Courbures, entrelacements et inégalités fonctionnelles pour quelques processus de Markov

Aldéric Joulin

INSTITUT de MATHEMATIQUES de TOULOUSE
UPS CNRS INSA UT1 UT2

Université de Toulouse

HABILITATION À DIRIGER DES RECHERCHES

Discipline : **Mathématiques**

Spécialités : **Probabilités et Statistiques, Analyse**

Courbures, entrelacements et inégalités fonctionnelles pour quelques processus de Markov

Document de synthèse présenté par

Aldéric JOULIN

Au vu des rapports de

Sergey BOBKOV	University of Minnesota
Pietro CAPUTO	Università Roma Tre
Olivier GUEDON	Université Paris-Est Marne-la-Vallée

Soutenue publiquement le 29 novembre 2019 devant le jury composé de

Sergey BOBKOV	University of Minnesota	Rapporteur
Pietro CAPUTO	Università Roma Tre	Rapporteur
Patrick CATTIAUX	Université Paul Sabatier, Toulouse 3	Parrain
Gersende FORT	Université Paul Sabatier, Toulouse 3	Examinatrice
Olivier GUEDON	Université Paris-Est Marne-la-Vallée	Rapporteur
Florent MALRIEU	Université François-Rabelais, Tours	Examinateur
Sylvie ROELLY	Universität Potsdam	Examinatrice

À mes sombres héros trop tôt disparus...

Remerciements

« Par une belle matinée de mai, une svelte amazone, montée sur une somptueuse jument alezane, parcourait, au milieu des fleurs, les allées du Bois... » C'est par ces quelques mots que Joseph Grand, l'un des personnages du roman La Peste d'Albert Camus, commençait l'ouvrage de sa vie, qui ne comportera malheureusement que cette seule phrase... Ah, qu'il est difficile de trouver l'inspiration, surtout lorsqu'il s'agit comme dans le cas présent de rédiger des remerciements. Et dire que ces quelques lignes sont écrites quelques semaines après avoir apposé un point final au charabia qui va suivre, quelle ironie ! Comme si l'on contait une histoire en débutant par la fin... Dans ce cas, le suspense étant quelque peu dissipé, je serai bref.

Mes premiers remerciements sont destinés aux trois rapporteurs de cette habilitation, Sergey Bobkov, Pietro Caputo et Olivier Guedon. Malgré mes (vaines ?) tentatives pour rendre la lecture du manuscrit aussi vivante et agréable que possible, un tel travail d'arbitrage de leur part n'est pas une mince affaire, surtout lorsque les travaux décrits s'étalent sur une grosse dizaine d'années et nécessitent par conséquent une mise en perspective par rapport au contexte scientifique de l'époque. Alors merci à vous trois de m'avoir fait l'honneur de rapporter ce mémoire d'habilitation. On ne présente plus Sergey Bobkov, un mathématicien exceptionnel, dont le nombre de citations de ses articles apparaissant dans ce manuscrit témoigne véritablement de l'influence qu'ont eu ses travaux sur mes orientations et intérêts scientifiques. Cher Sergey, soyez-en sincèrement remercié. J'adresse également mes remerciements à Pietro Caputo, que je connais depuis un moment déjà. Pietro, ta générosité, ton dynamisme et ta culture scientifique m'ont toujours impressionné. Enfin, je tiens à exprimer ma profonde reconnaissance envers Olivier Guedon, que j'ai rencontré plus récemment et avec lequel le contact est passé instantanément. En particulier, ayant eu le courage de rester jusqu'à la fin de mon exposé marathon proposé en mai dernier dans le cadre du groupe de travail C-TOP à l'IHP (quelle aventure !), j'en déduisis, mon cher Olivier, que tu ne pouvais être tout à fait réfractaire à nos entrelacements et autres futilités développés dans ce mémoire. Merci à toi pour le temps passé à déchiffrer tout cela.

Je désire remercier tout spécialement Gersende Fort, Florent Malrieu et Sylvie Roelly qui ont accepté de participer à ce jury. Gersende, ton expertise markovienne et ta bonne humeur permanente font de toi une collègue toulousaine avec laquelle il est toujours agréable de discuter. Florent, lorsque je relis les remerciements figurant dans mon manuscrit de thèse, je vois que tu y as déjà une place importante : comme quoi, y être de nouveau présent treize années plus tard témoigne parfaitement de la bienveillance dont tu as toujours fait preuve à mon égard. Sylvie, je me souviens avec plaisir de la première édition du Collège doctoral franco-allemand Nanterre/Potsdam 2006-2008 et de la visite du magnifique palais de Sanssouci ponctuant une journée d'exposés : science et culture combinées, une formule que tu affectionnes tant et qui fonctionne à merveille.

Parmi les mathématiciens que j'ai eu la chance de côtoyer, je souhaiterais en mentionner quelques-uns qui m'ont tant inspiré tout au long de ces années : Patrick Cattiaux, Michel Ledoux et Dominique Bakry. Des personnalités. Patrick, notre première rencontre date d'une conférence toulousaine organisée en avril 2006 sous l'égide de l'ANR IFO que tu portais à l'époque. En quelques mots, je te dois beaucoup (au-delà d'un immense respect et de quelques euros) et en particulier ma rencontre avec Christian Léonard qui m'a fait venir à Nanterre par

la suite. Je me souviens des discussions à bâtons rompus que nous avions à des horaires tardifs dans les couloirs alors désertés de l'entresol du laboratoire MODAL'X. Depuis, de l'eau a coulé sous les ponts (plus précisément la Garonne) et deux co-encadrements de thèse sont passés par là. Lorsque j'ai su qu'il me fallait trouver un parrain pour cette habilitation, j'ai immédiatement pensé à toi pour remplir ce rôle de coordinateur : un grand merci d'avoir accepté. Michel, il n'y a pas un seul mois de l'année sans que je n'aille fouiner dans l'un de tes articles pour y trouver un résultat remarquable dont j'ai besoin ou encore une référence historique précise. Par ailleurs, je te remercie pour ta disponibilité, tes conseils et ton aide précieuse concernant l'organisation de cette habilitation. Dominique, chaque fois que je me suis rendu à ton bureau pour te poser une question, je repartais avec une série d'autres questions auxquelles je n'avais absolument pas songé. Quant à la réponse à ma question initiale, encore aurait-il fallu que je la comprenne... Je détournerai ainsi le récent détournement par Clément Steiner d'une célèbre citation de Pierre Desproges : « on peut difficilement s'empêcher de penser que si Napoléon avait vécu cent-soixante ans de plus, il aurait adoré le critère Γ_2 ».

Étant un homme dont la pénibilité dans le travail est directement proportionnelle à la qualité des passements de jambes lorsque l'entorse est évitée, j'ai malgré tout découvert durant ces dix dernières années l'existence de quelques âmes charitables d'ici et d'ailleurs (notamment de l'empire du milieu) ayant accepté de travailler avec moi sur des projets scientifiques plus ou moins bien définis, dont certains ont donné lieu à des articles de recherche (certains diront que le miracle existe et n'est pas que gaussien). Qu'elles en soient présentement remerciées ! Parmi celles-ci, l'une est devenue au fil des années mon « âme frère » scientifique (la preuve, il soutient son habilitation la veille de la mienne). Il s'agit du collègue et néanmoins ami que je ne peux m'empêcher d'appeler dès que je pense avoir une idée intéressante pour aborder tel ou tel problème mathématique (certains me conseilleront de réduire mon forfait téléphonique) car il maîtrise aussi bien le disque de Poincaré que celui de l'Ultimate : Michel Bonnefont. Merci pour toutes choses passées, présentes et à venir. Pour moi, tu seras toujours sélectionnable en équipe de France des inégalités fonctionnelles...

À présent approche un moment délicat, celui de citer nommément les collègues de l'IMT éparpillés sur l'ensemble des sites toulousains, avec lesquels je prends plaisir à discuter, rire, déjeuner, épuiser les pots et quelquefois travailler. Bref, vous tous qui me donnez envie chaque matin de parcourir ces nombreux kilomètres en train-vélo indépendamment de la météo et du climat social. La liste étant non dénombrable, je ne prendrai aucun risque (je tiens à ma peau) en me gardant de citer quiconque (ou alors juste une allusion, car je ne peux y résister : mon bureau médiéval au département GMM de l'INSA est vraiment charmant). Je salue également les « fonctionnalistes » que j'ai pris l'habitude de rencontrer en dehors de la Ville Rose, par exemple lors de conférences mémorables organisées grâce aux projets ANR EVOL, GEMECOD et STAB. Au fil des années et des expériences communes, certains sont devenus des amis proches, notamment à travers les précieuses LSM : merci à vous pour votre bonne humeur contagieuse, votre notion du temps quelque peu altérée (surtout en fin de parcours) et, *last but not least*, votre envie d'en découdre. Parmi les camarades éminents de cette association à but hautement non lucratif, j'ai une pensée toute particulière pour toi, Fabulous Fab, angevin depuis 2016 et avec qui j'ai partagé par ailleurs un nombre incalculable de soirées montalbanaises. Quel duo topalogique nous formions à l'INSA ! Assurément, le LAREMA a de beaux jours devant lui... Quant à toi, Clément, tu es chez toi chez moi (ceci n'a rien à voir avec ton prochain

déménagement) comme je me sens chez moi chez toi (la famille Pellegrini l'a bien compris) : quel plaisir de nous retrouver en dehors des sentiers mathématiques ! Malgré l'ouragan qui va t'emporter prochainement, espérons que cela continue encore longtemps.

L'un des rôles fondamentaux d'un enseignant-chercheur est de transmettre des savoirs aux futures générations. N'en étant pas forcément le meilleur vecteur, je tiens ainsi à adresser mes remerciements à mes anciens étudiants dont j'ai co-encadré la thèse et qui, *in fine*, ont dû subir en quelque sorte mon apprentissage difficile, ainsi qu'à Bertrand Cloez dont j'ai encadré le Post-Doctorat. Claire, je te souhaite de t'épanouir pleinement à Bath dans le projet sur la tomographie médicale que tu as choisi. À vrai dire, je n'ai aucun doute sur le fait que tes qualités scientifiques et humaines te permettront une adaptation rapide à ton nouvel environnement. Chtimi, reste tel que tu es, un mathématicien dynamique et curieux ainsi qu'un ami dont le sourire permanent témoigne d'une réelle pointe de finesse dans ce monde de matheux (tiens, c'est de qui, ça ?). Enfin, Clément, j'éprouve beaucoup de plaisir à co-diriger ton travail de thèse. Comme l'a dit un jour un célèbre cadre picto-charentais d'une non moins célèbre entreprise commercialisant du café, « la route est droite mais la pente est forte ». Nous sommes donc sur le bon chemin, l'important étant d'avancer à la bonne vitesse.

Vers l'infini et au-delà, il y a le reste, c'est-à-dire ce qui n'est pas mathématique, donc à peu près tout. Les amis, la famille. Muriel, un immense merci pour toutes ces belles années passées à tes côtés. Que de choses ai-je apprises en te côtoyant. Quant aux Bobies, Raphaël et Aloïs, vous êtes décidément mes plus beaux théorèmes (certains diront les seuls).

Les maux de la fin. Cette année 2019 fût pour moi une année pleine d'émotions, sans doute l'une des plus intenses de ma vie. À la fois magnifique et tragique. C'est pourquoi je ne peux clore ces remerciements sans mentionner l'une des personnes qui m'a le plus inspiré tout au long de ma vie, et qui partit par une belle matinée de juillet : Dominique, dit Pépito. Un phare dans mon voyage au bout de la nuit.

Montauban, novembre 2019

N.D.A. : Finalement, je fus plus long que prévu : merci d'avoir parcouru ces quelques lignes jusqu'au bout et salut à toi, ô mon lecteur courageux.

Table des matières

Chapitre 1

Présentation

Ce manuscrit présente une synthèse de mes principaux travaux de recherche effectués depuis mon arrivée dans la Ville Rose en septembre 2007, s'étalant peu ou prou sur une grosse dizaine d'années. Durant cette période, je me suis principalement intéressé à des problèmes mathématiques situés à l'interface entre la théorie des probabilités et l'analyse fonctionnelle, en passant (très modestement) par la géométrie et l'analyse des équations aux dérivées partielles. Plus précisément, si l'une des questions fondamentales potentiellement au centre des préoccupations d'un probabiliste est de savoir si un phénomène aléatoire donné se stabilise en temps long vers un équilibre, mes travaux de recherche portent principalement autour de cette problématique pour des dynamiques markoviennes. En particulier, un outil puissant pour établir et quantifier la convergence de tels modèles est la théorie des inégalités fonctionnelles (en partie) développée depuis plusieurs décennies par l'école toulousaine.

1.1 Organisation du mémoire

Afin de fluidifier la lecture de ce document, j'ai opté pour un regroupement de mes travaux de recherche en deux parties distinctes, suivant plus ou moins un ordre chronologique dans leur élaboration et correspondant chacune à un chapitre du mémoire. Le chapitre 2 concerne la période 2008-2013 et est dédié aux diverses notions de courbures de chaînes et processus de Markov que j'ai étudiées, seul ou en collaboration, dans le cadre de leur convergence vers l'équilibre et du phénomène de concentration de leur mesure invariante et de leur loi instantanée. Le chapitre 3, quant à lui, est consacré à la notion d'entrelacement, un outil que j'ai développé depuis 2013 principalement avec Michel Bonnefont, dans le cadre d'un espace aussi bien discret que continu et qui est lui aussi étroitement lié aux propriétés d'ergodicité des dynamiques markoviennes considérées. S'en déduisent alors certaines conséquences intéressantes en matière d'inégalités spectrales et fonctionnelles pour la mesure invariante et l'opérateur sous-jacent.

1.2 Centres d'intérêts, publications et encadrement scientifique

Depuis mes premiers pas dans le monde de la recherche en septembre 2003 et le début de ma thèse de doctorat, je me suis intéressé à plusieurs problèmes mathématiques mélangeant à la fois la théorie des probabilités et l'analyse fonctionnelle, comme en témoigne la liste de mes centres d'intérêts scientifiques suivante :

[A] Autour des chaînes et processus de Markov :

- **[A1]** Propriétés spectrales des opérateurs markoviens.
- **[A2]** Inégalités fonctionnelles.
- **[A3]** Géométrie des semigroupes.
- **[A4]** Convergence vers l'équilibre.
- **[A5]** Concentration de la mesure.
- **[A6]** Méthode de Stein.

[B] Autour des processus de type Lévy :

- **[B1]** Théorèmes limites.
- **[B2]** Processus de type stable.
- **[B3]** Espaces de Wiener et de Poisson, calcul de Malliavin.
- **[B4]** Ordres stochastiques.

Pour la rédaction de ce mémoire, j'ai décidé de n'aborder que certaines de mes publications et prépublications (celles illustrant le mieux mes préoccupations passées et actuelles) qui seront détaillées dans les chapitres 2 et 3. Les autres travaux sont seulement listés en précisant pour chacun le thème auquel il se rattache principalement. Les publications [1] à [4] correspondent au travail effectué pendant la thèse.

0. *Concentration et fluctuations de processus stochastiques avec sauts*. 163 pages. Thèse de Doctorat de l'Université de La Rochelle, octobre 2006.

1. Functional inequalities for discrete gradients and applications to the geometric distribution, avec Nicolas Privault. *ESAIM Probability and Statistics*, 8 : 87-101, 2004. **[A1,A2,A5]**

2. A logarithmic Sobolev inequality for an interacting spin system under a geometric reference measure, avec Nicolas Privault. Actes de la 26ème conférence « Quantum probability and infinite dimensional analysis » (Levico, 2005). *Quantum Probability and White Noise Analysis*, XX : 267-273, 2007, World Scientific. **[A1,A2,A5]**

3. On maximal inequalities for stable stochastic integrals. *Potential Analysis*, 26 : 57-78, 2007. **[B1,B2]**

4. Poisson-type deviation inequalities for curved continuous time Markov chains. *Bernoulli*, 13 : 782-798, 2007. **[A3,A4,A5]**

5. A new Poisson-type deviation inequality for Markov jump processes with positive Wasserstein curvature. *Bernoulli*, 15 : 532-549, 2009. ↪ **Chapitre 2**

6. Curvature, concentration, and error estimates for Markov chain Monte Carlo, avec Yann Ollivier. *Annals of Probability*, 38 : 2418-2442, 2010. ↪ **Chapitre 2**

7. Upper bounds on Rubinstein distances on configuration spaces and applications, avec Laurent Decreusefond et Nicolas Savy. *Communications on Stochastic Analysis*, 4 :377-399, 2010. **[A2,A5,B3]**

8. Measure concentration through non-Lipschitz observables and functional inequalities, avec Arnaud Guillin. *Electronic Journal of Probability*, 18 : 1-26, 2013. ↪ **Chapitre 2**

9. Intertwining and commutation relations for birth-death processes, avec Djalil Chafaï. *Bernoulli*, 19 : 1855-1879, 2013. **[A2,A3,B4]**

10. Intertwining relations for one-dimensional diffusions and application to functional inequalities, avec Michel Bonnefont. *Potential Analysis*, 41 : 1005-1031, 2014. ↪ **Chapitre 3**

11. A note on convex ordering for stable stochastic integrals, avec Solym Mawaki Manou-Abi. *Stochastics*, 87 : 592-603, 2015. **[B2,B4]**

12. Spectral gap for spherically symmetric log-concave probability measures, and beyond, avec Michel Bonnefont et Yutao Ma. *Journal of Functional Analysis*, 270 : 2456-2482, 2016. ↪ **Chapitre 3**

13. A note on spectral gap and weighted Poincaré inequalities for some one-dimensional diffusions, avec Michel Bonnefont et Yutao Ma. *ESAIM Probability and Statistics*, 20 : 18-29, 2016. ↪ **Chapitre 3**

14. Intertwinings and generalized Brascamp-Lieb inequalities, avec Marc Arnaudon et Michel Bonnefont. *Revista Matemática Iberoamericana*, 34 : 1021-1054, 2018. ↪ **Chapitre 3**

15. Intertwinings, second-order Brascamp-Lieb inequalities and spectral estimates, avec Michel Bonnefont. *ArXiv 1710.08106*, 23 pages, 2017 (nouvelle version, 29 pages, 2019). ↪ **Chapitre 3**

16. A note on eigenvalues estimates for one-dimensional diffusion operators, avec Michel Bonnefont. *Arxiv 1906.02496*, 24 pages, 2019. ↪ **Chapitre 3**

Durant ces dernières années, j'ai eu l'opportunité de co-encadrer des thèses de doctorat avec des collègues de l'Institut de Mathématiques de Toulouse, ainsi que de recruter un jeune chercheur sur un support de Post-doctorat :

∘ Encadrement du Post-Doctorat de B. Cloez (2013-2014). Sujet proposé : méthode de Stein et inégalités fonctionnelles. **[A2,A4,A6]**

∘ Co-encadrement avec P. Cattiaux de la thèse de S.-M. Manou-Abi (2012-2016). Titre de la thèse : *Théorèmes limites et ordres stochastiques relatifs aux lois et processus stables*. **[B1,B2,B4]**

◦ Co-encadrement avec L. Miclo de la thèse de C. Delplancke (2013-2017). Titre de la thèse : *Méthodes quantitatives pour le comportement asymptotique de processus de Markov homogènes et non-homogènes*. **[A2,A4,A6]**

◦ Co-encadrement avec P. Cattiaux de la thèse de C. Steiner (2018-). Titre (provisoire) de la thèse : *Entrelacements et concentration de la mesure*. **[A2,A3,A5]**

Chapitre 2

Courbures, inégalités fonctionnelles et concentration de la mesure

2.1 Contexte et rappels

Dans tout ce premier chapitre, on se donne un espace polonais (E, d) muni de sa tribu borélienne, qui est considéré indifféremment discret, tel un ensemble fini ou dénombrable, un graphe muni de sa distance naturelle de graphe, ou continu avec suffisamment de structure, comme par exemple l'espace euclidien $\mathbb{R}^d$ ou une variété riemannienne. Les processus de Markov $(X_t)_{t\geq 0}$ à temps continu que l'on considère sont de la forme suivante :

∘ si E est discret, on travaille avec un processus dit de sauts purs sur E, c'est-à-dire qu'il est constant par morceaux et son générateur est donné par

$$Lf(x) \quad = \quad \int_E (f(y) - f(x))\, Q(x, dy), \quad x \in E,$$

où pour tout $x \in E$, $Q(x, \cdot)$ est une mesure positive finie correspondant aux taux de transition du processus. L'exemple typique que l'on a en tête est un processus simple par sa structure mais néanmoins très intéressant car rassemblant les difficultés inhérentes au cadre discret : le processus de naissance et de mort. Il s'agit d'un processus de Markov à valeurs dans $\mathbb{N}$, de sauts seulement de taille 1, dont le générateur L est donné par

$$Lf(x) \quad = \quad \lambda_x \, (f(x+1) - f(x)) + \nu_x \, (f(x-1) - f(x)), \quad x \in \mathbb{N},$$

où les taux de transition de naissance et de mort, respectivement λ et ν, sont des fonctions strictement positives avec $\nu_0 = 0$, assurant l'irréductibilité du processus, mais éventuellement non bornées. Sous des hypothèses adéquates, le processus admet une unique mesure de probabilité invariante μ, qui est réversible pour la dynamique.

∘ si E est continu, le processus que l'on considère est un processus de diffusion dont le générateur associé est un opérateur différentiel du second ordre sans coefficient constant. Par exemple sur l'espace euclidien $E = \mathbb{R}^d$, le processus $(X_t)_{t\geq 0}$ est l'unique solution forte de l'Équation Différentielle Stochastique (ÉDS) suivante :

$$dX_t \quad = \quad \sqrt{2}\, \sigma(X_t)\, dB_t + b(X_t)\, dt,$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien standard sur $\mathbb{R}^d$ indépendant de X_0, la matrice de diffusion σ, de taille $d \times d$, et le champ de vecteurs b satisfont les hypothèses de régularité appropriées. Le générateur est donné pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$ par

$$Lf \quad = \quad \sum_{i,j=1}^{d} \sigma\sigma^T_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i \frac{\partial f}{\partial x_i}.$$

De même, l'exemple auquel on pense immédiatement est le processus dit de Kolmogorov sur $\mathbb{R}^d$, solution de l'ÉDS

$$dX_t \quad = \quad \sqrt{2}\, dB_t - \nabla V(X_t)\, dt,$$

où $V : \mathbb{R}^d \to \mathbb{R}$ est un potentiel régulier tel que la mesure μ dont la densité de Lebesgue est proportionnelle à e^{-V} soit une mesure de probabilité. Le générateur est donné par

$$Lf = \Delta f - \nabla V \cdot \nabla f,$$

et la mesure μ est alors invariante et réversible pour cette dynamique.

Étant donné un processus de Markov $(X_t)_{t\geq 0}$ à valeurs dans l'espace polonais (E, d), qu'il soit de sauts purs ou de diffusion, sa loi instantanée converge-t-elle en temps long vers une mesure d'équilibre μ et si oui, dans quel sens et à quelle vitesse ? Parmi les pistes proposées pour aborder ce type de problème, une des méthodes principales nécessitant potentiellement peu d'informations sur l'éventuelle mesure invariante μ consiste à étudier ce que l'on appelle la courbure de la dynamique, dont l'analogie avec la courbure de Ricci pour des variétés riemanniennes est à présent bien comprise, depuis le travail fondateur de Bakry et Émery [13] dans les années 80 jusqu'aux travaux plus récents datant du milieu des années 2000 de Sturm et von Renesse [168] et de Kuwada [123] notamment, les différentes notions de courbure de processus introduites étant essentiellement équivalentes dans ce contexte continu. Notons par ailleurs la généralisation de ces travaux dans les espaces métriques mesurés due parallèlement à Sturm [166] et à Lott et Villani [140], ces contributions majeures étant principalement basées sur une approche par le transport optimal en lien avec la convexité de l'entropie relative dans l'ensemble des mesures de probabilités. En revanche, les espaces discrets sont dépourvus d'une géométrie convenable (pas de structure différentiable ni de notion de géodésique pour l'espace de Wasserstein associé), d'où la difficulté d'étudier la convergence vers l'équilibre des modèles discrets, malgré l'apparente simplicité de la dynamique. À l'époque, les tentatives pour définir une bonne notion de courbure dans des espaces discrets puis l'utiliser pour établir des inégalités fonctionnelles encodant la convergence en temps long du processus, reposaient principalement sur une adaptation du critère de courbure-dimension de Bakry-Émery, un critère par essence « continu » bien que compatible avec la structure inhérente d'un espace discret, cf. les articles [7, 163] pour des processus de Markov de sauts purs sur des graphes mais aussi les travaux de l'école italienne pour des systèmes de particules en interaction [43, 49].

2.1.1 Courbure de Wasserstein

Je ne fus pas le dernier à tenter ma chance en proposant au milieu des années 2000 une notion de courbure que j'espérais intrinsèque, c'est-à-dire permettant d'unifier les cadres discret et continu dans l'étude de la convergence en temps long de ces dynamiques markoviennes. Pour ce faire, dans l'article [111] paru en 2007, j'adaptai au cadre discret l'approche combinée couplage/transport optimal initiée par Sturm et von Renesse, en proposant de considérer une notion appelée courbure de Wasserstein, basée sur la distance de Wasserstein d'ordre 1. Pour d'autres concepts plus ou moins équivalents, avec quelques nuances (en temps discret ou continu et éventuellement au-delà du monde markovien), on pourra consulter l'article fondateur de Dobrushin [74], les travaux de Marton [142] ou encore ceux autour de Chen [64, 65] ainsi que les références plus récentes [73, 161] précédant de quelques années mon premier article [111] sur le sujet.

Avant tout, procédons à de brefs rappels à propos du transport optimal. Si $\mathcal{P}(E)$ désigne l'ensemble des mesures de probabilité sur E et $\mathcal{P}_1(E)$ le sous-ensemble des mesures de probabilité m admettant un moment d'ordre 1, c'est-à-dire telles qu'il existe $x_0 \in E$ tel que $\int_E d(x, x_0)\, m(dx) < \infty$, le coût de transport optimal entre deux mesures $m_1, m_2 \in \mathcal{P}_1(E)$ est défini par

$$W_1(m_1, m_2) \;=\; \inf_{\pi \in \Pi(m_1,m_2)} \int_E \int_E d(x,y)\, \pi(dx, dy),$$

où $\Pi(m_1, m_2)$ est l'ensemble des mesures de probabilité sur l'espace produit $E \times E$ ayant pour première et seconde marginales m_1 et m_2, respectivement. Ci-dessus, la distance $d(x, y)$ représente le coût de transport d'une unité de masse de x vers y et $\pi(dx, dy)$ la quantité déplacée de x en y. L'espace (E, d) étant supposé polonais, ce coût est alors atteint et est une distance sur $\mathcal{P}_1(E)$ appelée distance de Wasserstein d'ordre 1, ou encore distance de Kantorovich-Rubinstein. L'ensemble $(\mathcal{P}_1(E), W_1)$ est alors un espace polonais nommé l'espace de Wasserstein d'ordre 1 et la convergence en distance W_1 est équivalente à la convergence faible des mesures de probabilité et des moments d'ordre 1. En particulier, si d est la distance triviale sur E, $d(x, y) = 1_{\{x \neq y\}}$, alors W_1 coïncide avec la distance en variation totale définie pour toutes mesures de probabilité $m_1, m_2 \in \mathcal{P}(E)$ par

$$d_{\mathrm{VT}}(m_1, m_2) \;=\; \sup_{A \in \mathcal{B}(E)} |m_1(A) - m_2(A)|.$$

Plutôt que d'utiliser les distances de Wasserstein d'ordre supérieur, le choix de la distance W_1 est motivé d'une part car elle donne lieu à une notion de courbure à venir la plus faible possible et permet donc d'appréhender plus de cas, notamment dans des espaces discrets, et d'autre part elle est plus facile à contrôler grâce à la célèbre dualité de Kantorovich-Rubinstein qui, dans le cas de la distance de Wasserstein W_1, permet la représentation alternative suivante :

$$W_1(m_1, m_2) \;=\; \sup_{f \in \mathrm{Lip}_1} \int_E f\, dm_1 - \int_E f\, dm_2,$$

où Lip désigne l'espace des fonctions lipschitziennes sur l'espace métrique (E, d) muni de la seminorme

$$\|f\|_{\mathrm{Lip}} \;=\; \sup_{x \neq y \in E} \frac{|f(x) - f(y)|}{d(x, y)},$$

et $\mathrm{Lip}_1 \subset \mathrm{Lip}$ est le sous-ensemble des fonctions 1-lipschitziennes. Pour plus de détails à propos de ces distances entre mesures de probabilité, nous renvoyons à l'ouvrage de référence de Villani sur le transport optimal [177].

À présent, nous sommes en mesure d'introduire la notion de courbure de Wasserstein. Ci-dessous la notation $(P_t)_{t\geq 0}$ désigne le semigroupe associé au processus de Markov $(X_t)_{t\geq 0}$ et si $\nu \in \mathcal{P}(E)$, alors νP_t désigne la loi instantanée de X_t sachant que X_0 suit la loi ν. Enfin, la mesure δ_x est la masse de Dirac au point $x \in E$.

Définition 2.1.1. *Soit $(X_t)_{t\geq 0}$ un processus de Markov à valeurs dans (E, d) et de semigroupe associé $(P_t)_{t\geq 0}$ tel que $\delta_x P_t \in \mathcal{P}_1(E)$ pour tout $x \in E$ et tout $t \geq 0$. Sa courbure de Wasserstein est définie par la meilleure (i.e., la plus grande) constante σ_d telle que la contraction exponentielle suivante ait lieu :*

$$W_1(\delta_x P_t, \delta_y P_t) \;\leq\; e^{-\sigma_d t}\, d(x, y), \quad x, y \in E, \quad t \geq 0.$$

Bien évidemment, la courbure dépend fondamentalement de la distance d et changer la métrique revient à la modifier. Plutôt que de parler de courbure, qui est une notion locale par définition, on devrait parler ici de « courbure minorée » car si le processus est le mouvement brownien standard sur une variété riemannienne, la constante σ_d n'est rien d'autre que la borne inférieure sur la courbure de Ricci de la variété, cf. [168]. En particulier, cette notion de courbure est pertinente pour les trois espaces modèles « en courbure constante » et étroitement liés par les propriétés géométriques et fonctionnelles qu'ils partagent - grâce au Lemme de Poincaré (sur la limite gaussienne de projections euclidiennes de la mesure de volume normalisée sur une sphère en très grande dimension) et au Théorème Central Limite (TCL) - que sont :

∘ la sphère unité $E = \mathbb{S}^d \subset \mathbb{R}^{d+1}$ de dimension d munie de sa distance gédésique et de sa mesure de volume normalisée, qui se révèle être la mesure invariante du mouvement brownien standard sur cette variété riemannienne. La courbure de Wasserstein coïncide avec la plus grande borne inférieure sur la courbure de Ricci de cette variété, qui est constante dans le cas de la sphère et vaut $d - 1$.

∘ l'espace euclidien $E = \mathbb{R}^d$ muni de la mesure gaussienne standard γ (centrée et de matrice de covariance l'identité) : le processus de Markov sous-jacent est un processus de Kolmogorov particulier nommé le processus d'Ornstein-Uhlenbeck, de potentiel $V(x) = |x|^2/2$ et solution de l'ÉDS

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

ayant pour mesure invariante γ. Pour ce modèle, la courbure de Wasserstein vaut 1.

∘ l'hypercube $E = \{0, 1\}^d$ muni de la distance de Hamming, comptant le nombre de coordonnées différentes entre deux points de $\{0, 1\}^d$. Ici, la mesure de probabilité uniforme est associée à la marche aléatoire simple à temps continu sur cet hypercube, de loi instantanée produit

$$\delta_x P_t(\{y\}) \;=\; \frac{1}{2^d} \prod_{i=1}^{d} \left(1 + (-1)^{|x_i - y_i|}\, e^{-t/d}\right), \quad x, y \in \{0, 1\}^d.$$

La courbure de Wasserstein vaut $1/d$. En effet, cette courbure ayant de bonnes propriétés de tensorisation, le caractère produit de cette dynamique permet de ramener le calcul à la dimension 1, situation dans laquelle elle vaut trivialement 1.

Notons que par la dualité de Kantorovich-Rubinstein, une définition alternative de courbure de Wasserstein peut être proposée. En effet, la courbure de Wasserstein est la meilleure constante σ_d dans l'inégalité suivante : pour toute fonction $f \in \mathrm{Lip}$,

$$\|P_t f\|_{\mathrm{Lip}} \leq e^{-\sigma_d t} \|f\|_{\mathrm{Lip}}, \quad t \geq 0.$$

En particulier si cette courbure est strictement positive, alors le semigroupe est une contraction stricte sur l'espace Lip. On voit ainsi l'intérêt d'obtenir la stricte positivité d'une telle courbure au sens où elle est fortement reliée à l'ergodicité du processus.

Théorème 2.1.2. *Si la courbure de Wasserstein σ_d est strictement positive, alors il existe une unique mesure invariante $\mu \in \mathcal{P}_1(E)$ telle que l'on ait la convergence exponentielle en distance de Wasserstein W_1 suivante : pour toute loi initiale $\nu \in \mathcal{P}_1(E)$,*

$$W_1(\nu P_t, \mu) \leq e^{-\sigma_d t} W_1(\nu, \mu), \quad t \geq 0.$$

2.1.2 Inégalités fonctionnelles

Voyons à présent une autre façon d'appréhender l'ergodicité du processus. Pour ce faire, nous allons introduire les inégalités fonctionnelles principales dont nous aurons besoin dans la suite. Tout d'abord, définissons l'un des protagonistes principaux de ce mémoire, l'opérateur carré du champ Γ : il s'agit de la forme bilinéaire symétrique suivante : pour toutes fonctions $f, g : E \to \mathbb{R}$ suffisamment régulières,

$$\Gamma(f,g) = \frac{1}{2}\left(L(fg) - f\,Lg - g\,Lf\right).$$

Dans le cadre discret, il s'écrit

$$\Gamma(f,g)(x) = \frac{1}{2}\int_E (f(y)-f(x))\,(g(y)-g(x))\,Q(x,dy), \quad x \in E,$$

tandis qu'en continu,

$$\Gamma(f,g) = \sigma^T \nabla f \cdot \sigma^T \nabla g.$$

À présent, définissons l'opérateur carré du champ itéré Γ_2 comme la forme bilinéaire symétrique suivante : pour toutes fonctions réelles f, g suffisamment régulières sur E,

$$\Gamma_2(f,g) = \frac{1}{2}\left(L\Gamma(f,g) - \Gamma(Lf,g) - \Gamma(f,Lg)\right),$$

obtenue par un procédé itératif à partir du carré du champ en remplaçant le produit par Γ lui-même. Le célèbre critère de courbure-dimension $\mathrm{CD}(\rho,\infty)$ de Bakry-Émery [13] est le suivant.

Définition 2.1.3. *Étant donnée une constante réelle ρ, on dit que la dynamique markovienne satisfait le critère de courbure-dimension* $\mathrm{CD}(\rho,\infty)$ *si pour toute fonction réelle f suffisamment régulière sur E, on a l'inégalité ponctuelle*

$$\Gamma_2(f,f) \quad \geq \quad \rho\,\Gamma(f,f).$$

La meilleure constante (i.e., la plus grande) dans cette inégalité est notée ρ_{BE}.

Par exemple, dans le cadre du processus de Kolmogorov à valeurs dans l'espace euclidien $\mathbb{R}^d$, l'opérateur Γ_2 s'écrit

$$\Gamma_2(f,f) \quad = \quad \|\nabla^2 f\|_{\mathrm{HS}}^2 + \nabla f \cdot \nabla^2 V\, \nabla f,$$

de sorte que $\rho_{\mathrm{BE}} = \inf_{\mathbb{R}^d} \rho(\nabla^2 V)$. Ci-dessus, les notations $\rho(\nabla^2 V)$ et $\|A\|_{\mathrm{HS}} = \sqrt{\mathrm{Trace}(AA^T)}$ désignent respectivement la plus petite valeur propre de la matrice hessienne $\nabla^2 V$ et la norme de Hilbert-Schmidt d'une matrice A. Ce critère est optimal dans le cas gaussien standard pour lequel on a $\rho_{\mathrm{BE}} = 1$, une quantité coïncidant avec la courbure de Wasserstein.

Étant donnée une mesure de probabilité μ sur E invariante pour le processus, l'opérateur $-L$ est positif dans $L^2(\mu)$ car Γ est une forme bilinéaire symétrique positive. S'il est symétrique, *i.e.*, pour toutes fonctions $f,g : E \to \mathbb{R}$ suffisamment régulières,

$$\int_E f\,(-Lg)\,d\mu \quad = \quad \int_E (-Lf)\,g\,d\mu \;=\; \int_E \Gamma(f,g)\,d\mu,$$

l'opérateur $-L$ est (essentiellement) auto-adjoint sur l'espace $L^2(\mu)$ (car borné inférieurement) : on parle alors de réversibilité de la dynamique. Dans le cadre discret, la réversibilité se lit simplement sur la propriété

$$\mu(dx)\,Q(x,dy) \quad = \quad \mu(dy)\,Q(y,dx), \quad x,y \in E,$$

appelée communément la condition de balance détaillée (ou équilibre détaillé).

Consacrons-nous à présent aux inégalités fonctionnelles. La première dont nous allons nous servir est connue sous le nom d'inégalité de Poincaré.

Définition 2.1.4. *Étant donnée une constante $\lambda > 0$, on dit que l'inégalité de Poincaré de constante λ est satisfaite si pour toute fonction $f : E \to \mathbb{R}$ suffisamment régulière, on a*

$$\lambda\,\mathrm{Var}_\mu(f) \quad \leq \quad \int_E \Gamma(f,f)\,d\mu,$$

où $\mathrm{Var}_\mu(f)$ *désigne la variance de la fonction f sous μ, à savoir*

$$\mathrm{Var}_\mu(f) \quad = \quad \int_E f^2\,d\mu - \left(\int_E f\,d\mu\right)^2.$$

La constante optimale (i.e., la plus grande) dans l'inégalité de Poincaré est notée λ_1.

Comme on le verra dans la suite, il est intéressant quelquefois d'utiliser la variance sous sa forme dupliquée, à savoir

$$\mathrm{Var}_\mu(f) \;=\; \frac{1}{2}\int_E\int_E (f(x)-f(y))^2 \; \mu(dx)\,\mu(dy).$$

L'intérêt d'obtenir une telle inégalité de Poincaré est qu'elle est étroitement liée à la convergence exponentielle dans l'espace $L^2(\mu)$ du semigroupe associé.

Théorème 2.1.5. *L'inégalité de Poincaré de constante $\lambda > 0$ a lieu si et seulement si l'on a l'inégalité suivante : pour toute fonction $f \in L^2(\mu)$,*

$$\|P_t f - \mu(f)\|_{L^2(\mu)} \;\leq\; e^{-\lambda t}\,\|P_t f - \mu(f)\|_{L^2(\mu)}, \quad t \geq 0,$$

où $\mu(f)$ est l'intégrale de f sous μ.

Notons que sous l'hypothèse de réversibilité de la dynamique, on a la proposition suivante, très commode en pratique.

