Concentration et fluctuations de processus stochastiques avec sauts

Aldéric Joulin

THÈSE DE DOCTORAT de l'UNIVERSITÉ DE LA ROCHELLE

Discipline : Mathématiques

Spécialité : Probabilités

Présentée par **M. ALDÉRIC JOULIN**

pour obtenir le grade de **DOCTEUR de l'UNIVERSITÉ DE LA ROCHELLE.**

Titre de la thèse :

CONCENTRATION ET FLUCTUATIONS DE PROCESSUS STOCHASTIQUES AVEC SAUTS

soutenue publiquement le 6 octobre 2006 devant le jury composé de :

M. Piotr GRACZYK (Université d'Angers) - Examinateur

M. Christian HOUDRÉ (Georgia Institute of Technology - Atlanta, USA) - Rapporteur

M. Christian LÉONARD (Université de Paris X - Nanterre) - Rapporteur

M. Nicolas PRIVAULT (Université de La Rochelle) - Directeur de thèse

M. Emmanuel RIO (Université de Versailles) - Président du jury

M. Liming WU (Université de Clermont-Ferrand II) - Examinateur

Résumé

Cette thèse est constituée de deux parties indépendantes, le premier thème traitant du phénomène de concentration de la mesure pour des processus de naissance et de mort, tandis que le second est consacré aux fluctuations des intégrales stochastiques dirigées par des processus stables.
Dans la première partie de la thèse, nous explorons le phénomène de concentration des processus de naissance et de mort. Les différentes approches considérées sont d'une part les inégalités fonctionnelles ainsi que la méthode de Herbst, et d'autre part l'étude des propriétés du semigroupe associé et des techniques de martingales. En particulier, nous sommes amenés à introduire diverses notions de courbures de ces processus, analogues discrets du critère de courbure de Bakry-Emery dans le cadre des processus de diffusion.
Dans la deuxième partie de la thèse, nous étudions le comportement du processus supremum d'une intégrale stable stochastique en établissant des inégalités maximales que nous appliquons à des problèmes de temps de passage de processus symétriques stables. Enfin, nous démontrons un principe de domination convexe pour des intégrales stochastiques brownienne et stable corrélées.
Mots-clés : processus de naissance et de mort, phénomène de concentration, courbure discrète, intégrale stable stochastique, inégalité maximale, domination convexe.

Abstract

This PhD Thesis is divided into two independent parts, the first one dealing with the concentration of measure phenomenon for birth-death processes, whereas the second one is devoted to the analysis of the fluctuations of stochastic integrals driven by stable processes.
In the first part of the thesis, we explore the measure concentration for birth-death processes. The various approaches considered are on the one hand the functional inequalities together with the Herbst method, and one the other hand the study of the associated semigroup and martingales techniques. In particular, we are led to introduce some notions about curvatures of such processes, which are the discrete analogous of the Bakry-Emery curvature criterion given for diffusion processes.
In the second part of the thesis, we study the behavior of the supremum process of a stable stochastic integral by providing maximal inequalities which are applied to passage time problems of symmetric stable processes. Finally, we derive a convex domination principle for dependent Brownian and stable stochastic integrals.
Keywords : birth-death process, concentration phenomenon, discrete curvature, stable stochastic integral, maximal inequality, convex domination.

Remerciements

En premier lieu, je tiens à remercier profondément Nicolas Privault pour avoir encadré cette thèse de doctorat et m'avoir témoigné sa confiance. Durant ces trois années, j'ai largement bénéficié de ses conseils avisés ainsi que de son intuition pour les mathématiques. De plus, il m'a toujours encouragé à suivre la voie de l'autonomie, grâce notamment aux différents voyages scientifiques auxquels il m'a fait participer. Pour toutes ces raisons, je lui suis infiniment reconnaissant.

Je remercie sincèrement Christian Houdré et Chistian Léonard pour avoir accepté d'être les rapporteurs de cette thèse, ainsi que Piotr Graczyk, Emmanuel Rio et Liming Wu qui me font l'honneur de composer le jury.

Un grand merci au ténébreux Jean-Christophe Breton, que j'ai sollicité un nombre incalculable de fois pour qu'il relise patiemment mes textes parfois erronés, toujours un peu brouillon. Je souhaite également exprimer ma reconnaissance à Djelil Chafaï ainsi qu'à Florent Malrieu que j'ai eu le plaisir de rencontrer au tout début de ma thèse dans le désormais célèbre établissement "Le Viking" de Saint-Flour, et qui ont gentiment essayé de répondre aux questions que je leur avais posées.

À vous les Anthony, Delphine et Ma, avec qui j'ai partagé le bureau "probabilistes débutants" dont la porte exhibe encore nos portraits de malfrats, je compte sur vous pour continuer à faire vivre notre groupe de travail étudiant.

Je tiens évidemment à rendre hommage à mes amis thésards (et assimilés) avec qui j'ai partagé à de nombreuses reprises le(s) verre(s) de l'amitié aux Z'endich's (et ailleurs pour certains), et qui se reconnaîtront : Bobiche la frange, Csainj, Charles et Claire, Élodie, Fabrice coeur de rocker, le Forbin, fragile Granto, Henri petit Belgacom, Jf, Nicolas et Émilie, Pedrolito, Salvijus, ainsi que J.-C. et les footeux du mardi midi ...

Malgré la distance qui nous sépare, je ne peux oublier de mentionner mes vieux potes, dont la contribution indirecte à la préparation de cette thèse est monumentale : Anthony, Camille, le petit Draleb, Johanna, le beau Julien, le encore plus beau Julien, l'adorable Lucie, Spoon, la délicieuse Stéphanie et enfin seigneur Toch-Ap 21st.

Enfin, il est temps pour moi de clore ces quelques lignes en dédiant ce travail à ma petite Muriel à moi ainsi qu'à mes parents et à ma famille, d'ici et d'ailleurs, et sans qui tout ceci n'aurait aucun sens ...

À un certain A. M., pour l'éternité ...

Table des matières

Chapitre 0

Présentation

0.1 Concentration de processus de naissance et de mort

L'objectif de la première partie de cette thèse est d'étudier le phénomène de concentration des processus de naissance et de mort en tenant compte des difficultés posées par le cadre discret. Deux approches de ce problème sont envisagées : la première consiste à établir des inégalités fonctionnelles affaiblies et d'utiliser la méthode de Herbst, tandis que la deuxième repose sur l'analyse du semigroupe associé, à travers l'obtention de bornes sur les courbures discrètes du processus, et des techniques de martingales.

Depuis le début des années soixante-dix, l'étude des inégalités de déviation pour les mesures de probabilité, plus communément appelées phénomène de concentration de la mesure, a été source d'explorations diverses et variées. Popularisées par de nombreux auteurs au milieu des années quatre-vingt-dix et synthétisées dans le livre de Ledoux [62], les méthodes fonctionnelles et probabilistes utilisées afin d'établir de tels résultats font fortement intervenir les propriétés ergodiques des semigroupes de Markov.
Commençons par introduire quelques notions sur les processus et semigroupes de Markov. Soit $(E, \mathscr{F})$ un espace mesurable et soit $(X_t)_{t\geq 0}$ un processus de Markov à valeurs dans cet espace et de distribution stationnaire μ. On désigne par $(P_t)_{t\geq 0}$ le semigroupe associé et $\mathcal{L}$ son générateur infinitésimal défini sur un domaine dense de $L^2(\mu)$, noté $\mathrm{Dom}\,\mathcal{L}$ (on supposera par simplicité qu'il contient les fonctions constantes et qu'il est stable par les actions de $\mathcal{L}$, P_t et par composition avec les fonctions C^∞). Définissons les formes bilinéaires symétriques suivantes : pour des fonctions réelles $f, g \in \mathrm{Dom}\,\mathcal{L}$, on introduit l'opérateur *carré du champ* Γ par

$$\Gamma(f,g) := \frac{1}{2}\left(\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f\right), \tag{0.1.1}$$

et l'opérateur *carré du champ itéré* Γ_2 par

$$\Gamma_2(f,g) := \frac{1}{2}\left(\mathcal{L}\Gamma(f,g) - \Gamma(f,\mathcal{L}g) - \Gamma(g,\mathcal{L}f)\right).$$

On notera dans la suite $\Gamma f := \Gamma(f,f)$ et $\Gamma_2 f := \Gamma_2(f,f)$. La forme de Dirichlet associée est donnée par

$$\mathscr{E}_\mu(f,g) := \int \Gamma(f,g) d\mu.$$

Le processus est dit ergodique si sa distribution converge vers la mesure d'équilibre, c'est-à-dire pour toute fonction f suffisamment régulière, on a $\lim_{t\to+\infty} P_t f = \int f d\mu$. On peut déterminer dans quel espace et à quelle vitesse a lieu cette convergence en étudiant certaines inégalités fonctionnelles que nous introduisons maintenant.

