégalités de superconcentration fournissent des nouvelles bornes sur les variances.

Suite gaussienne stationnaire

Nous souhaitons démontrer un résultat annoncé plus tôt, nous redonnons son énoncé ci-dessous.

Théorème 4.2.4. *Soit $(X_i)_{i\geq 0}$ une suite gaussienne centrée stationnaire de fonction de covariance ϕ. Supposons que ϕ est décroissante et satisfait $\phi(1) < 1$. Alors, il existe $\alpha = \alpha(\phi) \in (0,1)$ et $c = c(\phi, \alpha) > 0$ tel que pour tout $n \geq 2$,*

$$\mathbb{P}\left(|M_n - \mathbb{E}[M_n]| > t\right) \leq 6e^{-ct/\sqrt{\max(\phi(n^\alpha), 1/\log n)}}, \quad t \geq 0. \tag{4.13}$$

Démonstration. Nous devons montrer que les hypothèses du théorème 4.2.2 sont satisfaites. Pour $0 < \alpha < 1$, on choisit $r_0 = \phi(\lfloor n^\alpha \rfloor)$ et $\mathcal{C}(r_0) = \{D_1, D_2, \ldots\}$ où $D_1 = \{1, \ldots, 2\lfloor n^\alpha \rfloor\}$, $D_2 = \{\lfloor n^\alpha \rfloor, \ldots, 3\lfloor n^\alpha \rfloor\}$ et ainsi de suite, avec $\lfloor \cdot \rfloor$ pour désigner la partie entière d'un nombre réel.

Il est facilement visible qu'un tel recouvrement vérifie les hypothèses du théorème 4.2.2. En effet, soit $i, j \in \{1, \ldots, n\}$ tels que $i < j$. Par définition de la fonction de covariance, $\mathrm{Cov}(X_{i+1}, X_j) = \phi(j-i)$. Donc si $\phi(j-i) \geq r_0 = \phi(\lfloor n^\alpha \rfloor)$, puisque ϕ est décroissante, nous devons avoir $j - i \leq \lfloor n^\alpha \rfloor$. Par construction, chaque indice est contenu dans, au plus, trois ensembles $D \in \mathcal{C}(r_0)$ différents. Donc i appartient à des ensembles D_{i_1}, D_{i_2} et D_{i_3} avec $i_1 < i_2 < i_3$ et j appartient à D_{j_1}, D_{j_2} et D_{j_3} avec $j_1 < j_2 < j_3$. Puisque $j - i \leq n^\alpha$ et que la longueur de n'importe quel ensemble D dans $\mathcal{C}(r_0)$ est $2\lfloor n^\alpha \rfloor$, il existe $s \in \{1,2,3\}$ tel que $D_{i_s} = D_{j_s}$ (faire un dessin pour s'en convaincre) et l'on peut choisir $D = D_{i_s} \in \mathcal{C}(r_0)$.

Il est clair que $\sum_{D \in \mathcal{C}(r_0)} 1_{\{I \in D\}} \leq C$, avec C une constante numérique ($C = 3$ par exemple). Pour conclure la preuve, nous devons majorer $\max_{D \in \mathcal{C}(r_0)} \mathbb{P}(I \in D)$. Soit $D \in \mathcal{C}(r_0)$, et pour simplifier les notations, considérons $D = \{1, \ldots, 2\lfloor n^\alpha \rfloor\}$ (la propriété importante est que D contient $2\lfloor n^\alpha \rfloor$ éléments). Alors

$$\mathbb{P}(I \in D) = \sum_{i=1}^{2\lfloor n^\alpha \rfloor} \mathbb{P}(I = i).$$

La concentration gaussienne standard entraine que (cf. [112] ou l'appendice de [48]), pour tout $i \in D$,

$$\begin{aligned} \mathbb{P}(I = i) &= \mathbb{P}(X_i = M_n) \\ &\leq \mathbb{P}(X_i \geq t) + \mathbb{P}(M_n \leq t) \\ &\leq 2e^{-\mathbb{E}[M_n]^2/2}. \end{aligned}$$

La dernière étape de l'argumentation est d'obtenir une minoration de $\mathbb{E}[M_n]$. Pour ce faire, nous utilisons la minoration de Sudakov (cf. [4] ou [117]). Soient $i, j \in D$, $i \neq j$,

$$\mathbb{E}\left[(X_i - X_j)^2\right] = 2 - 2\phi\big(|j-i|\big) \geq 2\big(1 - \phi(1)\big) = \delta$$

puisque $\phi\big(|j-i|\big) \leq \phi(1)$. Par hypothèses sur $\phi(1)$, $\delta > 0$, donc d'après la minoration de Sudakov il existe $c > 0$ (indépendant de n) telle que $\mathbb{E}[M_n] \geq c\delta\sqrt{\log n}$. En conséquence de ce qui précède,

$$\mathbb{P}(I=i) \leq \frac{2}{n^{(c\delta)^2/2}} = \frac{2}{n^{\epsilon}}\,.$$

D'où

$$\mathbb{P}(I \in D) \leq \frac{4}{n^{\epsilon-\alpha}} = \frac{4}{n^{\eta}}$$

où α est telle que $\eta = \epsilon - \alpha > 0$, c'est à dire $\alpha < (c\delta)^2/2 < 1$.

Les hypothèses du théorème 4.2.2 sont donc satisfaites avec $\rho(r_0) \leq 4/n^{\eta}$ concluant la démonstration du théorème 4.2.4. □

Remarque. Notons le fait suivant, la démonstration précédente reste valable (à des modifications mineures près) si l'on suppose que la fonction de covariance admet un comportement différent en l'infini. Ce qui permet d'obtenir les majorations suivantes, via le théorème de Chatterjee, ou en intégrant l'inégalité de superconcentration obtenue via le théorème 4.2.2, si jamais $\phi(n) = O(1/\log n)$ alors $\mathrm{Var}(M_n) \leq \frac{C}{\log n}$ ce qui correspond au théorème des extrêmes 3.2.2. Tandis que si $\phi(n)\log(n) \to \infty$ lorsque $n \to \infty$, comme dans le théorème 3.2.2 (version faible), alors $\mathrm{Var}(M_n) \leq \frac{C}{(\log n)^{\eta}}$, $0 < \eta < 1$.

Au niveau des inégalités de superconcentration, le cas $\phi(n) = O(1/\log n)$ est satisfaisant puisque : $\mathbb{P}\Big(N\big(]0,1]\big) = 0 \geq t\Big) \leq e^{-t+2\eta}$, $t \geq 0$. Tandis que lorsque $\phi(n) \sim_{n\to\infty} \frac{C}{(\log n)^{\eta}}$, $0 < \eta < 1$ la loi limite est gaussienne et la décroissance exponentielle ne convient pas. Nous verrons un peu plus loin comment obtenir le bon comportement.

Processus gaussien stationnaire

L'énoncé suivant est l'analogue du théorème 4.2.4 pour les processus gaussiens stationnaires, celui-ci quantifie correctement le théorème de convergence 3.2.4. Nous présentons cet énoncé dans un contexte plus abstrait, celui d'un champ gaussien stationnaire $(X_t)_{t\in\mathbb{R}^d}$ tel que $\mathbb{E}\left[X_t^2\right] = 1$, $t \in \mathbb{R}^d$, de fonction de covariance ϕ. D'après [48], très peu de choses semblent connues concernant les fluctuations du supremum d'un champ gaussien indexé par $\mathbb{R}^d$ lorsque la dimension d est plus grande que deux. Toutefois, des résultats spécifiques ont été obtenu pour des champ gaussien admettant des structures de corrélations logarithmiques (cf. , par exemple, [2, 73, 122, 70] et [74] pour un survol de ce sujet). Cependant, étendant de manière similaire les bornes sur la variance de [48], nous obtenons une inégalité de concentration pour le supremum, sur un sous ensemble A de $\mathbb{R}^d$, d'un champ gaussien. Le cas $d = 1$ correspond ainsi au théorème 3.2.4.

Théorème 4.2.5. *Soit $(X_t)_{t\in\mathbb{R}^d}$ un champ gaussien stationnaire euclidien, de fonction de covariance ϕ. Supposons que $t \mapsto \phi(t)$ est décroissante et que $\phi(1) < 1$. Si A est un sous ensemble de $\mathbb{R}^d$, désignons par $N(A) = N(A,1)$ le nombre minimal de boule de rayon 1 nécessaires pour recouvrir A. On pose $M(A) = \sup_{s\in A} X_s$. Alors, il existe $C = C(\phi,d) > 0$ et $c = c(\phi,d) > 0$ dépendant uniquement de ϕ et d telles que, pour tout $A \subset \mathbb{R}^d$ satisfaisant $N(A) > 1$,*

$$\mathbb{P}\left(|M_A - \mathbb{E}\left[M_A\right]| \geq t\right) \leq 6e^{-ct/\sqrt{K_{N(A)}}}, \quad t \geq 0,$$

où

$$K_{N(A)} = \max\left(\phi\left(N(A)^C\right), 1/\log N(A)\right).$$

Remarque. Lorsque $d = 1$ et $A = [0, T]$ pour $T > 0$, $N(A) = \frac{\lfloor T \rfloor}{2} + 1$ et sous les hypothèses du théorème 4.2.4, $K_{N(A)} = 1/\log(\frac{\lfloor T \rfloor}{2} + 1)$ pour T suffisament grand.

A nouveau, nous imitons la preuve de Chatterjee pour les bornes sur la variance de champs gaussiens euclidiens (cf. théorème 9.12 dans [48]). Pour tout ensemble borélien $A \subset \mathbb{R}^d$, on pose $m(A) = \mathbb{E}[M(A)] = \mathbb{E}[\sup_{s \in A} X_s]$. Sous les hypothèses du théorème, il est prouvé dans [48] que pour tout $A \subset \mathbb{R}^d$ tel que $N(A) > 1$,

$$c_1(\phi, d)\sqrt{\log N(A)} \leq m(A) \leq c_2(\phi, d)\sqrt{\log N(A)}$$

où $c_1 = c_1(\phi, d)$ and $c_2 = c_2(\phi, d)$ sont des constantes positives ne dépendant que la fonction de covariance ϕ et de la dimension d (et pas de l'ensemble A). On pose alors

$$s_0 = N(A)^{\frac{1}{8}(c_1/c_2)^2},$$

et supposons que $N(A)$ est suffisament grand de telle sorte que $s_0 > 2$. Soit $r_0 = \phi(s_0)$. Choisisons un réseau maximal de rayon s_0 of A (*i.e.* un ensemble de points qui sont mutuellement séparés les uns des autres par une distance strictement plus grande que s_0 et qui est maximal vis à vis de cette propriété), et considérons $\mathcal{C}(r_0)$ comme étant l'ensemble des boules de rayon $2s_0$ centrées en les points du réseau. En fait, $\mathcal{C}(r_0)$ est un recouvrement de A vérifiant les conditions du théorème 4.2.2. Par construction, il est facilement visible que $\sum_{D \in \mathcal{C}(r_0)} 1_{\{I \in D\}} \leq C$ pour $C \geq 1$ et

$$\max_{D \in \mathcal{C}(r_0)} \mathbb{P}(I \in D) \leq \frac{1}{N(A)^{C(\phi,d)}}.$$

La conclusion s'ensuit en appliquant le théorème 4.2.2 avec $\rho(r_0) \leq 1/N(A)^{C(\phi,d)}$.

Autres modèles gaussiens

Dans l'ouvrage [48], Chatterjee exhibe différents modèles gaussiens pour illustrer le phénomène de superconcentration. Pour chacun d'entre eux, il parvient à atteindre une meilleure borne sur la variance que celle fournie par la concentration de gaussiennes standards. Pour certains de ces exemples, nous avons obtenus une version exponentielle des résultats de Chatterjee. Tout d'abord au niveau de la proposition 2.3.4 et de la proposition 4.1.1.

Proposition 4.2.6. *Pour tout $n \geq 2$, soit $X = (X_1, \ldots, X_n)$ un vecteur gaussien centré tel que $\mathbb{E}[X_i]^2 = 1$ et $\mathbb{E}[X_i X_j] \leq \epsilon$ pour tout $i \neq j = 1, \ldots, n$ et $\epsilon \in]0, 1[$. Alors*

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| \geq t) \leq 6e^{-ct/\sqrt{K_n}}, \quad t \geq 0,$$

où $K_n = \max(\epsilon, 1/\log n)$ et $c > 0$ est une constante numérique.

Démonstration. soit $r_0 > \epsilon$ et posons $\mathcal{C}(r_0) = \{\{1\}, \ldots, \{N\}\}$. Il est facile de voir que le recouvrement $\mathcal{C}(r_0)$ satisfait les hypothèses du théorème 4.2.2. De plus, $\sum_{D \in \mathcal{C}(r_0)} 1_{\{I \in D\}} = 1$ et $\max_{D \in \mathcal{C}(r_0)} \mathbb{P}(I \in D) \leq 1/n^\eta$ pour $\eta > 0$. La conclusion s'ensuit. □

Comme autre illustration, on peut considérer le champ gaussien sur $\{-1, 1\}^n$, $n \geq 1$, définit de la façon suivante. Soient $X_1, \ldots, X_n$ des variables gaussiennes standards *i.i.d.*, et définissons $f : \{-1, 1\}^n \to \mathbb{R}$ par

$$f(\sigma) = \sum_{i=1}^{n} X_i \sigma_i.$$

Corollaire 4.2.7. *Il existe $S \subset \{-1,1\}^n$, tel que, pour tout $\sigma \neq \sigma' \in S$, $|\sigma \cdot \sigma'| \leq Cn^{2/3}$ et*

$$\mathbb{P}\left(\Big|\max_{\sigma\in S} f(\sigma) - \mathbb{E}\big[\max_{\sigma\in S} f(\sigma)\big]\Big| \geq t\right) \leq 6e^{-ct/n^{1/3}}, \quad t \geq 0.$$

Démonstration. Chatterjee a démontré dans [48] l'existence d'un ensemble $S \subset \{-1,1\}^n$ tel que $\mathrm{Card}(S) = \lfloor 2^{n^{1/3}} \rfloor$ sur lequel $|\sigma \cdot \sigma'| \leq Cn^{2/3}$ avec $C > 0$ independant de n. Puisque, $\mathrm{Cov}\big(f(\sigma), f(\sigma)\big) = \sigma \cdot \sigma'$, il suffit d'appliquer la proposition 4.2.6 à $f_{|S}$ (renormalisée par $n^{1/2}$) et $\epsilon = Cn^{-1/3}$ en observant que $K_n = 1/n^{1/3}$. Nous obtenons donc

$$\mathbb{P}\left(\Big|\max_{\sigma\in S} f(\sigma) - \mathbb{E}\big[\max_{\sigma\in S} f(\sigma)\big]\Big| \geq u\sqrt{n}\right) \leq 6e^{-cun^{1/6}}, \quad t \geq 0.$$

Il suffit alors de poser $t = u\sqrt{n}$ pour conclure.

□

Théorème de Darling-Erdös

Nous proposons une inégalité de concentration qui reflète le résultat de convergence du théorème 3.3.1 de Darling-Erdös. De plus, une fois intégrée, cette inégalité de concentration implique que $\mathrm{Var}(M_n) \leq \frac{C}{\log\log n}$. Nous rappelons, ci-dessous l'énoncé du théorème 3.3.1.

Théorème 4.2.8. *Soient $(X_i)_{i\geq 1}$ des variables gaussiennes standards indépendantes et de même loi. On définit, pour tout $k \in \mathbb{N}$, $S_k = X_1 + \cdots + X_k$ et l'on pose $Z_i = \frac{S_i}{\sqrt{i}}$ pour $i = 1, \ldots, n$. $Z = (Z_1, \ldots, Z_n)$ est un vecteur gaussien centré dans $\mathbb{R}^n$, admettant la structure de covariance suivante*

$$\mathrm{Cov}(Z_i, Z_j) = \frac{\min(i,j)}{\sqrt{ij}}, \quad i, j = 1, \ldots, n$$

Alors, nous avons le résultat de convergence en loi

$$a_n(M_n - b_n) \to \Lambda_0,$$

lorsque $n \to \infty$, avec $M_n = \max_{i=1,\ldots,n} Z_i$, $a_n = \sqrt{2\log\log n}$ et $b_n = a_n + \frac{1/2\log\log\log n - \log 2\sqrt{\pi}}{a_n}$.

Théorème 4.2.9. *Avec les notations précédentes, l'inégalité suivante est satisfaite, pour tout $n > e$,*

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| \geq t) \leq 6e^{-ct\sqrt{\log\log n}}, \quad t \geq 0,$$

Remarque. Le même résultat est satisfait si l'on remplace $M_n = \max_{i=1,\ldots,n} Z_i$ par $M_n = \sup_{t\in[1,n]} \frac{B_t}{\sqrt{t}}$.

Démonstration. Le résultat est obtenu facilement dès lors que l'on remarque le fait suivant : $(Z_1, \ldots, Z_n)$ a la même loi que $(X_{t_1}, \ldots, X_{t_n})$ où $t_i = \frac{1}{2}\log i$ pour $i = 1, \ldots, n$ avec $(X_t)_{t\geq 0}$ désignant le processus d'Ornstein-Uhlenbeck. Rappelons que celui-ci est un processus gaussien stationnaire centré admettant pour fonction de covariance $\mathrm{Cov}(X_s, X_t) = e^{-|t-s|}$ pour tout $s, t \geq 0$.

En effet, pour tout $i, j \in \{1, \ldots, n\}$ nous avons

$$\frac{\min(i,j)}{\sqrt{ij}} = e^{-|\frac{1}{2}\log i - \frac{1}{2}\log j|}.$$

C'est pourquoi $M_n = \max_{i=1,\dots,n} X_{t_i}$ en loi. La conclusion est obtenue via le théorème 4.2.5 appliqué au processus d'Ornstein-Uhlenbeck $(X_t)_{t\geq 0}$.

En effet, le processus d'Ornstein-Uhlenbeck vérifie les hypothèses du théorème, il suffit ensuite de choisir $A = \{t_1, \dots, t_n\}$ pour conclure puisque $N(A) = \frac{1}{2}\log n$ et $K_{N(A)} = \frac{1}{2}\log n$. $\square$

Remarque. Comme nous l'avons mentionné précédemment, le théorème de Darling-Erdös est valable pour des lois satisfaisants une condition d'intégrabilité relativement faible (grossièrement, un moment d'ordre trois suffit). Il serait donc intéressant de proposer une preuve de l'inégalité de concentration précédente ne reposant pas sur des outils gaussiens, afin d'atteindre l'universalité du théorème de Darling-Erdös.

Champ libre gaussien discret sur un groupe

Nous avons déjà évoqué les champs libres gaussiens discrets dans le chapitre trois. Une des propriétés importantes de tels objets est que le comportement de la marche aléatoire sous jacente détermine celui de la fonction de covariance. Par exemple sur $\mathbb{Z}^2$ la fonction de covariance est logarithmique, tandis que pour $\mathbb{Z}^d, d \geq 3$ elle se comportera comme l'inverse d'une fonction puissance (ceci correspondant bien aux différents comportements de la marche aléatoire symétrique). En fait, $\mathbb{Z}^d, d \geq 1$ est un groupe infini de type fini, d'après [138], il est possible de considérer des marches aléatoires symétriques sur un groupe $\mathcal{S}$ infini de type fini abstrait. Les travaux de Saloff-Coste attestent que des hypothèses sur la croissance de volume de boule permettent alors de déterminer le comportement de la fonction de Green associée à la marche aléatoire symétrique. Ceci entrainant donc, *a fortiori*, un contrôle de la fonction de covariance du champ libre gaussien discret indexé par $\mathcal{S}$. Dans le cas de $\mathbb{Z}^d, d \geq 3$ cela correspond à l'estimation suivante, présente dans [60],

$$\mathrm{Cov}(X_i, X_j) \leq \frac{C}{\|i-j\|_\infty^{d-2}},$$

avec $(X_i)_{i\in\mathbb{Z}^d}$, $d \geq 3$ correspondant au champ libre gaussien discret indexé par $\mathbb{Z}^d$, $C > 0$ étant une constante numérique. Nous proposons ci-dessous une inégalité de superconcentration reflétant le résultat de convergence du Théorème 3.2.7 (que nous rappelons ci-dessous), démontré dans [60].

Théorème 4.2.10. *Soient A un sous-ensemble de $\mathbb{Z}^d$ satisfaisant* $\mathrm{Card(A)} = n^q$, $q \geq 1$ *et* $(X_i)_{i\in\mathbb{Z}^d}, d \geq 3$ *le champ libre gaussien discret indexé par* $\mathbb{Z}^d$, $d \geq 3$*. On pose*

$$b_n = \left[\sqrt{2\log n^q} - \frac{\log\log n^q + \log 4\pi}{2\sqrt{2\log n^q}}\right] \quad \textit{and} \quad a_n = \sqrt{2\log n^q}.$$

Alors, nous avons la convergence en loi suivante

$$a_n\left(\max_{i\in A} X_i - b_n\right) \to \Lambda_0 \quad n\to\infty$$

Nous ferons la démonstration, de l'inégalité de concentration, dans le cas où $\mathcal{S} = \mathbb{Z}^d$, $d \geq 3$. La preuve serait identique dans un cadre plus général, les arguments essentiels de celle-ci sont les suivants :

1. Nous avons besoin d'un contrôle sur la covariance par une fonction ϕ qui décroit vers 0 à l'infini plus vite que $n \mapsto 1/\log n$. Dans le cas de $\mathbb{Z}^d$, $d \geq 3$ $\phi(n) = \frac{C}{n^{d-2}}$. Par exemple, pour un groupe satisfaisant des hypothèses de croissance polynomiale similaires à $\mathbb{Z}^d$, $d \geq 3$,

on aurait également $\phi(n) = \frac{C}{n^{d-2}}$ avec $C > 0$, $d > 2$ des constantes numériques.

2. Nous avons aussi besoin de contrôler le nombres de points de $\mathbb{Z}^d$, $d \geq 3$ contenu dans une boule de rayon $s > 0$. A nouveau, ceci est fournit par la croissance du volume de boule. Dans le cas de $\mathbb{Z}^d$, toute boule de rayon $s > 0$ contient au maximum s^d éléments de $\mathbb{Z}^d$.

Ces observations étant faites, nous pouvons énoncer notre inégalité de superconcentration pour le champ libre gaussien indexé par $\mathbb{Z}^d$, $d \geq 3$. Nous laissons aux lecteurs le soin d'apporter les modifications évidentes à la preuve ainsi qu'à l'énoncé du théorème lorsque l'on remplace $\mathbb{Z}^d$, $d \geq 3$ par un groupe $\mathcal{S}$ satisfaisant des hypothèses de croissance de volume de boules différentes.

Théorème 4.2.11. *Sous le cadre précédent, l'inégalité de concentration suivante est satisfaite par le champ libre gaussien discret $(X_i)_{i\in\mathbb{Z}^d}$ indexé par $\mathbb{Z}^d$, $d \geq 3$. Pour tout sous-ensemble $A \subset \mathbb{Z}^d$, $d \geq 3$ tel que $\mathrm{Card}(A) = n^q$, $q \geq 1$ et tout $n \geq 1$,*

$$\mathbb{P}(|M_A - \mathbb{E}[M_A]| \geq t) \leq 6e^{-ct\sqrt{\log n}}, \quad t \geq 0, \tag{4.14}$$

où $M_A = \max_{i\in A} X_i$, avec $c > 0$ une constante numérique ne dépendant que de $\mathbb{Z}^d$.

Nous aurons besoin du lemme suivant pour la preuve. Notons qu'il s'agit d'une adaptation du lemme 9.13 (page 101) de [48].

Lemme 4.2.12. *Soit $A \subset \mathbb{Z}^d$, $d \geq 3$. Alors il existe $C_1 > 0$ (ne dépendant que de la dimension d) telle que*

$$C_1\sqrt{\log \mathrm{Card}(A)} \leq \mathbb{E}[M_A].$$

Démonstration. La démonstration du lemme repose sur le principe de minoration de Sudakov. Puisque l'on a l'estimation suivante de la covariance

$$\mathrm{Cov}(X_i, X_j) \leq \frac{C}{\|i-j\|_\infty^{d-2}}, \quad i,j \in \mathbb{Z}^d$$

et que $\phi(n) = \frac{C}{n^{d-2}} \to 0$ lorsque $n \to \infty$, il est possible de trouver $s_0 \in \mathbb{R}$ tel $\phi(s_0) < 1$. Fixons un tel réel s_0 et considérons, tout d'abord, un réseau (pour la métrique $d(\cdot,\cdot)$ induite par $\|\cdot\|_\infty$) B de rayon 1 de l'ensemble A ainsi qu'un second réseau D de rayon s_0 de l'ensemble B. Pour n'importe quel élément $x \in D$, la boule de rayon s_0 centrée en x contient au plus k éléments de B (ce nombre k étant un nombre fixé, dépendant uniquement de s_0 et de la dimension d). Ceci entraine l'inégalité suivante

$$\mathrm{Card}(D) \geq \frac{\mathrm{Card}(B)}{k} = \frac{N(A)}{k}.$$

Enfin, remarquons que pour tout $i \neq j \in D$ nous avons $d(i,j) = \|i-j\|_\infty \geq s_0$.

D'où,

$$\mathbb{E}[(X_i - X_j)^2] = 2\big(1 - \mathrm{Cov}(X_i, X_j)\big) \geq 2\Big(1 - \phi\big(d(i,j)\big)\Big) \geq 2(1 - \phi(s_0)) > 0$$

Alors, d'après la minoration de Sudakov, il existe une constante $c_1 > 0$ telle que

$$\mathbb{E}[M_D] \geq c_1\sqrt{2\log \mathrm{Card}(D)}.$$

En conclusion,

$$\mathbb{E}[M_A] \geq \mathbb{E}[M_D] \geq c_1\sqrt{2\log\mathrm{Card}(D)} \geq C_1(s_0,d)\sqrt{\log N(A)}$$

Ce qui conclu la preuve. □

A présent, nous pouvons démontrer l'inégalité de concentration. Nous expliquerons brièvement les modifications à apporter à la preuve lorsque la croissance du volume des boules n'est pas polynomiale. Le schéma de preuve s'inspire d'une démonstration de Chatterjee (cf. [48] page 101).

Démonstration. Soit C_1 la constante fournie par le lemme précédent. On pose

$$s_0 := \mathrm{Card}(A)^{\frac{C_1^2}{8d}}/2 \quad \text{et} \quad r_0 = \phi(s_0)$$

où ϕ désigne la fonction majorant la covariance. Soit B un réseau maximal de rayon s_0 (pour le même C_1 que celui obtenu dans la démonstration du lemme précédent) de l'ensemble A. On, définit alors $C(r_0)$ comme étant l'ensemble des boules de rayon $2s_0$ centrées en les points du réseau B. Il est facile de voir, qu'un tel recouvrement satisfait les hypothèses du théorème 4.2.2.

En effet, soient $i,j \in A$ tels que $\mathrm{Cov}(X_i,X_j) \geq r_0 = \phi(s_0)$. Alors, puisque $\mathrm{Cov}(X_i,X_j) \leq \phi\big(d(i,j)\big) = \phi\big(\|i-j\|_\infty\big)$ et comme ϕ est décroissante, nous devons avoir $\|i-j\|_\infty \leq s_0$. Notons par x_i et x_j le centre des boules tel que $i \in B(x_i,s_0)$ et $j \in B(x_j,s_0)$. Puisque $\|i-j\|_\infty \leq s_0$, soit i et j sont dans la même boule, ainsi $x_i = x_j$ et $D = B(x_i,2s_0) \in C(r_0)$. Ou bien, $x_i \neq x_j$ mais alors $\|i-j\|_\infty \leq s_0$ entraine (faire un dessin) que l'on peut choisir $D = B(x_i,2s_0)$ ou $D = B(x_j,2s_0)$ dans $C(r_0)$.

Il reste donc à obtenir des bornes sur $\tau(r_0)$ et $\rho(r_0)$ afin d'appliquer le théorème 4.2.2.

Considérons donc $D \in \mathcal{C}(r_0)$, puisque $(X_i)_{i\in\mathbb{Z}^d}$ est composé de variables gaussiennes et que D contient au plus $(2s_0)^d$ éléments, nous avons l'estimation suivante

$$\mathbb{E}[M_D] = \mathbb{E}\big[\max_{i\in D} X_i\big] \leq \sqrt{2\log\mathrm{Card}(D)} \leq \sqrt{2\log(2s_0)^d} = \frac{C_1}{2}\sqrt{\log\mathrm{Card}(A)}, \tag{4.15}$$

par définition de s_0.

De plus, d'après le lemme 4.2.12, nous avons $\mathbb{E}[M_A] = \mathbb{E}\big[\max_{i\in A} X_i\big] \geq C_1\sqrt{\log\mathrm{Card}(A)}$.

Pour alléger les notations, nous noterons par $m(H)$, pour tout $H \subset \mathbb{Z}^d$, la quantité $\mathbb{E}\big[\max_{i\in H} X_i\big]$. En utilisant des inégalités de concentration standards, nous déduisons de ce qui précède les inégalités suivantes

$$\begin{aligned}
\mathbb{P}(I \in D) &\leq \mathbb{P}\left(\max_{i\in D} X_i \geq \frac{m(A)+m(D)}{2}\right) + \mathbb{P}\left(\max_{i\in A} X_i \leq \frac{m(A)+m(D)}{2}\right) \\
&= \mathbb{P}\left(\max_{i\in D} X_i \geq m(D) + \frac{m(A)-m(D)}{2}\right) \\
&+ \mathbb{P}\left(\max_{i\in A} X_i \leq m(A) - \frac{m(A)-m(D)}{2}\right) \\
&\leq 2\exp\left(-\frac{C_1^2\log\mathrm{Card}(A)}{32}\right),
\end{aligned}$$

puisque $m(A) - m(D) \geq \frac{C_1}{2}\sqrt{\log \mathrm{Card}(A)}$. La borne ci-dessus ne dépendant pas de D, permet de majorer $\rho(r_0)$, c'est à dire

$$\rho(r_0) \leq \frac{2}{\mathrm{Card}(A)^{C_1^2/32}}.$$

Il reste à majorer $\tau(r_0) = \sum_{D \in \mathcal{C}(r_0)} 1_{I \in D}$.

Soit $i \in A$, le centre de la boule $D \in \mathcal{C}(r_0)$ contenant i est également un point de $B(i, 2s_0)$. Puisque les centres des boules composant l'ensemble $\mathcal{C}(r_0)$ sont mutuellement séparés par une distance supérieure à s_0, le nombre d'éléments $D \in \mathcal{C}(r_0)$ contenant i est borné par $N\big(B(u, 2s_0), s_0\big)$ (faire un dessin), lequel, par changement d'échelle, vaut $N\big(B(0,2), 1\big)$. En conclusion $\tau(r_0) \leq c$, avec c une constante qui ne dépend que de $\mathbb{Z}^d$.

En combinant les bornes précédentes, le théorème 4.2.2 entraine que

$$\mathbb{P}(|M_A - \mathbb{E}[M_A]| \geq t) \leq 6e^{-ct\sqrt{\log \mathrm{Card}(A)}}, \quad t \geq 0,$$

puisque $\max\Big(r_0, 1/\sqrt{\log \mathrm{Card}(A)}\Big) = 1/\log \mathrm{Card}(A)$. □

Remarque. Lorsque les conditions de croissance de volume de boule ne sont pas polynomiale, il suffit de choisir s_0 de telle sorte que l'équation 4.15 soit satisfaite. Comme nous avons pu le voir dans le chapitre deux, le cas de la dimension $d = 2$ a été l'objet d'une longue étude et les méthodes hypercontractives, présentées ci-dessus, ne semblent pas permettre de retrouver les résultats optimaux obtenus dans [72, 44]... (principalement à cause du comportement logarithmique de la fonction de covariance).

4.3 Lemme de Slepian et superconcentration

Un outil très utile pour des variables gaussiennes est le lemme de comparaison de Slepian (cf. [117]) que nous rappelons ci-dessous.

Théorème 4.3.1. *(Slepian) Soient X et Y deux vecteurs gaussiens centrés de $\mathbb{R}^n$ tels que*

$$\mathbb{E}\left[X_i X_j\right] \leq \mathbb{E}\left[Y_i, Y_j\right], \quad \forall i \neq j \in \{1, \ldots, n\}$$

et

$$\mathbb{E}\left[X_i^2\right] = \mathbb{E}\left[Y_i^2\right] \quad \forall i \in \{1, \ldots, n\}$$

Alors, pour tout nombre $\lambda_i \in \mathbb{R}$, $i \leq n$

$$\mathbb{P}\left(\cup_{i=1}^n \left(Y_i > \lambda_i\right)\right) \leq \mathbb{P}\left(\cup_{i=1}^n \left(X_i > \lambda_i\right)\right).$$

Le but de cette section est d'obtenir un résultat analogue au lemme de Slepian pour obtenir de la superconcentration en comparant les structures de covariance. Nous proposons, ci-dessous, un outil de comparaison dans cet esprit. Soient X, Y, Z des vecteurs gaussiens centrés de $\mathbb{R}^n$, $n \geq 1$ de matrices de covariance respectives $\Gamma^X, \Gamma^Y, \Gamma^Z$.

Proposition 4.3.2. *Supposons que $\Gamma^Y \leq \Gamma^X \leq \Gamma^Z$ et que, pour un certain $m_n \in \mathbb{R}$ et tout $t > 0$, les inégalités suivantes soient satisfaites*

1. $\mathbb{P}\left(\max_{i=1,\dots,n} Y_i - m_n \leq -t\right) \leq \alpha_n(t)$

2. $\mathbb{P}\left(\max_{i=1,\dots,n} Z_i - m_n \geq t\right) \leq \beta_n(t)$.

Alors, pour tout $t \geq 0$ *et tout* $n \geq 1$,

$$\mathbb{P}\left(\left|\max_{i=1,\dots,n} X_i - m_n\right| \geq t\right) \leq 2\max\left(\alpha_n(t), \beta_n(t)\right) = \theta_n(t)$$

De plus, si l'on suppose que $t \mapsto \theta_n(t)$ *et* $t \mapsto t\theta_n(t)$ *sont intégrables sur* $[0, +\infty[$. *Alors,*

$$\left|\mathbb{E}\left[\max_{i=1,\dots,n} X_i\right| - m_n\right| \leq C$$

et

$$\mathrm{Var}\left(\max_{i=1,\dots,n} X_i\right) \leq C \int_0^\infty t\theta_n(t)dt.$$

Démonstration. Pour tout $t \in \mathbb{R}$, le lemme de Slepian 4.3.1 nous assure que

1. $\mathbb{P}\left(\max_{i=1,\dots,n} X_i \leq -t\right) \leq \mathbb{P}\left(\max_{i=1,\dots,n} Y_i \leq -t\right)$.

2. $\mathbb{P}\left(\max_{i=1,\dots,n} X_i \geq t\right) \leq \mathbb{P}\left(\max_{i=1,\dots,n} Z_i \geq t\right)$.

Ce qui implique, pour tout $t > 0$, que $\mathbb{P}\left(|\max_{i=1,\dots,n} X_i - m_n| \geq t\right) \leq 2\theta_n(t)$. Le résultat concernant l'écart entre m_n et $\mathbb{E}\left[\max_{i=1,\dots,n} X_i\right]$ est classique.

$$\begin{aligned}\left|\mathbb{E}\max_{i=1,\dots,n} X_i - m_n\right| &\leq \mathbb{E}\left[\left|\max_{i=1,\dots,n} X_i - m_n\right|\right]\\ &= \int_0^\infty \mathbb{P}\left(\left|\max_{i=1,\dots,n} X_i - m_n\right| > t\right)dt\\ &\leq \int_0^\infty \theta_n(t)dt < +\infty\end{aligned}$$

Quant à la dernière affirmation de la proposition, elle découle de ce qui suit

$$\begin{aligned}\mathbb{P}\left(\left|\max_{i=1,\dots,n} X_i - \mathbb{E}\left[\max_{i=1,\dots,n} X_i\right]\right| \geq t\right) &\leq \mathbb{P}\left(\left|\max_{i=1,\dots,n} X_i - m_n\right| \geq t - C\right)\\ &\leq 2\theta_n(t-C)\end{aligned}$$

Fournissant donc, pour tout $u > 0$, les inégalités suivantes

$$\mathbb{P}\left(\left|\max_{i=1,\dots,n} X_i - \mathbb{E}\left[\max_{i=1,\dots,n} X_i\right]\right| \geq u\right) \leq \hat{C}\theta_n(u)$$

Pour conclure, il suffit de remarquer que

$$\mathrm{Var}(\max_{i=1,\dots,n} X_i) \leq \mathbb{E}\big[(\max_{i=1,\dots,n} X_i - m_n)^2\big] = 2\int_0^\infty u\mathbb{P}\big(|M_n - m_n| \geq u\big)du$$

et de majorer l'intégrande par $\theta_n(u)$. □

4.3.1 Champ libre gaussien discret avec condition de bord

Nous allons illustrer la proposition 4.3.2 sur un champ libre gaussien discret avec des conditions de bords. Récemment, dans [60], Cipriani *et al.* ont obtenu un résultat de convergence en loi du maximum du champ libre gaussien discret $(X_i)_{i \in V_n}$ sur $V_n = [0, n-1]^d$ avec des conditions de bords, vers la loi de Gumbel Λ_0. Nous énonçons leur théorème ci-dessous

Théorème 4.3.3. *(Cipriani et al.) Soit V_n défini comme ci-dessus et $(X_i)_{i \in V_n}$ un champ libre gaussien discret avec des conditions de bords nulles en dehors de V_n. Avec les mêmes constantes a_n et b_n que celles du théorème 3.4 nous avons la convergence en loi suivante*

$$a_n \left(\max_{i \in V_n} X_i - b_n \right) \to \Lambda_0, \quad n \to \infty.$$

Ce résultat est de même nature que celui concernant les suites gaussiennes stationnaires et nous avons obtenu, via un argument de comparaison, une inégalité de concentration non-asymptotique reflétant cette convergence.

Théorème 4.3.4. *Pour tout $t > 0$, nous avons l'inégalité suivante, $n \geq 1$,*

$$\mathbb{P} \left(\left| \max_{i \in V_n} X_i - \mathbb{E} \left[\max_{i \in V_n} X_i \right] \right| \geq t \right) \leq C e^{-ct\sqrt{a_n}},$$

avec $C, c > 0$ des constantes numériques.

La démonstration du théorème 4.3.4 sera obtenue, via la proposition 4.3.2, après comparaison avec le champ libre gaussien discret sur $\mathbb{Z}^d, d \geq 3$ (sans conditions de bords) et le cas indépendant.

Démonstration. L'inégalité découle de l'application de la proposition 4.3.2. En effet, on peut choisir Y comme étant un vecteur gaussien standard dans $\mathbb{R}^n$ et Z le champ libre gaussien discret sur $\mathbb{Z}^d$, $d \geq 3$ puisque, d'après [60],

$$0 \leq \mathrm{Cov}(X_i, X_j) \leq \Gamma_{ij}^Z, \quad \forall i \neq j \in V_n$$

et l'on a déjà montré, cf. (4.8) et (4.14) les inégalités suivantes

1. $\mathbb{P} \left(\max_{i=1,\dots,n} Y_i - m_n \leq -t \right) \leq 6e^{-t\sqrt{a_n}}$
2. $\mathbb{P} \left(\max_{i=1,\dots,n} Z_i - m_n \geq t \right) \leq 6e^{-t\sqrt{a_n}}$

Les hypothèses de la proposition 4.3.2 sont donc satisfaites. □

Remarque. Ce genre d'arguments de comparaison a déjà été utilisé pour étudier des champ gaussiens log-corrélés (le champ libre gaussien discret en dimension deux par exemple [44, 68, 72]). La procédure est la même :

1. Ils obtiennent une inégalité de déviation (à gauche) pour la marche aléatoire branchante ainsi qu'une inégalité de déviation (à droite) pour la marche aléatoire branchante modifiée (« modified branching random walk »).