Proposition 2.1.6. *Étant donnée une constante $\lambda > 0$, on a dans le cadre réversible l'équivalence suivante : l'inégalité de Poincaré de constante λ est satisfaite si et seulement si pour toute fonction réelle f suffisamment régulière sur E,*

$$\int_E \Gamma_2(f,f)\,d\mu \;\geq\; \lambda \int_E \Gamma(f,f)\,d\mu.$$

Cette inégalité, qui n'est rien d'autre que le critère de courbure dimension $\mathrm{CD}(\lambda,\infty)$ intégré, peut être vue comme une inégalité duale de celle de Poincaré. De surcroît, toujours dans le cadre réversible, l'inégalité de Poincaré de constante $\lambda > 0$ est équivalente au fait que le spectre de l'opérateur $-L$ soit inclus dans l'ensemble $\{0\} \cup [\lambda,\infty[$. Dans ce cas, la constante optimale λ_1 est appelée le trou spectral de l'opérateur $-L$, la première valeur propre λ_0 étant nulle (les fonctions propres associées sont les constantes, la mesure μ étant finie). En général, estimer le trou spectral n'est pas une chose aisée : c'est l'un des objectifs principaux du chapitre 3 que nous aborderons ultérieurement.

Une autre inégalité fonctionnelle d'un grand intérêt est l'inégalité entropique, ou inégalité de Sobolev logarithmique modifiée, suivante.

Définition 2.1.7. *Étant donnée une constante $\rho > 0$, on dit que l'inégalité entropique de constante ρ est vérifiée si pour toute fonction $f : E \to]0,\infty[$ suffisamment régulière, on a*

$$\rho\,\mathrm{Ent}_\mu(f) \;\leq\; \int_E f\,(-L\log f)\,d\mu,$$

où $\mathrm{Ent}_\mu(f)$ désigne l'entropie de la fonction f sous μ :

$$\mathrm{Ent}_\mu(f) \;=\; \int_E f\,\log f\,d\mu - \int_E f\,d\mu\,\log\int_E f\,d\mu.$$

On note ρ_0 la constante optimale (i.e., la plus grande) dans cette inégalité.

Dans le cadre diffusif, l'inégalité entropique coïncide à une constante près avec la célèbre inégalité de Sobolev logarithmique de Gross apparue dans l'article fondateur [88],

$$\mathrm{Ent}_\mu(f^2) \quad \leq \quad C \int_E \Gamma(f,f)\, d\mu,$$

dans lequel il démontre l'équivalence de cette inégalité avec la propriété d'hypercontractivité du semigroupe. En revanche dans le cadre discret, la règle de dérivation de Leibniz n'est pas vérifiée et cela entraîne que l'inégalité entropique est différente de l'inégalité de Sobolev logarithmique étudiée initialement par Diaconis et Saloff-Coste [72] pour estimer le temps de mélange de chaînes de Markov sur un ensemble fini, d'où son appellation quelquefois d'inégalité de Sobolev logarithmique modifiée, cf. par exemple les articles [49, 50, 70] dans lesquels est établie l'inégalité entropique sous la condition de réversibilité pour divers systèmes de particules en interaction, cette inégalité se réécrivant alors sous la forme plus symétrique,

$$\begin{aligned} \rho\, \mathrm{Ent}_\mu(f) \quad &\leq \quad \int_E \Gamma(f, \log f)\, d\mu \\ &= \quad \frac{1}{2} \int_E \int_E (f(x) - f(y))\, (\log f(x) - \log f(y))\, Q(x,dy)\, \mu(dx). \end{aligned}$$

Citons encore les travaux [57, 182] tournant autour d'inégalités de ce type pour des mesures de Poisson, ainsi que le travail initial de Bobkov et Ledoux [29] sur les entiers naturels où d'autres inégalités modifiées sont introduites pour étudier le phénomène de concentration de la mesure. Dans le cadre des chaînes de Markov à temps continu réversibles sur un ensemble fini, l'article de Bobkov et Tetali [31] propose une hiérarchie entre les différentes inégalités de Sobolev logarithmiques modifiées et parmi celles-ci, il s'avère que la plus forte est justement l'inégalité entropique (elle reste malgré tout moins forte que l'inégalité de Sobolev logarithmique classique).

Parmi toutes les inégalités de Sobolev logarithmiques introduites, l'inégalité entropique est celle qui encode la décroissance exponentielle du semigroupe en entropie, cf. par exemple les articles [31, 72, 152] pour une preuve dans le cas discret non-réversible, un contexte moins fréquemment étudié dans la littérature.

Théorème 2.1.8. *L'inégalité entropique de constante $\rho > 0$ a lieu si et seulement si l'on a l'inégalité suivante : pour toute fonction $f : E \to]0,\infty[$ d'entropie finie, on a*

$$\mathrm{Ent}_\mu(P_t f) \quad \leq \quad e^{-\rho t}\, \mathrm{Ent}_\mu(f), \quad t \geq 0.$$

À présent, tentons de dégager une hiérarchie entre toutes ces constantes associées aux différentes propriétés d'ergodicité du processus. Tout d'abord, on a toujours l'inégalité

$$\lambda_1 \quad \geq \quad \frac{\rho_0}{2},$$

(prendre $f = 1 + \varepsilon g$ dans l'inégalité entropique puis faire tendre ε vers 0), c'est-à-dire que l'inégalité entropique est plus forte que l'inégalité de Poincaré. Dans le cadre diffusif, on a la fameuse inégalité due à Bakry et Émery [13],

$$\rho_0 \quad \geq \quad 2\, \rho_{\mathrm{BE}},$$

traduisant l'obtention de l'inégalité de Sobolev logarithmique en courbure strictement positive. Ces deux inégalités se révèlent optimales car ce sont des égalités dans le cas du processus d'Ornstein-Uhlenbeck de mesure invariante γ, pour lequel on a $\lambda_1 = \rho_{\mathrm{BE}} = 1$ et $\rho_0 = 2$. En revanche, le lien entre la courbure de Wasserstein et les constantes optimales λ_1 et ρ_0 n'est absolument pas clair dans le cadre discret. Néanmoins, on a le résultat suivant, sans doute dû à Chen même s'il est difficile à dater précisément dans son oeuvre foisonnante. En guise de référence, on pourra consulter le chapitre 9 de son livre [61]. Ci-dessous, l'hypothèse de réversibilité est due à l'approche spectrale utilisée par Chen.

Théorème 2.1.9. *On suppose que la courbure de Wasserstein σ_d d'un processus de Markov réversible est strictement positive. Alors l'inégalité de Poincaré de constante σ_d a lieu ou, en d'autres termes, $\lambda_1 \geq \sigma_d$.*

Au vu de l'ensemble des éléments que nous venons de décrire, reliant la courbure de Wasserstein aux diverses propriétés d'ergodicité du processus, il n'est pas surprenant d'envisager qu'une hypothèse de courbure de Wasserstein strictement positive puisse contribuer à la vitesse de convergence du théorème ergodique, ou Loi des Grands Nombres (LGN) markovienne, à savoir la convergence en temps long des mesures empiriques du type $L_t = t^{-1} \int_0^t \delta_{X_s}\, ds$ vers l'équilibre μ. En particulier, si l'on note $\mathbb{P}_\nu$ la probabilité $\mathbb{P}$ sachant que X_0 suit la loi ν et $L_t(f)$ l'intégrale d'une fonction donnée $f : E \to \mathbb{R}$ par rapport à la mesure empirique L_t, une façon classique de quantifier cette convergence est d'obtenir des bornes non-asymptotiques en temps sur les probabilités de déviation du type

$$\mathbb{P}_\nu \left(|L_t(f) - \mu(f)| \geq r \right), \quad r > 0,$$

en fonction des paramètres d'intérêt de la dynamique markovienne. Une telle étude sort complètement du cadre asymptotique de la théorie des grandes déviations : c'est l'objet de ce que l'on appelle communément le phénomène de concentration de la mesure.

2.1.3 Concentration de la mesure

Avant de rentrer plus en détail dans les résultats que nous allons exposer incessamment sous peu, procédons à quelques rappels sur la concentration de la mesure. Munissons l'espace polonais (E, d) d'une mesure de probabilité, disons ν. Le phénomène de concentration de la mesure (ou plutôt sa version fonctionnelle) a lieu lorsque cet espace métrique mesuré, observé à travers une bonne classe de fonctions réelles définies sur E, apparaît plus petit qu'il ne l'est véritablement. Mathématiquement, cela se traduit par l'existence d'une fonction $\alpha_\nu : [0, \infty) \to [0, 1]$ tendant vers 0 à l'infini telle que l'on ait l'inégalité suivante :

$$\nu \left(\{ x \in E : |f(x) - m_\nu(f)| > r \} \right) \quad \leq \quad 2\, \alpha_\nu(r), \quad r > 0,$$

où $m_\nu(f)$ est une médiane de f sous ν, qui peut être remplacée par $\nu(f)$, la moyenne de f sous ν, ou toute autre constante, quitte à translater le niveau de déviation r. La fonction f est classiquement supposée lipschitzienne par rapport à la distance d (et même 1-lipschitzienne pour une raison de normalisation, d'où le fait qu'elle n'apparaisse pas dans la fonction α_ν), ce choix étant fortement motivé par des questions isopérimétriques (on se réfèrera aux travaux précurseurs de Lévy datant de 1919). Lorsque l'on s'intéresse à la concentration du point de vue de la fonction f, on dit alors que f se concentre autour de l'une de ses médianes $m_\nu(f)$ à vitesse α_ν.

Le phénomène de concentration de la mesure a des conséquences et ramifications importantes en analyse fonctionnelle et en géométrie, comme en statistiques, une thématique dans laquelle le caractère hautement dimensionnel des objets manipulés est primordial. C'est là l'une des propriétés les plus intéressantes de la concentration de la mesure, amenant la célèbre observation de Talagrand [129, 172] que l'on peut traduire ainsi : « les fonctions d'un grand nombre de variables ayant de petites variations sont (presque) constantes ». Le phénomène de concentration de la mesure a été découvert par Milman au début des années 70, dans le cadre de ses travaux sur la géométrie des espaces de Banach. Il a notamment utilisé dans [151] de telles idées de concentration pour proposer une nouvelle démonstration du théorème de Dvoretzky sur les sections (presque) sphériques des corps convexes. Pour n'en citer que quelques-uns, s'ensuivirent à partir des années 80 les contributions majeures d'inspiration géométrique de Gromov [85, 86, 87] jusqu'aux travaux de Talagrand [171] dans les espaces produits, mêlant à la fois les approches combinatoire, fonctionnelle et probabiliste. Enfin, un nouveau point de vue sur la concentration de la mesure a été proposé au milieu des années 90 d'une part par Bobkov et Ledoux, cf. [28, 29, 126, 127], basé essentiellement sur la théorie des inégalités fonctionnelles de type entropique et la fameuse méthode de Herbst (due à Herbst en 1975 dans une lettre non puliée adressée à Gross), et d'autre part par Marton [142] suivie de près par Bobkov et Götze [25] à propos des liens étroits que la concentration entretient avec le transport optimal de mesures. Pour une introduction en douceur sur le sujet ainsi qu'une liste exhaustive de références historiques, le lecteur intéressé pourra consulter le livre de Ledoux [129], tandis que le *survey* plus récent de Gozlan-Léonard [83] propose un état de l'art des inégalités de transport reliées à la concentration de la mesure.

Si la fonction α_ν est du type $\alpha_\nu(r) = \exp(-c\,r^2)$, où $c > 0$ est une constante dépendant de certains paramètres associés à l'espace métrique mesuré (E, d, ν), on parle de concentration gaussienne, car elle est similaire à celle de la loi gaussienne γ. Dans le cadre continu, ce type de concentration s'obtient par exemple par la méthode de Herbst à partir d'une inégalité de Sobolev logarithmique, cf. [128] pour une présentation générale de l'argument. En travaillant un peu, cette approche peut être adaptée au cadre discret et plus précisément pour les ensembles finis (la réversibilité jouant alors un rôle important, car elle permet de contourner l'absence de la règle de Leibniz), cf. [31, 105], l'inégalité de Sobolev logarithmique n'étant en général pas vérifiée lorsque E est dénombrable. En revanche, si la décroissance de la fonction α_ν est plutôt en $\exp(-c\,r)$, il s'agit alors d'un régime exponentiel, établi par exemple pour les lois exponentielle et géométrique, cf. [28, 114]. L'inégalité fonctionnelle derrière ce comportement est l'inégalité de Poincaré, un phénomène mis en avant par Gromov et Milman [87] dans le cas continu et retrouvé par Aida et ses coauteurs dans [1, 2], entraînant l'intégrabilité exponentielle de la mesure ν. Dans le cas des graphes, cette analyse est due à Alon et Milman [4]. Le TCL embusqué, le bon comportement est en fait une décroissance mixte gaussienne/exponentielle du type $\alpha_\nu(r) = \exp(-\min\{c\,r^2, Cr\})$, un régime gaussien apparaissant pour les petits niveaux de déviation r. Par exemple, ce caractère mixte apparaît dans les travaux de Talagrand [170] et de Maurey [143], ainsi qu'ultérieurement dans les articles [28, 105, 114]. L'intérêt d'une approche de la concentration par de telles inégalités fonctionnelles réside dans l'aspect adimensionnel des résultats obtenus, ces inégalités ayant de bonnes propriétés de tensorisation. Bien évidemment, il existe d'autres types de concentration apparaissant fréquemment dans la littérature, comme la concentration poissonnienne typique des espaces discrets (car analogue à la queue de distribution d'une loi de Poisson) en $\exp(-c\, r \log(1 + C\, r))$, cf. [7, 29, 102, 104, 111, 115, 162, 182], dont certaines de ces inégalités sont issues d'inégalités de Sobolev logarithmiques modifiées, ou encore celle pour les mesures de probabilité à queue lourde pour lesquelles α_ν décroît seulement polynomialement, ceci étant dû bien sûr à un manque d'intégrabilité exponentielle. Citons par exemple les travaux de Houdré et ses coauteurs [46, 103] sur le sujet. Pour une approche statistique mathématique de la concentration des processus empiriques mettant en exergue ces divers comportements en fonction des paramètres importants du modèle considéré, dont certaines des inégalités associées ont pour nom les inégalités de Hoeffding, de Bernstein et de Bennett, on pourra consulter l'ouvrage de référence [42].

2.2 Quelques résultats

À présent, nous allons exposer quelques-uns des résultats que nous avons obtenus principalement dans les articles [89, 112, 113], correspondant chacun à une partie indépendante des deux autres. Néanmoins, le fil conducteur de ces travaux est le lien étroit entre courbures, inégalités fonctionnelles et concentration de la mesure réunissant à la fois les cadres discret et continu.

2.2.1 Courbure de Wasserstein et concentration de la moyenne empirique

Comme annoncé plus haut, il est tentant d'imaginer que la courbure de Wasserstein puisse avoir des conséquences intéressantes sur la LGN markovienne, et plus précisément en termes de concentration de la moyenne empirique $L_t(f)$ autour de l'équilibre $\mu(f)$ pour une certaine classe de fonctions f : c'est l'objet du premier travail que je vais présenter dans ce mémoire, tiré de l'article [112] datant de 2009, et qui unifie les cadres discret et continu.

Dans les dix années précédant ce travail, de nombreux résultats sur le sujet ont vu le jour, mettant en exergue des comportements différents des processus de Markov considérés sous diverses hypothèses. Tout d'abord, en utilisant une approche de type Feynman-Kac, Wu [181] a établi pour des processus de Markov aussi bien discrets que continus une borne de déviation pour des fonctions f intégrables par rapport à μ, exhibant une décroissance exponentielle en temps du type $\exp(-t\beta(r))$. Cette estimation est optimale dans le cadre réversible à la fois en temps long et par rapport au niveau de déviation r, la fonction β coïncidant ici avec la fonction de taux du Principe de Grandes Déviations (PGD) sous-jacent qui n'est malheureusement pas explicite. C'est pourquoi ce travail a ensuite fait l'objet d'améliorations par plusieurs auteurs ayant proposé, sous des hypothèses de « Lipschitzianité » ou de moment sur les fonctions f considérées, des inégalités de concentration gaussiennes ou exponentielles (voire même polynomiales) essentiellement pour des diffusions dont la mesure invariante satisfait certaines inégalités fonctionnelles comme celle de Poincaré, de Sobolev logarithmique ou encore des inégalités de transport, cf. par exemple les articles [53, 73, 90]. À l'époque, les techniques utilisées reposaient fortement sur l'aspect diffusif des processus et étaient par conséquent difficilement transposables à des processus de Markov de sauts purs. Mentionnons cependant le remarquable travail de Lezaud [136, 137] à partir de la fin des années 90, unifiant les milieux discret et continu en utilisant une approche originale, la théorie des perturbations de Kato. Les bornes de déviation qu'il a alors obtenues pour des fonctions f bornées sont poissonniennes, sous l'hypothèse d'un trou spectral satisfait par la dynamique qui de surcroît est supposée bornée.

Le contexte à présent exposé, décrivons le résultat principal contenu dans notre article [112]. Bien que l'approche adoptée unifie à la fois les cadres discret et continu, concentrons-nous essentiellement sur l'aspect discret. Le processus de Markov générique $(X_t)_{t\geq 0}$ à valeurs dans l'ensemble discret (E,d) est de sauts purs, la dynamique $x \to Q(x,E)$ étant éventuellement non bornée sur E. De surcroît, on suppose

$$V(x) \quad = \quad \frac{1}{2} \int_E d(x,y)^2 \, Q(x,dy) \quad < \infty,$$

cette quantité étant entendue comme un contrôle du carré du champ appliqué aux fonctions lipschitziennes, *i.e.*, pour toute fonction $f \in \mathrm{Lip}$, on a immédiatement l'inégalité ponctuelle $\Gamma(f,f) \leq V \, \|f\|_{\mathrm{Lip}}^2$. De plus, si V est intégrable par rapport à μ ou bornée, alors cela permet de dominer la variance de f sous la mesure invariante μ dès lors que l'inégalité de Poincaré a lieu, ce qui est le cas sous les hypothèses de réversiblité et de courbure de Wasserstein strictement positive, d'après le Théorème 2.1.9 : pour toute fonction $f \in \mathrm{Lip}$, on a

$$\mathrm{Var}_\mu(f) \quad \leq \quad \frac{\|V\|_\infty \, \|f\|_{\mathrm{Lip}}^2}{\sigma_d},$$

où $\| \cdot \|_\infty$ désigne la norme « infini » sur l'espace discret E.

Combinée à une méthode de tensorisation issue de la discrétisation temporelle de la moyenne empirique (la difficulté essentielle réside dans la spécificité de la loi jointe d'un vecteur extrait du processus de Markov, qui n'est assurément pas produit mais seulement conditionnellement produit), l'utilisation de la courbure de Wasserstein nous a permis d'établir de nouvelles inégalités de concentration poissonniennes non-asymptotiques en temps pour des fonctions lipschitziennes. En particulier, la réversibilité n'y joue aucun rôle spécifique.

Théorème 2.2.1. *On suppose que la courbure de Wasserstein σ_d du processus est strictement positive. De plus, on fait l'hypothèse que les sauts sont bornés par une constante $b > 0$ et que la fonction V est bornée. Alors pour toute fonction $f \in \mathrm{Lip}$, tout $x \in E$ et tout temps $t \geq 0$, on a : pour tout $r > 0$,*

$$\mathbb{P}_{\delta_x} \left(|L_t(f) - \mu(f)| > r + M_{t,x} \right) \quad \leq \quad 2 \, \exp \left(- \frac{tr}{2bC_{t,\sigma_d} \|f\|_{\mathrm{Lip}}} \log \left(1 + \frac{br}{2\|V\|_\infty C_{t,\sigma_d} \|f\|_{\mathrm{Lip}}} \right) \right), \tag{2.2.1}$$

où C_{t,σ_d} désigne la constante $C_{t,\sigma_d} = (1 - e^{-\sigma_d t})/\sigma_d$ et $M_{t,x}$ est le terme correctif dû au centrage, $M_{t,x} = C_{t,\sigma_d} \, \|f\|_{\mathrm{Lip}} \, W_1(\mu, \delta_x)/t$, qui tend vers 0 lorsque t tend vers l'infini.

Tout d'abord, notons que subséquemment à l'hypothèse de courbure, il est nécessaire de supposer la bornitude des sauts du processus. En effet, dans le cas contraire, considérons sur $(\mathbb{Z}, |\cdot|)$ le processus le plus simple auquel on puisse penser, celui qui à chaque instant $t > 0$ reste en l'état initial avec probabilité e^{-t} ou est distribué avec probabilité $(1 - e^{-t})$ selon une mesure de probabilité μ de support plein sur $\mathbb{Z}$, qui se révèle être son unique mesure invariante. En particulier, sa courbure de Wasserstein vaut 1. Néanmoins, si la mesure μ n'admet pas de moment exponentiel (par exemple, une loi à queue lourde fera l'affaire), la loi instantanée de X_t n'admet pas elle non plus de moment exponentiel, d'où l'impossibilité d'obtenir une inégalité de concentration même exponentielle de la moyenne empirique autour de l'équilibre : dans la situation présente, les sauts ne sont pas bornés.

Commentons à présent l'inégalité (2.2.1). Il s'avère qu'elle étend dans plusieurs directions celles proposées par Lezaud dans ses articles [136, 137]. Tout d'abord, les fonctions f considérées ici ne sont pas forcément bornées mais lipschitziennes, tout comme la dynamique $x \to Q(x, E)$ qui n'est pas supposée bornée. Enfin, aucune régularité particulière n'est requise sur la loi initiale (disons ν) du processus alors que Lezaud la suppose absolument continue par rapport à la mesure invariante μ, de densité (disons g) de carré intégrable dans $L^2(\mu)$, une hypothèse intrinsèquement liée à l'approche spectrale qu'il utilise. En particulier si E est fini et ν est une masse de Dirac δ_x, la quantité $\|g\|_{L^2(\mu)}$ vaut $1/\sqrt{\mu(\{x\})}$ et a donc vocation à exploser lorsque E devient grand. Néanmoins, le prix à payer pour que l'on obtienne de telles améliorations est de supposer une courbure de Wasserstein strictement positive, qui est une hypothèse plus forte en situation de réversibilité que celle mise en avant par Lezaud, à savoir l'existence d'un trou spectral, cf. le Théorème 2.1.9.

Dans le cadre diffusif pour lequel la technique peut être adaptée aisément, cette méthode est suffisamment robuste pour récupérer une décroissance gaussienne faisant intervenir le supremum essentiel (par rapport à la mesure de référence sur l'espace E considéré) du carré du champ à la place de la quantité $\|V\|_\infty \|f\|_{\mathrm{Lip}}^2$. Notons par ailleurs que le régime gaussien apparaît de nouveau, indépendamment de la nature discrète ou continue de l'espace E, lorsqu'on applique le TCL markovien (ceci correspond à prendre r petit, l'ordre de grandeur étant en $1/\sqrt{t}$). En revanche, contrairement au second travail de Lezaud [137], on ne récupère pas exactement la variance asymptotique $\sigma_f^2 = \lim_{t\to\infty} t \,\mathrm{Var}_\mu(L_t(f))$ de la fonction f mais simplement une borne supérieure sur cette quantité dans le cas réversible. En effet, on peut montrer en travaillant un peu l'inégalité

$$\sigma_f^2 \quad \leq \quad \frac{2\|V\|_\infty \|f\|_{\mathrm{Lip}}^2}{\sigma_d^2}.$$

À titre d'exemples, le Théorème 2.2.1 peut aussi bien être appliqué à des modèles multidimensionnels qu'aux processus de naissance et de mort à valeurs entières que nous allons maintenant considérer. L'exemple prototypique d'un tel processus est une célèbre file d'attente, nommée $M/M/\infty$ selon la classification de Kendall pour les processus de queues markoviens, ayant pour générateur

$$Lf(x) \quad = \quad \lambda\, (f(x+1) - f(x)) + x\, (f(x-1) - f(x)), \quad x \in \mathbb{N},$$

où les taux de transition de naissance et de mort sont respectivement constant et linéaire. La mesure invariante μ est la loi de Poisson de paramètre $\lambda > 0$. Bien que d'apparence simple, il s'agit d'une dynamique non bornée sur $\mathbb{N}$ pour laquelle aucun des résultats de déviation mentionnés jusqu'à présent n'était réellement disponible. Néanmoins, nous sommes en mesure de lui appliquer le Théorème 2.2.1. Voyons tout d'abord comment procéder dans le cas général. Si u est une fonction strictement positive sur $\mathbb{N}$ avec par convention $u_{-1} = 1$, on observe que la fonction bivariée

$$d(x,y) \quad = \quad \left| \sum_{k=0}^{x-1} u_k - \sum_{k=0}^{y-1} u_k \right|, \quad x, y \in \mathbb{N},$$

définit une distance sur les entiers naturels. L'intérêt d'introduire ce type de distance est qu'elle préserve la structure géodésique, c'est-à-dire qu'étant donnés plusieurs points distincts sur $\mathbb{N}$, la distance entre les deux plus éloignés correspond à la somme des distances intermédiaires, ce qui permet dans la détermination de la courbure de Wasserstein σ_d de réduire le problème aux points voisins. On démontre alors par un argument de couplage le résultat suivant.

Théorème 2.2.2. *La courbure de Wasserstein du processus définie en fonction de cette distance vaut*

$$\sigma_d \quad = \quad \inf_{x \in \mathbb{N}} \left\{ \nu_{x+1} + \lambda_x - \nu_x \frac{u_{x-1}}{u_x} - \lambda_{x+1} \frac{u_{x+1}}{u_x} \right\}.$$

Mentionnons que Van Doorn [173] a montré que le trou spectral λ_1, qu'il appelle le paramètre de décroissance, majorait cette constante σ_d pour certains choix de fonction u, sans qu'aucun lien n'ait été fait avec l'aspect métrique et une quelconque décroissance exponentielle en distance de Wasserstein. Plus tard, Chen [60] a établi par la méthode de couplage mentionnée ci-dessus que le trou spectral valait exactement le supremum de ces constantes σ_d pris sur l'ensemble de ces fonctions strictement positives u. Autrement dit, si ce supremum est atteint, il s'agit alors d'un cas d'égalité dans le Théorème 2.1.9, la distance d reliée à la fonction u réalisant le supremum s'écrivant alors $d(x,y) = |g_1(x) - g_1(y)|$, où g_1 est une fonction propre associée au trou spectral, qui dans ce cas s'avère être une vraie valeur propre.

Bref, il reste ainsi à trouver une distance d construite à partir d'une fonction strictement positive u (on rappelle que $u_x = d(x, x+1)$ pour tout $x \in \mathbb{N}$) telle que :

∘ la courbure de Wasserstein σ_d soit strictement positive ;

∘ les sauts du processus soient bornés relativement à la distance d, ce qui revient à supposer la bornitude de la fonction u ;

∘ la fonction V soit bornée, avec ici

$$\begin{aligned} V(x) \quad &= \quad \frac{1}{2} \left(\lambda_x \, d(x, x+1)^2 + \nu_x \, d(x, x-1)^2 \right) \\ &= \quad \frac{1}{2} \left(\lambda_x \, u_x^2 + \nu_x \, u_{x-1}^2 \right). \end{aligned}$$

Ces hypothèses sont souvent faciles à satisfaire en pratique, auquel cas le Théorème 2.2.1 peut s'appliquer pour des fonctions lipschitziennes par rapport à cette distance d. *In fine*, si l'on revient à la file d'attente $M/M/\infty$, on récupère alors une borne de déviation poissonnienne pour des fonctions f dont l'ordre de grandeur à l'infini est en $\sqrt{x}$, la distance d étant construite à partir d'une fonction u à croissance au plus $1/\sqrt{x}$ à l'infini. Bien que sûrement sous-optimal (on s'attend plutôt à ce que la fonction identité elle-même satisfasse cette borne de déviation), ce résultat est sans doute l'un des premiers, sinon le premier, à considérer de telles dynamiques non bornées en sortant du cadre des fonctions f bornées.

2.2.2 Courbure de Ricci grossière et méthodes MCMC

En 2009, Yann Ollivier a introduit dans un article retentissant [155] une notion de courbure pour des chaînes de Markov, qu'il a appelée la courbure de Ricci grossière (ou encore courbure de Ricci discrète, à présent connue sous le nom de courbure de Ricci-Ollivier) et qui ressemble dans l'esprit à la courbure de Wasserstein, la différence majeure au-delà de la distinction temps discret/continu étant qu'elle est plus locale car définie pour un couple de points. Plus précisément, la définition est la suivante.

Définition 2.2.3. *Étant donnée une chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ à valeurs dans (E,d) et de matrice de transition P telle que $\delta_x P \in \mathcal{P}_1(E)$ pour tout $x \in E$, sa courbure de Ricci grossière $\kappa(x,y)$ entre deux points distincts $x,y \in E$ est définie par la quantité*

$$\kappa(x,y) \quad = \quad 1 - \frac{W_1(\delta_x P, \delta_y P)}{d(x,y)}.$$

Ainsi, la chaîne de Markov a une courbure minorée par une constante réelle κ si l'inégalité

$$W_1(\delta_x P, \delta_y P) \quad \leq \quad (1-\kappa)\, d(x,y), \qquad x,y \in E,$$

a lieu. Autrement dit, cette inégalité est reliée à une version à temps discret de la courbure de Wasserstein (modulo le fait qu'elle n'est définie qu'à travers un seul pas de temps effectué par la chaîne) et c'est cette notion de « courbure minorée » qui est réellement utilisée en pratique. Dans le cas des chaînes de Markov finies, une minoration strictement positive de la courbure de Ricci grossière n'est rien d'autre qu'une contraction dans la méthode de couplage de chemins initiée par Bubley et Dyer [47]. On remarque qu'une courbure positive signifie que les lois $\delta_x P$ et $\delta_y P$ sont plus proches que les masses de Dirac δ_x et δ_y au sens de la distance de Wasserstein W_1. En particulier, la courbure de Ricci grossière a, lorsqu'elle est minorée par une constante strictement positive, de nombreuses conséquences en matière d'ergodicité, d'inégalités fonctionnelles et de concentration satisfaites par la loi instantanée de la chaîne de Markov, mais aussi par la mesure invariante. Par exemple, on a le résultat suivant, qui est l'analogue à temps discret du Théorème 2.1.2.

Théorème 2.2.4. *Si la courbure de Ricci grossière est minorée par une constante $\kappa > 0$, alors il existe une unique mesure invariante $\mu \in \mathcal{P}_1(E)$ telle que l'on ait la convergence exponentielle en distance de Wasserstein W_1 suivante : pour toute loi initiale $\nu \in \mathcal{P}_1(E)$,*

$$W_1(\nu P^n, \mu) \quad \leq \quad (1-\kappa)^n\, W_1(\nu, \mu), \qquad n \in \mathbb{N}. \tag{2.2.2}$$

Pour plus de détails sur les propriétés de la courbure de Ricci grossière, nous renvoyons à l'article fondateur [155].

Avec Yann Ollivier, nous nous sommes alors intéressés dans l'article [113] publié en 2010 à la généralisation dans plusieurs directions des bornes non-asymptotiques d'ergodicité mentionnées dans la Section 2.1.1, concernant la moyenne empirique à temps discret de la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$,

$$L_N = \frac{1}{N} \sum_{k=T_0+1}^{T_0+N} \delta_{X_k},$$

où T_0 est le « temps de chauffe », ou *burn-in*, de la chaîne. En particulier, l'idée est de choisir T_0 suffisamment grand, en général de l'ordre du temps de mélange

$$\tau = \inf\left\{ n \in \mathbb{N} : d_{\mathrm{VT}}(\delta_x P^n, \mu) \leq \frac{1}{4} \right\},$$

(le $1/4$ étant anecdotique, l'important étant que la valeur considérée soit strictement inférieure à $1/2$ pour des raisons de (presque) sous-multiplicativité de la distance ci-dessus par rapport à n) une quantité pour laquelle une borne supérieure est souvent accessible en situation de courbure strictement positive. Alors X_{T_0} suit approximativement la loi invariante μ et l'estimation de l'intégrale $\mu(f)$ par la moyenne empirique $L_N(f)$ a lieu sur les N pas suivants de la trajectoire markovienne. Dans cet article, nous avons obtenu sous l'hypothèse de courbure strictement positive des contrôles de l'erreur quadratique moyenne et des inégalités de concentration mixtes gaussiennes/exponentielles menant à des intervalles de confiance utilisables en pratique, et avons étudié l'influence du point de départ de la chaîne sur la variance. Tout en établissant des garanties théoriques sur ces convergences applicables sur de nombreux exemples, aussi bien discrets que continus en temps et/ou en espace (tels que la marche aléatoire paresseuse sur l'hypercube, la dynamique de Glauber à haute température pour le modèle d'Ising, les files d'attente markoviennes comme la file $M/M/\infty$ que nous avons abordée précédemment, les schémas d'Euler et enfin les diffusions sur des variétés riemanniennes à courbure de Ricci strictement positive), l'idée était de conserver un important aspect pratique à travers la modélisation issue des algorithmes MCMC (*Markov Chain Monte Carlo*). Rappelons que la méthode classique de Monte Carlo est de proposer un algorithme efficace de simulation d'intégrale du type $\int_E f \, d\mu$, où μ est une loi de probabilité sur E et $f \in L^1(\mu)$ est une fonction donnée. Essentiellement basée sur la LGN pour les variables ou vecteurs aléatoires i.i.d. de loi μ, elle devient inopérante dès lors que μ est partiellement connue (même à une constante - de normalisation - près, par exemple pour des raisons de haute dimensionnalité ou de très grand cardinal de l'espace E) alors que la méthode MCMC, reposant sur la LGN markovienne, permet de contourner cette difficulté. Bien évidemment, il y a un prix à payer qui est d'introduire des corrélations dans la modélisation en remplaçant l'échantillon i.i.d. par une chaîne de Markov construite de manière à ce que la probabilité cible μ soit son unique mesure invariante (par exemple en utilisant l'algorithme de Metropolis-Hastings). Pour une introduction instructive sur ces méthodes MCMC ainsi que des résultats et problèmes ouverts sur le sujet, on pourra consulter le *survey* de Roberts et Rosenthal [158].

À présent, exposons plus en détail quelques résultats que nous avons établis, dont l'idée principale des preuves repose tout particulièrement sur une extension de la méthode de tensorisation introduite dans l'article [112] décrit auparavant. Tout d'abord, en procédant à une décomposition classique sous la forme biais/variance,

$$\mathbb{E}_{\delta_x}\left[|L_N(f)-\mu(f)|^2\right] \;=\; |\mathbb{E}_{\delta_x}\left[L_N(f)\right]-\mu(f)|^2+\mathbb{E}_{\delta_x}\left[|L_N(f)-\mathbb{E}_{\delta_x}\left[L_N(f)\right]|^2\right],$$

ce dernier terme étant noté dans la suite $\mathrm{Var}_{\delta_x}(L_N(f))$, nous sommes en mesure d'estimer l'erreur quadratique moyenne par des bornes ne requérant pas d'information *a priori* sur la mesure invariante μ, toutes les quantités apparaissant étant très souvent calculables ou contrôlables en pratique.

Théorème 2.2.5. *On suppose que la courbure de Ricci grossière de la chaîne est minorée par une constante $\kappa > 0$. Alors pour toute fonction $f \in \mathrm{Lip}$, on a l'estimation du biais suivante : pour tout $x \in E$ et tout temps $N \in \mathbb{N}^*$,*

$$|\mathbb{E}_{\delta_x}\left[L_N(f)\right]-\mu(f)| \;\leq\; \frac{(1-\kappa)^{T_0+1}\,\|f\|_{\mathrm{Lip}}}{\kappa N}\,W_1(\delta_x,\mu), \tag{2.2.3}$$

et, sous réserve que l'inégalité suivante fasse sens, un contrôle de la variance : pour tout $x \in E$ et tout temps $N \in \mathbb{N}^*$,

$$\mathrm{Var}_{\delta_x}(L_N(f)) \;\leq\; \frac{1}{\kappa N}\left(1+\frac{1}{\kappa N}\,1_{\{T_0\neq 0\}}\right)\frac{\|f\|_{\mathrm{Lip}}^2}{\kappa}\,\sup_{x\in E}\,\sup_{g\in \mathrm{Lip}_1}\,\mathrm{Var}_{\delta_x P}(g). \tag{2.2.4}$$

Remarquons que l'estimation du biais (2.2.3) décroît exponentiellement en le *burn-in* T_0. Par ailleurs, en utilisant l'inégalité triangulaire puis la borne (2.2.2) avec $n = 1$ et $\nu = \delta_x$, on obtient aisément la majoration

$$W_1(\delta_x,\mu) \;\leq\; \frac{1}{\kappa}\,W_1(\delta_x,\delta_x P), \tag{2.2.5}$$

entraînant ainsi dans (2.2.3) une estimation du biais par une quantité en réalité ne dépendant pas de la mesure invariante μ. On aurait pu aussi penser à majorer cette quantité $W_1(\delta_x,\mu)$ par le diamètre de E, ce qui peut être intéressant dans certaines situations, mais rien ne nous dit *a priori* qu'il est fini, l'espace E étant éventuellement non borné. De même, la distance $W_1(\delta_x,\delta_x P)$ peut être bornée dès lors que les sauts de la chaîne sont supposés bornés, c'est-à-dire que la quantité

$$\sigma_\infty \;=\; \frac{1}{2}\sup_{x\in E}\,\mathrm{diam}\,\mathrm{supp}\,\delta_x P,$$

est finie, la notation diam supp désignant le diamètre du support d'une mesure donnée.

Commentons à présent l'inégalité (2.2.4). Tout d'abord, on voit apparaître une quantité de l'ordre du (carré du) diamètre observable selon Gromov [86] à travers la contribution de la variance des fonctions 1-lipschitziennes sous la loi $\delta_x P$. En pratique, on peut s'en débarrasser grâce à l'inégalité triviale

$$\sup_{g\in\mathrm{Lip}_1}\,\mathrm{Var}_{\delta_x P}(g) \;\leq\; \frac{1}{2}\int_E\int_E d(y,z)^2\,\delta_x P(dy)\delta_x P(dz),$$

obtenue en écrivant la variance sous sa forme dupliquée. Par ailleurs, mentionnons que les termes de l'inégalité (2.2.4) ont été regroupés de cette manière pour être en mesure de comparer notre estimation de la variance avec celle issue de la méthode de Monte Carlo classique, qui donnerait $\mathrm{Var}_\mu(L_N(f)) = \mathrm{Var}_\mu(f)/N$ dans le cadre i.i.d. de loi μ. En effet, si l'on sait échantillonner sous la mesure μ, alors l'inégalité suivante,

$$\mathrm{Var}_\mu(f) \quad \leq \quad \frac{\|f\|_{\mathrm{Lip}}^2}{\kappa} \sup_{x\in E} \sup_{g\in \mathrm{Lip}_1} \mathrm{Var}_{\delta_x P}(g),$$

que l'on démontre par récurrence puis passage à la limite ergodique, indique que la décroissance de la variance de la moyenne empirique est en $1/\kappa N$, seul un facteur $1/\kappa$ étant perdu lors du passage de l'indépendant au corrélé. En fait il s'avère que cette perte est assez naturelle au moins dans le cas réversible. En effet, si le *burn-in* T_0 est choisi de l'ordre du temps de mélange τ introduit précédemment, alors contrôler la variance dans le Théorème 2.2.5 revient à contrôler la variance de la moyenne empirique sous μ. En courbure strictement positive, les covariances décroissant exponentiellement avec taux $1-\kappa$, on montre ainsi que

$$\mathrm{Var}_\mu\left(L_N(f)\right) \quad \leq \quad \frac{2}{\kappa N}\,\mathrm{Var}_\mu(f),$$

justifiant la présence de cette constante $1/\kappa$ dans le préfacteur de l'inégalité (2.2.4).