Définition 0.1.1. *La mesure de probabilité μ satisfait une inégalité de Poincaré de constante $\lambda > 0$ si pour toute fonction f suffisamment régulière,*

$$\lambda \int \left(f - \int f d\mu \right)^2 d\mu \leq \mathscr{E}_\mu(f,f). \tag{0.1.2}$$

De même, μ vérifie une inégalité de Sobolev logarithmique de constante $C > 0$ si pour toute fonction f suffisamment régulière,

$$C \int f^2 \log f^2 d\mu - \int f^2 d\mu \log \int f^2 d\mu \leq \mathscr{E}_\mu(f,f). \tag{0.1.3}$$

Par exemple, une inégalité de Poincaré (resp. une inégalité de Sobolev logarithmique) satisfaite par la distribution stationnaire μ nous assure une vitesse de convergence exponentielle du semigroupe vers l'équilibre dans $L^2(\mu)$ (resp. une décroissance exponentielle entropique du semigroupe).
Dans la suite, on parlera d'inégalité fonctionnelle locale lorsqu'à $t > 0$ fixé, la mesure de probabilité $P_t(\cdot)(x)$ satisfait cette inégalité fonctionnelle uniformément en la condition initiale $x \in E$. Bien que plus difficiles à établir en général, il est plus intéressant d'obtenir des inégalités locales que des inégalités directement pour la distribution stationnaire car elles fournissent plus d'informations sur le comportement du processus.

Tout au long de cette partie, il nous a paru intéressant de faire le parallèle entre notre étude et celle concernant le cas continu des processus de diffusion, pour lesquels un grand nombre de résultats est disponible. Un processus de diffusion est un processus de Markov à trajectoires continues et dont le générateur infinitésimal satisfait la propriété de dérivation en chaîne suivante : pour toute fonction $\phi \in C^\infty(\mathbb{R}^d, \mathbb{R})$ et $f = (f_1, \ldots, f_d) \in (\mathrm{Dom}\, \mathcal{L})^d$,

$$\mathcal{L}\phi(f) = \sum_i \phi'_i(f) \mathcal{L} f_i + \sum_{i,j} \phi''_{i,j}(f) \Gamma(f_i, f_j). \tag{0.1.4}$$

Énonçons à présent le théorème fondamental suivant, tiré de l'article de Bakry [6]. Ce résultat fait le lien entre les diverses notions introduites précédemment : un calcul fonctionnel Γ_2, des relations de sous-commutation entre le semigroupe et l'opérateur Γ, et des inégalités fonctionnelles locales :

Théorème 0.1.2. *Soit ρ un nombre réel. Les assertions suivantes sont équivalentes :*

(i) Pour toute fonction $f \in \mathrm{Dom}\,\mathcal{L}$*, on a* $\Gamma_2 f \geq \rho \Gamma f$*.*

(ii) Pour toute fonction $f \in \mathrm{Dom}\,\mathcal{L}$ *et tout* $t > 0$*, la relation faible de sous-commutation* $\Gamma P_t f \leq e^{-2\rho t} P_t \Gamma f$ *est satisfaite.*

(iii) Le semigroupe vérifie une inégalité de Poincaré locale de constante $\frac{\rho}{1-e^{-2\rho t}}$*.*

Si de plus le processus est de diffusion, alors les assertions précédentes sont encore équivalentes à :

(iv) Pour toute fonction $f \in \mathrm{Dom}\,\mathcal{L}$*, on a* $\Gamma_2 f - \rho \Gamma f \geq \Gamma \sqrt{\Gamma f}$*.*

(v) Pour toute fonction $f \in \mathrm{Dom}\,\mathcal{L}$ *et tout* $t > 0$*, la relation forte de sous-commutation* $\sqrt{\Gamma P_t f} \leq e^{-\rho t} P_t \sqrt{\Gamma f}$ *est satisfaite.*

(vi) Le semigroupe vérifie une inégalité de Sobolev logarithmique locale de constante $\frac{\rho}{2(1-e^{-2\rho t})}$*.*

Lorsque ρ *est nul, on remplacera les constantes par leur limite lorsque* ρ *tend vers* 0*.*

Le processus de diffusion de base dans cette étude est le processus dit de Kolmogorov sur $\mathbb{R}^d$ solution de l'équation de Langevin

$$Y_t = Y_0 + \sqrt{2} B_t - \int_0^t \nabla V(Y_s) ds, \quad t \geq 0,$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien standard sur $\mathbb{R}^d$ et V est un potentiel de classe C^2 tel que $\int \exp(-V(x))dx = 1$. C'est un processus de Markov ergodique dont la distribution (réversible) stationnaire est la mesure de Boltzmann γ de densité $\exp(-V)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$. Son générateur infinitésimal $\mathcal{L}$ est donné par

$$\mathcal{L} f = \Delta f - \nabla V \cdot \nabla f,$$

et l'opérateur Γ est le carré de la norme euclidienne du gradient $\Gamma f = |\nabla f|^2$, d'où la dénomination de *carré du champ* (de gradient). Enfin, l'opérateur markovien Γ_2 a l'expression

$$\Gamma_2 f = \|\nabla^2 f\|^2 + \nabla f \cdot \nabla^2 V \nabla f,$$

où $\nabla^2 f$ est la matrice hessienne de f. Ainsi, dans le cas des processus de Kolmogorov, les assertions équivalentes (i), (ii) et (v) du théorème 0.1.2 se réécrivent respectivement de la façon suivante :

(i') ρ minore le spectre de la matrice $\nabla^2 V$.

(ii') Pour toute fonction $f \in \mathrm{Dom}\,\mathcal{L}$ et tout $t > 0$, on a $|\nabla P_t f|^2 \leq e^{-2\rho t} P_t |\nabla f|^2$.

(v') Pour toute fonction $f \in \mathrm{Dom}\,\mathcal{L}$ et tout $t > 0$, on a $|\nabla P_t f| \leq e^{-\rho t} P_t |\nabla f|$.

Ainsi, l'assertion (i') traduit la comparaison, en termes de convexité de potentiel, d'un processus de Kolmogorov avec le processus d'Ornstein-Uhlenbeck le plus proche, de potentiel donné par $V(x) := \rho|x|^2/2$. Remarquons par ailleurs que ce dernier processus vérifie la relation de commutation

$$\nabla P_t f = e^{-\rho t} P_t \nabla f. \tag{0.1.5}$$

On dit alors que le nombre ρ est la courbure exacte de ce processus de diffusion.