2. Ensuite, il montre que la structure de covariance du champ libre gaussien discret est comparable à celle de la marche aléatoire branchante et de la la marche aléatoire branchante modifiée.

Alors que les hypothèses de la proposition 4.3.2 imposent que les inégalités de déviation satisfaites par Y et Z soient valables pour tout $t > 0$. Dans l'article [166] ils en obtiennent pour $0 < t < c_n$ et utilisent la concentration gaussienne standard pour $t > c_n$. Néanmoins, la précision de la première inégalité, pour les petites valeurs de t, est suffisante pour retrouver, une fois intégrée, le fait que la variance du maximum du champ libre gaussien discret sur $\mathbb{Z}^2$ est d'ordre $O(1)$.

4.3.2 Suite gaussienne stationnaire avec décroissance lente

Notre inégalité de comparaison va également nous permettre d'obtenir une inégalité de déviation (à droite) pour des suites gaussiennes stationnaires (lorsque la fonction de covariance décroit plus lentement que $n \mapsto 1/\log n$ à l'infini). Cela va notamment permettre d'obtenir une inégalité pertinente, qu'on ne peut pas obtenir via des méthodes hypercontractives, vis à vis du théorème 3.2.3. Celui-ci assurant la convergence en loi suivante :

Théorème 4.3.5. *(Mittal-Ylvisaker) Soit $(X_i)_{i\geq 0}$ une suite gaussienne stationnaire avec, pour fonction de covariance, ϕ telle que $\phi(n)\log n \to \infty$ as $n \to \infty$ et $\phi(n) \to 0$ monotonement. Alors*

$$c_n\left(M_n - \left(1-c_n^{-2}\right)^{1/2} b_n\right) \underset{n\to\infty}{\longrightarrow} X,$$

où X suit une loi gaussienne standard et $c_n = \sqrt{\phi(n)^{-1}}$ et $b_n = (2\log n)^{1/2} - \frac{1}{2}(2\log n)^{-1/2}(\log\log n + \log 4\pi)$.

Inégalité de déviation à droite

Nous nous restreignons à la version faible du théorème de Mittal et Ylvisaker, les calculs semblent trop difficiles dans le cas général. Plus précisément, en suivant l'exposition faite dans [108] (page 138) nous considérons la suite convexe $\big(\phi(n)\big)_{n\geq 0}$ avec $\phi : \mathbb{N} \to \mathbb{R}$ telle que $\phi(0) = 1$ et $\phi(n) = \frac{C}{(\log n)^\delta}$ pour $n \geq n_0 \geq 2$ avec $0 < \delta < 1$ et $C > 0$. Remarquons que $\phi(n)\log n \to \infty$ lorsque $n \to \infty$. D'après [108], le critère de Polya entraine que $\big(\phi(n)\big)_{n\geq 0}$ est suite de covariance. On désigne alors par $(X_i)_{i\geq 0}$ la suite gaussienne stationnaire admettant une telle fonction de covariance et par $M_n = \max_{i=1,\ldots,n} X_i$. Nous proposons l'inégalité de déviation suivante :

Théorème 4.3.6. *Avec les notations précédentes, l'inégalité suivante est valable pour tout $t \in [0, c\log(n)^{(1-\delta)/2}]$, $n \geq N_0$,*

$$\mathbb{P}\left(c_n\left(M_n - \left(1-c_n^{-2}\right)^{1/2} b_n\right) \geq t\right) \leq C e^{-t^2/2},$$

avec $c_n = \sqrt{\phi(n)^{-1}}$, $b_n = (2\log n)^{1/2} - \frac{1}{2}(2\log n)^{-1/2}(\log\log n + \log 4\pi)$ et $C > 0$ une constante numérique.

Démonstration. Puisque $\phi(n)\log n \to \infty$ est convexe, d'après [108] (page 138) on peut également considérer une autre fonction de covariance définie comme suit

$$\psi(k) = \begin{cases} \frac{\phi(k)-\phi(n)}{1-\phi(n)}, & 0 \leq k \leq n \\ 0, \quad k > n \end{cases}$$

Soit $(Y_i')_i$ une suite gaussienne stationnaire admettant $\big(\psi(n)\big)_{n\geq 0}$ comme structure de covariance, similairement nous posons $M_n' = \max_{i=1\ldots,n} Y_i$. L'égalité en loi suivante est alors essentielle pour la démonstration du théorème.

$$M_n = (1-\phi(n))^{1/2} M_n' + \phi(n)^{1/2} X, \tag{4.16}$$

avec X une variables aléatoire gaussienne standard, indépendante de $(X_i)_{i\geq 0}$ et $(Y_i)_{i\geq 0}$.

Soient $t \geq 0$ et $n \geq N_0$ tel que $1-\phi(N_0) \leq 1/4$, l'inégalité de Chernoff combinée à (4.16) entraine le fait suivant

$$\mathbb{P}\Big(\phi(n)^{-1/2}\big(M_n - (1-\phi(n))^{1/2} b_n\big) \geq t\Big) \leq e^{-\lambda t + \lambda^2/2} \mathbb{E}[e^{\frac{\lambda}{2}\phi(n)^{-1/2}[M_n' - b_n]}], \tag{4.17}$$

avec $\lambda \geq 0$.

En intégrant par partie

$$\mathbb{E}[e^{\lambda/2\phi(n)^{-1/2}[M_n'-b_n]}] \leq 1 + \frac{\lambda}{2}\int_0^\infty e^{\lambda r}\mathbb{P}(\phi(n)^{-1/2}[M_n'-b_n] \geq r)dr$$

De plus, pour tout $k \in \mathbb{N}$, $\psi(k) \geq 0$, alors l'inégalité de comparaison (4.3.2) entraine que

$$\mathbb{P}([M_n' - b_n] \geq r\phi(n)^{1/2}) \leq \mathbb{P}([M_n(0) - b_n] \geq r\phi(n)^{1/2}),\ r > 0$$

où $M_n(0)$ désigne le maximum de n gaussiennes standards indépendantes.

Puisque, d'après 4.8, nous avons pour tout $r > 0$

$$\mathbb{P}(|M_n(0) - b_n| \geq r) \leq Ce^{-cra_n},$$

avec $a_n = \sqrt{2\log n}$ (la constante de renormalisation du cas indépendant). Ceci implique

$$\mathbb{P}([M_n' - b_n] \geq r\phi(n)^{1/2}) \leq Ce^{-cr(\log n)^{(1-\delta)/2}},$$

avec $C, c > 0$ des constantes numériques. Finalement, pour $\lambda \leq c(\log n)^{(1-\delta)/2}$, ceci permet de majorer la transformée de Laplace

$$\mathbb{E}[e^{\frac{\lambda}{2}\phi(n)^{-1/2}[M_n'-b_n]}] \leq 1 + \frac{C\lambda}{2(c(\log n)^{(1-\delta)/2} - \lambda)}.$$

En particulier, pour tout $0 \leq \lambda \leq c(\log n)^{(1-\delta)/2}/2$

$$\mathbb{E}[e^{\frac{\lambda}{2}\phi(n)^{-1/2}[M_n'-b_n]_+}] \leq 1 + \frac{C}{2} = C_1.$$

En d'autres termes, l'inégalité initiale

$$\mathbb{P}\Big(\phi(n)^{-1/2}\big(M_n - (1-\phi(n))^{1/2} b_n\big) \geq t\Big) \leq e^{-\lambda t + \lambda^2/2}\mathbb{E}[e^{\lambda/2\phi(n)^{-1/2}[M_n'-b_n]}],$$

est majorée, pour tout $0 \leq \lambda \leq c(\log n)^{(1-\delta)/2}/2$. C'est à dire

$$\mathbb{P}\Big(\phi(n)^{-1/2}\big(M_n - (1-\phi(n))^{1/2} b_n\big) \geq t\Big) \leq Ce^{-\lambda t + \lambda^2/2}.$$

Optimisons alors en $\lambda = t$, pour tout $0 \leq t \leq c(\log n)^{(1-\delta)/2}/2$, afin d'obtenir

$$\mathbb{P}\Big(c_n\big(M_n-(1-c_n^{-2})^{1/2}b_n\big)\geq t\Big)\leq Ce^{-t^2/2}.$$

Ce qui conclut la démonstration et founit une inégalité de déviation non asymptotique, reflétant le théorème 3.2.3 □

4.4 Superconcentration pour des lois non-gaussiennes

Cette section présente l'utilisation de l'extension de l'inégalité de Talagrand 1.3.8, vis à vis de la superconcentration, pour des mesures hypercontractives non-gaussiennes. Comme observé par Cordero-Erausquin et Ledoux dans [62], l'inégalité de Talagrand 1.3.6 peut-être obtenue dans un cadre plus abstrait de diffusions markoviennes (cf. [16]). Les propriétés essentielles de ce cadre est de considérer un processus de Markov $(X_t)_{t\geq 0}$ avec sa mesure invariante μ, ainsi que son générateur infinitésimal L. Il est supposé que μ est une mesure de probabilité (sur un espace mesurable $(E,\mathcal{A})$) hypercontractive et que l'énergie de Dirichlet $\mathcal{E}_\mu$, liée à L, peut-être décomposée suivant des directions Γ_i le long desquelles une propriété de commutation avec le semi groupe de $(X_t)_{t\geq 0}$ est satisfaite. Plus précisément, pour $f:E\to\mathbb{R}$ suffisamment régulière, cela signifie

$$\mathcal{E}_\mu(f)=\int_E f(-Lf)d\mu=\sum_{i=1}^n\int_E\Gamma_i(f)^2d\mu\quad\Gamma_i(P_tf)\leq e^{\kappa t}P_t(\Gamma_i f)\quad t\geq 0,\,\kappa\in\mathbb{R}\tag{4.18}$$

Rappelons, que la mesure standard gaussienne γ_n satisfait ces hypothèses avec le semi groupe d'Ornstein-Uhlenbeck et son générateur $L=\Delta-x\cdot\nabla$ avec $\Gamma_i=\partial_i$ et $\kappa=-1$. En effet, on peut calculer $\mathcal{E}_{\gamma_n}(f)=\int_{\mathbb{R}^n}|\nabla f|^2d\gamma_n$, pour f dans le domaine $\mathcal{D}(L)$, et observer que $\mathcal{E}_{\gamma_n}$ peut-être décomposée comme suit :

$$\mathcal{E}_{\gamma_n}(f)=\sum_{i=1}^n\int_{\mathbb{R}^n}(\partial_i f)^2d\gamma_n=\sum_{i=1}^n\int_{\mathbb{R}^n}\Gamma_i(f)^2d\gamma_n\quad\text{et}\quad\Gamma_i(P_tf)=e^{-t}P_t(\Gamma_i f).$$

Sous les conditions (4.18), le résultat principal de [62] est le suivant

Théorème 4.4.1. *Pour $f:E\to\mathbb{R}$ suffisamment régulière, $n\geq 1$,*

$$\mathrm{Var}_\mu(f)\leq C\sum_{i=1}^n\frac{\|\Gamma_i(f)\|_2^2}{1+\log\left(\frac{\|\Gamma_i(f)\|_2}{\|\Gamma_i(f)\|_1}\right)}$$

Notons que ce théorème correspond bien à une extension de 1.3.6 puisqu'il peut être simplement appliqué à des mesures log-concaves $d\mu(x)=e^{-V(x)}dx$ où $x\mapsto V(x)$ est un potentiel strictement convexe (*i.e.* $V''\geq c>0$) et $E=\mathbb{R}^n$. Pour ce type de mesure, on trouve la même énergie de Dirichlet que le cas gaussien $\mathcal{E}_\mu(f)=\int_{\mathbb{R}^n}|\nabla f|^2d\mu$, pour f suffisamment régulière et donc la même décomposition suivant les directions Γ_i. Pour cet exemple, la propriété de commutation s'exprime de la manière suivante $\Gamma_i(P_tf)\leq e^{-ct}P_t(\Gamma_i f)$. Autrement dit, $\kappa=-c$ où c correspond à la constante minorant la hessienne du potentiel V.

Dans ce qui suit, nous proposons deux applications de cette extension (de l'inégalité de Talagrand) à la théorie des extrêmes. Ceci s'inscrit dans le cadre de la superconcentration et permet d'exhiber deux exemples non-gaussiens.

4.4.1 Maximum des coordonnées d'un vecteur de loi uniforme sur la sphère

La mesure uniforme sur la sphère $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ vérifie les hypothèses du théorème 4.4.1. Notons que cet exemple est une mesure non-produit. Considérons, pour tout $i, j = 1, \ldots, n$ $D_{ij} = x_i \partial_j - x_j \partial_i$. il s'agit des directions Γ_{ij} le long desquelles nous pouvons décomposer l'énergie de Dirichlet associée à la mesure uniforme sur la sphère. En effet,

$$\mathcal{E}_{\sigma_{n-1}}(f) = \int_{\mathbb{S}^{n-1}} f(-\Delta f) d\sigma_{n-1} = \frac{1}{2} \sum_{i,j=1}^{n} \int_{\mathbb{S}^{n-1}} (D_{ij} f)^2 d\sigma_{n-1}.$$

Ces opérateurs D_{ij} commutent avec le laplacien sphérique $\Delta = \frac{1}{2} \sum_{i,j=1}^n D_{ij}^2$ (autrement dit la propriété de commutation est vérifiée avec $\kappa = 0$). Enfin, la constante de Sobolev logarithmique est connue et vaut $n - 1$ (cf. [62, 16, 111, 17]).

Au niveau de la variance, cela fournit l'inégalité suivante

Théorème 4.4.2. *Pour toute fonction $f : \mathbb{S}^{n-1} \to \mathbb{R}$*

$$\mathrm{Var}_{\sigma_{n-1}}(f) \leq \frac{4e}{n} \sum_{i,j=1}^{n} \frac{\|D_{ij} f\|_2^2}{1 + \log(\|D_{ij} f\|_2 / \|D_{ij} f\|_1)}.$$

Remarque. L'inégalité de Poincaré pour la mesure uniforme sur la sphère $\mathbb{S}^{n-1}$, appliquée à la fonction $f(x) = \max x_{i=1,\ldots,n} = \sum_{i=1}^n x_i 1_{x_i = \max_{j=1,\ldots,n} x_j}$ entraine que

$$\mathrm{Var}(\max_{i=1,\ldots,n} U_i) \leq \frac{C}{n}$$

Le Théorème 3.2.5 suggère que $\mathrm{Var}(\max_{i=1,\ldots,n} U_i) \leq \frac{C}{n \log n}$, nous allons voir qu'une telle borne peut-être obtenue via l'extension de l'inégalité de Talagrand.

En fait, nous pouvons faire beaucoup mieux et obtenir une inégalité de concentration exponentielle reflétant la taille de la variance du maximum. Pour cela, il suffit de combiner le schéma de preuve de [62] avec certains arguments du théorème 4.2.2 de Tanguy.

Théorème 4.4.3. *Soit $n \geq 3$, sous le cadre précédent et en notant par $M_n = \max_{i=1,\ldots,n} U_i$ où $U = (U_1, \ldots, U_n)$ et $\mathcal{L}(U) = \sigma_{n-1}$ la mesure uniforme sur la sphère $\mathbb{S}^{n-1}$.*

$$\mathrm{Var}(e^{\theta M_n / 2}) \leq e \frac{\theta^2}{4n \log n} \mathbb{E}[e^{\theta M_n}], \quad \theta \in \mathbb{R}$$

En particulier, ceci entraine que $\mathbb{P}(|M_n - \mathbb{E}[M_n]| \geq t) \leq 6e^{-ct\sqrt{n \log n}}$ $t \geq 0$ avec $c > 0$ une constante numérique

Démonstration. La démonstration de l'article [62] entraine la majoration suivante de la variance, valable pour tout $T > \frac{1}{2(n-1)}$ et toute fonction f suffisamment régulière,

$$\mathrm{Var}_{\sigma_{n-1}}(f) \leq \frac{e}{n} \sum_{i,j=1}^{n} \int_0^T \int_{\mathbb{S}^{n-1}} (P_t D_{ij} f)^2 d\sigma_{n-1} dt.$$

□

Notons que si $f(x) = e^{\max_{i=1,\dots,n} x_i \theta/2}$, $\theta \in \mathbb{R}$, alors

$$D_{ij} = \frac{\theta}{2}(x_i 1_A j - x_j 1 A_i) e^{\frac{\theta}{2}\max_{i=1,\dots,n} x_i}, \quad \forall i,j \in \{1,\dots,n\},$$

avec $A_i = \{x_i = \max_{j=1,\dots,n} x_j\}, \quad i = 1,\dots,n$. Nous obtenons donc

$$\mathrm{Var}(e^{\frac{\theta}{2}M_n}) \leq \frac{2e}{n}\frac{\theta^2}{4}\sum_{i,j=1}^n \int_0^T \mathbb{E}\Big[P_t^2\big(U_i 1_{A_j} e^{\frac{\theta}{2}M_n}\big)\Big] + \mathbb{E}\Big[P_t^2\big(U_j 1_{A_i} e^{\frac{\theta}{2}M_n}\big)\Big] dt.$$

En outre, pour $i,j \in \{1,\dots,n\}$ fixés, par hypercontractivité du semi-groupe, nous trouvons que

$$\mathbb{E}\Big[P_t^2\big(U_i 1_{A_j} e^{\frac{\theta}{2}M_n}\big)\Big] \leq \mathbb{E}\Big[\big(U_i^p 1_{A_j} e^{\frac{p\theta}{2}M_n}\big)\Big]^{2/p},$$

avec $p = p(t) = 1 + e^{-2(n-1)t}$. Ensuite, l'inégalié de Hölder entraîne que

$$\mathbb{E}\Big[\big(U_i^p 1_{A_j} e^{\frac{p\theta}{2}M_n}\big)\Big]^{2/p} \leq \mathbb{E}\Big[\big(U_i^2 1_{A_j} e^{\theta M_n}\big)\Big]\sigma_{n-1}(A_j)^{(2-p)/p}.$$

De plus, $A_i = \{U_i = \max_{j=1,\dots,n} U_j\} = \{\frac{X_i}{|X|} = \max_{j=1,\dots,n} \frac{X_j}{|X|}\}$ avec X_i des gaussiennes standards indépendantes et $|X|$ la norme euclidienne du vecteur $X = (X_1,\dots,X_n)$. Donc, par symétrie, $\sigma_{n-1}(A_i) = 1/n$ pour tout $i = 1,\dots,n$.

Finalement,

$$\mathrm{Var}(e^{\frac{\theta}{2}M_n}) \leq \frac{e\theta^2}{2n}\bigg(\int_0^T \Big(\frac{1}{n}\Big)^{(2-p)/p} dt\bigg)\bigg(\sum_{i,j=1}^n \mathbb{E}\Big[\big(U_i^2 1_{A_j} e^{\theta M_n}\big)\Big] + \mathbb{E}\Big[\big(U_j^2 1_{A_i} e^{\theta M_n}\big)\Big]\bigg).$$

Pour conclure, remarquons que d'une part

$$\sum_{i,j=1}^n \mathbb{E}\Big[\big(U_i^2 1_{A_j} e^{\theta M_n}\big)\Big] = \sum_{j=1}^n \mathbb{E}\Big[1_{A_j} e^{\theta M_n}\Big] = \mathbb{E}\big[e^{\theta M_n}\big],$$

puisque $\sum_{i=1}^n U_i^2 = 1$ et que $(A_j)_{j=1,\dots,n}$ est une partition de $\mathbb{S}^{n-1}$. Tandis que d'autre part, pour $T = \frac{\log\big(2(n-1)\big)}{2(n-1)}$ (puisque $n \geq 3$, un tel T vérifie bien $T \geq \frac{1}{2(n-1)}$), on trouve que

$$\begin{aligned} I &= \int_0^T \Big(\frac{1}{n}\Big)^{(2-p)/p} dt \\ &= \int_{1+e^{-2(n-1)T}}^2 \Big(\frac{1}{n}\Big)^{\frac{2-v}{v}} \frac{dv}{2(n-1)(v-1)} \end{aligned}$$

en posant $v = 1 + e^{-2(n-1)t}$. Observons alors que $\frac{1}{2(n-1)(v-1)} \leq 1$ pour un tel choix de T, ainsi nous obtenons que

$$\int_0^T \Big(\frac{1}{n}\Big)^{(2-p)/p} dt \leq \int_1^2 \Big(\frac{1}{n}\Big)^{\frac{2-v}{v}} dv \leq \int_0^1 e^{-u\log n} du \leq \frac{1}{\log n}$$

.

Remarque. Notons que ce résultat fournit une inégalité de concentration non-asymptotique reflétant le théorème 3.2.5. En particulier, après intégration, cette inégalité entraine que

$$\mathrm{Var}(M_n) \leq \frac{C}{n \log n}$$

avec $C > 0$.

4.4.2 Lois log-concaves

Faisons quelques remarques vis à vis de l'efficacité et de l'utilité de l'inégalité 4.4.1 pour des mesures log-concaves. Le cas gaussien et la fonction maximum fournissent des exemples optimaux comme nous avons pu le voir plus tôt dans ce chapitre. Cependant, cette méthode ne fournit pas le bon ordre de grandeur pour la médiane. Similairement, le cas des lois log-concaves dont le potentiel est plus convexe que celui de la mesure gaussienne n'est pas satisfaisant. En effet, pour $V(x) = |x|^\alpha/\alpha$, $\alpha > 2$, l'inégalité (4.4.1) entraine que la variance est plus petite que $C/\log n$. Néanmoins, le résultat de convergence 3.1.6, suggère que la variance du maximum est encore plus petite que cette quantité. Nous verrons que notre approche de la superconcentration par la théorie du transport optimal permettra de régler ces problèmes et d'obtenir les bons ordres de grandeur.

Ajoutons cependant quelques remarques supplémentaires pour certaines mesures log-concaves, dont la dérivée seconde du potentiel ne peut être bornée inférieurement, uniformément, par une constante strictement positive. Par exemple, la mesure Gamma, vérifie les hypothèses pour appliquer le théorème 4.4.1. En effet, considérons la mesure Gamma sur $\mathbb{R}_+$, de paramètres $\alpha > 0$,

$$d\mu_\alpha(x) = \gamma_\alpha^{-1} x^{\alpha-1} e^{-x} dx,$$

rappelons qu'il s'agit de la mesure invariante de l'opérateur de Laguerre L_α défini par

$$L_\alpha f = x f'' + (\alpha - x) f'$$

Comme présenté dans [16], on peut calculer l'énergie de Dirichlet $\mathcal{E}_{\mu_\alpha}$ associée à L_α et obtenir $\mathcal{E}_{\mu_\alpha}(f) = \int_{\mathbb{R}_+} x f'^2 d\mu_\alpha$, pour f suffisamment régulière. En considérant la mesure produit μ_α^n dans $\mathbb{R}^n$ ceci entraine que $\mathcal{E}(f) = \sum_{i=1}^n x_i (\partial_i f)^2 d\mu_\alpha^n$, il est également connu (cf. [16]) que cette mesure satisfait une inégalité de Sobolev logarithmique (pour l'énergie de Dirichlet citée plus tôt), c'est-à-dire, pour f suffisament régulière,

$$\mathrm{Ent}_{\mu_\alpha^n}(f^2) \leq \sum_{i=1}^n \mathbb{E}_{\mu^\alpha}[x_i (\partial_i f)^2],$$

ce qui est équivalent à l'hypercontractivité du semi groupe associé. Remarquons de plus que l'énergie de Dirichlet se décompose suivant les directions $\Gamma_i(f) = \sqrt{x_i}(\partial_i f)$ et que celles-ci satisfont une relation de commutation $\Gamma_i(P_t f) \leq e^{-t} P_t(\Gamma_i(f))$. D'où, d'après le théorème 4.4.1

$$\mathrm{Var}_{\mu_\alpha^n}(f) \leq C \sum_{i=1}^n \frac{\|\Gamma_i(f)\|_2^2}{1 + \log\left(\|\Gamma_i(f)\|_2 / \Gamma_i(f)\|_1\right)},$$

avec $\Gamma_i(f) = \sqrt{x_i}\partial_i f$ et $\| \cdot \|_p$ désignent les normes L^p par rapport à la ν_α^n. En particulier, cette inégalité appliquée à $f(x) = \max_{i=1,\dots,n} x_i$ fournit la majoration suivante $\mathrm{Var}_{\mu_\alpha^n}(M_n) \leq C$, avec $M_n = \max_{i=1,\dots,n} X_i$ et $C > 0$. En effet, par indépendance

$$\|\Gamma_i(M_n)\|_2 = \sqrt{\frac{\mathbb{E}[M_n]}{n}}, \quad \|\Gamma_i(M_n)\|_1 = \frac{\mathbb{E}[M_n]}{n}$$

et par l'inégalité de Jensen

$$\|\Gamma_i(M_n)\|_2/\|\Gamma_i(M_n)\|_1 \geq \sqrt{n}.$$

Ainsi,

$$\mathrm{Var}(M_n) \leq C\frac{\mathbb{E}[M_n]}{\log n} = O(1).$$

Ce qui est en adéquation avec la théorie des extrêmes, puisque $M_n - b_n$, $b_n = \ln n + (\alpha-1)\ln\ln n - \ln\Gamma(\alpha)$ (ici Γ désigne la fonction Γ de Euler), converge en loi vers la distribution de Gumbel Λ_0.

En comparaison, la mesure Gamma satisfait aussi une inégalité de Poincaré par rapport à l'énergie de Dirichlet $\mathcal{E}_{\mu_\alpha}(f) = \int_{\mathbb{R}_+} x f'^2 d\mu_\alpha$. Ce qui implique, par tensorisation, pour toute fonction régulière

$$\mathrm{Var}_{\mu_\alpha^n}(f) \leq C \sum_{i=1}^{n} \int_{\mathbb{R}_+^n} x_i (\partial_i f)^2 d\mu_\alpha^n.$$

Et l'on constate que cette inégalité est sous optimale pour la fonction maximum, justifiant l'emploi des méthodes hypercontractives décrites ci-dessus. Toutefois, comme mentionné dans le chapitre un, un résultat de Bobkov *et al* (cf. [16]) affirme que toute mesure log-concave sur $\mathbb{R}^n$ satisfait une inégalité de Poincaré traditionnelle (par rapport à l'énergie de Dirichlet $\mathcal{E}_{\mu_\alpha^n}(f) = \int_{\mathbb{R}_+^n} |\nabla f|^2 d\mu_\alpha^n$). Autrement dit, pour toute fonction suffisamment régulière,

$$\mathrm{Var}_{\mu_\alpha^n}(f) \leq C \sum_{i=1}^{n} \int_{\mathbb{R}_+^n} [(\partial_i f)^2] d\mu_\alpha^n.$$

Il se trouve que cette inégalité fournie, plus simplement, le bon ordre de grandeur de la variance du maximum correspondant à la théorie des extrêmes et couvre l'exemple de la loi Gamma (et en particulier le cas exponentiel) mais aussi la loi de Gumbel (qui est également log-concave avec le potentiel $V(x) = x + e^{-x}$) pour laquelle le maximum recentré, $M_n - \ln n$, converge vers la loi de Gumbel. Remarquons également que l'inégalité de Poincaré, satisfaite par la loi Gamma et la loi de Gumbel, fournit immédiatement la décroissance exponentielle des queues de distributions du maximum, reflétant ainsi les asymptotiques de la loi de Gumbel. Plus précisément,

$$\mathbb{P}(|\max_{i=1,\dots,n} X_i - \mathbb{E}(\max_{i=1,\dots,n} X_i)| \geq t) \leq 6e^{-t/\sqrt{C_P}}, \quad t \geq 0,$$

avec $C_P > 0$ la constante de Poincaré, on obtient facilement la même inégalité avec le termes de recentrage b_n à la place de $\mathbb{E}(\max_{i=1,\dots,n} X_i)$.

Néanmoins, ces observations ne sont pas pertinentes vis à vis de la mesure uniforme sur un intervalle. Soient $U_1, \dots, U_n$ des variables aléatoires indépendantes, de même loi uniforme sur$[0,1]$ (le potentiel V associé, peut être défini comme suit $V(x) = 1$ si $x \in [0,1]$ et $V(x) = +\infty$ sinon). Rappelons (cf. chapitre trois) que la théorie des extrêmes fournit les constantes de renormalisation suivantes $a_n = n$ $b_n = 1$, afin que le maximum des U_i, $i = 1, \dots, n$ converge en loi vers la loi de Weibull (rappelons la définition de sa fonction de répartition $\Psi_1(x) = e^x$ pour

$x < 0$, $\Psi_a(x) = 1$ sinon). Ceci suggérant alors que $\mathrm{Var}(\max_{i=1,\dots,n} U_i) \leq 1/n$. Une telle borne ne pouvant être obtenue par des méthodes hypercontractives, il serait intéressant d'explorer plus encore le phénomène de superconcentration pour les mesures log-concaves et de proposer une démonstration unifiée pour obtenir le bon ordre de grandeur de la variance du maximum, puis les inégalités de déviation reflétant la convergence des extrêmes. Nous verrons que les méthodes de transports du chapitre sept permettent de répondre partiellement à ces questions.

4.5 Application en test statistique

Dans cette section nous allons illustrer l'inégalité de superconcentration (4.8), obtenue pour un échantillon gaussien, sur un problème de test statistique étudié dans [3]. Dans cet article les auteurs utilisent des résultats de concentration classique, à savoir (1.13), pour obtenir une région d'acceptation de leur test. En utilisant le matériel développé durant cette thèse, la conclusion de leur article peut-être renforcé dans certains cas sous la forme d'une borne de superconcentration.

En suivant le cadre et les notations de [3], nous observons un vecteur gaussien $X = (X_1, \dots, X_n)$ et étudions le test d'hypothèses suivant :

— Sous l'hypothèse nulle H_0, les coordonnées de X sont indépendantes et identiquement distribuées selon une loi normale standard. Notons alors par $\mathbb{P}_0$ et $\mathbb{E}_0$ respectivement les mesures de probabilités et les espérances sous jacentes, sous l'hypothèse H_0.

— Pour décrire l'hypothèse alternative H_1, considérons une classe $\mathcal{C} = \{S_1, \dots, S_N\}$ de N ensembles d'indices tels que $S_k \subset \{1, \dots, n\}$ pour tout $k = 1, \dots, N$. Sous H_1, il existe un ensemble $S \in \mathcal{C}$ tel que

$$X_i \text{ a pour loi} \begin{cases} \mathcal{N}(0,1) & \text{si } i \in S^c, \\ \mathcal{N}(\mu,1) & \text{si } i \in S, \end{cases}$$

avec $\mu > 0$ un paramètre strictement positif. Les coordonnées de X sont supposées indépendantes sous H_1 également. Pour tout $S \in \mathcal{C}$, notons alors par $\mathbb{P}_S$ et $\mathbb{E}_S$ respectivement les mesures de probabilités et les espérances sous jacentes, sous l'hypothèse H_1. Nous supposerons aussi que tous les ensembles $S \in \mathcal{C}$ possèdent le même nombre d'éléments $\mathrm{Card}(S) = K$.

Rappelons qu'un test est une fonction à valeurs binaire $f : \mathbb{R}^n \to \{0,1\}$. Si $f(X) = 0$, le test accepte l'hypothèse nulle, tandis qu'elle rejette H_0 si $f(X) = 1$. Comme dans [3], mesurons le risque associé au test f par

$$R(f) = \mathbb{P}_0\big(f(X) = 1\big) + \frac{1}{N} \sum_{S \in \mathcal{C}} \mathbb{P}_S\big(f(X) = 0\big).$$

Les auteurs de [3] cherchent à déterminer, ou au moins à estimer, la valeur de μ sous laquelle le risque peut-être rendu petit. Parmi d'autres résultats, ils utilisent un test basé sur le maximum, appellé le « test de scan », pour lequel ils montrent que

$$f(X) = 1 \iff 2 \max_{S \in \mathcal{C}} X_S \geq \mu K + \mathbb{E}_0\Big[\max_{S \in \mathcal{C}} X_S\Big]$$

où $X_S = \sum_{i \in S} X_i$ pour un ensemble $S \in \mathcal{C}$. Ils obtiennent le résultat suivant.

Proposition 4.5.1. *Le risque du test f, associé au maximum, satisfait $R(f) \leq \delta$ à condition que*

$$\mu \geq \frac{1}{K} \mathbb{E}_0 \left[\max_{S \in \mathcal{C}} X_S \right] + 2\sqrt{\frac{2}{K} \log \frac{2}{\delta}}.$$

En utilisant (4.8), cette borne peut-être améliorée en reflétant l'ordre de grandeur correct de la variance du maximum d'un vecteur gaussien standard.

Proposition 4.5.2. *Pour tout $n \geq 2$, le risque du test f vérifie $R(f) \leq \delta$ lorsque*

$$\mu \geq \frac{1}{K} \mathbb{E}_0 \left[\max_{S \in \mathcal{C}} X_S \right] + \log \frac{6}{\delta} \times \frac{2}{c\sqrt{K \log N}}$$

avec $c > 0$ une constante numérique.

Démonstration. Nous suivons la démonstration de [3] pour obtenir simultanément, pour tout $t \geq 0$,

$$\mathbb{P}_0 \left(\max_{S \in \mathcal{C}} X_S \geq \mathbb{E}_0 \left[\max_{S \in \mathcal{C}} X_S \right] + t \right) \leq 3e^{-ct\sqrt{\log(N)/K}}$$

et, pour tout $S \in \mathcal{C}$,

$$\begin{aligned} \mathbb{P}_S \left(\max_{S' \in \mathcal{C}} X_{S'} \leq \mu K - t \right) &\leq \mathbb{P}_S \left(\max_{S' \in \mathcal{C}} X_{S'} \leq \mathbb{E}_{S'} \left[\max_{S' \in \mathcal{C}} X_{S'} \right] - t \right) \\ &\leq 3e^{-ct\sqrt{\log(N)/K}} \end{aligned}$$

puisque, sous H_1, pour un ensemble $S \in \mathcal{C}$ fixé, $\mu K = \mathbb{E}_S [X_S] \leq \mathbb{E}_S [\max_{S' \in \mathcal{C}} X_{S'}]$ et X_S est de même loi que n'importe quel $X_{S'}$. Posons alors t tel que

$$2t = \mu K - \mathbb{E} \left[\max_{S \in \mathcal{C}} X_S \right]$$

qui est positif d'après les hypothèses. A l'aide des inégalités précédentes, nous obtenons une majoration du risque

$$R(f) \leq 6e^{-ct\sqrt{\log(N)/K}}$$

qui permet d'obtenir $R(f) \leq \delta$ comme annoncé dans la proposition. □

La nouvelle borne obtenue dans la proposition 4.5.2 est aussi bonne que celle de la proposition 4.5.1 lorsque $\delta \simeq 1/N^\alpha$ avec $\alpha > 0$. Elle est meilleure que celle de la proposition 4.5.1 lorsque $\delta \simeq 1/\log^\alpha(N)$ avec $\alpha > 0$. Néanmois, la proposition 4.5.1 est plus performante que la proposition 4.5.2 lorsque $\delta \simeq e^{-N^\alpha}$, avec $\alpha > 0$. Cependant, il n'est peut-être pas pertinent de considérer un risque si petit.

Chapitre 5

Extremalité

Ce chapitre reprend une thématique présente dans le livre de Chatterjee [48]. Dans son ouvrage, Chatterjee propose une méthode alternative à l'hypercontractivité pour obtenir de la superconcentration. Il s'agit de la notion d'extremalité, reposant sur la taille de l'espérance du maximum. Après une brève introduction à cette notion, nous proposons d'utiliser des outils hypercontractifs pour étendre, au niveau exponentiel, les résultats de Chatterjee. Une inégalité de concentration similaire à été obtenue dans [69], nous comparons notre résultat au leur en dernier lieu.

5.1 Introduction

Le cadre est le suivant : nous considérons un vecteur gaussien centré $X = (X_1, \dots, X_n)$ dans $\mathbb{R}^n$, $n \geq 1$ de matrice de covariance Γ et nous supposons que $\mathrm{Var}(X_i) = 1, \forall i \in \{1, \dots, n\}$. A nouveau, le maximum des coordonnées du vecteur X sera noté par $M_n = \max_{i=1,\dots,n} X_i$. D'après Chatterjee, [48], le vecteur X est dit extremal si $\mathbb{E}[M_n]$ se comporte comme le cas indépendant, autrement dit $\mathbb{E}[M_n] \simeq \sqrt{2\log n}$. De nombreux modèles gaussiens, présentés dans le chapitre deux, vérifient cette propriété : modèles de percolation, le modèle de verres de spin de Sherrington et Kirpatrick, le champ libre gaussien discret sur $\mathbb{Z}^2$, suite gaussienne stationnaire,... (cf. [48]). Par la suite, nous noterons par $\alpha \in]0,1[$ le rapport

$$\frac{\mathbb{E}[M_n]}{\sqrt{2\log n}} = \alpha.$$

Dans son livre [48], Chatterjee a obtenu le théorème suivant

Théorème 5.1.1. *Pour $n \geq e$ et tout $\alpha \in]0,1[$, nous avons la majoration suivante*

$$\mathrm{Var}(M_n) \leq C\left[\sqrt{1-\alpha} + \left(\frac{\log\log n}{\log n}\right)^{1/4}\right]$$

avec $C > 0$.

Remarque. Cette approche est une manière alternative à la méthode hypercontractive pour améliorer l'inégalité de Poincaré satisfaite par la mesure gaussienne γ_n sur $\mathbb{R}^n$. Rappelons, une fois de plus, que cette inégalité fonctionnelle s'exprime commes suit, pour $f : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulière,

$$\mathrm{Var}_{\gamma_n}(f) \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n. \tag{5.1}$$

Appliquée à la fonction $f(x) = \max_{i=1,\dots,n} x_i$, elle fournie

$$\mathrm{Var}(M_n) \leq 1.$$

Comme nous le savons, ceci est loin d'être optimal. Chatterjee suggère dans son livre que l'extremalité du vecteur X entraine que M_n est superconcentré. En effet, si $\alpha = 1 - \epsilon_n$ avec $\epsilon_n \to 0$ lorsque $n \to \infty$ alors le théorème (5.1.1) entraine la majoration suivante , pour tout $n \geq e$,

$$\mathrm{Var}(M_n) \leq C \max\left(\sqrt{\epsilon_n}, \left(\frac{\log\log n}{\log n}\right)^{1/4}\right).$$

Cette borne étant meilleure que celle obtenue via (5.1), nous obtenons donc de la superconcentration pour M_n.

Cette notion d'extrémalité est également utilisée dans l'article [69], dans lequel les auteurs obtiennent l'inégalité de concentration suivante

Théorème 5.1.2. *Pour tout $0 \leq t \leq C_\alpha \sqrt{\log n}$, avec $C_\alpha = \alpha/100$, $\alpha \in]0,1[$ et $n \geq 2$, on a*

$$\mathbb{P}\big(|M_n - \mathbb{E}[M_n]| \geq t\big) \leq C e^{-\frac{t^2}{2[1-c(\alpha)]}}, \tag{5.2}$$

avec $c(\alpha) = 2\alpha^2 10^{-6}$, $C > 0$.

Remarque. Il est important de faire quelques commentaires sur le théorème précédent. Celui-ci doit être comparé avec la concentration gaussienne classique (1.13) (cf. [112]). Celle-ci fournissant, pour le choix de la fonction lipschitzienne $f(x) = \max_{i=1,\dots,n} x_i$, le résultat suivant

$$\mathbb{P}\big(|M_n - \mathbb{E}[M_n]| \geq t\big) \leq 2e^{-t^2/2}, \quad t \geq 0. \tag{5.3}$$

Il est évident que l'inégalité (5.2) est meilleure que celle fournie par la concentration classique (cf. par exemple (5.3)) pour les petites valeurs de t (*i.e.* lorsque $t \leq C\sqrt{\log n}$). Ceci suggère donc que l'extremalité de X renforce la concentration du maximum. Nous comparerons à la fin de cette section l'inégalité (5.2) avec notre résultat.