Dans un second temps, une contribution importante de ce travail est de proposer un autre contrôle de la variance affranchi du supremum en espace apparaissant dans (2.2.4). En effet, même si en pratique l'hypothèse souvent utilisée $\sigma_\infty < \infty$ (par exemple pour la concentration que nous allons évoquer dans un instant) entraîne la bornitude de la fonction $x \to \mathrm{Var}_{\delta_x P}(g)$ pour $g \in \mathrm{Lip}_1$, ce supremum peut se révéler problématique dans les passages à la limite en temps continu, comme dans le cas simple d'un modèle d'urne d'Ehrenfest, de la file d'attente markovienne $M/M/\infty$ ou encore pour des solutions d'ÉDS dont la matrice de diffusion σ n'est pas bornée. Le résultat annoncé, englobant ces exemples avec notamment dans le cas continu une matrice de diffusion dont le carré de la norme d'opérateur est autorisé à être, ou majoré par, une fonction lipschitzienne, est le suivant. Pour ne pas surcharger l'inégalité ci-dessous, on suppose ici que le *burn-in* T_0 est nul.

Théorème 2.2.6. *Supposons la courbure de Ricci grossière de la chaîne minorée par une constante $\kappa > 0$. De surcroît, on fait l'hypothèse qu'il existe une fonction $h \in \mathrm{Lip}$ telle que*

$$\frac{1}{\kappa} \sup_{g\in \mathrm{Lip}_1} \mathrm{Var}_{\delta_x P}(g) \quad \leq \quad h(x), \qquad x \in E.$$

Alors pour toute fonction $f \in \mathrm{Lip}$, la variance de la moyenne empirique est cette fois contrôlée de la manière suivante : pour tout $x \in E$ et tout temps $N \in \mathbb{N}^*$,

$$\mathrm{Var}_{\delta_x}(L_N(f)) \quad \leq \quad \frac{\|f\|_{\mathrm{Lip}}^2}{\kappa N} \left(\mu(h) + \frac{\|h\|_{\mathrm{Lip}}}{\kappa N} W_1(\delta_x, \mu) \right).$$

Lorsque h est une fonction constante, on retrouve le résultat du Théorème 2.2.5. Sinon, le supremum en l'état initial est remplacé par la moyenne de h sous μ et la borne supérieure se comporte pour N grand en $\|f\|_{\mathrm{Lip}}^2 \mu(h)/\kappa N$, avec une correction d'ordre $1/(\kappa N)^2$ dépendant de l'état initial x. De nouveau, les termes faisant intervenir la mesure invariante μ peuvent être contrôlés grâce à l'inégalité (2.2.5), l'intégrale $\mu(h)$ de la fonction lipschitzienne h étant alors majorée de la manière suivante :

$$\mu(h) \quad \leq \quad h(z) + \frac{\|h\|_{\mathrm{Lip}}}{\kappa} W_1(\delta_z, \delta_z P), \quad z \in E,$$

le point z pouvant être choisi ou non égal à l'état initial x de la chaîne.

Avant de nous intéresser aux bornes de déviation, mentionnons que tous nos résultats s'adaptent immédiatement au cas d'un état initial X_0 aléatoire. En effet, un terme supplémentaire apparaît alors dans l'erreur quadratique moyenne, *i.e.*, si $\nu \in \mathcal{P}_1(E)$ désigne la loi initiale de la chaîne, alors

$$\begin{aligned} \mathbb{E}_\nu \left[|L_N(f) - \mu(f)|^2 \right] \quad &= \quad |\mathbb{E}_\nu \left[L_N(f) \right] - \mu(f)|^2 + \int_E \mathrm{Var}_{\delta_x}(L_N(f)) \, \nu(dx) \\ &\quad + \mathrm{Var}_\nu(\mathbb{E}[L_N(f) \mid X_0 = \cdot]). \end{aligned}$$

La fonction $P^k f$ étant lipschitzienne de constante $(1-\kappa)^k \|f\|_{\mathrm{Lip}}$, on montre alors que

$$\mathrm{Var}_\nu(\mathbb{E}[L_N(f) \mid X_0 = \cdot]) \quad \leq \quad \frac{(1-\kappa)^{2(T_0+1)}}{2\kappa^2 N^2} \int_E \int_E d(x,y)^2 \, \nu(dx)\nu(dy),$$

et le tour est joué. Notons également que la quantité $W_1(\delta_x, \mu)$ dans l'inégalité sur le biais (2.2.3) peut être remplacée par la distance $W_1(\nu, \mu)$, permettant de réduire considérablement l'estimation proposée si d'aventure la loi initiale ν est proche de la loi invariante μ.

Intéressons-nous désormais aux inégalités de concentration mixtes gaussiennes/exponentielles que nous avons obtenues pour des fonctions lipschitziennes, et qui peuvent être utilisées pour en déduire des intervalles de confiance utiles en pratique. Quitte à légèrement modifier le niveau de déviation en lui incluant l'information sur le biais donnée dans le Théorème 2.2.5, étudier la concentration de la moyenne empirique $L_N(f)$ autour de la moyenne $\mu(f)$ ou de son espérance $\mathbb{E}_{\delta_x}[L_N(f)]$ revient au même.

Théorème 2.2.7. *On suppose que la courbure de Ricci grossière de la chaîne est minorée par une constante $\kappa > 0$ et que les sauts de la chaîne sont bornés, i.e., $\sigma_\infty < \infty$. De plus, on fait l'hypothèse qu'il existe une fonction $h \in \mathrm{Lip}$ telle que*

$$\frac{1}{\kappa} \sup_{g \in \mathrm{Lip}_1} \mathrm{Var}_{\delta_x P}(g) \quad \leq \quad h(x), \quad x \in E.$$

Pour tout $N \in \mathbb{N}^*$, on note V_N la fonction qui à tout $x \in E$ associe la quantité

$$V_N(x) \quad = \quad \frac{1}{\kappa N} \left(1 + \frac{T_0}{N} \right) \mu(h) + \frac{\|h\|_{\mathrm{Lip}} W_1(\delta_x, \mu)}{\kappa^2 N^2}.$$

Alors pour toute fonction $f \in \mathrm{Lip}_1$, tout $x \in E$ et tout temps $N \in \mathbb{N}^*$, on a l'inégalité de concentration mixte gaussienne/exponentielle suivante :

$$\mathbb{P}_{\delta_x}\left(|L_N(f) - \mathbb{E}_{\delta_x}[L_N(f)]| \geq r \right) \quad \leq \quad \begin{cases} 2 \exp\left(-\frac{r^2}{16 V_N(x)} \right) & si \quad r \in (0, r_{\max}), \\ 2 \exp\left(-\frac{\kappa N r}{4 \max\{2\|h\|_{\mathrm{Lip}}, 3\sigma_\infty\}} \right) & si \quad r \geq r_{\max}, \end{cases}$$

où la transition continue a lieu en $r_{\max} = 4V_N(x)\kappa N/\max\{2\|h\|_{\mathrm{Lip}}, 3\sigma_\infty\}$.

Ci-dessus, la fonction V_N joue le rôle d'une estimation de la variance comme proposée dans le Théorème 2.2.6. De plus, l'hypothèse de sauts bornés est nécessaire pour la même raison que celle évoquée dans le Théorème 2.2.1, la différence essentielle avec cette inégalité de concentration poissonnienne résidant dans l'aspect « variance bornée ou non ». Enfin, les mêmes commentaires à propos du TCL markovien s'appliquent ici aussi.

Pour conclure la présentation du travail [113], regardons ce que nos résultats garantissent sur deux exemples emblématiques : la marche aléatoire paresseuse sur l'hypercube et une solution d'ÉDS dont la matrice de diffusion n'est pas bornée. Commençons par la situation discrète. Soit $E = \{0,1\}^d$ l'hypercube de dimension d muni de la distance de Hamming d_H (comptant le nombre de coordonnées différentes entre deux points de $\{0,1\}^d$) et de la mesure uniforme, $\mu(\{x\}) = 2^{-d}$ pour tout $x \in \{0,1\}^d$. La chaîne de Markov que l'on considère consiste à chaque étape à choisir à Pile ou Face si l'on reste en place ou pas et, dans ce dernier cas, on remplace une seule coordonnée choisie uniformément au hasard : pour tout $x \in \{0,1\}^d$, on a $\delta_x P(\{x\}) = 1/2$ et $\delta_x P(\{y\}) = 1/2d$ si y est un voisin de x, *i.e.*, $d_H(x,y) = 1$. Alors un calcul élémentaire entraîne que la courbure de Ricci grossière est constante et vaut $\kappa = 1/d$. Les quantités d'intérêt sont calculées ou estimées dans [155] :

$$\sup_{x\in E}\sup_{g\in\mathrm{Lip}_1} \mathrm{Var}_{\delta_x P}(g) \;\leq\; \frac{1}{2}, \quad W_1(\delta_x,\mu) \;=\; \frac{d}{2} \quad \text{et} \quad \sigma_\infty \;=\; 1.$$

Ainsi, on obtient du Théorème 2.2.6 le résultat suivant.

Théorème 2.2.8. *Pour toute fonction* $f \in \mathrm{Lip}$ *sur l'hypercube* $\{0,1\}^d$*, on a l'estimation du biais*

$$\begin{aligned}
|\mathbb{E}_{\delta_x}[L_N(f)] - \mu(f)| &\leq \frac{d^2}{2N}\left(1-\frac{1}{d}\right)^{T_0+1}\|f\|_{\mathrm{Lip}} \\
&\leq \frac{d^2}{2N}\exp\left(-\frac{T_0}{d}\right)\|f\|_{\mathrm{Lip}},
\end{aligned}$$

tandis que pour la variance, dans le cas d'un burn-in T_0 *non nul,*

$$\mathrm{Var}_{\delta_x}(L_N(f)) \;\leq\; \frac{d^2}{2N}(1+d/N)\|f\|_{\mathrm{Lip}}^2.$$

Ainsi, choisir $T_0 = \mathcal{O}(d\log d)$ permet d'assurer un biais petit, en accord avec l'estimation du temps de mélange τ connue pour être du même ordre de grandeur, cf. [72]. Par ailleurs, on peut obtenir une inégalité de concentration par le Théorème 2.2.7, avec la fonction h constante égale à $d/2$ (auquel cas bien sûr $\|h\|_{\mathrm{Lip}} = 0$). Pour le régime gaussien, on a alors la borne $2\exp(-r^2N/8d^2)$ pour $r \in]0, 2d/3[$. Notons que le résultat de Lezaud [136] fait apparaître ce régime gaussien pour r assez petit, avec quelques différences néanmoins dont la principale est que son préfacteur devant l'exponentielle, traduisant la régularité de la loi initiale ν par rapport à la mesure invariante requise par sa méthode, vaut $2^{d/2}$ lorsque ν est une masse de Dirac, une quantité explosant exponentiellement en grande dimension. Ainsi, notre inégalité de concentration améliore son résultat d'un facteur $2^{d/2}$, dû au fait que la « densité » de la loi initiale ne joue aucun rôle prépondérant dans notre approche. Notons que le même type de phénomène apparaît dans les travaux de Rudolf [159, 160] sur l'estimation de la variance, d'où une attention particulière portée sur l'influence du choix de la distribution initiale et du *burn-in* T_0 dans ses résultats.

À présent, passons à l'exemple d'une diffusion à valeurs dans $\mathbb{R}^d$ dont la matrice de diffusion σ n'est pas bornée. Étant donnée une matrice A, on note sa norme d'opérateur $\|A\|_{\mathrm{op}} = \sup\left\{|Ax|/|x| : x \in \mathbb{R}^d \backslash \{0\}\right\}$. On considère le processus $(X_t)_{t\geq 0}$ solution de l'ÉDS suivante sur $\mathbb{R}^d$:

$$dX_t \quad = \quad \sqrt{2}\,\sigma(X_t)\,dB_t + b(X_t)\,dt,$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien standard sur $\mathbb{R}^d$ et la matrice de diffusion σ et le champ de vecteurs b sont supposés réguliers lipschitziens. De surcroît, on suppose l'hypothèse de stabilité suivante, assimilée à une force de rappel pour la diffusion : il existe une constante $\kappa > 0$ telle que

$$\|\sigma(x) - \sigma(y)\|_{\mathrm{HS}}^2 + (x-y)\cdot(b(x) - b(y)) \;\leq\; -\kappa\,|x-y|^2, \quad x, y \in \mathbb{R}^d.$$

Alors le processus admet une unique mesure de probabilité invariante μ vers laquelle il converge en loi exponentiellement vite, cf. [73]. Ainsi, en utilisant par exemple le Théorème 2.2.7 pour le schéma d'Euler associé à la diffusion puis en passant soigneusement à la limite (en particulier, κ joue le rôle de la courbure de Ricci grossière minorée et la borne σ_∞ sur la taille des sauts disparaît), on obtient le résultat suivant.

Théorème 2.2.9. *Sous les hypothèses mentionnées ci-dessus, on suppose de surcroît qu'il existe une fonction $h \in \mathrm{Lip}$ telle que l'on ait l'inégalité ponctuelle,*

$$\frac{2\,\|\sigma\|_{op}^2}{\kappa} \quad \leq \quad h.$$

Pour tout $t > 0$, notons V_t la fonction qui à tout $x \in E$ associe

$$V_t(x) \quad = \quad \frac{\mu(h)}{\kappa t} + \frac{\|h\|_{\mathrm{Lip}} W_1(\delta_x, \mu)}{\kappa^2 t^2}.$$

Alors pour toute fonction $f \in \mathrm{Lip}_1$, tout $x \in E$ et tout temps $t > 0$, on a l'inégalité de concentration mixte gaussienne/exponentielle suivante pour la moyenne empirique à temps continu $L_t(f)$:

$$\mathbb{P}_{\delta_x}\left(|L_t(f) - \mathbb{E}_{\delta_x}[L_t(f)]| \geq r\right) \quad \leq \quad \begin{cases} 2\,\exp\left(-\frac{r^2}{16 V_t(x)}\right) & si \quad r \in (0, r_{\max}), \\ 2\,\exp\left(-\frac{\kappa t r}{8\,\|h\|_{\mathrm{Lip}}}\right) & si \quad r \geq r_{\max}, \end{cases}$$

où la transition continue a lieu en $r_{\max} = 2V_t(x)\kappa t/\|h\|_{\mathrm{Lip}}$.

Par exemple, cette inégalité de concentration peut être utilisée lorsque la matrice de diffusion se comporte comme $\sqrt{|x|}$ à l'infini, des modèles pour lesquels de tels résultats n'étaient manifestement pas disponibles jusqu'à présent. C'est donc avec cette bonne nouvelle que nous achevons la présentation de l'article [113].

2.2.3 Concentration de la mesure invariante et inégalités fonctionnelles

La concentration de la moyenne empirique autour de l'équilibre ayant été étudiée dans les Sections 2.2.1 et 2.2.2, nous allons nous focaliser dans cette partie sur le phénomène de concentration des mesures invariantes de processus de Markov, c'est-à-dire étudier le comportement des probabilités de déviation de la moyenne

$$\mu\left(\{x \in E : |f(x) - \mu(f)| > r\}\right), \quad r > 0,$$

lorsque μ est la mesure invariante d'un processus de Markov ergodique, la fonction f appartenant à une certaine classe de fonctions. On peut aborder cette question en prenant en compte l'aspect dynamique permettant d'exploiter pleinement les informations relatives à la loi instantanée du processus, propriétés ayant vocation en temps long à se propager à la mesure invariante : c'est l'approche par semigroupe développée avec succès par Ledoux pour établir les inégalités fonctionnelles désirées sous la mesure invariante (telles que les inégalités de Poincaré ou de Sobolev logarithmique), entraînant des estimées de concentration robustes en grande dimension. Dans ce cadre markovien, la distance naturellement considérée est associée au carré du champ Γ. Cette distance ne coïncidant pas nécessairement avec la distance originelle sur E (par exemple pour des processus de sauts purs dont la dynamique n'est pas bornée ou encore des diffusions dont la matrice de diffusion n'est pas bornée), ceci donne lieu à une notion de « Lipschitzianité » associée au processus, c'est-à-dire que dans ce contexte une fonction $f : E \to \mathbb{R}$ appartient à l'espace Γ-Lip si elle est Γ-lipschitzienne au sens où $\Gamma(f, f) \in L^\infty(\mu)$. Or il est de plus en plus souvent requis dans les applications de sortir du cadre lipschitzien et c'est pourquoi dans l'article [89], écrit en collaboration avec Arnaud Guillin et publié en 2013, nous affaiblissons dans le cadre réversible cette hypothèse de « Γ-Lipschitzianité » en autorisant des fonctions f ayant par exemple une croissance polynomiale à l'infini, cette étude prenant tout son sens sur un espace E non borné, qu'il soit discret ou continu. À notre connaissance, aucun résultat de concentration de ce type n'avait été démontré jusque là par une telle approche dynamique. Pour ce faire, notre idée repose sur l'introduction dans les inégalités fonctionnelles considérées d'une condition de Lyapunov.

Définition 2.2.10. *Soient $a \geq 0$ et $b > 0$ deux constantes quelconques. Une fonction $f : E \to \mathbb{R}$ suffisamment régulière appartient à l'ensemble $\mathcal{L}(a, b)$ s'il existe une fonction test suffisamment régulière $h : E \to]0, \infty[$ telle que l'on ait l'inégalité μ-p.p. suivante :*

$$\Gamma(f, f) \quad \leq \quad -a\frac{Lh}{h} + b. \tag{2.2.6}$$

Différente des critères de Lyapunov à la Meyn-Tweedie [145] classiquement utilisés pour établir des résultats d'ergodicité en variation totale (à poids ou non), cette condition de Lyapunov dépend de la fonction f et donc n'est pas intrinsèque à la dynamique. Le fait que la constante a puisse être strictement positive nous permet de sortir du cas Γ-lipschitzien, c'est-à-dire que le carré du champ associé $\Gamma(f, f)$ n'est plus forcément borné mais dépend à présent de la dynamique sous-jacente : moralement, plus la quantité $-Lh/h$ est grande à l'infini, plus la classe de fonctions f considérées dans cette étude est importante. En pratique, on trouve une fonction test h telle que le ratio $-Lh/h$ soit le plus grand possible (ceci se fait au cas par cas, en fonction des attributs de la dynamique) et ensuite seulement l'on s'intéresse aux fonctions f vérifiant $\Gamma(f,f) \approx -a\, Lh/h$ à l'infini.

La raison principale justifiant la forme particulière d'une telle condition de Lyapunov réside dans ses bonnes propriétés par rapport à l'intégration par partie sous la mesure invariante. Ceci donne lieu au lemme clé de notre travail, qui est bien connu des gens ayant l'habitude de « grand dévier ». On rappelle que dans toute cette partie, l'espace E est indifféremment discret ou continu et le processus de Markov considéré sur cet espace est supposé réversible.

Lemme 2.2.11. *Pour toutes fonctions réelles g, h suffisamment régulières sur E avec de surcroît $h > 0$, on a l'inégalité suivante :*

$$\int_E \frac{-Lh}{h}\, g^2\, d\mu \quad \leq \quad \int_E \Gamma(g,g)\, d\mu.$$

Preuve. La démonstration étant brève dans le cas diffusif, concentrons-nous sur cette situation, le cas discret étant plus difficile à établir car la règle de Leibniz n'est pas vérifiée sur les espaces discrets. L'opérateur carré du champ satisfaisant la règle de Leibniz en chacun de ses arguments, on a pour toute fonction réelle g suffisamment régulière sur E,

$$\begin{aligned}
\int_E \frac{-Lh}{h}\, g^2\, d\mu &= \int_E \Gamma\left(h, \frac{g^2}{h}\right) d\mu \\
&= \int_E \frac{2\, g\, \Gamma(h,g)}{h}\, d\mu - \int_E \frac{g^2\, \Gamma(h,h)}{h^2}\, d\mu \\
&\leq \int_E \Gamma(g,g)\, d\mu,
\end{aligned}$$

où l'inégalité de Cauchy-Schwarz pour le carré du champ Γ est utilisée pour obtenir la dernière ligne. ■

L'explication probabiliste de cette inégalité est due au PGD sous-jacent, au sens où l'intégrale $\int_E \Gamma(g,g)\, d\mu$ n'est rien d'autre que l'entropie de Donsker-Varadhan de la mesure de probabilité ν de densité par rapport à la mesure invariante μ proportionnelle à g^2, c'est-à-dire la fonction de taux dans le PGD pour la moyenne empirique L_t en temps long, qui est justement donnée par le supremum de l'intégrale $\int_E -Lh/h\, d\nu$ sur une classe de fonctions h adéquate. Nous renvoyons par exemple à l'article de Wu [183] dans lequel cette propriété est établie en toute généralité pour les processus de Markov réversibles, incluant les processus de sauts purs.

En particulier, en prenant $g = 1$ dans le Lemme 2.2.11, on en déduit deux choses : premièrement, il est vain de tenter de trouver une fonction test h telle que le ratio $-Lh/h$ soit toujours positif. Ainsi, le paramètre b dans la condition de Lyapunov est présent pour compenser le manque de positivité dans certaines régions bien localisées en espace. Deuxièmement, si $f \in \mathcal{L}(a,b)$, alors

$$\int_E \Gamma(f,f)\, d\mu \quad \leq \quad b. \tag{2.2.7}$$

Autrement dit, la fonction non Γ-lipschitzienne f satisfaisant la condition de Lyapunov (2.2.6) est Γ-lipschitzienne en moyenne.

À présent, voyons comment le Lemme 2.2.11 peut être utilisé à bon escient dans des calculs faisant intervenir la transformée de Laplace.

Lemme 2.2.12. *Soit $f \in \mathcal{L}(a,b)$. Alors on a l'inégalité suivante :*

$$\int_E \Gamma\left(e^{\lambda f/2}, e^{\lambda f/2}\right) d\mu \quad \leq \quad \frac{\lambda^2 b}{4-\lambda^2 a} \int_E e^{\lambda f}\, d\mu, \quad 0 < \lambda < \frac{2}{\sqrt{a}}.$$

Preuve. La démonstration étant rapide, proposons-la dans le cadre discret car elle illustre bien comment la réversibilité permet de résoudre le problème de l'absence de la règle de Leibniz. Le cas diffusif est encore plus bref. En utilisant la réversibilité, on a

$$\begin{aligned}
\int_E \Gamma\left(e^{\lambda f/2}, e^{\lambda f/2}\right) d\mu &= \frac{1}{2}\int_E \int_E \left(e^{\lambda f(x)/2} - e^{\lambda f(y)/2}\right)^2 Q(x,dy)\,\mu(dx) \\
&= \int \int_{f(x)>f(y)} \left(1 - e^{-\lambda(f(x)-f(y))/2}\right)^2 e^{\lambda f(x)}\, Q(x,dy)\,\mu(dx) \\
&\leq \frac{\lambda^2}{4} \int_E \Gamma(f,f)\, e^{\lambda f}\, d\mu \\
&\leq \frac{\lambda^2}{4} \int_E \left(-a\,\frac{Lh}{h} + b\right) e^{\lambda f}\, d\mu \\
&\leq \frac{\lambda^2}{4} \left(a \int_E \Gamma\left(e^{\lambda f/2}, e^{\lambda f/2}\right) d\mu + b \int_E e^{\lambda f}\, d\mu\right),
\end{aligned}$$

où l'on a utilisé l'inégalité triviale $1 - e^{-x} \leq x$, $x \geq 0$ et le Lemme 2.2.11 pour obtenir les troisième et dernière lignes, respectivement. On en déduit alors la conclusion désirée en réarrangeant les termes. ∎

La présence dans l'inégalité ci-dessus de l'intégrale $\int_E \Gamma\left(e^{\lambda f/2}, e^{\lambda f/2}\right) d\mu$ et de la transformée de Laplace $\int_E e^{\lambda f}\, d\mu$ de la fonction f n'est pas innocente : il s'agit des quantités auxquelles on a affaire en général lorsque l'on désire montrer un résultat de concentration à partir d'inégalités fonctionnelles telles que l'inégalité de Poincaré, auquel cas la méthode dite d'Aida-Stroock [2] est utilisée, ou encore l'inégalité entropique et ses variantes, les différentes inégalités de Sobolev logarithmique modifiées, pour lesquelles la méthode de Herbst, reposant sur une inégalité différentielle en λ satisfaite par la transformée de Laplace, donne de bons résultats, cf. par exemple [28, 29, 31].

Tournons-nous maintenant vers les résultats de concentration que nous avons obtenus pour des fonctions satisfaisant la condition de Lyapunov (2.2.6). Comme dans les travaux précédemment cités, le point de départ est de supposer que la mesure invariante et réversible μ satisfait une certaine inégalité fonctionnelle. Notons que dans le cadre diffusif cette hypothèse n'est pas si gourmande au sens où s'il existe une fonction $f \in \mathcal{L}(a,b)$ telle que son carré du champ $\Gamma(f,f)$ soit minoré à l'infini, et c'est le cas dans la majorité des exemples intéressants, alors l'inégalité de Poincaré est satisfaite. Pour s'en convaincre, le lecteur intéressé pourra consulter les articles de Cattiaux, Guillin et cie [11, 54, 55, 56] faisant le lien entre les inégalités fonctionnelles et les diverses conditions de Lyapunov existant dans la littérature. Ainsi, les inégalités fonctionnelles que nous considérons dans notre article sont l'inégalité entropique, la plus forte, et les inégalités de Beckner, une famille d'inégalités moins standard interpolant entre la première et l'inégalité de Poincaré, cf. [21]. Dans la suite, nous n'énonçons par souci de concision que le résultat de concentration mixte gaussienne/exponentielle établi sous l'hypothèse de l'inégalité entropique, le régime obtenu par les inégalités de Beckner étant quelque peu similaire mais avec une dépendance sous-optimale en certains paramètres (la différence principale réside dans la classe des fonctions pour laquelle la concentration est valable).

Théorème 2.2.13. *On suppose que la dynamique réversible satisfait l'inégalité entropique de constante optimale ρ_0. Alors pour toute fonction $f \in \mathcal{L}(a,b)$, on a l'inégalité de concentration mixte gaussienne/exponentielle*

$$\mu\left(\{x \in E : |f(x) - \mu(f)| \geq r\}\right) \;\leq\; \begin{cases} 2\,\exp\left(-\frac{3\rho_0 r^2}{16b}\right) & si \quad r \in (0, r_{\max}), \\ 2\,\exp\left(-\frac{r}{2\sqrt{a}}\right) & si \quad r \geq r_{\max}, \end{cases}$$

où la transition continue a lieu en $r_{\max} = 8b/3\rho_0\sqrt{a}$.

Preuve. Dans l'article [89], nous proposons deux démonstrations très simples donnant lieu à ce même résultat (aux constantes numériques près). Celle que nous n'exposons pas présentement est inspirée de la méthode à la Otto-Villani [156], démontrant dans le cas diffusif que l'inégalité de Sobolev logarithmique implique l'inégalité de Talagrand (une inégalité de transport quadratique), qui entraîne à son tour par l'argument de Marton la concentration gaussienne. Dans celle que nous présentons maintenant, l'idée est d'exploiter le Lemme 2.2.12 dans la méthode de Herbst que nous détaillons (enfin !) ci-dessous après l'avoir évoquée à maintes reprises. De nouveau, établissons la preuve dans le cadre disret. On a pour tout $\lambda > 0$,

$$\begin{aligned} \frac{\partial}{\partial \lambda}\left(\frac{1}{\lambda}\log \int_E e^{\lambda f}\,d\mu\right) &= \frac{1}{\lambda^2 \int_E e^{\lambda f}\,d\mu}\,\mathrm{Ent}_\mu\left(e^{\lambda f}\right) \\ &\leq \frac{1}{\rho_0\,\lambda^2 \int_E e^{\lambda f}\,d\mu} \int_E \Gamma\left(\lambda f, e^{\lambda f}\right)\,d\mu \\ &\leq \frac{1}{\rho_0 \int_E e^{\lambda f}\,d\mu} \int_E \Gamma(f,f)\,e^{\lambda f}\,d\mu \\ &\leq \frac{1}{\rho_0 \int_E e^{\lambda f}\,d\mu} \int_E \left(-a\,\frac{Lh}{h} + b\right) e^{\lambda f}\,d\mu \\ &\leq \frac{1}{\rho_0 \int_E e^{\lambda f}\,d\mu} \left(a \int_E \Gamma\left(e^{\lambda f}, e^{\lambda f}\right)\,d\mu + b \int_E e^{\lambda f}\,d\mu\right). \end{aligned}$$

Ci-dessus, on a utilisé la réversibilité comme dans la preuve du Lemme 2.2.12 pour obtenir la troisième ligne et ensuite le Lemme 2.2.11 pour établir la dernière inégalité. Si maintenant l'on restreint les valeurs du paramètre λ à l'intervalle $(0, 1/\sqrt{a})$, alors en appliquant le Lemme 2.2.12 puis en intégrant, il vient l'estimation suivante sur la transformée de Laplace :

$$\int_E e^{\lambda f}\,d\mu \;\leq\; \exp\left(\lambda\,\mu(f) + \frac{4b\lambda^2}{3\rho_0}\right), \quad 0 < \lambda < \frac{1}{\sqrt{a}}.$$

Enfin, l'inégalité de Chebyshev suivie d'une optimisation en λ donne le résultat annoncé. ■

Dans le cas usuel des fonctions $f \in \Gamma$-Lip, c'est-à-dire que $a = 0$ dans la condition de Lyapunov (2.2.6), il s'agit du résultat classique de concentration gaussienne. Même s'il est inutile de le mentionner, précisons que notre résultat ne peut pas être obtenu en toute généralité en appliquant naïvement cette concentration gaussienne pour une fonction Γ-lipschitzienne puis en changeant de fonction par une transformation pertinente, éventuellement monotone. Bien que ce procédé grossier, s'il peut s'appliquer dans certaines situations, ne permette pas en général de récupérer les bonnes dépendances en les paramètres importants du modèle étudié, en particulier lorsque la moyenne de la fonction n'est pas explicitement calculable, il indique néanmoins le régime de concentration auquel on devrait s'attendre pour les grands niveaux de déviation.

Ainsi, lorsque l'on sort du cadre Γ-lipschitzien, la concentration n'a plus de raison d'être gaussienne pour les grands niveaux de déviation (et indépendante de la dimension), la croissance à l'infini des fonctions considérées n'étant alors plus contrôlée de la même manière que dans le cas Γ-lipschitzien. Le régime exponentiel est optimal pour les fonctions f pour lesquelles on a le comportement asymptotique $\Gamma(f,f) \approx -a\, Lh/h$, la fonction test h étant choisie de telle sorte que le ratio $-Lh/h$ soit le plus grand possible. Néanmoins, ce n'est qu'un exemple de vitesse de concentration que nous obtenons, le caractère exponentiel pour les grands niveaux de déviation pouvant être renforcé pour des fonctions g pour lesquelles on a $\Gamma(g,g) \ll \Gamma(f,f)$ à l'infini. En revanche, comme dans le cas des moyennes empiriques évoquées précédemment, le régime gaussien est incontournable pour les petits niveaux de déviation, le TCL étant caché derrière cette affaire. Notons également que l'ordre de grandeur du régime gaussien est optimal par rapport aux paramètres d'intérêt du problème, à savoir ρ_0 et b, la variance de f étant contrôlée de la manière suivante (utiliser la borne (2.2.7), l'inégalité de Poincaré enfin l'inégalité $\lambda_1 \geq \rho_0/2$) :

$$\mathrm{Var}_\mu(f) \quad \leq \quad \frac{2b}{\rho_0}.$$

Le Théorème 2.2.13 peut être appliqué à de nombreuses situations aussi bien discrètes que continues. Par exemple dans le cas discret, il est efficace pour des processus de naissance et de mort à dynamique non bornée comme la file d'attente $M/M/\infty$, et leur pendant spatial avec interactions, *via* certaines dynamiques de Glauber à haute température sur des espaces de configuration non bornés du type $\mathbb{N}^\Lambda$ avec Λ un sous-ensemble borné de $\mathbb{Z}^d$, satisfaisant l'inégalité entropique. Pour avoir une idée plus précise à ce sujet, on pourra consulter les articles [49, 50, 62, 70, 71, 182] ainsi que le travail que nous avons effectué avec Djalil Chafaï [58] sur la notion d'entrelacement, faisant ainsi le lien avec le prochain chapitre.

Dans le cas continu, ce phénomène de concentration mixte gaussienne/exponentielle a lieu typiquement lorsque μ est la mesure gaussienne standard γ sur $\mathbb{R}^d$, pour laquelle $\rho_0 = 2$, et f est la forme quadratique $f(x) = x^T A\, x$, où A est une matrice symétrique définie positive, auquel cas en choisissant judicieusement la fonction test h dans la condition de Lyapunov (2.2.6), on retrouve la célèbre inégalité de Hanson-Wright [91] avec $a = \mathcal{O}(\|A\|_{\mathrm{op}}^2)$ et $b = \mathcal{O}(\|A\|_{\mathrm{HS}}^2)$.

Un autre contexte diffusif pour lequel le Théorème 2.2.13 est pertinent, est le suivant. Le carré du champ des processus de Kolmogorov étant donné par $\Gamma(f,f) = |\nabla f|^2$, certaines de ces dynamiques ne satisfont pas l'inégalité entropique (réécrite généralement dans ce cadre diffusif sous la forme de l'inégalité de Sobolev logarithmique), mais plutôt une inégalité de Sobolev logarithmique avec un poids (non borné), disons σ^2 :

$$\mathrm{Ent}_\mu(f^2) \;\leq\; C \int_{\mathbb{R}^d} \sigma^2 \, |\nabla f|^2 \, d\mu.$$

Par exemple si le potentiel V est radial (*i.e.*, invariant par rotation) de la forme $V(x) = |x|^\alpha$ sur $\mathbb{R}^d$ avec $1 < \alpha < 2$, on sait que la mesure invariante μ (dite de Subbotin [169] car sa densité de Lebesgue est proportionnelle à $\exp(-|x|^\alpha)$; par souci de normalisation, on appellera dans la suite loi de Subbotin la mesure dont la densité de Lebesgue est plutôt proportionnelle à $\exp(-|x|^\alpha/\alpha))$, associée à la dynamique $L = \Delta - \nabla V \cdot \nabla$, satisfait une inégalité de Poincaré (c'est le cas si et seulement si $\alpha \geq 1$) mais pas l'inégalité de Sobolev logarithmique (ce serait le cas si et seulement si $\alpha \geq 2$). En revanche elle satisfait l'inégalité de Sobolev logarithmique avec poids $\sigma(x)^2 = (1 + |x|^{2-\alpha})$, cf. [55]. Tout en conservant la même mesure invariante et réversible μ, l'idée est alors de modifier la dynamique judicieusement afin de considérer cette inégalité à poids comme une vraie inégalité de Sobolev logarithmique pour un nouveau carré du champ Γ^σ, qui n'est plus celui d'un processus de Kolmogorov car il comporte dans son expression la matrice de diffusion $\sigma\, I$. Dans la situation présente, ce carré du champ est $\Gamma^\sigma(f,f) = \sigma^2\, |\nabla f|^2$ et il est associé à l'opérateur

$$L^\sigma f \;=\; \sigma^2 \, \Delta f + \big(\nabla(\sigma^2) - \sigma^2 \, \nabla V\big) \cdot \nabla f.$$

Ainsi, on est en mesure d'appliquer le Théorème 2.2.13 pour des fonctions f satisfaisant la condition de Lyapunov (2.2.6),

$$\Gamma^\sigma(f,f) \;\leq\; -a \, \frac{L^\sigma h}{h} + b, \tag{2.2.8}$$

dans laquelle on a remplacé L et Γ par L^σ et Γ^σ, respectivement. Un choix convenable de la fonction test h permet alors d'obtenir une inégalité de concentration mixte gaussienne/exponentielle pour des fonctions du type $f(x) = |x|^\alpha$. Ce résultat est attendu au vu du travail de Latala et Oleszkiewicz [124], démontrant que la concentration pour des fonctions lipschitziennes $f \in \mathrm{Lip}$ est d'ordre $\exp(-c\, r^\alpha)$.

Le même raisonnement peut s'appliquer pour des processus de Kolmogorov avec un potentiel non convexe tel que celui associé à la loi de Cauchy généralisée, *i.e.*, $V(x) = \beta \log(1 + |x|^2)$. Dans cet exemple, μ est une mesure à queue lourde de densité de Lebesgue proportionnelle à $(1 + |x|^2)^{-\beta}$, $\beta > d/2$. Dans le même article [55], les auteurs couvrent cet exemple en raffinant l'inégalité de Sobolev logarithmique à poids initialement démontrée par Bobkov et Ledoux dans [30]. Plus précisément, ils montrent une inégalité avec le poids $(1 + |x|^2)\, \log(e + |x|^2)$, l'inégalité de Bobkov et Ledoux exhibant un poids plus grand, $\sigma(x)^2 = (1 + |x|^2)^2$ (avec la restriction supplémentaire $\beta \geq (d+1)/2$ et $\beta > 1$, néanmoins la constante devant le terme d'énergie ne dépend pas de la dimension). En procédant de la même manière que précédemment, une fonction test polynomiale h dans la condition de Lyapunov (2.2.8) permet alors d'établir une concentration mixte gaussienne/exponentielle pour des fonctions du type $f(x) = \log|x|$: il s'agit-là d'une illustration quelque peu originale du phénomène de concentration des mesures à queue lourde, cf. [30, 52].

Pour achever la présentation de l'article [89], revenons sur la question des fonctions lipschitziennes. On a vu que la distance intrinsèquement liée au carré du champ n'était pas forcément équivalente à la distance d munissant initialement l'espace métrique (E, d), comme par exemple pour des processus de diffusion dont la matrice de diffusion σ dans le générateur L n'est pas bornée. L'idée est alors de fixer les fonctions lipschitziennes considérées pour la concentration du Théorème 2.2.13 et de faire porter la condition de Lyapunov sur la matrice de diffusion. Plus précisément, soit $f \in \mathrm{Lip}_1$ une fonction 1-lipschitzienne sur $\mathbb{R}^d$. Si la matrice de diffusion σ satisfait la condition de Lyapunov

$$\|\sigma\|_{\mathrm{op}}^2 \quad \leq \quad -a\,\frac{Lh}{h} + b,$$

alors $f \in \mathcal{L}(a, b)$ et l'inégalité de concentration mixte gaussienne/exponentielle du Théorème 2.2.13 s'applique à cette fonction lipschitzienne f, sous réserve que l'inégalité entropique soit satisfaite. Cette condition de Lyapunov est intéressante au sens où elle est intrinsèque à la dynamique. En pratique, la mesure de probabilité μ et la classe de fonctions lipschitziennes Lip étant données, il s'agit alors de trouver un bon processus de Markov réversible par rapport à μ, de générateur L dont la matrice de diffusion σ satisfait à l'infini $\|\sigma\|_{\mathrm{op}}^2 \approx -a\, Lh/h$ pour une fonction test h bien choisie. Mentionnons que ce type d'hypothèse sur la croissance de la matrice de diffusion apparaît dans le travail d'Ollivier [155] sur la concentration en courbure de Ricci grossière strictement positive (et est repris pour les moyennes empiriques dans notre article commun [113] décrit dans la Section 2.2.2). Si l'on traduit dans notre langage diffusif le résultat de concentration qu'il obtient pour les fonctions lipschitziennes, on observe qu'il est comparable à celui émanant du Théorème 2.2.13. En effet, les deux hypothèses principales qu'il fait sont les suivantes : d'une part le critère de courbure-dimension de Bakry-Émery doit être satisfait avec constante strictement positive (correspondant à une courbure de Ricci grossière minorée par un nombre strictement positif) et d'autre part la quantité $\|\sigma\|_{\mathrm{op}}^2$ doit être majorée par une fonction lipschitzienne. Ainsi, on observe que nos résultats sont qualitativement équivalents dans le cadre réversible : il se peut que l'un ou l'autre soit à privilégier selon l'exemple considéré.