Sous l'inégalité de Sobolev logarithmique de la condition (vi) du théorème 0.1.2, l'utilisation de la méthode de Herbst entraîne pour les processus de diffusion le résultat de concentration ci-dessous :

Corollaire 0.1.3. *Supposons qu'un processus de diffusion $(X_t)_{t\geq 0}$ vérifie une inégalité de Sobolev logarithmique locale de constante $\frac{\rho}{2(1-e^{-2\rho t})}$. Soit $f : \mathbb{R}^d \to \mathbb{R}$ une fonction lipschitzienne de constante $\|f\|_{\mathrm{Lip}} \leq 1$. Alors pour tout niveau de déviation $y > 0$ et tout $t > 0$, l'inégalité de concentration gaussienne locale suivante est satisfaite :*

$$\mathbb{P}_x(|f(X_t) - \mathbb{E}_x[f(X_t)]| \geq y) \leq 2\exp\left(-\frac{\rho y^2}{2(1-e^{-2\rho t})}\right).$$

En particulier, si $\rho > 0$, alors par ergodicité en passant à la limite lorsque t tend vers l'infini, on a sous la distribution stationnaire μ l'inégalité de concentration :

$$\mu\left(\left|f - \int f d\mu\right| \geq y\right) \leq 2\exp\left(-\frac{\rho y^2}{2}\right), \quad y > 0.$$

Ainsi, l'obtention d'une inégalité de Sobolev logarithmique locale pour les processus de diffusion permet d'établir des résultats de concentration point par point, qui sont valables pour la distribution stationnaire dans le cas où le paramètre ρ du théorème 0.1.2 est strictement positif. En particulier, l'équivalence entre les assertions (i) et (vi) donne dans le cas des processus de diffusion une caractérisation explicite de l'inégalité de Sobolev logarithmique locale. L'idée principale de la démonstration de cette inégalité repose sur la propriété de diffusion (0.1.4) du générateur infinitésimal. Cependant, les gradients discrets ne satisfaisant pas la règle de dérivation en chaîne, cette propriété n'est pas valable dans le cas des processus de naissance et de mort. D'autre part, lorsque l'espace d'état n'est pas fini mais seulement dénombrable, il est bien connu que certaines distributions stationnaires de ces processus, comme par exemple la mesure de Poisson pour la file d'attente markovienne $M/M/\infty$, peuvent se concentrer plus faiblement que la distribution gaussienne. La méthode de Herbst pouvant être adaptée à des gradients aux différences en utilisant des inégalités du type accroissements finis, il n'est pas surprenant que l'inégalité de Sobolev logarithmique (0.1.3) soit mise en défaut. Ainsi, il est nécessaire d'élaborer d'autres critères adaptés au cadre discret nous permettant d'établir un certain nombre de résultats autour de concentration non gaussienne.

Processus de naissance et de mort

Considérons un processus de naissance et de mort $(X_t)_{t\geq 0}$, à valeurs dans $E = \mathbb{N}$ ou dans l'ensemble $E = \{0, \ldots, n\}$, et de distribution stationnaire π. Son générateur infinitésimal est donné pour toute fonction $f : E \to \mathbb{R}$ par

$$\mathcal{L}f(x) = \lambda_x \left(f(x+1) - f(x)\right) + \nu_x \left(f(x-1) - f(x)\right), \quad x \in E, \tag{0.1.6}$$

où les fonctions de transition λ et ν sont strictement positives, sauf en l'état 0 qui est réfléchissant, c'est-à-dire $\nu_0 = 0$, conditions assurant l'irréductibilité de la chaîne (dans le cas fini, l'état n est aussi supposé réfléchissant). La fonction de transition ν est assimilée à une force de rappel car elle permet de ramener le processus vers 0. L'opérateur carré du champ Γ donné par (0.1.1) vérifie pour tout $x \in E$,

$$2\Gamma(f,g) = \lambda_x(f(x+1) - f(x))(g(x+1) - g(x)) + \nu_x(f(x-1) - f(x))(g(x-1) - g(x)).$$

Désignons par $(P_t)_{t\geq 0}$ le semigroupe homogène dont les probabilités de transition sont données pour tout $x \in E$ par

$$P_t(x,y) = \begin{cases} \lambda_x t + o(t) & \text{si } y = x+1, \\ \nu_x t + o(t) & \text{si } y = x-1, \\ 1 - (\lambda_x + \nu_x)t + o(t) & \text{si } y = x. \end{cases}$$

Le semigroupe $(P_t)_{t\geq 0}$ est contractif dans $L^\infty(\pi)$ et dans $L^1(\pi)$, donc par interpolation dans tous les espaces $L^p(\pi)$, $p \in [1, +\infty]$. En particulier, si la distribution stationnaire satisfait la condition de moment

$$\sum_{y\in E} \varrho(x,y)\pi(y) < +\infty, \quad x \in E, \tag{0.1.7}$$

où $\varrho : E \times E \to [0, +\infty)$ est une distance quelconque sur E, alors le semigroupe est bien défini sur l'espace Lip_ϱ des fonctions ϱ-lipschitziennes $f : E \to \mathbb{R}$, muni de la seminorme de Lipschitz

$$\|f\|_{\mathrm{Lip}_\varrho} := \sup_{x,y\in E} \frac{|f(x) - f(y)|}{\varrho(x,y)} < +\infty.$$

Si ϱ est la distance classique d donnée par $d(x,y) := |x-y|$, $x, y \in E$, on notera dans la suite $\mathrm{Lip} := \mathrm{Lip}_d$ et $\|\cdot\|_{\mathrm{Lip}} := \|\cdot\|_{\mathrm{Lip}_d}$.

Courbures discrètes

D'après le théorème 0.1.2, vérifier que la propriété fonctionnelle (i) est satisfaite revient à établir une inégalité de Poincaré locale, qui est strictement plus faible qu'une inégalité de Sobolev logarithmique locale pour des gradients aux différences. Ce type d'argument ne nous permet donc pas d'utiliser cette dernière inégalité ainsi que la méthode de Herbst afin d'établir des résultats de concentration pour des processus avec sauts. Ainsi, l'idée est d'introduire de nouveaux critères adaptés au cadre discret et faisant intervenir comme dans le théorème 0.1.2 des semigroupes, des gradients aux différences et des opérateurs fonctionnels markoviens. Ces critères sont nommés les *courbures discrètes* des processus avec sauts. Cependant, la multitude de gradients aux différences que l'on peut considérer nous impose de définir plusieurs courbures discrètes qui ne sont pas comparables. Dans tout ce qui suit, nous allons nous concentrer sur le cas des processus de naissance et de mort, bien que l'on puisse aussi définir ce type de courbures discrètes pour des processus plus généraux. Commençons tout d'abord par la courbure de Wasserstein associée à une distance ϱ sur l'espace d'état E.

Définition 0.1.4. *On suppose que la distribution stationnaire π vérifie la condition de moment (0.1.7) avec une distance ϱ. La ϱ-courbure de Wasserstein au temps $t > 0$ du processus de naissance et de mort $(X_t)_{t\geq 0}$ est définie par*

$$\alpha_t := -\frac{1}{t} \sup \left\{ \log \left(\frac{\|P_t f\|_{\mathrm{Lip}_\varrho}}{\|f\|_{\mathrm{Lip}_\varrho}} \right) : f \in \mathrm{Lip}_\varrho,\ f \neq \text{constante} \right\} \quad \in [-\infty, +\infty),$$

et elle est dite minorée par $\alpha \in \mathbb{R}$ si $\inf_{t>0} \alpha_t \geq \alpha$. Autrement dit, pour toute fonction $f \in \mathrm{Lip}_\varrho$ et tout $t > 0$, le semigroupe P_t est ϱ-lipschitzien au sens suivant :

$$\|P_t f\|_{\mathrm{Lip}_\varrho} \leq e^{-\alpha t} \|f\|_{\mathrm{Lip}_\varrho}. \tag{0.1.8}$$

La stricte positivité de α implique que le processus $(X_t)_{t\geq 0}$ possède de bonnes propriétés d'ergodicité. En effet, par le théorème de dualité de Kantorovich-Rubinstein, cf. le théorème 5.10 de [29], la ϱ-courbure de Wasserstein du processus est minorée par α si et seulement si

$$W_\varrho(P_t(x,\cdot), P_t(y,\cdot)) \leq e^{-\alpha t} \varrho(x,y), \quad x, y \in E, \quad t > 0, \tag{0.1.9}$$

où W_ϱ est la distance de Wasserstein sur l'ensemble des mesures de probabilité sur E munie de la fonction de coût ϱ. Il en résulte par le théorème 5.23 de [29] que si α est strictement positif, le semigroupe $(P_t)_{t\geq 0}$ converge à vitesse exponentielle vers la distribution stationnaire π dans la distance de Wasserstein W_ϱ.
Notons qu'une version de l'inégalité (0.1.9) est introduite par Marton dans [66] pour des chaînes de Markov à temps discret, avec la métrique triviale $\varrho(x,y) = 1_{x\neq y}$, et par Djellout, Guillin et Wu à travers la condition (C1) de [35], afin d'établir des inégalités de transport pour des suites de variables aléatoires faiblement dépendantes.