En combinant des arguments hypercontractifs à la démonstration de Chatterjee (du théorème 5.1.1), nous obtenons le théorème suivant

Théorème 5.1.3. *Sous le cadre précédent, pour tout $\theta \in \mathbb{R}$ et tout $n \geq e$, nous avons*

$$\mathrm{Var}(e^{\theta M_n/2}) \leq \frac{\theta^2}{4} C\left(\sqrt{1-\alpha} + \left(\frac{\log\log n}{\log n}\right)^{1/4}\right)\mathbb{E}[e^{\theta M_n}].$$

En particulier, cela entraine que, pour tout $t \geq 0$,

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| \geq t) \leq 6e^{-ct/\sqrt{\sqrt{1-\alpha}+\epsilon_n}}, \tag{5.4}$$

avec $c > 0$ et $\epsilon_n = \frac{1}{\log\log n}$.

Remarque. On constate facilement que l'intégration de (5.4) entraine, pour $n \geq e$,

$$\mathrm{Var}(M_n) \leq C\left(\sqrt{1-\alpha} + \frac{1}{\log \log n}\right)$$

qui est similaire au résultat de Chatterjee (théorème 5.1.1) et montre que l'extremalité entraine de la superconcentration. Toutefois cette méthode est clairement non-optimale, comme on peut le constater sur l'exemple classique du maximum des coordonnées d'un vecteur gaussien indépendant, et n'entraine qu'une légère amélioration de l'inégalité de Poincaré. Néamoins, elle conserve un intérêt puisqu'il est en général difficile d'obtenir des bornes optimales sur la variance de M_n (rappelons par exemple le cas du champ libre discret gaussien sur $\mathbb{Z}^2$ et les articles de Zeitouni *et al.*), tandis que l'estimation de l'espérance de M_n peut-être obtenue de manière plus algorithmique. Nous renvoyons pour cela le lecteur, lorsqu'une structure d'arbre est sous-jacente, au cours de Kistler [104] qui expose de quelle manière les méthodes de ≪ second moment modifiée ≫ peuvent facilement fournir des asympotiques précises de $\mathbb{E}[M_n]$. D'un point de vue plus théorique, l'étude de supremum de variables gaussiennes à été achevée par Talagrand dans [153, 156] via des méthodes de chainage et de mesures majorantes.

Le reste de ce chapitre est organisé de la manière suivante : tout d'abord nous présentons la démonstration du théorème 5.1.3, ensuite nous ferons un rapide survol de la littérature concernant les champs gaussiens extremaux et comparerons notre résultat à celui de [69].

5.2 Démonstration du théorème 5.1.3

La tâche principale de la démonstration est d'adapter les arguments de Chatterjee au niveau exponentiel, le début de la preuve est donc similaire au résultat principal de [157]. Nous proposerons ensuite un lemme permettant d'utiliser les estimations obtenues par Chatterjee dans [48] pour conclure.

Soit $n \geq e$, on pose $\delta = \frac{\sqrt{\log(m_n^4/2)}}{m_n} \in (0,1)$ et $r_0 = r_0(t) = \sqrt{1-\gamma^2}$, pour tout $t \geq 0$ (avec le même choix de γ que Chatterjee, cf. [48], Chapitre 8 page 76) . Comme annoncé, le schéma de preuve suit [48]. Le point de départ est la formule de représentation

$$\mathrm{Var}(f) = \int_0^\infty e^{-t}\mathbb{E}\left[\nabla f \cdot P_t \nabla f\right] dt$$

pour la variance d'une fonction (régulière) $f : \mathbb{R}^n \to \mathbb{R}$ le long du semi-groupe d'Ornstein-Uhlenbeck $(P_t)_{t\geq 0}$. Rappelons les ouvrages de référence [17, 16] concernant le semi groupe d'Ornstein-Uhlenbeck et des semi groupes de Markov plus généraux, ainsi que des méthodes d'interpolations.

En suivant [48], étant donné un vecteur gaussien centré X de matrice de covariance $\Gamma = (\Gamma_{ij})_{1\leq i,j\leq n}$, on applique (4.10) à (une approximation régulière) $f = e^{\theta M/2}$ où $\theta \in \mathbb{R}, F(x) = \max_{i=1,\dots,n}(Mx)_i$ et $\Gamma = M\,{}^tM$. Cela entraine, que pour tout $\theta \in \mathbb{R}$,

$$\mathrm{Var}\left(e^{\theta M_n/2}\right) = \frac{\theta^2}{4}\int_0^\infty e^{-t}\,\mathbb{E}\left[\sum_{i,j=1}^n \Gamma_{ij}\,1_{\{X\in A_i\}}e^{\theta M_n/2}1_{\{X^t\in A_j\}}e^{\theta M_n^t/2}\right]dt.$$

Ici $\{X \in A_i\} = \{I = i\}$, pour $i = 1,\dots,n$, et $X^t = e^{-t}X + \sqrt{1-e^{-2t}}\,Y$, où Y est une copie indépendante de X, avec son maximum correspondant M_n^t.

$$\mathbb{E}\left[|Q_t f(X)|^q\right]^{1/q} = \|P_t g\|_q \leq \|g\|_p = \mathbb{E}\left[|f(X)|^p\right]^{1/p},$$

où $\|\cdot\|_r$, pour tout $r \geq 1$, désigne la norme L^r par rapport à la mesure gaussienne standard sur $\mathbb{R}^n$.
De plus, pour tout $t \geq 0$, avec $I^t = \mathrm{argmax}_{i=1,\dots,n} X_i^t$,

$$\begin{aligned}
\mathcal{I} &= \mathbb{E}\Big[\sum_{i,j=1}^n \Gamma_{ij}\, 1_{\{X\in A_i\}} e^{\theta M_n/2} 1_{\{X^t\in A_j\}} e^{\theta M_n^t/2}\Big] \\
&= \mathbb{E}\Big[e^{\theta(M_n+M_n^t)/2} \sum_{i,j=1}^n \Gamma_{ij}\, 1_{\{I=i\}} 1_{\{I^t=j\}}\Big] \\
&= \mathbb{E}\left[e^{\theta\left(M_n+M_n^t\right)/2}\, \Gamma_{II^t}\right] \\
&= \sum_{k\geq 0} \mathbb{E}\left[e^{\theta\left(M_n+M_n^t\right)/2}\, \Gamma_{II^t} 1_{\{2^{-k-1}\leq \Gamma_{II^t}\leq 2^{-k}\}}\right].
\end{aligned}$$

Pour alléger les notations, on notera Γ_{II^t} par Γ et on posera $F^t = M_n + M_n^t$. Let $k_0 = \min\{k \geq 0\,; r_0 \leq 2^{-k-1}\}$. En coupant la somme précédente en deux, on obtient

$$\begin{aligned}
\mathcal{I} &\leq \sum_{k=0}^{k_0} 2^{-k}\, \mathbb{E}\left[e^{\theta F^t/2} 1_{\{\Gamma\geq r_0\}}\right] + \sum_{k=k_0+1}^{\infty} 2^{-k}\, \mathbb{E}\left[e^{\theta F^t/2} 1_{\{2^{-k-1}\leq\Gamma\leq 2^{-k}\}}\right] \\
&\leq 2 \sum_{D\in\mathcal{C}(r_0)} \mathbb{E}\left[e^{\theta F^t/2} 1_{\{I,I^t\in D\}}\right] + \sum_{k\geq 0} r_0\, \mathbb{E}\left[e^{\theta F^t/2} 1_{\{2^{-k-1}\leq\Gamma\leq 2^{-k}\}}\right] \\
&= 2 \sum_{D\in\mathcal{C}(r_0)} \mathbb{E}\left[e^{\theta F^t/2} 1_{\{I,I^t\in D\}}\right] + r_0\, \mathbb{E}\left[e^{\theta F^t/2}\right].
\end{aligned}$$

Puis, d'après l'inégalité de Cauchy-Schwartz et l'invariance rotationnelle gaussienne

$$\mathbb{E}\left[e^{\theta F^t/2}\right] \leq \mathbb{E}\left[e^{\theta M_n}\right].$$

Par simplicité de notation, nous désignerons par $m_n = \mathbb{E}[M_n]$. Pour contrôler Le deuxième terme

$$\sum_{k=0}^{k_0} 2^{-k}\, \mathbb{E}\left[e^{\theta F^t/2} 1_{\{\Gamma\geq r_0\}}\right],$$

nous introduisons, comme Chatterjee, les ensembles aléatoires suivants, pour tout $0 \leq r \leq 1$ et tout $t \geq 0$ fixés, $B_r = \{i\,; \Gamma_{Ii} \geq r\}$ et $D_\delta = \{i\,; |X_i - e^{-t} m_n| \leq \delta M_n\}$, où $\delta = \frac{\sqrt{\log(m_n^4/2)}}{m_n} \in]0,1[$. Remarquons que D_δ correspond aux indices $1 \leq i \leq n$ pour lequels X_i est proche de $e^{-t} m_n$. Notons également que $B_r \subset B_{r'}$ si $r \leq r'$. On pose à présent $B = B_{r_0(t)}$ et $D = D_\delta$.

Alors, en suivant la preuve de Chatterjee,

$$\begin{aligned}
\mathbb{E}[e^{\theta F^t/2} 1_{\{\Gamma_{II^t}\geq r_0\}}] &= \mathbb{E}[e^{\theta F^t/2} 1_{I^t\in B}] \\
&\leq \mathbb{E}[e^{\theta F^t/2} 1_{I^t\in B\cap D}] + \mathbb{E}[e^{\theta F^t/2} 1_{I^t\notin D}]
\end{aligned}$$

Il reste deux étapes pour achever la démonstration, celles-ci sont présentées dans deux lemmes différents. Le premier correspond à une adaptation, que nous avons obtenu, au niveau exponentiel des arguments de Chatterjee, le deuxième correspond à certaines majorations obtenues dans [48] que nous employerons pour achever la preuve.

Lemme 5.2.1. *Soit A un borélien de $\mathbb{R}^n$ alors, pour tout $\theta \in \mathbb{R}$ et tout $t \geq 0$,*

$$\mathbb{E}[e^{\theta M_n/2} Q_t(e^{\theta M_n/2} 1_{\{I \in A\}})] \leq \mathbb{P}(I^t \in A)^{(2-p)/p} \mathbb{E}[e^{\theta M_n}],$$

avec $p = p(t) = 1 + e^{-t} < 2$.

Démonstration. La preuve repose sur l'inégalité de Hölder et la propriété d'hypercontractivité de $(Q_t)_{t \geq 0}$. En effet, appliquons l'inégalité d'Hölder avec $\frac{1}{p} + \frac{1}{q} = 1$

$$\mathbb{E}\Big[e^{\theta M_n/2} Q_t\big(e^{\theta M_n/2} 1_{\{I \in A\}}\big)\Big] \leq \mathbb{E}[e^{\theta p M_n/2}]^{1/p} \mathbb{E}\Big[Q_t^q\Big(e^{\theta M_n/2} 1_{\{I \in A\}}\Big)\Big]^{1/q}.$$

Alors, par la propriété d'hypercontractivité,

$$\mathbb{E}[Q_t^q(e^{\theta M_n/2} 1_{\{I \in A\}})]^{1/q} \leq \mathbb{E}[e^{\theta p M_n/2} 1_{\{I \in A\}}],$$

sous la condition que $e^{2t} = \frac{q-1}{p-1}$, c'est à dire $p = 1 + e^{-t} < 2$. En conséquence,

$$\mathbb{E}[e^{\theta M_n/2} Q_t(e^{\theta M_n} 1_{\{I \in A\}})] \leq \mathbb{E}[e^{\theta p M_n/2} 1_{\{I \in A\}}] \mathbb{E}[e^{\theta p M_n/2}]^{1/p}$$

et une nouvelle application de l'inégalité d'Hölder sur le membre de droite mène à

$$\mathbb{E}\Big[e^{\theta M_n/2} Q_t\big(e^{\theta M_n/2} 1_{\{I \in A\}}\big)\Big] \leq \mathbb{P}(I \in A)^{(2-p)/p} \mathbb{E}[e^{\theta M_n}] = \mathbb{P}(I^t \in A)^{(2-p)/p} \mathbb{E}[e^{\theta M_n}],$$

puisque I et I^t sont de même loi pour tout $t \geq 0$. □

Le lemme précédent, entraine alors que

$$\mathbb{E}\big[e^{\theta F^t/2} 1_{\{\Gamma_{II^t} \geq r_0\}}\big] \leq \mathbb{E}[e^{\theta F^t/2}] \mathbb{P}(I^t \in B \cap D)^{(2-p)/p} + \mathbb{E}[e^{\theta F^t/2}] \mathbb{P}(I^t \notin D)^{(2-p)/p}$$

Il suffit d'utiliser des estimées obtenues par Chatterjee pour terminer la preuve. Nous les résumons dans le lemme suivant.

Lemme 5.2.2. *(Chatterjee) En conservant les notations précédentes, nous avons les majorations suivantes, pour tout $t \geq 0$,*

1. $\mathbb{P}(I^t \in B \cap D) \leq 3e^{-\delta^2 m_n^2/8}$
2. $\mathbb{P}(I^t \notin D) \leq 4e^{-\delta^2 m_n^2/4}$

Ainsi, le choix de δ et le lemme précédent entraine que

$$\mathbb{E}[e^{\theta F^t/2} 1_{\{\Gamma_{II^t} \geq r_0\}}] \leq \mathbb{E}[e^{\theta M_n}] \left(\frac{C}{\sqrt{m_n}}\right)^{(2-p)/p},$$

avec $C = 7 \times 2^{1/8}$. Ceci fournit la majoration suivante, pour tout $t \geq 0$,

$$\mathbb{E}\big[e^{\theta M_n/2}\Gamma_{II^t}\big] \leq \left(r_0(t) + \left(\frac{C}{\sqrt{m_n}}\right)^{(2-p)/p}\right)\mathbb{E}[e^{\theta M_n}].$$

Jusqu'à présent, nous avons établit

$$\mathrm{Var}(e^{\theta M_n/2}) \leq \left[\int_0^\infty e^{-2t}\left(r_0(t) + \left(\frac{C}{\sqrt{m_n}}\right)^{(2-p)/p}\right)dt\right]\mathbb{E}[e^{\theta M_n}].$$

De plus, pour tout $t \geq 0$, $r_0(t) \leq \frac{\sqrt{1-\alpha^2}}{\sqrt{1-e^{-2t}}} + \frac{1}{\sqrt{1-e^{-2t}}}\left(\frac{C \log\log n}{\log n}\right)^{1/4}$. Ainsi,

$$\mathrm{Var}(e^{\theta M_n/2}) \leq \left(\sqrt{1-\alpha} + \left(\frac{C\log\log n}{\log n}\right)^{1/4} + \int_0^1 \left(\frac{C}{\sqrt{m_n}}\right)^t dt\right)\mathbb{E}[e^{\theta M_n}].$$

Ce qui conclut la preuve, puisque $\int_0^1 \left(\frac{C}{\sqrt{m_n}}\right)^t dt \leq \frac{C}{\log\log n}$.

5.3 Comparaison

Tout d'abord nous allons faire quelques commentaires sur l'étude du phénomène d'extremalité, il est important de comparer ceci à l'approche hypercontractive développée durant cette thèse. Nous comparerons également le théorème 5.1.3 avec le résultat (5.2) de [69].

5.3.1 Hypercontractivité et extrémalité

L'approche hypercontractive à l'avantage de ne pas utiliser explicitement $\mathbb{E}[M_n]$ pour obtenir une borne sur la variance. En effet, la formule

$$\mathrm{Var}(f) = \int_0^\infty e^{-t}\mathbb{E}\left[\nabla f \cdot P_t \nabla f\right] dt,$$

où $f(x) = \max_{i=1,\ldots,n} x_i$, n'utilise pas $\mathbb{E}[M_n]$ mais plutôt les ensembles $A_i = \{X_i = \max_j X_j\}$, pour $i = 1,\ldots,n$, intervenant dans les dérivées partielles de f. Ces ensembles interviennent dans la majoration obtenue par hypercontracivité, sous la forme de $\mathbb{P}(A_i)^p$, $p \in \{1/2, 1\}$ (voir (2.3.3) ou (1.3.6)). En pratique, il est donc nécessaire de majorer uniformément en $i = 1,\ldots,n$ $\mathbb{P}(A_i)$ pour obtenir une borne sur $\mathrm{Var}(M_n)$. Ceci peut-être fait de la manière suivante, pour tout $1 \leq i \leq n$, la concentration gaussienne classique entraine que

$$\mathbb{P}(A_i) = \mathbb{P}(X_i = \max_{j=1,\ldots,n} X_j) \leq 2e^{-\mathbb{E}[M_n]^2/2}$$

Ainsi, si $\mathbb{E}[M_n] \geq \alpha\sqrt{2\log n}$, pour $0 < \alpha < 1$ (par exemple, ceci peut-être obtenue via le principe de minoration de Sudakov) alors $\mathbb{P}(A_i) \leq 2/\log(n^{\alpha^2})$, $\forall i \in \{1,\ldots,n\}$. Bien que ceci soit réminiscent du phénomène d'extrémalité, le coeur de la méthode reste basé sur l'hypercontractivité du semi groupe d'Ornstein-Uhlenbeck et la représentation dynamique de la variance le long de celui-ci.

L'approche de la superconcentration de Chatterjee via l'extrémalité repose sur deux ingrédients : la taille de $\mathbb{E}[M_n] \sim \sqrt{2\log n}$, la représentation dynamique de la variance et les outils de concentration de la mesure classique. L'avantage de ceci est de proposer une méthode alternative à l'hypercontractivité pour obtenir de la superconcentration, autrement dit une amélioration de l'inégalité de Poincaré. Pour étendre le théorème 5.1.1 au niveau exponentiel, nous avons du recourir à l'hypercontractivié et nos estimées fournissent un facteur $1/\log\log n$ sous optimal (si l'on considère l'exemple de n variables aléatoires gaussiennes indépendantes). Toutefois, le théorème 5.1.1 était, lui aussi, initialement sous-optimal (pour le même exemple). Il est donc préférable de voir ces résultats de la manière suivante : si l'espérance du maximum se comporte comme le cas indépendant il est possible d'améliorer légérement les bornes fournies par la concentration classique. Bien que sous-optimal, nous allons tout de même voir que notre résultat reflète, partiellement, l'asymptotique de la loi limite des extrêmes dans le cas indépendant.

5.3.2 Comparaison avec l'inégalité 5.1.2 de Ding et al.

Nous allons à présent comparer notre résultat avec celui de [69] pour α fixé et lorsque $\alpha \to 1$. Nous allons voir que dans le second cas notre inégalité conserve, contrairement à la leur, un sens. Pour que la comparaison soit plus visible, nous rappelons les deux inégalités mise jeu.

$$\mathbb{P}(|M_n - m_n| \geq t) \leq 6e^{-ct/\sqrt{K_n^\alpha}}, \quad t \geq 0, \tag{5.5}$$

avec $K_n^\alpha = \sqrt{1-\alpha} + \frac{1}{\log\log n}$.

et celle de [69] , pour tout $0 \leq t \leq C_\alpha\sqrt{\log n}$, avec $C_\alpha = \alpha/100$,

$$\mathbb{P}(|M_n - m_n| \geq t) \leq Ce^{-t^2/(2-c(\alpha))}, \tag{5.6}$$

où $c(\alpha) = 2\alpha^2 10^{-6}$.

Soient $\alpha \in]0,1[$ et $t \in [0, C_\alpha\sqrt{\log n}]$ fixés. On constate facilement que (5.6) est toujours meilleure que (5.5). En effet, l'intervalle $I \subset [0, C_\alpha\sqrt{\log n}]$ pour lequel (5.5) est meilleur que (5.6) est le suivant

$$0 \leq t \leq \frac{c\big[2-c(\alpha)\big]}{K_n^\alpha} = O(1)$$

Cependant, contrairement à (5.5), l'inégalité (5.6) n'implique pas le résultat de Chatterjee (théorème 5.1.1). En effet, après intégration, (5.6) entraine que

$$\mathrm{Var}(M_n) \leq \big[2 - c(\alpha)\big](1 - n^{-\eta}) + n^{-\eta},$$

avec $\eta = \frac{C_\alpha^2}{2-c(\alpha)}$. Ce résultat est très différent de celui de Chatterjee et est du même ordre de grandeur que la borne fournie par l'inégalité de Poincaré. Tandis que l'inégalité (5.6), entraine que

$$\mathrm{Var}(M_n) \leq C\left(\sqrt{1-\alpha} + \frac{1}{\log\log n}\right),$$

avec $C > 0$. Ceci est de même nature que la borne obtenue par Chatterjee (même si celle que nous obtenons est légèrement plus grande).

Supposons, à présent, que $\alpha = \alpha_n \to 1$, lorsque $n \to \infty$, ceci pouvant s'exprimer par $\alpha = \alpha_n = 1 - \epsilon_n$ avec $\epsilon_n \to 0$ lorsque $n \to \infty$. Typiquement, $\epsilon_n = C\frac{\log\log n}{\log n}$ avec $C > 0$. L'inégalité (5.5) devient alors,

$$\mathbb{P}(|M_n - m_n| \geq t) \leq 6e^{-ct/\sqrt{\max(\epsilon_n^{1/2},\delta_n)}},\ t \geq 0, \tag{5.7}$$

avec $\delta_n = \frac{1}{\log\log n}$. Tandis que l'inégalité (5.6) s'exprime par

$$\mathbb{P}(|M_n - m_n| \geq t) \leq Ce^{-t^2/(2(1-10^{-6})},\ 0 \leq t \leq C\sqrt{\log n}, \tag{5.8}$$

avec $C = 2/100 > 0$.

Il semble plus pertinent de comparer ces deux inégalités sur des exemples explicites, faisons le pour le cas indépendant (ce qui va suivre serait aussi valable pour une suite gaussienne stationnaire dont la fonction de covariance décroit aussi vite que $1/\log n$ lorsque $n \to \infty$). Plus précisément, on considère $X_1, \ldots, X_n$ des variables aléatoires gaussiennes standards indépendantes et de même loi. Rappelons les faits suivants, provenant de la théorie des extrêmes,

$$\sqrt{2\log n}(M_n - b_n) \to \Lambda_0, \quad n \to \infty,$$

en loi et, au vu de la taille de $m_n = \mathbb{E}[M_n]$, $\alpha = 1 - \frac{\log\log n}{2\log n} + o\left(\frac{1}{\log\log n}\right)$. De plus, nous avons prouvé l'inégalité de concentration suivante

$$p(\sqrt{\log n}|M_n - m_n| \geq t) \leq 6e^{-ct}, \tag{5.9}$$

Notre nouvelle inégalité (5.5) se réécrit de la manière suivante

$$\mathbb{P}(\sqrt{\log n}|M_n - m_n| \geq t) \leq 6e^{-ct\sqrt{\log\log n}/\sqrt{\log n}}. \tag{5.10}$$

Il est clair que la dépendance en n est mauvaise dans l'inégalité précédente, toutefois elle conserve du sens si on prend $t = c\sqrt{\log n}$. De plus, elle présente les mêmes asymptotiques que la loi de Gumbel avec une décroissance exponentielle. Similairement, nous pouvons réexprimer (5.6) par

$$\mathbb{P}(\sqrt{\log n}|M_n - m_n| \geq t) \leq Ce^{-t^2/(2-c(1))\log n}.$$

Il est évident que la dépendance en n est plus mauvaise encore que celle de l'inégalité (5.10). De plus, la décroissance gaussienne ne reflète pas le comportement des queues de distributions de la loi de Gumbel et l'inégalité précédente est vide de sens si on l'évalue en $t = c\sqrt{\log n}$.

Nous pourrions faire le même raisonnement avec le champ libre gaussien discret sur $\mathbb{Z}^2$. Rappelons que Ding et Zeitouni ont obtenu les résultats suivants :

$$\mathbb{E}[M_n] = m_n + O(1),$$

avec $m_n = 1\sqrt{2/\pi}(\log n - 3\log\log n/8) + O(1)$. Ainsi que les inégalités de déviations suivantes, valables pour tout $0 \leq t \leq (\log n)^{2/3}$,

$$Ce^{-ct} \leq \mathbb{P}(M_n \geq m_n + t) \leq Ce^{-ct}$$

et

$$Ce^{-Ce^{Ct}} \leq \mathbb{P}(M_n \leq m_n - t) \leq Ce^{-ce^{ct}},$$

avec $C, c > 0$.

En conclusion, l'inégalité (5.5) est meilleure que celle obtenue dans [69] lorsque l'on fait tendre α vers 1. Néanmoins, elle reste sous-optimale lorsqu'on l'applique à des exemples connus et ne semble utile qu'en théorie. A ce sujet, comme dans [69], si l'on considère un vecteur gaussien dans $\mathbb{R}^n$, dont l'espérance du maximum se comporte de la même manière que le cas indépendant, est-il possible d'améliorer significativement les résultats fournis par la théorie classique de la concentration ?

Chapitre 6

Inégalité de Talagrand d'ordre supérieur, critère de courbure dimension inverse

Cette partie s'inspire d'un résultat de Ledoux sur le développement de la variance de la mesure gaussienne standard via des développements de Taylor [110] et d'un résultat de Chatterjee portant sur une méthode d'interpolation par semi-groupe. Nous verrons que le schéma de preuve permet d'obtenir une inégalité de Talagrand aux ordres supérieurs sur le cube discret $C_n = \{-1,1\}^n$ et nous permettra de proposer une application sur l'influence d'une coordonnée en analyse booléenne. Les résultats que nous allons présenter sont une partie majeure des travaux de cette thèse, un article reprenant ce chapitre est en cours de rédaction.

6.1 Introduction

Sommairement, dans son article, Ledoux combine des développements de Taylor sur l'intervalle $[0,t]$ avec des méthodes d'interpolation pour retrouver un développement de la variance. Plus précisément, il obtient la formule (déjà présente dans l'article [93]) suivante.

Proposition 6.1.1. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ suffisament régulière, de classe $C^p(\mathbb{R}^n)\, p \geq 1$ dont les dérivées partielles successives appartiennent à $L^2(\gamma_n)$, on obtient alors la formule suivante*

$$\mathrm{Var}_{\gamma_n}(f) = \sum_{k=1}^{p-1} \frac{(-1)^k}{k!} \int_{\mathbb{R}^n} |\nabla^k f|^2 d\gamma_n - \frac{(-1)^{p-1}}{(p-1)!} \int_0^\infty 2e^{-2pt} \int_{\mathbb{R}^n} |P_t(\nabla^p f)|^2 d\gamma_n dt.$$

Tandis qu'une relecture des travaux de Chatterjee, présentés dans le chapitre dix de son livre [48], permet d'extraire un procédé d'interpolation par semi groupes entre $[t,\infty[$. L'utilisation de ses deux arguments va nous permettre d'obtenir une décomposition de la variance originale. Nous verrons par la suite que les travaux de Chatterjee sont pertinents dans le cadre des modèles de verres de spins et s'apparentent à un critère de courbure-dimension inversé.

Nous allons tout d'abord présenter le développement de la variance, à partir duquel nous obtiendrons une formule de « Taylor » avec un reste. Ce reste s'exprimant sous la forme d'une inégalité de Talagrand L1/L2 (1.3.6) avec des dérivées partielles d'ordre supérieur. Ensuite, après de brefs rappels sur les inégalités de courbures dimension provenant de la théorie de Bakry et Émery, nous reformulerons les travaux de Chatterjee pour exhiber un critère de courbure dimension (intégré) inverse. Nous illustrerons cette approche sur le modèle du REM à différentes températures pour retrouver des résultats de Bovier *et al.*, puis nous rappelerons le résultat de Chatterjee concernant le modèle de verres de spin de Sherrington et Kirckpatrick.

Nous présenterons ensuite comment nos travaux peuvent sétendre au cube discret $\{-1,1\}^n$, afin d'obtenir de nouveaux résultats sur l'influence de coordonnées de fonctions booléennes ainsi qu'une inégalité de Talagrand d'ordre supérieur. Il est évident que le cadre de l'article de Cordero-Erausquin et Ledoux [62] fournit une direction propice pour étendre les travaux présentés dans ce chapitre, nous faisons le choix (par pédagogie et un soucis de clarté) de ne pas entrer dans ce degré de généralités.

6.2 Développement de la variance

Voici le résultat principal de cette section.

Théorème 6.2.1. *Soit $f : \mathbb{R}^n \to \mathbb{R}$, $n \geq 1$, on suppose qu'il existe $m \geq 1$ tel que $f \in \mathcal{C}^m(\mathbb{R}^n)$ et que f et ses dérivées successives appartiennent toutes à $L^2(\gamma_n)$, où γ_n désigne la mesure gaussienne standard sur $\mathbb{R}^n$. Alors, pour tout $1 \leq p \leq m-1$, on a la formule de représentation suivante de la variance de f, sous la mesure γ_n, le long du semi-groupe d'Ornstein-Uhlenbeck $(P_t)_{t\geq 0}$.*

$$\mathrm{Var}_{\gamma_n}(f) = \sum_{k=1}^{p} \frac{1}{k!}\left| \int_{\mathbb{R}^N} \nabla^k f d\gamma_n \right|^2 + \frac{2}{p!}\int_0^\infty e^{-2t}\left(1-e^{-2t}\right)^p \int_{\mathbb{R}^n} \left|P_t(\nabla^{p+1} f)\right|^2 d\gamma_n dt, \tag{6.1}$$

Remarque. Notons, qu'en faisant tendre $p \to \infty$, la formule précédente redonne le développement, après intégration par parties successives, d'une fonction $L^2(\gamma_n)$ le long de la base des polynômes d'Hermite (cf. (1.7) chapitre un). On peut aussi procéder comme dans l'article de Ledoux avec l'entropie à la place de la variance. Seulement, les formules ne semblent pas s'exprimer aussi facilement que pour le cas de la variance. Regardons tout de même ce qu'il se passe lors de la première itération de la formule. On trouve, pour $f : \mathbb{R}^n \to \mathbb{R}$ telle que $f > 0$,

$$2\mathrm{Ent}_{\gamma_n}(f) = \frac{\left| \int_{\mathbb{R}^n} \nabla f d\gamma_n \right|^2}{\int_{\mathbb{R}^n} f d\gamma_n} + \int_0^\infty e^{-2u}(1-e^{-2u}) \int_{\mathbb{R}^n} k_u d\gamma_n du$$

avec $k_u = (P_u f)^{-3}\left|P_u(\nabla f)^t P_u(\nabla f) - P_u f P_u(\nabla^2 f)\right|^2$. Puisque $k_u \geq 0$ pour tout $u \geq 0$, on obtient la minoration suivante, valable pour toute fonctions vérifiant $\int_{\mathbb{R}^n} f d\gamma_n = 1$,

$$2\mathrm{Ent}_{\gamma_n}(f^2) \geq \left| \int_{\mathbb{R}^n} \nabla f d\gamma_n \right|^2.$$

Cette minoration correspondant à l'inégalité de Sobolev logarithmique inverse.

Démonstration. Le point de départ est la représentation dynamique de la variance le long du semi-groupe d'Ornstein-Uhlenbeck.

$$\mathrm{Var}_{\gamma_n}(f) = 2\int_0^\infty e^{-2t} \int_{\mathbb{R}^n} \left|P_t(\nabla f)\right|^2 d\gamma_n dt.$$

Nous adoptons la notation suivante, pour $t \geq 0$

$$K_1(t) = \int_{\mathbb{R}^n} |P_t(\nabla f)|^2 d\gamma_n.$$

Ainsi, d'après le théorème fondamental de l'analyse, pour tout $s \geq t \geq 0$

$$K_1(t) = K_1(s) - \int_t^s K_1'(u) du$$

et en utilisant le fait que $\nabla P_u f = e^{-u} P_u \nabla f$ et une intégration par partie, on obtient

$$K_1'(u) = \frac{d}{du} \int_{\mathbb{R}^n} |P_u \nabla f|^2 d\gamma_n = -2 \int_{\mathbb{R}^n} e^{-2u} |P_u \nabla^2 f|^2 d\gamma_n$$

Finalement, pour tout $s \geq t \geq 0$,

$$K_1(t) = K_1(s) + 2 \int_t^s e^{-2u} \int_{\mathbb{R}^N} |P_u \nabla^2 f|^2 d\gamma_N du,$$

ce qui entraine, lorsque $s \to \infty$,

$$K_1(t) = \left| \int_{\mathbb{R}^n} \nabla f d\gamma_n \right|^2 + 2 \int_t^\infty e^{-2u} \int_{\mathbb{R}^n} |P_u \nabla^2 f|^2 d\gamma_n du,$$

par ergodicité de $(P_t)_{t \geq 0}$. En substituant l'expression de K_1 dans la formule de représentation de la variance, on obtient

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

Enfin, par Fubini,

$$\mathrm{Var}_{\gamma_n}(f) = \left| \int_{\mathbb{R}^n} \nabla f d\gamma_n \right|^2 + 2 \int_0^\infty e^{-2u}(1 - e^{-2u}) \int_{\mathbb{R}^n} |P_u(\nabla^2 f)|^2 d\gamma_N du.$$

il suffit ensuite d'itérer à nouveau le procédé : en posant

$$K_2(u) = \int_{\mathbb{R}^n} |P_u(\nabla^2 f)|^2 d\gamma_n,$$

il vient que

$$K_2(u) = \left| \int_{\mathbb{R}^n} \nabla^2 f d\gamma_n \right|^2 + 2 \int_u^\infty e^{-2t} \int_{\mathbb{R}^n} |P_t \nabla^3 f|^2 d\gamma_n dt.$$

Puis après substitution, il suffit de calculer

$$\left| \int_{\mathbb{R}^N} \nabla^2 f d\gamma_n \right|^2 \left[2 \int_0^\infty e^{-2u}(1 - e^{-2u}) du \right] = \frac{1}{2} \left| \int_{\mathbb{R}^n} \nabla^2 f d\gamma_n \right|^2$$

et

$$4 \int_0^\infty e^{-2t} \left(\int_{\mathbb{R}^n} |P_t \nabla^3 f|^2 d\gamma_n \right) \left[\int_0^t e^{-2u}(1 - e^{-2u}) du \right] dt.$$

Un calcul immédiat entraine que

$$2\int_0^t e^{-2u}(1-e^{-2u})du = \left(1-e^{-2t}\right)^2$$

Il suffit de procéder par récurrence pour conclure. En effet, les coefficients que nous obtenons à chaque itération se définissent par récurrence assez aisément. En effet, on pose $a_0(t) = 2e^{-2t}$ et $a_1 = \int_0^\infty a_0(t)dt$. Ensuite, pour $k \geq 1$, $a_k(t) = a_0(t)\int_0^t a_{k-1}(u)du$ et $a_k = \int_0^\infty a_k(t)dt$. Il n'est pas difficile de montrer que, pour tout $k \geq 0$ et tout $t \geq 0$,

$$a_k(t) = \frac{2}{k!}e^{-2t}\left(1-e^{-2t}\right)^k.$$

Ce qui entraine alors que, pour tout $k \geq 0$, $a_k = \frac{1}{k!}$. □

Remarque. Comme me l'a indiqué Houdré lors du workshop « High dimensional probability VIII », ce résultat existait déjà dans la littérature (cf. [95]). Les auteurs de l'article [95] obtiennent la même formule de représentations, pour des lois infiniments divisibles, à l'aide d'arguments similaires à ceux présentés plus haut (bien que le semi-groupe d'Ornstein-Uhlenbeck n'apparaisse pas explicitement). Toutefois, les inégalités de Talagrand d'ordre supérieur que nous en déduisons dans la section suivante semblent nouvelles.

6.3 Développement de Taylor de la variance avec reste

Nous présentons la formule précédente pour $p = 1$ et $p = 2$, afin de voir ce qu'il est possible d'obtenir à partir d'une telle représentation.

6.3.1 Ordre 1

Notons le fait suivant : pour $p = 1$, la formule de représentation de la variance nous indique que

$$\mathrm{Var}_{\gamma_n}(f) = \left|\int_{\mathbb{R}^n}\nabla f d\gamma_n\right|^2 + 2\int_0^\infty e^{-2t}(1-e^{-2t})\int_{\mathbb{R}^n}|P_t\nabla^2 f|^2 d\gamma_n dt.$$

Le deuxième terme étant strictement positif, cela permet de retrouver le contenu de l'inégalité de Poincaré inverse

$$\mathrm{Var}_{\gamma_n}(f) \geq \left|\int_{\mathbb{R}^n}\nabla f d\gamma_n\right|^2.$$

On peut également tirer parti de $2\int_0^\infty e^{-2t}(1-e^{-2t})\int_{\mathbb{R}^n}|P_t\nabla^2 f|^2 d\gamma_n dt$ de la manière suivante :

$$\mathrm{Var}_{\gamma_n}(f) = \left|\int_{\mathbb{R}^n}\nabla f d\gamma_n\right|^2 + 2\int_0^\infty e^{-2t}(1-e^{-2t})\int_{\mathbb{R}^n}|P_t\nabla^2 f|^2 d\gamma_n dt.$$

A partir de cette formule, on peut mettre en place la démonstration hypercontractive de Talagrand. En effet,

$$2\int_0^\infty e^{-2t}(1-e^{-2t})\int_{\mathbb{R}^n}|P_t\nabla^2 f|^2 d\gamma_n dt = 2\sum_{i,j=1}^n\int_0^\infty e^{-2t}(1-e^{-2t})\int_{\mathbb{R}^n}\left(P_t(\partial^2_{ij}f)\right)^2 d\gamma_n dt,$$

Autrement dit, nous avons l'expression $I = 2\sum_{i,j=1}^n \int_0^\infty e^{-2t}(1-e^{-2t})\|P_t(\partial_{ij}f)\|_2^2 dt$ à majorer. On peut alors utiliser la propriété d'hypercontractivité du semi groupe $(P_t)_{t\geq 0}$, pour toute fonction $g : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulière,

$$\|P_t(g)\|_2^2 \leq \|g\|_{1+e^{-2t}}^2, \quad t \geq 0.$$

En appliquant ceci à $g = \partial_{ij}f$, pour tout $i,j = 1,\ldots,n$, puis en posant $v = 1+e^{-2t}$, on obtient

$$I \leq \sum_{i,j=1}^n \int_0^2 (2-v)\|\partial_{ij}f\|_v^2 dv.$$

De plus, l'inégalité de Hölder entraine que $\|\partial_{ij}f\|_v \leq \|\partial_{ij}f\|_1^\theta \|\partial_{ij}f\|_2^{1-\theta}$, avec $\theta = \theta(v)$ satisfaisant la relation suivante $\frac{1}{v} = \frac{\theta}{1} + \frac{1-\theta}{2}$, pour tout $v \in [1,2]$. Ceci entraine donc la majoration suivante

$$I \leq \sum_{i,j=1}^n \|\partial_{ij}f\|_2^2 \int_1^2 (2-v)\left(\frac{\|\partial_{ij}f\|_1}{\|\partial_{ij}f\|_2}\right)^{2\theta} dv.$$

Si l'on pose $\alpha = \frac{\|\partial_{ij}f\|_1}{\|\partial_{ij}f\|_2} \leq 1$, il n'est pas difficile de montrer (après un changement de variable) que, pour tout $i,j = 1,\ldots,n$,

$$\int_1^2 (2-v)\left(\frac{\|\partial_{ij}f\|_1}{\|\partial_{ij}f\|_2}\right)^{2\theta} dv = \int_0^1 u e^{-\frac{2u}{2-u}\log(1/\alpha)} du.$$

Nous observons alors que $\int_0^1 u e^{-\frac{2u}{2-u}\log(1/\alpha)} du \leq C \frac{1}{\left[1+\log(1/\alpha)\right]^2}$ avec $C > 0$ une constante numérique. Finalement, nous avons obtenu la majoration de la variance suivante

$$\mathrm{Var}_{\gamma_n}(f) \leq \left|\int_{\mathbb{R}^n} \nabla f d\gamma_n\right|^2 + C\sum_{i,j=1}^n \frac{\|\partial_{ij}^2 f\|_2^2}{\left[1+\log\frac{\|\partial_{ij}^2 f\|_2}{\|\partial_{ij}^2 f\|_1}\right]^2},$$

avec $C > 0$, une constante numérique. En conclusion, nous avons démontré le théorème suivant

Théorème 6.3.1. *Dans le cadre qui précède, pour toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulière,*

$$\mathrm{Var}_{\gamma_n}(f) \leq \left|\int_{\mathbb{R}^n} \nabla f d\gamma_n\right|^2 + C\sum_{i,j=1}^n \frac{\|\partial_{ij}^2 f\|_2^2}{\left[1+\log\frac{\|\partial_{ij}^2 f\|_2}{\|\partial_{ij}^2 f\|_1}\right]^2},$$

6.3.2 Ordre 2

Les mêmes arguments fonctionnent à l'ordre $p = 2$. D'une part, ils permettent d'obtenir une puissance supplémentaire au dénominateur via l'argument de Talagrand. En effet, nous avons

$$\begin{aligned}\mathrm{Var}_{\gamma_n}(f) &= \left|\int_{\mathbb{R}^n} \nabla f d\gamma_n\right|^2 + \frac{1}{2}\left|\int_{\mathbb{R}^n} \nabla^2 f d\gamma_n\right|^2 \\ &+ 2\int_0^\infty \frac{1}{2}e^{-2t}\big(1-e^{-2t}\big)^2 \int_{\mathbb{R}^n} |P_t\nabla^3 f|^2 d\gamma_n dt.\end{aligned}$$

Pour les mêmes raisons que précédemment, nous trouvons que

$$2\int_0^\infty \frac{1}{2}e^{-2t}\big(1-e^{-2t}\big)^2\int_{\mathbb{R}^n}|P_t\nabla^3 f|^2 d\gamma_n dt \leq C\sum_{i,j,l=1}^n \|\partial_{ijl}f\|_2^2\int_1^2(2-v)^2\left(\frac{\|\partial_{ijl}f\|_1}{\|\partial_{ijl}f\|_2}\right)^{2\theta}dv.$$

Similairement à l'ordre 1, nous obtenons l'estimation suivante, valable pour tout $i,j,l=1,\dots,n$,

$$\int_1^2(2-v)^2\left(\frac{\|\partial_{ijl}f\|_1}{\|\partial_{ijl}f\|_2}\right)^{2\theta}dv \leq C\frac{1}{\Big[1+\log(1/\alpha)\Big]^3},$$

où cette fois $\alpha=\frac{\|\partial_{ijl}f\|_1}{\|\partial_{ijl}f\|_2}$.