2.3 Perspectives

Closons ce second chapitre en mentionnant quelques perspectives suscitées par mes travaux essentiellement sur les thèmes des courbures dans les espaces discrets que j'ai décrits précédemment, la partie « Concentration de la mesure » étant illustrée dans les perspectives du prochain chapitre sur les entrelacements.

2.3.1 Différentes notions de courbures discrètes

Ma dernière contribution au cadre discret date de 2013 avec l'article écrit avec Djalil Chafaï [58] sur les entrelacements pour les processus de naissance et de mort, un concept que je vais introduire dans le prochain chapitre. Depuis, je travaille essentiellement sur ces objets des entrelacements dans le cadre diffusif et, par conséquent, force est de constater que je m'éloigne quelque peu des situations discrètes évoquées dans la Section 2.2. Néanmoins, je me suis toujours tenu au courant des progrès effectués sur le sujet pour au moins deux raisons : d'une part, je suis toujours intéressé par ce thème de l'analyse fonctionnelle et géométrique dans les espaces discrets, avec l'espoir d'y revenir un jour prochain en y contribuant de nouveau, et d'autre part j'accepte une grande partie des sollicitations et demandes diverses d'expertises de travaux soumis pour publication dans des journaux scientifiques dont le thème principal tourne autour de ces considérations discrètes. Il faut dire que les avancées sur le sujet des courbures discrètes sont considérables depuis quelques années, un nombre croissant de chercheurs travaillant sur cette thématique transverse. En effet, plusieurs notions de courbures dans les espaces discrets ont été depuis introduites par plusieurs auteurs, avec des conséquences intéressantes en matière d'inégalités fonctionnelles. Citons entre autres la contribution importante apportée initialement par Maas [141] ainsi que l'article [78], étendant au cadre discret l'approche à la Lott-Sturm-Villani et l'interprétation de type flot-gradient proposée par Jordan, Kinderlehrer et Otto [109], ayant donné lieu ultérieurement à la définition de la courbure de Ricci entropique, une notion à présent plus ou moins stabilisée dans la littérature. Mentionnons également les travaux [35, 84, 98, 99, 134, 135, 147, 148] reposant sur des variantes de cette approche. Par ailleurs, Veysseire a défini dans [175] l'exact analogue de la courbure de Ricci grossière dans une version en temps continu, raffinant ainsi par un passage du global au local la notion de courbure de Wasserstein que j'avais introduite dans [111]. Très récemment, les travaux de Villemonais [178] d'une part et de Liming Wu et ses coauteurs [66] d'autre part poussent encore un peu plus loin l'étude de cette version à temps continu de la courbure de Ricci grossière dans le cadre des processus de Markov de sauts purs et des systèmes de particules en interaction. Néanmoins, les liens et comparaisons entre toutes ces courbures discrètes sont encore difficiles à formuler à l'heure actuelle, même si quelques travaux récents ont fait avancer les choses dans cette direction, notamment en ce qui concerne la courbure de Ricci grossière et le retour (renouveau ?) du critère de courbure-dimension de Bakry-Émery sur les espaces discrets, cf. [79, 110, 139] mais aussi [119]. Ainsi, il serait intéressant de tenter de relier entre eux ces différents concepts, afin d'éventuellement en dégager une hiérarchie claire basée sur certains critères encore à définir.

2.3.2 Courbure de Ricci grossière non strictement positive

Concernant la courbure de Ricci grossière (incluant celle à temps continu définie par Veysseire), l'idée à présent est de tenter d'affaiblir l'hypothèse classique de stricte positivité par une hypothèse moins gourmande telle qu'une courbure strictement positive en moyenne, voire une courbure autorisée à être positive ou nulle (ou même négative) seulement dans une certaine partie de l'espace. Pour la courbure de Ricci entropique, cette question a été récemment abordée par Erbar et Fathi [77], un travail dans lequel certaines inégalités fonctionnelles de type Poincaré ou entropique sont établies en courbure positive ou nulle. Néanmoins, il n'existe que peu de travaux à ce sujet sur la courbure de Ricci grossière. Un exemple très simple mais cependant difficile à appréhender d'un point de vue dynamique est celui de la file d'attente markovienne $M/M/1$, un processus de naissance et de mort dont les taux de transition sont (presque) constants, i.e., le taux de naissance est constant et est donné par un paramètre $\lambda \in]0,1[$ et le taux de mort vaut $1_{\mathbb{N}^*}$, la loi invariante étant la distribution géométrique sur $\mathbb{N}$ de paramètre λ. Ce processus se comporte en dehors de l'origine comme une marche aléatoire asymétrique dérivant vers 0, de courbure nulle (pour la métrique usuelle sur les entiers naturels) sur $\mathbb{N}^*$, la réflection en le point attractif 0 assurant une courbure strictement positive à l'origine seulement. Tout comme son pendant continu, le mouvement brownien avec dérive orientée vers l'origine et réfléchi en 0, de mesure invariante la loi exponentielle sur $\mathbb{R}^+$, un tel modèle discret satisfait une inégalité de Poincaré de constante optimale $(1-\sqrt{\lambda})^2$, et donc converge en loi exponentiellement vite vers l'équilibre. Néanmoins, il ne me semble pas qu'il existe à l'heure actuelle de preuve basée sur la courbure de Ricci grossière (pour la métrique usuelle sur $\mathbb{N}$, sinon le Théorème 2.1.9 combiné au Théorème 2.2.2 résout le problème) permettant de retrouver une telle inégalité fonctionnelle. Un premier pas a été effectué dans ce sens par Ollivier [155] et ensuite Veysseire [176] avec leurs résultats de concentration exponentielle sous l'hypothèse de courbure de Ricci grossière positive ou nulle avec au moins un point attractif. Une autre approche déjà suggérée à l'époque par Ollivier et concrétisée par Paulin [157] consiste à définir la courbure de Ricci grossière non pas sur un seul pas de temps de la chaîne de Markov, mais sur des pas multiples de taille k, menant à une famille de courbures de Ricci grossières indexées par k. Si la courbure initiale est strictement positive seulement sur une grosse partie de l'espace mais pas globalement, alors il arrive quelquefois que la courbure négative ou nulle dispararaisse pour k suffisamment grand, auquel cas les résultats spectraux et de concentration établis dans [113, 155] peuvent (partiellement) s'appliquer. Cependant, en dehors de ces quelques travaux, il n'existe pas pour le moment de traitement synthétique de la courbure de Ricci grossière dans des espaces discrets permettant d'établir des résultats valables au-delà de l'hypothèse de courbure strictement positive sur tout l'espace. Pour ce faire, il faudrait mieux prendre en compte les variations spatiales de ces courbures discrètes en envisageant par exemple une approche plus probabiliste basée sur le couplage, voire sur les entrelacements. Ces derniers ont été étudiés pour ce type de modèles discrets dans notre travail commun avec Chafaï [58], sans pour autant que les résultats obtenus en situation de courbure positive ou nulle soient tout à fait satisfaisants. Néanmoins, il serait intéressant de poursuivre dans cette direction à l'avenir.

2.3.3 La conjecture de Peres-Tetali

Je souhaiterais terminer cette partie sur les perspectives sur les courbures discrètes en mentionnant un problème ouvert, connu depuis peu sous le nom de conjecture de Peres et Tetali. Comme on l'a vu dans le cadre diffusif, une inégalité de Sobolev logarithmique est satisfaite dès lors que le critère de courbure-dimension $\mathrm{CD}(\rho,\infty)$ de Bakry-Émery est vérifié avec une constante ρ strictement positive. Récemment, ce type de résultat a été généralisé par Maas et Erbar [78, 141] dans des espaces discrets. Plus précisément, ils ont démontré pour des chaînes de Markov réversibles sur un ensemble fini une version discrète de l'inégalité HWI d'Otto et Villani [156] sous une hypothèse de courbure entropique strictement positive, entraînant alors l'inégalité entropique comme dans le cadre diffusif (dans ce cadre à temps discret sur un espace fini, l'inégalité entropique est définie de la même manière que dans le cas continu, la mesure $Q(x,\cdot)$ correspondant aux transitions étant remplacée par la matrice de transition $\delta_x P$ de la chaîne, le lien étant $L = P - I$). En revanche, comme nous l'avons souligné dans la Section 2.1.2, les liens entre la courbure de Ricci grossière et l'inégalité entropique ne sont pas clairs du tout. Ceci a amené Peres et Tetali à poser la conjecture suivante, dont on peut trouver l'énoncé dans l'article [76]. De manière un peu plus générale que pour le cas de l'hypercube étudié précédemment, on définit ci-dessous une marche aléatoire paresseuse sur un graphe fini (non orienté, connexe et muni de sa distance naturelle de graphe) comme une chaîne de Markov réversible restant en place avec probabilité 1/2 ou sautant sur l'un de ses voisins choisi selon une probabilité dépendant du poids de l'arête parcourue. Le point important dans ces modèles réside dans la bornitude des sauts indépendamment du diamètre du graphe : on a $\sigma_\infty = 1$.

Conjecture 2.3.1 (Peres-Tetali)**.** *Considérons la marche aléatoire paresseuse sur un graphe fini. Si la courbure de Ricci grossière de la chaîne est minorée par une constante $\kappa > 0$, alors l'inégalité entropique a lieu et la constante optimale ρ_0 satisfait $\rho_0 \geq C\,\kappa$, où $C > 0$ est une constante universelle.*

Comme nous l'avons vu, établir une inégalité entropique permet d'obtenir la décroissance exponentielle du semigroupe en entropie entraînant dans la foulée des estimations du temps de mélange éventuellement plus fines que celles obtenues par l'approche spectrale, une question ayant motivé le célèbre article de Diaconis et Saloff-Coste [72]. Non démontrée à ce jour, il existe quelques indices troublants menant à une telle conjecture. Par exemple, comme cela a été remarqué dans [76], ces deux objets entraînent la concentration gaussienne de la mesure invariante pour des fonctions lipschitziennes, cf. [76, 79] dans le cas d'une courbure de Ricci grossière minorée (par $\kappa > 0$) et [161] pour l'inégalité entropique, les constantes obtenues par l'approche duale des inégalités de transport étant du même ordre de grandeur en κ et ρ_0. Citons de surcroît les articles [33, 73], antérieurs aux deux premiers et qui obtiennent le même type de résultat, ainsi que le travail [31] reposant sur l'utilisation dans l'inégalité entropique de la méthode de Herbst permise par la réversibilité. Mentionnons également que les processus de naissance et de mort, qui sont une version à temps continu et à espace d'états dénombrable de ces marches aléatoires, fournissent un autre indice nous permettant de penser que cette conjecture est vraie. En effet, on sait d'après les travaux de Caputo, Dai Pra et Posta [49] que si les taux de transition de naissance et de mort sont respectivement décroissants et croissants, alors l'inégalité entropique est satisfaite et l'on a

$$\rho_0 \quad \geq \quad \inf_{x \in \mathbb{N}} \nu_{x+1} - \nu_x + \lambda_x - \lambda_{x+1},$$

sous réserve que la quantité apparaissant à droite soit strictement positive (elle est déjà positive ou nulle par les hypothèses de monotonie). Or, le Théorème 2.2.2 indique que cette quantité n'est rien d'autre que la courbure de Wasserstein par rapport à la distance usuelle sur $\mathbb{N}$ (prendre u constant égal à 1), qui est la version à temps continu de la courbure de Ricci grossière minorée. On en déduit que pour ces modèles, la conjecture de Peres-Tetali est vérifiée dès lors que l'on arrive à s'affranchir de ces hypothèses de monotonie, ce qui n'est sans doute pas déraisonnable, ces hypothèses entraînant en réalité une inégalité strictement plus forte que l'inégalité entropique, d'après l'approche adoptée dans l'article [49]. À ma connaissance, cela n'a pas encore été fait et c'est pourquoi l'utilisation des récents résultats de Liming Wu et de ses coauteurs [66] pourrait peut-être permettre une avancée sur le sujet.

Chapitre 3

Entrelacements et inégalités spectrales et fonctionnelles

3.1 Contexte et rappels

À présent, focalisons-nous sur mes travaux les plus récents concernant les entrelacements et ses conséquences en matière d'inégalités spectrales et fonctionnelles. Bien que ces objets soient tout à fait pertinents dans le cadre discret, comme en témoigne le travail avec Djalil Chafaï [58] pour les processus de naissance et de mort à valeurs entières, nous ne nous consacrons dans cette partie qu'à des situations continues, c'est-à-dire que les processus de Markov sous-jacents vivent sur la droite réelle ou dans l'espace euclidien $\mathbb{R}^d$. Ainsi, on considère un processus de Kolmogorov $(X_t)_{t\geq 0}$ à valeurs dans $\mathbb{R}^d$, solution de l'ÉDS

$$dX_t \;=\; \sqrt{2}\, dB_t - \nabla V(X_t)\, dt,$$

où $V : \mathbb{R}^d \to \mathbb{R}$ est un potentiel suffisamment régulier tel que la mesure μ dont la densité de Lebesgue est proportionnelle à e^{-V} soit une mesure de probabilité. Le générateur est donné pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$ par

$$Lf \;=\; \Delta f - \nabla V \cdot \nabla f,$$

et la mesure μ est alors invariante et réversible pour cette dynamique. Enfin, on note $(P_t)_{t\geq 0}$ le semigroupe associé.

3.1.1 Quelques éléments de théorie spectrale

Procédons tout d'abord à quelques rappels de théorie spectrale. L'opérateur $-L$ étant (essentiellement) auto-adjoint et positif dans $L^2(\mu)$, son spectre $\sigma(-L)$ est inclus dans l'intervalle $[0, \infty)$ et est divisé en deux parties : le spectre essentiel $\sigma_{\mathrm{ess}}(-L)$ correspondant à l'ensemble des points limites de $\sigma(-L)$ et des valeurs propres de multiplicité infinie, et le spectre discret $\sigma_{\mathrm{disc}}(-L)$, *i.e.*, le complémentaire du spectre essentiel, constitué des valeurs propres isolées de multiplicité finie. Dans la suite, tous les éléments du spectre seront appelés valeurs propres par abus de langage. Comme on l'a vu au chapitre 2, la valeur propre λ_0 correspondant au bas du spectre est nulle et de multiplicité finie (égale à 1 ici, les seules fonctions propres associées étant justement les constantes, la mesure invariante μ étant supposée finie). Ainsi, son éventuelle appartenance au spectre discret ou essentiel dépend de son caractère isolé ou non, c'est-à-dire de la possible stricte positivité de la valeur propre suivante, définie par la formule variationnelle

$$\lambda_1 \quad = \quad \inf_{f \in \mathbb{1}^\perp} \frac{\int_{\mathbb{R}^d} f\,(-Lf)\,d\mu}{\int_{\mathbb{R}^d} f^2\,d\mu} \;=\; \inf_{f \in \mathbb{1}^\perp} \frac{\int_{\mathbb{R}^d} |\nabla f|^2\,d\mu}{\int_{\mathbb{R}^d} f^2\,d\mu},$$

où $\mathbb{1}^\perp$ désigne le sous-espace du domaine de l'opérateur orthogonal aux constantes. On retrouve là la constante optimale dans l'inégalité de Poincaré. Ainsi dans notre cadre on a $\lambda_1 > 0$ si et seulement si $0 \in \sigma_{\mathrm{disc}}(-L)$, d'où la dénomination de trou spectral pour la quantité λ_1 car dans ce cas $\sigma(-L) \subset \{0\} \cup [\lambda_1, \infty)$. Bien que la valeur explicite du trou spectral soit difficile à obtenir en général, l'enjeu est de pouvoir le minorer convenablement, motivé par le contrôle de la vitesse de convergence dans $L^2(\mu)$ de la dynamique markovienne sous-jacente.

Lorsque le trou spectral d'un opérateur de type laplacien sur un domaine euclidien ou d'une variété riemannienne existe, les tentatives pour l'estimer quantitativement remontent à des temps assez anciens. Plutôt que de proposer une liste de références forcément incomplète sur le sujet, nous renvoyons le lecteur intéressé à la section 4.11 de l'impressionnant ouvrage de référence de Bakry, Gentil et Ledoux [14], dans laquelle est proposée une liste considérable de références précises émanant de travaux issus de l'analyse des équations aux dérivées partielles à la théorie des probabilités, en passant par la géométrie et l'analyse fonctionnelle, le tout saupoudré de commentaires permettant peu ou prou de retracer la longue histoire de ce problème. Dans le cadre qui nous intéresse, celui des processus de Kolmogorov à valeurs dans l'espace euclidien $\mathbb{R}^d$ tout entier, l'un des premiers résultats sur le sujet se lit à travers l'inégalité dite de Brascamp-Lieb [44] : si le potentiel V est strictement convexe sur $\mathbb{R}^d$ (au sens où la matrice hessienne $\nabla^2 V$ est définie positive en chaque point de l'espace), alors pour toute fonction $f : \mathbb{R}^d \to \mathbb{R}$ suffisamment régulière,

$$\mathrm{Var}_\mu(f) \quad \leq \quad \int_{\mathbb{R}^d} \nabla f \cdot (\nabla^2 V)^{-1}\, \nabla f\, d\mu. \tag{3.1.1}$$

En particulier si V est uniformément convexe, *i.e.*, la plus petite valeur propre $\rho(\nabla^2 V)$ de la matrice $\nabla^2 V$ est minorée par un nombre strictement positif, alors la formule variationnelle du trou spectral donne immédiatement la borne

$$\lambda_1 \quad \geq \quad \inf_{\mathbb{R}^d} \rho(\nabla^2 V),$$

que Bakry et Émery [13] ont retrouvé quelques années plus tard *via* le critère de courbure-dimension. Bien qu'optimale dans le cadre gaussien standard, la convexité uniforme du potentiel est loin d'être nécessaire pour obtenir l'existence ainsi qu'une estimation du trou spectral et n'est d'ailleurs pas vérifiée dans beaucoup de cas intéressants. Ainsi, de nombreuses études ont été proposées ces trente dernières années pour tenter d'affaiblir cette hypothèse et parmi celles-ci, retenons la méthode de perturbation dite d'Holley-Stroock [100] permettant d'obtenir une inégalité de Poincaré (et de Sobolev logarithmique) dans le cas d'un potentiel à perturbation bornée, donc une estimation sur le trou spectral. Néanmoins, il s'agit d'un résultat considéré plutôt comme qualitatif que quantitatif, les constantes obtenues dans certains modèles explosant exponentiellement avec la dimension. Mentionnons également une autre contribution importante, le critère dit d'Helffer-Nier [96] utilisant l'approche basée sur le laplacien de Witten. Plus précisément, si le potentiel régulier V est de la forme suivante pour x assez grand :

$$V(x) \quad \approx \quad \sum_{i=1}^{N} |x|^{\alpha_i} \, \varphi_i \left(\frac{x}{|x|} \right), \tag{3.1.2}$$

où les fonctions φ_i sont toutes positives ou nulles sur $\mathbb{S}^{d-1}$ et les paramètres α_i sont tous strictement positifs, alors l'existence d'un trou spectral est assurée si et seulement si d'une part il existe au moins un $\alpha_i \geq 1$ et d'autre part toutes les fonctions φ_i associées aux $\alpha_i \geq 1$ ne s'annulent pas en un même point de $\mathbb{S}^{d-1}$. Ainsi, on retrouve le fait que pour les lois de Subbotin (c'est-à-dire associées aux potentiels du type $V(x) = |x|^\alpha/\alpha$), on a $\lambda_1 > 0$ si et seulement si $\alpha \geq 1$, un résultat déjà mentionné dans la Section 2.2.3. De surcroît, si V s'écrit comme une somme de monômes tous positifs ou nuls, alors le trou spectral existe si et seulement si V tend vers l'infini à l'infini. Néanmoins, ces résultats de caractérisation ne donnent malheureusement pas d'information sur la valeur du trou spectral. Plus récemment [11, 12], l'approche par les conditions de Lyapunov à la Meyn-Tweedie s'est révélée pertinente dans l'étude qualitative du trou spectral, redémontrant en quelques lignes l'existence d'un trou spectral dans le cas d'une mesure log-concave (*i.e.*, le potentiel V est convexe) établie par Kannan, Lovász et Simonovits [116] d'une part et Bobkov [23] d'autre part (en utilisant la méthode de perturbation d'Holley-Stroock, il suffit alors d'un potentiel convexe à l'infini pour que l'existence du trou spectral soit assurée). Néanmoins le caractère quantitatif souffre du même mal dimensionnel que celui obtenu par la méthode de perturbation d'Holley-Stroock. Enfin, signalons un autre résultat devenu classique depuis la note aux CRAS de Veysseire [174], dans laquelle ce dernier établit dans un cadre riemannien une borne de type moyenne harmonique pour le trou spectral. Dans notre cadre euclidien log-concave, l'adaptation de son travail donne lieu à l'inégalité suivante :

$$\lambda_1 \quad \geq \quad \frac{1}{\int_{\mathbb{R}^d} \frac{d\mu}{\rho(\nabla^2 V)}}, \tag{3.1.3}$$

valable pour tout potentiel V strictement convexe. Ainsi, l'intérêt de cette inégalité est de mieux prendre en compte les variations spatiales de la matrice hessienne de V que sous une hypothèse trop brutale de convexité uniforme.

Justement, tentons d'aller un peu plus loin dans le cas log-concave en mentionnant la désormais célèbre conjecture KLS, énoncée en 1995 par Kannan, Lovász et Simonovits [116] et toujours non résolue à ce jour. Concernant initialement une certaine propriété isopérimétrique pour la mesure uniforme sur les corps convexes de $\mathbb{R}^d$ (du point de vue fonctionnel, il s'agit d'une inégalité dite de Poincaré $L^p - L^q$ avec $p = q = 1$, une telle inégalité signifiant que la norme L^p des fonctions recentrées est majorée par une constant fois la norme L^q de la norme euclidienne de leur gradient), un résultat de Ledoux [131] montre que cette conjecture peut être reformulée de manière équivalente en terme de trou spectral pour l'opérateur $-L$, c'est-à-dire *via* l'inégalité de Poincaré classique (correspondant à l'inégalité de Poincaré $L^2 - L^2$ dans la formulation ci-dessus). Ajoutons que Milman a ensuite établi [149] dans ce cadre log-concave l'équivalence de toutes les inégalités de Poincaré $L^p - L^q$, $1 \leq p \leq q \leq \infty$, à une constante universelle près dépendant de p et q.

Conjecture 3.1.1 (KLS). *On suppose que la mesure μ est log-concave sur $\mathbb{R}^d$. Alors il existe une constante universelle $C > 0$ (i.e., indépendante de la mesure μ et de la dimension d) telle que l'on ait l'inégalité*

$$\lambda_1 \quad \geq \quad \frac{C}{\|\mathrm{Cov}_\mu\|_{\mathrm{op}}},$$

où Cov_μ *désigne la matrice de covariance $d \times d$ de la mesure μ.*

En d'autres termes, la conjecture KLS énonce que le trou spectral est de l'ordre de celui de l'opérateur $-L$ restreint aux fonctions linéaires. Cette conjecture étant « invariante » par transformation linéaire, on ne perd rien en se restreignant aux mesures log-concaves isotropes, c'est-à-dire centrées et de matrice covariance identité, auquel cas l'inégalité spectrale de la conjecture KLS devient simplement $\lambda_1 \geq C$. Ainsi, le résultat obtenu par la méthode de localisation issue de [116] s'écrit dans ce cadre spectral sur l'espace $\mathbb{R}^d$ tout entier,

$$\lambda_1 \quad \geq \quad \frac{C}{\mathrm{Trace}\,(\mathrm{Cov}_\mu)},$$

un résultat notamment retrouvé par Bobkov [23]. Sous l'hypothèse additionnelle d'isotropie, on obtient $\lambda_1(-L) \geq C/d$. Depuis le début du vingt-et-unième siècle, des avancées ont été réalisées à propos de cette conjecture. Dernièrement, Lee et Vempala [133] ont utilisé la méthode de localisation stochastique développée par Eldan [75] afin d'améliorer l'estimation du trou spectral de la manière suivante :

$$\lambda_1(-L) \geq \frac{C}{\|\mathrm{Cov}_\mu\|_{\mathrm{HS}}},$$

donnant dans le cas isotrope l'inégalité $\lambda_1(-L) \geq C/\sqrt{d}$. Par ailleurs, au-delà des mesures produits $\mu = \otimes\, \mu_i$ pour lesquelles la conjecture est vraie grâce à la propriété de tensorisation de l'inégalité de Poincaré, *i.e.*,

$$\lambda_1 \quad = \quad \min_{i=1,\ldots,d} \lambda_1^i,$$

où λ_1^i est le trou spectral associé à la dynamique unidimensionnelle dont la mesure invariante sur $\mathbb{R}$ est μ_i, elle a été démontrée dans quelques cas particuliers, parmi lesquels les mesures log-concaves radiales, un résultat dû à [24] à travers la borne

$$\lambda_1 \quad \geq \quad \frac{d}{12\, \int_{\mathbb{R}^d} |x|^2\, \mu(dx)}. \tag{3.1.4}$$

Citons aussi les mesures uniformes sur les boules unités de l_p, $p \geq 1$, cf. [125, 164], sur le simplexe régulier [20] ou encore sur certains convexes de révolution [107] et boules d'Orlicz [122]. Enfin, dans le cas des mesures uniformes sur les corps convexes inconditionnels, c'est-à-dire symétriques par rapport aux hyperplans de coordonnées, Klartag [118] a établi la borne

$$\lambda_1(-L) \geq \frac{C}{(\log d)^2\, \|\mathrm{Cov}_\mu\|_{\mathrm{op}}},$$

qui est donc la borne attendue, à un facteur logarithmique près. Notons enfin que ce résulat a été étendu dans [16] à d'autres types de symétries. Pour en savoir plus sur cette fameuse conjecture et les liens qu'elle entretient avec d'autres conjectures (comme celle de la variance par exemple), le lecteur intéressé pourra consulter l'ouvrage d'Alonso-Gutiérrez et Bastero [5].

Sur la droite réelle, le trou spectral est connu dans quelques rares situations, comme pour le cas gaussien pour lequel on rappelle que l'on a $\lambda_1 = 1$ ou encore pour la loi de Laplace (connue aussi sous le nom de la loi double-exponentielle sur $\mathbb{R}$; c'est la loi de Subbotin de paramètre $\alpha = 1$) avec $\lambda_1 = 1/4$, une valeur que vérifie aussi le trou spectral pour la loi logistique, de potentiel $V(x) = 2\log(1+e^x) - x$, cf. [15]. Dans le cas de mesures dont le support est un intervalle borné, étendant l'exemple classique de la mesure uniforme sur un intervalle, l'article récent de Barthe, Iooss et Roustant [17] produit de nouvelles formulations explicites du trou spectral, notamment pour la version tronquée de quelques-unes des mesures précédemment citées. Plus généralement, le critère de Muckenhoupt [153], datant des années 70 et basé sur les inégalités de type Hardy, établit dans le cas unidimensionnel une complète caractérisation sous la forme de critères intégraux sur le potentiel V pour que l'inégalité de Poincaré soit satisfaite, avec un contrôle du trou spectral à un facteur 8 près. Mentionnons également les travaux [25, 26, 146] permettant d'améliorer quelque peu ce critère. Une élégante analyse des cas pour lesquels le critère de Muckenhoupt est satisfait est proposée dans la Section 8.4 de l'excellent livre toulousain [6]. Le cas des mesures log-concaves sur $\mathbb{R}$, quant à lui, a été revisité par Fougères [80] à l'aide de ce critère. Par ailleurs, une approche par transport de mesures en dimension 1 a été proposée dans [27] pour estimer la constante de Cheeger, *i.e.*, la meilleure constante dans l'inégalité de Poincaré $L^1 - L^1$, revenant donc dans le cas log-concave à estimer le trou spectral. Enfin, le résultat unidimensionnel sans doute le plus facile d'utilisation est la célèbre formule dite de Chen-Wang [65] obtenue par un argument de couplage,

$$\lambda_1 \quad \geq \quad \sup_h \inf_{\mathbb{R}} \frac{(-Lh)'}{h'}, \tag{3.1.5}$$

le supremum étant pris sur les fonctions régulières $h : \mathbb{R} \to \mathbb{R}$ de dérivée ne s'annulant pas. L'intérêt d'une telle formule est que tout bon choix de la fonction test h donne lieu à une estimation du trou spectral. De surcroît, sa structure fait clairement apparaître les cas d'égalité : dès que le trou spectral appartient au spectre discret de l'opérateur $-L$, le supremum est alors un maximum : il suffit de prendre $h = g_1$, où g_1 est une fonction propre associée au trou spectral, dont on sait que la dérivée ne s'annule pas, cf. [65].

Les résultats en dimension 1 sont essentiels au sens où de nombreux modèles éventuellement complexes, au-delà des cas produits classiques, peuvent se réduire à l'étude d'une certaine dynamique sur la droite réelle, tels que des systèmes de spins non bornés issus de la mécanique statistique. Par exemple si le potentiel V est de la forme

$$V(x) \quad = \quad \sum_{i=1}^d U(x_i) + \varphi(x), \quad x \in \mathbb{R}^d,$$

la mesure μ est une mesure produit (correspondant à la mesure de probabilité unidimensionnelle de densité de Lebesgue proportionnelle à e^{-U} sur $\mathbb{R}$) perturbée par un potentiel d'interaction φ.

Dans les applications, le potentiel régulier unidimensionnel U, appelé phase, peut être supposé convexe ou non et φ modélise souvent des interactions aux plus proches voisins, *i.e.*, $\varphi(x) = J\sum_{i=1}^d \phi(x_i, x_{i+1})$, $x_{d+1} = x_1$ avec ϕ une fonction lisse et J est le coefficient d'interaction en général assez petit (on parle de régime perturbatif). De nouveau basés sur le laplacien de Witten, les travaux d'Helffer [93, 94] pour estimer le trou spectral de tels modèles reposent sur l'existence d'un trou spectral pour les lois conditionnelles unidimensionnelles uniforme en les autres coordonnées, fixées. Plus précisément, si l'on note $\widetilde{\nabla^2 V}$ la matrice hessienne de V remplie de zéros sur la diagonale, le résultat d'Helffer est le suivant :

$$\lambda_1 \quad \geq \quad \inf_{\mathbb{R}^d} \rho\left(\widetilde{\nabla^2 V}\right) + \min_{i=1,\dots,d} \inf_{\hat{x_i}\in\mathbb{R}^{d-1}} \lambda_1^{\hat{x_i}}, \tag{3.1.6}$$

où pour tout $x \in \mathbb{R}^d$ fixé, $\hat{x_i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d)$ désigne le vecteur de $\mathbb{R}^{d-1}$ extrait de x en enlevant la i-ème coordonnée et $\lambda_1^{\hat{x_i}}$ est le trou spectral de la loi unidimensionnelle de la i-ème coordonnée sous la mesure μ, les autres étant fixées égales à $\hat{x_i}$. Ce type d'approche peut s'appliquer par exemple au cas d'une phase correspondant à un double-puits [93], $U(x) = |x|^4 - \beta|x|^2$ avec $\beta > 0$, voire même à une phase seulement uniformément convexe à l'infini [94], Helffer optant dans ses deux articles pour un potentiel d'interaction quadratique du type $\phi(x_i, x_{i+1}) = |x_i - x_{i+1}|^2$. Par ailleurs, les articles [32, 130] considèrent le cas d'une phase uniformément convexe à l'infini associée à des interactions aux plus proches voisins plus générales du type $\phi(x_i, x_{i+1}) = x_i\, x_{i+1}$ ou $\phi(x_i, x_{i+1}) = \widetilde{\phi}(x_i - x_{i+1})$ avec des conditions de symétrie et de croissance sur $\widetilde{\phi}$, tandis que le travail de Gentil et Roberto [81] permet dans ce cadre d'aller au-delà de l'hypothèse de convexité uniforme à l'infini de la phase en exploitant le critère de Muckenhoupt mentionné plus haut.

À présent, consacrons-nous aux valeurs propres supérieures situées sous le spectre essentiel, que l'on organise en ordre croissant et compte avec leur multiplicité. Pour ce faire, le théorème du min-max de Courant-Fisher, dont on peut trouver une version dans l'ouvrage d'Helffer [95], nous dit que la quantité définie par la formule variationnelle

$$\lambda_n \quad = \quad \sup_{g_0,\dots,g_{n-1}\in L^2(\mu)} \ \inf_{f\perp g_i,\, i=0,\dots,n-1} \frac{\int_{\mathbb{R}^d} f\,(-Lf)\, d\mu}{\int_{\mathbb{R}^d} f^2\, d\mu},$$

désigne soit la n-ième valeur propre strictement positive du spectre discret située sous le spectre essentiel, soit le bas du spectre essentiel lui-même, toutes les λ_m coïncidant avec λ_n pour $m > n$, et il y a alors au plus $n-1$ valeurs propres strictement positives du spectre discret strictement inférieures à λ_n. Ci-dessus, la notation $\perp$ désigne l'orthogonalité dans $L^2(\mu)$. Le supremum est réalisé lorsque les fonctions g_i sont les fonctions propres associées, si elles existent. C'est par exemple le cas si le spectre est discret, *i.e.*, $\sigma_{\mathrm{ess}}(-L) = \emptyset$, comme dans le cas d'un potentiel uniformément convexe ou encore s'il existe $\theta \in]0,1[$ tel que

$$\lim_{|x|\to\infty} \theta|\nabla V(x)|^2 - \Delta V(x) \quad = \quad +\infty,$$

cf. [95]. Le cas gaussien standard vérifiant la propriété de convexité uniforme, le spectre est discret et vaut exactement $\sigma(-L) = \mathbb{N}$, les fonctions propres associées étant les ploynômes d'Hermite. Plus généralement pour les lois de Subbotin de paramètre $\alpha \geq 1$, le second critère indique que le spectre est discret si $\alpha > 1$, le spectre discret étant réduit à $\{0\}$ avec de surcroît $\sigma_{\mathrm{ess}}(-L) = [1/4, \infty[$ dans le cas $\alpha = 1$. Le cas du double-puits est aussi couvert par le second critère, indépendamment de la valeur du paramètre $\beta > 0$. Enfin, dans le même article déjà évoqué plus haut pour le trou spectral, Helffer et Nier [96] établissent une version de leurs résultats pour l'obtention d'un spectre discret. Tout d'abord, si le potentiel V s'écrit comme une somme de monômes tous positifs ou nuls, alors on a $\sigma_{\mathrm{ess}}(-L) = \emptyset$ si et seulement si V tend vers l'infini à l'infini. Autrement dit, d'après ce que nous avons énoncé auparavant, cela signifie que dans cette situation bien particulière, l'existence d'un trou spectral est équivalente à la condition *a priori* bien plus forte que le spectre soit discret, faisant de cette condition nécessaire et suffisante un résultat remarquable. Enfin, lorsque l'on reprend un potentiel V du type (3.1.2), une condition suffisante permettant d'obtenir un spectre discret est de supposer d'une part qu'il existe au moins un $\alpha_i > 1$ et d'autre part que toutes les fonctions φ_i associées aux $\alpha_i > 1$ ne s'annulent pas en un même point de $\mathbb{S}^{d-1}$: c'est le second critère d'Helffer-Nier (il s'agit d'une condition suffisante similaire à celle du premier critère, avec la différence notable que les inégalités ≥ 1 sont remplacées par > 1).

3.1.2 L'entrelacement classique

À présent, introduisons la notion d'entrelacement. L'entrelacement classique entre gradient et opérateur est de la forme suivante : pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$,

$$\nabla L f \;=\; \widetilde{\mathcal{L}}\, \nabla f, \tag{3.1.7}$$

où il faut donner un sens à l'opérateur $\widetilde{\mathcal{L}}$ agissant sur les gradients et plus généralement sur les champs de vecteurs réguliers de $\mathbb{R}^d$. Cet opérateur est de type Schrödinger, *i.e.*, $\widetilde{\mathcal{L}} = \mathcal{L} - \nabla^2 V$ avec

$$\mathcal{L} \;=\; \begin{pmatrix} L & & \\ & \ddots & \\ & & L \end{pmatrix},$$

vu comme un opérateur de diffusion matriciel $d \times d$ diagonal et la matrice hessienne $\nabla^2 V$ agit de manière multiplicative sur les champs de vecteurs. Ce type d'entrelacement apparut anciennement en géométrie riemannienne à travers la formule de Weitzenböck pour les 1-formes différentielles, mettant en scène l'opérateur de Laplace-Beltrami (le générateur du mouvement brownien sur la variété), le laplacien de Hodge-de Rham (commutant avec la différentielle extérieure) et la courbure de Ricci (le potentiel multiplicatif, ou opérateur d'ordre 0), cf. [106]. Au fond, l'entrelacement est une notion réminiscente du calcul Γ_2 de Bakry et Émery, *via* la formule de Bochner dans notre cadre euclidien :

$$\begin{aligned} \Gamma_2(f,f) \;&=\; \frac{1}{2}\, L(|\nabla f|^2) - \nabla f \cdot \nabla L f \\ &=\; \frac{1}{2}\, L(|\nabla f|^2) - \nabla f \cdot \widetilde{\mathcal{L}}\, \nabla f \\ &=\; \underbrace{\frac{1}{2}\, L(|\nabla f|^2) - \nabla f \cdot \mathcal{L}\, \nabla f} + \nabla f \cdot \nabla^2 V\, \nabla f \end{aligned}$$

$$= \qquad \|\nabla^2 f\|_{\mathrm{HS}}^2 \qquad + \nabla f \cdot \nabla^2 V \, \nabla f,$$

la quantité $\|\nabla^2 f\|_{\mathrm{HS}}^2$ étant interprétée comme une sorte de carré du champ associé à l'opérateur matriciel $\mathcal{L}$ agissant dans ses deux arguments sur le gradient ∇f.