En comparant avec le cas des diffusions, cf. le théorème 0.1.2 et le corollaire 0.1.3, il est naturel de penser que la stricte positivité de α pourraît être un critère permettant d'obtenir des bornes de concentration locale fournissant de l'information en temps grand, c'est-à-dire pouvant s'étendre à la distribution stationnaire π du processus de naissance et de mort $(X_t)_{t\geq 0}$.

D'autre part, on remarque que l'inégalité (0.1.8) n'est pas la version renforcée de la relation forte de sous-commutation apparaissant dans l'assertion (v) du théorème 0.1.2.

Ceci est dû à l'utilisation d'un gradient discret, induit par la distance ϱ, qui diffère du gradient donné par l'opérateur carré du champ. En effet, la seminorme de Lipschitz $\|\cdot\|_\varrho$ induit le gradient

$$|\nabla_\varrho f(x)| := \frac{|f(x+1)-f(x)|}{\varrho(x,x+1)}, \quad x \in E, \tag{0.1.10}$$

qui n'est pas comparable au gradient discret

$$\sqrt{\Gamma f(x)} = \sqrt{\lambda_x (f(x+1)-f(x))^2 + \nu_x (f(x-1)-f(x))^2}, \quad x \in E,$$

où les fonctions de transition du générateur infinitésimal jouent un rôle fondamental. Notons que dans le cas des processus de Kolmogorov, le gradient $\sqrt{\Gamma f}$ coïncide avec la longueur euclidienne $|\nabla f|$ du gradient usuel ∇.

Introduisons à présent la courbure discrète du processus de naissance et de mort $(X_t)_{t\geq 0}$ liée à l'opérateur carré du champ Γ. C'est l'analogue discret de l'assertion (v) du théorème 0.1.2.

Définition 0.1.5. *La Γ-courbure au temps $t > 0$ du processus de naissance et de mort $(X_t)_{t\geq 0}$ est définie par*

$$\rho_t := -\frac{1}{t} \sup \left\{ \log \left\| \frac{(\Gamma P_t f)^{1/2}}{P_t\,(\Gamma f)^{1/2}} \right\|_\infty : f \in \mathrm{Dom}\,\mathcal{L},\ f \neq \text{constante} \right\} \quad \in [-\infty, +\infty),$$

et elle est dite minorée par $\rho \in \mathbb{R}$ si $\inf_{t>0} \rho_t \geq \rho$. De manière équivalente, pour toute fonction $f \in \mathrm{Dom}\,\mathcal{L}$ et tout $t > 0$, la relation de sous-commutation suivante est vérifiée :

$$(\Gamma P_t f)^{1/2} \leq e^{-\rho t} P_t\,(\Gamma f)^{1/2}. \tag{0.1.11}$$

Remarquons que contrairement à l'inégalité (0.1.8), où sont comparées des seminormes de Lipschitz, la relation (0.1.11) est une inégalité entre opérateurs, donc donnée point par point. En d'autres termes, il s'agit d'une vraie relation de sous-commutation entre la racine carrée de l'opérateur carré du champ et le semigroupe markovien.

Courbures des files d'attente markoviennes $M/M/\infty$ et $M/M/1$

Dans le cadre des processus de naissance et de mort, un exemple de base que nous allons considérer tout au long de cette étude est la file d'attente markovienne $M/M/\infty$ sur $\mathbb{N}$ (la chaîne d'Ehrenfest à temps continu joue ce rôle dans le cas fini $\{0,\dots,n\}$). Tout d'abord, décrivons ce processus de queue. On considère un système de guichets de vente où chaque client arrivant est immédiatement servi par l'un des guichetiers. En notant X_t le nombre de clients dans le système à l'instant $t > 0$, on suppose que le processus d'arrivée des clients est un processus de Poisson d'intensité $\lambda > 0$ et que conditionnellement à l'événement $\{X_s = x\}$, le temps de service $T := \inf\{t > s : X_t \neq X_s\} - s$ suit une loi exponentielle de paramètre $\lambda + \nu x$, $\nu > 0$. Alors le processus aléatoire $(X_t)_{t\geq 0}$ est une file d'attente $M/M/\infty$. C'est un processus de naissance et de mort dont les fonctions de transition du générateur infinitésimal sont données par $\lambda_x = \lambda$ et $\nu_x = \nu x$, $x \in \mathbb{N}$. La chaîne est ergodique de mesure réversible la loi de Poisson $\mathscr{P}(\sigma)$ sur $\mathbb{N}$ de paramètre $\sigma := \lambda/\nu$, et sa distribution est donnée par la formule de convolution de type Mehler

$$\mathcal{L}(X_t|X_0 = x) = \mathscr{B}\left(x, e^{-\nu t}\right) * \mathscr{P}\left(\sigma(1-e^{-\nu t})\right), \quad t \geq 0, \tag{0.1.12}$$

où $\mathscr{B}(n,p)$ désigne une loi binomiale de paramètres $n \in \mathbb{N}$ et $p \in (0,1)$.
L'intérêt d'introduire cette file d'attente dans l'étude de la concentration des processus de naissance et de mort réside dans le fait que les calculs sont effectués de manière simple et explicite. En effet, par la convolution (0.1.12), on a pour tout $\tau > 0$,

$$\begin{aligned} \mathbb{E}_x\left[e^{\tau(X_t - \mathbb{E}_x[X_t])}\right] &= \exp\left\{x\log\left(1+e^{-\nu t}(e^\tau - 1)\right) - \tau x e^{-\nu t} + \sigma(1-e^{-\nu t})\left(e^\tau - \tau - 1\right)\right\} \\ &\leq \exp\left\{\left(xe^{-\nu t} + \sigma(1-e^{-\nu t})\right)\left(e^\tau - \tau - 1\right)\right\} \\ &= \exp\left\{\mathbb{E}_x[X_t]\ \left(e^\tau - \tau - 1\right)\right\}, \end{aligned}$$

où est utilisée dans l'inégalité précédente la relation $\log(1+x) \leq x$, $x > 0$. Ainsi, par l'inégalité de Chebychev, on obtient pour tout $y > 0$ l'inégalité de déviation locale :

$$\begin{aligned} \mathbb{P}_x\left(X_t - \mathbb{E}_x[X_t] \geq y\right) &\leq \inf_{\tau > 0} e^{-\tau y}\, \mathbb{E}_x\left[e^{\tau(X_t - \mathbb{E}_x[X_t])}\right] \\ &\leq \exp\left\{y - \left(\mathbb{E}_x[X_t] + y\right)\log\left(1 + \frac{y}{\mathbb{E}_x[X_t]}\right)\right\}, \end{aligned}$$

qui entraîne par ergodicité lorsque $t \to +\infty$ l'estimée

$$\mathbb{P}\left(X - \mathbb{E}[X] \geq y\right) \leq \exp\left\{y - (\sigma + y)\log\left(1 + \frac{y}{\sigma}\right)\right\},$$

où X désigne une variable aléatoire de Poisson de paramètre σ. Cette inégalité pouvant être immédiatement étendue aux fonctions d-lipschitziennes, on retrouve pour la loi de Poisson l'inégalité de déviation (13) établie par Bobkov et Ledoux dans [19].