Par récurrence immédiate, nous en déduisons le théorème suivant. Notons que le deuxième terme (avec le gain logarithmique) s'apparente à un « reste » d'un développement de Taylor de la variance .

Théorème 6.3.2. *D'après ce qui précède, pour tout $p\geq 1$, nous avons*

$$\mathrm{Var}_{\gamma_n}(f)\leq\sum_{k=1}^p\frac{1}{k!}\left|\int_{\mathbb{R}^n}\nabla^k f d\gamma_N\right|^2+C\sum_{i_1,\dots,i_{p+1}}^n\frac{\|\partial_{i_1,\dots,i_{p+1}}f\|_2^2}{\left[1+\log\left(\frac{\|\partial_{i_1,\dots,i_{p+1}}f\|_2}{\|\partial_{i_1,\dots,i_{p+1}}f\|_1}\right)\right]^{p+1}}.$$

Remarque. Observons à nouveau que la somme $\sum_{k=1}^p\frac{1}{k!}\left|\int_{\mathbb{R}^n}\nabla^k f d\gamma_n\right|^2$ est précisément le début du développement d'une fonction $L^2(\gamma_n)$ sur la base des polynômes d'Hermite.

6.4 Critère de courbure dimension intégré inverse

Cette partie montre que nous pouvons combiner la méthode d'interpolation présentée dans la section précédente avec les inégalités de courbure dimension provenant de la théorie de Bakry-Émery. Cela nous permet d'obtenir des inégalités de courbure dimension inverse, ce qui semble être nouveau dans ce domaine. Nous débutons cette section par quelques rappels.

6.4.1 Introduction

Nous souhaitons faire de brefs rappels concernant la théorie de Bakry-Émery et sur les inégalités de courbure-dimension intégrées dans le cas gaussien. Pour plus de détails, nous renvoyons le lecteur vers [16] ou le cours [111].

Considérons le processus d'Ornstein-Uhlenbeck $(X_t)_{t\geq 0}$ défini comme suit :

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

avec X et Y des vecteurs gaussiens standard, indépendants, de même loi dans $\mathbb{R}^n$. A ce processus est associé un semi-groupe $(P_t)_{t\geq 0}$ agissant sur un ensemble de fonctions. Ce semi-groupe possède une représentation explicite, pour f suffisamment régulière,

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

où γ_n désigne la mesure gaussienne standard sur $\mathbb{R}^n$. On peut facilement calculer le générateur infinitésimal d'un tel semi-groupe, celui-ci est donné par

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

agissant sur une certaine classe de fonction $\mathcal{D}(L)$. Ce semi-groupe admet γ_n comme mesure invariante et réversible.

Suivant Meyer, Bakry et Emery ont introduit les opérateurs « carré du champ » et leurs itérés de la manière suivante :

$\Gamma_0(f,g) = fg$ et $\Gamma_n(f,g) = \frac{1}{2}\Big[L\Gamma_{n-1}(f,g) - \Gamma_{n-1}(f,Lg) - \Gamma_{n-1}(Lf,g)\Big]$ pour $n \geq 1$. Afin d'alléger les notations, nous utiliserons $\Gamma_n(f)$ pour désigner $\Gamma_n(f,f)$, $n \geq 0$.

Certaines de ces formes bilinéaires symétriques jouent un rôle important pour obtenir, de manière élémentaire, des inégalités fonctionnelles. Par exemple, une inégalité du type $\Gamma_2 \geq \rho\Gamma_1$ avec $\rho > 0$ permet d'obtenir facilement, lorsqu'une hypothèse supplémentaire d'ergodicité est satisfaite, une inégalité de Sobolev logarithmique de constante $1/\rho$ pour la mesure invariante d'une diffusion (cf. [16] pour plus de détails). Dans la littérature, ce genre de relation, entre Γ_2 et Γ_1, est appelée : une inégalité de courbure-dimension $CD(\rho,\infty)$. Nous avons déjà briévement mentionné ce genre de résultat dans le chapitre un, en présentant, de manière équivalente, l'inégalité de Poincaré satisfaite par la mesure gaussienne, en terme d'inégalité entre Γ_1 et Γ_2 (cf. (1.24)).

Dans le cas du semi-groupe d'Ornstein-Uhlenbeck et son générateur infinitésimal L, il est facile de montrer que, pour toute fonction f suffisamment régulière,

$$\Gamma_1(f) = |\nabla f|^2 \quad \Gamma_2(f) = \|\mathrm{Hess} f\|_2^2 + |\nabla f|^2.$$

Il est alors évident que $\Gamma_2 \geq \Gamma_1$. Il se trouve que ces opérateurs apparaissent naturellement (cf. [16]) dans la formule de représentation de la variance (1.23) que nous avons employée durant une majeure partie de notre travail :

$$\mathrm{Var}_{\gamma_n}(f) = 2\int_0^\infty \int_{\mathbb{R}^n} |\nabla P_t f|^2 d\gamma_n dt,$$

avec γ_n la mesure gaussienne standard dans $\mathbb{R}^n$. Posons, pour tout $t \geq 0$,

$$I(t) = \int_{\mathbb{R}^n} |\nabla P_t f|^2 d\mu.$$

Au vu de ce qui précède, on remarque que

$$I(t) = \int_{\mathbb{R}^n} \Gamma_1(P_t f) d\gamma_n$$

et

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

Il est alors possible d'exprimer le critère de courbure dimension $\Gamma_2 \geq \Gamma_1$, le long du semi-groupe, sous une forme intégrée, à l'aide des fonctions $t \mapsto I(t)$ et $t \mapsto I'(t)$. Autrement dit

$$2I + I' \leq 0 \Leftrightarrow \int_{\mathbb{R}^n} \Gamma_2(P_t f)\mathrm{d}\mu \geq \int_{\mathbb{R}^n} \Gamma(P_t f) d\mu$$

Remarque. Notons le fait suivant, en intégrant l'inégalité $2I + I' \leq 0$ entre s et t (pour $0 \leq s < t$) nous obtenons

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

Il est alors classique de choisir $s = 0$, ceci permet d'obtenir facilement l'inégalité de Poincaré gaussienne. Observons que l'on peut garder s fixé et faire tendre t vers l'infini. Puisque $e^{2t}I(t) \to \left| \int_{\mathbb{R}^n} \nabla f d\gamma_n \right|^2$, nous en déduisons

$$I(s) \geq e^{-2s} \left| \int_{\mathbb{R}^n} \nabla f d\gamma_n \right|^2, \ s \geq 0.$$

Ce qui permet de retrouver l'inégalité de poincaré inverse satisfaite par la mesure gaussienne $\mathrm{Var}_{\gamma_n}(f) \geq \left| \int_{\mathbb{R}^n} \nabla f d\gamma_n \right|^2$.

6.4.2 Inégalité de courbure dimension, intégrée, inverse

Dans ce qui suit, nous montrons qu'il est possible d'obtenir des informations pertinentes sur la variance d'une fonction, si l'on peut obtenir une majoration (au lieu de la minoration classique) de $\int_{\mathbb{R}^n} \Gamma_2(P_t f)\mathrm{d}\mu$. Nous présentons un résultat abstrait, sur lequel nous proposerons une illustration sur un modèle de verres de spin. Rappelons que la méthodologie et les illustrations sont inspirées du chapitre dix du livre de Chatterjee, que nous réecrivons sous une forme plus manipulable.

Ces nouvelles inégalités de courbures dimension, intégrées, renversées (6.2) permettent d'obtenir la proposition suivante dans le cas du semi-groupe d'Ornstein-Uhlenbeck.

Proposition 6.4.1. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ fixée et supposons qu'il existe une fonction $\psi : \mathbb{R}_+ \to \mathbb{R}$ telle que*

1. *pour tout $t \geq 0$,*

$$\int_{\mathbb{R}^n} \Gamma_2(P_t f) d\gamma_n \leq \int_{\mathbb{R}^n} \Gamma(P_t f) d\gamma_n + \psi(t), \tag{6.2}$$

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

Alors,

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

Démonstration. L'equation 6.2 peut être reformulée comme suit :

$$I' \geq -2(I+\psi),$$

avec, rappelons le, $I(t) = \int_{\mathbb{R}^n} |\nabla P_t f|^2 d\gamma_n$, $t \geq 0$. Posons $I(t) = K(t)e^{-2t}$, l'inégalité précédente devient alors

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

En intégrant l'inégalité différentielle précédente entre s et t nous obtenons,

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

Ce qui entraine, lorsque $t \to \infty$,

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

Pour conclure, remarquons que

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

Nous avons donc obtenu, pour tout $t \geq 0$

$$I(t) = \int_{\mathbb{R}_n} \Gamma(P_t f) d\gamma_n \leq e^{-2t}\left(\left|\int_{\mathbb{R}^n} \nabla f d\gamma_n\right|^2 + 2\int_t^\infty e^{2s}\psi(s)ds\right), \tag{6.3}$$

des calculs élémentaires achèvent la démonstration.

□

Proposition 6.4.2. *Toujours dans le cas du semi-groupe d'Ornstein-Uhlenbeck, l'équation 6.2 est satisfaite ponctuellement pour la fonction, dite, d'énergie libre* $f_\beta(x) = \frac{1}{\beta}\log\left(\sum_{i=1}^n e^{\beta x_i}\right)$, $\beta > 0$ *avec* $\psi(t) = 2\beta^2 e^{-2t} I(t)$.

Le lemme technique suivant nous sera utile pour démontrer la proposition précédente.

Lemme 6.4.3. *Supposons qu'il existe une famille de fonctions* $(u_i)_{i=1,\dots,n}$*, avec* $u_i : \mathbb{R}^n \to \mathbb{R}$ *pour tout* $i = 1,\dots,n$*, satisfaisant*

— $0 \leq u_i(x) \leq 1$ $\forall i = 1,\dots,n, \forall x \in \mathbb{R}^n$.

— $\sum_{i=1}^n u_i(x) = 1$ $\forall x \in \mathbb{R}^n$.

Alors, pour toute fonction $v : \mathbb{R}^n \to \mathbb{R}_+$ *et toute mesure de probabilités* μ*, nous avons*

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

Démonstration. D'après le théorème de Fubini, nous obtenons

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

C'est pourquoi,

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

□

Nous pouvons alors démontrer la proposition 6.4.2. Il s'agit de démontrer que l'inégalité suivante est satisfaite :

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

Dans le cas du processus d'Ornstein-Uhlenbeck, cela peut s'exprimer plus simplement :

$$\int_{R_n}|\mathrm{Hess}P_tf_\beta|^2d\gamma_n \leq 2\beta^2e^{-2t}\int_{\mathbb{R}^n}|\nabla P_tf_\beta|^2d\gamma_n, \quad t\geq 0. \tag{6.4}$$

Démonstration. (proposition 6.4.2).
Notons que, ponctuellement, l'équation (6.4) est équivalente (grâce à la propriété de commutation entre ∇ et $(P_t)_{t\geq 0}$) à

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

Remarquons ensuite, que pour tout $i=1,\dots,n$, tout $\beta>0$,

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

et, pour tout $j=1,\dots,n$,

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

Nous obtenons, pour tout $t\geq 0$,

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

Il suffit donc d'ignorer les termes croisés qui sont toujours négatifs et d'appliquer le lemme précédent au troisième terme (à i fixé) pour conclure.

En effet, d'après le lemme 6.4.3, pour tout $i=1,\dots,n$,

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

D'où,

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

□

Ceci permet d'obtenir une majoration de la variance de l'énergie libre pour le modèle de verres de spin REM. Sans modification majeure, ce qui précède reste vérifié en remplaçant les vecteurs gaussiens X et Y (intervenant dans la définition du semi-groupe d'Ornstein-Uhlenbeck) par $\sqrt{n}X$ et $\sqrt{n}Y$ (avec cette fois-ci, X et Y des vecteurs gaussiens standards et indépendants de $\mathbb{R}^{2^n}$) pour $n \geq 1$. Notons alors que, pour tout $i = 1, \ldots, 2^n$, $\int_{\mathbb{R}^{2^n}} \partial_i f d\gamma_{2^n} = 2^{-n}$. Il s'ensuit donc, via la proposition précédente,

$$\mathrm{Var}_{\gamma_{2^n}}(f_\beta) \leq \frac{e^{-n(\log 2 - 2\beta^2)}}{\beta^2}, \quad \beta > 0.$$

En particulier, cette borne est du même ordre de grandeur que celle obtenue par Bovier *et al.* pour le REM lorsque $\beta \in]0, \sqrt{(\log 2)/2}]$ (cf. le théorème 2.2.12).

6.4.3 Inégalité de courbure dimension, intégrée, inverse partielle

Au vu de ce qui vient d'être présenté, deux questions se posent : que se passe-t-il si l'inégalité de courbure-dimension inverse est satisfaite pour $0 \leq t \leq T < \infty$, pour un certain $T > 0$, au lieu de $0 \leq t \leq \infty$? Cela permettrait-il de retrouver des estimées de l'énergie libre pour des grandes valeurs de β ? Afin d'atteindre ce genre de résultat, nous aurons besoin de deux lemmes techniques.

Tout d'abord, nous présentons un résultat de Cordero-Erausquin et Ledoux [62]. Celui-ci repose sur le fait que la mesure γ_n satisfait une inégalité de Poincaré et permet de restreindre le domaine d'intégration, de la formule de représentation de la variance, à un certain temps $T > 0$ fixé. Puisque une inégalité de Poincaré, entraine la décroissance exponentielle de la variance de $(P_t)_{t\geq 0}$ (cf. [16]).

Lemme 6.4.4. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulière, alors pour tout $T > 0$*

$$\mathrm{Var}_{\gamma_n}(f) \leq \frac{2}{1 - e^{-2T}} \int_0^T I(t)dt \tag{6.5}$$

Démonstration.

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

□

Ensuite, nous utiliserons le fait que le générateur infinitésimal $(-L)$ du processus d'Ornstein-Uhlenbeck $(X_t)_{t\geq 0}$ admet une décomposition spectrale discrète. Il est alors possible de considérer dE la mesure spectrale, discrète, associée. Ceci fournit une nouvelle représentation de $t \mapsto I(t)$, pour f fixée, :

$$I(t) = \int_0^\infty e^{-2t} dE(f), \quad t \geq 0$$

C'est pourquoi, en suivant les arguments de Baudoin et Wang dans [19], on peut montrer que $t \mapsto I(t)$ satisfait une inégalité du type Hölder. Autrement dit, pour tout $T > 0$,

$$I(s) \leq I(0)^{1-s/T} I(T)^{s/T}, \quad 0 \leq s \leq T \tag{6.6}$$

Tout ceci, permet d'obtenir une nouvelle majoration de la variance.

Proposition 6.4.5. *Soit $f : \mathbb{R}^n \to \mathbb{R}$, suffisamment régulière, fixée. Supposons qu'il existe $T > 0$ et $\psi : \mathbb{R}_+ \to \mathbb{R}$ tels que*

1. *pour tout $t \geq T$,*

$$\int_{\mathbb{R}^n} \Gamma_2(P_t f) d\gamma_n \leq \int_{\mathbb{R}^n} \Gamma(P_t f) d\gamma_n + \psi(t), \tag{6.7}$$

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

Alors,

$$\mathrm{Var}_{\gamma_n}(f) \leq \frac{2T}{1 - e^{-2T}} \frac{\int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n}{\log\left(\frac{\int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n}{|\int_{\mathbb{R}^n} \nabla f d\gamma_n|^2 + \alpha(T)}\right) + 2T},$$

avec $\alpha(T) = \int_T^\infty e^{2s} \psi(s) ds$

Démonstration. La démonstration est élémentaire. Il suffit d'utiliser le lemme 6.4.4 et l'inégalité 6.6 avec le schéma de preuve de la proposition 6.4.1 (permettant de contrôler $t \mapsto I(t)$ sur l'intervalle $[T, +\infty[$). Ceci mène à la majoration suivante de la variance

$$\int_0^T I(0)^{1-t/T} I(T)^{t/T} dt = I(0) \int_0^T e^{-\frac{t}{T} \ln\left(\frac{I(0)}{I(T)}\right)} dt = \frac{[I(0) - I(T)]}{\ln I(0)/I(T)}$$

$\square$

Tout ceci, permet d'étudier le modèle du REM à basse température (β suffisamment grand). Nous avons déja montré que, pour tout $t \geq 0$,

$$e^{2t} I(t) \leq \frac{1}{2^n} e^{2\beta^2 e^{-2t} n}.$$

On peut alors choisir $T = T_\beta = \frac{1}{2}\log(2\beta^2)$ de telle sorte que, pour $t \geq T$, nous ayons $e^{2\beta^2 e^{-2t} n} \leq e^n$. En utilisant la proposition précédente, ceci fournit

$$\mathrm{Var}(F_{\beta,n}) \leq \frac{2T}{1 - e^{-T}} \frac{n}{n \log(2/e)} \simeq C_\beta, \quad \beta \geq 0$$

Néanmoins, nous ne pouvons pas faire tendre β vers l'infini $\beta \to \infty$ et retrouver le fait que model $\mathrm{Var}(\max_{i=1,\dots 2^N} X_i) = O(1)$ (avec de telles renormalisations). Toutefois, la borne que nous obtenons correspond à celle du Bovier *et al.* 2.2.12 lorsque $\beta > \sqrt{2\ln 2}$.

Remarque. Le cas de $\beta = \infty$ peut-être traité par l'inégalité de Talagrand 1.3.6.

D'après Chatterjee [48], il est possible de traiter le modèle de Sherrington et Kirkpatrick (ainsi que d'autres modèles plus complexes comme les p-spin) par cette méthodologie. En effet, il a prouvé le résultat suivant

Proposition 6.4.6. *(Chatterjee) Soit X un vecteur gaussien centré de $\mathbb{R}^n$ de matrice de covariance Γ et Y une copie indépendante de X. Par la suite, on utilisera le processus d'Ornstein-Uhlenbeck (généralisé) sur $\mathbb{R}^n$, defini à l'aide de ces deux vecteurs gaussiens. Il suppose de plus que $\Gamma_{ij} \geq 0$ pour tout $i, j \in \{1, \dots, n\}$. Alors, l'inégalité qui suit est satisfaite, pour tout $t \geq 0$,*

$$I(t) \leq \sum_{i,j=1}^{n} \Gamma_{ij} e^{2\beta^2 e^{-2t}\Gamma_{ij} - 2t} \nu_i \nu_j, \tag{6.8}$$

avec $\nu_i = \mathbb{E}[\partial_i F_{\beta,n}]$ pour tout $i = 1, \dots, n$ et $F_\beta(x) = \frac{1}{\beta} \log\left(\sum_{i=1}^n e^{\beta x_i}\right)$, $\beta > 0$ la fonction d'énergie libre.

Remarque. La démonstration de Chatterjee n'est pas présentée dans le langage des opérateurs Γ_1 et Γ_2. Nous avons souhaité clarifier son schéma de preuve ainsi que certains de ses arguments.

Concernant le modèle plus complexe de Sherrington et Kirkpartrick, Chatterjee obtient une série d'inégalités différentielles de la forme $0 \leq g_r'(x) \leq 2\beta^2 g_{r+1}(x)$ pour tout $r \geq 1$ et tout $x \in [0, 1]$. La signification du r provient des puissances successives de la matrice de covariance. Il parvient à tirer profit de cette série d'inégalités pour obtenir une majoration de la fonction suivante

$$h_r(t) = \mathbb{E}[\sum_{i,j=1}^{n} (\Gamma_{ij})^r \partial_i f P_t \partial_j f],\ t \geq 0$$

Nous souhaitons souligner que la présence de Γ_{ij} dans l'exponentielle est essentiel. C'est ce qui lui permet, dans le cas du modèle SK, de relier la fonction $t \mapsto I(t)$ à une fonctionnelle de variables aléatoires de Rademacher. En particulier, en raisonnant avec l'inégalité de type Hölder il obtient la majoration suivante pour le modèle SK

$$\mathrm{Var}(F_{n,\beta}) \leq \frac{C_\beta n}{\log n}, \quad \beta \geq 0$$

En fait, il est possible, sans utiliser l'inégalité de Hölder, d'obtenir la borne suivante à partir des travaux de Chatterjee

$$\mathrm{Var}(F_{n,\beta}) \leq \frac{C_\beta}{4\beta^2}, \quad 0 < \beta < 1/2$$

En effet, l'inégalité (6.8) entraine que

$$\mathrm{Var}(F_{n,\beta}) \leq \int_0^\infty \sum_{\sigma,\sigma' \in C_n} \Gamma_{\sigma\sigma'} e^{2\beta^2 e^{-2t}\Gamma_{\sigma\sigma'} - 2t} \nu_\sigma \nu_{\sigma'} dt.$$

On peut ensuite appliquer le théorème de Fubini pour obtenir

$$\mathrm{Var}(F_{n,\beta}) \leq \frac{1}{4\beta^2} \sum_{\sigma,\sigma' \in C_n} e^{2\beta^2 \Gamma_{\sigma,\sigma'}} \nu_\sigma \nu_{\sigma'} = \frac{1}{4\beta^2} \mathbb{E}_{\sigma,\sigma'} \left[e^{2\beta^2 \frac{1}{n} \left(\sum_{i=1}^n \sigma_i \sigma_i' \right)^2} \right],$$

puisque $\Gamma_{\sigma\sigma'} = \frac{1}{n}(\sigma \cdot \sigma')^2$ et $\nu_\sigma = \mathbb{E}\left[\frac{e^{-\beta H_n(\sigma)}}{\sum_{\sigma'} e^{-\beta H_n(\sigma')}}\right] = \frac{1}{2^n}$ pour tout $\sigma \in C_n$ d'après [48]. Autrement dit,

$$\mathrm{Var}(F_{n,\beta}) \leq \frac{1}{4\beta^2} \mathbb{E}_{\sigma,\sigma'} \left[e^{2\beta^2 \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \right)^2} \right]$$

avec Y_i une variable de Rademacher pour tout $i = 1, \ldots, n$. Ainsi, pour tout $0 < \beta < 1/2$

$$\mathrm{Var}(F_{n,\beta}) \leq \frac{C_\beta}{4\beta^2}.$$

Ceci permet donc de retrouver à un résultat connu sur le modèle SK : la variance de l'énergie libre est bien bornée par une constante indépendante de la dimension pour tout $0 < \beta < 1/2$ (cf. [154]). Le véritable challenge est d'obtenir le même résultat à haute température (lorsque $\beta > 1$). Le résultat précédent permet toutefois d'obtenir de la superconcentration pour le « ground state ». En effet, d'après 4.1.5,

$$\mathrm{Var}\big(\max_{\sigma \in C_n} H_n(\sigma) \big) \leq 3\mathrm{Var}(F_{n,\beta}) + 6 \left(\frac{\log n}{\beta} \right)^2, \ \beta > 0$$

En choisissant $\beta = 1/4$, on en déduit que $\mathrm{Var}\big(\max_{\sigma \in C_n} H_n(\sigma)\big) \leq C_\beta (\log n)^2$ au lieu de la majoration en n fournie par l'inégalité de Poincaré.

6.5 Remarques diverses

La méthodologie précédente pourrait se généraliser en considérant le cadre présenté par Cordero-Erausquin et Ledoux dans leur article [62]. Nous ne souhaitons pas entrer dans ce degré d'abstraction ici.

Nous posons la question suivante : est-il possible d'exhiber une autre fonction f telle que pour un certain $T > 0$ et une certaine fonction ψ, l'inégalité

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

est satisfaite pour tout $t \geq T$?

Notons également le fait suivant, déjà présent dans le livre de Chatterjee, que l'on peut facilement minorer la variance d'une fonction à l'aide de l'égalité (en fait, on peut le faire directement via la décomposition de f sur la base des polynômes d'Hermite)

$$\mathrm{Var}_{\gamma_n}(f) = \sum_{k=1}^p \frac{1}{k!} \left| \int_{\mathbb{R}^N} \nabla^k f d\gamma_n \right|^2 + \frac{2}{p!} \int_0^\infty e^{-2t} \big(1 - e^{-2t}\big)^p \int_{\mathbb{R}^n} \left| P_t(\nabla^{p+1} f) \right|^2 d\gamma_n dt,$$

Pour obtenir que la variance de f est supérieure à $\frac{1}{k!} \left| \int_{\mathbb{R}^N} \nabla^k f d\gamma_N \right|^2$, pour n'importe quel $k \geq 1$. Ce terme correspond à projection de f sur le chaos gaussien d'ordre k (il suffit de procéder à une intégration par partie afin de faire apparaitre les polynômes d'Hermite). La conséquence de ceci est que le phénomène de superconcentration est difficilement réalisable pour une fonction f. En effet, l'heuristique de ceci est que, nécessairement, chacune des projections de f sur les différents chaos doit « tendre vers 0 lorsque $n \to \infty$ ». Sans cela, nous pourrions minorer la variance par une constante et montrer que la variance est exactement du même ordre de grandeur que la borne fournie par l'inégalité de Poincaré.

6.6 Cube discret

6.6.1 Introduction

Le cube discret $C_n = \{-1,1\}^n$ est un exemple intéressant sur lequel les méthodes de semi-groupes peuvent fournir des inégalités intéréssantes. Rappelons quelques propriétés de cet espace muni de la mesure produit μ^n, où $\mu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. L'espace $L^2(C_n, \mu^n)$ est un espace de Hilbert pour le produit scalaire usuel, il est possible d'exhiber une base orthonormée de cet espace : il s'agit des polynômes de Walsh. Ces polynômes sont définis commes suit : pour tout sous-ensemble $S \subset \{1, \ldots, n\}$ on définit le polynôme W_S par $W_S(x) = \prod_{i \in S} x_i$. L'orthogonalité de ces polynômes est facile à obtenir, en effet, par indépendance des x_i, $i = 1, \ldots, n$, nous avons, pour tous S et T des sous-ensembles de $\{1, \ldots, n\}$,

$$\langle W_S, W_T \rangle_{L^2(C_n,\mu^n)} = \mathbb{E}[W_S(X)W_T(X)] = \prod_{i \in S\Delta T} \mathbb{E}[X_i] = \delta(S,T),$$

où δ désigne le symbole de Kronecker et $(S,T) \mapsto S\Delta T$ la différence symétrique de deux ensembles. Cette base orthonormée est réminiscente de celle, formée par les polynômes d'Hermite, que l'on a présenté pour $\mathbb{R}^n$ muni de la mesure gaussienne standard γ_n. Ainsi, de manière similaire au cas gaussien, beaucoup de formules de Fourier restent valables dans ce cadre discret. Par exemple, il est possible de décomposer une fonction $f \in L^2(C_n, \mu^n)$ sur cette base

$$f = \sum_{S \subset \{1,\ldots,n\}} \hat{f}(S) W_S,$$

avec $\hat{f}(S) = \langle W_S, f \rangle_{L^2(C_n,\mu^n)}$. Il est aussi possible d'obtenir une formule de Plancherel pour deux fonctions $f, g \in L^2(C_n, \mu^n)$,

$$\mathbb{E}_{\mu^n}[fg] = \sum_{S \subset \{1,\ldots,n\}} \hat{f}(S)\hat{g}(S),$$

à partir de laquelle on obtient facilement que $\mathrm{Var}_{\mu^n}(f) = \sum_{S \neq \emptyset} \hat{f}^2(S)$, puisque $\mathbb{E}_{\mu^n}[f] = \hat{f}(\emptyset)$.

Comme pour le cas gaussien, il est possible de construire un opérateur différentiel discret auquel nous associerons un semi-groupe. Ce semi-groupe, dit de Bonami-Beckner, admet comme mesure invariante et réversible la mesure μ^n. Plus précisément, on définit le Laplacien « discret » par $L = \frac{1}{2}\sum_{i=1}^n D_i$, où D_i correspond à la dérivée partielle (discrète) en la i-ème coordonnée d'une fonction. C'est à dire, pour $f : C_n \to \mathbb{R}$, $D_i f = \frac{1}{2}\Big(f\big(\tau_i(x)\big) - f(x)\Big)$ avec $\tau_i(x) = (x_1, \ldots, -x_i, \ldots, x_n) \in C_n$ pour tout $i = 1, \ldots, n$ et tout $x = (x_1, \ldots, x_n)$. A l'aide de cet opérateur différentiel on peut définir une forme de Dirichlet sur C_n, pour toute fonction f, g définie sur le cube discret nous avons

$$\mathcal{E}(f,g) = \int_{C_n} f(-Lg)d\mu^n = \int_{C_n} Df \cdot Dg d\mu^n,$$

avec $Dh = (D_1 h, \ldots, D_n h)$ pour toute fonction $h : C_n \to \mathbb{R}$. Il est facile de voir que les éléments de la base de Fourier-Walsh sont des vecteurs propres de l'opérateur L. En effet, pour tout $S \subset \{1, \ldots, n\}$,

$$LW_S(x) = -\frac{1}{2}\sum_{i=1}^{n}\left(W_S(x) - W_S\big(\tau_i(x)\big)\right) = -\mathrm{Card}(S)W_S(x),$$

car $W_S(x) - W_S\big(\tau_i(x)\big) = 2 1_{i\in S} W_S(x)$. L'opérateur L engendre un semi-groupe, dit de Bonami-Beckner, $(Q_t = e^{tL})_{t\geq 0}$. Nous présentons ci-dessous quelques propriétés essentielles de ce semi-groupe.

Proposition 6.6.1. *On peut représenter le semi-groupe de la manière suivante, pour toute fonction $f : C_n \to \mathbb{R}$ et tout $t \geq 0$:*

$$Q_t(f) = \sum_{S\subset\{1,\ldots,n\}} e^{-t\mathrm{Card}(S)} \hat{f}(S) W_S.$$

$(Q_t)_{t\geq 0}$ est markovien et admet μ^n comme mesure invariante et réversible. Autrement dit, pour tout $t \geq 0$, $Q_t 1 = 1$ et

$$\int_{C_n} f Q_t g d\mu^n = \int_{C_n} g Q_t f d\mu^n,$$

pour toute fonction f, g définie sur le cube discret. A l'instar du semi-groupe d'Ornstein-Uhlenbeck, le semi-groupe de Bonami-Beckner admet une représentation intégrale explicite, pour tout $t \geq 0$,

$$Q_t f(x) = \int_{C_n} f(y) \prod_{i=1}^{n}(1 + e^{-t} x_i y_i) d\mu^n(y), \quad x \in C_n$$

Remarque. La représentation intégrale du semi-groupe permet d'établir la relation de commutation suivante $Q_t D = D Q_t$, $t \geq 0$ entre le semi-groupe et le gradient discret.

Il a été démontré par Bonami et Beckner, que ce semi-groupe vérifie une propriété d'hypercontractivité qui, rappelons-le, est équivalente à une inégalité de Sobolev logarithmique.

Théorème 6.6.2. *(Bonami-Beckner) Le semi groupe $(Q_t)_{t\geq 0}$ est hypercontractif. Autrement dit, pour toute fonction $f : C_n \to \mathbb{R}$, tout $t \geq 0$ et tout $p \geq 1$*

$$\|Q_t f\|_q \leq \|f\|_p, \tag{6.9}$$

avec $p = p(t) = 1 + (q-1)e^{-2t}$.

Remarque. L'inégalité (6.9) est équivalente à l'inégalité de Sobolev logarithmique

$$\mathrm{Ent}_{\mu^n}(f^2) \leq \int_{C_n} |Df|^2 d\mu^n,$$

avec $|\cdot|$ désignant la norme euclidienne. Bien qu'il soit possible de le démontrer de manière spectrale, via les polynômes de Walsh, le résultat précédent entraine également que la mesure μ^n satisfait une inégalité de Poincaré. Plus précisément, pour toute fonction f suffisamment régulière, nous avons l'inégalité suivante

$$\mathrm{Var}_{\mu^n}(f) \leq \frac{1}{2}\int_{C_n} |Df|^2 d\mu^n.$$

6.6.2 Influence

Dans ce paragraphe nous faisons de brefs rappels concernant la notion d'influence pour des fonctions booléennes $f : C_n \to \{0,1\}$. Pour plus de détails, nous renvoyons le lecteur vers les ouvrages suivants [128, 85].

Soit $f : C_n \to \{-1,1\}$, $n \geq 1$, il est possible d'imaginer qu'une telle fonction booléenne représente un système de vote. En effet, un point $x \in \{-1,1\}^n$ représenterait les votes de n personnes qui auraient choisi le candidat « -1 » ou bien le candidat « 1 » et la valeur de $f(x)$ désignerait le candidat gagnant. Il est naturel de s'interroger sur l'influence du i-ème vote pour $i \in \{1,\ldots,n\}$. Notamment est-il possible que la modification du i-ème vote change le résultat de l'élection ? Ceci amène à la définition suivante

Définition 6.6.3. Soit $f : C_n \to \{-1,1\}$, $n \geq 1$. La i-ème coordonnée, pour tout $i \in \{1,\ldots,n\}$, est dit pivotale, pour un certain $x \in C$ si $f(x) \neq f\big(\tau_i(x)\big)$ où $\tau_i(x) = (x_1,\ldots,-x_i,\ldots,x_n)$ correspond au vecteur x auquel on aurait changer de signe la i-ème coordonnée.

Ceci permet alors de définir la notion d'influence d'une coordonnée d'une fonction booléenne.

Définition 6.6.4. Soit $f : C_n \to \{-1,1\}$, $n \geq 1$. L'influence de la i-ème coordonnée de la fonction f est donnée par

$$I_i(f) = \mathbb{P}\Big(f(X) \neq f\big(\tau_i(X)\big)\Big)$$

avec $\mathcal{L}(X) = \mu^n$ la loi uniforme sur C_n.

Remarque. Notons que $I_i(f)$ correspond à la probabilité que la coordonnée i soit pivotale. Par la suite, nous adopterons la même définition d'influence pour des fonctions à valeurs dans $\{0,1\}$.

Il est utile de remarquer que l'influence de la i-ème coordonnée d'une fonction booléenne peut s'exprimer sous la forme de norme $L^p, p \geq 1$ des dérivées partielles de f. En effet, au prix de constantes numériques, nous avons l'identité suivante

$$I_i(f) = \|D_i f\|_1 = \|D_i f\|_p^p,\ i \in \{1,\ldots,n\}$$

Nous renvoyons à nouveau le lecteur vers les ouvrages [128, 85] pour des exemples de calculs d'influence pour des fonctions booléennes standards (fonction majorité, fonction tribues,...). Un résultat essentiel dans l'analyse booléenne est le théorème de Kahn, Kalai et Linial [101] dont nous fournissons l'énoncé ci-dessous.

Théorème 6.6.5. *[Kahn-Kalai-Linial] Soit $f : C_n \to \{-1,1\}$, $n \geq 1$. Alors il existe $i \in \{1,\ldots,n\}$, $c > 0$ tel que*

$$I_i(f) \geq c\mathrm{Var}_{\mu^n}(f)\frac{\log n}{n}.$$

Remarque. Il est notamment possible de déduire ce résultat à partir de l'inégalité de Talagrand (originelle) sur le cube discret (cf. [147]). Des extensions de ceci ont été proposées par différents auteurs, nous renvoyons le lecteur vers l'article [62] et les références de celui-ci pour plus de détails. Il a été démontré par Ben Or et Linial (cf. [20]) que la fonction tribues a toutes ses influences d'ordre $\log n/n$, ils conjecturent que ce résultat est optimal. Ainsi, le théorème de Kahn, Kalai et Linail assure donc cette optimalité.

Dans ce qui va suivre, nous proposons une définition d'influence « double » d'un couple $(i,j) \in \{1,\dots,n\}^2$ généralisant la définition 6.6.4. Nous proposerons un théorème dans le même esprit que celui de Kahn, Kalai et Linial sur les influences doubles. Ce théorème découlera d'une inégalité de Talagrand, sur le cube discret, à l'ordre deux. Le schéma de preuve consistera à adapter le cas gaussien présenté au début du chapitre six.

Comme nous l'avons vu précédemment, nous pouvons exprimer l'influence d'une coordonnée d'une fonction à partir de la norme L^p de la dérivée discrète de cette même fonction. Par analogie, nous allons définir les influences doubles $I_{ij}(f)$ à partir de la norme L^2 de $D_{ij}f$ où $D_{ij} = D_i \circ D_j$.

Il n'est pas difficile de montrer que, pour tout $(i,j) \in \{1,\dots,n\}^2$ et toute fonction $f : \mathcal{C}_n \to \{-1,1\}$,

$$D_{ij}f = \frac{1}{4}\Big(f(x) + f\big(\tau_{ij}(x)\big) - f\big(\tau_i(x)\big) - f\big(\tau_j(x)\big)\Big),$$

avec $\tau_{ij} = \tau_i \circ \tau_j$. Lorsque $i = j$ on trouve que $D_{ii}f = D_i f$.

Définition 6.6.6. Sans perdre de généralité, une fonction booléenne $f : C_n \to \{0,1\}$, $n \geq 1$ peut s'écrire sous la forme $f = 1_A$ avec $A \subset C_n$. Pour $(i,j) \in \{1,\dots,n\}^2$, l'influence du couple (i,j) est donnée par

$$I_{ij}(f) = \frac{1}{4}\Big[\mathbb{P}\big(X \in A, \tau_i(X) \notin A, \tau_j(X) \notin A\big) + \mathbb{P}\big(X \notin A, \tau_i(X) \in A, \tau_j(X) \in A\big)\Big],$$

où $\mathcal{L}(X) = \mu^n$.