Au niveau des semigroupes, l'entrelacement classique s'écrit de la manière suivante : pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$,

$$\nabla P_t f \;=\; \widetilde{\mathcal{P}_t} \nabla f. \tag{3.1.8}$$

Ci-dessus, la famille d'opérateurs matriciels $(\widetilde{\mathcal{P}_t})_{t\geq 0}$ agissant sur les gradients, ou plus généralement sur les champs de vecteurs, est aussi un semigroupe mais n'est en général pas conservatif, *i.e.* $\widetilde{\mathcal{P}_t}\mathbb{1} \neq \mathbb{1}$ pour $t > 0$, $\mathbb{1}$ désignant le vecteur constant de $\mathbb{R}^d$ ayant toutes ses coordonnées égales à 1 et de surcroît les propriétés usuelles de contraction (*via* la norme euclidienne) dans les espaces de Lebesgue ne sont plus forcément satisfaites. Néanmoins, étudier de tels semigroupes définis sur les champs de vecteurs de $\mathbb{R}^d$ s'avère très intéressant pour de multiples raisons, la première étant leur représentation probabiliste sous la forme Feynman-Kac, permettant d'en déduire des propriétés importantes d'ergodicité du semigroupe originel $(P_t)_{t\geq 0}$. En effet, cette interprétation probabiliste, que l'on peut faire remonter dans le cadre riemannien à Airault [3] au milieu des années 70, a notamment été utilisée par Bakry en 1987 dans son étude sur les transformées de Riesz [9] et fait élégamment le lien avec les célèbres formules de Bismut et le calcul de Malliavin, cf. [22]. Citons de nouveau le livre de Hsu [106] pour une présentation claire et détaillée de l'ensemble de ces objets définis dans le cadre riemannien. Dans notre cadre euclidien, le semigroupe matriciel $(\widetilde{\mathcal{P}_t})_{t\geq 0}$ admet précisément la représentation probabiliste suivante : pour tout champ de vecteurs de $\mathbb{R}^d$ suffisamment régulier,

$$\widetilde{\mathcal{P}_t} F(x) \;=\; \mathbb{E}_{\delta_x}\left[J_t \, F(X_t)\right], \quad t \geq 0, \quad x \in \mathbb{R}^d,$$

où $(J_t)_{t\geq 0}$ est solution matricielle de l'équation différentielle

$$\partial_t J_t + J_t \, \nabla^2 V(X_t) \;=\; 0, \quad J_0 = I.$$

Par exemple, nous retrouvons dans le cas gaussien standard la relation de commutation

$$\nabla P_t f(x) \;=\; e^{-t}\, \mathcal{P}_t \nabla f(x) \;=\; e^{-t}\, \mathbb{E}_{\delta_x}\left[\nabla f(X_t)\right],$$

classiquement obtenue par la formule de Mehler,

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

De retour au cas général, notons que la représentation probabiliste peut être obtenue en utilisant la méthode du processus tangent, c'est-à-dire en différentiant par rapport à la condition initiale le processus de Kolmogorov $(X_t)_{t\geq 0}$. Ce flot de matrices $(J_t)_{t\geq 0}$ joue le rôle du poids de Feynman-Kac $\exp\left(-\int_0^t V''(X_s)\, ds\right)$ que l'on obtient classiquement en dimension 1, un cadre dans lequel les choses se simplifient, la raison évidente étant que le gradient d'une fonction réelle définie sur la droite réelle n'est rien d'autre qu'une fonction de $\mathbb{R}$ dans $\mathbb{R}$. En particulier si $\rho : \mathbb{R}^d \to \mathbb{R}$

désigne la fonction correspondant à la plus petite valeur propre de la matrice symétrique $\nabla^2 V$, alors $\|J_t\|_{\mathrm{op}} \leq \exp\left(-\int_0^t \rho(X_s)\, ds\right)$ et en recollant l'ensemble, on en déduit l'inégalité suivante : pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$,

$$|\nabla P_t f| \quad \leq \quad P_t^\rho \left(|\nabla f|\right),$$

où $(P_t^\rho)_{t\geq 0}$ est le (véritable) semigroupe de Feynman-Kac agissant sur les fonctions régulières,

$$P_t^\rho f(x) \quad = \quad \mathbb{E}_{\delta_x}\left[f(X_t)\, \exp\left(-\int_0^t \rho(X_s)\, ds\right)\right], \quad t \geq 0, \quad x \in \mathbb{R}^d.$$

Cette inégalité, retrouvée très simplement à partir de l'entrelacement classique (3.1.8), n'est rien d'autre que la relation de sous-commutation renforcée de Bakry et Émery obtenue en suant quelque peu à partir du calcul Γ_2, à une amélioration près : la fonction ρ elle-même apparaît au lieu de sa borne inférieure, ouvrant la voie à des approches sans doute plus probabilistes des inégalités fonctionnelles pour des mesures dont le potentiel associé V n'est pas uniformément convexe. Mentionnons que le critère de courbure-dimension $\mathrm{CD}(\rho, \infty)$ originel de Bakry-Émery peut être aisément adapté dans le cas où la constante réelle ρ est en fait une fonction, modulo l'utilisation de tels semigroupes de Feynman-Kac. Cette observation, que Bakry avait déjà formulée dans son cours de Saint-Flour [10], a été récemment exploitée avec succès par Sturm et ses collaborateurs [45, 167] dans le cadre des espaces métriques mesurés.

À présent, relions l'entrelacement classique (3.1.7) au trou spectral de l'opérateur $-L$, que l'on note dans la suite $\lambda_1(-L)$ (plutôt que λ_1), l'analyse spectrale à venir mettant en scène plusieurs opérateurs différents. Ce lien, sous-jacent aux travaux d'Helffer et Sjöstrand [97] autour des laplaciens de Witten et de la décroissance des corrélations de certains modèles en mécanique statistique, a été formalisé par Johnsen dans l'article [108]. Citons par exemple l'ouvrage d'Helffer [95] pour un état de l'art sur ce vaste sujet. L'opérateur de type Schrödinger $-\widetilde{\mathcal{L}}$ est symétrique et positif dans $L^2(\mu)$ (présentement et dans la suite, on note aussi $L^2(\mu)$ l'espace des champs de vecteurs sur $\mathbb{R}^d$ de norme de carré intégrable par rapport à μ). Il est donc (essentiellement) auto-adjoint sur cet espace. La même propriété est vérifiée lorsque l'opérateur est restreint au sous-espace $\nabla_I = \{\nabla f \in L^2(\mu) : f \text{ régulière sur } \mathbb{R}^d\}$ engendré par les gradients de fonctions. Le résultat déterminant de Johnsen est alors le suivant.

Théorème 3.1.1. *Sous l'hypothèse de trou spectral $\lambda_1(-L) > 0$, les opérateurs (essentiellement) auto-adjoints $-L_{|\mathbb{1}^\perp}$ et $-\widetilde{\mathcal{L}}_{|\nabla_I}$, restreints respectivement aux sous-espaces $\mathbb{1}^\perp$ et ∇_I, sont unitairement équivalents, la transformation unitaire pour passer de l'un à l'autre étant la transformée de Riesz $U = \nabla(-L)^{-1/2} : \mathbb{1}^\perp \to \nabla_I$:*

$$-\widetilde{\mathcal{L}}_{|\nabla_I} \quad = \quad U\,(-L)_{|\mathbb{1}^\perp}\, U^*.$$

Par conséquent, les spectres coïncident : $\sigma(-L_{|\mathbb{1}^\perp}) = \sigma(-\widetilde{\mathcal{L}}_{|\nabla_I})$, tout comme les sous-parties constituées des spectres essentiels et discrets.

Dès lors que l'on a $\lambda_1(-L) > 0$, le semigroupe $(P_t)_{t\geq 0}$ converge à vitesse exponentielle $e^{-\lambda_1 t}$ vers l'équilibre μ et par conséquent l'opérateur ci-dessus $(-L)^{-1/2}$, ou plus généralement l'opérateur $(-L)^{-\alpha}$ pour $\alpha > 0$, correspondant au potentiel de Riesz

$$(-L)^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} P_t \, dt,$$

est bien défini comme opérateur borné du sous-espace $\mathbb{1}^\perp$ vers lui-même. Ici, Γ désigne la célèbre fonction Gamma, $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} \, dt$. Par ailleurs, une analyse similaire permet de correctement définir sur le sous-espace des gradients l'opérateur $(-\widetilde{\mathcal{L}})^{-\alpha}$, $\alpha > 0$, en fonction du semigroupe $(\widetilde{\mathcal{P}_t})_{t\geq 0}$ sous l'hypothèse $\lambda_0(-\widetilde{\mathcal{L}}_{|\nabla_I}) > 0$, une valeur propre correspondant au bas du spectre de l'opérateur de type Schrödinger $-\widetilde{\mathcal{L}}_{|\nabla_I}$ restreint aux gradients définie par la formule variationnelle

$$\lambda_0(-\widetilde{\mathcal{L}}_{|\nabla_I}) = \inf_{F \in \nabla_I} \frac{\int_{\mathbb{R}^d} F \cdot (-\widetilde{\mathcal{L}} F) \, d\mu}{\int_{\mathbb{R}^d} |F|^2 \, d\mu}.$$

Ainsi, grâce à l'entrelacement au niveau des semigroupes (3.1.8), il en résulte que pour tout $\alpha > 0$,

$$\nabla(-L)^{-\alpha} = (-\widetilde{\mathcal{L}})^{-\alpha} \, \nabla,$$

et en particulier la transformée de Riesz du Théorème 3.1.1 s'écrit aussi $U = (-\widetilde{\mathcal{L}})^{-1/2} \nabla$. Le lien entre ces deux valeurs propres se fait justement par l'une des conséquences du Théorème 3.1.1 traduite par l'égalité

$$\lambda_1(-L) = \lambda_0(-\widetilde{\mathcal{L}}_{|\nabla_I}),$$

la valeur propre nulle ayant pour fonctions propres associées les constantes (on a donc $\sigma(-L_{|\mathbb{1}^\perp}) = \sigma(-L) \backslash \{0\}$). Autrement dit, le trou spectral de l'opérateur de diffusion $-L$ vaut la valeur propre correspondant au bas du spectre de l'opérateur de type Schrödinger $-\widetilde{\mathcal{L}}_{|\nabla_I}$ restreint aux gradients. Notons que l'inégalité $\lambda_1(-L) \geq \lambda_0(-\widetilde{\mathcal{L}}_{|\nabla_I})$ avait été observée par Helffer dans [93] comme conséquence immédiate de l'identité

$$\mathrm{Var}_\mu(f) = \int_{\mathbb{R}^d} \nabla f \cdot (-\widetilde{\mathcal{L}})^{-1} \, \nabla f \, d\mu, \tag{3.1.9}$$

obtenue grâce à l'entrelacement classique (3.1.7), sans pour autant que l'égalité soit établie. Ce n'est que deux ans plus tard que Johnsen l'établira. Dans l'esprit, cette approche ressemble à la méthode L^2 d'Hörmander [101] que ce dernier a mise en œuvre pour résoudre l'équation de Poisson associée à l'opérateur $\bar{\partial}$ une grosse trentaine d'années avant le travail d'Helffer et dans le contexte différent de l'analyse complexe. Ainsi, l'opérateur de diffusion matriciel diagonal $-\mathcal{L}$ étant positif, un petit argument de monotonie dans la formule variationnelle de $\lambda_0(-\widetilde{\mathcal{L}}_{|\nabla_I})$ nous permet d'une part de retrouver l'inégalité de Brascamp-Lieb (3.1.1) si le potentiel V est strictement convexe et d'autre part l'inégalité $\lambda_1(-L) \geq \inf_{\mathbb{R}^d} \rho(\nabla^2 V)$ dans le cas uniformément convexe. Remarquons qu'en exploitant plus finement la positivité de l'opérateur $-\mathcal{L}$, cette approche laisse manifestement de la place pour de potentielles améliorations des estimations du trou spectral lorsque le critère de courbure-dimension de Bakry-Émery ne s'applique pas, comme dans le cas d'une mesure log-concave générale (voire même lorsque V n'est pas convexe, cf. l'exemple du double-puits). En particulier, il est tentant de se poser la question suivante : l'interprétation spectrale de l'entrelacement classique peut-elle contribuer à faire avancer les choses concernant la conjecture KLS ? Sans être en mesure de répondre directement à cette question, nous rodons sans cesse autour de ces considérations log-concaves, les travaux que nous menons depuis 2013 s'inscrivant souvent dans ce cadre, comme nous allons le voir à présent.

3.2 Quelques résultats

3.2.1 L'entrelacement à poids

Rentrons maintenant dans le vif du sujet en décrivant quelques-uns de nos travaux sur les entrelacements en lien avec les inégalités spectrales et fonctionnelles. Faisant suite aux travaux en discret effectués avec Djalil Chafaï [58] en 2013, l'idée fondatrice dans le cadre diffusif, que nous avons eue principalement avec Michel Bonnefont, a été formalisée tout d'abord sur la droite réelle dans l'article [37] publié en 2014, puis généralisée en dimension quelconque dans le travail [8] en collaboration avec Marc Arnaudon, datant de 2016 et publié en 2018. Depuis, nous travaillons toujours sur le sujet, comme en témoignent les deux prépublications récentes [38, 39] que nous avons écrites Michel et moi. Suivant la présentation proposée dans [8], l'idée dont nous parlons repose sur l'introduction dans l'entrelacement classique (3.1.7) d'un poids matriciel $x \mapsto A(x) \in \mathrm{GL_d}(\mathbb{R})$ (dans la suite, on notera $A \in \mathrm{GL_d}(\mathbb{R})$ pour ne pas trop surcharger la notation), dépendant de la variable d'espace d'une manière régulière et correspondant à une notion de h-transformée de Doob matricielle avec $h = A^{-1}$: « on multiplie par h à l'intérieur de l'opérateur matriciel $\widetilde{\mathcal{L}}$ et par h^{-1} à l'extérieur ». Cette transformation (au sens classique et non matricielle), familière aux stochasticiens lorsqu'ils désirent déterminer la loi de processus markoviens conditionnés, laisse le carré du champ inchangé mais modifie la mesure invariante selon le poids A. Plus précisément, si l'on note $\widetilde{\mathcal{L}_A}$ l'opérateur matriciel agissant sur les champs de vecteurs F suffisamment réguliers de $\mathbb{R}^d$ défini par

$$\widetilde{\mathcal{L}_A} F \;=\; A\,\widetilde{\mathcal{L}}\,(A^{-1}\,F),$$

alors partant de l'entrelacement classique (3.1.7), on a l'entrelacement à poids suivant : pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$,

$$A\,\nabla Lf \;=\; \widetilde{\mathcal{L}_A}\,(A\,\nabla f). \tag{3.2.1}$$

Comme dans le cas de l'entrelacement classique (3.1.7) (correspondant ici au choix d'une matrice A constante et inversible), l'opérateur matriciel $\widetilde{\mathcal{L}_A}$ est de type Schrödinger, c'est-à-dire qu'il se décompose sous la forme $\widetilde{\mathcal{L}_A} = \mathcal{L}_A - M_A$, où $\mathcal{L}_A$ est l'opérateur de diffusion matriciel

$$\mathcal{L}_A F \;=\; \mathcal{L}F + 2A\,\nabla A^{-1} \cdot \nabla F,$$

et M_A est la matrice

$$M_A \;=\; A\,\left(\nabla^2 V - \mathcal{L}A^{-1}\,A\right)\,A^{-1},$$

correspondant à un opérateur multiplicatif. Ci-dessus, le gradient ∇F est défini comme le vecteur formé des coordonnées vectorielles ∇F_i et si la matrice A^{-1} a pour entrées les $a^{i,j}$, alors ∇A^{-1} désigne la matrice des gradients $\nabla a^{i,j}$. Enfin, ces gradients agissent par contraction pour définir le vecteur $\nabla A^{-1} \cdot \nabla F$:

$$(\nabla A^{-1} \cdot \nabla F)_i \quad = \quad \sum_{j=1}^{d} \nabla a^{i,j} \cdot \nabla F_i.$$

Comme l'opérateur $-\widetilde{\mathcal{L}}$ est (essentiellement) auto-adjoint dans l'espace $L^2(\mu)$, on en déduit que l'opérateur $-\widetilde{\mathcal{L}_A}$ est au moins symétrique dans l'espace $L^2_A(\mu)$ constitué des champs de vecteurs F suffisamment réguliers sur $\mathbb{R}^d$ satisfaisant la condition d'intégrabilité

$$\int_{\mathbb{R}^d} F \cdot \left(A\,A^T\right)^{-1} F\, d\mu \quad < \quad \infty.$$

De plus, l'entrelacement à poids (3.2.1) entraîne la positivité de $-\widetilde{\mathcal{L}_A}$ lorsqu'il est restreint au sous-espace des gradients à poids $\nabla_A = \{A\,\nabla f \in L^2_A(\mu) : f \text{ régulière sur } \mathbb{R}^d\}$. Ainsi, l'opérateur $-\widetilde{\mathcal{L}_A}_{|\nabla_A}$ est lui aussi (essentiellement) auto-adjoint sur ∇_A. En revanche, les symétrie et positivité dans $L^2_A(\mu)$ (évaluées indifféremment sur l'espace des champs de vecteurs suffisamment réguliers sur $\mathbb{R}^d$ ou sur ∇_A) de l'opérateur de diffusion matriciel $-\mathcal{L}_A$ ne sont pas automatiques et l'hypothèse sur A nécessaire et suffisante pour ce faire est la suivante :

$$(H_{\text{sym}}) \qquad \text{la matrice } \mathcal{L}A^{-1}\,A \text{ est symétrique.}$$

Récapitulons en énonçant la proposition suivante.

Proposition 3.2.1. *Étant donné un poids matriciel $A \in \mathrm{GL_d}(\mathbb{R})$ régulier, l'opérateur de type Schrödinger matriciel $-\widetilde{\mathcal{L}_A}$ est d'une part symétrique sur $L^2_A(\mu)$ et d'autre part (essentiellement) auto-adjoint et positif sur le sous-espace ∇_A engendré par les gradients à poids.*
L'opérateur de diffusion matriciel $-\mathcal{L}_A$, quant à lui, est (essentiellement) auto-adjoint et positif dans $L^2_A(\mu)$ (donc sur ∇_A) si et seulement si l'hypothèse (H_{sym}) est satisfaite.

Notons que c'est le cas par exemple si A est diagonale, ce qui permet, comme nous le verrons plus tard, d'appréhender dans notre étude spectrale à venir des cas de potentiels V non uniformément convexes, voire même non convexes.

Du point de vue des semigroupes, l'entrelacement à poids s'écrit pour toute fonction réelle f suffisamment régulières sur $\mathbb{R}^d$,

$$A\,\nabla P_t f \quad = \quad \widetilde{\mathcal{P}_{A,t}}(A\,\nabla f), \tag{3.2.2}$$

où le semigroupe matriciel de type Feynman-Kac $(\widetilde{\mathcal{P}_{A,t}})_{t\geq 0}$ agit sur les champs de vecteurs réguliers. Contrairement au cas du semigroupe issu de l'entrelacement classique (3.1.8), l'interprétation probabiliste sous-jacente n'est pas claire, la raison principale étant due au caractère éventuellement non diagonal de l'opérateur de diffusion matriciel $\mathcal{L}_A$. En revanche lorsque la matrice A est diagonale, $A = \mathrm{diag}\,(a_i)$ avec les a_i des fonctions régulières sur $\mathbb{R}^d$ ne s'annulant pas, cet opérateur devient diagonal, *i.e.*,

$$\mathcal{L}_A = \begin{pmatrix} L_{a_1} & & \\ & \ddots & \\ & & L_{a_d} \end{pmatrix},$$

où sur chaque ligne apparaît le vrai opérateur de diffusion suivant : pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$,

$$L_{a_i} f \;=\; Lf + 2a_i \, \nabla a_i^{-1} \cdot \nabla f,$$

la mesure invariante μ_{a_i} du processus de Markov sous-jacent $(X_{a_i,t})_{t\geq 0}$ ayant donc sa densité de Lebesgue proportionnelle à $a_i^{-2}\, e^{-V}$ (il n'y a aucune raison *a priori* qu'elle soit de masse finie, donc renormalisable en une mesure de probabilité). Si de surcroît la matrice A est un multiple de l'identité, *i.e.*, les poids a_i sont tous égaux à un poids générique a, alors le semigroupe $(\widetilde{\mathcal{P}_{A,t}})_{t\geq 0}$ admet la représentation suivante : pour tout champ de vecteurs F suffisamment régulier de $\mathbb{R}^d$,

$$\widetilde{\mathcal{P}_{A,t}} F(x) \;=\; \mathbb{E}_{\delta_x}\left[J_{a,t}\, F(X_{a,t})\right], \quad t \geq 0, \quad x \in \mathbb{R}^d,$$

où $(J_{a,t})_{t\geq 0}$ est solution matricielle de l'équation différentielle

$$\partial_t J_{a,t} + J_{a,t}\, \left(\nabla^2 V - a\, La^{-1}\, I\right)(X_{a,t}) \;=\; 0, \quad J_{a,0} = I.$$

Par exemple dans le cas unidimensionnel, l'entrelacement à poids (3.2.2) donne lieu à une approche stochastique simple utilisant le théorème de Girsanov. En effet, l'entrelacement classique (3.1.8) puis le théorème de Girsanov entraînent l'identité

$$(P_t f)'(x) \;=\; \mathbb{E}_{\delta_x}\left[f'(X_{a,t})\, \exp\left(-\int_0^t V''(X_{a,s})\, ds\right)\, M_t^{(a)}\right],$$

où $(X_{a,t})_{t\geq 0}$ est le processus de générateur L_a et $(M_t^{(a)})_{t\geq 0}$ est la martingale exponentielle (sous quelques propriétés supplémentaires sur a et ses dérivées)

$$M_t^{(a)} \;=\; \frac{a(X_{a,t})}{a(x)}\, \exp\left(-\int_0^t \frac{L_a a(X_{a,s})}{a(X_{a,s})}\, ds\right),$$

qui, lorsque l'on considère l'opérateur L plutôt que L_a, se réécrit

$$M_t^{(a)} = \frac{a(X_{a,t})}{a(x)}\, \exp\left(+\int_0^t a(X_{a,s})\, La^{-1}(X_{a,s})\, ds\right).$$

Ainsi, on en déduit directement la représentation probabiliste de l'entrelacement à poids (3.2.2) :

$$a(x)\,(P_t f)'(x) = \mathbb{E}_{\delta_x}\left[a(X_{a,t})\, f'(X_{a,t})\, \exp\left(-\int_0^t \left(V'' - a\, La^{-1}\right)\,(X_{a,s})\, ds\right)\right].$$

3.2.2 Inégalités spectrales

Les différents objets d'intérêt à présent introduits, décrivons quelques résultats que nous avons établis dans les travaux [8, 39] à l'aide de ces entrelacements à poids, le cas unidimensionnel [37, 38, 40] étant traité plus tard. Tout d'abord, le Théorème 3.1.1 présenté auparavant, dû à Johnsen et basé sur l'entrelacement classique (3.1.7), peut être généralisé en y intégrant l'entrelacement à poids (3.2.1), la raison principale étant que la h-transformée de Doob matricielle est une opération unitaire sur les espaces appropriés. Le résultat, figurant seulement dans notre prépublication récente [39] car nous n'avons découvert l'existence de l'article de Johnsen [108] que tardivement, est le suivant.

Théorème 3.2.2. *Soit $A \in \mathrm{GL_d}(\mathbb{R})$ une matrice régulière. Sous l'hypothèse de trou spectral $\lambda_1(-L) > 0$, les opérateurs (essentiellement) auto-adjoints $-L_{|\mathbb{1}^\perp}$ et $-\widetilde{\mathcal{L}_A}_{|\nabla_A}$, restreints respectivement aux sous-espaces $\mathbb{1}^\perp$ et ∇_A, sont unitairement équivalents, la transformation unitaire pour passer de l'un à l'autre étant la multiplication par la matrice A avec la transformée de Riesz, i.e., $U_A = A\,\nabla(-L)^{-1/2} : \mathbb{1}^\perp \to \nabla_A$:*

$$-\widetilde{\mathcal{L}_A}_{|\nabla_A} \;=\; U_A\,(-L)_{|\mathbb{1}^\perp}\,U_A^*.$$

Comme précédemment, les spectres de ces deux opérateurs coïncident, à savoir

$$\sigma(-L)\backslash\{0\} \;=\; \sigma(-\widetilde{\mathcal{L}_A}_{|\nabla_A}),$$

tout comme les sous-parties constituées des spectres essentiels et discrets. En terme de trou spectral de l'opérateur $-L$, on a

$$\lambda_1(-L) \;=\; \lambda_0(-\widetilde{\mathcal{L}_A}_{|\nabla_A}).$$

Ainsi, s'intéresser au trou spectral de l'opérateur $-L$ revient à étudier le bas du spectre de l'opérateur de type Schrödinger $-\widetilde{\mathcal{L}_A}$ restreint aux gradients à poids, une analyse se révélant être moins difficile grâce à la présence du terme d'ordre 0. Comparé au théorème de Johnsen initial, l'apport du poids matriciel A offre un éventail de possibilités incomparablement plus grand pour attaquer le problème de l'estimation du trou spectral, le degré de liberté portant sur le choix de ce poids, nous permettant éventuellement de considérer une plus grande variété de potentiels V au-delà du cas uniformément convexe, voire même convexe, comme nous allons le voir maintenant. Du point de vue des inégalités fonctionnelles, l'interprétation spectrale de l'entrelacement à poids apparaissant dans le Théorème 3.2.2 peut être partiellement encodé par des inégalités de type Brascamp-Lieb [44] et de Poincaré à poids, un travail que nous avons fait dans le cadre unidimensionnel tout d'abord dans [37] puis en toute généralité dans [8]. Plus précisément, l'entrelacement à poids (3.2.1) permet d'obtenir une inégalité de Brascamp-Lieb généralisée, valable pour toute matrice régulière $A \in \mathrm{GL_d}(\mathbb{R})$ satisfaisant l'hypothèse ci-dessous, plus forte que (H_{sym}) :

$$(H_{\mathrm{pos}}) \qquad \text{la matrice } \nabla^2 V - \mathcal{L}A^{-1}\,A \text{ est symétrique définie positive,}$$

L'énoncé de notre résultat est le suivant.

Théorème 3.2.3. *Soit $A \in \mathrm{GL_d}(\mathbb{R})$ une matrice régulière satisfaisant l'hypothèse (H_{pos}). Alors l'inégalité de Brascamp-Lieb généralisée a lieu : pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$,*

$$\mathrm{Var}_\mu(f) \leq \int_{\mathbb{R}^d} \nabla f \cdot \left(\nabla^2 V - \mathcal{L} A^{-1} A\right)^{-1} \nabla f \, d\mu.$$

Preuve. La démonstration de ce résultat étant brève, proposons-la. L'idée est d'introduire le poids matriciel A dans l'identité (3.1.9) :

$$\begin{aligned}
\mathrm{Var}_\mu(f) &= \int_{\mathbb{R}^d} \nabla f \cdot (-\widetilde{\mathcal{L}})^{-1} \nabla f \, d\mu \\
&= \int_{\mathbb{R}^d} (A \nabla f) \cdot \left(A A^T\right)^{-1} A (-\widetilde{\mathcal{L}})^{-1} (A^{-1} \cdot) (A \nabla f) \, d\mu \\
&= \int_{\mathbb{R}^d} (A \nabla f) \cdot \left(A A^T\right)^{-1} (-\widetilde{\mathcal{L}_A})^{-1} (A \nabla f) \, d\mu \\
&\leq \int_{\mathbb{R}^d} (A \nabla f) \cdot \left(A A^T\right)^{-1} M_A^{-1} (A \nabla f) \, d\mu \\
&= \int_{\mathbb{R}^d} \nabla f \cdot \left(\nabla^2 V - \mathcal{L} A^{-1} A\right)^{-1} \nabla f \, d\mu,
\end{aligned}$$

car d'une part on a $(-\widetilde{\mathcal{L}_A})^{-1} F = A (-\widetilde{\mathcal{L}})^{-1} (A^{-1} F)$ pour tout $F \in \nabla_A$ et d'autre part les opérateurs de diffusion matriciel $-\mathcal{L}_A$ et multiplicatif M_A sont tous deux (essentiellement) auto-adjoints et positifs dans $L^2_A(\mu)$ (donc aussi sur ∇_A) grâce à l'hypothèse (H_{pos}). La preuve du théorème est à présent achevée. ■

Mentionnons qu'une autre démonstration du même type passe par une version renforcée de la Proposition 2.1.6 que l'on démontre facilement. L'énoncé est le suivant.

Proposition 3.2.4. *S'il existe une matrice symétrique définie positive R en chaque point de l'espace telle que pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$, on ait l'inégalité renforcée*

$$\int_{\mathbb{R}^d} \Gamma_2(f,f) \, d\mu \;\geq\; \int_{\mathbb{R}^d} \nabla f \cdot R \nabla f \, d\mu,$$

alors on a l'inégalité de type Brascamp-Lieb suivante : pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$,

$$\mathrm{Var}_\mu(f) \;\leq\; \int_{\mathbb{R}^d} \nabla f \cdot R^{-1} \nabla f \, d\mu.$$

L'idée est alors d'appliquer la Proposition 3.2.4 à la matrice $R = \nabla^2 V - \mathcal{L}(A^{-1}) A$ pour retrouver le Théorème 3.2.3. Pour ce faire, notons que l'entrelacement à poids (3.2.1) donne lieu à l'inégalité

$$\int_{\mathbb{R}^d} \Gamma_2(f,f) \, d\mu \;=\; \int_{\mathbb{R}^d} A \nabla f \cdot \left(A A^T\right)^{-1} \left(-\mathcal{L}_A + M_A\right) A \nabla f \, d\mu$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^d} \nabla(A\nabla f)\cdot \left(A\,A^T\right)^{-1}\,\nabla(A\nabla f)\,d\mu + \int_{\mathbb{R}^d} \nabla f\cdot A^{-1}\,M_A\,A\nabla f\,d\mu \\
&\geq \int_{\mathbb{R}^d} \nabla f\cdot \left(\nabla^2 V - \mathcal{L}A^{-1}\,A\right)\,\nabla f\,d\mu,
\end{aligned}
$$

où l'on a utilisé les symétrie et positivité de l'opérateur de diffusion matriciel $-\mathcal{L}_A$ sur le sous-espace ∇_A, l'hypothèse (H_{pos}) entraînant (H_{sym}). D'où le résultat désiré.

En particulier, l'intérêt d'une telle démonstration est qu'elle nous permet de voir comment la méthode de l'entrelacement à poids est utilisée pour extraire de la positivité dans l'opérateur $-\mathcal{L}$, comme suggéré à la fin de la Section 3.1.2. En effet, en écrivant par l'entrelacement classique l'intégrale du Γ_2 comme

$$
\int_{\mathbb{R}^d} \Gamma_2(f,f)\,d\mu \;=\; \int_{\mathbb{R}^d} \nabla f\cdot\left(-\mathcal{L}+\nabla^2 V\right)\,\nabla f\,d\mu,
$$

on déduit de l'inégalité précédente la proposition suivante (ci-dessous, seule l'hypothèse (H_{sym}) prévaut, l'hypothèse plus forte (H_{pos}) étant seulement requise pour revenir à l'inégalité de Brascamp-Lieb généralisée).

Proposition 3.2.5. *Soit $A \in \mathrm{GL_d}(\mathbb{R})$ une matrice régulière satisfaisant l'hypothèse (H_{sym}). Alors on a l'inégalité suivante : pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$,*

$$
\int_{\mathbb{R}^d} \nabla f\cdot(-\mathcal{L})\,\nabla f\,d\mu \;\geq\; \int_{\mathbb{R}^d} \nabla f\cdot\left(-\mathcal{L}A^{-1}\,A\right)\,\nabla f\,d\mu.
$$

Autrement dit, il s'agit d'une version du Lemme 2.2.11 apparaissant dans la Section 2.2.3 pour l'opérateur de diffusion matriciel $-\mathcal{L}$ agissant sur les gradients.

Revenons au Théorème 3.2.3 en faisant quelques commentaires sur l'inégalité de Brascamp-Lieb généralisée. Tout d'abord, le potentiel V n'est pas forcément supposé strictement convexe (voire même convexe), la présence de la matrice $-\mathcal{L}A^{-1}\,A$ permettant de corriger l'éventuel manque de convexité dans certaines régions de l'espace. L'inégalité de Brascamp-Lieb classique (3.1.1) est obtenue en choisissant la matrice $A = I$, c'est-à-dire en considérant simplement l'entrelacement classique (3.1.7) comme le fait Helffer [93], les fonctions extrémales étant données par $f = \nabla V\cdot c$, où $c \in \mathbb{R}^d$ est une constante quelconque. Dans la situation présente avec le poids A, l'optimalité d'une telle inégalité n'est pas claire. En revanche si l'on contraint A à être de la forme

$$
A \;=\; (J_H^T)^{-1},
$$

où H est un difféomorphisme suffisamment régulier de $\mathbb{R}^d$ et J désignant la matrice jacobienne, le produit scalaire $A\,A^T$ induit par l'espace $L^2_A(\mu)$ ayant alors une structure particulière,

$$
\left(A\,A^T\right)^{-1}_{i,j} \;=\; (J_H\,J_H^T)_{i,j} \;=\; \nabla H_i\cdot\nabla H_j,
$$

alors l'entrelacement classique (3.1.7) interprété au niveau matriciel, *i.e.*,

$$
J^T_{\mathcal{L}H} \;=\; (\mathcal{L}-\nabla^2 V)\,(J_H^T),
$$

nous donne alors l'identité

$$\nabla^2 V - \mathcal{L}A^{-1}\, A \;\; = \;\; -J^T_{\mathcal{L}H}\, (J^T_H)^{-1}.$$

Ainsi, on en déduit que l'inégalité de Brascamp-Lieb généralisée est saturée pour les fonctions de la forme $f = \mathcal{L}H \cdot c$ avec $c \in \mathbb{R}^d$ une constante, dès lors que ce choix de difféomorphisme H satisfait (H_{pos}), *i.e.*, tel que la matrice $-J^T_{\mathcal{L}H}\, (J^T_H)^{-1}$ soit symétrique définie positive. Dans le cas de l'inégalité de Brascamp-Lieb classique, ce choix correspond à $H(x) = x$, $x \in \mathbb{R}^d$.

En pratique, on peut tenter de déterminer la matrice A en fonction d'un difféomorphisme bien choisi H, ou alors tester directement une matrice A diagonale car d'une part la condition (H_{sym}) est forcément satisfaite, et d'autre part l'hypothèse (H_{pos}) est souvent plus facile à vérifier. Dans le cas où le modèle considéré comporte certaines symétries, comme dans le cas radial, la matrice A peut même être un multiple de l'identité, $A = a\, I$ où a est une fonction suffisamment régulière sur $\mathbb{R}^d$ et ne s'annulant pas. Par exemple si l'on s'intéresse aux lois de Subbotin de paramètre $\alpha > 1$, alors en calculant la plus petite valeur propre de la matrice $\nabla^2 V$, l'inégalité de Brascamp-Lieb classique entraîne l'inégalité de Poincaré à poids suivante : pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$,

$$\mathrm{Var}_\mu(f) \;\; \leq \;\; \int_{\mathbb{R}^d} \frac{|\nabla f(x)|^2}{\min\{1, \alpha - 1\}\; |x|^{\alpha-2}}\, \mu(dx), \tag{3.2.3}$$

tandis qu'un bon choix de fonction poids a nous permet d'en déduire une autre inégalité de Poincaré à poids grâce à l'inégalité de Brascamp-Lieb généralisée du Théorème 3.2.3 : il existe une constante $C_{d,\alpha} > 0$ explicite ne dépendant que de α et de la dimension d (≥ 2) telle que pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$, on ait

$$C_{d,\alpha}\, \mathrm{Var}_\mu(f) \;\; \leq \;\; \int_{\mathbb{R}^d} \frac{|\nabla f(x)|^2}{\min\{1, \alpha - 1\}\; |x|^{\alpha-2} + |x|^{2(\alpha-1)}}\, \mu(dx). \tag{3.2.4}$$

Ci-dessus, la constante se comporte comme $C_{d,\alpha} \approx \min\{1, \alpha - 1\}/d$ pour d assez grand. Dans le cas gaussien standard, la première inégalité de Poincaré à poids (3.2.3) est réduite à l'inégalité de Poincaré classique tandis que la seconde inégalité (3.2.4) donne avec la constante $C_{d,\alpha}$ écrite explicitement,

$$\frac{8\sqrt{d-1}}{\left(\sqrt{d-1} + \sqrt{d+3}\right)^3}\, \mathrm{Var}_\mu(f) \;\; \leq \;\; \int_{\mathbb{R}^d} \frac{|\nabla f(x)|^2}{1 + |x|^2}\, \gamma(dx). \tag{3.2.5}$$

En particulier, on observe en testant sur les fonctions linéaires que cette inégalité se révèle asymtotiquement optimale en grande dimension.

En terme de trou spectral, l'inégalité de Brascamp-Lieb classique donne lieu à une borne inférieure sur $\lambda_1(-L)$ seulement dans le cadre d'un potentiel uniformément convexe, comme on l'a vu précédemment. En revanche, utiliser l'inégalité de Brascamp-Lieb généralisée du Théorème 3.2.3 permet dans les exemples de relâcher cette hypothèse de convexité uniforme (voire même de convexité tout court), sous réserve de nouveau que l'on sache choisir judicieusement la matrice A. On rappelle que ρ désigne la fonction correspondant à la plus petite valeur propre de la matrice symétrique considérée.

Théorème 3.2.6. *On a l'estimation suivante sur le trou spectral :*

$$\lambda_1(-L) \quad \geq \quad \sup_H \inf_{\mathbb{R}^d} \rho\left(-J_{\mathcal{L}H}^T\,(J_H^T)^{-1}\right),$$

où le supremum est pris sur l'ensemble des difféomorphismes H suffisamment réguliers de $\mathbb{R}^d$ satisfaisant l'hypothèse (H_{pos}).

Bien évidemment, cette borne inférieure sur le trou spectral n'est intéressante qu'à partir du moment où cette fonction ρ est minorée par une constante strictement positive. Ce résultat est la version multidimensionnelle de la célèbre formule de Chen-Wang (3.1.5) établie sur la droite réelle [65], faisant clairement ressortir les cas d'optimalité. Présentée comme telle (*i.e.*, avec le difféomorphisme H plutôt que la matrice A), notre borne sur le trou spectral révèle elle-aussi les cas potentiels d'optimalité, dont les conditions sont manifestement plus difficiles à établir. En effet, l'égalité a lieu dès lors que l'espace propre $E_{\lambda_1(-L)}$ associé au trou spectral, lorsqu'il existe, est de dimension d et les fonctions propres associées forment un difféomorphisme H de $\mathbb{R}^d$ suffisamment régulier (auquel cas H vérifie bien la condition (H_{pos})). En-dehors du cas trivial des opérateurs associés à une même mesure unidimensionnelle tensorisées d fois (auquel cas le trou spectral $\lambda_1(-L)$ vaut exactement celui en dimension 1, les d fonctions propres associées sur chacune des coordonnées formant alors une base de $E_{\lambda_1(-L)}$), un exemple intéressant concernant le problème de la multiplicité du trou spectral a été mis récemment en exergue par Barthe et Klartag [18]. En effet, si le potentiel V est strictement convexe et possède les symétries du cube, c'est-à-dire que V est invariant par toute permutation et tout changement de signe des coordonnées, alors $\dim E_{\lambda_1(-L)} = d$. Ce résultat repose en partie sur une observation de Barthe et Cordero-Erausquin [16] généralisant aux potentiels V strictement convexes sur $\mathbb{R}^d$ le résultat originel de Klartag [118] pour le laplacien de Neumann sur un corps convexe, énonçant que l'application qui à toute fonction propre $f \in E_{\lambda_1(-L)}$ associe la quantité $\int_{\mathbb{R}^d} \nabla f \, d\mu \in \mathbb{R}^d$ est injective, avec pour conséquence immédiate que la dimension de l'espace propre $E_{\lambda_1(-L)}$ est au plus d.

Pour appréhender la pertinence du Théorème 3.2.6, intéressons-nous au cas de mesures de probabilité produits perturbées par un potentiel d'interaction possiblement non convexe, une situation que nous avons étudiée dans l'article [39]. Comme déjà évoqué auparavant, on considère un potentiel V de la forme

$$V(x) \quad = \quad \sum_{i=1}^d U_i(x_i) + \varphi(x), \qquad x \in \mathbb{R}^d,$$

où les fonctions unidimensionnelles suffisamment régulières $U_i : \mathbb{R} \to \mathbb{R}$ sont telles que les e^{-U_i} soient intégrables par rapport à la mesure de Lebesgue sur $\mathbb{R}$ et le potentiel d'interaction $\varphi : \mathbb{R}^d \to \mathbb{R}$ est suffisamment régulier et tel que la mesure μ soit une mesure de probabilité. Notons que l'on ne suppose pas les U_i convexes. Si l'on désire appliquer le Théorème 3.2.6, il nous faut trouver un bon difféomorphisme H sur $\mathbb{R}^d$ satisfaisant l'hypothèse (H_{pos}) et même un peu plus : l'infimum sur $\mathbb{R}^d$ de la fonction $\rho\left(-J_{\mathcal{L}H}^T\,(J_H^T)^{-1}\right)$ doit être strictement positif. L'idée est alors de choisir H comme un champ de vecteurs « diagonal », la i-ème coordonnée ne dépendant que de la variable x_i, auquel cas la matrice sous-jacente A devient diagonale et l'hypothèse

$(H_{\rm sym})$ est automatiquement satisfaite. D'une certaine manière, le problème se réduit alors à une étude fine unidimensionnelle, ce choix de difféomorphisme H étant en adéquation avec la forme « presque produit » du problème, sous réserve que le potentiel d'interaction ne chahute pas trop la structure produit de la mesure μ. On obtient alors le résultat suivant. Ci-dessous, la notation ∂_{x_i} désigne la dérivée partielle en la variable x_i.