D'autre part, la file d'attente $M/M/\infty$ est considérée comme l'analogue discret d'un processus d'Ornstein-Uhlenbeck réel. En effet, ce dernier peut être construit comme la limite fluide de files d'attente $M/M/\infty$ renormalisées, cf. [72]. Si l'on désigne par d^+ le gradient *forward* $\mathrm{d}^+ f(x) := f(x+1) - f(x)$, $x \in \mathbb{N}$, Chafaï a remarqué dans [25] que le semigroupe de ce processus de queue satisfaisait la relation de commutation

$$\mathrm{d}^+ P_t f = e^{-\nu t} P_t \mathrm{d}^+ f, \tag{0.1.13}$$

qui est la version discrète de la relation (0.1.5). On dit dans ce cas que le paramètre ν est la courbure exacte de ce processus de naissance et de mort. Par conséquent, comme lorsque l'on compare un processus de Kolmorogov avec le processus d'Ornstein-Uhlenbeck le plus proche dans le but d'établir des résultats de concentration, il est naturel dans notre cadre discret de considérer la file d'attente $M/M/\infty$ comme point de référence et de chercher des critères permettant de comparer des processus de naissance et de mort généraux avec ce processus de queue.

D'un autre côté, il existe des cas simples où les arguments développés dans le théorème 0.1.2 pour les diffusions ne peuvent pas être utilisés afin de donner des résultats de concentration locale fournissant de l'information en temps grand. En effet, il est d'un intérêt fondamental d'obtenir des inégalités locales dont les constantes se comportent bien lorsque le paramètre de temps s'approche de l'infini. Par exemple, dans le cas du processus de Kolmogorov sur $\mathbb{R}^d$ solution de l'équation différentielle stochastique

$$Y_t = Y_0 + B_t - \int_0^t \mathrm{sign}(Y_s)ds, \tag{0.1.14}$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien standard sur $\mathbb{R}^d$, les assertions du théorème 0.1.2 sont satisfaites avec le paramètre $\rho = 0$. Ainsi, le corollaire 0.1.3 nous donne pour toute fonction lipschitzienne $f : \mathbb{R}^d \to \mathbb{R}$ la concentration locale

$$\mathbb{P}_x(|f(Y_t) - \mathbb{E}_x[f(Y_t)]| \geq y) \leq 2\exp\left(-\frac{y^2}{4t}\right), \quad y > 0, \quad t > 0,$$

inégalité qui ne peut pas être étendue à la distribution stationnaire exponentielle

$$\mu(dx) = 2^{-d}\lambda^d e^{-\lambda|x|}dx, \quad x \in \mathbb{R}^d.$$

Afin de contourner cette difficulté, la stratégie utilisée par Bobkov et Ledoux dans [18] est de travailler directement sur cette loi en exploitant ses propriétés spécifiques. En particulier, sa queue de distribution tendant vers 0 plus lentement que la queue gaussienne, l'inégalité de Sobolev logarithmique (0.1.3) est mise en défaut par la distribution exponentielle, ce qui a amené les auteurs à étudier des versions affaiblies de cette inégalité, appelées *inégalités de Sobolev logarithmiques modifiées*. En utilisant alors la méthode de Herbst, ils retrouvent la version fonctionnelle des résultats de concentration de Talagrand établis dans [80] pour des mesures exponentielles produits.

La version discrète du processus de diffusion $(Y_t)_{t\geq 0}$ ci-dessus est la file d'attente $M/M/1$. En effet, le processus solution de l'équation différentielle stochastique (0.1.14) peut être construit comme la limite fluide d'une suite de files d'attente $M/M/1$ renormalisées, cf. [72]. La file d'attente $M/M/1$ peut d'ailleurs être modélisée de la façon suivante. Considérons un unique guichet de vente devant lequel des clients font la queue. Le flot des arrivées au guichet est simulé par un processus de Poisson d'intensité $\lambda > 0$ et dès qu'un client est servi, il sort de la file et laisse la place au client suivant. On suppose que les durées de service des clients sont indépendantes et identiquement distribuées, de loi exponentielle de paramètre $\nu > 0$. Alors, si l'on désigne par X_t le nombre de clients dans la file d'attente à l'instant t, le processus aléatoire $(X_t)_{t\geq 0}$ est une file d'attente $M/M/1$. C'est un processus de naissance et de mort sur $E = \mathbb{N}$ dont les fonctions de transition du générateur infinitésimal sont données par $\lambda_x = \lambda$, $\nu_x = \nu 1_{\{x\neq 0\}}$, $x \in \mathbb{N}$. D'après la relation de commutation entre gradient et générateur infinitésimal

$$\mathrm{d}^+\mathcal{L}f = \mathcal{L}\mathrm{d}^+ f,$$

on en déduit que la courbure exacte de ce processus est nulle. Ainsi, en comparant avec le cas continu de la loi exponentielle, l'approche de la concentration locale de la file d'attente $M/M/1$ par les courbures discrètes ne paraît pas être appropriée afin d'obtenir des informations sur sa distribution stationnaire en régime ergodique, c'est-à-dire la loi géométrique sur $\mathbb{N}$ de paramètre $\lambda/\nu < 1$. Il est donc intéressant d'adapter dans le cadre des processus de naissance et de mort à courbure nulle la méthode utilisée par Bobkov et Ledoux pour établir des résultats de concentration pour les distributions stationnaires associées.

0.2 Fluctuations des intégrales stables stochastiques

L'objectif de la seconde partie de la thèse est d'étudier les fluctuations des intégrales stochastiques dirigées par des processus stables, à travers l'analyse des trajectoires du processus supremum ainsi qu'en termes d'estimations de temps de passage de processus voisins.

La théorie des fluctuations des processus stochastiques consiste à analyser les trajectoires de ces processus en étudiant le comportement de leurs extrema. En particulier, le cas des processus de Lévy avec sauts pose de nombreuses difficultés et est d'un intérêt majeur dans les applications, notamment en finance mathématique. Depuis les années soixante, un grand nombre de chercheurs se sont penchés sur l'étude fine des fluctuations des processus de Lévy et en particulier des processus stables, citons entre autres Khintchine, Blumenthal, Getoor, Fristedt, Rogozin, Pruitt, Taylor, Doney et Bertoin. La difficulté principale dans l'analyse des processus stables réside dans l'absence d'une expression explicite des densités, rendant l'étude de leurs trajectoires à la fois délicate et passionnante. Donnons tout d'abord quelques résultats classiques de régularité sur les processus de Lévy avec sauts. Dans l'article fondateur [14] du début des années soixante, Blumenthal et Getoor ont introduit des indices liés aux queues de distribution des mesures de Lévy afin de déterminer en fonction de ces indices le comportement en temps petit des trajectoires des processus de Lévy associés. Plus précisément, désignons par $(X_t)_{t\geq 0}$ un processus de Lévy sur $\mathbb{R}^d$ sans partie gaussienne et de caractéristiques $(0, b, \nu)$, où ν est sa mesure de Lévy et b son drift, et introduisons l'indice

$$\beta := \inf \left\{ \delta > 0 : \int_{|x|\leq 1} |x|^\delta \nu(dx) < +\infty \right\}.$$

Alors, Blumenthal et Getoor ont prouvé le

Théorème 0.2.1. *On a la caractérisation*

$$\limsup_{t\to 0} t^{-1/\eta}|X_t| = 0 \quad \textit{ou} \quad +\infty \quad \textit{p.s.,} \quad \textit{selon que} \quad \eta > \beta \quad \textit{ou} \quad \eta < \beta.$$

Par exemple, dans le cas $\eta < \beta$, ce résultat signifie que le processus pénètre infiniment souvent la région $\{(t, x) \in \mathbb{R}_+ \times \mathbb{R} : |x| > yt^{1/\eta}\}$ en temps arbitrairement petit, et ce pour tout $y > 0$. Ainsi, le théorème 0.2.1 peut être interprété comme un résultat de régularité du processus de Lévy $(X_t)_{t\geq 0}$ autour de 0.