Remarque. Ceci est en adéquation avec les influences d'ordre un, puisque l'on a, pour tout $i \in \{1,\dots,n\}$ et pour $f = 1_A$,

$$2I_i(f) = \mathbb{P}\big(X \in A, \tau_i(X) \notin A\big) + \mathbb{P}\big(X \notin A, \tau_i(X) \in A\big).$$

De plus, comme signifié auparavant, $I_i(f) = \|D_i f\|_2^2$ (aux constantes près). Il est également possible de montrer que

$$\|D_{ij}f\|_2^2 = 4\Big[\mathbb{P}\big(X \in A, \tau_i(X) \notin A, \tau_j(X) \notin A\big) + \mathbb{P}\big(X \notin A, \tau_i(X) \in A, \tau_j(X) \in A\big)\Big].$$

Comme nous le verrons dans la section suivante, il sera primordial de pouvoir comparer $\|D_{ij}f\|_2^2$ avec $\|D_{ij}f\|_1$ afin d'obtenir un résultat sur les influences à partir d'une inégalité de Talagrand d'ordre supérieur. En faisant une disjonction de cas (suivant que $x \in A, \tau_i(x) \in A, \tau_j(x) \in A, \tau_{ij}(x) \in A \dots$) il est possible de montrer que

$$\|D_{ij}f\|_1 \leq 4\|D_{ij}f\|_2^2 \leq 2\|D_{ij}f\|_1,\ (i,j) \in \{1,\dots,n\}^2.$$

Heuristiquement, ceci s'explique de la manière suivante : pour tout $(i,j) \in \{1,\dots,n\}^2$, $4D_{ij}1_A(x) \in \{-2,-1,0,1,2\}$, $x \in C_n$. Ainsi, pour tout $p \geq 1$, nous obtenons

$$|4D_{ij}1_A(x)| \leq |4D_{ij}1_A(x)|^p \leq 2^p|4D_{ij}1_A(x)|,\ x \in C_n,$$

permettant de comparer les normes L^p des $D_{ij}1_A$ entre elles.

6.6.3 Inégalités de Talagrand d'ordre supérieur et influence d'ordre deux

Nous allons voir qu'il est possible, au prix de constante numérique, de retrouver le même schéma de preuve que le cas gaussien. Une des différences majeures est le fait que le semi-groupe de Bonami-Beckner commute exactement avec le gradient discret. Autrement dit, pour tout $t \geq 0$, $DQ_t = Q_t D$ alors que cette commutation fait apparaitre un facteur e^{-t} dans le cas gaussien. Au cours de la démonstration qui va suivre, nous utiliserons le fait suivant (il s'agit d'une formulation équivalente de l'inégalité de Poincaré) : pour toute fonction $f : C_n \to \mathbb{R}$ de moyenne nulle sous la mesure μ^n, nous avons, pour tout $t \geq 0$

$$\|Q_t f\|_2^2 \leq e^{-2t} \|f\|_2^2.$$

Ceci sera notamment utilisé pour des fonctions de la forme $D_i f$ ou encore $D_{ij} f$ pour $i, j = 1, \dots, n$. Il n'est pas difficile de vérifier que de telles fonctions sont de moyenne nulle puisque, pour tout $i = 1, \dots, n$,

$$\int_{C_n} f(x) d\mu^n = \int_{C_n} f\big(\tau_i(x)\big) d\mu^n.$$

Notons également le fait que, pour tout $i = 1, \dots, n$, $D_{ii} f = D_i f$. Ces remarques étant faites nous pouvons nous attacher à la démonstration du résultat suivant.

Théorème 6.6.7. *Soit $f : \{-1,1\}^n \to \mathbb{R}$, $n \geq 1$, alors il existe $0 < s_0 < 1$ (fixé) telle que la majoration suivante est satisfaite*

$$\mathrm{Var}_{\mu^n}(f) \leq C \Bigg(\sum_{i=1}^n \|D_i f\|_{1+e^{-2s_0}}^2 + \sum_{i \neq j} \frac{\|D_{ij} f\|_2^2}{\left[1 + \log\left(\frac{\|D_{ij} f\|_2}{\|D_{ij} f\|_1}\right)\right]^2} \Bigg) \tag{6.10}$$

avec $C > 0$ une constante numérique.

Le théorème précédent permet d'obtenir un nouveau résultat, dans le même esprit que le théorème 6.6.5, portant sur l'influence de coordonnées de fonctions booléennes. En effet, si $f : C_n \to \{0,1\}$ le résultat précédent entraine une majoration de la variance de f en terme des influences de ses coordonnées.

Corollaire 6.6.8. *Sous le cadre précédemment décrit, il existe $0 < \eta < 1$ tel que*

$$\mathrm{Var}_{\mu^n}(f) \leq C_\eta \Bigg(\sum_{i=1}^n I_i(f)^{1+\eta} + \sum_{i \neq j} \frac{I_{ij}(f)}{\left[1 + \log\left(\frac{1}{16 I_{ij}(f)}\right)\right]^2} \Bigg).$$

Lorsque $f = 1_A$, avec $\mu^n(A) = 1/2$ (plus généralement, ce qui suit fonctionne également pour $0 < \epsilon < \mu^n(A) \leq 1 - \epsilon < 1$, pour tout $0 < \epsilon < 1$ indépendant de n), l'inégalité précédente entraine alors l'alternative suivante : soit il existe $i \in \{1, \dots, n\}$ tel que

$$I_i(A) \geq C \left(\frac{1}{n}\right)^{1/(1+\eta)}$$

ou bien il existe $(i,j) \in \{1, \dots, n\}^2$, $i \neq j$ tel que

$$I_{i,j}(A) \geq C \left(\frac{\log n}{n}\right)^2.$$

Démonstration. (Théorème 6.6.7) Le schéma de preuve débute comme le cas gaussien avec la représentation de la variance le long du semi-groupe de Bonami-Beckner $(Q_t)_{t\geq 0}$ (cf. [62, 27]) :

$$\mathrm{Var}_{\mu^n}(f) = 2\int_0^{\infty}\sum_{i=1}^{n}\int_{C_n} Q_t^2(D_i f)d\mu^n dt. \tag{6.11}$$

On pose alors $2s = t$ et utilisons le fait que $\int_{C_n} Q_{2s}^2(D_i f)d\mu^n = \|Q_s \circ Q_s(D_i f)\|_2^2 \leq e^{-2s}\|Q_s D_i f\|_2^2$ pour tout $i = 1, \ldots, n$. Puisque, par invariance de la mesure μ^n par rapport au semi-groupe de Bonami-Beckner, $\int_{C_n} Q_s(D_i f)d\mu^n = \int_{C_n} D_i f d\mu^n = 0$ pour tout $i = 1, \ldots, n$. Ceci mène à la majoration suivante,

$$\mathrm{Var}_{\mu^n}(f) \leq 4\int_0^{\infty} e^{-2s}\sum_{i=1}^{n}\int_{C_n} Q_s^2(D_i f)d\mu^n ds.$$

Comme dans le cas gaussien on pose, pour tout $s \geq 0$, $K(s) = \int_{C_n} |Q_s(Df)|^2 d\mu^n$. La formule d'intégration par partie discrète combinée au théorème fondamental de l'analyse entraine alors

$$K(s) = K(\infty) + 2\int_s^{\infty} |Q_u(D^2 f)|^2 d\mu^n du,$$

pour tout $s \geq 0$. Puisque le semi-groupe de Bonami Beckner est ergodique

$$K(\infty) = \left|\int_{C_n} Df d\mu^n\right|^2 = \sum_{i=1}^{n}\big(\int_{C_n} D_i f d\mu^n\big)^2 = 0.$$

Nous avons donc montré l'égalité suivante, pour tout $s \geq 0$,

$$K(s) = 2\sum_{i,j=1}^{n}\int_s^{\infty}\int_{C_n} Q_u^2(D_{ij} f)d\mu^n du.$$

Finalement, après substitution et utilisation du théorème de Fubini, ceci fournit

$$\mathrm{Var}_{\mu^n}(f) \leq 4\sum_{i,j=1}^{n}\int_0^{\infty}(1-e^{-2u})\int_{C_n} Q_u^2(D_{ij} f)d\mu^n du.$$

En posant à nouveau $2s = u$ puis en utilisant la décroissance exponentielle du semi-groupe dans $L^2(\mu^n)$, c'est-à-dire $\|Q_{2s} D_{ij} f\|_2^2 \leq e^{-2s}\|Q_s f\|_2^2$ $\forall i, j \in \{1, \ldots, n\}$, la majoration précédente devient

$$\mathrm{Var}_{\mu^n}(f) \leq 8\sum_{i,j=1}^{n}\int_0^{\infty} e^{-2s}(1-e^{-4s})\int_{C_n} Q_s^2(D_{ij} f)d\mu^n ds.$$

Nous allons découper la somme précédente suivant que $i = j$ ou non, afin de tirer profit du fait que $D_{ii} = D_i$ pour tout $i = 1, \ldots, n$. La variance de la fonction f est donc contrôlée par deux termes,

$$8\sum_{i=1}^{n}\int_0^{\infty} e^{-2s}(1-e^{-4s})\int_{C_n} Q_s^2(D_i f)d\mu^n ds + 8\sum_{i\neq j}\int_0^{\infty} e^{-2s}(1-e^{-4s})\int_{C_n} Q_s^2(D_{ij} f)d\mu^n ds.$$

Introduisons un paramètre $s_0 > 0$ que nous choisirons ultérieurement, nous pouvons alors traiter le premier terme de l'équation précédente de la manière suivante,

$$\begin{aligned}
8\sum_{i=1}^{n}\int_{0}^{\infty} e^{-2s}(1-e^{-4s})\int_{C_n} Q_s^2(D_i f)d\mu^n ds &= 8\sum_{i=1}^{n}\int_{0}^{s_0} e^{-2s}(1-e^{-4s})\int_{C_n} Q_s^2(D_i f)d\mu^n ds \\
&+ 8\sum_{i=1}^{n}\int_{s_0}^{\infty} e^{-2s}(1-e^{-4s})\int_{C_n} Q_s^2(D_i f)d\mu^n ds
\end{aligned}$$

Il n'est pas difficile de voir que,

$$\begin{aligned}
8\sum_{i=1}^{n}\int_{0}^{s_0} e^{-2s}(1-e^{-4s})\int_{C_n} Q_s^2(D_i f)d\mu^n ds &\leq 8\sum_{i=1}^{n}\int_{0}^{s_0} 4s\int_{C_n} Q_s^2(D_i f)d\mu^n ds \\
&\leq 32 s_0\int_{0}^{\infty}\sum_{i=1}^{n}\int_{C_n} Q_s^2(D_i f)d\mu^n ds = 16 s_0 \mathrm{Var}_{\mu^n}(f)
\end{aligned}$$

où la dernière égalité découle de la formule de représentation dynamique de la variance (6.11). Si l'on choisit s_0 tel que $s_0 16 \leq 1/2$, tout ceci mène donc à l'inégalité

$$\begin{aligned}
\frac{1}{2}\mathrm{Var}_{\mu^n}(f) &\leq 8\int_{s_0}^{\infty} e^{-2s}(1-e^{-4s})\sum_{i=1}^{n}\int_{C_n} Q_s^2(D_i f)d\mu^n ds \\
&+ 8\sum_{i\neq j}\int_{0}^{\infty} e^{-2s}(1-e^{-4s})\int_{C_n} Q_s^2(D_{ij} f)d\mu^n ds.
\end{aligned}$$

Occupons nous à présent d'estimer $8\int_{s_0}^{\infty} e^{-2s}(1-e^{-4s})\sum_{i=1}^{n}\int_{C_n} Q_s^2(D_i f)d\mu^n ds$. A cet effet, remarquons que la contractivité du semi-groupe de Bonami-Beckner entraine la majoration suivante, valable pour tout $i = 1,\ldots,n$ et tout $s \geq s_0$,

$$\|Q_s D_i f\|_2^2 = \|Q_{s-s_0}\circ Q_{s_0} D_i f\|_2^2 \leq \|Q_{s_0}(D_i f)\|_2^2.$$

Ainsi,

$$\begin{aligned}
8\int_{s_0}^{\infty} e^{-2s}(1-e^{-4s})\sum_{i=1}^{n}\int_{C_n} Q_s^2(D_i f)d\mu^n ds &\leq 8\sum_{i=1}^{n}\|Q_{s_0} D_i f\|_2^2\int_{s_0}^{\infty} e^{-2s}(1-e^{-4s})ds \\
&\leq 4\sum_{i=1}^{n}\|Q_{s_0} D_i f\|_2^2 \\
&\leq 4\sum_{i=1}^{n}\|D_i f\|_{1+e^{-2s_0}}^2
\end{aligned}$$

ou la dernière majoration provient de la propriété d'hypercontractivité du semi-groupe de Bonami-Beckner. Il ne reste plus qu'à majorer la somme lorsque $i \neq j = 1,\ldots,n$. Ceci est assez simple puisque la majoration suivante nous ramène aux mêmes estimations que le cas gaussien

$$8\sum_{i\neq j}\int_{0}^{\infty} e^{-2s}(1-e^{-4s})\int_{C_n} Q_s^2(D_{ij} f)d\mu^n ds \leq 16\sum_{i\neq j}\int_{0}^{\infty} e^{-2s}(1-e^{-2s})\int_{C_n} Q_s^2(D_{ij} f)d\mu^n ds.$$

Il suffit alors d'appliquer les mêmes arguments que ceux utilisés dans le théorème 6.3.1 pour montrer que

$$16\sum_{i\neq j}\int_0^\infty e^{-2s}(1-e^{-2s})\int_{C_n} Q_s^2(D_{ij}f)d\mu^n ds \leq C\sum_{i\neq j}\frac{\|D_{ij}f\|_2^2}{\left[1+\log\left(\frac{\|D_{ij}f\|_2}{\|D_{ij}f\|_1}\right)\right]^2}$$

□

Remarque. Il n'est pas difficile de voir que la preuve précédente fonctionne également pour les mesures biaisées $\mu_p^n = \{p\delta{-1}+q\delta 1\}^{\otimes n}$ avec $p\in]0,1[$ et $p+q=1$. Il suffit de modifier légérement certains objets : le générateur L ainsi que la constante de l'inégalité de Sobolev logarithmique apparaissant dans la propriété d'hypercontractivité. Nous renvoyons le lecteur vers [128, 115] pour trouver les informations nécessaires à ces modifications. Signalons à nouveau que la méthodologie employée ici semble aisément s'étendre au cadre proposé dans l'article [115].

Remarque. Notons toutefois que la démonstration précédente fonctionnerait aux ordres supérieurs. Il faudrait toutefois être prudent pour décomposer la somme faisant intervenir des dérivées de la forme D_{ijk}. Par exemple, sur la diagonale ($i=j=k$) nous obtiendrons des D_i ; si jamais deux indices sont égaux ($i=j$, $k\neq i\neq j$ par exemple) nous aurons des D_{ik}. Il ne restera plus qu'à appliquer l'argument hypercontractif aux dérivées de la forme D_{ijk}, avec $i\neq j\neq k$.

Maintenant que la démonstration du théorème 6.6.7 est faite, nous pouvons proposer une preuve du corollaire 6.6.8 sur les influences. Nous nous ne ferons que le cas particulier $f=1_A$ pour un ensemble $A\subset C_n$, la démonstration lorsque $f: C_n\to\{-1,1\}$ est identique.

Démonstration. (Corollaire 6.6.8) Soit $A\subset C_n$ et $f=1_A$. Il suffit d'appliquer l'inégalité (6.10) du théorème 6.6.7 pour faire apparaitre les influences. Tout d'abord, notons que pour une telle fonction f, $\|D_i f\|^2_{1+e^{-2s_0}} = I_i(A)^{2/(1+e^{-2s_0})}$ pour tout $i=1,\dots,n$. Notons que $\frac{2}{1+e^{-2s_0}}\in]1,2[$ pour tout $s_0>0$ et donc peut ceci peut se réécrire sous la forme $1+\eta$ avec $\eta\in]0,1[$.

Comme nous avons pu le faire remarquer dans la section précédente, les normes $\|D_{ij}f\|_2$ et $\|D_{ij}f\|_1$ avec $f=1_A$ sont comparables, pour tout $i\neq j=1,\dots,n$, et correspondent à des influences doubles $I_{ij}(A)$. Plus précisément, pour tout $i\neq j=1,\dots,n$

$$\frac{1}{2}\|D_{ij}f\|_1^{1/2}\leq\|D_{ij}f\|_2\leq\frac{1}{\sqrt{2}}\|D_{ij}f\|_1^{1/2}$$

et $\|D_{ij}f\|_2 = 4I_{ij}(A)$. D'après ce qui précède, nous obtenons donc

$$\mu^n(A)\big(1-\mu^n(A)\big)\leq C\sum_{i=1}^n I_i(A)^{1+\eta} + C\sum_{i\neq j}\frac{I_{ij}(A)}{\left[1+\log\left(\frac{1}{16I_{ij}(A)}\right)\right]^2}.$$

A présent, soit le premier terme domine l'autre et l'inégalité précédente entraine que

$$\mu^n(A)\big(1-\mu^n(A)\big)\leq C\sum_{i=1}^n I_i(A)^{1+\eta}.$$

Ainsi, il existe $i\in\{1,\dots,n\}$ tel que $I_i(A)^{1+\eta}\geq\frac{\mu^n(A)\left(1-\mu^n(A)\right)}{Cn}$. Ou bien le deuxième terme domine l'autre et nous obtenons

$$\mu^n(A)\big(1-\mu^n(A)\big) \leq C\sum_{i\neq j}\frac{I_{ij}(A)}{\left[1+\log\left(\frac{1}{16I_{ij}(A)}\right)\right]^2}.$$

Il suffit alors de faire le raisonnement suivant : soit il existe $(i,j)\in\{1,\dots,n\}$, $i\neq j$ tel que $I_{ij}(A)\geq \frac{\mu^n(A)\left(1-\mu^n(A)\right)}{16n}$ et le résultat s'ensuit. Soit, pour tout $(i,j)\in\{1,\dots,n\}$, $i\neq j$, nous avons $I_{ij}(A)\leq \frac{\mu^n(A)\left(1-\mu^n(A)\right)}{16n}$, permettant d'obtenir la majoration suivante

$$\mu^n(A)\big(1-\mu^n(A)\big) \leq C\sum_{i\neq j}\frac{I_{ij}(A)}{\left[1+\log\left(\frac{n}{\mu^n(A)\left(1-\mu^n(A)\right)}\right)\right]^2}.$$

Il existe donc un couple $(i,j)\in\{1,\dots,n\}$, $i\neq j$ tel que

$$\frac{\mu^n(A)\big(1-\mu^n(A)\big)}{Cn(n-1)} \leq \frac{I_{ij}(A)}{\left[1+\log\left(\frac{n}{\mu^n(A)\left(1-\mu^n(A)\right)}\right)\right]^2},$$

ce qui permet d'établir la conclusion voulue. □

A titre indicatif, nous énonçons ci dessous l'inégalité de Talagrand à l'ordre p que nous pouvons obtenir en itérant la procédure. Contrairement au cas gaussien, la dépendance en p est mauvaise (principalement à cause des majorations successives permettant d'imiter le cas gaussien) et ne permet de faire tendre p vers l'infini. Nous laissons au lecteur le soin de compléter la démonstration.

Théorème 6.6.9. *Soit $f : C_n \to \mathbb{R}$, $n \geq 1$. Pour tout $p \geq 0$ nous avons la majoration suivante de la variance*

$$\mathrm{Var}_{\mu^n}(f) \leq \frac{2^{p+2+\sum_{k=0}^{p}k}}{p!}\int_0^\infty e^{-2s}(1-e^{-2s})^p\int_{C_n}\sum_{i_1,\dots,i_{p+1}=1}^{n} Q_s^2(D_{i_1,\dots,i_{p+1}}f)d\mu^n ds. \tag{6.12}$$

En particulier, l'argument hypercontractif entraine le résultat suivant

$$\mathrm{Var}_{\mu^n}(f) \leq C_p\sum_{i_1,\dots,i_{p+1}=1}^{n}\frac{\|D_{i_1,\dots,i_{p+1}}f\|_2^2}{\left[1+\log\left(\frac{\|D_{i_1,\dots,i_{p+1}}f\|_2}{\|D_{i_1,\dots,i_{p+1}}f\|_1}\right)\right]^{p+1}},$$

avec $C_p > 0$.

Chapitre 7

Bornes sur la variance et inégalité de déviation non asymptotique par transport optimal

7.1 Introduction

Ce chapitre propose d'aborder la superconcentration par le biais du transport optimal. Les inégalités de Poincaré à poids ont été étudiées par différents auteurs (cf. [18, 26, 87, 8]), nous verrons qu'elles s'avèrent être des outils performants pour obtenir de la superconcencentration pour des mesures produits. Nous montrerons également que certains problèmes isopérimétriques peuvent également être reliés à la superconcentration et permettent d'obtenir des inégalités de déviations pertinentes. L'idée essentielle, déjà présente dans [87], des travaux qui vont suivre est la suivante : étant donnée une mesure de probabilité ν satisfaisant une inégalité fonctionnelle, nous choisissons de transporter cette mesure sur une mesure μ afin d'obtenir de la superconcentration pour μ. L'étape essentielle est d'estimer précisément le comportement de l'application de transport. Signalons tout de même que nous n'utiliserons pas de fonction de coût ou d'inégalités fonctionelles, comme celle de Pinsker, issues de la théorie du transport optimal par la suite. Peut-être que la mot transfert, illustrant de quelle manière une mesure peut hériter des propriétés satisfaites par une autre, eût été plus adapté pour décrire la partie qui va suivre. Cependant nous trouvions cette terminologie moins évocatrice et avons préféré conserver l'appellation de transport optimal. Les résultats exposés dans ce chapitre seront également rédiger sous forme d'article.

Dans ce qui suit, nous illustrerons nos théorèmes abstraits par des résultats provenant de la théorie des extrêmes. Nous renverrons régulièrement le lecteur vers le chapitre trois à cet effet. Étant donné des mesures de probabilités, absolument continues par rapport à la mesure de Lebesgue sur $\mathbb{R}$, μ et ν nous désignerons, respectivement, par $X_1, \ldots, X_n$ une famille de $n \geq 1$ variables aléatoires indépendantes de même loi (*i.e.* $\mathcal{L}(X_1) = \mu$) et par $Y_1, \ldots, Y_n$ une famille de $n \geq 1$ variables inépendantes de loi ν. On pose alors $M_n = \max_{i=1,\ldots,n} X_i$.

Comme nous l'avons mentionné, nos travaux sont réminiscent des travaux de [87, 18] sur les inégalités de Poincaré à poids sur $\mathbb{R}$. A titre d'exemple, dans [18], les auteurs obtiennent le théorème suivant (cf. [8])

Théorème 7.1.1. *(Barthe-Roberto) Soit μ une mesure de probabilité sur $\mathbb{R}$ de médiane m et considérons $d\nu(x) = g(x)dx$ une mesure sur $\mathbb{R}$. La meilleure constante $C_P > 0$, telle que pour* *toute fonction localement lipschitzienne $f : \mathbb{R} \to \mathbb{R}$ nous ayons l'inégalité suivante*

$$\mathrm{Var}_\mu(f) \le C_P \int_{\mathbb{R}} (f')^2 d\nu,$$

vérifie $\max(B_+, B_-) \le C_P \le 4\max(B_+, B_-)$*, où*

$$B_+ = \sup_{x>m} \nu\big([x,+\infty)\big) \int_m^x \frac{dt}{g(t)}, \quad B_- = \sup_{x<m} \nu\big((-\infty,x]\big) \int_x^m \frac{dt}{g(t)}. \tag{7.1}$$

La preuve de ce résultat repose sur l'inégalité de Hardy et le critère de Muckenhoupt. Notre approche est, par nature, similaire à ceci ; toutefois notre schéma de preuve s'appuie sur des arguments de transport optimal et se rapproche de ceux de Gozlan. L'idée majeure de [87] est le fait suivant : si l'on modifie la métrique euclidienne, il est possible de renforcer certains résultats classiques de concentration. L'article de Gozlan peut-être résumé, sommairement, de la manière suivante : étant donné une mesure μ^n ainsi qu'une fonction ω (satisfaisant des conditions de régularités) Gozlan considère la mesure poussée $\omega^{\#}\mu^n$. Similairement au théorème 7.1.1, ω est choisie comme étant la plus « grande » fonction possible (au sens où la meilleure constante de Poincaré C_P puisse obtenir, en transportant $\omega^{\#}\mu^n$ sur μ^n, doit satisfaire une condition impliquant les constantes B_+ et B_- données par (7.1)) afin que $\omega^{\#}\mu^n$ satisfasse encore une inégalité de Poincaré. La conséquence de ceci est de permettre de renforcer la concentration satisfaite, initialement, par la mesure μ^n. Plus précisément, Gozlan a obtenu le théorème suivant :

Lemme 7.1.2. *(Gozlan) Les assertions suivantes sont équivalentes.*

1. $\omega^{\#}\mu^n$ *vérifie une inégalité de Poincaré de constante* $C_P > 0$ *: pour toute fonction* f *suffisamment régulière*

$$\mathrm{Var}_{\omega^{\#}\mu^n}(f) \le C_P \int_{\mathbb{R}^n} |\nabla f|^2 d\omega^{\#}\mu^n.$$

2. μ^n *vérifie une inégalité de Poincaré à poids : pour toute fonction* f *suffisamment régulière,*

$$\mathrm{Var}_{\mu^n}(f) \le C \sum_{i=1}^n \int_{\mathbb{R}^n} \frac{1}{\omega'(x_i)} (\partial_i f)^2 d\mu^n$$

Remarque. Grâce au critère de Muckenhoupt (cf. [16] par exemple), il propose également des conditions suffisantes pour qu'une mesure de la forme $d\mu = e^{-V(x)}dx$, avec V convexe satisfasse ce genre d'inégalité. Gozlan utilise ce genre d'inégalités pour démontrer des résultats de concentration (cf. [87]). Il étudie également les liens reliant ces inégalités de Poincaré à poids avec des inégalités de transport et d'inf-convolution, nous ne développerons pas ce point par la suite.

Nous souhaitons faire remarquer au lecteur que, bien que notre point de vue soit légérement différent, notre méthodologie est essentiellement la même que celle de Gozlan. En effet, nous préferons choisir une mesure ν vérifiant une inégalité de Poincaré afin de la transporter, par une application monotone T, sur la mesure d'intérêt μ. Finalement, avec les notations de Gozlan, tout ceci revient à considérer $T = \omega^{-1}$. Nous verrons dans ce chapitre que certaines propriétés de la mesure exponentielle produit (comme l'inégalité de Poincaré ou un résultat isopérimétrique) permettent d'améliorer certaines inégalités fonctionnelles satisfaites par la mesure μ^n. Nous illustrerons cette approche de la superconcentration en proposant des bornes sur la variance ainsi que des inégalités de déviation. Ces résultats seront non asymptotiques et reflèteront le comportement des extrêmes de la loi μ.

7.2 Notations et rappels de transport optimal

Soient μ et ν deux mesures de probabilités sur $\mathbb{R}$. Supposons que ces deux mesures soient absolument continues par rapport à la mesure de Lebesgue sur $\mathbb{R}$. Plus précisément,

$$d\mu(x) = h(x)dx, \quad d\nu(x) = g(x)dx$$

Soit X une variables aléatoire loi μ et Y une variables aléatoire de loi ν. Notons H (respectivement G) la fonction de répartition de X (respectivement Y). Et définissons la fonction de risque associée à une mesure de probabilité μ par

$$\kappa_\mu(x) = \frac{h(x)}{1 - H(x)}, \quad x \in \mathrm{supp}(\mu) \subset \mathbb{R}$$

De plus, nous supposerons que ν satisfait une inégalité de Poincaré sur $\mathbb{R}$ de constante $C_P > 0$. Autrement dit, pour $f : \mathbb{R} \to \mathbb{R}$ suffisamment régulière,

$$\mathrm{Var}_\nu(f) \leq C_P \int_{\mathbb{R}} f'^2 d\nu.$$

Par tensorisation, cela entraine que la mesure produit ν^n satisfait une inégalité de Poincaré sur $\mathbb{R}^n$ avec la même constante C_P. Nous souhaitons souligner le fait que nous choisirons principalement ν comme étant la mesure exponentielle sur $\mathbb{R}_+$ (ou bien comme étant la mesure exponentielle symétrique sur $\mathbb{R}$) comme mesure de référence à partir de laquelle nous améliorerons la concentration satisfaite par la mesure d'intérêt μ. Cependant, lors de l'énoncé de nos résultats principaux, nous ne spécifierons pas les mesures μ et ν.

Rappelons que le transport monotone de ν sur μ (cf. chapitre un ou [162] pour plus de détails) est obtenu par une application $t : \mathbb{R} \to \mathbb{R}$ telle que, pour tout $x \in \mathbb{R}$,

$$\int_{-\infty}^{x} d\nu = \int_{-\infty}^{t(x)} d\mu.$$

De manière équivalente, en termes des fonctions de répartition,

$$G(x) = \mathbb{P}(Y \leq x) = \mathbb{P}\big(X \leq t(x)\big) = H\big(t(x)\big), \quad x \in \mathbb{R}. \tag{7.2}$$

Ce qui mène, après dérivation, à

$$g(x) = h\big(t(x)\big)t'(x), \quad x \in \mathbb{R} \tag{7.3}$$

On définit à présent $T : \mathbb{R}^n \to \mathbb{R}^n$ par $T(x) = \big(t(x_1), \ldots, t(x_n)\big)$, pour tout $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. L'application T transporte ν^n sur μ^n.

C'est pourquoi, pour toutes applications $f : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulières,

$$\mathrm{Var}_{\mu^n}(f) = \mathrm{Var}_{\nu^n}\big(f \circ T\big),$$

7.3 Résultat principal

Voici notre résultat principal, à partir duquel nous obtiendrons des bornes sur la variance et des inégalités de déviations non-asymptotique.

Théorème 7.3.1. *Avec les notations précédentes, pour toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ suffisamment régulière, $n \geq 1$, nous avons*

$$\operatorname{Var}\big(f(X)\big) \leq C_P \sum_{i=1}^{n} \mathbb{E}\left[(\partial_i f)^2 \circ T(Y)\left(\frac{\kappa_\nu(Y_i)}{\kappa_\mu\big(t(Y_i)\big)}\right)^2\right], \tag{7.4}$$

avec $Y = (Y_1, \ldots, Y_n)$ où $Y_1, \ldots, Y_n$ sont des copies indépendantes de $Y, \mathcal{L}(Y) = \nu$ et $X = (X_1, \ldots, X_n)$ où $X_1, \ldots, X_n$ sont des copies indépendantes de $X, \mathcal{L}(X) = \mu$.

Démonstration. Par définition de l'application de transport T, nous avons

$$\operatorname{Var}_{\mu^n}(f) = \operatorname{Var}_{\nu^n}\big(f \circ T\big).$$

Nous pouvons alors appliquer l'inégalité de Poincaré satisfaite par ν^n à la fonction $f \circ T$:

$$\operatorname{Var}_{\nu^n}(f \circ T) \leq C_P \sum_{i=1}^{n} \int_{\mathbb{R}^n} (\partial_i f)^2 \circ T(x) t'^2(x_i) d\nu^n(x).$$

De plus, la relation (7.3) entraine que

$$t'(x) = \frac{g(x)}{1 - G(x)} \times \frac{1 - H\big(t(x)\big)}{h\big(t(x)\big)} = \frac{\kappa_\nu(x)}{\kappa_\mu\big(t(x)\big)}, \quad x \in \mathbb{R}$$

à condition que $h(x) > 0, x \in \mathbb{R}$.

□

7.4 Bornes non-asymptotiques de variances

Nous allons illustrer le Théorème 7.3.1 pour obtenir des bornes pertinentes sur la variance de fonctionnelles de lois de probabilités usuelles. Nous commencerons par la théorie des extrêmes et les fluctuations du maximum, ensuite nous traiterons le cas des normes $l^p, p \geq 2$ d'un vecteur gaussien standard.

7.4.1 Domaine d'attraction des lois de Weibull et de Fréchet

Nous débutons par des lois usuelles, appartenant au domaine d'attraction de la loi de Weibull ou de Fréchet. Pour cela, nous choisissons ν^n comme étant le produit de n mesures exponentielles standards (rappelons que, dans ce cas, $H(x) = 1 - e^{-x}$ si $x \geq 0$, $H(x) = 0$ sinon). Nous obtenons alors le corollaire suivant.

Corollaire 7.4.1. *Si Y suit une loi exponentielle standard sur $\mathbb{R}_+$ alors, pour toute fonction f suffisamment régulière $f : \mathbb{R}^n \to \mathbb{R}$, pour $n \geq 1$,*

$$\operatorname{Var}\big(f(X)\big) \leq 4 \sum_{i=1}^{n} \mathbb{E}\left[\left(\frac{\partial_i f(X)}{\kappa_\mu(X_i)}\right)^2\right], \tag{7.5}$$

où $X_1, \ldots, X_n$ sont des variables aléatoires indépendantes de loi μ.

En particulier, pour (une approximation régulière de) $f(x) = \max_{i=1,\ldots,n} x_i$,

$$\mathrm{Var}\big(M_n\big) \leq C\mathbb{E}\bigg[\Big(\frac{1}{\kappa_\mu(M_n)}\Big)^2\bigg], \tag{7.6}$$

où $M_n = \max_{i=1,\ldots,n} X_i$ et $C > 0$ est une constante numérique.

Il ne reste plus qu'à spécifier la loi μ pour illustrer ce nouveau corollaire 7.4.1.

Mesure uniforme

Si μ désigne la mesure uniforme sur $[0,1]$ alors $\kappa_\mu(x) = 1_{x\in[0,1]}\frac{1}{1-x}$. Ainsi,

$$\mathrm{Var}(M_n) \leq 4\mathbb{E}[(1-M_n)^2] = O(1/n^2).$$

Rappelons que, $n(M_n - 1)$ converge en loi vers la distribution de Weibull.

Démonstration. En effet, pour tout $t \in [0,1]$, $\mathbb{P}(M_n \leq t) = t^n$ ceci impliquant que la loi de M_n admet $t \mapsto nt^{n-1}1_{[0,1]}$ comme densité par rapport à la mesure de Lebesgue.

Ainsi,

$$\mathbb{E}[M_n] = \int_0^1 nt^n dt = \frac{n}{n+1}$$

tandis que

$$\mathbb{E}[M_n^2] = \int_0^1 nt^{n+1} dt = \frac{n}{n+2}.$$

C'est pourquoi $\mathrm{Var}(M_n) = \frac{n}{n+1} - \frac{n^2}{(n+1)^2} = \frac{2n}{(n+2)(n+1)^2} = O(1/n^2)$. En comparaison, la méthode de transport (corollaire 7.4.1) entraine que

$$\mathrm{Var}(M_n) \leq 4\mathbb{E}[(1-M_n)^2]$$

et la majoration obtenue est bien du bon ordre de grandeur par ce qui précède. $\square$

Loi beta

Plus généralement, si μ désigne la loi bêta de paramètres $a, b > 0$ il n'est pas difficile de montrer que, pour tout $x \in [0,1]$,

$$\begin{aligned}
\frac{1}{\kappa_\mu(x)} &= \frac{\int_x^1 (1-t)^{b-1} t^{a-1} dt}{x^{a-1}(1-x)^{b-1}} \\
&\leq \min\left(\frac{1}{a}\frac{(1-x^a)}{x^{a-1}}, \frac{1}{b}\frac{(1-x)}{x^{a-1}}\right)
\end{aligned}$$

En particulier, lorsque $a = b = 1$ nous retrouvons l'estimation obtenue pour la mesure uniforme. Dans le cas général, les calculs semblent trop complexes en comparaison de l'obtention d'une éventuelle majoration non-asymptotique, de l'ordre de $n^{-1/b}$, de la variance du maximum.

Loi de Paréto

Rappelons que la théorie des extrêmes assure que $n^{-1/\alpha}M_n$ converge en loi vers la distribution de Fréchet. Il est donc pertinent de s'interroger sur l'efficacité de la méthode de transport pour des lois à queues lourdes.

Si μ correspond à la loi de Pareto de paramètre $\alpha > 3$ (afin que $\mathrm{Var}(X_1)$ soit finie) sur l'intervalle $[1, +\infty)$ alors $H(x) = 1 - \frac{1}{x^\alpha}$ si $x \geq 1$ et $H(x) = 0$ sinon. Des calculs élémentaires montrent que $\kappa_\mu(x) = \frac{\alpha}{x}$, ainsi le corollaire 7.4.1 entraine que

$$\mathrm{Var}(M_n) \leq \frac{4}{\alpha}\mathbb{E}[M_n^2]$$

Cette inégalité est mauvaise lorsque $\alpha \leq 4$ car la majoration obtenue est moins bonne que la borne triviale $\mathrm{Var}(M_n) \leq \mathbb{E}[M_n^2]$; lorsque $\alpha > 4$ le gain n'est vraiment pas significatif et semble peu pertinent car l'estimation directe de $\mathbb{E}[M_n^2]$ suffit à déterminer la taille de la variance du maximum.

7.4.2 Domaine d'attraction de la loi de Gumbel

Pour des mesures log-concaves sur $\mathbb{R}$ (comme la mesure gaussienne), il est commode de choisir ν comme étant la mesure exponentielle symétrique sur $\mathbb{R}$. Rappelons que ν admet la densité suivante $g(x) = \frac{1}{2}e^{-|x|}$ par rapport à la mesure de Lebesgue, et admet pour fonction de répartition $G(x) = \frac{1}{2}e^x$ si $x \leq 0$, $G(x) = 1 - \frac{1}{2}e^{-x}$ sinon. Des calculs élémentaires montrent que

$$\kappa_\nu(x) = \begin{cases} 1, & x > 0, \\ \frac{1}{2e^{-x}-1}, & x \leq 0. \end{cases} \tag{7.7}$$

Ceci entraine le corollaire suivant

Corollaire 7.4.2. *Si Y suit la loi exponentielle symétrique sur $\mathbb{R}$ alors, pour toute fonction régulière $f : \mathbb{R}^n \to \mathbb{R}$,*

$$\mathrm{Var}\big(f(X)\big) \leq 4\sum_{i=1}^n \mathbb{E}\left[\partial_i^2 f(X)\left(\frac{\kappa_\nu\big(t^{-1}(X_i)\big)}{\kappa_\mu(X_i)}\right)^2\right]. \tag{7.8}$$

avec $X = (X_1, \ldots, X_n)$ de loi μ^n. Rappelons que la constante 4 correspond à la constante de Poincaré de la mesure exponentielle symétrique.

Pour illustrer le résultat précédent, nous aurons besoin du lemme technique suivant. Ceci permettra d'obtenir des majorations pertinentes de la variance du maximum de loi symétrique log-concave $d\mu(x) = e^{-V(x)}dx$ (*e.g.* $V(x) = |x|^\alpha/\alpha$, $\alpha > 1$).

Lemme 7.4.3. *Dans le cas du transport de la mesure exponentielle symétrique vers la mesure $d\mu(x) = e^{-V(x)}dx$, où $V(x) = |x|^\alpha/\alpha$, $\alpha > 1$, l'application de transport $t : \mathbb{R} \to \mathbb{R}$ satisfait*

$$|t' \circ t^{-1}(x)| \leq \frac{C}{V'(|x|) + 1}, \quad x \in \mathbb{R}$$

avec $C > 0$ une constante numérique ne dépendant que de α.

Démonstration. Nous souhaitons majorer, pour tout $x \in \mathbb{R}$, le rapport suivant

$$t' \circ t^{-1}(x) = \frac{\kappa_\nu\big(t^{-1}(x)\big)}{\kappa_\mu(x)}, \tag{7.9}$$

avec κ_ν définie par (7.7) et $\kappa_\mu(x) = e^{-V(x)} \int_x^\infty e^{-V(t)} dt$, $x \in \mathbb{R}$. Rappelons que

$$t^{-1}(x) = G^{-1} \circ H(x),\ x \in \mathbb{R}$$

avec

$$G^{-1}(y) = \begin{cases} \ln(2y), & 0 \le y \le 1/2, \\ \ln\left(\frac{1}{2(1-y)}\right), & 1/2 \le y \le 1. \end{cases} \tag{7.10}$$

Soit $A > 0$ suffisamment grand, pour $x > A$ l'équation (7.9) est facilement majorée via des estimées classiques (cf. [8]), on obtient

$$|t' \circ t^{-1}(x)| = e^{V(x)} \int_x^\infty e^{-V(t)} dt \le \frac{C}{V'(x)},$$

avec $C > 0$.