Théorème 3.2.7. *Supposons qu'il existe des constantes* $\varepsilon_1, \ldots, \varepsilon_d \in [0,1[$ *et* $\rho > 0$ *telles que l'on ait*

$$\nabla^2\varphi(x) + diag\left\{(1-\varepsilon_i)\left(U_i''(x_i) + \varepsilon_i\, U_i'(x_i)^2\right) + \varepsilon_i\, \partial_{x_i}\varphi(x)\, U_i'(x_i)\right\} \ \geq \ \rho\, I. \qquad (3.2.6)$$

Alors on a la borne inférieure suivante sur le trou spectral :

$$\lambda_1(-L) \ \geq \ \rho.$$

Ainsi, sous certaines hypothèses sur la matrice hessienne $\nabla^2\varphi$, les dérivées U_i', U_i'' et enfin sur le signe des produits $\partial_{x_i}\varphi(x)\, U_i'(x_i)$, on est en mesure de minorer convenablement le trou spectral une fois l'optimisation en les paramètres ε_i effectuée (notons que le cas $\varepsilon_i = 0$ pour tout $i = 1, \ldots, d$, correspond à un potentiel V uniformément convexe). Voyons ce que ce résultat quelque peu abstrait donne sur deux exemples intéressants, manifestement non couverts par les différents travaux précédemment cités reposant sur la méthode d'Helffer, requérant des trous spectraux uniformes pour les lois conditionnelles unidimensionnelles : le modèle gaussien standard perturbé par un potentiel d'interaction quartique et la loi de Subbotin de paramètre $\alpha \in [1,2]$ perturbée par un potentiel d'interaction convexe et lipschitzien.

∘ le modèle gaussien standard perturbé par un potentiel d'interaction quartique : le potentiel V est de la forme

$$V(x) \ = \ \sum_{i=1}^{d} \underbrace{\frac{x_i^2}{2}}_{U_i(x_i)} + \underbrace{J\sum_{i=1}^{d} x_i^2\, x_{i+1}^2}_{\varphi(x)}, \quad x \in \mathbb{R}^d,$$

où par convention $x_{d+1} = x_1$ et le coefficient d'interaction J est positif ou nul. Le spectre de l'opérateur associé $-L$ est discret car il satisfait le second critère d'Helffer-Nier présenté dans la Section 3.1.1. Néanmoins, il n'est pas difficile de voir que le potentiel V n'est pas convexe (il en est même très loin, sa matrice hessienne n'étant pas minorée uniformément en espace). De surcroît, si l'on regarde attentivement ce modèle, on s'aperçoit que la loi unidimensionnelle de la i-ème coordonnée sous la mesure μ, les autres étant fixées égales à $\hat{x_i}$ (on rappelle que $\hat{x_i}$ désigne le vecteur $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d)$ de $\mathbb{R}^{d-1}$ issu de x en enlevant la i-ème coordonnée), est gaussienne centrée et de variance $1/(1+2J\,(x_{i-1}^2 + x_{i+1}^2))$, auquel cas le trou spectral associé $\lambda_1^{\hat{x_i}}$ vaut exactement $\lambda_1^{\hat{x_i}} = 1 + 2J\,(x_{i-1}^2 + x_{i+1}^2)$, l'inverse de la variance. Ainsi, non seulement l'estimation d'Helffer (3.1.6) ne peut être appliquée dans ce cadre, mais c'est aussi le cas pour son raffinement,

$$\lambda_1(-L) \ \geq \ \inf_{x\in\mathbb{R}^d} \rho\left(\widetilde{\nabla^2 V}(x) + \mathrm{diag}\,(\lambda_1^{\hat{x_i}})\right), \qquad (3.2.7)$$

obtenu par l'approche de type Brascamp-Lieb figurant dans la Proposition 3.2.4, cf. [16, 18] et aussi l'article plus ancien [63] (ci-dessus, on rappelle que la matrice $\widetilde{\nabla^2 V}$ désigne la matrice hessienne de V remplie de zéros sur la diagonale). En effet, la matrice apparaissant dans l'inégalité (3.2.7) coïncide justement avec la matrice $\nabla^2 V$. Néanmoins, on est en mesure d'appliquer le Théorème 3.2.7 pour un coefficient d'interaction J suffisamment petit.

∘ la loi de Subbotin de paramètre $\alpha \in [1, 2]$ perturbée par un potentiel d'interaction convexe lipschitzien : le potentiel V est cette fois de la forme

$$V(x) \quad = \quad \sum_{i=1}^{d} \underbrace{\frac{|x_i|^\alpha}{\alpha}}_{U_i(x_i)} + \underbrace{J \sum_{i=1}^{d} |x_i - x_{i+1}|}_{\varphi(x)}, \quad x \in \mathbb{R}^d,$$

où par convention $x_{d+1} = x_1$ et $J \geq 0$. Par le second critère d'Helffer-Nier, le spectre de l'opérateur est discret pour $\alpha \in]1, 2]$, le cas $\alpha = 1$ vérifiant seulement l'existence d'un trou spectral et contenant une partie de spectre continu, tout comme le cas sans interaction $J = 0$ déjà évoqué. Le potentiel d'interaction est convexe lipschitzien (la dégénérescence sur les diagonales n'est pas vraiment un problème, quitte à régulariser). Déterminer explicitement les trous spectraux des lois conditionnelles unidimensionnelles, voire même seulement les estimer convenablement en fonction des coordonnées gelées, paraît bien difficile. De surcroît, une autre difficulté en dehors du cas gaussien $\alpha = 2$ réside dans l'absence de convexité uniforme à l'infini, un point important dans les travaux menés jusqu'à présent pour ces modèles de mesures produits perturbées. Néanmoins, le Théorème 3.2.7 peut s'appliquer là-aussi pour un coefficient d'interaction J suffisamment petit. Récapitulons les résultats que nous avons obtenus pour ces deux exemples.

Théorème 3.2.8. *Considérons le modèle gaussien standard perturbé par un potentiel d'interaction quartique ou la loi de Subbotin de paramètre $\alpha \in [1, 2]$ perturbée par un potentiel d'interaction convexe lipschitzien. Alors il existe deux constantes explicites $J_0 > 0$ et $C > 0$ telles que pour tout $J \in [0, J_0]$, on ait*

$$\lambda_1(-L) \quad \geq \quad C.$$

Dans chacun de ces deux cas, les constantes J_0 et C sont des constantes numériques universelles indépendantes de la dimension.

Pour terminer notre étude du trou spectral par cette approche par entrelacement à poids dans le cadre multidimentionnel, exposons un résultat que nous avons obtenu dans l'article [8], quelque peu dans le même esprit que ce qui précède. Plus précisément, nous étendons le résultat de Veysseire [174] mentionné *via* l'inégalité (3.1.3).

Théorème 3.2.9. *Soit $A \in \mathrm{GL_d}(\mathbb{R})$ une matrice régulière vérifiant (H_{pos}). De surcroît, si la matrice $(A\,A^T)^{-1}$ est minorée et majorée uniformément (au sens des matrices symétriques) respectivement par deux constantes strictement positives α et β, alors le trou spectral vérifie l'inégalité suivante :*

$$\lambda_1(-L) \quad \geq \quad \left(\int_{\mathbb{R}^d} \frac{1}{\rho_A}\, d\mu + \frac{\left(1 - \frac{\alpha}{\beta}\right)}{\inf\limits_{\mathbb{R}^d} \rho_A} \right)^{-1},$$

où $\rho_A : \mathbb{R}^d \to]0,\infty[$ *désigne la plus petite valeur propre de la matrice* $\nabla^2 V - \mathcal{L}A^{-1}\,A$.

Grâce à la présence du poids matriciel A, il s'agit-là d'une sorte de perturbation du résultat obtenu par Veysseire, au sens où cette inégalité peut être appliquée à des potentiels V non convexes, le prix à payer étant d'avoir un terme supplémentaire faisant intervenir la borne inférieure uniforme sur ρ_A dont on attend qu'elle soit strictement positive (sinon notre inégalité n'apporte aucune information). En l'absence de poids, *i.e.*, $A = I$, le second terme dans la parenthèse n'a plus lieu d'être et nous retrouvons par la méthode des entrelacements son inégalité (3.1.3), l'hypothèse (H_{pos}) étant réduite à la stricte convexité du potentiel V.

À présent, intéressons-nous aux valeurs propres supérieures. La version à poids du théorème de Johnsen comportant la totalité de l'information spectrale, nous avons récemment poussé plus loin la méthode des entrelacements à poids en nous consacrant à l'estimation des valeurs propres supérieures de l'opérateur $-L$ situées en dessous (voire même correspondant au bas) du spectre essentiel. Dans le cadre multidimensionnel, nous proposons dans le papier [39] une borne inférieure sur la $(d+1)$-ème valeur propre strictement positive, notée $\lambda_{d+1}(-L)$. Nos travaux englobant des situations pour lesquelles le potentiel V est strictement convexe, il n'est donc pas étonnant, pour les raisons de multiplicité évoquées auparavant, que la valeur propre pertinente à considérer après le trou spectral soit justement $\lambda_{d+1}(-L)$, les valeurs propres étant comptées avec leur multiplicité. L'idée repose sur l'établissement d'inégalités de Brascamp-Lieb généralisées du second ordre. Tout d'abord, rappelons que les inégalités de Brascamp-Lieb ont connu un regain d'intérêt récemment, avec de nombreuses tentatives d'améliorations dans des cadres divers et variés, comme en témoignent les articles sur le sujet [30, 34, 92, 120, 121, 154] reposant sur des méthodes de preuve assez différentes. En particulier, Cordero-Erausquin a utilisé dans l'article [68] une approche par le transport optimal pour établir l'inégalité de Brascamp-Lieb du second ordre suivante pour un potentiel V strictement convexe : pour toute fonction réelle f suffisamment régulière telle que $f \perp x_i$ pour tout $i = 1, \ldots, d$, on a

$$\mathrm{Var}_\mu(f) \;\leq\; \int_{\mathbb{R}^d} \nabla f \cdot \left(\nabla^2 V + \lambda_1(-L)\, I\right)^{-1} \nabla f \, d\mu.$$

En particulier, si le potentiel V est uniformément convexe, alors par le théorème du min-max de Courant-Fisher, on obtient la borne suivante :

$$\lambda_{d+1}(-L) \;\geq\; \lambda_1(-L) + \inf_{\mathbb{R}^d} \rho(\nabla^2 V),$$

qui est optimale dans le cas gaussien standard, pour lequel on a $\lambda_{d+1}(-L) = 2$. Cette approche spectrale avait déjà été utilisée dans ce cadre gaussien par Cordero-Erausquin, Fradelizi et Maurey [69], faisant le lien avec la fameuse (B)-conjecture (un problème qu'avait posé Banaszczyk, popularisé ensuite par Latala lors de son exposé donné au Congrès International des Mathématiciens de Pékin en 2002). Dès lors que l'on désire sortir du cadre uniformément convexe, l'idée est de reprendre l'approche par l'entrelacement à poids des inégalités de Brascamp-Lieb généralisées et de voir comment certaines améliorations peuvent amener une information pertinente sur cette plus grande valeur propre. Le problème de l'estimation des plus grandes valeurs propres a une longue histoire, sans doute aussi longue que celle du trou spectral, les liens qu'elles entretiennent avec le noyau du semigroupe sous-jacent par la formule de trace étant connus depuis longtemps, notamment dans le cas du laplacien de Neumann sur une variété riemannienne compacte à courbure de Ricci positive ou nulle, cf. par exemple [138]. En 2000, ce problème a été étudié par Wang [179] sous l'angle des inégalités dites de super-Poincaré qu'il venait tout juste d'introduire. Notons par ailleurs le travail récent de Milman [150] procédant à une comparaison des valeurs propres en termes d'espaces modèles. Néanmoins, le résultat que nous allons exposer dans un instant n'est pas réellement comparable à ceux contenus dans chacune des références mentionnées ci-dessus, les hypothèses requises dans ces dernières et les approches utilisées étant de nature bien différente des nôtres.

Soit $A \in \mathrm{GL_d}(\mathbb{R})$ une matrice régulière vérifiant l'hypothèse (H_{sym}). Commençons par définir la notion de trou spectral pour l'opérateur de diffusion matriciel $-\mathcal{L}_A$ restreint au sous-espace ∇_A des gradients à poids (on rappelle que sous l'hypothèse (H_{sym}), cet opérateur est (essentiellement) auto-adjoint et positif sur $L^2_A(\mu)$, donc sur ∇_A). Pour ce faire, on définit $\int_{\mathbb{R}^d} \left(A\,A^T\right)^{-1} d\mu$ la matrice dont les entrées sont l'intégrale sous μ des entrées de la matrice $\left(A\,A^T\right)^{-1}$, sous réserve que ces intégrales soient toutes bien définies. Si cette matrice est inversible, alors la bonne notion de moyenne dans l'espace $L^2_A(\mu)$ est la suivante : pour tout champ de vecteurs $F \in L^2_A(\mu)$,

$$m_A(F) \quad = \quad \left(\int_{\mathbb{R}^d} \left(A\,A^T\right)^{-1} d\mu\right)^{-1} \int_{\mathbb{R}^d} \left(A\,A^T\right)^{-1} F\, d\mu,$$

car c'est l'unique vecteur constant de $\mathbb{R}^d$ minimisant la fonctionnelle

$$c \in \mathbb{R}^d \quad \to \quad \int_{\mathbb{R}^d} (F-c)\cdot \left(A\,A^T\right)^{-1} (F-c)\, d\mu.$$

En particulier, tout vecteur constant $c \in \mathbb{R}^d$ appartient à $L^2_A(\mu)$ et l'on a $m_A(c) = c$. Si le champ de vecteurs $F \in L^2_A(\mu)$ est tel que $m_A(F) = 0$, cela signifie que F est orthogonal aux constantes de $\mathbb{R}^d$ dans l'espace $L^2_A(\mu)$.

De tout ce qui précède, nous pouvons en déduire que sous les hypothèses annoncées plus haut, la première valeur propre $\lambda_0(-\mathcal{L}_A)$ de l'opérateur de diffusion matriciel $-\mathcal{L}_A$ est nulle, les fonctions propres associées étant les vecteurs constants de $\mathbb{R}^d$. De même, la première valeur propre $\lambda_0(-\mathcal{L}_{A\,|\nabla_A})$ de l'opérateur $-\mathcal{L}_A$ restreint au sous-espace ∇_A des gradients à poids vaut 0. Ainsi, ceci nous amène à considérer la valeur propre suivante, c'est-à-dire le trou spectral de l'opérateur de diffusion matriciel $-\mathcal{L}_{A\,|\nabla_A}$ restreint au sous-espace ∇_A :

$$\lambda_1(-\mathcal{L}_{A\,|\nabla_A}) \quad = \quad \inf\left\{ \frac{\int_{\mathbb{R}^d} F\cdot \left(A\,A^T\right)^{-1} (-\mathcal{L}_A F)\, d\mu}{\int_{\mathbb{R}^d} F\cdot (A\,A^T)^{-1}\, F\, d\mu} : F \in \nabla_A;\, m_A(F) = 0 \right\}.$$

À présent, nous sommes en mesure d'énoncer l'inégalité de Brascamp-Lieb généralisée du second ordre. Dans le résultat ci-dessous, nous conservons à la fois les éléments A et H pour simplifier la notation, le lien entre les deux étant de nouveau l'identité $A = (J_H^T)^{-1}$. On rappelle qu'alors on a $\left(A\,A^T\right)^{-1} = J_H\,J_H^T$ et $\nabla^2 V - \mathcal{L}A^{-1}\,A = -J_{\mathcal{L}H}^T\,(J_H^T)^{-1}$.

Théorème 3.2.10. *Soit H un difféomorphisme suffisamment régulier de $\mathbb{R}^d$, que l'on suppose d'une part vérifiant l'hypothèse (H_{pos}) et d'autre part tel que la matrice $\int_{\mathbb{R}^d} J_H\, J_H^T\, d\mu$ ait toutes* *ses entrées bien définies et soit inversible. Alors pour toute fonction réelle f suffisamment régulière sur $\mathbb{R}^d$ telle que $f \perp H_i$ pour tout $i = 1, \ldots, d$, on a l'inégalité*

$$\mathrm{Var}_\mu(f) \quad \leq \quad \int_{\mathbb{R}^d} \nabla f \cdot \left(\lambda_1(-\mathcal{L}_{A\,|\nabla_A})\, I - J_{\mathcal{L} H}^T\, (J_H^T)^{-1}\right)^{-1} \nabla f \, d\mu.$$

Si de plus la matrice $-J_{\mathcal{L} H}^T\, (J_H^T)^{-1}$ est minorée uniformément par une constante strictement positive, alors on a l'estimation suivante sur la valeur propre plus grande

$$\lambda_{d+1}(-L) \quad \geq \quad \lambda_1(-\mathcal{L}_{A\,|\nabla_A}) + \inf_{\mathbb{R}^d} \rho\left(-J_{\mathcal{L} H}^T\, (J_H^T)^{-1}\right).$$

Le point clé de cette démonstration réside dans l'égalité

$$\begin{aligned} m_A(A\,\nabla(-L)^{-1} f) \quad &= \quad \int_{\mathbb{R}^d} \left(A\, A^T\right)^{-1} A\,\nabla(-L)^{-1} f \, d\mu \\ &= \quad \int_{\mathbb{R}^d} \left(A^T\right)^{-1} \nabla(-L)^{-1} f \, d\mu \\ &= \quad \int_{\mathbb{R}^d} J_H\, \nabla(-L)^{-1} f \, d\mu \\ &= \quad \int_{\mathbb{R}^d} H\,(-L)\,(-L)^{-1} f \, d\mu \\ &= \quad \int_{\mathbb{R}^d} H\, f \, d\mu. \end{aligned}$$

Ainsi, la condition d'orthogonalité sur la fonction (centrée) f se révèle être équivalente à la condition de centrage $m_A(A\,\nabla(-L)^{-1} f) = 0$ qui émerge naturellement par l'approche L^2 lorsque l'on désire contrôler certaines quantités par le trou spectral $\lambda_1(-\mathcal{L}_{A\,|\nabla_A})$. Par ailleurs, notons que l'optimalité est de nouveau atteinte dans le Théorème 3.2.10 lorsque l'espace propre $E_{\lambda_1(-L)}$ associé au trou spectral, lorsqu'il existe, est de dimension d et les fonctions propres associées forment un difféomorphisme H de $\mathbb{R}^d$ suffisamment régulier. En effet, on a obtient alors l'inégalité

$$\lambda_{d+1}(-L) \quad \geq \quad \lambda_1(-\mathcal{L}_{A\,|\nabla_A}) + \lambda_1(-L),$$

l'égalité étant justifiée par le Théorème 3.2.2. Notons que l'opérateur $-\mathcal{L}_A$ dépend alors de ces fonctions propres. Enfin, observons que sous nos hypothèses, il se peut que le trou spectral $\lambda_1(-\mathcal{L}_{A\,|\nabla_A})$ soit nul, auquel cas le Théorème 3.2.10 ne nous apporte aucune information supplémentaire par rapport aux Théorèmes 3.2.3 et 3.2.6. Ainsi, il est attendu en pratique que ce trou spectral soit strictement positif, soulignant alors la pertinence du Théorème 3.2.10.

Pour conclure sur la présentation des deux articles [8, 39] utilisant la méthode des entrelacements à poids pour établir quelques inégalités spectrales, reprenons les deux exemples de mesures produits perturbées que nous avons étudiés précédemment pour le trou spectral. Une analyse quelque peu similaire à celle ayant débouché sur le Théorème 3.2.7 mais néanmoins plus approfondie nous a permis de montrer le résultat suivant, figurant dans [39].

Théorème 3.2.11. *Considérons le modèle gaussien standard perturbé par un potentiel d'interaction quartique ou la loi de Subbotin de paramètre* $\alpha \in]1,2]$ *perturbée par un potentiel d'interaction convexe lipschitzien. Alors il existe deux constantes explicites* $\widetilde{J}_0 > 0$ *et* $\widetilde{C} > 0$ *telles que pour tout* $J \in [0, \widetilde{J}_0]$*, on ait*

$$\lambda_{d+1}(-L) \quad \geq \quad \widetilde{C}.$$

Dans chacun des deux cas, les constantes $\widetilde{J}_0$ *et* $\widetilde{C}$ *sont indépendantes de la dimension, satisfaisant* $\widetilde{J}_0 < J_0$ *et* $\widetilde{C} > C$*, où* J_0 *et* C *sont les constantes apparaissant dans le Théorème 3.2.8.*

La conclusion de ce résultat peut être interprétée de la manière suivante : le $\widetilde{C} > C$ signifie que l'on obtient une borne sur la valeur propre $\lambda_{d+1}(-L)$ meilleure que celle obtenue directement par le Théorème 3.2.8 et l'inégalité triviale $\lambda_{d+1}(-L) \geq \lambda_1(-L)$. Cependant, le prix à payer pour une telle amélioration est de réduire strictement l'intervalle dans lequel vit le coefficient d'interaction J. Notons que contrairement au Théorème 3.2.8, le cas $\alpha = 1$ dans le deuxième exemple n'est pas couvert par ce résultat, ou plutôt si : on obtiendrait alors $\widetilde{J}_0 = 0$ et $\widetilde{C} = C$, un résultat tout à fait attendu car le trou spectral correspondant au bas du spectre essentiel, on a donc $\lambda_{d+1}(-L) = \lambda_1(-L)$. Enfin, mentionnons également que pour ces deux exemples, les constantes des Théorèmes 3.2.8 et 3.2.11 appliqués aux potentiels sans interaction $J = 0$ et avec $\alpha = 2$ dans le second cas sont alors optimales car il s'agit-là du cas gaussien standard pour lequel nous retrouvons les constantes $\lambda_1(-L) = 1$ et $\lambda_{d+1}(-L) = 2$.

3.2.3 Le cas des mesures radiales

Dans cette partie, nous nous focalisons plus précisément sur le cas particulier des mesures radiales. On rappelle que le potentiel V est radial lorsqu'il est invariant par rotation : il ne dépend alors que du rayon et s'écrit $V(x) = U(|x|)$ où $U : \mathbb{R}_+ \to \mathbb{R}$. Ainsi, le potentiel V est convexe si et seulement si la fonction U est convexe croissante sur $\mathbb{R}_+$. Beaucoup d'exemples classiques de mesures invariantes de dynamiques markoviennes rentrent dans ce cadre radial, auxquels nous avons déjà eu affaire dans les parties précédentes : la loi uniforme sur la boule unité $\mathcal{B}_d$ de $\mathbb{R}^d$, où le potentiel V est nul et les conditions au bord sont celles de Neumann, les lois de Subbotin de paramètre $\alpha \geq 1$, mais aussi les lois de Cauchy généralisées ou encore le double-puits, ces deux derniers exemples ne donnant pas lieu à des mesures log-concaves. Dans l'article écrit avec Michel Bonnefont et Yutao Ma [41] paru en 2016, nous estimons le trou spectral $\lambda_1(-L)$ pour des potentiels radiaux sur $\mathbb{R}^d$, $d \geq 2$, éventuellement non convexes. L'idée est d'adapter l'approche proposée par Bobkov dans [24] lui ayant permis de montrer que les mesures log-concaves radiales vérifiaient la conjecture KLS, avec de surcroît la borne (3.1.4). Plus précisément, en décomposant tout élément de $\mathbb{R}^d$ en le produit de ses parties radiale et angulaire, il s'agit de tirer parti du fait que les mesures radiales sont « presque » des mesures produits des lois sous-jacentes radiale et uniforme sur la sphère unité $\mathbb{S}^{d-1}$ de $\mathbb{R}^d$, et d'utiliser l'information sur leur trou spectral respectif. Ci-dessous, l'opérateur L_{rad} désigne la dynamique radiale unidimensionnelle associée, *i.e.*,

$$L_{\text{rad}} f(r) \quad = \quad f''(r) - \left(U'(r) - \frac{d-1}{r} \right) f'(r), \quad r > 0,$$

dont la mesure invariante et réversible ν sur $]0,\infty[$ a sa densité de Lebesgue proportionnelle à la fonction $r \to r^{d-1}\, e^{-U(r)}$. Le résultat est le suivant.

Théorème 3.2.12. *Les inégalités suivantes entre trous spectraux ont lieu :*

$$\min\left\{\lambda_1(-L_{\mathrm{rad}}), \frac{d-1}{\int_{\mathbb{R}^d} |x|^2\, \mu(dx)}\right\} \;\leq\; \lambda_1(-L) \;\leq\; \min\left\{\lambda_1(-L_{\mathrm{rad}}), \frac{d}{\int_{\mathbb{R}^d} |x|^2\, \mu(dx)}\right\}.$$

En particulier, lorsque la mesure μ admet un moment d'ordre deux, l'existence d'un trou spectral pour $-L$ est assurée si et seulement si elle l'est pour la partie radiale $-L_{\mathrm{rad}}$. Ci-dessus, le facteur $d-1$ dans la borne de gauche n'est rien d'autre que le trou spectral du laplacien sur la sphère $\mathbb{S}^{d-1}$, tandis que la borne de droite est trivialement obtenue en testant sur les fonctions radiales et linéaires la formule variationnelle définissant le trou spectral.

Dans le cas où la mesure μ est log-concave, l'inégalité de Veysseire (3.1.3) appliquée à la dynamique radiale nous permet de conclure que l'on a toujours l'inégalité

$$\lambda_1(-L_{\mathrm{rad}}) \;\;\geq\;\; \frac{d-1}{\int_{\mathbb{R}^d} \|x\|^2 \mu(dx)},$$

la mesure ν sur la demi-droite réelle étant « plus log-concave » que la mesure μ sur $\mathbb{R}^d$, une observation déjà exploitée sous une autre forme par Bobkov dans son article. Ainsi, dans le cadre log-concave nous tirons de cette inégalité et du Théorème 3.2.12 le résultat suivant.

Théorème 3.2.13. *Dans le cas log-concave, on a l'encadrement :*

$$\frac{d-1}{\int_{\mathbb{R}^d} \|x\|^2\, \mu(dx)} \;\leq\; \lambda_1(-L) \;\leq\; \frac{d}{\int_{\mathbb{R}^d} \|x\|^2\, \mu(dx)}.$$

Bien que qualitativement équivalent au résultat de Bobkov, ce résultat l'améliore numériquement et est asymptotiquement optimal en grande dimension. Par exemple, pour les lois de Subbotin de paramètre $\alpha \geq 1$, on obtient les inégalités

$$\frac{(d-1)\, \Gamma(d/\alpha)}{\alpha^{2/\alpha}\, \Gamma((d+2)/\alpha)} \;\leq\; \lambda_1(-L) \;\leq\; \frac{d\, \Gamma(d/\alpha)}{\alpha^{2/\alpha}\, \Gamma((d+2)/\alpha)},$$

donnant en grande dimension $\lambda_1(-\mathcal{L}) \approx d^{1-2/\alpha}$. Pour la mesure uniforme sur la boule unité $\mathcal{B}_d$ de $\mathbb{R}^d$, la dynamique sous-jacente correspondant au mouvement brownien vivant dans $\mathcal{B}_d$ et réfléchi au bord (encodant les condition de Neumann au bord pour le laplacien), on obtient l'encadrement

$$\frac{(d-1)\,(d+2)}{d} \;\leq\; \lambda_1(-L) \;\leq\; d+2,$$

d'où le trou spectral vaut approximativement d. Sa valeur exacte, connue depuis longtemps et que l'on peut trouver par exemple dans un article de Szegö ou de Weinberger, cf. [180], est donnée par le carré du premier zéro strictement positif de la dérivée (en un certain sens) d'une fonction de Bessel de première espèce ; en particulier, elle n'est pas du tout explicite, d'où l'intérêt de proposer une borne non asymptotique en la dimension.

Discutons maintenant du trou spectral de la dynamique radiale unidimensionnelle L_{rad} associée à la loi de Subbotin de paramètre $\alpha \geq 1$. Manifestement, sa valeur exacte n'est pas accessible en général, sauf dans le cas gaussien standard pour lequel on a $\lambda_1(-L_{\mathrm{rad}}) = 2$, correspondant à la fonction propre $f(r) = r^2 - d$. Néanmoins, son comportement par rapport à la dimension peut être étudié et l'on trouve $\lambda_1(-L_{\mathrm{rad}}) \approx d^{1-2/\alpha}$ lorsque d est grand, tout comme le trou spectral $\lambda_1(-L)$. Dans le cas $\alpha = 1$, cette estimation concorde avec la borne $\lambda_1(-L_{\mathrm{rad}}) \geq 1/4d$ obtenue en utilisant le fait que ν est la loi Gamma de paramètres d et 1 qui est la convolution de d lois exponentielles de même paramètre 1, pour lesquelles on sait que le trou spectral vaut 1/4. Enfin, pour la mesure uniforme sur $\mathcal{B}_d$, la mesure ν a pour densité de Lebesgue $d\,r^{d-1}$ sur $]0,1[$, pour laquelle on démontre que $\lambda_1(-L_{\mathrm{rad}}) \approx d^2$, un comportement donc différent de $\lambda_1(-L)$ qui est d'ordre d. Pour ce dernier exemple, cette différence avait déjà été observée par Bobkov et Ledoux dans [23, 128].

L'un des points forts de la méthode utilisée pour démontrer le Théorème 3.2.12 est qu'elle est suffisamment robuste pour être généralisée à des mesures radiales pour lesquelles il n'y a pas de trou spectral, *i.e.*, $\lambda_1(-L) = 0$, telles que les mesures à queue lourde qui sont loin d'être log-concaves. L'idée est alors de modifier convenablement l'opérateur $-L$ en rectifiant la métrique naturelle associée à l'opérateur carré du champ de la même manière que dans la Section 2.2.3. Si σ désigne une fonction régulière strictement positive sur $\mathbb{R}^d$, l'opérateur

$$L^\sigma f \;=\; \sigma^2\,\Delta f + \left(\nabla(\sigma^2) - \sigma^2\,\nabla V\right)\cdot\nabla f,$$

conserve la même mesure invariante et réversible μ. Ainsi, l'estimation du trou spectral $\lambda_1(-L^\sigma)$ donne lieu à la détermination de la meilleure constante dans l'inégalité de Poincaré à poids

$$\lambda\,\mathrm{Var}_\mu(f) \;\leq\; \int_{\mathbb{R}^d} \sigma^2\,|\nabla f|^2\,d\mu, \tag{3.2.8}$$

l'opérateur carré du champ étant ici $\Gamma^\sigma(f,f) = \sigma^2\,|\nabla f|^2$. On note L^σ_{rad} l'opérateur radial unidimensionnel associé. Le résultat est alors le suivant.

Théorème 3.2.14. *Les inégalités suivantes entre trous spectraux ont lieu :*

$$\min\left\{\lambda_1(-L^\sigma_{\mathrm{rad}}),\, \frac{d-1}{\int_{\mathbb{R}^d}\frac{|x|^2}{\sigma(x)^2}\,\mu(dx)}\right\} \;\leq\; \lambda_1(-L^\sigma) \;\leq\; \min\left\{\lambda_1(-L^\sigma_{\mathrm{rad}}),\, \frac{d\,\int_{\mathbb{R}^d}\sigma(x)^2\,\mu(dx)}{\int_{\mathbb{R}^d}|x|^2\,\mu(dx)}\right\}.$$

Contrairement au cas log-concave, les ordres de grandeur des bornes inférieure et supérieure sur le trou spectral peuvent être sensiblement différents. Regardons ce qu'il en est par exemple pour la loi de Cauchy généralisée dont nous avons déjà parlé, *i.e.*, $V(x) = \beta\log(1+|x|^2)$ avec $\beta > d/2$, la mesure μ ayant sa densité de Lebesgue proportionnelle à $(1+|x|^2)^{-\beta}$. Dans l'article [30], Bobkov et Ledoux ont établi l'inégalité de Poincaré à poids (3.2.8) avec $\sigma(x)^2 = 1+|x|^2$, sous la contrainte $\beta \geq d$, la constante explicite qu'ils ont obtenue n'étant pas optimale. Dans le cas $\beta \geq d+1$, Nguyen [154] a trouvé la bonne constante, à savoir $2(\beta-1)$. De notre côté, la méthode développée ci-dessus nous a permis de couvrir une partie des valeurs manquantes entre $d/2$ et $d+1$ pour le paramètre β. Dans ce cadre, l'opérateur L^σ s'écrit

$$L^\sigma f(x) \;=\; \left(1+|x|^2\right)\,\Delta f(x) + 2\,(1-\beta)\,x\,\nabla f(x), \quad x\in\mathbb{R}^d.$$

Ci-dessous, les transitions continues en fonction du paramètre β sont essentiellement dues à la perte d'intégrabilité L^2 des fonctions propres associées, lorsqu'elles ont été identifiées.

Théorème 3.2.15. *Dans le cas $d \geq 3$, le trou spectral de la dynamique markovienne associée à la loi de Cauchy généralisée de paramètre $\beta > d/2$ satisfait :*

- $\lambda_1(-L^\sigma) = \lambda_1(-L^\sigma_{\text{rad}}) = \left(\beta - \frac{d}{2}\right)^2$ *si* $d/2 < \beta \leq d/2 + 2$;
- $\lambda_1(-L^\sigma) = \lambda_1(-L^\sigma_{\text{rad}}) = 4\,\left(\beta - \frac{d}{2} - 1\right)$ *si* $2 + d/2 < \beta \leq d\,(d+2)/(d+1)$;
- $\frac{2\,\beta\,(d-1)}{d} \leq \lambda_1(-L^\sigma) \leq \lambda_1(-L^\sigma_{\text{rad}}) = 4\left(\beta - \frac{d}{2} - 1\right)$ *si* $d\,(d+2)/(d+1) < \beta < d+1$;
- $\lambda_1(-L^\sigma) = 2(\beta - 1) < \lambda_1(-L^\sigma_{\text{rad}}) = 4\,\left(\beta - \frac{d}{2} - 1\right)$ *si* $\beta \geq d+1$.

De plus, le même type d'estimations a lieu en dimension 2.

On peut procéder de même en revenant au cas gaussien standard et en intégrant au modèle un poids σ^2. On obtient alors les résultats suivants.

Théorème 3.2.16. *Si μ désigne la mesure gaussienne standard γ, alors pour le poids $\sigma(x)^2 = 1 + |x|^2$, on a l'encadrement*

$$d - 1 \;\leq\; \lambda_1(-L^\sigma) \;\leq\; d + 1,$$

tandis que si le poids est plutôt $\sigma(x)^2 = 1/\left(1 + |x|^2\right)$, alors on a

$$\frac{d-1}{d\,(d+3)} \;\leq\; \lambda_1(-L^\sigma) \;\leq\; \frac{1}{d-2} \wedge 1.$$

En termes d'inégalité de Poincaré à poids, celle dans le second cas est plus exigeante que l'inégalité de Poincaré classique, qui elle-même est plus forte que l'inégalité de Poincaré à poids induite dans le premier cas, les poids ayant tendance respectivement à s'écraser à l'infini et à tendre vers l'infini. Il n'est donc pas surprenant d'obtenir de tels ordres de grandeur du trou spectral dans ces deux cas. Notons enfin que la borne inférieure dans le second cas est comparable à celle obtenue par la méthode des entrelacements induite par l'inégalité de Poincaré à poids (3.2.5). Ceci clôt le résumé de l'article [41].

3.2.4 Sur la droite réelle

Nous terminons la présentation succincte de nos résultats en considérant avec attention le cas unidimensionnel, les choses étant plus simples mais non moins intéressantes, beaucoup de résultats importants en dimension supérieure se réduisant à une étude fine d'un problème voisin sur la droite réelle, comme nous l'avons vu auparavant. De surcroît, l'analyse spectrale et fonctionnelle par les entrelacements à poids peut être poussée plus loin que dans le cadre multidimensionnel général. La raison principale repose sur l'observation triviale que les gradients de fonctions définies sur la droite réelle sont eux aussi des fonctions réelles, d'où les opérateurs apparaissant dans l'entrelacement sont de même dimension. Ci-dessous, nous décrivons très succinctement les travaux menés dans les articles [37, 38, 40] correspondant à deux publications parues en 2014 et 2016, ainsi qu'à une prépublication datant de 2019. Bien qu'il y ait beaucoup de choses à présenter dans ce cadre unidimensionnel, nous avons décidé par souci de concision de ne détailler qu'un (voire deux) résultat(s) figurant dans chacun de ces trois articles.

Dans cette partie, nous considérons le processus de Kolmogorov $(X_t)_{t\geq 0}$ sur $\mathbb{R}$, de semigroupe $(P_t)_{t\geq 0}$ et de générateur

$$Lf \;=\; f'' - V'\, f',$$

où $V : \mathbb{R} \to \mathbb{R}$ est un potentiel suffisamment régulier, la mesure invariante et réversible μ, de densité de Lebesgue proportionnelle à e^{-V}, étant supposée être une mesure de probabilité. Néanmoins, mentionnons que dans ce cadre unidimensionnel, tous les résultats que l'on va exposer admettent une version pour les processus de diffusion de générateur

$$L^\sigma f \;=\; \sigma^2\, f'' + \left((\sigma^2)' - \sigma^2\, V' \right)\, f',$$

étudiés dans les Sections 2.2.3 et 3.2.3, dont le coefficient de diffusion σ^2 n'est pas constant. Contrairement au cas multidimensionnel pour lequel le choix du générateur L^σ induit forcément la réversibilité de la dynamique (la matrice de diffusion est particulière, au sens où c'est un multiple de l'identité), la forme de ce générateur n'est en rien réductrice sur la droite réelle car toutes les diffusions unidimensionnelles sont réversibles. Avant de commencer à énoncer nos résultats, rappelons tout d'abord les entrelacements (3.1.7) et (3.2.1) dans leur version unidimensionnelle. L'entrelacement classique est donné pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ suffisamment régulière par l'identité :

$$(Lf)' \;=\; (L - V'')\, f'.$$

Notons que l'opérateur L apparaît des deux côtés de la formule, ce qui n'est pas sans conséquence dans l'analyse qui va suivre. Étant donné un poids $a : \mathbb{R} \to \mathbb{R}$ suffisamment régulier et ne s'annulant pas, on note $\partial_a f = a\, f'$. L'entrelacement à poids, qui est ramené à l'entrelacement classique précédent lorsque a est constant non nul, s'écrit quant à lui :

$$\partial_a L f \;=\; (L_a - M_1^a)\, \partial_a f, \tag{3.2.9}$$

où L_a est l'opérateur de diffusion $L_a f = f'' - V_a\, f'$ de potentiel $V_a = V + \log(a^2)$ et M_1^a est le potentiel multiplicatif

$$M_1^a \;=\; V'' - a\, La^{-1} \;=\; V'' + \frac{L_a(a)}{a} \;=\; \frac{(-Lh)'}{h'},$$

avec $a = 1/h'$ seulement dans la troisième égalité, la fonction $h : \mathbb{R} \to \mathbb{R}$ étant suffisamment régulière et de dérivée ne s'annulant pas sur $\mathbb{R}$. La mesure invariante et réversible de l'opérateur L_a, notée μ_a, a sa densité de Lebesgue proportionnelle à $a^{-2}\, e^{-V}$. Néanmoins, il se peut qu'elle soit de masse infinie. Dans ce cadre réel, le processus sous-jacent au générateur L_a est noté $(X_{a,t})_{t\geq 0}$, de semigroupe $(P_{a,t})_{t\geq 0}$. Enfin, si $U : \mathbb{R} \to \mathbb{R}$ est un potentiel suffisamment régulier, alors on note $(P_{a,t}^U)_{t\geq 0}$ le semigroupe associé à l'opérateur de Schrödinger $L_a - U$.