Ces travaux ont été généralisés au tout début des années quatre-vingt par Pruitt dans [69] en utilisant une approche duale de celle de Blumenthal et Getoor. En effet, considérons la fonction h définie sur $\mathbb{R}_+$ par

$$h(r) := \int_{|x|>r} \nu(dx) + r^{-2} \int_{|x|\leq r} |x|^2 \nu(dx) + r^{-1} \left| b + \int_{1<|x|\leq r} x\nu(dx) - \int_{r<|x|\leq 1} x\nu(dx) \right|, \tag{0.2.1}$$

où $r > 0$ est un niveau de troncature du support de la mesure de Lévy, et définissons les indices suivants :

$$\beta_1 := \inf \left\{ \delta > 0 : \limsup_{r\to 0} r^\delta h(r) = 0 \right\}, \quad \beta_2 := \inf \left\{ \delta > 0 : \liminf_{r\to 0} r^\delta h(r) = 0 \right\}.$$

Dans la plupart des cas, les indices β et β_1 coïncident. En notant $(X_t^*)_{t\geq 0}$ le processus supremum $X_t^* = \sup_{0\leq s\leq t} |X_s|$, Pruitt a établi le

Théorème 0.2.2. *On a les comportements suivants :*

$$\limsup_{t\to 0} t^{-1/\eta} X_t^* = 0 \quad \textit{ou} \quad +\infty \quad \textit{p.s.,} \quad \textit{selon que} \quad \eta > \beta_1 \quad \textit{ou} \quad \eta < \beta_1.$$

$$\liminf_{t\to 0} t^{-1/\eta} X_t^* = 0 \quad \textit{ou} \quad +\infty \quad \textit{p.s.,} \quad \textit{selon que} \quad \eta > \beta_2 \quad \textit{ou} \quad \eta < \beta_2.$$

Remarquons que le cas où η est égal à l'un des indices définis plus haut n'est pas traité. Cependant, si $(X_t)_{t\geq 0}$ est un processus stable, alors on a les égalités $\alpha = \beta = \beta_1 = \beta_2$, et le cas critique $\eta = \alpha$ est couvert par l'assertion (iii) du théorème VIII.5 de [10], résultat dû à Khintchine :

Théorème 0.2.3. *Si $f : (0, +\infty) \to (0, +\infty)$ est une fonction continue et strictement croissante, alors on a le test intégral*

$$\limsup_{t\to 0} \frac{X_t^*}{f(t)} = 0 \quad \textit{ou} \quad +\infty \quad \textit{p.s.,} \quad \textit{selon que} \quad \int_{0+} \frac{dt}{f(t)^\alpha} < +\infty \quad \textit{ou} \quad = +\infty.$$

Inégalités maximales et temps de sortie

Le point déterminant dans les théorèmes précédents est la donnée d'estimées asymptotiques ou d'inégalités (maximales) satisfaites par le processus supremum $(X_t^*)_{t\geq 0}$. En effet, l'estimation de la probabilité de déviation du processus supremum

$$\mathbb{P}(X_t^* \geq x), \quad t \geq 0, \quad x > 0,$$

est essentielle afin d'établir via le lemme de Borel-Cantelli des résultats de convergence presque sûre. Par exemple, Pruitt obtient dans [69] les inégalités maximales suivantes pour un processus de Lévy $(X_t)_{t\geq 0}$ sur $\mathbb{R}^d$:

Théorème 0.2.4. *Il existe une constante $C_d > 0$ dépendant seulement de la dimension d telle que pour tout $t > 0$ et tout $x > 0$,*

$$\mathbb{P}(X_t^* \geq x) \leq C_d t h(x), \quad \mathbb{P}(X_t^* \leq x) \leq \frac{C_d}{th(x)},$$

où la fonction h est définie en (0.2.1).

En particulier, ces inégalités maximales fournissent des informations sur les temps de sortie de boules des processus. En effet, la distribution du processus supremum étant étroitement liée à celle de la variable aléatoire

$$T_x := \inf\{t \geq 0 : |X_t| > x\},$$

désignant le premier instant de sortie du processus $(X_t)_{t\geq 0}$ de la boule centrée en 0 de rayon $x > 0$, le théorème 0.2.4 entraîne immédiatement le

Théorème 0.2.5. *Il existe une constante $C_d > 0$ dépendant seulement de la dimension d telle que pour tout $x > 0$,*

$$\frac{1}{C_d h(x)} \leq \mathbb{E}T_x \leq \frac{C_d}{h(x)}.$$

Ainsi, le temps moyen mis par un processus de Lévy pour sortir de boules centrées en 0 est déterminé par le comportement de la mesure de Lévy associée, à travers la fonction h. Dans le cas des processus spectralement négatifs, c'est-à-dire n'ayant pas de sauts positifs, en notant

$$\tau_x := \inf\{t \geq 0 : X_t > x\}$$

le premier instant de sortie du processus $(X_t)_{t\geq 0}$ de l'intervalle $(-\infty, x]$, des techniques de martingales combinées à la nullité de l'overshoot $X_{\tau_x} - x = 0$ permettent de trouver la transformée de Laplace de la variable aléatoire τ_x, cf. le théorème VII.1 de [10]. Cependant, il est en général difficile de donner des formules explicites lorsque le processus possède à la fois des sauts positifs et négatifs comme par exemple un processus stable symétrique. Dans ce cas, il est pertinent d'établir des bornes pour estimer la distribution de ces temps de sortie.

Par ailleurs, il est bien connu que les inégalités maximales interviennent dans l'étude de la convergence des processus. Dans le cas des intégrales stochastiques dirigées par un processus réel symétrique α-stable $(Z_t)_{t\in[0,1]}$ avec $\alpha \in (0,2)$, Giné et Marcus ont montré dans l'article [39] que la condition $(H_t)_{t\geq 0} \in L^\alpha(\mathbb{P} \times dt)$ est suffisante pour que l'intégrale $(\int_0^t H_s dZ_s)_{t\in[0,1]}$ soit bien définie. La clé du raisonnement repose sur une inégalité maximale du type

$$\sup_{x>0} x^\alpha \, \mathbb{P}\left(\sup_{t\in[0,1]} \left|\int_0^t H_s dZ_s\right| \geq x\right) \leq \frac{C_\alpha}{\alpha(2-\alpha)^2} \int_0^1 \mathbb{E}|H_t|^\alpha dt, \tag{0.2.2}$$

où C_α est une constante dépendant seulement de α et des poids de la mesure de Lévy stable. Ceci leur permet d'établir dans l'espace de Skorohod des fonctions *càdlàg* un théorème de la limite centrale pour ces intégrales stables stochastiques. Cependant, afin d'établir l'inégalité (0.2.2), les auteurs utilisent à plusieurs reprises une formule d'isométrie faisant apparaître un facteur $(2-\alpha)^2$ au dénominateur de la borne supérieure. Ainsi, lorsque α est proche de 2, on observe une perte d'information par rapport au cas asymptotique récemment étudié par Hult et Lindskog dans [47], où ces derniers ont obtenu le résultat suivant : il existe une constante D_α bornée en α telle que

$$\lim_{x\to+\infty} x^\alpha \, \mathbb{P}\left(\sup_{t\in[0,1]} \left|\int_0^t H_s dZ_s\right| \geq x\right) = \frac{D_\alpha}{\alpha}\int_0^1 \mathbb{E}|H_t|^\alpha dt. \tag{0.2.3}$$

D'autre part, les estimées précédentes concernant des grands niveaux de déviation x, il est intéressant de se demander si le comportement du processus supremum est identique lorsque le niveau de déviation x est petit. Autrement dit, la queue de distribution de la mesure de Lévy gouverne-t-elle encore le comportement du processus supremum dans un voisinage de 0 ? Si la réponse à cette question est négative, peut-on alors espérer retrouver les inégalités maximales gaussiennes classiques lorsque le paramètre de stabilité α tend vers 2 ?