Pour x appartenant au compact $[0, A]$, on obtient facilement que $|t' \circ t^{-1}(x)| \le C$. Finalement, ceci peut se résumer par

$$|t' \circ t^{-1}(x)| \le \frac{C}{V'(x)+1},\ x > 0$$

Pour $x = 0$ on $|t' \circ t^{-1}(x)| = 1$ puisque $t^{-1}(0) = G^{-1} \circ H(0) = G^{-1}(1/2) = 0$ par symétrie.

Si maintenant, $x < -A$, on obtient

$$|t' \circ t^{-1}(x)| \le \frac{2e^{V(x)}}{2e^{-t^{-1}(x)} - 1}$$

Puisque $\frac{1}{\kappa_\mu(x)} = \frac{e^{V(x)}}{\int_x^\infty e^{-V(t)} dt} \le 2e^{V(x)}$ pour $x \ge 0$.

Il suffit donc de majorer $t^{-1}(x)$ pour $x < -A$ afin de conclure. En utilisant la symétrie de la loi μ, on obtient

$$\begin{aligned} t^{-1}(x) \le \ln\big(2H(x)\big) &= \ln\Big(2\big[1 - H(-x)\big]\Big) \\ &\le \ln\left[\frac{2e^{-V(-x)}}{V'(-x)}\right]. \end{aligned}$$

Ce qui entraine la majoration suivante, pour $x < -A$,

$$|t' \circ t^{-1}(x)| \le \frac{2e^{V(-x)}}{V'(-x)e^{V(-x)} - 1} \le \frac{C}{V'(-x)}.$$

Si $-A \le x \le 0$, on obtient facilement que $|t' \circ t^{-1}(x)| \le C$
Finalement, tout ceci peut s'écrire

$$\left|\frac{\kappa_\nu(t^{-1}(x))}{\kappa_\nu(x)}\right| \le \frac{C}{V'(|x|)+1},$$

avec $C > 0$. □

Loi normale

En combinant le dernier lemme obtenu avec le corollaire 7.4.2, nous obtenons le résultat suivant pour la mesure gaussienne standard sur $\mathbb{R}^n$.

Proposition 7.4.4. *Pour toute fonction f suffisamment régulière,*

$$\mathrm{Var}_{\gamma_n}(f) \leq C \sum_{i=1}^{n} \mathbb{E}\left[(\partial_i f)^2(X)\left(\frac{1}{1+|X_i|}\right)^2\right]$$

En particulier, appliquée à (une approximation régulière de) $f(x) = \max_{i=1,\ldots,n} x_i$, nous obtenons, $n \geq 1$,

$$\mathrm{Var}(M_n) \leq C\mathbb{E}\left[\frac{1}{1+M_n^2}\right] \leq \frac{C}{1+\log n} \tag{7.11}$$

Remarque. Signalons que cette inégalité est déjà présente dans l'article [32] lorsque $n = 1$, les auteurs de l'article l'obtiennent en utilisant une version renforcée de l'inégalité de Brascamp-Lieb.

Démonstration. En effet, pour la fonction maximum, $\partial_i f = 1_{A_i}$, $i = 1, \ldots, n$ avec $A_i = \{X_i = \max_{j=1,\ldots,n} X_j\}$. Rappelons que $(A_i)_{i=1,\ldots,n}$ est une partition de $\mathbb{R}^n$. C'est pourquoi,

$$\begin{aligned}
\sum_{i=1}^{n} \mathbb{E}\left[(\partial_i f)^2(X)\left(\frac{1}{1+|X_i|}\right)^2\right] \leq \mathbb{E}\left[\frac{1}{1+M_n^2}\right] &\leq \frac{1}{1+\log n} + \mathbb{P}(M_n \leq \sqrt{\log n}) \\
&\leq \frac{1}{1+\log n} + \left(1 - \frac{\sqrt{\log n}}{1+\log n} e^{-\log n/2}\right)^n \\
&\leq \frac{C}{1+\log n}
\end{aligned}$$

Puisque, pour tout $t \geq 0$, $\mathbb{P}(M_n \leq t) = \big(1 - \mathbb{P}(X_1 > t)\big)^n$ avec X_1 une variable aléatoire gaussienne standard. Nous pouvons alors employer la borne suivante (cf. [131] (lemme 2.5) ou l'annexe de [48]) pour majorer la quantité précédente : pour tout $t \geq 0$,

$$\mathbb{P}(X_1 > t) \geq \frac{t}{\sqrt{2\pi}(1+t^2)} e^{-t^2/2}.$$

□

Rappelons que $\sqrt{2\log n}(M_n - b_n)$ converge, lorsque $n \to \infty$, en loi vers la loi de Gumbel (la valeur précise de b_n importe peu ici, nous renvoyons vers le chapitre 3 ou les ouvrages [67, 108]). Ainsi, $\mathrm{Var}(M_n) \leq \frac{C}{\log n}$.

Remarque. Le schéma de preuve précédent fonctionne également pour la fonction $f(x) = \mathrm{Med}(x_1, \ldots, x_n)$ entraîne que, $n \geq 1$,

$$\begin{aligned}
\mathrm{Var}\big(\mathrm{Med}(X)\big) &\leq \frac{C}{1+n} + C\mathbb{P}(\mathrm{Med}(X) \leq \sqrt{n})^{n/2} \\
&\leq \frac{C}{1+n} + o\left(\frac{1}{1+n}\right) \leq \frac{C}{1+n}
\end{aligned}$$

Ceci correspondant à l'ordre de grandeur attendu de la médiane (cf. [37]).

Loi Log-concave

Plus généralement, si l'on considère la mesure log-concave $d\mu(x) = e^{-V(x)}\frac{dx}{Z}$, avec $V(x) = |x|^\alpha/\alpha, \quad \alpha > 1$ et Z une constante de renormalisation. La même preuve, avec le lemme 7.4.3, entraine

Corollaire 7.4.5.

$$\mathrm{Var}_{\mu^n}(M_n) \leq C\sum_{i=1}^n \mathbb{E}_{\mu^n}\left[(\partial_i f)^2(X)\left(\frac{1}{1+V'(|X_i|)}\right)^2\right]$$

En particulier, appliquée à (une approximation régulière de) $f(x) = \max_{i=1,\dots,n} x_i$, cela fournit, pour $n \geq N_0$ suffisamment grand,

$$\mathrm{Var}(M_n) \leq C\mathbb{E}\left[\frac{1}{V'^2(M_n)+1}\right] \leq \frac{C}{1+C_\alpha \ln(n)^{2(\alpha-1)/\alpha}}, \tag{7.12}$$

avec $C_\alpha > 0$ et $C > 0$ des constantes numériques.

Démonstration.

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{1+|M_n|^{2(\alpha-1)}}\right] &\leq \frac{1}{1+(\log n)^{2(\alpha-1)/\alpha}} + \mathbb{P}(M_n \leq (\ln n)^{1/\alpha}) \\
&\leq \frac{1}{1+(\log n)^{2(\alpha-1)/\alpha}} + [1-\mathbb{P}(X_1 \geq (\ln n)^{1/\alpha})]^n \\
&\leq \frac{1}{1+(\log n)^{2(\alpha-1)/\alpha}} + \left(1-\frac{1}{2(\log n)^{(\alpha-1)/\alpha}n^{1/\alpha}}\right)^n \\
&\leq \frac{C}{1+C_\alpha(\log n)^{2(\alpha-1)/\alpha}}
\end{aligned}$$

Puisque, si X_1 désigne une variable aléatoire de loi μ, nous pouvons procéder comme dans le cas gaussien. En effet, $\mathbb{P}(X_1 \geq t) \sim \frac{1}{t^{\alpha}-1}e^{-t^\alpha/\alpha}$ lorsque $t \to \infty$. En particulier, pour t assez grand, ceci entraine que $\mathbb{P}(X_1 \geq t) \geq \frac{1}{2t^{\alpha-1}}e^{-t^\alpha/\alpha}$. □

Rappelons que nous avons démontré le fait suivant (cf. chapitre trois, proposition 3.1.6),

$$a_n(M_n - b_n) \to \Lambda_0,$$

en loi, lorsque $n \to \infty$, avec $a_n = \sqrt{\alpha(\log n)^{2(\alpha-1)/\alpha}}$ et $b_n = (\log n)^{1/\alpha} - \frac{\log(\alpha Z)+\frac{\alpha-1}{\alpha}\log\log n}{(\log n)^{(\alpha-1)/\alpha}}$.

Le corollaire redonne bien une majoration de la variance reflétant cette convergence. Soulignons également que ce type de majoration ne peut pas être obtenue par des méthodes hypercontractives (lorsque $\alpha > 2$) comme nous l'avons expliqué dans le chapitre quatre.

7.4.3 Bornes sur la variance de la norme $l^p, p \geq 2$ d'un vecteur gaussien standard

Comme illustration supplémentaire de notre approche, nous proposons d'obtenir des majorations de la norme $l^p, p \geq 1$ d'un vecteur gaussien standard, à partir de la proposition 7.4.4. Nous adoptons les notations suivantes, étant donné un vecteur $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ nous notons par $\|x\|_p^p = \sum_{i=1}^n |x_i|^p$sa norme et γ_n désignera la mesure gaussienne standard sur $\mathbb{R}^n$ dans ce qui suit. Dans l'article de Paouris *et al.* [131], les auteurs ont remarqué que la variance de $\|X\|_p$ n'est pas précisément estimée par la théorie classique de la concentration. Plus précisément, des outils classiques de la théorie de la concentration comme l'inégalité de Poincaré ou encore l'inégalité isopérimétrique gaussienne fournissent les majorations suivantes

$$\mathrm{Var}(\|X\|_p) \leq \max(n^{2/p-1}, 1).$$

Selon [131], cette borne n'est optimale qu'à condition que $1 \leq p \leq 2$. Les auteurs de [131] améliorent cette borne en utilisant des estimées précises ainsi que l'inégalité de Sobolev logarithmique (au travers de l'inégalité de Talagrand ainsi que des estimations précises des moments de fonctionnelles d'un vecteur gaussien standard). Ils obtiennent également des inégalités de concentration reflétant ces nouvelles bornes sur la variance (cf. chapitre deux). Nous proposons, ici, d'illustrer notre approche de la superconcentration par le transport optimal pour retrouver certains de leurs résultats (théorème 2.2.7 de [131]) sur la variance de $\|X\|_p$:

$$\mathrm{Var}(\|X\|_p) \leq \begin{cases} C\frac{2^p}{p}n^{2/p-1}, & 2 < p \leq c\log n, \\ C/\log n, & p > c\log n, \end{cases}$$

avec $C, c > 0$ des constantes numériques indépendantes de n et de p.

Pour obtenir ceci, nous utiliserons certaines estimations de l'article [131]. Similairement à [131], nous adoptons la notation suivante : si X_1 est une variable aléatoire gaussienne standard, nous posons $\sigma_p^p = \mathbb{E}[|X_1|^p]$ pour tout $p > 0$. Nous noterons également par $I_r^r = \int_{\mathbb{R}^n} \|x\|^r d\gamma_n(x)$ pour tout $r \in \mathbb{R}$. Par exemple, pour $p > 0$

$$I_{-p}^{-p} = \int_{\mathbb{R}^n} \frac{1}{\|x\|^p} d\gamma_n(x)$$

Voici quelques estimations, obtenues par les auteurs de [131] (page 7 et page 17), qui nous seront utile par la suite. Notons que l'équation (7.14) est conséquence de l'inégalité de Sobolev logarithmique satisfaite par la mesure gaussienne standard sur $\mathbb{R}^n$.

Lemme 7.4.6. *Avec les notations introduites ci-dessus, nous avons les résultats suivants*

$$\sigma_p^p = \frac{2^{p/2}}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right) \sim \sqrt{2}\left(\frac{p}{e}\right)^{p/2}, \quad p \to \infty \tag{7.13}$$

$$I_{-2(p-1)}^{-2(p-1)} \leq n^{2/p-2}/\sigma_p^{2(p-1)} \quad p \leq c\log n \tag{7.14}$$

Nous utiliserons également les estimées classiques suivantes (cf. [8, 112]), pour tout $\delta > 0$,

1. $\mathbb{P}(X \in B_\infty(0,\delta)) \leq [1 - \frac{\delta}{1+\delta^2}e^{-\delta^2/2}]^n$.

2. $\mathbb{P}(|X_1| \geq \delta) \leq e^{-\delta^2/2}$.

Nous aurons aussi besoin de la proposition suivante par la suite (cf. [38]).

Proposition 7.4.7. *[Harris] Soient $f : \mathbb{R}^n \to \mathbb{R}$ une fonction croissante et $g : \mathbb{R}^n \to \mathbb{R}$ une fonction décroissante (au sens où la fonction considérée est croissante, respectivement décroissante, en chacune de ses coordonnées lorsque les autres sont fixées). Soient $X_1, \ldots, X_n$ des variables indépendantes réelles et posons $X = (X_1, \ldots, X_n)$. Alors*

$$\mathbb{E}\big[f(X)g(X)\big] \leq \mathbb{E}\big[f(X)\big]\mathbb{E}\big[g(X)\big]. \tag{7.15}$$

Nous allons d'abord traiter le cas $2 < p \leq c\log n$, afin d'obtenir :

Proposition 7.4.8.

$$\mathrm{Var}(\|X\|_p) \leq C\frac{2^p}{p} n^{2/p-1}, \quad 2 < p \leq c \log n,$$

avec $C > 0$ une constante indépendante de p et de n.

Démonstration. D'après la proposition 7.11,

$$\begin{aligned}
\mathrm{Var}(\|X\|_p) &\leq C \sum_{i=1}^{n} \int_{\mathbb{R}^n} \frac{|x_i|^{2(p-1)}}{1+|x_i|^2} \frac{d\gamma_n(x)}{\|x\|_p^{2(p-1)}} \\
&\leq n \left(\int_{\mathbb{R}} \frac{|x|^{2(p-1)}}{1+|x|^2} d\gamma_1(x) \right) \times I_{-2(p-1)}^{-2(p-1)}
\end{aligned}$$

où nous avons utilisé l'inégalité d'association négative de Harris (cf. proposition 7.4.7 ou [38]) pour les deux fonction suivantes $F_i(x_1, \dots, x_n) = \frac{|x_i|^{2(p-1)}}{1+|x_i|^2}$ pour $i = 1, \dots, n$ qui est croissante en chaque coordonnée tandis que $x \mapsto \frac{1}{\|x\|_p^{2(p-1)}}$ est décroissante. Ensuite, nous avons utilisé le fait que les variables X_i's sont indépendantes et de mêmes lois

Nous obtenons les majorations suivantes

$$\begin{aligned}
\int_{\mathbb{R}} \frac{|x|^{2(p-1)}}{1+|x|^2} d\gamma_1(x) &= \int_{\mathbb{R}} \frac{|x|^2}{1+|x|^2} |x|^{2(p-2)} d\gamma_1(x) \\
&\leq \int_{\mathbb{R}} |x|^{2(p-2)} d\gamma_1(x)
\end{aligned}$$

Ceci entraine finalement que

$$\mathrm{Var}(\|X\|_p) \leq n \sigma_{2(p-2)}^{2(p-2)} I_{-2(p-1)}^{-2(p-1)},$$

qui peut s'exprimer par

$$\mathrm{Var}(\|X\|_p) \leq C \frac{\sigma_{2(p-2)}^{2(p-2)}}{\sigma_p^{2(p-1)}} n^{2/p-1},$$

grâce à (7.14) (puisque $p \leq c \log n$). Remarquons que $C \frac{\sigma_{2(p-2)}^{2(p-2)}}{\sigma_p^{2(p-1)}} n^{2/p-1} \leq C \frac{2^p}{p} n^{2/p-1}$ pour conclure. □

Nous nous tournons à présent vers la deuxième majoration de la variance.

Proposition 7.4.9. *Pour $n \geq N_0$, nous avons l'inégalité suivante*

$$\mathrm{Var}(\|X\|_p) \leq \frac{C}{\log n}, \quad p > c \log n,$$

avec $C > 0$ une constante numérique indépendant de p et de n.

Démonstration. Lorsque $p > c\log n$ les estimées (7.14) ne sont plus valides, il est donc nécessaire de procéder autrement. Soit alors $\delta > 0$ que nous choisirons ultérieurement et notons par $B_\infty(0,\delta) = \{x \in \mathbb{R}^n,\ \|x\|_\infty < \delta\}$. Ainsi,

$$\begin{aligned}
\mathrm{Var}(\|X\|_p) &\leq C\sum_{i=1}^n\left(\int_{B_\infty(0,\delta)}\frac{|x_i|^{2(p-1)}}{1+|x_i|^2}\frac{1}{\|x\|_p^{2(p-1)}}d\gamma_n(x) + \int_{B_\infty^c(0,\delta)}\frac{|x_i|^{2(p-1)}}{1+|x_i|^2}\frac{1}{\|x\|_p^{2(p-1)}}d\gamma_n(x)\right)\\
&= C\left(\sum_{i=1}^n \mathcal{I}_i + \mathcal{J}_i\right)
\end{aligned}$$

Nous rappelons les relations suivantes entre les normes l^p et l^q, pour $p < q$, provenant de l'inégalité de Hölder, que nous utiliserons librement par la suite,

$$\|x\|_q \leq \|x\|_p \leq n^{1/p-1/q}\|x\|_q, \quad \forall x \in \mathbb{R}^n$$

D'une part, puisque $p < 2(p-1)$,

$$\begin{aligned}
\sum_{i=1}^n \mathcal{I}_i &\leq \int_{B_\infty(0,\delta)}\frac{\|x\|_{2(p-1)}^{2(p-1)}}{\|x\|_p^{2(p-1)}}d\gamma_n(x)\\
&\leq \mathbb{P}(X \in B_\infty(0,\delta))
\end{aligned}$$

D'autre part, puisque $p < 2(p-2)$,

$$\begin{aligned}
\sum_{i=1}^n \mathcal{J}_i &\leq \int_{B_\infty^c(0,\delta)}\frac{\|x\|_{2(p-2)}^{2(p-2)}}{\|x\|_p^{2(p-1)}}d\gamma_n(x) = \int_{B_\infty^c(0,\delta)}\left(\frac{\|x\|_{2(p-2)}}{\|x\|_p}\right)^{2(p-2)}\frac{1}{\|x\|_p^2}d\gamma_n(x)\\
&\leq \int_{B_\infty^c(0,\delta)}\frac{d\gamma_n(x)}{\|x\|_p^2}\\
&\leq \frac{1}{\delta^2}\mathbb{P}(X \in B_\infty^c(0,\delta))
\end{aligned}$$

De plus, remarquons que la majoration suivante est satisfaite

$$\begin{aligned}
\mathbb{P}(X \in B_\infty^c(0,\delta)) = \mathbb{P}(\max_{i=1,\dots,n}|X_i| \geq \delta) &= \mathbb{P}(\exists j \in \{1,\dots,n\},\ |X_j| \geq \delta)\\
&\leq n\mathbb{P}(|X_1| \geq \delta) \leq 2ne^{-\delta^2/2}.
\end{aligned}$$

Nous avons donc obtenu la majoration suivante,

$$\mathrm{Var}(\|X\|_p) \leq C\left(\left[1 - \frac{\delta}{1+\delta^2}e^{-\delta^2/2}\right]^n + \frac{2ne^{-\delta^2/2}}{\delta^2}\right)$$

Il suffit alors de choisir $\delta = \sqrt{2\log n}$ (avec n assez grand) pour conclure. En effet, nous avons d'une part

$$\mathbb{P}\big(X \in B_\infty(0,\delta)\big) \leq (1 - e^{-\delta^2/3})^n \sim e^{-n^{1/3}}$$

et d'autre part

$$\frac{2ne^{-\delta^2/2}}{\delta^2} = \frac{1}{\log n}.$$

En d'autres termes

$$\mathrm{Var}(\|X\|_p) \leq C\Big(o(\frac{1}{\log n}) + \frac{1}{\log n}\Big) \leq \frac{C}{\log n},$$

qui est le résultat annoncé. □

7.5 Inégalités de déviation

Comme mentionné dans l'introduction, la flexibilité du théorème 7.3.1 permet d'obtenir des inégalités de déviations. C'est le contenu du théorème suivant

Théorème 7.5.1. *Supposons qu'il existe une fonction $x \mapsto \psi(x)$ allant de $\mathbb{R}$ dans $\mathbb{R}$, décroissante telle que, pour tout $x \in \mathbb{R}$*

$$\frac{\kappa_\nu\big(t^{-1}(x)\big)}{\kappa_\mu(x)} \leq \psi(x)$$

et qu'il existe ϵ_n tel que

$$\mathbb{E}\Big[\psi(M_n)^2\Big] \leq \epsilon_n.$$

Alors, pour tout $t \geq 0$ et $n \geq 1$,

$$\mathbb{P}(\sqrt{\epsilon_n}\big(M_n - \mathbb{E}_{\nu^n}[M_n]\big) \geq t) \leq 3e^{-t}.$$

Remarque. Notons que ce théorème n'est pertinent que si la loi limite μ appartient au domaine d'attraction de la distribution de Gumbel. En effet, la queue de droite de la loi de Gumbel se comporte précisément comme $t \mapsto e^{-t}$ (tandis que la queue à gauche décroit plus vite en e^{-e^t}).

La démonstration de ce résultat découle de la combinaison des résultats suivants (cf. [112, 66, 65]). Un lemme provenant de la concentration, que nous avons déjà employé dans ce manuscrit de thèse.

Lemme 7.5.2. *Soit X une variables aléatoire centrée telle que, pour tout $0 < \theta < \frac{1}{2\sqrt{K_n}}$,*

$$\mathrm{Var}(e^{\theta X/2}) \leq \frac{\theta^2}{4} K_n \mathbb{E}[e^{\theta X}],$$

alors $\mathbb{P}(X \geq t\sqrt{K_n}) \leq 3e^{-ct}$ pour tout $t \geq 0$, avec $c > 0$ une constante numérique.

Que nous combinerons avec l'inégalité d'association négative de Harris 7.4.7 que nous avons énoncée plus tôt.

Démonstration. (du théorème 7.5.1) Appliquons successivement le théorème 7.5.2, puis l'inégalité de Harris (7.4.7)

$$\begin{aligned}\mathrm{Var}(e^{\theta M_n/2}) &\leq C\frac{\theta^2}{4}\mathbb{E}\Big[e^{\theta M_n}\psi(M_n)^2\Big] \\ &\leq C\frac{\theta^2}{4}\mathbb{E}_{\nu^n}[e^{\theta M_n}]\mathbb{E}\big[\psi(M_n)^2\big] \leq C\frac{\theta^2}{4}\epsilon_n\mathbb{E}[e^{\theta M_n}]\end{aligned}$$

puisque, pour tout $i = 1, \ldots, n$,

$$x_i \mapsto f(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, n) = \max(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$$

est croissante. Ceci implique que, pour tout $\theta > 0$,

$$x_i \mapsto e^{\theta \max(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)}$$

l'est aussi. De plus, $\psi \circ f$ est décroissante, par composition, au vu des hypothèses portant sur ψ. Les hypothèses pour appliquer l'inégalité de Harris sont donc satisfaites et le reste de la démonstration s'ensuit facilement. □

Comme application direct du théorème 7.5.1 nous obtenons, à l'aide des estimées sur la variance (7.11) et (7.12), le corollaire suivant. En effet, la combinaison d'un lemme de concentration avec l'inégalité d'association négative de Harris (cf. [38]) implique une inégalité de déviation (de la moyenne) à droite qui préserve les bornes obtenues sur la variance du maximum. Formellement,

Théorème 7.5.3. *1. Pour γ_n la mesure gaussienne standard sur $\mathbb{R}^n$, l'inégalité suivante est satisfaite, pour tout $n \geq 1$ et tout $t \geq 0$,*

$$\mathbb{P}(\sqrt{\log n}\big(M_n - \mathbb{E}_{\nu^n}[M_n]\big) \geq t) \leq 3e^{-Ct},$$

avec $C > 0$ une constante numérique.

2. Plus généralement, si μ^n désigne la mesure log-concave de potentiel $|x|^\alpha, \alpha > 1$, alors pour tout $n \geq 1$ et tout $t \geq 0$ nous obtenons

$$\mathbb{P}(\sqrt{(\log n)^{2(\alpha-1)/\alpha}}\big(M_n - \mathbb{E}_{\nu^n}[M_n]\big) \geq t) \leq 3e^{-C_\alpha t},$$

avec $C_\alpha > 0$ une constante numérique ne dépendant que de α.

Remarque. Soulignons le fait qu'une telle inégalité de déviation ne peut-être obtenue par des moyens classiques. En effet, la concentration gaussienne classique (cf. e.g. [112]) produit typiquement

$$\mathbb{P}\left(|M_n - \mathbb{E}\left[M_n\right]| \geq t\right) \leq 2e^{-t^2/2}, \quad t \geq 0. \tag{7.16}$$

Bien que la décroissance soit gaussienne, de telles bornes ne reflète pas les fluctuations du maximum M_n. Il est aussi bien connu (cf. [112], chapitre un) que les lois strictement log-concaves telles que $d\mu(x) = e^{-|x|^\alpha/\alpha)}\frac{dx}{Z}$, $\alpha > 1$ satisfont des inégalités analogues, celles-ci ne reflétant toujours pas la convergence des extrêmes.

7.6 Modèle de Ginibre complexe

Cette section propose un exemple d'application provenant de la théorie des matrices aléatoires. Nous choisissons de séparer cette partie de ce qui précède car, dans cet exemple, les facteurs de la mesure produit $\mu_1 \otimes \cdots \otimes \mu_n$ sont de lois différentes.

A présent, plaçons dans le même contexte que celui présenté dans le théorème 2.2.4 du chapitre deux. Nous allons voir qu'il n'est pas difficile d'obtenir une majoration de la variance et une inégalité de déviation à droite pour la module de la plus grande valeur propre $|z_{(1)}|$ de l'ensemble de Ginibre complexe. Le point important est la formule de représentation suivante (cf. [47]) :

$$|z_{(1)}| = \max_{i=1,\ldots,n} R_i$$

en loi, avec $R_1, \ldots, R_n$ des variables aléatoires indépendantes mais de loi différentes. D'après [47], pour tout $k = 1, \ldots, n$, R_k admet une densité par rapport à la mesure de Lebesgue proportionnelle à

$$t \mapsto t^{2k-1} e^{-nt^\alpha} 1_{[0,\infty)}(t),\ \alpha \geq 1.$$

Il est donc possible de transporter la mesure produit exponentielle vers la mesure $\mu_1 \otimes \cdots \otimes \mu_n$ avec $\mu_k = \mathcal{L}(R_k)$ pour tout $k = 1, \ldots, n$. En constatant que, pour tout $k = 1, \ldots, n$, μ_k est une loi log-concave sur $\mathbb{R}_+$, de potentiel associé $W_k(x) = nt^\alpha - (2k-1)\log t$, il n'est pas difficile de montrer (à l'aide des estimées de [8]) que

$$\frac{1}{\kappa_{\mu_k}(x)} \leq \frac{C_\alpha}{nx^{\alpha-1}+1},\ x > 0$$

avec $C_\alpha > 0$ une constante numérique. L'utilisation de 7.4.1 fournit donc

$$\begin{aligned}
\mathrm{Var}(|z_{(1)}|) \leq \frac{C_\alpha}{n^2}\mathbb{E}\left[\frac{1}{|z_{(1)}|^{2(\alpha-1)}}\right] &\leq \frac{C_\alpha}{n\log n} + C_\alpha \mathbb{P}\left(|z_{(1)}| \leq \left(\frac{\log n}{n}\right)^{1/2(\alpha-1)}\right) \\
&\leq \frac{C_\alpha}{n\log n} + \prod_{i=1}^n \left[1 - \mathbb{P}\left(R_i \geq \frac{\log n}{n}\right)^{1/2(\alpha-1)}\right)\right] \\
&\leq \frac{C_\alpha}{n\log n} + o\left(\frac{1}{n\log n}\right) \\
&\leq \frac{C_\alpha}{n\log n}
\end{aligned}$$

Aussi, le théorème 7.5.1 permet d'obtenir l'inégalité de déviation suivante

$$\mathbb{P}\left(\sqrt{n\log n}\big(|z_{(1)}| - \mathbb{E}[|z_{(1)}|\big) \geq t\right) \leq 6e^{-t},\ t \geq 0.$$

En conclusion nous avons obtenu une inégalité de déviation et une majoration de la variance non asymptotique en accord avec le théorème 2.2.4.

Proposition 7.6.1. *Soit $\{z_1, \ldots, z_n\}$ un gaz coulombien de densité de probabilité proportionnelle à*

$$(z_1,\dots,z_n)\mapsto\prod_{j=1}^n e^{-nQ(z_j)}\prod_{1\le j<k\le n}|z_j-z_k|^2,$$

avec $Q=V(|z|)$ et $V(t)=t^\alpha$, $\alpha\ge 1$. Alors, pour tout $n>1$, les inégalités suivantes sont satisfaites

$$\mathrm{Var}(|z_{(1)}|)\le\frac{C_\alpha}{n\log n}$$

et

$$\mathbb{P}\Big(\sqrt{n\log n}\big(|z_{(1)}|-\mathbb{E}[|z_{(1)}|)\ge t\Big))\le 6e^{-t},\ t\ge 0.$$

7.7 Transport et isopérimétrie

Comme nous l'avons vu dans ce qui précède, certaines propriétés satisfaites par la mesure exponentielle permettent de renforcer la concentration d'autres mesures. En fait, ce phénomène n'est pas nouveau, Talagrand a observé dans son article [151] que des méthodes de transport permettent de renforcer la concentration gaussienne. Plus précisément, comme présenté dans [151] ou [112], Talagrand a démontré une inégalité isopérimétrique pour la mesure exponentielle qui permet d'obtenir le renforcement sus-mentionné. Soit ν^n la mesure produit sur $\mathbb{R}^n$ dont chaque facteur correspond à la mesure ν de densité $\frac12 e^{-|x|}$ par rapport à la mesure de Lebesgue. Alors, pour tout boréliens A tels que $\nu^n(A)\ge\frac12$ et tout $r\ge 0$,

$$\nu^n(A+\sqrt{r}B_2+rB_1)\ge 1-e^{-r/K}$$

avec $K>0$ une constante numérique, où B_2 désigne la boule unité euclidienne et B_1 celle pour la norme l^1 dans $\mathbb{R}^n$. A présent, considérons l'application croissante $t:\mathbb{R}\to\mathbb{R}$ qui transporte ν sur la mesure gaussienne γ_1. On peut facilement vérifier que cette application vérifie

$$|t(x)-t(y)|\le C\min\big(|x-y|,|x-y|^{1/2}\big),\quad x,y\in\mathbb{R}\tag{7.17}$$

pour $C>0$ une constante numérique. On définit alors, l'application $T:\mathbb{R}^n\to\mathbb{R}^n$ by $T(x)=\big(t(x_1),\dots,t(x_n)\big)$ qui, par construction, transporte ν^n sur γ_n. Considérons à présent, un borélien A de $\mathbb{R}^n$ tel que $\gamma_n(A)\ge\frac12$. Alors, pour tout $r\ge 0$

$$\gamma_n\Big(T\big(T^{-1}(A)+\sqrt{r}B_2+rB_1\big)\Big)=\nu^n\big(T^{-1}(A)+\sqrt{r}B_2+rB_1\big)\ge 1-e^{-r/K}.$$

Cependant, il s'ensuit, d'après (7.17), que

$$T\big(T^{-1}(A)+\sqrt{r}B_2+rB_1\big)\subset A+C'\sqrt{r}B_2.$$

Ainsi (7.17) améliore la concentration gaussienne classique : si $\gamma_n(A)\ge\frac12$ alors $\gamma_n(A+rB_2)\ge\Phi(r)\ge 1-e^{-r^2/2}$, où Φ désigne la fonction de répartition de la loi gaussienne standard. Afin d'illustrer ce renforcement, soit

$$A=\{x\in\mathbb{R}^n\,;\,\max_{i=1,\dots,n}|x_i|\le m\}$$

avec $m=m(n)$ choisit de telle sorte que $\gamma_n(A)\ge\frac12$ (ceci signifiant que $m(n)$ est d'ordre $\sqrt{\log n}$). Alors, lorsque $r\ge 1$ est petit comparé à $\log n$, on voit facilement que

$$
\begin{aligned}
T\big(T^{-1}(A)+\sqrt{r}B_2+rB_1\big) &\subset A+C'\Big(\frac{\sqrt{r}}{\sqrt{\log n}}B_2+\frac{r}{\sqrt{\log n}}B_1\Big)\\
&\subset A+C_2\sqrt{\frac{r}{\log n}}\sqrt{r}B_2
\end{aligned}
$$

Tout ceci, se retranscrit de manière équivalente en terme d'inégalité de concentration, laquelle apparait dans l'article récent [157] :

$$
\mathbb{P}(\max_{i=1,\dots,n}|X_i|-\sqrt{\log n}\geq C\frac{t}{\sqrt{\log n}})\leq Ce^{-ct},\quad t\geq 0
$$

D'après [131], $\max_{i=1,\dots,n}|X_i|$ converge en loi vers une distribution de Gumbel avec les mêmes constantes que le cas classique (sans les valeurs absolues). L'inégalité précédente reflète donc, de manière non-asymptotique, ce résultat des extrêmes. Il se trouve qu'un autre résultat isopérimétrique du à Bobkov [24] permet de préciser ce genre d'inégalité et d'obtenir des inégalités de déviations dont la vitesse de décroissance reflète parfaitement le comportement de la loi de Gumbel (c'est à dire, une décroissance en e^{-r} à droite et une décroissance en $e^{-e^{-r}}$ à gauche).

7.7.1 Isopérimérie et élargissement uniforme

La loi exponentielle étant symétrique, il ne semble pas possible que le résultat de Talagrand permette de faire la distinction entre les différents asymptotiques de la loi de Gumbel. Dans [151], Talagrand considère des grossissements particuliers d'un ensemble, il utilise un mélange de boule l^1 et l^2. Dans [24], Bobkov étudie le problème isopérimétrique pour la mesure exponentielle sur $\mathbb{R}_+$ avec des grossissements uniformes (ceux-ci sont de la forme rB_∞, $r\geq 0$ avec $B_\infty=[0,1]^n$, au lieu de $\sqrt{r}B_2+rB_1$). Il choisit également de restreindre son étude à des ensembles appelés idéaux. Ces ensembles sont définit comme suit A est un idéal de $\mathbb{R}_+^n$ s'il satisfait la condition suivante

si $x=(x_1,\dots,x_n)\in A$, $y=(y_1,\dots,y_n)\in\mathbb{R}_+^n$, $y_i\leq x_i$ pour tout $i=1,\dots,n$, alors $y\in A$

Rappelons qu'il existe une vaste littérature concernant les problèmes isopérimétriques et les ensembles extremaux y apparaisant. En particulier, en dimension un, il existe des conditions suffissantes pour lesquelles les ensembles extremaux sont des demi-espaces $(-\infty,a]$ [24, 33, 29]. Par exemple, la mesure symétrique exponentielle et la mesure gaussienne standard vérifient ce critère. Ce n'est pas le cas de la mesure exponentielle sur $\mathbb{R}_+$, notamment à cause de son manque de symétrie. Le résultat de Bobkov s'énonce comme suit, avec ν^n la mesure produit dont chaque facteur correspond à la mesure exponentielle sur $\mathbb{R}_+$,

Théorème 7.7.1. *(Bobkov) Pour tout idéal non vide $A\subset\mathbb{R}_+^n$ tel que $\nu^n(B_\infty)=\nu^n(A)$ et tout $r\geq 0$, l'inégalité suivante est satisfaite :*

$$
\nu^n(A+rB_\infty)\geq\nu^n(B+rB_\infty),
$$

autrement dit,

$$
\nu^n(A+rB_\infty)\geq\Big[e^{-r}\big[\nu^n(A)\big]^{1/n}+(1-e^{-r})\Big]^n.
$$

Remarque. Si $n\to\infty$ et $\nu^n(A)=p$ est constant, le membre de droite de l'inégalité précédente décroit et converge vers la fonction double exponentielle. C'est à dire

$$
\nu^n(A+rB_\infty)\geq\exp(-e^{-r}\log(1/p)).
$$

Comme présenté dans [24], il est intéressant de considérer un certains types de mesures μ, lesquelles peuvent être atteinte en transportant la mesure exponentielle sur $\mathbb{R}_+$ ν. Ces mesures doivent vérifier deux conditions :

1. La fonction de répartition H de la mesure μ, $H(x) = \mu[0, x]$, est continue et strictement croissante sur $[0, b_H)$, où $b_H = \sup\{x : H(x) < 1\}$.

2. $$\lim_{t\to\infty} \sup_{0\leq x<b_H} \frac{1 - H(x+t)}{1 - H(x)} = 0.$$

D'une certaine manière les travaux de Bobkov vont dans le même sens que ceux de Gozlan [87]. En effet, dans l'article [24], Bobkov montre que l'application de transport monotone induit une métrique d sur $\mathbb{R}$. Il montre notamment que cette métrique est sous-euclidienne, au sens où $d(x, y) \leq C|x - y|$ pour tout $x, y \in \mathbb{R}$. Ainsi, le renforcement de la concentration est obtenu par un changement de métrique.

Remarque. Cette famille de lois, que nous noterons $\mathcal{F}$, est pertinente pour notre problème, puisqu'il s'agit de mesure pouvant être vue comme des mesures images de la loi exponentielle sur $\mathbb{R}_+$. D'après l'article de Bobkov, pour $H \in \mathcal{F}$ le fait suivant est satisfait : il existe $C > 0$ telle que, pour tout t suffisament grand

$$1 - F(t) \leq \exp^{-t/C}. \tag{7.18}$$

Nous allons plutôt restreindre notre étude à l'ensemble $\mathcal{F}_0 \subset \mathcal{F}$, correspondant aux fonctions de répartition H pouvant se représenter sous la forme

$$F(t) = 1 - \exp\big(-V(t)\big), t \in \mathbb{R}$$

avec $t \mapsto V(t)$ une fonction convexe, continue et strictement croissante sur $[0, b_H)$ satisfaisant $V(0) = 0$ et $\lim_{t\to b_H} V(t) = +\infty$. Il est clair que la valeur absolue de gaussienne standard, la mesure exponentielle symétrique sur $\mathbb{R}$ ou encore les lois Gammas appartiennent à $\mathcal{F}_0$. Il est aisé de voir que les lois appartenant à $\mathcal{F}_0$ font parties du domaine d'attraction de la loi de Gumbel, au prix de conditions supplémentaires mineures.