Comme nous l'avons vu précédemment, il s'agit d'un semigroupe de Feynman-Kac admettant une représentation probabiliste simple, à savoir

$$P_{a,t}^{U} f(x) \;=\; \mathbb{E}_{\delta_x} \left[f(X_{a,t}) \, \exp\left(- \int_0^t U(X_{a,s}) \, ds \right) \right],$$

sous réserve que cet objet soit bien défini. Une condition suffisante assurant cela est que U soit minorée, auquel cas l'on obtient l'inégalité ponctuelle

$$\left| P_{a,t}^{U} f \right| \;\leq\; e^{-t \inf_{\mathbb{R}} U} \, P_{a,t} \left(|f| \right).$$

Commençons par l'article [37], écrit avec Michel Bonnefont et publié en 2014 à la suite du travail effectué avec Djalil Chafaï [58] dans le cadre discret des processus de naissance et de mort, ces derniers ayant de nombreux points communs avec les diffusions unidimensionnelles. Ce fût le premier travail effectué autour des entrelacements à poids dans des situations continues. Les grands principes ayant déjà été exposés, nous ne reviendrons pas sur les résultats spectraux unidimensionnels figurant dans cet article. En revanche, un aspect n'ayant pas encore été évoqué jusqu'à maintenant est celui des conséquences de ces entrelacements en matière d'autres inégalités fonctionnelles, à la saveur moins spectrale telles que les inégalités de type Sobolev logarithmiques. Ces inégalités fonctionnelles, comme celle dite de Beckner, sont basées par définition sur la donnée d'une certaine fonctionnelle convexe. Afin d'être en mesure d'établir de telles inégalités par la théorie des entrelacements, notamment lorsque l'on sort du cadre spectral de l'inégalité de Poincaré, la première chose à faire est peut-être d'étudier les propriétés de stabilité de ces entrelacements par l'action de fonctions convexes. Ainsi, soit $\mathcal{I} \subset \mathbb{R}$ un intervalle ouvert stable par l'action du semigroupe $(P_t)_{t \geq 0}$, *i.e.*, si f est à valeurs dans $\mathcal{I}$ alors $P_t f$ aussi, et ce pour tout $t \geq 0$, et soit $\mathcal{C}_{\mathcal{I}}$ l'ensemble des fonctions convexes régulières $\varphi : \mathcal{I} \to \mathbb{R}$ telles que $\varphi'' > 0$, $\varphi''' \leq 0$ et $-1/\varphi''$ est convexe sur $\mathcal{I}$. Par exemple, pour l'inégalité de Poincaré, on prend $\mathcal{I} = \mathbb{R}$ et $\varphi(x) = x^2$, pour l'inégalité entropique, $\mathcal{I} =]0, \infty[$ et $\varphi(x) = x \log x$ et enfin pour l'inégalité de Beckner, $\mathcal{I} =]0, \infty[$ et $\varphi(x) = x^p$, $p \in]1, 2[$. Pour une fonction $\varphi \in \mathcal{C}_{\mathcal{I}}$ donnée, définissons la fonction bivariée $\Psi : \mathcal{I} \times \mathbb{R} \to [0, \infty[$ par $\Psi(x, y) = \varphi''(x) \, y^2$. Alors Ψ est convexe sur $\mathcal{I} \times \mathbb{R}$, cf. [57]. Le résultat de convexité faisant intervenir l'entrelacement à poids (3.2.9) est le suivant.

Théorème 3.2.17. *Supposons qu'il existe un poids $a : \mathbb{R} \to \mathbb{R}$ suffisamment régulier ne s'annulant pas et monotone tel que le potentiel M_1^a soit minoré. Alors pour toute fonction $f : \mathbb{R} \to \mathcal{I}$ suffisamment régulière ayant le même sens de monotonie que a, l'inégalité suivante a lieu : pour tout $t \geq 0$,*

$$\Psi \left(P_t f, \partial_a P_t f \right) \;\leq\; P_{a,t}^{2M_1^a} \left(\Psi(f, \partial_a f) \right). \tag{3.2.10}$$

Formellement, on interpole « à la Bakry-Émery », c'est-à-dire que si l'on note

$$\phi_t(s) \;=\; P_{a,s}^{2M_1^a} \left(\Psi \left(P_{t-s} f, \partial_a P_{t-s} f \right) \right), \quad s \in [0, t],$$

alors établir l'inégalité désirée revient à montrer que $\phi_t(0) \leq \phi_t(t)$, ce qui est le cas dès lors que ϕ_t est croissante sur $[0, t]$. Le principe de la démonstration repose d'une part sur cette observation, la convexité de la fonction bivariée Ψ faisant fonctionner la méthode, et d'autre part sur le fait que les fonctions f et $P_t f$ partagent la même monotonie (*i.e.*, l'une est croissante (resp. décroissante) si et seulement si l'autre l'est pour tout $t \geq 0$; cette propriété, qui peut se voir trivialement à travers la représentation probabiliste de l'entrelacement classique, est une conséquence d'un argument de couplage fonctionnant sur la droite réelle), justifiant l'hypothèse quelque peu étrange de monotonie de l'énoncé. Notons également que si le poids a est constant non nul, alors la condition de monotonie sur f n'a plus lieu d'être et comme dans ce cas nous avons $M_1^a = V''$, l'hypothèse en force dans le Théorème 3.2.17 est tout simplement le critère de courbure-dimension de Bakry-Émery, où la constante $\rho\,(= \inf_{\mathbb{R}} V'')$ dans $\mathrm{CD}(\rho, \infty)$ est remplacée par la fonction V'', un gain apporté par l'approche Feynman-Kac.

Ainsi, dans le cas $\mathcal{I} =]0, \infty[$ et $\varphi(x) = x \log x$, l'inégalité (3.2.10) s'écrit :

$$\frac{|\partial_a P_t f|^2}{P_t f} \leq P_{a,t}^{2M_1^a} \left(\frac{|\partial_a f|^2}{f} \right), \quad t \geq 0,$$

et en particulier,

$$\frac{|\partial_a P_t f|^2}{P_t f} \leq e^{-2\rho_a t} \, P_{a,t} \left(\frac{|\partial_a f|^2}{f} \right), \quad t \geq 0,$$

où $\rho_a = \inf_{\mathbb{R}} M_1^a$. On en déduit alors une inégalité entropique pour les fonctions monotones.

Théorème 3.2.18. *Supposons qu'il existe un poids $a : \mathbb{R} \to \mathbb{R}$ suffisamment régulier ne s'annulant pas et croissant (resp. décroissant) tel que $\rho_a > 0$. Alors, l'inégalité entropique de constante $2\rho_a$ restreinte aux fonctions croissantes (resp. décroissantes) est satisfaite.*

Preuve. La démonstration étant brève et instructive de la méthode par semigroupe développée par Bakry et Ledoux pour établir ces inégalités fonctionnelles, exposons-la. Quitte à changer le sens de monotonie, on suppose que le poids a en question est croissant. Pour $f : \mathbb{R} \to]0, \infty[$ suffisamment régulière et croissante, on a par le Théorème 3.2.17 :

$$\begin{aligned}
\mathrm{Ent}_\mu(f) &= -\int_{\mathbb{R}} \int_0^\infty \partial_t \left(P_t f \, \log P_t f \right) \, dt \, d\mu \\
&= -\int_0^\infty \int_{\mathbb{R}} \log P_t f \, L P_t f \, d\mu \, dt \\
&= \int_0^\infty \int_{\mathbb{R}} \frac{|\partial_a P_t f|^2}{P_t f} \, d\mu_a \, dt \\
&\leq \int_0^\infty \int_{\mathbb{R}} P_{a,t}^{2M_1^a} \left(\frac{|\partial_a f|^2}{f} \right) d\mu_a \, dt \\
&\leq \int_0^\infty e^{-2\rho_a t} \int_{\mathbb{R}} P_{a,t} \left(\frac{|\partial_a f|^2}{f} \right) d\mu_a \, dt \\
&= \int_0^\infty e^{-2\rho_a t} dt \int_{\mathbb{R}} \frac{|\partial_a f|^2}{f} \, d\mu_a \\
&= \frac{1}{2\rho_a} \int_{\mathbb{R}} \Gamma(f, \log f) \, d\mu,
\end{aligned}$$

où l'on a utilisé l'invariance de μ_a pour le semigroupe $(P_{a,t})_{t \geq 0}$. ∎

Ainsi, ce type de résultat peut être appliqué à des potentiels V non uniformément convexes comme pour la loi de Subbotin de paramètre $\alpha > 2$, voire même non convexes tels que le cas du double-puits. Les constantes obtenues, qui n'ont aucune raison d'être optimales en général, peuvent être quelquefois comparées aux constantes entropiques apparaissant dans la littérature, notamment à travers la version « entropique » du critère de Muckenhoupt établie par Bobkov et Götze [25]. Citons aussi [6, 19] pour quelques améliorations.

Tournous-nous très brièvement vers l'article [40] rédigé avec Michel Bonnefont et Yutao Ma et publié en 2016. Dans ce travail très court, la méthode des entrelacements à poids nous permet d'obtenir non seulement de nouvelles estimations du trou spectral à la suite de celles établies dans l'article [37] pour la loi de Subbotin de paramètre $\alpha \geq 1$, mais aussi de déterminer les constantes optimales dans les inégalités de Poincaré à poids (3.2.8) que nous avons déjà vues pour la dynamique gaussienne standard ou pour la loi de Cauchy généralisée. D'une certaine manière, ce travail peut être interprété comme la version unidimensionnelle de l'article [41] sur les mesures radiales en dimension $d \geq 2$ que nous avons présenté auparavant. Dans le cadre de ces inégalités de Poincaré avec un poids σ^2 dans le terme d'énergie, deux ingrédients importants alimentent notre stratégie : d'une part la constante optimale dans l'inégalité de Poincaré à poids (3.2.8) est réinterprétée comme le trou spectral de l'opérateur L^σ, et d'autre part l'entrelacement à poids appliqué à l'opérateur L^σ, qui fournit un nouveau point de vue sur la formule de Chen-Wang (3.1.5),

$$\lambda_1(-L^\sigma) \quad \geq \quad \sup_a \inf_{\mathbb{R}} \, M_1^a \qquad \left(= \; \sup_h \inf_{\mathbb{R}} \frac{(-L^\sigma h)'}{h'} \right),$$

le supremum étant pris sur les fonctions régulières $a : \mathbb{R} \to \mathbb{R}$ ne s'annulant pas (le second est pris sur les fonctions régulières $h : \mathbb{R} \to \mathbb{R}$ de dérivée ne s'annulant pas, le lien entre les deux étant $a = 1/h'$), permet *in fine* d'estimer précisément ce trou spectral lorsque ce dernier appartient au spectre discret de l'opérateur mais la fonction propre associée est inconnue, ou lorsqu'il est dans le spectre essentiel. Pour ce faire, un exemple de poids souvent pertinent est $a = \sigma^2 \exp(-\varepsilon V)$, où $\varepsilon \in]0,1[$ est un paramètre optimisé en fonction de l'exemple traité, permettant d'obtenir l'estimation suivante à la source de nos résultats :

$$\lambda_1(-L^\sigma) \quad \geq \quad \sup_{\varepsilon \in]0,1[} \inf_{\mathbb{R}} \, (1-\varepsilon)\, \sigma^2 \, \left(V'' + \varepsilon \, (V')^2 \right). \tag{3.2.11}$$

Dans le cadre des lois de Subbotin de paramètre $\alpha \geq 1$, le trou spectral de l'opérateur $-L$ est inconnu, sauf dans les cas $\alpha = 1$ (il vaut $1/4$ pour cette loi de Laplace) et $\alpha = 2$ ($\lambda_1(-L) = 1$ dans ce cas gaussien standard), comme nous l'avons mentionné à plusieurs reprises. Ainsi, nous proposons dans cette courte note de nouvelles bornes sensiblement différentes selon que le paramètre $\alpha \in [1,2]$ ou $\alpha > 2$. En effet, une telle variation est justifiée par le fait que le cas gaussien est critique pour les propriétés de convexité du potentiel V : l'infimum de V'' est nul et est « atteint » en l'infini lorsque $\alpha \in [1,2[$ et à l'origine pour $\alpha > 2$. Le résultat initialement démontré faisant intervenir la fonction Gamma, nous exposons ci-dessous des bornes simplifiées pour une meilleure compréhension de nos estimations en fonction du paramètre α.

Théorème 3.2.19. *Le trou spectral associé à la loi de Subbotin de paramètre $\alpha \geq 1$ satisfait les bornes suivantes :*

○ *si* $\alpha \in [1,2[$, *on a l'encadrement :*

$$\frac{\alpha^2}{4} \;\leq\; \lambda_1(-L) \;\leq\; 2^{1-2/\alpha};$$

○ *si* $\alpha \geq 2$, *alors on a*

$$\frac{2(1+\alpha)^{1-2/\alpha}}{\alpha} \;\leq\; \lambda_1(-L) \;\leq\; 3^{1-2/\alpha}.$$

En particulier, nos bornes sont optimales lorsque α est proche de 2 et de 1 (sauf la borne supérieure dans ce dernier cas, qui donne 1/2 et non 1/4).

À présent, exposons rapidement les constantes optimales dans les inégalités de Poincaré à poids obtenues dans les cas gaussien standard et loi de Cauchy généralisée. On rappelle que le trou spectral $\lambda_1(-L^\sigma)$ de l'opérateur $-L^\sigma$ est la constante optimale dans l'inégalité de Poincaré à poids (3.2.8).

Théorème 3.2.20. *Dans le cas gaussien standard, on se donne le poids* $\sigma_b(x)^2 = (1+bx^2)^{-1}$, *où* b *est un paramètre quelconque strictement positif. Alors on a la dichotomie suivante :*

$$\lambda_1(-L^{\sigma_b}) = \begin{cases} 1-b & si \quad 0 < b < 1/2; \\ 1/4b & si \quad b \geq 1/2. \end{cases}$$

Pour la loi de Cauchy généralisée de paramètre $\beta > 1/2$, *le poids considéré est* $\sigma(x)^2 = 1 + x^2$. *Alors le trou spectral vaut*

$$\lambda_1(-L^\sigma) = \begin{cases} 2(\beta-1) & si \quad \beta > 3/2; \\ (\beta-1/2)^2 & si \quad 1/2 < \beta \leq 3/2. \end{cases}$$

Mentionnons que dans les cas $b \in]0,1/2[$ et $\beta > 3/2$, on est en mesure de déterminer explicitement la fonction propre associée au trou spectral. En effet, il s'agit d'un phénomène particulier de la dimension 1, stipulant que si l'on trouve une fonction propre d'un opérateur de diffusion dont la dérivée ne s'annule pas (elle est donc strictement monotone), alors elle appartient forcément à l'espace propre $E_{\lambda_1(-L)}$ qui est de dimension 1, les valeurs propres étant simples dans le cas unidimensionnel (cet argument a déjà été utilisé pour déterminer dans certains cas le trou spectral de l'opérateur radial dans la Section 3.2.3). Tout comme le cas de la loi de Cauchy généralisée multidimensionnelle évoqué plus haut, les transitions continues en fonction des paramètres b et β sont justifiées par l'absence soudaine d'intégrabilité dans l'espace $L^2(\mu)$: c'est dans cette situation que l'utilisation de la formule (3.2.11) devient appropriée. Pour terminer, précisons également que les poids choisis ci-dessus sont optimaux, au sens où si l'on désire considérer les poids $\sigma(x)^2 = (1+x^2)^{-a}$, $a > 1$, et $\sigma(x)^2 = (1+x^2)^a$, $0 < a < 1$, dans les situations gaussienne et Cauchy respectivement, alors on montre que le trou spectral n'existe pas, *i.e.*, $\lambda_1(-L^\sigma) = 0$, ou encore que l'inégalité de Poincaré à poids (3.2.8) n'est pas vérifiée.

À présent, nous allons conclure la présentation de nos travaux en commentant de manière plus détaillée le dernier projet que nous avons achevé récemment avec Michel Bonnefont. Il s'agit d'une prépublication [38] datant de 2019, traitant du problème de l'estimation des valeurs propres supérieures en utilisant la méthode des entrelacements à poids de manière quelque peu similaire à l'étude effectuée dans [39] et résumée dans la Section 3.2.2. Néanmoins nous allons beaucoup plus loin, la dimension 1 permettant une analyse plus poussée qui bloque pour l'instant en dimension supérieure. Plus précisément, nous obtenons pour une large classe de dynamiques markoviennes (au-delà de celles dont le potentiel est uniformément convexe) des bornes inférieures et supérieures optimales, données par des formules variationnelles, sur toutes les valeurs propres situées sous le (voire correspondant au) bas du spectre essentiel. Ainsi, ce résultat généralise d'une part la formule de Chen-Wang (3.1.5) pour le trou spectral au cas des valeurs propres supérieures données par le théorème du min-max de Courant-Fisher, et d'autre part les bornes obtenues (dans une version unidimensionnelle) par Milman [150] pour des potentiels uniformément convexes.

Commençons par énoncer le résultat de Milman dans sa version unidimensionnelle. On rappelle que sous l'hypothèse de convexité uniforme du potentiel V, le spectre de l'opérateur $-L$ est discret.

Théorème 3.2.21. *Supposons que le potentiel soit uniformément convexe. Alors pour tout $n \in \mathbb{N}^*$, la n-ième valeur propre strictement positive $\lambda_n(-L)$ de l'opérateur $-L$ satisfait les bornes*

$$n \inf_{\mathbb{R}} V'' \ \leq \ \lambda_n(-L) \ \leq \ n \sup_{\mathbb{R}} V'',$$

où il est entendu que le supremum ci-dessus est infini si d'aventure la dérivée seconde du potentiel V n'est pas majorée.

Ainsi, ce résultat de Milman (ou plutôt sa version multidimensionnelle), obtenu en utilisant le célèbre théorème de contraction de Caffarelli [48] issu du transport optimal, nous dit que dans le cas uniformément convexe, les valeurs propres sont comparables à celles de l'opérateur associé au processus d'Ornstein-Uhlenbeck le plus proche, c'est-à-dire que si l'on note L_{OU_p} le générateur $L_{OU_p} f(x) = f''(x) - pxf'(x)$, alors on a

$$\lambda_n(L_{OU_{p_1}}) \ \leq \ \lambda_n(-L) \ \leq \ \lambda_n(L_{OU_{p_2}}),$$

avec $p_1 = \inf_{\mathbb{R}} V''$ et $p_2 = \sup_{\mathbb{R}} V''$. Dans ce cadre unidimensionnel, le Théorème 3.2.21 se révèle être une simple conséquence du Théorème 3.1.1 établi par Johnsen à l'aide de l'entrelacement classique, qui d'ailleurs se simplifie dans sa formulation, la restriction au sous-espace des gradients étant inutile d'une part (toute fonction régulière étant la dérivée de ses primitives) et les valeurs propres étant toutes simples d'autre part.

Théorème 3.2.22. *Sous l'hypothèse de trou spectral $\lambda_1(-L) > 0$, les spectres des opérateurs $-L_{|\mathbb{1}^\perp}$ et $-L+V''$ coïncident, i.e., $\sigma(-L)\backslash\{0\} = \sigma(-L+V'')$, tout comme les spectres essentiels et discrets. De surcroît, les valeurs propres données par le théorème du min-max de Courant-Fisher, vérifient pour tout $n \in \mathbb{N}^*$,*

$$\lambda_n(-L) \ = \ \lambda_{n-1}(-L+V'').$$

Ainsi, en utilisant alors la formule variationnelle définissant les valeurs propres de l'opérateur de Schrödinger $-L+V''$, on a pour tout $n \in \mathbb{N}^*$,

$$\lambda_n(-L) \ = \ \lambda_{n-1}(-L+V'') \ \geq \ \lambda_{n-1}(-L) + \inf_{\mathbb{R}} V'',$$

d'où la borne inférieure de Milman par une simple récurrence. La borne supérieure s'obtient de la même manière. Notons que cette approche par les entrelacements permet d'aller un peu plus loin, au sens où le résultat obtenu en réalité est : pour tout $n \in \mathbb{N}^*$,

$$\inf_{\mathbb{R}} V'' \; \leq \; \lambda_n(-L) - \lambda_{n-1}(-L) \; \leq \; \sup_{\mathbb{R}} V'', \tag{3.2.12}$$

c'est-à-dire que le Théorème 3.2.22 fournit une information sur l'écart entre les valeurs propres successives. Néanmoins, cette stratégie n'est plus appropriée dès lors que le potentiel V n'est plus uniformément convexe. L'idée est alors de considérer les entrelacements à poids et de tenter de les itérer pour récupérer de l'information sur les plus grandes valeurs propres en dessous du (voire même correspondant au) bas du spectre essentiel. Pour ce faire, l'entrelacement à poids doit être effectué à chaque étape sur un opérateur du type L_a et pas seulement sur l'opérateur originel L. Étant données deux fonctions $a, b : \mathbb{R} \to \mathbb{R}$ suffisamment régulières et ne s'annulant pas, l'entrelacement à poids général que l'on considère est

$$\partial_b L_a f \;\; = \;\; (L_{ab} - M_a^b)\, \partial_b f,$$

où L_{ab} est l'opérateur de diffusion défini à l'égal de L_a, en remplaçant simplement a par le produit ab et M_a^b est le potentiel multiplicatif

$$M_a^b \;\; = \;\; V_a'' - b\, L_a b^{-1} \;\; = \;\; V_a'' + \frac{L_{ab}(b)}{b} \;\; = \;\; \frac{(-L_a h)'}{h'},$$

où dans la troisième égalité, $b = 1/h'$, la fonction $h : \mathbb{R} \to \mathbb{R}$ étant suffisamment régulière et de dérivée ne s'annulant pas. La notation pour le potentiel M_a^b est importante car elle nous permet de mieux comprendre l'entrelacement à poids, dont on rappelle qu'il s'agit d'une composition (commutative en dimension 1) de deux opérations : l'entrelacement classique et la h-transformée de Doob : le a en indice correspond à l'opérateur à partir duquel on fait l'entrelacement classique, tandis que le b en exposant traduit la h-transformée de Doob avec $h = 1/b$, correspondant au poids choisi (d'où la notation M_1^a utilisée pour l'entrelacement à poids initial (3.2.9)). La raison pour laquelle le produit ab apparaît est due à la structure de groupe sous-jacente à la h-transformée de Doob (si h et k sont deux fonctions régulières ne s'annulant pas, alors la h-transformée de la k-transformée n'est rien d'autre que la hk-transformée, et les opérateurs sont les mêmes si et seulement si h est constante non nulle). En revanche, l'entrelacement et la h-transformée de Doob étant deux concepts distincts, les potentiels $M_a^b, M_1^{ab}, M_{ab}^1$ et M_b^a provenant de différentes combinaisons de ces deux opérations n'ont aucune raison de coïncider.

La précédente analyse donne lieu à la version à poids du Théorème 3.2.22. Comme l'objet d'intérêt est le trou spectral de l'opérateur $-L_a$, la mesure μ_a doit être supposée finie pour que 0 soit valeur propre.

Théorème 3.2.23. *Soit $a, b : \mathbb{R} \to \mathbb{R}$ deux fonctions régulières ne s'annulant pas et telle que μ_a soit une mesure de probabilité. Sous l'hypothèse de trou spectral $\lambda_1(-L_a) > 0$, les spectres (resp. spectres discrets, resp. spectres essentiels) des opérateurs $-L_{a|\mathbb{1}^\perp}$ et $-L_{ab} + M_a^b$ coïncident. En particulier, les valeurs propres données par le théorème du min-max de Courant-Fisher, vérifient pour tout $n \in \mathbb{N}^*$,*

$$\lambda_n(-L_a) \;\; = \;\; \lambda_{n-1}(-L_{ab} + M_a^b).$$

À présent, tous les éléments sont en place pour que nous puissions énoncer le résultat principal de l'article [38]. On rappelle que par la formule de Chen-Wang (3.1.5) adaptée à notre cadre,

$$\lambda_1(-L_a) \;\geq\; \sup_b \inf_{\mathbb{R}} M_a^b \quad \left(= \sup_h \inf_{\mathbb{R}} \frac{(-L_a h)'}{h'} \right),$$

le supremum étant pris sur les fonctions régulières $b : \mathbb{R} \to \mathbb{R}$ ne s'annulant pas (le second est pris sur les fonctions régulières $h : \mathbb{R} \to \mathbb{R}$ de dérivée ne s'annulant pas, le lien entre les deux étant $b = 1/h'$), une condition suffisante nous assurant l'existence d'un trou spectral est qu'il existe un tel poids b tel que le potentiel M_a^b soit minoré par une constante strictement positive. Dans la suite, on introduit le poids artificiel $a_0 = 1$ pour que les quantités apparaissant ci-dessous soient correctement définies (en particulier, on a $L_{a_0} = L$ et $\mu_{a_0} = \mu$).

Théorème 3.2.24. *Soit $n \in \mathbb{N}^*$. Alors on a l'encadrement suivant pour la n-ième valeur propre strictement positive de l'opérateur $-L$:*

$$\sup_{a_1,\ldots,a_n} \sum_{i=1}^n \inf_{\mathbb{R}} M_{a_0\ldots a_{i-1}}^{a_i} \;\leq\; \lambda_n(-L) \;\leq\; \inf_{a_1,\ldots,a_n} \sum_{i=1}^n \sup_{\mathbb{R}} M_{a_0\ldots a_{i-1}}^{a_i}, \tag{3.2.13}$$

les extrema étant pris sur les fonctions régulières $a_i : \mathbb{R} \to \mathbb{R}$ ne s'annulant pas et telles que pour tout $i = 1, \ldots, n$, la mesure $\mu_{a_0\ldots a_{i-1}}$ soit une mesure de probabilité et le trou spectral de l'opérateur $-L_{a_0\ldots a_{i-1}}$ existe, i.e., $\lambda_1(-L_{a_0\ldots a_{i-1}}) > 0$. Ci-dessus, la convention est de valoir l'infini si d'aventure l'un des potentiels $M_{a_0\ldots a_{i-1}}^{a_i}$ n'est pas borné inférieurement dans le terme de gauche ou supérieurement dans celui de droite.

De surcroît, lorsque les $\lambda_i(-L), i = 1, \ldots, n$, appartiennent au spectre discret de l'opérateur $-L$, ces inégalités deviennent des égalités et l'on a l'identité

$$\lambda_n(-L) \;=\; \sum_{i=1}^n \lambda_1(-L_{a_0\ldots a_{i-1}}), \tag{3.2.14}$$

pour les poids a_i déterminés par le système récursif bien défini $\partial_{a_i} \ldots \partial_{a_1} g_i = 1$, où chaque g_i est une fonction propre associée à la valeur propre $\lambda_i(-L)$.

La preuve des deux inégalités n'est pas très difficile : il s'agit d'invoquer de manière récursive le Théorème 3.2.23 et d'utiliser à chaque étape les propriétés de monotonie induites par les formules variationnelles définissant ces valeurs propres : avec $L = L_{a_0}$, on a en choisissant des poids $a_1, \ldots, a_n$ qui satisfont les contraintes de l'énoncé,

$$\begin{aligned}
\lambda_n(-L) &= \lambda_{n-1}(-L_{a_0a_1} + M_{a_0}^{a_1}) \\
&\geq \lambda_{n-1}(-L_{a_0a_1}) + \inf_{\mathbb{R}} M_{a_0}^{a_1} \\
&= \lambda_{n-2}(-L_{a_0a_1a_2} + M_{a_0a_1}^{a_2}) + \inf_{\mathbb{R}} M_{a_0}^{a_1} \\
&\geq \lambda_{n-2}(-L_{a_0a_1a_2}) + \inf_{\mathbb{R}} M_{a_0a_1}^{a_2} + \inf_{\mathbb{R}} M_{a_0}^{a_1} \\
&= \ldots \\
&\geq \lambda_0(-L_{a_0\ldots a_n}) + \sum_{i=1}^n \inf_{\mathbb{R}} M_{a_0\ldots a_{i-1}}^{a_i}
\end{aligned}$$

$$\geq \sum_{i=1}^{n} \inf_{\mathbb{R}} M_{a_0\ldots a_{i-1}}^{a_i}.$$

Notons que la valeur propre $\lambda_0(-L_{a_0\ldots a_n})$ peut ne pas être nulle, la mesure $\mu_{a_0\ldots a_n}$ n'ayant pas été supposée finie. Bien évidemment, la borne supérieure est obtenue de la même manière.

En revanche, le résultat d'optimalité est bien plus technique et utilise de nouveau le Théorème 3.2.23. Les points importants de la démonstration reposent sur le fait que lorsque les $\lambda_i(-L), i = 1, \ldots, n$, appartiennent au spectre discret de l'opérateur $-L$, alors les trous spectraux $\lambda_1(-L_{a_0\ldots a_{i-1}}), i = 1, \ldots, n$, appartiennent au spectre discret de l'opérateur $-L_{a_0\ldots a_{i-1}}$, où chaque a_i est défini par $1/(g_1^i)'$, g_1^i étant une fonction propre de l'opérateur $-L_{a_0\ldots a_{i-1}}$ associée au trou spectral $\lambda_1(-L_{a_0\ldots a_{i-1}})$, donc de dérivée ne s'annulant pas. Choisis tels quels, les a_i satisfont bien les contraintes apparaissant dans les extrema des inégalités (3.2.13). L'égalité (3.2.14) établie, il reste à montrer que les poids a_i ainsi définis satisfont bien le système récursif $\partial_{a_i} \ldots \partial_{a_1} g_i = 1$, et l'argument clé pour ce faire est de voir que les valeurs propres $\lambda_i(-L)$ peuvent être obtenues par un entrelacement à poids d'ordre supérieur, *i.e.*,

$$\partial_{a_i} \ldots \partial_{a_1} Lf \;=\; L_{a_0\ldots a_i}(\partial_{a_i} \ldots \partial_{a_1} f) - \lambda_i(-L)\, \partial_{a_i} \ldots \partial_{a_1} f.$$

Apportons quelques éclairages sur le Théorème 3.2.24. Tout d'abord, il s'agit d'une généralisation de la formule de Chen-Wang au cas des valeurs propres supérieures, leur formule apparaissant dans le cas $n = 1$ (complétée par la borne supérieure) :

$$\sup_a \inf_{\mathbb{R}} M_1^a \;\leq\; \lambda_1(-L) \;\leq\; \inf_a \sup_{\mathbb{R}} M_1^a,$$

tandis que si tous les poids a_i sont supposés constants non nuls, alors on retrouve le Théorème 3.2.21 sous l'hypothèse de convexité uniforme du potentiel V. Pour $n = 2$, la borne inférieure de l'inégalité (3.2.13) est peu ou prou la version unidimensionnelle de l'estimation de $\lambda_{d+1}(-L)$ décrite dans le Théorème 3.2.10. Par ailleurs, mentionnons également que derrière le système récursif $\partial_{a_i} \ldots \partial_{a_1} g_i = 1$ se cache une propriété intéressante des oscillations des fonctions propres, connue dans la théorie de Sturm-Liouville sous le nom de système de Chebyshev, ou T-système, cf. par exemple l'ouvrage de Karlin-Studden [117], que l'on retrouve donc sur la droite réelle par la technique des entrelacements. Enfin, l'identité d'optimalité (3.2.14) peut être réécrite comme

$$\lambda_n(-L) - \lambda_{n-1}(-L) \;=\; \lambda_1(-L_{a_0\ldots a_{n-1}}), \tag{3.2.15}$$

où les a_i vérifient le système récursif $\partial_{a_i} \ldots \partial_{a_1} g_i = 1$ pour $i = 1, \ldots, n-1$. En particulier, cette formule reste vraie si la valeur propre $\lambda_n(-L)$ n'est plus dans le spectre discret de l'opérateur $-L$, auquel cas la valeur propre $\lambda_1(-L_{a_0\ldots a_{n-1}})$ elle-même n'est plus dans le spectre discret de l'opérateur $-L_{a_0\ldots a_{n-1}}$. Ainsi, le résultat d'optimalité se révèle de nouveau satisfait dès lors que l'on montre la formule suivante,

$$\lambda_1(-L_{a_0\ldots a_{n-1}}) \;=\; \sup_{a_n} \inf_{\mathbb{R}} M_{a_0\ldots a_{n-1}}^{a_n},$$

c'est-à-dire l'égalité dans la formule de Chen-Wang lorsque la valeur propre appartient au spectre essentiel. Ne l'ayant malheureusement pas trouvée dans les travaux foisonnants de ces deux auteurs sur le sujet, nous l'avons démontrée en utilisant le Théorème 3.2.23.

Le Théorème 3.2.24 peut être appliqué dans de nombreuses situations pour lesquelles le potentiel V n'est pas uniformément convexe, voire même convexe. De surcroît, le cas de certains potentiels dégénérés, *i.e.*, $\liminf_{|x|\to\infty} V''(x) = -\infty$, peut être considéré. Si l'on revient sur l'exemple typique de la loi de Subbotin de paramètre $\alpha > 1$ (on rappelle que pour $\alpha > 1$ le spectre est discret tandis que pour $\alpha = 1$ et la loi de Laplace, le trou spectral existe et vaut $1/4$, mais correspond exactement au bas du spectre essentiel de l'opérateur), on obtient le résultat suivant. Comme pour le trou spectral, nous nous concentrons plutôt sur la borne inférieure car c'est elle en général qui est difficile à établir.

Théorème 3.2.25. *Dans le cadre de la loi de Subbotin de paramètre $\alpha > 1$, on a :*

∘ *si $\alpha \in]1,2[$, alors pour tout $\varepsilon > 0$, il existe une constante $C_{\alpha,\varepsilon} > 0$ explicite en fonction des paramètres α et ε telle que pour tout $n \in \mathbb{N}^*$, on ait*

$$\lambda_n(-L) \;\; \geq \;\; C_{\alpha,\varepsilon} n^{2-2/\alpha-\varepsilon}.$$

∘ *si $\alpha \geq 2$, alors il existe $C_\alpha > 0$ explicite telle que pour tout $n \in \mathbb{N}^*$,*

$$\lambda_n(-L) \;\; \geq \;\; C_\alpha \, n^{2-2/\alpha}.$$

Ces résultats sont en accord avec la loi de Weyl donnant le comportement asymptotique de la fonction de comptage des valeurs propres pour tout $\alpha > 1$, cf. [150] :

$$\# \{k \geq 1 : \lambda_k(-L) \leq \lambda\} \underset{\lambda\to\infty}{\approx} C_\alpha \, \lambda^{\alpha/2(\alpha-1)}.$$

En effet, dans le cas $\alpha \geq 2$ il s'agit exactement du résultat attendu tandis que dans le cas $\alpha \in]1,2[$, on obtient le bon ordre de grandeur à ε près. Néanmoins, nous soupçonnons qu'un meilleur choix des poids a_i permettrait de corriger cela.

Concluons la présentation de l'article [38] en énonçant le dernier résultat que nous souhaiterions évoquer. L'identité (3.2.15) ouvre la voie, comme dans le cas trivial d'un potentiel uniformément convexe, cf. (3.2.12), à une estimation de l'écart entre les valeurs propres, la difficulté résidant dans le caractère non explicite des poids a_i car dépendant des fonctions propres g_i qui sont en général inconnues. Malgré cette anicroche, l'écart au moins entre les deux premières valeurs propres strictement positives peut être estimé en utilisant crucialement l'identité (3.2.15) avec $n = 2$,

$$\lambda_2(-L) - \lambda_1(-L) \;\; = \;\; \lambda_1(-L_a),$$

où $a = 1/g_1'$ (l'entier n étant ici explicite, il devient inutile de conserver le poids artificiel a_0 dans la notation). Notons que cette question n'a d'intérêt que si $\lambda_1(-L) \in \sigma_{\mathrm{disc}}(-L)$ (en revanche il se peut que $\lambda_2(-L) \in \sigma_{\mathrm{ess}}(-L)$). Dans le cas présent, la densité de la mesure invariante μ_a de l'opérateur $-L_a$ est proportionnelle à $(g_1')^2 \, e^{-V}$, c'est-à-dire qu'elle dépend de la fonction propre g_1 inconnue. Néanmoins, si le potentiel associé V_a est uniformément convexe, alors on a l'inégalité $\lambda_1(-L_a) \geq \inf_{\mathbb{R}} V_a''$ et le tour est joué. Ainsi, le travail consiste à montrer cette propriété de convexité uniforme sous certaines hypothèse sur V et, pour ce faire, le Théorème 3.2.23 joue un rôle important car en choisissant le bon poids il permet de ramener le problème à une propriété de log-concavité uniforme de l'état fondamental (ou *ground state*) d'un opérateur de Schrödinger de type « moins laplacien plus potentiel », ce potentiel (disons U) valant $V''/2 + (V')^2/4$. D'après le fameux résultat de Brascamp et Lieb [44], cette propriété de log-concavité uniforme de l'état fondamental est satisfaite dès lors que U est uniformément convexe. Ainsi, on obtient le résultat suivant.

Théorème 3.2.26. *Supposons que le trou spectral $\lambda_1(-L)$ appartienne au spectre discret de l'opérateur $-L$. Si le potentiel $V''/2 + (V')^2/4$ a sa dérivée seconde minorée par une constante strictement positive ρ, alors l'écart entre les deux premières valeurs propres satisfait :*

$$\lambda_2(-L) - \lambda_1(-L) \quad \geq \quad \sqrt{2\rho}.$$

Cette condition sur le potentiel V, ma foi assez inhabituelle et hautement sous-optimale car provenant de la méthode utilisée, a le mérite d'inclure parmi les exemples intéressants certains potentiels non uniformément convexes, comme la loi de Subbotin de paramètre $\alpha = 4$, voire même non convexes tels que le double-puits, le spectre de ce dernier étant lui aussi discret.

3.3 Perspectives

La théorie des entrelacements à poids étant assez récente, il y aurait beaucoup à dire au sujet des questions et problèmes encore non résolus que nous souhaiterions aborder à l'avenir. Néanmoins, terminons ce mémoire par une sélection de quelques ouvertures possibles.

3.3.1 Trou spectral dans le cas des mesures produits perturbées

Dans le travail de Barthe et Klartag [18] que nous avons déjà évoqué, les auteurs étudient le trou spectral pour des mesures produits perturbées. Parmi les nombreux résultats qu'ils établissent dans ce travail, l'un d'entre eux nous intéresse tout particulièrement car il est proche de celui que nous avons obtenu dans ce cadre.