Ainsi, il est naturel d'étudier le comportement du processus supremum à travers des inégalités maximales du type (0.2.2) lorsque le paramètre de stabilité α est proche de 2, c'est-à-dire lorsque le processus stable sous-jacent est proche d'un mouvement brownien standard, et de regarder quelles sont les conséquences de ces inégalités dans l'estimation de temps de passage de ces processus.

Principe de domination convexe

Commençons par introduire le principe de domination convexe entre variables aléatoires.

Définition 0.2.6. *Soient X et Y deux variables aléatoires réelles. On dit que X est dominée au sens convexe par Y si pour toute fonction convexe ϕ suffisamment régulière,*

$$\mathbb{E}[\phi(X)] \leq \mathbb{E}[\phi(Y)].$$

Le principe de domination convexe peut être interprété comme une généralisation des inégalités de moments développées dans la théorie générale des processus, et fournit des informations intéressantes dans la comparaison de variables aléatoires. Les applications de la domination convexe sont nombreuses et permettent de retrouver un grand nombre de résultats bien connus comme par exemple les inégalités de Doob et de Burkholder, cf. [32], mais aussi des inégalités de déviation et de concentration. Récemment, Klein a montré dans sa thèse [58] que sous quelques conditions de bornitude, une intégrale stochastique dirigée par un processus ponctuel compensé est dominée au sens convexe par une variable aléatoire de Poisson centrée, dont le paramètre dépend des caractéristiques de l'intégrale stochastique précédente. Ce type de résultats a été étendu dans l'article [59], où un principe de domination convexe est établi dans le cadre général des martingales à sauts bornés et des variables aléatoires admettant une représentation en termes d'intégrales stochastiques dirigées par un mouvement brownien corrélé avec une mesure aléatoire de Poisson compensée. Cependant, ce travail ne concerne pas le cas des processus à variation infinie avec sauts non bornés tel qu'un processus stable symétrique, et il serait intéressant de regarder si ce type de processus satisfait un principe de domination convexe.

Construction des intégrales stables stochastiques

Sur un espace de probabilité filtré $\mathbf{\Omega} := (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$, on considère $(Z_t)_{t\geq 0}$ un processus stable réel *càdlàg* d'index $\alpha \in (0,2)$ sans partie gaussienne. Sa fonction caractéristique est donnée par

$$\varphi_{Z_t}(u) = \exp t \left(iub + \int_{-\infty}^{+\infty} \left(e^{iuy} - 1 - iuy\, 1_{\{|y|\leq 1\}} \right) \nu(dy) \right), \quad u > 0, \tag{0.2.4}$$

où ν désigne la mesure de Lévy stable sur $\mathbb{R} \setminus \{0\}$:

$$\nu(dy) = \left(c_-\, 1_{\{y<0\}} + c_+\, 1_{\{y>0\}} \right) \frac{dy}{|y|^{\alpha+1}}, \quad c_-, c_+ \geq 0, \quad c_- + c_+ > 0. \tag{0.2.5}$$

Étant un processus de Lévy, le processus stable $(Z_t)_{t\geq 0}$ est une semimartingale dont la décomposition de Lévy-Itô est donnée par

$$Z_t = bt + \int_0^t \int_{|y|\leq 1} y\, (\mu - \sigma)(dy, ds) + \int_0^t \int_{|y|>1} y\, \mu(dy, ds), \quad t \geq 0, \tag{0.2.6}$$

où μ est une mesure de Poisson aléatoire sur $\mathbb{R} \times [0, +\infty)$ d'intensité $\sigma(dy, dt) = \nu(dy) \otimes dt$ et b est le drift. Dans le cas où $c := c_+ = c_-$, le processus est symétrique et sa fonction caractéristique (0.2.4) se réécrit comme

$$\varphi_{Z_t}(u) = e^{-t\rho_\alpha |u|^\alpha},$$

où la constante ρ_α est donnée par la formule

$$\rho_\alpha := \frac{\sqrt{\pi}\,\Gamma((2-\alpha)/2)}{\alpha 2^\alpha \Gamma((1+\alpha)/2)}\, 2c.$$

Introduisons maintenant la méthode de troncature du support de la mesure de Lévy stable (0.2.5). Afin de contrôler la taille des sauts de la partie martingale, on introduit un niveau de troncature $R > 1$ de la manière suivante : on définit les processus de Lévy indépendants $(Z_t^{(R+)})_{t\geq 0}$ et $(Z_t^{(R-)})_{t\geq 0}$ par

$$Z_t^{(R-)} := \int_0^t \int_{|y|\leq R} y\, (\mu - \sigma)(dy, ds), \quad Z_t^{(R+)} := \int_0^t \int_{|y|>R} y\, \mu(dy, ds), \quad t \geq 0,$$

le premier processus étant une martingale de carré intégrable et à sauts bornés par R, tandis que le second est un processus de Poisson composé de sauts plus grand que R. Ainsi, la décomposition de Lévy-Itô (0.2.6) se réécrit comme

$$Z_t = b_R t + Z_t^{(R-)} + Z_t^{(R+)}, \quad t \geq 0, \tag{0.2.7}$$

où $b_R := b + \int_{1<|y|\leq R} y\, \nu(dy)$ est le drift modifié. Si $(H_t)_{t\geq 0}$ est un processus prévisible, on note dans la suite

$$\|H\|_{(p,t)} := \|H\|_{L^p(\mathbf{\Omega}\times[0,t])} = \left(\int_0^t \mathbb{E}\left[|H_s|^p\right] ds\right)^{\frac{1}{p}}, \quad t \geq 0, \quad p > 0,$$

et on définit $\mathscr{P}_p$ (resp. $\mathscr{B}_p$) comme l'espace des processus prévisibles $(H_t)_{t\geq 0}$ tels que pour tout $t \geq 0$, $\|H\|_{(p,t)} < +\infty$ (resp. $\|H\|_{L^\infty(\mathbf{\Omega}, L^p([0,t]))} < +\infty$). En particulier, le processus $(H_t)_{t\geq 0}$ est dit de carré intégrable si $H \in \mathscr{P}_2$. En notant

$$X_t^{(R-)} := \int_0^t H_s dZ_s^{(R-)}, \quad X_t^{(R+)} := \int_0^t H_s dZ_s^{(R+)}, \quad A_t^R := b_R \int_0^t H_s ds, \quad t \geq 0,$$

où la première intégrale est une martingale de carré intégrable et les deux autres sont construites dans le sens de Lebesgue-Stieltjes, on définit l'intégrale stable $(X_t)_{t\geq 0}$, qui ne dépend pas du choix de R, comme la somme de ces trois intégrales :

$$X_t := \int_0^t H_s dZ_s = A_t^R + X_t^{(R-)} + X_t^{(R+)}, \quad t \geq 0. \tag{0.2.8}$$

Remarquons que dans le cas où $(Z_t)_{t\geq 0}$ est symétrique et à variation infinie, c'est-à-dire $\alpha \in (1,2)$, la partie drift $(A_t^R)_{t\geq 0}$ est nulle et le processus $(X_t)_{t\geq 0}$ est une martingale par la Proposition 2.1 de [8].

0.3 Panorama des résultats obtenus

À présent, nous allons décrire les résultats obtenus dans cette thèse. Chacun des chapitres 1 à 6 correspond à une publication, une prépublication soumise ou à un travail en cours de finalisation. Les références apparaissant entre parenthèses renvoient aux résultats principaux des chapitres concernés.
Les annexes A et B n'étant pas voués à être soumis pour publication, nous n'en parlerons pas dans cette introduction.