Proposition 7.7.2. *Soit $H \in \mathcal{F}_0$ et supposons que $V \in C^2(\mathbb{R})$ tel que $\frac{V''(t)}{V'(t)^2} \to 0$ lorsque $t \to b_H$. Alors H appartient au domaine d'attraction de la loi de Gumbel.*

Démonstration. D'après [108], le critère de von Mises reposant sur H afin d'appartenir au domaine d'attraction de la loi de Gumbel est le suivant

$$\lim_{t\to b_H} \frac{H''(t)\big(1 - H(t)\big)}{H'(t)^2} = -1$$

Il est aisé de montrer que c'est le cas si $H(t) = 1 - \exp\big(-V(t)\big)$ avec V satisfaisant les hypothèses de la proposition. $\square$

Le choix particulier de retreindre le problème isopérimétrique, de la mesure exponentielle, aux idéaux, permet d'obtenir des inégalités de déviations pour le supremum. Plus précisément, en suivant l'article de Bobkov, soit $(Y_n)_{n\geq1}$ des variables aléatoires indépendantes et de même loi commune $H \in \mathcal{F}_\prime$. Soit $L_{\mathcal{F}_0}$ l'ensemble des familles de variables aléatoires X que l'on peut représenter sous la forme d'une série presque sûrement convergente,

$$X = \sum_{n=1}^{\infty} a_n Y_n,$$

avec $a_n \geq 0$ pour tout $n \geq 1$. Puisque H admet des moments exponentiels finis, la convergence presque sûr de la série précédente est équivalente à la convergence de la série $\sum_{n\geq 1} a_n$. Etant donné un ensemble T, les travaux de Bobkov fournissent des inégalités de déviations pour le processus stochastique, presque sûrement borné, $(X_t)_{t\in T}$, consistant en des variables de l'ensemble $L_{\mathcal{F}_0}$ et leur supremum

$$M_T = \sup_{t\in T} X_t.$$

Nous restreignons notre étude au cas $T = \{1, \ldots, n\}$ et, pour tout $i = 1, n$, $X_i = Y_i$ (*i.e.* $a_n^i = 1$ si $n = i$, 0 sinon). Ainsi, $M_T = M_n = \max_{i=1,\ldots,n} X_i$ où X_i sont des variables indépendantes de même loi commune $H \in \mathcal{F}_0$. Dans ce cas particulier, le résultat de Bobkov peut-être formulé comme suit, avec l'utilisation de la remarque 7.18,

Théorème 7.7.3. *(Bobkov) Pour tout p, $0 < p < 1$, tout $t \geq 0$,*

$$\mathbb{P}(M_n - m_p \geq t) \geq C \log(1/p) \exp(-ct), \tag{7.19}$$

$$\mathbb{P}(M_n - m_p < -t) \leq C \exp\big(-e^{tc}\log(1/p)\big), \tag{7.20}$$

avec m_p désignant le quantile d'ordre p de M_n et H la fonction de répartition de la variables aléatoire X.

En particulier, si l'on choisit $p^{1/n} = F^{-1}(1-1/n)$, m_p correspondant au recentrage apparaissant dans la théorie des extrêmes. Par exemple, pour le cas exponentiel ou celui de la loi gamma, nous obtenons

Proposition 7.7.4. *Pour tout $t \geq 0$*

$$\mathbb{P}(M_n - \log n \geq t) \leq Ce^{-ct}$$

et

$$\mathbb{P}(M_n - \log n \leq -t) \leq Ce^{-e^{ct}}$$

Démonstration. La preuve est quasiment immédiate, d'après le choix de p. De plus, $\ln(1/p) = 1 + o(1)$. Nous avons aussi utilisé le fait que $1 - e^{-x} \leq x$ pour $x > 0$, ici $x = 2e^{-t}$. $\square$

Ces inégalités de déviations expriment bien, de manière non asymptotique, la convergence en loi de la mesure exponentielle (respectivement mesure Gamma) vers la loi de Gumbel. On retrouve bien la différence de comportements entre les queues de distributions à droite et à gauche. De plus, de telles inégalités impliquent que $\mathbb{P}(|M_n - \log n| \geq t) \leq Ce^{-ct}$, que l'on peut intégrer afin d'obtenir $\mathrm{Var}(M_n) \leq C$.

Tout ceci peut également être obtenu pour le maximum de valeur absolue de variables aléatoires gaussiennes indépendantes et de même loi. Nous laissons au lecteur le soin de compléter les détails. Rappelons qu'une inégalité similaire à déja été obtenue par Schetchtman dans [139].

Il pourrait être intéressant d'étudier un problème isopérimétrique, similaire à celui de Bobkov, pour des lois à queues plus lourdes. Par exemple, si $d\nu = e^{-x^q}dx$, $0 < q < 1$ sur $\mathbb{R}_+$, est-il possible de résoudre le problème isopérimétrique pour des élargissements uniformes. Peut-on obtenir des inégalités de déviations correspondant à un domaine d'attraction différent de le loi de Gumbel ? Il pourrait être aussi pertinent de tirer profit du théorème de Bobkov portant sur les déviations du supremum de variables de la forme $X = \sum_{n=1}^{\infty} a_n Y_n$ afin de pouvoir traiter les suites gaussiennes stationnaires et de sortir du cadre des mesures produits.

7.8 Comparaison de la méthode de transport avec la littérature existante

Dans cette section nous allons comparer notre méthode de transport pour obtenir de la superconcentration pour des variables indépendantes de même loi. La comparaison avec l'approche hypercontractive est immédiate, celle-ci permet de traiter des exemples gaussiens présentant des corrélations. Toutefois, elle ne permet pas d'atteindre une décroissance plus rapide que $1/\log n$ et ne propose qu'une décroissance exponentielle au niveau des inégalités de concentration. Par exemple, il ne semble pas possible de montrer, via l'hypercontractivité, que la variance de la médiane d'un échantillon gaussien est de l'ordre $1/n$ ni d'obtenir le bon ordre de grandeur pour des mesures log-concaves de potentiels $V(x) = |x|^\alpha$ lorsque $\alpha > 2$; elle ne permet pas de traiter le cas $0 < \alpha < 1$, puisque la propriété d'hypercontractivité n'est pas satisfaite.

7.8.1 Représentation de Renyi et statistiques d'ordre

Les auteurs de [37] ont combiné trois arguments différents pour borner la variance (ou obtenir des inégalités de déviations) des statistiques d'ordre d'un échantillon de variables indépendantes et de mêmes loi. Plus précisément, soient $X_1, \ldots X_n$ des variables aléatoires réelles indépendantes et de même loi. Notons les statistiques d'ordres associées par

$$X_{(1)} > \ldots > X_{(n)}.$$

Dans leur article [37], les auteurs obtiennent le résultat suivant

$$\mathrm{Var}(X_{(k)}) \leq \frac{2}{k}\mathbb{E}\left[\frac{1}{\kappa_\mu(X_{(k+1)})^2}\right], \quad k = 1, \ldots, n.$$

Leur schéma de preuve repose sur la représentation de Renyi, qui permet d'exprimer les statistiques d'ordres sous la forme de somme renormalisée de variables exponentielles standards indépendantes et de même loi. A cette représentation ils combinent l'inégalité d'Efron-Stein (cf. [38]) et l'inégalité d'association négative de Harris (afin d'utiliser cette dernière inégalité, ils doivent supposer que la fonction κ_μ est décroissante) pour majorer la variance de $X_{(k)}$, $k = 1, \ldots, n$.

Ils obtiennent également des inégalités de déviation (autour de la moyenne) à droite. Plus précisément, si $X_i = |Y_i|$ avec $\mathcal{L}(Y_i) = \mathcal{N}(0,1)$ pour tout $i = 1, \ldots n$ et $U(s) = \Phi^{-1}(1 - 1/(2s))$, avec Φ la fonction de répartition d'une variable aléatoire gaussienne standard.

$$\mathbb{P}\Big(X_{(1)} - \mathbb{E}[X_{(1)}] \leq t/(3U(n) + \sqrt{t}/U(n) + \delta_n\Big) \leq e^{-t},\ t \geq 0$$

avec $\delta_n > 0$ et $[U(n)]^3\delta_n \to \frac{\pi^2}{12}$ lorsque $n \to \infty$.

L'inconvénient de cette approche est qu'elle ne s'applique qu'aux statistiques d'ordre. Notre méthode semble plus flexible et permet de retrouver une inégalité de Poincaré pour la mesure μ^n lorsque l'application de transport T est lipschitzienne (c'est notamment le cas pour $\mu^n = \gamma_n$). Il est également clair que l'hypothèse de décroissance de la fonction κ_μ n'est pas nécessaire pour obtenir des majorations de la variance, nous avons seulement utilisé cette hypothèse pour obtenir des inégalités de déviations exponentielles. A ce propos, l'inégalité de déviation de type Bernstein de [37] est plus précise que ce que nous obtenons, mais ne permet pas de retrouver une majoration pertinente de la variance après intégration. Il est également surprenant que les auteurs de [37] ne traitent pas le cas, plus classique, de la loi gaussienne standard (sans présence de valeurs absolues).

7.8.2 Comparaison avec l'inégalité de Talagrand

Cette section propose de comparer deux inégalités fonctionnelles gaussiennes, pour $f : \mathbb{R}^n \to \mathbb{R}$ suffisament régulière :

$$\mathrm{Var}_{\gamma_n}(f) \leq C \sum_{i=1}^n \frac{\|\partial_i f\|_2^2}{1 + \log\left(\frac{\|\partial_i f\|_2}{\|\partial_i f\|_1}\right)}, \tag{7.21}$$

et

$$\mathrm{Var}_{\gamma_n}(f) \leq C \sum_{i=1}^n \int_{\mathbb{R}^n} (\partial_i f)^2(x) \left(\frac{1}{1+|x_i|}\right)^2 d\gamma_n(x) \tag{7.22}$$

Nous allons voir que ces deux inégalités ne sont pas comparables, à cet effet il suffit de traiter le cas de la dimension un. Soit $M > 0$ et définissons la fonction f_M par

$$f_M(x) = \left(\int_0^x e^{t^2/4} 1_{[-M,M]}(t)dt\right)/\|f'_M\|_1, \quad x \in \mathbb{R}.$$

On constate facilement que $\|f'_M\|_1 < \infty$ et $(f'_M)^2(x) = e^{x^2/2} 1_{[-M,M]}(x)/\|f'_M\|_1$. Ainsi, avec γ_1 la mesure gaussienne standard sur $\mathbb{R}$, nous avons

$$\int_R \frac{(f'_M)^2(x) d\gamma_1(x)}{1+x^2} = \frac{1}{\|f'_M\|_1} \int_{-M}^M \frac{dx}{1+x^2} \to \pi/\|f'_M\|_1 < \infty, \quad M \to \infty,$$

tandis que $\int_{\mathbb{R}} (f'_M)^2 d\gamma_1(x) = \int_{-M}^M dx \to \infty$ as $M \to \infty$. Ceci implique que

$$\frac{\|f'_M\|_2^2}{1 + \log\left(\frac{\|f'_M\|_2}{\|f'_M\|_1}\right)} \to \infty, \quad M \to \infty$$

Cet exemple exhibe une fonction pour laquelle (7.22) est finie alors que (7.21) ne l'est pas.

A présent, considérons la fonction f, définie par, pour tout $0 < \epsilon < 1$,

$$f_\epsilon(x) = \begin{cases} \frac{|x|}{\epsilon} + 1, & |x| \leq \epsilon \\ 0, & |x| > \epsilon, \end{cases}$$

Donc, par construction, $\|f'_\epsilon\|_1 \sim 1$ et $\|f'_\epsilon\|_2^2 \sim 1/\epsilon$. En choisissant $\epsilon = 1/2n, n \geq 1$ nous obtenons que (7.22) est d'ordre n tandis que (7.21) est d'ordre $n/\log n$. C'est pourquoi, en choisissant n suffisamment grand nous obtenons que (7.21) est strictement plus petit que (7.22).

Chapitre 8

Perspectives

Cette thèse a permis d'explorer un peu plus le phénomène de superconcentration. Toutefois les réponses que nous avons pu obtenir durant cette thèse soulèvent de nouvelles questions. Nous proposons, ci dessous, une modeste liste des pistes de recherche qui nous semblent pertinentes.

1. Est-il possible d'obtenir une condition suffisante et nécessaire sur la matrice de covariance d'un vecteur gaussien $(X_1, \ldots, X_n)$ assurant $\max_{i=1,\ldots,n} X_i$ soit superconcentré ? Est-il possible d'exhiber une classe de fonction $f : \mathbb{R}^n \to \mathbb{R}$ satisfaisant un phénomène de superconcentration ?

2. Concernant l'exemple de la marche aléatoire branchante, est-il possible de combiner des arguments de semi-groupe et de second moment pour prouver que $\mathrm{Var}(\max_{\pi \in V_n} X_\pi) = O(1)$. Il pourrait-être intéressant de voir si l'on peut adapter les arguments de l'article de Kistler-Schmidt pour y arriver. Ne vaudrait-il pas mieux commencer par traiter le modèle du GREM avant la marche aléatoire branchante ?

3. Peut-on exhiber de nouvelles fonction satisfaisant un critère de courbure dimension inverse dans un cadre gaussien. Il serait intéressant d'explorer plus encore cette idée d'interpolation entre l'infini et un point t.

4. Un travail sur l'inégalité de Talagrand sur le cube discret (aux ordres supérieurs) permettrait-il des avancées pertinentes en percolation ? Dans la même thématique, il serait pertinent de voir de quelle manière les inégalités de Talagrand d'ordre deux (pour le cube discret muni de la mesure $\mu^n = (p\delta_{-1} + (1-p)\delta_1)^{\otimes n}$) peuvent se combiner avec le lemme de Russo-Margulis (cf. [84] par exemple) pour établir des phénomènes de seuil. Il n'est pas très difficile de montrer que le lemme de Russo-Margulis sétend à l'ordre deux. Plus précisément, en dérivant une deuxième fois par rapport à p nous obtenons facilement la somme des influences d'ordre deux. En revanche, un travail supplémentaire sera nécessaire pour en déduire une phénomène de seuil dans l'esprit du résultat de Friedgut et Kalai. Il semblerait aussi pertinent de vérifier si un passage à la limite, via un argument de théorème de la limite centrale, permettrait d'obtenir une nouvelle inégalité pour la mesure gaussienne.

5. Les inégalités de Talagrand d'ordre deux ainsi que les inégalités de courbure dimension intégrées inverses semblent s'inscrire dans la continuité de certains travaux de Götze (cf.

[86, 25]) concernant la concentration de second ordre. A ce propos, la fonction d'énergie libre semble être un exemple pertinent de concentration de second ordre tandis que les U-statistiques présentent un cadre favorable à étudier pour obtenir des exemples concrets d'application des inégalités de Talagrand d'ordre deux. Au vu des travaux de Götze, il est également naturel de s'interroger sur le potentiel lien entre les inégalités de Talagrand d'ordre deux et l'obtention d'inégalités de concentration de second ordre.

6. Beaucoup de choses reste à faire concernant l'approche de la superconcentration via le transport optimal. Il faudrait étendre certains résultats obtenu durant cette thèse à des mesures non-produits afin d'être à même de traiter des modèles admettant des corrélations (suite gaussiennes stationnaire par exemple).

7. D'un point de vue géométrique, l'approche de la superconcentration par le transport optimal soulève des questions intéressante : est-il possible d'exprimer l'inégalité de Talagrand comme une inégalité de transport ? Pourrait-t-on utiliser le fait que l'inégalité de Talagrand peut s'exprimer à l'aide du norme d'Orlicz et que le carré du champ correspond à l'information de Fisher ? L'article de Bobkov [24] met en relief le fait que des résultats isopérimétriques particuliers permettent d'obtenir des inégalité de concentration fine pour le maximum. Il pourrait être utile de mieux comprendre quels types de grossissements permettent d'obtenir des inégalités de superconcentration.

8. Toujours d'un point de vue géométrique, l'étude du problème isopérimétrique uniforme (comme dans [24]) pour des lois à queues lourdes ($d\mu = Ze^{-x^\delta}dx$ avec $0 < \delta < 1$ par exemple) permettrait-il d'obtenir de nouvelles inégalités de déviation reflétant le comportement asymptotique de loi des extrêmes (autre que la loi de Gumbel) ? Sinon, est-il possible d'obtenir une méthodologie unifiée permettant de prouver des inégalités de concentration reflétant la convergence des extrèmes (pour n'importe quelle loi limite). Le cas de la loi de Gumbel est partiellement compris (mis à part, certaines difficultés pour les déviations à gauche) mais il reste beaucoup de choses à comprendre pour la loi de Weibull ou de Fréchet.
9. Il serait également intéressant de poursuivre l'étude de la concentration pour le modèle étudié par Chafaï et Péché ([47]) et ceux présentés dans [96]. Il faudrait obtenir une inégalité de déviation pour le maximum des coordonnées d'un vecteur $(X_1, \ldots, X_n)$ avec des variables aléatoires X_i, $i = 1, \ldots, n$ indépendantes de de loi $\Gamma(1, i)$. Pour cela, on pourrait surement commencer par écrire en détail le cas de la déviation à gauche du maximum de n gaussiennes standards indépendantes et de même loi à l'aide des outils isopérimétrique de Bobkov. Cet exemple, plus simple, permettrait de mieux comprendre comment traiter le maximum de loi $\Gamma(1, i)$, $i = 1, \ldots, n$. Il serait ensuite possible d'étendre tout ceci à des modèles de matrices non hermitiennes satisfaisant certaines conditions.

10. Durant cette thèse nous avons mentionné le théorème de Darling-Erdös (cf. chapitre deux) concernant la convergence du maximum de somme partielle renormalisée vers la loi de Gumbel. Nous avons obtenu une inégalité de concentration, non asymptotique, satisfaisante dans le cas gaussien (cf. chapitre quatre). Cependant, le théorème de Darling-Erdös est satisfait par une grande classe de lois de probabilité satisfaisant une condition de moment. Il sera intéressant d'étendre nos travaux dans cette direction de manière à obtenir une preuve ne reposant plus sur des arguments propres aux variables aléatoires gaussiennes.

Bibliographie

[1] *The Handbook in the Geometry of Banach Spaces.* Elsevier, 2001.

[2] J. A. Acosta. Tightness of the recentered maximum of log-correlated gaussian field. *Electronic Journal of Probability, Vol. 19, Article 90*, 2014.

[3] L. Addario-Berry, N. Broutin, L. Devroye, and G. Lugosi. On combinatorial testing problems. *Ann. Statist. 38 (2010), no. 5*, 2010.

[4] R. J. Adler. *An introduction to continuity, extrema and related topics for general Gaussian processes.* Institut of Mathematical Statistics Lectures Notes - Monograph Series, 12, 1990.

[5] S. Aida and D. Stroock. Moment estimates derived from poincaré inequalities and logarithmic sobolev inequalities. *Math. Res. Lett*, 1 :75–86, 1994.

[6] E. Aïdekon. Convergence in law of the minimum of a branching random walk. *Ann. Prob.*, 41 :1362–1426, 2013.

[7] G.W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices.* Cambridge University Press, Cambridge, 2010.

[8] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G.Scheffer. *Sur les inégalités de Sobolev logarithmiques.* Société mathématiques de France, 2000.

[9] L. P. Arguin. Extrema of log-correlated random variables : Principles and examples. *Arxiv (preprint) 1601.00582*, 2016.

[10] G. Aubrun. A sharp small deviation inequality for the largest eigenvalue of a random matrix. *In Séminaire de Probabilités XXXVIII*, 1857 :320–337, 2005.

[11] A. Auffinger and W. K. Chen. Universality of chaos and ultrametricity in mixed p-spin models. *Comm. Pure Appl. Math.*, 69(11) :2107–2130, 2016.

[12] A. Auffinger, M. Damron, and J. Hanson. 50 years of first passage percolation. *Arxiv (preprint) 1511.03262*, 2016.

[13] J.-M. Azaïs and M. Wschebor. *Level Sets and Extrema of Random Processes and Fields.* Wiley, 2009.

[14] Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices.* Springer, New York, 2010.

[15] J. Baik and J. O. Lee. Fluctuations of the free energy of the spherical sherrington-kirkpatrick model. *J. Stat. Phys.*, 165(2) :185–224, 2016.

[16] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators.* Grundlehren der Mathematischen Wissenschaften, 348, 2014.

[17] D. Bakry, M. Ledoux, and L. Saloff-Coste. *Markov semigroups at Saint-Flour.* Probability at Saint-Flour. Springer, Heidelberg, 2012.

[18] F. Barthe and C. Roberto. Modified logarithmic sobolev inequalities on ℝ. *Potential Analysis*, 2008.

[19] F. Baudoin and J. Wang. Curvature dimension inequalities and subelliptic heat kernel gradient bounds on contact manifolds. *Potential Analysis*, 40 :163–193, 2014.

[20] M. Ben-Or and N. Linial. Collective coin flipping. *Randomness and Computation Academic Press*, 70, 1990.

[21] M. Benaïm and R. Rossignol. Exponential concentration for first passage percolation through modified poincaré inequalities. *Ann. Inst. Henri Poincaré Probab. Stat. 44 (2008), no. 3*, 2008.

[22] I. Benjamini, G. Kalai, and O. Schramm. *First passage percolation has sublinear distance variance.* Sel. Works Probab. Stat. Springer, New York, 2011.

[23] J. Bertoin. Darling-erdös theorems for normalized sums of i.i.d. *Ann. Prob. 26, 832-852,*.

[24] S. Bobkov. Isoperimetric inequalities for distributions of exponential type. *The Annals of Probability, Vol. 22, No 2, 978-994*, 1994.

[25] S. Bobkov, G. Chistyakov, and F. Götze. Second order concentration on the sphere. *Arxiv (preprint) 1502.04178*, 2016.

[26] S. Bobkov and F. Götze. Exponential integrability and transportation cost related ot logarithmic sobolev inequlities. *J. Funct. Anal.*, 163(1) :1–28, 1999.

[27] S. Bobkov, F. Götze, and C. Houdré. On gaussian and bernoulli covariance representations. *Bernoulli*, 7(3) :439–451, 2001.

[28] S. Bobkov and C. Houdré. Converse poincaré-type inequalities for convex functions. *Statistics and Probability Letters*, 34(1) :37–42, 1997.

[29] S. Bobkov and C. Houdré. *Some connection between isoperimetric and Sobolev-type inequalities.* 1997.

[30] S. Bobkov and C. Houdré. A converse gaussian poincaré-type inequality for convec functions. *Statistics and Probability Letters*, 44(3) :281–290, 1999.

[31] S. Bobkov, I I. Gentil, and M. Ledoux. Hypercontractivity of hamilton-jacobi equations. *J. Math. Pures Appl.*, 80(7) :669–696, 2001.

[32] S. Bobkov and M. Ledoux. Weighted poincare-type inequalities for cauchy and other convex measures. *The Annals of Probability*, 37(2) :403–427, 2009.

[33] S. Bobkov and B. Zegarlinski. *Entropy bounds and isoperimetry.* Mem. Am. Math. Soc, 2005.

[34] E. Bolthausen, J-D. Deuschel, and G. Giacomin. Entropic repulsion and the maximum of the two-dimensional harmonic crystal. *The Annals of Probability*, 29(4) :1670–1692, 2001.

[35] E. Bolthausen, J-D. Deuschel, and O. Zeitouni. Recursions and tightness for the maximum of the discrete, two dimensional gaussian free field. *Electron. Commun. Probab.*, 16 :114–119, 2011.

[36] C. Borel. The brunn-minkoswki inequality in gauss space. *Invent. Math. 30 (2), 207-216*, 1975.

[37] S. Boucheron and M. Thomas. Concentration inequalities for order statistics. *Electronic Communications in Probability*, 2012.

[38] T. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities : a nonasymptotic theory of independance.* Oxford University Press, 2013.

[39] A. Bovier. *Gaussian Processes on Trees : frome Spin Glasses to Branching Brownian Motion*, volume 163. Cambridge sudies in advenced mathematics, 2016.

[40] A. Bovier and I. Kurkova. Derrida's generalised random energy models. i. models with finitely many hierarchies. *Ann. Inst. Henri Poincaré Prob. Stat.*, 40(4) :439–480, 2004.

[41] A. Bovier and I. Kurkova. Derrida's generalized random energy models. ii. models with continuous hierarchies. *Ann. Inst. Henri Poincaré Prob. Stat.*, 40(4) :481–495, 2004.

[42] A. Bovier, I. Kurkova, and M. Löwe. Fluctuations of the free energy in the mixed p-spin models with external field.rem and the p-spin sk models. *The Annals of Probability*, 30(2) :605–651, 2002.

[43] M. Bramson, J. Ding, and O. Zeitouni. Convergence in law of the maximum of nonlattice branching random walk. *Arxiv (preprint) 1404.3423*, 2016.

[44] M. Bramson, J. Ding, and O. Zeitouni. Convergence in law of the maximum of the two-dimensional discrete gaussian free field. *Comm. Pure Appl. Math.*, 69(1) :62–123, 2016.

[45] M. Bramson and O. Zeitouni. Tightness of the recentered maximum of the two-dimensional discrete gaussian free field. *Comm. Pure Appl. Math.*, 65(1) :1–20, 2012.

[46] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math. 44, 375-417*, 1991.

[47] D. Chafaï and S. Péché. A note on the second order universality at the edge of coulomb gases on the plane. *J. Stat. Phys.*, 2 :368–383, 2014.

[48] S. Chatterjee. *Superconcentration and related topics.* Springer, 2014.

[49] L. H. Y. Chen. An inequality for the multivariate normal distribution. *J. Multuvariate Anal. 12, 306-315*, 1982.

[50] W. K. Chen. Disorder chaos in the sherrington-kirkpatrick model with external field. *Ann. Prob.*, 41(5) :3345–3391, 2013.

[51] W. K. Chen. Chaos in the mixed even-spin models. *Comm. Math. Phys.*, 328(3) :867–901, 2014.

[52] W. K. Chen, P. Dey, and D. Panchenko. Fluctuations of the free energy in the mixed p-spin models with external field. *Probab. Theory Related Fields (to appear)*, 2017.

[53] W. K. Chen, M. Handschy, and G. Lerman. On the energy landscape of the mixed even p-spin model. *Preprint, http ://arxiv.org/abs/1609.04368*, 2016.

[54] W. K. Chen, H. W. Hsieh, C. R. Hwang, and Y. C. Sheu. Disorder chaos in the spherical mean field model. *J. Stat. Phys.*, 160(2) :417–429, 2015.

[55] W. K. Chen and D. Panchenko. An approach to chaos in some mixed p-spin models. *Probab. Theory Related Fields*, 151(1) :389–404, 2013.

[56] W. K. Chen and D. Panchenko. Temperature chaos in some spherical p-spin models. *Arxiv (preprint) 1605.01716*, 2016.

[57] W. K. Chen and A. Sen. Parisi formula, disorder chaos and fluctuation for the ground state energy in the spherical mixed p-spin models. *Comm. Math. Phys. (to appear)*, 2017.

[58] A. Chiarini, A. Cipriani, and R. S. Hazra. A note on the extremal process of the supercritical gaussian free field. *Electron. Commun. Probab.*, 20(74), 2015.

[59] A. Chiarini, A Cipriani, and R. S. Hazra. Extremes of some gaussian random interfaces. *J. Stat. Phys.*, 165(3) :521–544, 2016.

[60] A. Chiarini, A Cipriani, and R. S. Hazra. Extremes of the supercritical gaussian free field. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13(2) :711–724, 2016.

[61] D. Cordero-Erausquin. Some applications of mass transport to gaussian type inequalities. *Arch. Rational Mech. Anal. 161, 257-269*, 2002.

[62] D. Cordero-Erausquin and M. Ledoux. Hypercontractive measures, talagrand's inequality, and influences. *Geometric aspects of functional analysis, 169-189, Lectures Notes in Math 2050*, 2012.

[63] S. Dallaporta. Quelques aspects de l'étude quantitative de la fonction de comptage et des valeurs propres de matrices aléatoires. *Thèse*, 2012.

[64] S. Dallaporta. Eigenvalue variance bounds for covariance matrices. *Markov Process. Related Fields*, 21(1) :145–175, 2015.

[65] M. Damron, J. Hanson, and P. Sosoe. Subdiffusive concentration in first-passage percolation. *Electronic Journal of Probability*, 19(109), 2014.

[66] M. Damron, J. Hanson, and P. Sosoe. Sublinear variance in first-passage percolation for general distributions. *Probab. Theory Related Fields*, 163(1-2) :223–258, 2015.

[67] L. De Haan and A. Ferreira. *Extreme Value Theory.* Springer Series in Operations Research and Financial Engineering, 2006.

[68] J. Ding. Exponential and double exponential tails for maximum of two-dimensional discret gaussian free field. *Probab. Theory Related Fields 157, no. 1-2*, 157(1-2) :285–299, 2013.

[69] J. Ding, R. Eldan, and A. Zhai. On multiple peaks and moderate deviations for supremum of gaussian field. *Ann. Prob.*, 43(6) :3468–3493, 2015.

[70] J. Ding, R. Roy, and O. Zeitouni. Convergence of the centered maximum of log-correlated gaussian fields. *Preprint, http ://arxiv.org/abs/1503.04588*, 2015.

[71] J. Ding, R. Roy, and O. Zeitouni. Convergence of the centered maximum of log-correlated gaussian fields. *Arxiv (preprint) 1503.04588*, 2016.

[72] J. Ding and O. Zeitouni. Extreme values for two-dimensional discrete gaussian free field. *The Annals of Probability*, 42(4) :1480–1515, 2014.

[73] B. Duplantier, R. Rhodes, S. Sheffield, and V. Vargas. Critical gaussian multiplicative chaos : Convergence of the derivative martingale. *Ann. Prob. 42, no. 5,*, 2014.

[74] B. Duplantier, R. Rhodes, S. Sheffield, and V. Vargas. Log-correlated gaussian fields : an overview. *Preprint, http ://arxiv.org/abs/1407.5605*, 2014.

[75] A. Dvoretzky. Some results on convex bodies and banach spaces. *Proc. Symp. on Linear spaces*, pages 123–160, 1961.

[76] A. Ehrhard. Symmétrisation dans l'espace de gauss. *Math. Scand.*, 53 :281–381, 1983.

[77] U. Einmahl. The darling-erdös theorem for sums of i.i.d. random variables. *Probab. Theory Related Fields 82 214-257*, 1989.

[78] P. Embrechts, C. Klüpperlberg, and T. Mikosch. *Modelling extremal events.* Springer-Verlag, Berlin, 1997.

[79] P. Erdös and D. A. Darling. A limit theorem for the maximum of normalized sums of independant random variables. *Duke Math. Journal*, 1956.

[80] D. Falik and A. Samorodnitsky. Edge-isoperimetric inequalities and influences. *Arxiv (preprint) 0512636*, 2005.

[81] O. Feldheim and S. Sodin. A universality result for the smallest eigenvalues of certain sample covariance matrix. *Geom. Funct. Anal.*, 20(1) :88–123, 2010.

[82] X. Fernique. Intégrabilités des vecteurs gaussiens. *C. R. Acad. Sci. Paris 270, 1698-1699*, 1970.

[83] X. Fernique. Fonctions aléatoires gaussiennes, les résultats récents de m. talagrand. *Seminaire Bourbaki, exp 660. Astérisque 145-146, 177-186*, 1987.

[84] E. Friedgut and G. Kalai. Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.*, 124(10), 1996.

[85] C. Garban and J. E. Steif. *Noise sensitivity of Boolean functions and percolation.* Institute of Mathematical Statistics Textbooks. Cambridge University Press, New York, 2015.

[86] F. Götze and H. Sambale. Second order concentration via logarithmic sobolev inequalities. *Arxiv (preprint) 1605.08635*, 2016.

[87] N. Gozlan. Poincaré inequalities and dimension free concentration of measure. *Ann. Inst. Henri Poincaré Prob. Stat. 46 (2010), no. 3*, 2010.

[88] N. Gozlan and C. Léonard. Transport inequalities. a survey. *Markov Process. Related Fields 16, no. 4, 635-736*, 2010.

[89] B. T. Graham. Sublinear variance for directed last-passage percolation. *J. Theoretical Prob.*, 25(3) :687–702, 2012.

[90] M. Gromov and V. D. Milman. A topological application of the isoperimetric inequality. *Amer. J. Math.*, 105 :843–854, 1983.

[91] L. Gross. Logarithmic sobolev inequalities. *Amer. J. Math.*, 1975.

[92] C. Houdré. Some applications of covariance identities and inequalities to functions of multivariate normal variables. *J. Amer. Statist. Assoc. 90, no. 431, 965-968*, 1995.

[93] C. Houdré and A. Kagan. Variance inequalities for functions of gaussian variables. *J. Theoretical Prob. Vol. 8, 23-30*, 1995.

[94] C. Houdré and V. Perez-Abreu. Covariance identities and inequalities for functionals on wiener and poisson spaces. *Ann. Prob.*, 1992.

[95] C. Houdré, V. Perez-Abreu, and D. Surgailis. Interpolation, correlation identities, and inequalities for infinitely divisible variables. *Journal of Fourier Analysis and Applications*, 4(6) :651–668, 1998.

[96] Tiefeng Jiang and Yongcheng Qi. Spectral radii of large non-hermitian random matrices. *J. Theoret. Probab.*, 30(1) :326–364, 2017.

[97] K. Johansson. On fluctuations of eigenvalues of random hermitian matrices. *Duke Math. J.*, 91(1) :151–204, 1998.

[98] K. Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.*, 209 :437–476, 2000.

[99] I. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.*, 29(2) :295–327, 2001.

[100] Z Kabluchko. Extremes value analysis of standardized gaussian increments. *Arxiv (preprint) 0706.1849v3*, 2008.

[101] J. Kahn, G. Kalai, and N. Linial. The influence of variables on boolean functions. *Proc. 29th Ann. Symp. on Fondations of Comp. Sci., Computer Society Press*, 62 :68–80, 1988.

[102] R. Kannan, L. Lovász, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom*, 13(3-4) :541–559, 1995.

[103] O. Khorunzhiy. High moments of large wigner random matrices and asymptotic properties of the spectral norm. *Random Oper. Stoch. Equ.*, 20(1) :25–68, 2012.

[104] N. Kistler. *Derrida's random energy models. From spin glasses to the extremes of correlated random fields*, volume 2143 of *Lecture Notes in Math.* Springer, Cham, 2015.

[105] N. Kistler and M. Schmidt. From derrida's random energy model to branching random walks : from 1 to 3. *Electronic Communications in Probability*, 20(47), 2015.

[106] T. Kumagaï. Fluctuations of maxima of discrete gaussian free fields on a class of recurrent graphs. *Electron. Commun. Probab.*, 18(75), 2013.

[107] S. Kwapien. *A remark on the median and the expectation of convex functions of Gaussian vectors*, volume 33 of *Progr. Probab.* Birkhäuser Boston, Boston, MA, 1994.

[108] M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes.* Springer Series in Statistics., 1983.

[109] M. Ledoux. Semigroup proofs of the isoperimetric inequality in euclidean and gauss space. *Bull. Sci.math*, 1992.

[110] M. Ledoux. L'algèbre de lie des gradients itérés d'un générateur markovien - développements de moyenne et entropies. *Ann. Sci. Ec. Norm. Super 28(4), 435-460*, 1995.

[111] M. Ledoux. The geometry of markov diffusions operators. *Ann. Fac. Sci. Toulouse Math. (6) 9, no.2, 305-366*, 2000.

[112] M. Ledoux. *The concentration of measure phenomenon.* Mathematical Surveys and Monographs, 89, 2001.

[113] M. Ledoux. Measure concentration, transportation cost and functional inequalites. 2002.

[114] M. Ledoux. Deviation inequalities on largest eigenvalue. *Geometric Aspects of Functional Analysis, Israel Seminar 2004-2005. Lecture Notes in Math. 1910, 167-219.*, 2007.

[115] M Ledoux and D Cordero-Erausquin. Hypercontractive measures, talagrand's inequality, and influences. *Geometric aspects of functional analysis, 169-189, Lectures Notes in Math 2050*, 2012.

[116] M. Ledoux and B. Rider. Small deviations for beta ensemble. *Electronic Journal of Probability, Vol. 15, Article 41*, 2010.

[117] M. Ledoux and M. Talagrand. *Probability in Banach spaces. Isoperimetry and processes.* Springer-Verlage, Berlin, 2011.

[118] J. Lee and J. Yin. A necessary and sufficient condition for edge universality of wigner matrice. *Duke Math. J.*, 163(1) :117–173, 2014.

[119] P. Levy. Problème concerts d'analyse fonctionelle. *Gauthier-Villars*, 1951.

[120] L. Lusternik. Die brunn-minkowskische ungleichung für beliebige messbare mengen. *C. R. Acad. Sci. URSS, 8 p. 55-58 5*, 1935.

[121] A. Lytova and K. Tikhomirov. The variance of the l_p^n-norm of the gaussian vector and dvoretzky's theorem. *Arxiv (preprint) 1705.05052*, 2017.

[122] T. Madaule. Maximum of a log-correlated gaussian field. *Ann. Inst. Henri Poincaré Prob. Stat.*, 51(4) :1369–1431, 2015.

[123] K. Marton. A measure concentration inequality for contracting markov chains. *Geom. Funct. Anal. 6, 556-571*, 1997.

[124] K. Marton. A measure concentration for a class of random processes. *Probab. Theory Related Fields 110, 427-439*, 1998.

[125] R. J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. Journal 80, 309-323*, 1995.

[126] V. Milman. New proof of the theorem of dvoretzky on sections of convex bodies. *Funct. Anal. Appl. 5, 28-37*, 1971.

[127] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math. 80, no4, 931-954*, 1958.

[128] Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, New York, 2014.

[129] D. Panchenko. *The Sherrington-Kirkpatrick model*. Springer Monographs in Mathematics., 2013.

[130] G. Paouris and P. Valettas. A small deviation inequality for convex functions. *Arxiv (preprint) 1611.01723v1*, 2016.

[131] G. Paouris, P. Valettas, and J. Zinn. Random version of dvoretzky's theorem in l_p^n. 2015.

[132] S. Péché. Universality results for the largest eigenvalues of some sample covariance matrix ensembles. *Probab. Theory Related Fields*, 143(3-4) :481–516, 2009.

[133] S. Péché and A. Soshnikov. Wigner random matrices with non-symmetrically distributed entries. *J. Stat. Phys.*, 129(5-6), 2007.

[134] M. S. Pinsker. Information and information stability of random variables and processes. *Holden-Day, San Francisco*, 1964.

[135] B. Rider. Order statistics and ginibre's ensembles. *J. Stat. Phys.*, 114(3-4) :1139–1148, 2004.

[136] A. Ruzmaikina. Universality of the edge distribution of eigenvalues of wigner random matrices with polynomially decaying distributions of entries. *Comm. Math. Phys.,*, 261(2) :277–296, 2006.

[137] L. Saloff-Coste. Random walks on finite groups. *Springer*, 2004.

[138] L. Saloff-Coste and Hebisch W. Gaussian estimates for markov chains and random walks on groups. *The Annals of Propability*, 1993.

[139] G. Schechtmann. The random version of dvoretsky's theorem in l_∞^n. *GAFA Seminar 2004-2005*, 1910 :265–270, 2007.

[140] E. Schmidt. Die brunn-minkowskische ungleichung und ihr spiegelbild sowie die isoperimetrische eigenschaft der kugel in der euklidischen und nichteulkidischen geometrie. *Math. Nach. 1, 81-157*, 1948.

[141] M. Schmuckenschläger. Martingales, poincaré type inequalities and deviations inequalities. *J. Funct. Anal.*, 155 :303–323, 1998.

[142] A. Soshnikov. Universality at the edge of the spectrum in wigner random matrices. *Comm. Math. Phys.,*, 207(3) :697–733, 1999.

[143] E. Subag and O. Zeitouni. The extremal process of critical points of the pure p-spin spherical spin glass model. *Arxiv (preprint) 1509.03098*, 2016.

[144] V. N. Sudakov and B. S. Tsirel'son. Extremal properties of half-space for spherically invariant measures. *J. Soviet. Math. 9, 9-18*, 1978.

[145] M. Talagrand. A new isoperimetric inequality for product measure, and the concentration of measure phenomenon. *Geometric aspects of functional analysis, 91-124, Lectures Notes in Math 1469*, 1991.

[146] M. Talagrand. A new isoperimetric inequality for product measure, and the tails of sums of independant random variables. *Geom. Funct. Anal. 1, 211-223*, 1991.

[147] M. Talagrand. On russo's approximate zero-one law. *Ann. Prob. 22, 1576-1587,*, 1994.

[148] M. Talagrand. Sharper bounds for gaussian and empirical processes. *Ann. Prob. 22, 28-76*, 1994.

[149] M Talagrand. The supremum of some canonical processes. *Amer. J. Math.*, 1994.

[150] M. Talagrand. Transportation cost for gaussian and other product measures. *Geom. Funct. Anal. 6, 587-600*, 1996.

[151] M. Talagrand. An isoperimetric theorem on the cube and the khintchine-kahane inequalities. *Proc. Amer. Math. Soc. 104, 905-909*, 1998.

[152] M. Talagrand. *Spin glasses : a challenge for mathematicians*, volume 46 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, 2003.

[153] M. Talagrand. *The generic chaining. Upper and lower bounds of stochastic processes.* Springer Monographs in Mathematics., 2005.