Théorème 3.3.1. *On condidère le potentiel V de type « Subbotin perturbé »* défini par

$$V(x) \quad = \quad \sum_{i=1}^{d} \frac{|x_i|^\alpha}{\alpha} + \varphi(x), \quad x \in \mathbb{R}^d,$$

où $\alpha \in [1, 2]$ et le potentiel d'interaction φ est pair et convexe. Alors il existe une constante universelle $C > 0$ telle que l'on ait l'estimation du trou spectral suivante :

$$\lambda_1(-L) \quad \geq \quad \frac{C}{(\log d)^{1-2/\alpha}}.$$

Pour démontrer ce résultat, les auteurs utilisent plusieurs techniques issues du monde log-concave et des mixtures gaussiennes ainsi que les propriétés fines de parité du problème, la formule variationnelle définissant le trou spectral pouvant ainsi être restreinte aux fonctions impaires dès lors que la mesure μ est paire. Il est difficile de comparer un tel résultat à notre Théorème 3.2.7 car ils ne sont pas vraiment de même nature, le Théorème 3.3.1 couvrant sans doute plus de situations que le notre mais exhibant un aspect dimensionnel absent dans notre résultat, ce dernier étant calibré pour une estimation du trou spectral ne dépendant pas de la dimension. Dès lors que le potentiel d'interaction φ est convexe, notre résultat met en jeu dans la matrice diagonale de l'inégalité (3.2.6) une compétition entre la partie mesure produit et l'influence de l'interaction se manifestant par la quantité clé $\partial_{x_i}\varphi(x)\, U_i'(x_i)$. La prépublication [18] de Barthe et Klartag n'ayant été disponible que très récemment (en juillet 2019, alors que ces quelques lignes sont écrites début septembre 2019), nous n'avons pas encore eu le temps de regarder plus en détail comment leur étude pourrait nous permettre de dégager les bonnes hypothèses de symétrie sur le potentiel V afin que notre Théorème 3.2.7 établi par les entrelacements à poids puisse couvrir une classe plus générale de mesures produits perturbées. Cette question fera sans doute l'objet d'une attention particulière dans les mois à venir.

3.3.2 Entrelacements à poids dans le cas riemannien

Une autre question intéressante serait d'étendre les entrelacements à poids aux variétés riemanniennes, le cadre sans doute le plus naturel associé à ces objets notamment par la géométrie qu'ils encodent, l'entrelacement classique apparaissant déjà à travers la formule de Weitzenböck. En effet, au-delà de l'intérêt évident d'une telle extension, cela nous permettrait aussi de considérer des exemples euclidiens exhibant une matrice de diffusion σ non constante, nous évitant alors une analyse fastidieuse des opérateurs de type Schrödinger encombrés de cette matrice de diffusion. Pour ce faire, l'idée est de changer la métrique pour se ramener à l'étude d'un opérateur de type « laplacien + dérive » sur la variété riemannienne $(\mathbb{R}^d, (\sigma\sigma^T)^{-1})$ et de voir ce que les entrelacements à poids peuvent donner dans ce cadre. Ce type d'approche par changement de métrique, dans l'esprit des études effectuées dans les Sections 2.2.3 et 3.2.3 où l'on interprète une inégalité fonctionnelle à poids comme sa version classique pour une dynamique modifiée dont la matrice de diffusion est justement donnée par ce poids, a été exploitée dans le récent article de Gentil et Zugmeyer [82] avec le critère de courbure-dimension ou encore par Kolesnikov et Milman dans [120] *via* des inégalités de Brascamp-Lieb. De surcroît, cette extension au cadre riemannien nous permettrait de prendre en compte par exemple le cas de l'espace modèle qu'est la sphère (unité) $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ de dimension d (pour laquelle le spectre du laplacien $\Delta_{\mathbb{S}^d}$ est entièrement connu) qui, lorsqu'on la projette d'une certaine manière sur l'hyperplan horizontal $\mathbb{R}^d$, est ramenée à certains modèles euclidiens intéressants pour lesquels les entrelacements à poids n'ont pas encore été étudiés. Par exemple s'il s'agit de la projection stéréographique, alors le modèle sous-jacent sur $\mathbb{R}^d$ est la loi de Cauchy généralisée de paramètre $\beta = d$, obtenue comme la mesure image par cette projection de la mesure de volume normalisée sur $\mathbb{S}^d$. Bref, une telle extension au cadre riemannien est attendue pour de multiples raisons et il serait sans doute pertinent de commencer à étudier le problème de plus près.

3.3.3 Interprétation probabiliste des entrelacements à poids

Sur la droite réelle, la représentation probabiliste du semigroupe $(\widetilde{\mathcal{P}_{A,t}})_{t\geq 0}$ intervenant dans l'entrelacement à poids ne pose pas de problème car il s'agit en réalité d'un vrai semigroupe de Feynman-Kac, comme nous l'avons vu dans la Section 3.2.1. Dans le cas multidimensionnel, cette interprétation probabiliste est encore disponible dès lors que le poids A est un multiple de l'identité, $A = a\, I$, où a est une fonction régulière ne s'annulant pas. En particulier, il s'agit de l'élément clé permettant d'établir la propriété de contraction suivante, cf. [8] :

$$|\widetilde{\mathcal{P}_{aI,t}}\,(a\nabla f)\,| \quad \leq \quad P_{a,t}^{\rho_{aI}}\,(|a\nabla f|)\,,$$

où $(P_{a,t}^{\rho_{aI}})_{t\geq 0}$ désigne le semigroupe de Feynman-Kac suivant, agissant sur les fonctions :

$$P_{a,t}^{\rho_{aI}}\, g(x) \quad = \quad \mathbb{E}_{\delta_x}\left[g(X_{a,t})\, \exp\left(- \int_0^t \rho_{aI}(X_{a,s})\, ds \right) \right],$$

le processus $(X_{a,t})_{t\geq 0}$ ayant pour générateur $L_a = L + 2a\, \nabla a^{-1} \cdot \nabla$ et ρ_{aI} étant la plus petite valeur propre de la matrice symétrique $M_{aI} = \nabla^2 V - a\, La^{-1}\, I$. En revanche, pour une matrice A plus générale, il n'existe pas pour le moment de représentation probabiliste du semigroupe $(\widetilde{\mathcal{P}_{A,t}})_{t\geq 0}$. Pour commencer à étudier cette question, on pourrait supposer la matrice A diagonale, $A = \mathrm{diag}\,(a_i)$, auquel cas l'opérateur de diffusion matriciel $\mathcal{L}_A$ est encore diagonal, ce qui simplifie grandement les choses. La raison motivant cette éventuelle extension réside dans les conséquences qu'elle aurait en matière d'inégalités de covariance et de Sobolev logarithmique. En effet, dans le cas quelque peu contraignant où A est un multiple de l'identité, cette approche stochastique a été utilisée de manière fondamentale dans notre article [8] commun avec Arnaudon et Bonnefont pour établir certaines inégalités de covariance de type $L^1 - L^\infty$, connues sous le nom d'inégalités de Brascamp-Lieb asymétriques depuis les travaux de Menz et Otto [144] d'une part et de Carlen, Cordero-Erausquin et Lieb [51] d'autre part. L'approche stochastique que nous avons utilisée est justifiée par le fait que la propriété de contraction énoncée ci-dessus est requise pour conclure notre preuve de l'inégalité de covariance désirée. Rappelons que les inégalités de covariance sont notamment à la base de la méthode développée par Houdré et ses coauteurs [46, 102, 104] pour établir des propriétés de concentration de la mesure et peuvent être vues comme une alternative à l'approche par les inégalités fonctionnelles de type entropique. Ainsi, si une telle représentation stochastique du semigroupe $(\widetilde{\mathcal{P}_{A,t}})_{t\geq 0}$ venait à être établie au moins dans le cas un peu plus général d'une matrice A diagonale, elle pourrait être exploitée pour obtenir des inégalités de covariance de type Brascamp-Lieb asymétrique qui, à leur tour, permettraient d'envisager des résultats de stabilité intéressants pour la concentration grâce à la présence du poids A. Par ailleurs, la même analyse a lieu si l'on doit s'intéresser à l'inégalité de Sobolev logarithmique. En effet, cette interprétation probabiliste nous amènerait vers une version matricielle du théorème de Girsanov, une approche stochastique pouvant se révéler prometteuse suite aux premiers travaux effectués sur la droite réelle par Clément Steiner, mon étudiant en thèse depuis septembre 2018 (et co-encadré par Patrick Cattiaux). Ces questions autour de l'éventuelle représentation stochastique du semigroupe matriciel $(\widetilde{\mathcal{P}_{A,t}})_{t\geq 0}$ issu de l'entrelacement à poids et ses conséquences en matière d'inégalités fonctionnelles et de concentration de la mesure constituent quelques-unes de ses préoccupations actuelles.

3.3.4 Itération des entrelacements et entrelacements d'ordre supérieur

Pour conclure ces quelques lignes sur les perspectives envisagées, je souhaiterais proposer une dernière piste de recherche. On a vu dans la Section 3.2.4 qu'en dimension 1 l'estimation des valeurs propres situées en dessous du bas du spectre essentiel pouvait être effectuée à partir des entrelacements à poids, l'élément clé dans la démonstration étant de pouvoir les itérer comme on le désire (c'est-à-dire qu'à chaque étape, l'entrelacement à poids opère sur le nouvel opérateur de diffusion apparaissant dans l'opérateur de Schrödinger issu de l'entrelacement à poids précédent). Néanmoins, deux obstructions surviennent lorsque l'on s'intéresse à la dimension supérieure : d'une part la restriction aux gradients (éventuellement à poids) ne peut être évitée et d'autre part chaque entrelacement fait intervenir un opérateur de type Schrödinger sur un espace toujours plus gros. Néanmoins, on devrait être en mesure d'itérer au moins l'entrelacement classique, nous permettant de retrouver simplement le résultat euclidien original de Milman [150], c'est-à-dire la version multidimensionnelle du Théorème 3.2.21. Dans le même esprit, savoir entrelacer non pas avec un gradient (éventuellement à poids) mais avec les dérivées successives (on parle d'entrelacement d'ordre supérieur) pourrait être tout à fait intéressant pour ce que l'on appelle communément la méthode de Stein, permettant de considérer des problèmes d'approximation en loi. Rappelons que les écarts entre probabilités sont souvent mesurés par des distances de couplage satisfaisant un principe de dualité du type

$$d(\mu,\nu) \;=\; \sup_{g\in\mathcal{C}} \int_{\mathbb{R}^d} g\, d(\nu-\mu),$$

où $\mathcal{C}$ est la boule unité d'une « bonne » classe de fonctions (par exemple, l'ensemble des fonctions mesurables bornées est associé à la distance en variation totale d_{VT} tandis que la dualité de Kantorovich-Rubinstein fait correspondre la distance de Wasserstein W_1 aux fonctions lipschitziennes). Le principe de la méthode de Stein, d'après les travaux précurseurs initiés dans les années 70 par Stein [165] pour l'approximation gaussienne, est alors le suivant : étant donnée la mesure invariante μ d'un processus de diffusion à valeurs dans $\mathbb{R}^d$ et de générateur L, il s'agit dans un premier temps d'étudier la régularité de l'éventuelle solution h d'une équation de Poisson du type $-Lh = g$, où g est une fonction centrée donnée ayant une certaine régularité liée à la classe de fonctions $\mathcal{C}$, c'est-à-dire dépendant de la qualité de l'approximation désirée. Puis dans un second temps, l'étude de la régularité de h permet d'estimer quantitativement la distance $d(\mu,\nu)$ en contrôlant l'intégrale de $-Lh$ sous ν. Ainsi, la fonction h admettant la représentation intégrale

$$h \;=\; (-L)^{-1}g \;=\; \int_0^{+\infty} P_t g\, dt,$$

borner ses dérivées successives revient à borner celles du semigroupe. À ma connaissance, seules quelques études multidimensionnelles concernent une telle généralité de mesure cible μ, une des difficultés principales résidant dans la bornitude de ces dérivées successives. Citons par exemple l'article de Ledoux, Nourdin et Peccati [132] dans lequel ils utilisent l'opérateur Γ_3, le carré du champ itéré d'ordre 3, pour être en mesure de contrôler la matrice hessienne du semigroupe, ainsi que le travail de Bonis [36] proposant une approche directe basée sur le critère de courbure-dimension $\mathrm{CD}(\rho,\infty)$ pour borner toutes ses dérivées d'ordre supérieur. C'est pourquoi il serait intéressant d'observer ce que pourrait apporter la théorie des entrelacements à la méthode de Stein, une idée n'ayant pas encore été exploitée à ce jour dans ce contexte diffusif. Il y a quelques années, la méthode de Stein ayant un sens dans le cadre discret (il s'agit de la méthode dite de Chen-Stein, Chen ayant adapté dans la foulée des travaux de Stein la méthode au cadre discret, cf. [59]), Claire Delplancke, l'étudiante en thèse que nous encadrions avec Laurent Miclo, a travaillé sur ce sujet pour les processus de naissance et de mort en exploitant les entrelacements à poids que nous avions initiés dans ce cadre discret dans l'article [58]. Le travail qu'elle a mené en commun avec Bertrand Cloez [67] sur les entrelacements d'ordre supérieur a permis de poser les bases d'une nouvelle approche de la méthode de Chen-Stein, plus générale et moins dépendante des exemples considérés.

Bibliographie

[1] S. Aida, T. Masuda, and I. Shigekawa. Logarithmic Sobolev inequalities and exponential integrability. *J. Funct. Anal.*, 126 :83–101, 1994.

[2] S. Aida and D. Stroock. Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. *Math. Res. Lett.*, 1 :75–86, 1994.

[3] H. Airault. Subordination de processus dans le fibré tangent et formes harmoniques. *C. R. Acad. Sci. Paris Sér. I Math. Sér. A*, 282 :1311–1314, 1976.

[4] N. Alon and V. Milman. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *J. Combinatorial Theory, Series B*, 38 :73–88, 1985.

[5] D. Alonso-Gutiérrez and J. Bastero. *Approching the Kannan-Lovász-Simonovits and variance conjectures*, volume 2131 of *Lecture Notes in Math.* Springer, 2015.

[6] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses.* Société Mathématique de France, 2000.

[7] C. Ané and M. Ledoux. On logarithmic Sobolev inequalities for continuous time random walks on graphs. *Probab. Theory Related Fields*, 116 :573–602, 2000.

[8] M. Arnaudon, M. Bonnefont, and A. Joulin. Intertwinings and generalized Brascamp-Lieb inequalities. *Rev. Math. Ibero.*, 34 :1021–1054, 2018.

[9] D. Bakry. Étude des transformations de Riesz dans les variétés riemanniennes à courbure de Ricci minorée. In *Séminaire de probabilités, XXI*, volume 1247 of *Lecture Notes in Math.*, pages 137–172. Springer, 1987.

[10] D. Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer, 1994.

[11] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electron. Commun. Probab.*, 13 :60–66, 2008.

[12] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes : Lyapunov versus Poincaré. *J. Funct. Anal.*, 254 :727–759, 2008.

[13] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, 1985.

[14] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften.* Springer, 2013.

[15] F. Barthe, C. Bianchini, and A. Colesanti. Isoperimetry and stability of hyperplanes for product probability measures. *Annali di Matematica*, 192 :165–190, 2013.

[16] F. Barthe and D. Cordero-Erausquin. Invariances in variance estimates. *Proc. Lond. Math. Soc.*, 106 :33–64, 2013.

[17] F. Barthe, B. Iooss, and O. Roustant. Poincaré inequalities on intervals - application to sensitivity analysis. *Electron. J. Statist.*, 11 :3081–3119, 2017.

[18] F. Barthe and B. Klartag. Spectral gaps, symmetries and log-concave perturbations. ArXiv 1907.01823, 2019.

[19] F. Barthe and C. Roberto. Sobolev inequalitiesfor probability measures on the real line. *Studia Math.*, 159 :481–497, 2003.

[20] F. Barthe and P. Wolff. Remarks on non-interacting conservative sppin systems : the case of gamma distributions. *Stochastic Process. Appl.*, 119 :2711–2723, 2009.

[21] W. Beckner. A generalized Poincaré inequality for Gaussian measures. *Proc. Amer. Math. Soc.*, 105 :397–400, 1989.

[22] J.-M. Bismut. *Large deviations and the Malliavin calculus*, volume 45 of *Progress in Math.* Birkhäuser, 1984.

[23] S.G. Bobkov. Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Probab.*, 27 :1903–1921, 1999.

[24] S.G. Bobkov. Spectral gap and concentration for some spherically symmetric probability measures. In *Geometric aspects of functional analysis*, volume 1807 of *Lecture Notes in Math.*, pages 37–43. Springer, 2003.

[25] S.G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163 :1–28, 1999.

[26] S.G. Bobkov and F. Götze. Hardy type inequalities via Riccati and Sturm-Liouville equations. In *Sobolev Spaces In Mathematics I*, volume 8 of *International Mathematical Series*, pages 69–86. Springer, 2009.

[27] S.G. Bobkov and C. Houdré. Isoperimetric constants for product probability measures. *Ann. Probab.*, 25 :184–205, 1997.

[28] S.G. Bobkov and M. Ledoux. Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, 107 :383–400, 1997.

[29] S.G. Bobkov and M. Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.*, 156 :347–365, 1998.

[30] S.G. Bobkov and M. Ledoux. Weighted Poincaré-type inequalities for Cauchy and other convex measures. *Ann. Probab.*, 37 :403–427, 2009.

[31] S.G. Bobkov and P. Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theor. Probab.*, 19 :289–336, 2006.

[32] T. Bodineau and B. Helffer. The log-Sobolev inequalities for unbounded spin systems. *J. Funct. Anal.*, 166 :168–178, 1999.

[33] F. Bolley and G. Blower. Concentration inequalities on product spaces with applications to Markov processes. *Studia Math.*, 175 :47–72, 2006.

[34] F. Bolley, I. Gentil, and A. Guillin. Dimensional improvements of the logarithmic Sobolev, Talagrand and Brascamp-Lieb inequalities. *Ann. Probab.*, 45 :261–301, 2018.

[35] A.-I. Bonciocat and K.-T. Sturm. Mass transportation and rough curvature bounds for discrete spaces. *J. Funct. Anal.*, 256 :2944–2966, 2009.

[36] T. Bonis. Rates in the Central Limit Theorem and diffusion approximation via Stein's method. ArXiv 1506.06966, 2015.

[37] M. Bonnefont and A. Joulin. Intertwining relations for one-dimensional diffusions and application to functional inequalities. *Pot. Anal.*, 41 :1005–1031, 2014.

[38] M. Bonnefont and A. Joulin. A note on eigenvalues estimates for one-dimensional diffusion operators. ArXiv 1906.02496, 2019.

[39] M. Bonnefont and A. Joulin. Intertwinings, second-order Brascamp-Lieb inequalities and spectral estimates. ArXiv 1710.08106, Preliminary version : 2017. Last version : 2019.

[40] M. Bonnefont, A. Joulin, and Y. Ma. A note on spectral gap and weighted Poincaré inequalities for some one-dimensional diffusions. *ESAIM Probab. Stat.*, 20 :2016, 18–29.

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

[42] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities : a nonasymptotic theory of independence*. Oxford University Press, 2013.

[43] A.-S. Boudou, P. Caputo, P. DaiPra, and G. Posta. Spectral gap estimates for interacting particle systems via a Bochner-type identity. *J. Funct. Anal.*, 232 :222–258, 2006.

[44] H.J. Brascamp and E.H. Lieb. On extensions of the Brunn-Minkovski and Prékopa-Leindler theorems, including inequalities for log-concave functions, and with an application to the diffusion equation. *J. Funct. Anal.*, 22 :366–389, 1976.

[45] M. Braun, K. Habermann, and K.-T. Sturm. Optimal transport, gradient estimates, and pathwise Brownian coupling on spaces with variable Ricci bounds. ArXiv 1906.09186, 2019.

[46] J.-C. Breton, C. Houdré, and N. Privault. Dimension-free and infinite variance tail estimates on Poisson space. *Acta Appl. Math.*, 95 :151–203, 2007.

[47] R. Bubley and M.E. Dyer. Path coupling : a technique for proving rapid mixing in Markov chains. *Proc. of 38th FOCS*, pages 223–231, 1997.

[48] L.A. Caffarelli. Monotonicity properties of optimal transportation and the FKG and related inequalities. *Comm. Math. Phys.*, 214 :547–563, 2000.

[49] P. Caputo, P. DaiPra, and G. Posta. Convex entropy decay via the Bochner-Bakry-Émery approach. *Ann. Inst. H. Poincaré Probab. Statist.*, 45 :734–753, 2009.

[50] P. Caputo and G. Posta. Entropy dissipation estimates in a zero-range dynamics. *Probab. Theory Relat. Fields*, 139 :65–87, 2007.

[51] E. Carlen, D. Cordero-Erausquin, and E. Lieb. Asymmetric covariance estimates of Brascamp-Lieb type and related inequalities for log-concave measures. *Ann. Inst. H. Poincaré Probab. Statist.*, 49 :1–12, 2013.

[52] P. Cattiaux, N. Gozlan, A. Guillin, and C. Roberto. Functional inequalities for heavy tailed distributions and application to isoperimetry. *Electron. J. Probab.*, 15 :346–385, 2010.

[53] P. Cattiaux and A. Guillin. Deviation bounds for additive functionals of Markov processes. *ESAIM Probab. Stat.*, 12 :12–29, 2008.

[54] P. Cattiaux, A. Guillin, and L. Wu. A note on Talagrand transportation inequality and logarithmic Sobolev inequality. *Probab. Theory Related Fields*, 148 :285–304, 2010.

[55] P. Cattiaux, A. Guillin, and L. Wu. Some remarks on weighted logarithmic Sobolev inequality. *Indiana University Math. J.*, 60 :1885–1904, 2011.

[56] P. Cattiaux, A. Guillin, and P.-A. Zitt. Poincaré inequalities and hitting times. *Ann. Inst. H. Poincaré Probab. Statist.*, 49 :95–118, 2013.

[57] D. Chafaï. Binomial-Poisson entropic inequalities and the $M/M/\infty$ queue. *ESAIM Probab. Stat.*, 10 :317–339, 2006.

[58] D. Chafaï and A. Joulin. Intertwining and commutation relations for birth-death processes. *Bernoulli*, 19 :1855–1879, 2013.

[59] L.H.Y. Chen. Poisson approximation for dependent trials. *Ann. Probab.*, 3 :534–545, 1975.

[60] M.F. Chen. Estimation of spectral gap for Markov chains. *Acta Math. Sinica*, 12 :337–360, 1996.

[61] M.F. Chen. *From Markov chains to non-equilibrium particle systems*. World Scientific Publishing Co. Inc., second edition, 2004.

[62] M.F. Chen. Exponential convergence rate in entropy. *Front. Math. China*, 2 :329–358, 2007.

[63] M.F. Chen. Spectral gap and logarithmic Sobolev constant for continuous spin systems. *Acta Math. Sin. New Ser.*, 24 :705–736, 2008.

[64] M.F. Chen and S.F. Li. Coupling methods for multidimensional diffusion processes. *Ann. Probab.*, 17 :151–177, 1989.

[65] M.F. Chen and F.Y. Wang. Estimation of spectral gap for elliptic operators. *Trans. Amer. Math. Soc.*, 349 :1239–1267, 1997.

[66] L. Cheng, R. Li, and L. Wu. Ricci curvature and W_1 exponential convergence of Markov processes on graphs. ArXiv 1907.11036, 2019.

[67] B. Cloez and C. Delplancke. Intertwinings and stein's magic factors for birth-death processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 55 :341–377, 2019.

[68] D. Cordero-Erausquin. Transport inequalities for log-concave measures, quantitative forms and applications. *Canad. J. Math.*, 69 :481–501, 2017.

[69] D. Cordero-Erausquin, M. Fradelizi, and B. Maurey. The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems. *J. Funct. Anal.*, 214 :410–427, 2004.

[70] P. DaiPra, A.M. Paganoni, and G. Posta. Entropy inequalities for unbounded spin systems. *Ann. Probab.*, 30 :1959–1976, 2002.

[71] P. DaiPra and G. Posta. Entropy decay for interacting systems via the Bochner-Bakry-Émery approach. *Electron. J. Probab.*, 18 :1–21, 2013.

[72] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6 :695–750, 1996.

[73] H. Djellout, A. Guillin, and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.*, 32 :2702–2732, 2004.

[74] R.L. Dobrushin. Prescribing a system of random variables by conditional distributions. *Theory Probab. Appl.*, 15 :458–486, 1970.

[75] R. Eldan. Thin shell implies spectral gap up to polylog via a stochastic localization scheme. *Geom. Funct. Anal.*, 23 :532–569, 2013.

[76] R. Eldan, J.R. Lee, and J. Lehec. Transport-entropy inequalities and curvature in discrete-space Markov chains. In *A Journey Through Discrete Mathematics*, pages 391–406. Springer, 2017.

[77] M. Erbar and M. Fathi. Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature. *J. Funct. Anal.*, 274 :3056–3089, 2018.

[78] M. Erbar and J. Maas. Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.*, 206 :997–1038, 2012.

[79] M. Fathi and Y. Shu. Curvature and transport inequalities for Markov chains in discrete spaces. *Bernoulli*, 24 :672–698, 2018.

[80] P. Fougères. Spectral gap for log-concave probability measures on the real line. In *Séminaire de probabilités, XXXVIII*, volume 1857 of *Lecture Notes in Math.*, pages 95–123. Springer, 2005.

[81] I. Gentil and C. Roberto. Spectral gaps for spin systems : some non-convex phase examples. *J. Funct. Anal.*, 180 :66–84, 2001.

[82] I. Gentil and S. Zugmeyer. A family of Beckner inequalities under various curvature-dimension conditions. ArXiv 1903.00214, 2019.

[83] N. Gozlan and C. Léonard. Transport inequalities. A survey. *Markov Process. Related Fields*, 16 :635–736, 2010.

[84] N. Gozlan, C. Roberto, P.-M. Samson, and P. Tetali. Displacement convexity of entropy and related inequalities on graphs. *Probab. Theory Related Fields*, 160 :47–94, 2014.

[85] M. Gromov. Paul Lévy's isoperimetric inequality. Preprint I.H.E.S., 1979.

[86] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Math.* Birkhäuser, 1999.

[87] M. Gromov and V. Milman. A topological application of the isoperimetric inequality. *Amer. J. Math.*, 105 :843–854, 1983.

[88] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97 :1061–1083, 1975.

[89] A. Guillin and A. Joulin. Measure concentration through non-Lipschitz observables and functional inequalities. *Electron. J. Probab.*, 18 :1–26, 2013.

[90] A. Guillin, C. Léonard, L. Wu, and N. Yao. Transportation inequalities for Markov processes. *Probab. Theory Related Fields*, 144 :669–695, 2009.

[91] D.L. Hanson and F.T. Wright. A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Statist.*, 42 :1079–1083, 1971.

[92] G. Hargé. Reinforcement of an inequality due to Brascamp and Lieb. *J. Funct. Anal.*, 254 :2008, 267–300.

[93] B. Helffer. Remarks on decay of correlations and Witten Laplacians, Brascamp-Lieb inequalities and semiclassical limit. *J. Funct. Anal.*, 155 :571–586, 1998.

[94] B. Helffer. Remarks on decay of correlations and Witten Laplacians III. Applications to logarithmic Sobolev inequalities. *Ann. Inst. H. Poincaré Probab. Statist.*, 35 :483–508, 1999.

[95] B. Helffer. *Semiclassical analysis, Witten Laplacians, and statistical mechanics*, volume 1 of *Series in Partial Differential Equations and Applications*. World Scientific Publishing, River Edge, 2002.

[96] B. Helffer and F. Nier. Criteria to the Poincaré inequality associated with Dirichlet forms in $\mathbb{R}^d$, $d \geq 2$. *Int. Math. Res. Not.*, 22 :1199–1223, 2003.

[97] B. Helffer and J. Sjöstrand. On the correlation for Kac-like models in the convex case. *J. Stat. Phys.*, 74 :349–409, 1994.

[98] E. Hillion. Concavity of entropy along binomial convolutions. *Electron. Commun. Probab.*, 17 :1–9, 2012.

[99] E. Hillion. Contraction of measures on graphs. *Pot. Anal.*, 41 :679–698, 2014.

[100] R. Holley and D.W. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Stat. Phys.*, 46 :1159–1194, 1987.

[101] L. Hörmander. L^2 estimate and existence theorems for the $\bar{\partial}$ operator. *Acta Math.*, 113 :89–152, 1965.

[102] C. Houdré. Remarks on deviation inequalities for functions of infinitely divisible random vectors. *Ann. Probab.*, 30 :1223–1237, 2002.

[103] C. Houdré and P. Marchal. On the concentration of measure phenomenon for stable and related random vectors. *Ann. Probab.*, 32 :1496–1508, 2004.

[104] C. Houdré and N. Privault. Concentration and deviation inequalities in infinite dimensions via covariance representations. *Bernoulli*, 8 :697–720, 2002.

[105] C. Houdré and P. Tetali. Concentration of measure for products of Markov kernels and graph products via functional inequalities. *Combin. Probab. Comput.*, 10 :1–28, 2001.

[106] E.P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002.

[107] N. Huet. Spectral gap for some invariant log-concave probability measures. *Mathematika*, 57 :51–62, 2011.

[108] J. Johnsen. On the spectral properties of Witten-Laplacians, their range projections and Brascamp-Lieb's inequality. *Integr. equ. oper. theory*, 36 :288–324, 2000.

[109] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. math. Anal.*, 29 :1–17, 1998.

[110] J. Jost and S. Liu. Ollivier's Ricci curvature, local clustering and curvature-dimension inequalities on graphs. *Discrete Comput. Geom.*, 51 :300–322, 2014.

[111] A. Joulin. Poisson-type deviation inequalities for curved continuous-time Markov chains. *Bernoulli*, 13 :782–798, 2007.

[112] A. Joulin. A new Poisson-type deviation inequality for Markov jump processes with positive Wasserstein curvature. *Bernoulli*, 15 :532–549, 2009.

[113] A. Joulin and Y. Ollivier. Curvature, concentration and error estimates for Markov chain Monte-Carlo. *Ann. Probab.*, 38 :2418–2442, 2010.

[114] A. Joulin and N. Privault. Functional inequalities for discrete gradients and applications to the geometric distribution. *ESAIM Probab. Stat.*, 8 :87–101, 2004.

[115] A. Joulin and N. Privault. A logarithmic Sobolev inequality for an interacting spin system under a geometric reference measure. In *Proc. of 26th QPIDA (Levico, 2005)*, volume XX of *Quantum Probability and White Noise Analysis*, pages 267–273. World Scientific, 2007.

[116] E. Kannan, L. Lovász, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.*, 13 :541–559, 1995.

[117] S. Karlin and W.J. Studden. *Tchebycheff systems : with applications in analysis and statistics*, volume 15 of *Pure and applied mathematics*. Interscience Publishers, 1966.

[118] B. Klartag. A Berry-Esseen type inequality for convex bodies with an inconditional basis. *Probab. Theory Related Fields*, 145 :1–33, 2009.

[119] B. Klartag, G. Kozma, P. Ralli, and P. Tetali. Discrete curvature and abelian groups. *Canad. J. Math.*, 68 :655–674, 2016.

[120] A.V. Kolesnikov and E. Milman. Riemannian metrics on convex sets with applications to Poincaré and log-Sobolev inequalities. *Calc. Var. Partial Differ. Equ.*, 55 :77, 2016.

[121] A.V. Kolesnikov and E. Milman. Brascamp-Lieb-type inequalities on weighted Riemannian manifolds with boundary. *J. Geom. Anal.*, 27 :1680–1702, 2017.

[122] A.V. Kolesnikov and E. Milman. The KLS isoperimetric conjecture for generalized Orlicz balls. *Ann. Probab.*, 46 :3578–3615, 2018.

[123] K. Kuwada. Duality on gradient estimates and Wasserstein controls. *J. Funct. Anal.*, 258 :3758–3774, 2010.

[124] R. Latala and K. Oleszkiewicz. Between Sobolev and Poincaré. In *Geometric aspects of functional analysis*, volume 1745 of *Lecture Notes in Math.*, pages 147–168. Springer, 2000.

[125] R. Latala and J.O. Wojtaszczyk. On the infimum convolution inequality. *Studia Math.*, 189 :147–187, 2008.

[126] M. Ledoux. Remarks on logarithmic Sobolev constants, exponential integrability and bounds on the diameter. *J. Math. Kyoto Univ.*, 35 :211–220, 1995.

[127] M. Ledoux. On Talagrand's deviation inequalities for product measures. *ESAIM Probab. Statist.*, 1 :63–87, 1996.

[128] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 120–216. Springer, 1999.

[129] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2001.

[130] M. Ledoux. Logarithmic Sobolev inequalities for unbounded spin systems revisited. In *Séminaire de probabilités, XXXV*, volume 1755 of *Lecture Notes in Math.*, pages 167–194. Springer, 2001.

[131] M. Ledoux. Spectral gap, logarithmic Sobolev constant, and geometric bounds. In *Surveys in Differential Geometry*, volume 9, pages 209–240. International Press, 2004.

[132] M. Ledoux, I. Nourdin, and G. Peccati. Stein's method, logarithmic Sobolev and transport inequalities. *Geom. Funct. Anal.*, 25 :256–306, 2015.

[133] Y.T. Lee and S.S. Vempala. Eldan's stochastic localization and the KLS conjecture : isoperimetry, concentration and mixing. ArXiv 1612.01507, 2019 (preliminary version : 2016).

[134] C. Léonard. Lazy random walks and optimal transport on graphs. *Ann. Probab.*, 44 :1864–1915, 2016.

[135] C. Léonard. On the convexity of the entropy along entropic interpolations. In *Measure Theory in Non-Smooth Spaces*, Partial Differential Equations and Measure Theory, pages 195–242. De Gruyter, 2017.

[136] P. Lezaud. Chernoff-type bound for finite Markov chains. *Ann. Appl. Probab.*, 8 :849–867, 1998.

[137] P. Lezaud. Chernoff and Berry-Esséen inequalities for Markov processes. *ESAIM Probab. Statist.*, 5 :183–201, 2001.

[138] P. Li and S.T. Yau. On the parabolic kernel of the Schrödinger operator. *Acta Math.*, 156 :153–201, 1986.

[139] Y. Lin and S.T. Yau. Ricci curvature and eigenvalue estimate on locally finite graphs. *Math. Res. Lett.*, 17 :343–356, 2010.

[140] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. Math.*, 169 :903–991, 2009.

[141] J. Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal*, 261 :2250–2292, 2011.

[142] K. Marton. A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.*, 6 :556–571, 1996.

[143] B. Maurey. Some deviations inequalities. *Geom. Funct. Anal.*, 1 :188–197, 1991.

[144] G. Menz and F. Otto. Uniform logarithmic Sobolev inequalities for conservative spin systems with super-quadratic single-site potential. *Ann. Probab.*, 41 :2182–2224, 2013.

[145] S.P. Meyn and R.L. Tweedie. Stability of Markovian processes III : Foster-Lyapunov criteria for continuous-time processes. *Adv. Appl. Probab.*, 25 :518–548, 1993.

[146] L. Miclo. Quand est-ce que des bornes de Hardy permettent de calculer une constante de Poincaré exacte sur la droite. *Ann. Fac. Sci. Toulouse Math.*, 17 :121–192, 2008.

[147] A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24 :1329–1346, 2011.

[148] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Part. Diff. Equ.*, 48 :1–31, 2013.

[149] E. Milman. On the role of convexity in isoperimetry, spectral gap and concentration. *Invent. Math.*, 177 :1–43, 2009.

[150] E. Milman. Spectral estimates, contractions and hypercontractivity. *J. Spectr. Theory*, 8 :669–714, 2018.

[151] V.D. Milman. A new proof of the theorem of A. Dvoretzky on sections of convex bodies. *Funct. Anal. Appl.*, 5 :28–37, 1971.

[152] R. Montenegro and P. Tetali. Mathematical aspects of mixing times in Markov chains. *Found. Trends Theor. Comput. Sci.*, 1 :237–354, 2006.

[153] B. Muckenhoupt. Hardy's inequality with weights. *Studia Math.*, 44 :31–38, 1972.

[154] V.H. Nguyen. Dimensional variance inequalities of Brascamp-Lieb type and a local approach to dimensional Prékopa's theorem. *J. Funct. Anal.*, 266 :931–955, 2014.

[155] Y. Ollivier. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.*, 256 :810–864, 2009.

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

[157] D. Paulin. Mixing and concentration by Ricci curvature. *J. Funct. Anal.*, 270 :1623–1662, 2016.

[158] G.O. Roberts and J.S. Rosenthal. General state space Markov chains and MCMC algorithms. *Probab. Surv.*, 1 :20–71, 2004.

[159] D. Rudolf. Explicit error bounds for lazy reversible Markov chain Monte Carlo. *J. Complexity*, 25 :11–24, 2009.

[160] D. Rudolf. Error bounds for computing the expectation by Markov chain Monte Carlo. *Monte Carlo Meth. Appl.*, 16 :323–342, 2010.

[161] M.D. Sammer. Aspects of mass transportation in discrete concentration inequalities. PhD Thesis, Georgia Institute of Technology. https ://smartech.gatech.edu/handle/1853/7006, 2005.

[162] M. Schmuckenschläger. Martingales, Poincaré type inequalities, and deviation inequalities. *J. Funct. Anal.*, 155 :303–323, 1998.

[163] M. Schmuckenschläger. Curvature of nonlocal Markov generators. In *Convex geometric analysis*, volume 34 of *Math. Sci. Res. Inst. Publ.*, pages 189–197. Cambridge Univ. Press, 1999.

[164] S. Sodin. An isoperimetric inequality on the l_p balls. *Ann. Inst. H. Poincaré Probab. Statist.*, 44 :362–373, 2008.

[165] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proc. 6th Berkeley Symp. on Math. Statist. and Prob.*, volume 2, pages 583–602. Univ. of Calif. Press, 1972.

[166] K.-T. Sturm. On the geometry of metric measure spaces. *Acta. Math.*, 196 :65–177, 2006.

[167] K.-T. Sturm. Metric measure spaces with variable Ricci bounds and couplings of Brownian motions. In *Festschrift Masatoshi Fukushima*, volume 17 of *Interdisciplinary Mathematical Sciences*, pages 553–575. World Scientific, 2015.

[168] K.-T. Sturm and M.K. VonRenesse. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58 :923–940, 2005.

[169] M.T. Subbotin. On the law of frequency of error. *Mat. Sb.*, 31 :296–301, 1923.

[170] M. Talagrand. A new isoperimetric inequality and the concentration of measure phenomenon. In *Geometric aspects of functional analysis*, volume 1469 of *Leture Notes in Math.*, pages 94–124. Springer, 1991.

[171] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. volume 81 of *Inst. Hautes Études Sci. Publ. Math.*, pages 73–205. 1995.

[172] M. Talagrand. A new look at independence. *Ann. Probab.*, 24 :1–34, 1996.

[173] E. VanDoorn. Conditions for exponential ergodicity and bounds for the decay parameter of a birth-death process. *Adv. Appl. Probab.*, 17 :514–530, 1985.

[174] L. Veysseire. A harmonic mean bound for the spectral gap of the Laplacian on Riemannian manifolds. *C. R. Acad. Sci. Paris Sér. I*, 348 :1319–1322, 2010.

[175] L. Veysseire. Coarse Ricci curvature for continuous-time Markov processes. ArXiv 1202.0420, 2012.

[176] L. Veysseire. A concentration theorem for the equilibrium measure of Markov chains with non negative coarse Ricci curvature. ArXiv 1203.2897, 2012.

[177] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, 2003.

[178] D. Villemonais. Lower bound for the coarse Ricci curvature of continuous-time pure-jump processes. *J. Theor. Probab.*, First online :1–38, 2019.

[179] F.Y. Wang. Functional inequalities, semigroup properties and spectrum estimates. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 3 :263–295, 2000.

[180] H.F. Weinberger. An isoperimetric inequality for the n-dimensional free membrane problem. *J. Ration. Mech. Anal.*, 5 :633–636, 1956.

[181] L. Wu. A deviation inequality for non-reversible Markov processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 36 :435–445, 2000.

[182] L. Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Related Fields*, 118 :427–438, 2000.

[183] L. Wu. Uniformly integrable operators and large deviations for Markov processes. *J. Funct. Anal.*, 172 :301–376, 2000.