[154] M. Talagrand. *Mean field models for spin glasses. Volume I.* Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]. Springer-Verlag, Berlin, 2011.

[155] M. Talagrand. *Mean field models for spin glasses. Volume II : Advanced replica-symmetry and low temperature*, volume 55 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer, Heidelberg, 2011.

[156] M. Talagrand. *Upper and lower bounds for stochastic processes. Modern methods and classical problems.* Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, 60, 2014.

[157] K. Tanguy. Some superconcentration inequalities for extrema of stationary gaussian processes. *Statistics and Probability Letters*, 2015.

[158] T. Tao and V. Vu. Random matrices : sharp concentration of eigenvalues. *Random Matrices Theory Appl.*, 2(3), 2013.

[159] K. Tikhomirov. Superconcentration, and randomized dvoretzky's theorem for spaces with 1-unconditional bases. *Arxiv (preprint)*, 2017.

[160] A. W. Van der Vaart. *Asymptotic Statistics.* Cambrige University Press, Cambridge, 1998.

[161] R. Vershynin. *Introduction to the non-asymptotic analysis of random matrices.* Cambridge Univ. Press, Cambridge, 2012.

[162] C. Villani. *Topics in Optimal Transportation.* AMS, 2003.

[163] C. Villani. *Optimal Transport.* Springer, 2009.

[164] K. Yosida. *Functional Analysis.* Springer, 190.

[165] O. Zeitouni. Branching random walks and gaussian fields. *Notes for Lectures*, 2012.

[166] O. Zeitouni and J. Ding. Extremes values for two-dimensional discret gaussian free field. *Ann. Probab. 42, no. 4*, 2014.

Statistics and Probability Letters 106 (2015) 239–246

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Some superconcentration inequalities for extrema of stationary Gaussian processes

Kevin Tanguy *

University of Toulouse, France
Institute of Mathematics of Toulouse (CNRS UMR 5219), Université Paul Sabatier, 31062 Toulouse, France

ARTICLE INFO

Article history:
Received 23 March 2015
Received in revised form 20 July 2015
Accepted 21 July 2015
Available online 29 July 2015

Keywords:
Superconcentration
Concentration inequalities
Stationary Gaussian process
Hypercontractivity
Theory of extremes

ABSTRACT

This note is concerned with concentration inequalities for extrema of stationary Gaussian processes. It provides non-asymptotic tail inequalities which fully reflect the fluctuation rate, and as such improve upon standard Gaussian concentration. The arguments rely on the hypercontractive approach developed by Chatterjee for superconcentration variance bounds. Some statistical illustrations complete the exposition.

1. Introduction

1.1. Convergence of extremes

As an introduction, we recall some classical facts about weak convergence of extrema of stationary Gaussian sequences and processes.

1.1.1. Stationary Gaussian sequences

Let $(X_i)_{i\geq 0}$ be a centered stationary Gaussian sequence such that $\mathbb{E}\left[X_i^2\right] = 1$, $i \geq 0$, with covariance $\mathrm{Cov}(X_i, X_j) = \phi(|i-j|)$, $i, j \geq 0$, where $\phi : \mathbb{N} \to \mathbb{R}$. An extensive study has been developed in this setting towards the asymptotic behavior of the maximum $M_n = \max_{i=1,\ldots,n} X_i$ (Berman, Mittal, Pickands, etc. cf. e.g. Leadbetter et al., 1983). A sample result is the following theorem (see Leadbetter et al., 1983).

Theorem 1. *Let $(X_i)_{i\geq 0}$ be a stationary Gaussian sequence with covariance function ϕ such that $\phi(n)\log n \to 0$ as $n \to \infty$. Then*

$$a_n(M_n - b_n) \underset{n\to\infty}{\longrightarrow} G$$

in distribution where $a_n = (2\log n)^{1/2}$,

$$b_n = (2\log n)^{1/2} - \frac{1}{2}(2\log n)^{-1/2}(\log\log n + \log 4\pi)$$

and the random variable G has a Gumbel distribution (for all $t \in \mathbb{R}$, $\mathbb{P}(G \leq t) = e^{-e^{-t}}$).

* Correspondence to: Institute of Mathematics of Toulouse (CNRS UMR 5219), Université Paul Sabatier, 31062 Toulouse, France.
E-mail address: ktanguy@math.univ-toulouse.fr.

http://dx.doi.org/10.1016/j.spl.2015.07.028
0167-7152/

1.1.2. Stationary Gaussian processes

Let $(X_t)_{t\geq 0}$ be a centered stationary Gaussian process such that $\mathbb{E}\left[X_t^2\right] = 1,\ t \geq 0$, with covariance function $\mathrm{Cov}(X_s, X_t) = \phi\big(|t-s|\big)$, for $t, s \geq 0$, where $\phi : [0, +\infty) \to \mathbb{R}$. Consider two behaviors, as $t \to 0$, of the covariance function ϕ:

$$\phi(t) = 1 - \frac{\lambda_2 t^2}{2} + o(t^2) \tag{1.1}$$

$$\phi(t) = 1 - C|t|^\alpha + o\big(|t|^\alpha\big). \tag{1.2}$$

The first case ensures that X_t is differentiable and $\lambda_2 = -\phi''(0)$ is a spectral moment, whereas the second case concerns non-differentiable (but continuous) processes (such as the Ornstein–Uhlenbeck process). For more details about this topic, see Leadbetter et al. (1983). For any $T > 0$, set $M_T = \sup_{t\in[0,T]} X_t$.

Theorem 2. *Let $(X_t)_{t\geq 0}$ be a stationary Gaussian process such that $\phi(t)\log t \to 0$ as $t \to \infty$. Then*

$$a_T(M_T - b_T) \underset{T\to\infty}{\longrightarrow} G$$

in distribution where $a_T = (2\log T)^{1/2}$, b_T depends on the hypothesis (1.1) or (1.2) and G has a Gumbel distribution.

The aim of this note is to quantify the preceding asymptotic statements into sharp concentration inequalities fully reflecting the fluctuation rate of the maximum (respectively the supremum). Such variance bounds with the correct scale were first obtained in this context by Chatterjee in Chatterjee (2014) (Section 9.6), as part of the superconcentration phenomenon. The results presented here strengthen the variance bounds into exponential tail inequalities. One main result is the following statement.

Theorem 3. *Let $(X_i)_{i\geq 0}$ be a centered stationary Gaussian sequence with covariance function ϕ. Assume that ϕ is non-increasing and satisfies $\phi(1) < 1$. Then, there exist $\alpha = \alpha(\phi) \in (0,1)$ and $c = c(\phi, \alpha) > 0$ such that for all $n \geq 2$,*

$$\mathbb{P}\left(|M_n - \mathbb{E}[M_n]| > t\right) \leq 6e^{-ct/\sqrt{\max(\phi(n^\alpha),1/\log n)}}, \quad t \geq 0. \tag{1.3}$$

Remark. Under the hypothesis of Theorem 1, $\max\left(\phi(n^\alpha), 1/\log n\right) = 1/\log n$ for n large enough, which is exactly the fluctuation rate. Observe furthermore that integrating (1.3) recovers the variance bounds of Chatterjee (2014). In particular, in the framework of Theorem 1, this variance is of order $\frac{1}{\log n}$ which is optimal (see Corollary 1.9 in Ding et al., 2013). One last observation, under the hypothesis of Theorem 1, $\phi(n) \to 0$ as $n \to \infty$. This implies that $\sup_{n\geq 1}|\phi(n)| < 1$ (according to Leadbetter et al., 1983 p. 86), so in that particular case the hypothesis $\phi(1) < 1$ can be removed.

It is important to compare Theorem 3 to classical Gaussian concentration (see e.g. Ledoux, 2001) which typically produces

$$\mathbb{P}\left(|M_n - \mathbb{E}\,[M_n]\,| \geq t\right) \leq 2e^{-t^2/2}, \quad t \geq 0. \tag{1.4}$$

While of Gaussian tail, such bounds do not reflect the fluctuations of the extremum M_n of Theorem 1. Moreover, with respect to this Gaussian bound, Theorem 3 actually provides the correct tail behavior of the maximum M_n in the form of a superconcentration inequality in accordance with the fluctuation result and the limiting Gumbel distribution (since $\mathbb{P}(G > t) \sim e^{-t}$ as $t \to \infty$). On the other hand, the method emphasized here does not produce any lower bound. Let us also emphasizes that Theorem 3 covers the classical independent case, when all the X_i's are independent, by taking $\phi = 0$ and provides the following concentration inequality which will be useful in the last section,

$$\mathbb{P}\left(|M_n - \mathbb{E}\,[M_n]| \geq t\right) \leq 6e^{-ct\sqrt{\log n}}, \quad t \geq 0. \tag{1.5}$$

Observe in addition that Theorem 3 expresses a concentration property of the maximum around its mean whereas, in the regime of the convergence of extremes of Theorem 1, the centerings are produced by explicit values b_n. Actually, up to numerical constants, the same inequalities hold true around the constant b_n instead of the mean. To this task, it is enough to prove that $\sup_n \mathbb{E}\,[|a_n(M_n - b_n)|] < \infty$. Set $Z_n = a_n(M_n - b_n)$. Letting M_n' be an independent copy of M_n, set similarly $Z_n' = a_n(M_n' - b_n)$. Now, integrating (1.3), $\sup_n \mathbb{E}\,[a_n|M_n - \mathbb{E}[M_n]|] < \infty$. Hence $\sup_n \mathbb{E}\left[|Z_n - Z_n'|\right] < \infty$ from which it easily follows that $\sup_n \mathbb{E}\,[|Z_n|] < \infty$.

The next statement is the analogue of Theorem 3 for stationary Gaussian processes, suitably quantifying Theorem 2. It is presented in the more general context of a centered stationary Gaussian field $(X_t)_{t\in\mathbb{R}^d}$ such that $\mathbb{E}\left[X_t^2\right] = 1$, $t \in \mathbb{R}^d$, with covariance function ϕ. According to Chatterjee (2014), very little seems actually to be known on the asymptotic fluctuations of the supremum of stationary Gaussian processes indexed by $\mathbb{R}^d$ when the dimension d is greater than two. Some recent specific results are available for Gaussian fields with logarithmic correlation (see e.g. Acosta, 2014; Duplantier et al., 2014; Madaule, 2014; Ding et al., 2015; Duplantier et al., 2014, for an overview on the subject). However, extending similarly the variance bounds of Chatterjee (2014), we obtain a concentration inequality for the supremum of a Gaussian field over a subset A of $\mathbb{R}^d$. The case $d = 1$ thus covers Theorem 2.

Theorem 4. *Let $(X_t)_{t\in\mathbb{R}^d}$ be a stationary Gaussian Euclidean field with covariance function ϕ. Assume that $t \mapsto \phi(t)$ is non-increasing and $\phi(1) < 1$. If A is a subset of $\mathbb{R}^d$, denote by $N(A) = N(A, 1)$ the minimal number of balls with radius 1 needed to cover A. Set $M(A) = \sup_{s\in A} X_s$. Then, there exist $C = C(\phi, d) > 0$ and $c = c(\phi, d) > 0$ only depending on ϕ and d such that, for all $A \subset \mathbb{R}^d$ with $N(A) > 1$,*

$$\mathbb{P}\left(|M_A - \mathbb{E}[M_A]| \geq t\right) \leq 6e^{-ct/\sqrt{K_{N(A)}}}, \quad t \geq 0,$$

where

$$K_{N(A)} = \max\left(\phi\left(N(A)^C\right), 1/\log N(A)\right).$$

Remark. When $d = 1$ and $A = [0, T]$ for some $T > 0$, $N(A) = \frac{\lfloor T \rfloor}{2} + 1$ and under the hypotheses of Theorem 2, $K_{N(A)} = 1/\log(\frac{\lfloor T \rfloor}{2} + 1)$ for T large enough.

The proofs of Theorems 3 and 4 are based on the hypercontractive approach developed by Chatterjee (2014) towards variance bounds. The task will be to adapt the argument to reach exponential tail inequalities at the correct fluctuation regime. This will be achieved via the corresponding variance bound at the level of Laplace transforms.

The paper is organized as follows. In the next section, we present the semigroup tools used to prove the main results. The third section is devoted to the proof of Theorems 3 and 4. In Section 4, we present some illustrations to further Gaussian models. In the final section, we present an illustration in statistical testing.

2. Framework and tools

2.1. Notation

For any $n \geq 1$, denote by $X = (X_1, \ldots, X_n)$ a centered Gaussian vector, with maximum $M_n = \max_{i=1,\ldots,n} X_i$. Set $I = \operatorname{argmax}_{i=1,\ldots,n} X_i$.

The main technical part of this work is provided by the following extension of Chatterjee's approach (Chatterjee, 2014) adapted to exponential concentration bounds. The very basis of the argument is the hypercontractivity of the Ornstein–Uhlenbeck semigroup together with the exponential version of Poincaré inequality (cf. Bakry et al., 2012, Bakry et al., 2014, Chatterjee, 2014).

Theorem 5. *Let $X = (X_1, \ldots, X_n)$ be a centered Gaussian vector with covariance matrix Γ. Assume that for some $r_0 \geq 0$, there exists a non trivial covering $\mathcal{C}(r_0)$ of $\{1, \ldots, n\}$ verifying the following properties:*

- *for all $i, j \in \{1, \ldots, n\}$ such that $\Gamma_{ij} \geq r_0$, there exists $D \in \mathcal{C}(r_0)$ such that $i, j \in D$;*
- *there exists $C \geq 1$ such that, a.s., $\sum_{D\in\mathcal{C}(r_0)} 1_{\{I\in D\}} \leq C$.*

Let $\rho(r_0) = \max_{D\in\mathcal{C}(r_0)} \mathbb{P}(I \in D)$.

Then, for every $\theta \in \mathbb{R}$,

$$\operatorname{Var}\left(e^{\theta M_n/2}\right) \leq C\,\frac{\theta^2}{4}\left(r_0 + \frac{1}{\log(1/\rho(r_0))}\right)\mathbb{E}\left[e^{\theta M_n}\right]. \tag{2.1}$$

In particular,

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| \geq t) \leq 6e^{-ct/\sqrt{K_{r_0}}}, \quad t \geq 0,$$

where $K_{r_0} = \max\left(r_0, \frac{1}{\log(1/\rho(r_0))}\right)$ and $c > 0$ is a universal constant.

Remark. By monotone convergence, we can obtain the same result for a supremum instead of a maximum. This fact will be useful to obtain an application of Theorem 5 for a Gaussian Euclidean field.

Proof. As announced, the scheme of proof follows Chatterjee (2014). The starting point is the representation formula

$$\operatorname{Var}(f) = \int_0^\infty e^{-t}\,\mathbb{E}\left[\nabla f \cdot P_t \nabla f\right] dt \tag{2.2}$$

for the variance of a (smooth) function $f : \mathbb{R}^n \to \mathbb{R}$ along the Ornstein–Uhlenbeck semigroup $(P_t)_{t\geq 0}$. General references on the Ornstein–Uhlenbeck semigroup and more general Markov semigroups, as well as such semigroup interpolation formulas are Bakry et al. (2012) and Bakry et al. (2014).

Following Chatterjee (2014), given a centered Gaussian vector X with covariance matrix $\Gamma = (\Gamma_{ij})_{1\le i,j\le n}$, apply (2.2) to (a smooth approximation of) $f = e^{\theta M/2}$ where $\theta \in \mathbb{R}$, $M(x) = \max_{i=1,\dots,n}(Bx)_i$ and $\Gamma = B\,{}^tB$. It yields that, for all $\theta \in \mathbb{R}$,

$$\mathrm{Var}\left(e^{\theta M_n/2}\right) = \frac{\theta^2}{4}\int_0^\infty e^{-t}\,\mathbb{E}\Bigg[\sum_{i,j=1}^n \Gamma_{ij}\,1_{\{X\in A_i\}}e^{\theta M_n/2}1_{\{X^t\in A_j\}}e^{\theta M_n^t/2}\Bigg]dt.$$

Here $\{X \in A_i\} = \{I = i\}$, for $i = 1,\dots,n$, and $X^t = e^{-t}X + \sqrt{1-e^{-2t}}\,Y$, where Y is an independent copy of X, with corresponding maximum M_n^t. This process shares the same hypercontractivity property as the Ornstein–Uhlenbeck process (cf. Chatterjee, 2014). Indeed, assuming without loss of generality that B is invertible (otherwise restrict to a subspace of $\mathbb{R}^n$), define the semigroup $(Q_t)_{t\ge0}$ by

$$Q_tf(x) = \mathbb{E}\left[f\left(e^{-t}x + \sqrt{1-e^{-2t}}Y\right)\right],\quad x\in\mathbb{R}^n$$

for $f : \mathbb{R}^n \to \mathbb{R}$ smooth enough. Let Z be a standard Gaussian vector on $\mathbb{R}^n$ so that $Y = BZ$ in law. Thus, by defining $g : \mathbb{R}^n \to \mathbb{R}$ as $g(x) = f(Bx)$, we have, for all $x \in \mathbb{R}^n$,

$$\begin{aligned}Q_tf(x) &= \mathbb{E}\left[f\left(e^{-t}x + \sqrt{1-e^{-2t}}Y\right)\right]\\ &= \mathbb{E}\left[f\left(e^{-t}BB^{-1}x + \sqrt{1-e^{-2t}}BZ\right)\right]\\ &= \mathbb{E}\left[g\left(e^{-t}B^{-1}x + \sqrt{1-e^{-2t}}Z\right)\right] = P_tg(B^{-1}x).\end{aligned}$$

Therefore, by the hypercontractive property of $(P_t)_{t\ge0}$ with the same exposants p and q as the Ornstein–Uhlenbeck semigroup,

$$\mathbb{E}\left[|Q_tf(X)|^q\right]^{1/q} = \|P_tg\|_q \le \|g\|_p = \mathbb{E}\left[|f(X)|^p\right]^{1/p}, \tag{2.3}$$

where $\|\cdot\|_r$, for any $r \ge 1$, stands for the L^r-norm with respect to the standard Gaussian measure on $\mathbb{R}^n$.

Now, for every $t \ge 0$, with $I^t = \mathrm{argmax}_{i=1,\dots,n}X_i^t$,

$$\begin{aligned}\mathcal{I} &= \mathbb{E}\Bigg[\sum_{i,j=1}^n \Gamma_{ij}\,1_{\{X\in A_i\}}e^{\theta M_n/2}1_{\{X^t\in A_j\}}e^{\theta M_n^t/2}\Bigg]\\ &= \mathbb{E}\Bigg[e^{\theta(M_n+M_n^t)/2}\sum_{i,j=1}^n \Gamma_{ij}\,1_{\{I=i\}}1_{\{I^t=j\}}\Bigg]\\ &= \mathbb{E}\left[e^{\theta\left(M_n+M_n^t\right)/2}\,\Gamma_{II^t}\right]\\ &= \sum_{k\ge0}\mathbb{E}\left[e^{\theta\left(M_n+M_n^t\right)/2}\,\Gamma_{II^t}1_{\{2^{-k-1}\le\Gamma_{II^t}\le2^{-k}\}}\right].\end{aligned}$$

To ease the notation, denote Γ_{II^t} by Γ and set $F^t = M_n + M_n^t$. Let $k_0 = \min\{k \ge 0\,;\, r_0 \le 2^{-k-1}\}$. Cutting the preceding sum into two parts, we obtain

$$\begin{aligned}\mathcal{I} &\le \sum_{k=0}^{k_0}2^{-k}\,\mathbb{E}\left[e^{\theta F^t/2}1_{\{\Gamma\ge r_0\}}\right] + \sum_{k=k_0+1}^{\infty}2^{-k}\,\mathbb{E}\left[e^{\theta F^t/2}1_{\{2^{-k-1}\le\Gamma\le2^{-k}\}}\right]\\ &\le 2\sum_{D\in\mathcal{C}(r_0)}\mathbb{E}\left[e^{\theta F^t/2}1_{\{I,I^t\in D\}}\right] + \sum_{k\ge0}r_0\,\mathbb{E}\left[e^{\theta F^t/2}1_{\{2^{-k-1}\le\Gamma\le2^{-k}\}}\right]\\ &= 2\sum_{D\in\mathcal{C}(r_0)}\mathbb{E}\left[e^{\theta F^t/2}1_{\{I,I^t\in D\}}\right] + r_0\,\mathbb{E}\left[e^{\theta F^t/2}\right].\end{aligned}$$

By the Cauchy–Schwarz inequality and Gaussian rotational invariance,

$$\mathbb{E}\left[e^{\theta F^t/2}\right] \le \mathbb{E}\left[e^{\theta M_n}\right].$$

On the other hand, by Hölder's inequality with $\frac{1}{p} + \frac{1}{q} = 1$,

$$\begin{aligned}\mathbb{E}\left[e^{\theta F^t/2}1_{\{I,I^t\in D\}}\right] &= \mathbb{E}\left[e^{\theta M_n/2}1_{\{I\in D\}}\,e^{\theta M_n^t/2}1_{\{I^t\in D\}}\right]\\ &\le \mathbb{E}\left[e^{\theta pM_n/2}1_{\{I\in D\}}\right]^{1/p}\,\mathbb{E}\left[e^{\theta qM_n^t/2}1_{\{I^t\in D\}}\right]^{1/q}.\end{aligned}$$

Next, by the hypercontractivity property (2.3),

$$\mathbb{E}\left[e^{\theta q M_n^t/2}1_{\{I^t\in D\}}\right]^{1/q} \le \mathbb{E}\left[e^{\theta p M_n/2}1_{\{I\in D\}}\right]^{1/p}$$

provided that $e^{2t} = \frac{q-1}{p-1}$, that is $p = 1+e^{-t} < 2$. As a consequence,

$$\mathbb{E}\left[e^{\theta F^t/2}1_{\{I,I^t\in D\}}\right] \le \mathbb{E}\left[e^{\theta p M_n/2}1_{\{I\in D\}}\right]^{2/p}$$

and by a further use of Hölder's inequality on the right hand side of the preceding inequality,

$$\mathbb{E}\left[e^{\theta F^t/2}1_{\{I,I^t\in D\}}\right] \le \mathbb{P}(I\in D)^{\frac{2-p}{p}}\,\mathbb{E}\left[e^{\theta M_n}1_{\{I\in D\}}\right].$$

Combining the preceding with the assumptions, we get that

$$\mathcal{I} \le \left(r_0 + C\rho(r_0)^{\frac{2-p}{p}}\right)\mathbb{E}\left[e^{\theta M_n}\right].$$

Finally,

$$\mathrm{Var}\left(e^{\theta M_n/2}\right) \le C\,\frac{\theta^2}{4}\left(r_0 + \int_0^\infty e^{-t}\rho(r_0)^{\tanh(t/2)}dt\right)\mathbb{E}\left[e^{\theta M_n}\right].$$

The announced inequality (2.1) then follows since

$$\int_0^\infty e^{-t}\,(\rho(r_0))^{\tanh(t/2)}\,dt \le \frac{1}{\log\left(1/\rho(r_0)\right)}.$$

To end the proof of Theorem 5 and produce exponential tail bounds, we combine (2.1) with the following lemma (see Ledoux, 2001 p. 50).

Lemma 6. *Let Z be a random variable and $K > 0$. Assume that for every $|\theta| \le 2/\sqrt{K}$,*

$$\mathrm{Var}\left(e^{\theta Z/2}\right) \le K\,\frac{\theta^2}{4}\,\mathbb{E}\left[e^{\theta Z}\right]. \tag{2.4}$$

Then,

$$\mathbb{P}\left(|Z-\mathbb{E}[Z]| > t\right) \le 6e^{-ct/\sqrt{K}},$$

for all $t \ge 0$, with $c > 0$ a numerical constant.

Remark. Such a lemma has been used in prior works Benaïm and Rossignol (2008) and Damron et al. (2014) in order to obtain exponential concentration tail bounds for first passage percolation models. □

3. Proofs of the main results

3.1. Proof of Theorem 3

The task is to show that the hypotheses of Theorem 5 are satisfied. For $0 < \alpha < 1$, choose $r_0 = \phi(\lfloor n^\alpha\rfloor)$ and $\mathcal{C}(r_0) = \{D_1, D_2, \ldots\}$ where $D_1 = \{1,\ldots,2\lfloor n^\alpha\rfloor\}$, $D_2 = \{\lfloor n^\alpha\rfloor,\ldots,3\lfloor n^\alpha\rfloor\}$ and so on, where $\lfloor\cdot\rfloor$ denote the integer part of a real number.

It is easily seen that this covering satisfies the hypothesis of Theorem 5. Indeed, take $i,j \in \{1,\ldots,n\}$ such that $i<j$. By stationarity, $\mathrm{Cov}(X_{i+1},X_j) = \phi(j-i)$. So if $\phi(j-i) \ge r_0 = \phi(\lfloor n^\alpha\rfloor)$, since ϕ is non-increasing, we must have $j-i \le \lfloor n^\alpha\rfloor$. By construction, any index i is at most in three different elements D of $\mathcal{C}(r_0)$. So i belongs to some D_{i_1}, D_{i_2} and D_{i_3} with $i_1<i_2<i_3$ and j belong to D_{j_1}, D_{j_2} and D_{j_3} with $j_1<j_2<j_3$. Since $j-i \le n^\alpha$ and the length of any D in $\mathcal{C}(r_0)$ is $2\lfloor n^\alpha\rfloor$, there exists $s \in \{1,2,3\}$ such that $D_{i_s} = D_{j_s}$ (draw a picture) and we can choose $D = D_{i_s} \in \mathcal{C}(r_0)$.

Clearly $\sum_{D\in\mathcal{C}(r_0)} 1_{\{I\in D\}} \le C$, where C is a universal constant (say $C=3$). To end the proof, we have to bound $\max_{D\in\mathcal{C}(r_0)}\mathbb{P}(I\in D)$. Let $D\in\mathcal{C}(r_0)$, and to ease the notation, consider $D = \{1,\ldots,2\lfloor n^\alpha\rfloor\}$ (the important feature is that D contains $2\lfloor n^\alpha\rfloor$ elements). Then

$$\mathbb{P}(I\in D) = \sum_{i=1}^{2\lfloor n^\alpha\rfloor}\mathbb{P}(I=i).$$

By standard Gaussian concentration (see Ledoux, 2001 or the appendix of Chatterjee, 2014), for all $i \in D$,

$$\begin{aligned}\mathbb{P}(I = i) &= \mathbb{P}(X_i = M_n) \\ &\le \mathbb{P}(X_i \ge t) + \mathbb{P}(M_n \le t) \\ &\le 2e^{-\mathbb{E}[M_n]^2/2}.\end{aligned}$$

The final step of the argument is to achieve a lower bound on $\mathbb{E}[M_n]$. To this task, we make use of Sudakov's minoration (see Adler, 1990 or Ledoux and Talagrand, 2011). Letting $i, j \in D$, $i \neq j$,

$$\mathbb{E}\left[(X_i - X_j)^2\right] = 2 - 2\phi\big(|j-i|\big) \ge 2\big(1 - \phi(1)\big) = \delta$$

since $\phi\big(|j-i|\big) \le \phi(1)$. By the assumption on $\phi(1)$, $\delta > 0$, and thus by Sudakov's minoration there exists $c > 0$ (independent of n) such that $\mathbb{E}[M_n] \ge c\delta\sqrt{\log n}$. As a consequence of the preceding,

$$\mathbb{P}(I = i) \le \frac{2}{n^{(c\delta)^2/2}} = \frac{2}{n^{\epsilon}}.$$

Hence

$$\mathbb{P}(I \in D) \le \frac{4}{n^{\epsilon-\alpha}} = \frac{4}{n^{\eta}}$$

where α is such that $\eta = \epsilon - \alpha > 0$, namely $\alpha < (c\delta)^2/2 < 1$.

The hypotheses of Theorem 5 are therefore fulfilled with $\rho(r_0) \le 4/n^{\eta}$ concluding the proof of Theorem 3.

3.2. *Proof of Theorem 4*

We again follow Chatterjee's proof of variance bounds for Gaussian Euclidean fields (see Theorem 9.12 in Chatterjee, 2014). For any Borel set $A \subset \mathbb{R}^d$, set $m(A) = \mathbb{E}[M(A)] = \mathbb{E}\left[\sup_{s\in A} X_s\right]$. Under the hypothesis of the theorem, it is proved in Chatterjee (2014) that for any $A \subset \mathbb{R}^d$ such that $N(A) > 1$,

$$c_1(\phi, d)\sqrt{\log N(A)} \le m(A) \le c_2(\phi, d)\sqrt{\log N(A)}$$

where $c_1 = c_1(\phi, d)$ and $c_2 = c_2(\phi, d)$ are positive constants that depend only on the covariance ϕ and the dimension d (and not on the subset A). Set then

$$s_0 = N(A)^{\frac{1}{8}(c_1/c_2)^2},$$

and assume that $N(A)$ is large enough so that $s_0 > 2$. Let $r_0 = \phi(s_0)$. Take a maximal s_0-net of A (i.e. a set of points that are mutually separated from each other by a distance strictly greater than s_0 and is maximal with respect to this property), and let $\mathcal{C}(r_0)$ be the set of $2s_0$-ball around the points in the net. In fact, $\mathcal{C}(r_0)$ is a covering of A satisfying the condition of Theorem 5. With this construction, it is easily seen that $\sum_{D\in\mathcal{C}(r_0)} 1_{\{I\in D\}} \le C$ for some $C \ge 1$ and

$$\max_{D\in\mathcal{C}(r_0)} \mathbb{P}(I \in D) \le \frac{1}{N(A)^{C(\phi,d)}}.$$

The conclusion then directly follows from Theorem 5 with $\rho(r_0) \le 1/N(A)^{C(\phi,d)}$.

4. Other Gaussian models

In Chatterjee (2014), Chatterjee exhibits different models in order to illustrate the so-called superconcentration phenomenon. For each of them, he managed to reach a better variance bound than the one produced by standard Gaussian concentration. We present for some of these models an exponential version.

Proposition 7. *For any $n \ge 2$, let $X = (X_1, \ldots, X_n)$ be a centered Gaussian vector with $\mathbb{E}[X_i]^2 = 1$ and $\mathbb{E}\left[X_iX_j\right] \le \epsilon$ for all $i, j = 1, \ldots, n$ for some $\epsilon > 0$. Then*

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| \ge t) \le 6e^{-ct/\sqrt{K_n}}, \quad t \ge 0,$$

where $K_n = \max(\epsilon, 1/\log n)$.

Proof. Take $r_0 > \epsilon$ and define $\mathcal{C}(r_0) = \{\{1\}, \ldots, \{N\}\}$. It is easy to see that the covering $\mathcal{C}(r_0)$ enters the setting of Theorem 5. Furthermore, $\sum_{D\in\mathcal{C}(r_0)} 1_{\{I\in D\}} = 1$ and $\max_{D\in\mathcal{C}(r_0)} \mathbb{P}(I \in D) \le 1/n^{\eta}$ for some $\eta > 0$. The conclusion follows. □

As another instance, consider the Gaussian field on $\{-1, 1\}^n$, $n \geq 1$, defined in the following way. Let $X_1, \ldots, X_n$ be independent identically distributed standard Gaussian random variables, and define $f : \{-1, 1\}^n \to \mathbb{R}$ as

$$f(\sigma) = \sum_{i=1}^{n} X_i \sigma_i.$$

Corollary 8. *There exists $S \subset \{-1, 1\}^n$, such that, for all $\sigma \neq \sigma' \in S$, $|\sigma \cdot \sigma'| \leq Cn^{2/3}$ and*

$$\mathbb{P}\left(\left|\max_{\sigma \in S} f(\sigma) - \mathbb{E}\big[\max_{\sigma \in S} f(\sigma)\big]\right| \geq t\right) \leq 6e^{-ct/n^{2/3}}, \quad t \geq 0.$$

Proof. Chatterjee proved in Chatterjee (2014) the existence of $S \subset \{-1, 1\}^n$ with $|S| = \lfloor 2^{n^{1/3}} \rfloor$ on which $|\sigma \cdot \sigma'| \leq Cn^{2/3}$ for some $C > 0$ independent of n. Since, $\mathrm{Cov}\big(f(\sigma), f(\sigma)\big) = \sigma \cdot \sigma'$, it suffices to apply Proposition 7 to $f_{|S}$ and $\epsilon = Cn^{2/3}$. □

5. Statistical testing

In this last section, we illustrate the preceding results on a statistical testing problem analyzed in Addario-Berry et al. (2010). There, the authors used standard Gaussian concentration to obtain an acceptance region for their test. Using some of the material developed here, the conclusion may be reinforced in some instances in the form of a superconcentration bound.

Following the framework and notation of Addario-Berry et al. (2010), we observe an n dimensional Gaussian vector $X = (X_1, \ldots, X_n)$ and raise the following hypothesis problem:

- Under the null hypothesis H_0, the components of X are independent identically distributed standard normal random variables. Denote then by $\mathbb{P}_0$ and $\mathbb{E}_0$ respectively the underlying probability measure and expectation under H_0.
- To describe the alternative hypothesis H_1, consider a class $\mathcal{C} = \{S_1, \ldots, S_N\}$ of N sets of indices such that $S_k \subset \{1, \ldots, n\}$ for all $k = 1, \ldots, N$. Under H_1, there exists an $S \in \mathcal{C}$ such that

  $$X_i \text{ has distribution } \begin{cases} \mathcal{N}(0, 1) & \text{if } i \in S^c, \\ \mathcal{N}(\mu, 1) & \text{if } i \in S, \end{cases}$$

  where $\mu > 0$ is a positive parameter. The components of X are independent under H_1 as well. For any $S \in \mathcal{C}$, denote then by $\mathbb{P}_S$ and $\mathbb{E}_S$ respectively the underlying probability measure and expectation under H_1. We will also assume that every $S \in \mathcal{C}$ has the same cardinality $|S| = K$.

We recall that a test is a binary-valued function $f : \mathbb{R}^n \to \{0, 1\}$. If $f(X) = 0$, the test accepts the null hypothesis, otherwise H_0 is rejected. As in Addario-Berry et al. (2010), consider the risk of a test f measured by

$$R(f) = \mathbb{P}_0\big(f(X) = 1\big) + \frac{1}{N} \sum_{S \in \mathcal{C}} \mathbb{P}_S\big(f(X) = 0\big).$$

The authors of Addario-Berry et al. (2010) are interested in determining, or at least estimating, the value of μ under which the risk can be made small. Among others results, they used a test based on maxima, called the scan-test, for which they showed that

$$f(X) = 1 \iff 2 \max_{S \in \mathcal{C}} X_S \geq \mu K + \mathbb{E}_0\Big[\max_{S \in \mathcal{C}} X_S\Big]$$

where $X_S = \sum_{i \in S} X_i$ for some $S \in \mathcal{C}$. They obtain the following result.

Proposition 9. *The risk of the maximum test f satisfies $R(f) \leq \delta$ whenever*

$$\mu \geq \frac{1}{K} \mathbb{E}_0\Big[\max_{S \in \mathcal{C}} X_S\Big] + 2\sqrt{\frac{2}{K} \log \frac{2}{\delta}}.$$

Together with (1.5), this bound may be improved in accordance with the correct magnitude of the maximum of a standard Gaussian vector.

Proposition 10. *For any $n \geq 2$, the risk of the maximum test f satisfies $R(f) \leq \delta$ whenever*

$$\mu \geq \frac{1}{K} \mathbb{E}_0\Big[\max_{S \in \mathcal{C}} X_S\Big] + \log \frac{6}{\delta} \times \frac{2}{c\sqrt{K \log N}}$$

where $c > 0$ is a universal constant.

Proof. We follow the proof of Addario-Berry et al. (2010) to get that both, for every $t \geq 0$,

$$\mathbb{P}_0\left(\max_{S\in\mathcal{C}} X_S \geq \mathbb{E}_0\left[\max_{S\in\mathcal{C}} X_S\right] + t\right) \leq 3e^{-ct\sqrt{\log(N)/K}}$$

and, for any $S \in \mathcal{C}$,

$$\mathbb{P}_S\left(\max_{S'\in\mathcal{C}} X_{S'} \leq \mu K - t\right) \leq \mathbb{P}_S\left(\max_{S'\in\mathcal{C}} X_{S'} \leq \mathbb{E}_{S'}\left[\max_{S'\in\mathcal{C}} X_{S'}\right] - t\right)$$
$$\leq 3e^{-ct\sqrt{\log(N)/K}}$$

since, under H_1, for a fixed $S \in \mathcal{C}$, $\mu K = \mathbb{E}_S\left[X_S\right] \leq \mathbb{E}_S\left[\max_{S'\in\mathcal{C}} X_{S'}\right]$ and X_S has the same law as any $X_{S'}$. Set then t such that

$$2t = \mu K - \mathbb{E}\left[\max_{S\in\mathcal{C}} X_S\right]$$

which is positive according to the hypothesis. With the previous inequalities, we obtain a bound on the risk

$$R(f) \leq 6e^{-ct\sqrt{\log(N)/K}}$$

which may then be turned into $R(f) \leq \delta$ as in the statement. $\square$

The new bound of Proposition 10 is as good as the one of Proposition 9 when $\delta \simeq 1/N^{\alpha}$ for some $\alpha > 0$. It is better than Proposition 9 when $\delta \simeq 1/\log^{\alpha}(N)$ with $\alpha > 0$. Nevertheless, Proposition 9 is better than Proposition 10 when $\delta \simeq e^{-N^{\alpha}}$, for $\alpha > 0$. However, it might not be relevant to consider such small risk.

Acknowledgments

I thank my Ph.D. advisor M. Ledoux for introducing this problem to me and for fruitful discussions. I also thank the anonymous referee for helpful comments in improving the exposition.

References

Acosta, J.A., 2014. Tightness of the recentered maximum of log-correlated gaussian field. Electron. J. Probab. 19 (90).
Addario-Berry, L., Broutin, N., Devroye, L., Lugosi, G., 2010. On combinatorial testing problems. Ann. Statist. 38 (5), 2010.
Adler, R.J., 1990. An introduction to continuity, extrema and related topics for general Gaussian processes. In: Institut of Mathematical Statistics Lectures Notes. In: Monograph Series, vol. 12.
Bakry, D., Gentil, I., Ledoux, M., 2014. Analysis and geometry of Markov diffusion operators. Grundlehren Math. Wiss. 348.
Bakry, D., Ledoux, M., Saloff-Coste, L., 2012. Markov semigroups at Saint-Flour. In: Probability at Saint-Flour. Springer, Heidelberg.
Benaïm, M., Rossignol, R., 2008. Exponential concentration for first passage percolation through modified poincaré inequalities. Ann. Inst. Henri Poincaré Probab. Stat. 44 (3).
Chatterjee, S., 2014. Superconcentration and Related Topics. Springer.
Damron, M., Hanson, J., Sosoe, P., 2014. Subdiffusive concentration in first-passage percolation. Electron. J. Probab..
Ding, J., Eldan, R., Zhai, A., 2013. On multiple peaks and moderate deviations for supremum of gaussian field. Preprint, http://arxiv.org/abs/1311.5592v2.
Ding, J., Roy, R., Zeitouni, O., 2015. Convergence of the centered maximum of log-correlated gaussian fields. Preprint, http://arxiv.org/abs/1503.04588.
Duplantier, B., Rhodes, R., Sheffield, S., Vargas, V., 2014. Critical gaussian multiplicative chaos: Convergence of the derivative martingale. Ann. Probab. 42 (5).
Duplantier, B., Rhodes, R., Sheffield, S., Vargas, V., 2014. Log-correlated gaussian fields: an overview. Preprint, http://arxiv.org/abs/1407.5605.
Leadbetter, M.R., Lindgren, G., Rootzén, H., 1983. Extremes and Related Properties of Random Sequences and Processes. In: Springer Series in Statistics.
Ledoux, M., 2001. The concentration of measure phenomenon. Math. Surveys Monogr. 89.
Ledoux, M., Talagrand, M., 2011. Probability in Banach spaces. Isoperimetry and processes. Springer-Verlag, Berlin.
Madaule, T., 2014. Maximum of a log-correlated gaussian field. Arxiv (preprint).