Chapitre 1 : inégalités fonctionnelles autour de la loi géométrique

Le premier chapitre de la thèse est constitué de l'article [55] publié avec Nicolas Privault et traite d'inégalités fonctionnelles et de concentration autour de la distribution géométrique. En s'inspirant du travail de Bobkov et Ledoux [18] sur la loi exponentielle, nous établissons en particulier les constantes optimales dans les inégalités isopérimétrique et de Poincaré, ainsi qu'une inégalité de Sobolev logarithmique modifiée et des estimées de déviation dans le cas multidimensionnel.

Introduisons les inégalités fonctionnelles que nous allons étudier dans le premier chapitre.

Définition 0.3.1. *Soit $(E, \mathscr{F}, \mu)$ un espace de probabilité muni d'un opérateur de gradient ∇. La mesure μ satisfait une inégalité isopérimétrique de constante $h > 0$ si pour toute fonction $f : E \to \mathbb{R}$ suffisamment régulière,*

$$h \int |f - m(f)| d\mu \leq \int |\nabla f| d\mu,$$

où $m(f)$ est une médiane de f sous μ. En particulier, on note h_μ la constante optimale

$$h_\mu := \inf_{f \neq \text{constante}} \frac{\int |\nabla f| d\mu}{\int |f - m(f)| d\mu},$$

et λ_μ la constante optimale dans l'inégalité de Poincaré (0.1.2) d'énergie $\mathscr{E}_\mu(f, f) := \int |\nabla f|^2 d\mu$:

$$\lambda_\mu := \inf_{f \neq \text{constante}} \frac{\int |\nabla f|^2 d\mu}{\int \left(f - \int f d\mu\right)^2 d\mu}.$$

Dans le cas d'une mesure μ sur $\mathbb{N}$ et du gradient forward d^+, un résultat de co-aire nous permet d'établir l'équivalence entre les inégalités isopérimétrique et de Poincaré :

Proposition 0.3.2. ***(Proposition 1.2.2)*** *On a les inégalités suivantes :*

$$\left(\sqrt{1 + h_\mu} - 1\right)^2 \leq \lambda_\mu \leq 2h_\mu.$$

On désigne par π la distribution géométrique de paramètre $p \in (0, 1)$, i.e.

$$\pi(k) := (1 - p)p^k, \quad k \in \mathbb{N}.$$

En utilisant une formule d'intégration par parties satisfaite par la loi géométrique ainsi que la proposition 0.3.2, on obtient les expressions explicites des constantes optimales dans les inégalités isopérimétrique et de Poincaré, avec l'énergie induisant le gradient forward d^+ :

Proposition 0.3.3. ***(Propositions 1.3.2 and 1.3.3)*** *Sous la distribution géométrique π, on établit les constantes optimales*

$$h_\pi = \frac{1 - p}{p}, \quad \lambda_\pi = \frac{(1 - \sqrt{p})^2}{p}.$$

À notre connaissance, l'expression de h_π apparaît ici pour la première fois, tandis que celle de λ_π est établie par Van Doorn au début des années quatre-vingt et est redémontrée par Chen dans [28] en utilisant des formules variationnelles sur le trou spectral du générateur infinitésimal. Ainsi, nous donnons une autre preuve très simple de ce résultat. De plus, par un passage à la limite, nous retrouvons les constantes optimales pour la loi exponentielle munie du gradient usuel sur $\mathbb{R}$, cf. [18].

L'intérêt de se limiter à établir certaines inégalités fonctionnelles en dimension 1 repose dans leur propriété de tensorisation, le résultat en dimension supérieure se déduisant automatiquement du cas unidimensionnel. Désignons par $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$ et soit $f : \mathbb{N}^n \to \mathbb{R}$. On définit le gradient multidimensionnel $\mathrm{d}_i^+ f(x) := f(x+e_i) - f(x)$, $i = 1, \ldots, n$, où $x \in \mathbb{N}^n$, ainsi que la norme

$$\|\mathrm{d}^+ f\| := \left(\sum_{i=1}^n |\mathrm{d}_i^+ f|^2 \right)^{1/2}.$$

Alors, sous la distribution géométrique multidimensionnelle $\pi^{\otimes n}$, nous démontrons une inégalité de Sobolev logarithmique modifiée au sens de Bobkov et Ledoux [18] :

Théorème 0.3.4. ***(Theorem 1.3.6, cas multidimensionnel)***
Soit $c \in (0, -\log p)$ et soit $f : \mathbb{N}^n \to \mathbb{R}$ une fonction vérifiant la condition $\max_{i=1,\ldots,n} |\mathrm{d}_i^+ f| \leq c$. Alors la distribution géométrique multidimensionnelle $\pi^{\otimes n}$ satisfait l'inégalité de Sobolev logarithmique modifiée

$$\int f e^f d\pi^{\otimes n} - \int e^f d\pi^{\otimes n} \log \left(\int e^f d\pi^{\otimes n} \right) \leq \frac{pe^c}{(1-p)(1-\sqrt{pe^c})} \int \|\mathrm{d}^+ f\|^2 e^f d\pi^{\otimes n}. \quad (0.3.1)$$

En utilisant la méthode de Herbst, l'inégalité de Sobolev logarithmique modifiée (0.3.1) entraîne l'inégalité de déviation adimensionnelle suivante :

Corollaire 0.3.5. ***(Corollary 1.3.7)*** *Soit $c \in (0, -\log p)$ et soit $f : \mathbb{N}^n \to \mathbb{R}$ une fonction vérifiant les conditions $\max_{i=1,\ldots,n} |\mathrm{d}_i^+ f| \leq \beta$ et $\|\mathrm{d}^+ f\|^2 \leq \alpha^2$, où α et β sont des constantes positives. Alors la distribution géométrique multidimensionnelle $\pi^{\otimes n}$ satisfait l'inégalité de déviation*

$$\pi^{\otimes n}(f - \mathbb{E}_{\pi^{\otimes n}}[f] \geq r) \leq \exp\left(-\min\left(\frac{c^2 r^2}{4a_{p,c}\alpha^2\beta^2}, \frac{rc}{2\beta} \right) \right), \quad (0.3.2)$$

où $a_{p,c}$ est la constante donnée dans l'inégalité de Sobolev logarithmique modifiée (0.3.1).

Ainsi, on remarque que la concentration de la distribution géométrique est semblable à la version fonctionnelle de la concentration de la loi exponentielle établie par Talagrand dans [80]. Notons aussi que ce type de résultat est obtenu par Houdré et Tetali dans l'article [45] en utilisant l'approche par les chaînes de Markov réversibles et divers gradients discrets.

Dans le cas général d'une mesure produit $\mu^{\otimes n}$ sur $\mathbb{N}^n$ et du gradient $(\mathrm{d}_i^+)_{i=1,\ldots,n}$, nous montrons aussi qu'une condition suffisante pour qu'elle admette une concentration de type (0.3.2) est que la marginale μ vérifie une inégalité de Poincaré (0.1.2) avec le gradient forward d^+.

Chapitre 2 : inégalités fonctionnelles dans le cas avec interactions

Une mesure produit correspondant à la distribution stationnaire d'un vecteur de processus de Markov indépendants, il est naturel d'introduire de la dépendance en considérant un potentiel d'interaction associé à une mesure de Gibbs. Ainsi, nous étendons brièvement les résultats du chapitre 1 aux systèmes de spins discrets dont la mesure de référence est géométrique, ce qui fait l'objet de la courte note à paraître [54].

Tout d'abord, introduisons quelques notations. L'espace de configurations que l'on considère est $\mathbb{N}^{\mathbb{Z}^d}$. Étant donné un potentiel d'interaction borné et de portée finie $\Phi = \{\Phi_R : R \subset \mathbb{Z}^d\}$, ainsi qu'un ensemble fini Λ du réseau $\mathbb{Z}^d$, on définit l'Hamiltonien du système par

$$H_\Lambda(\eta) := \sum_{R \bigcap \Lambda \neq \emptyset} \Phi_R(\eta_R),$$

où η_R désigne la restriction de la configuration η à $\mathbb{N}^R$, $R \subset \mathbb{Z}^d$. La mesure de Gibbs en volume fini π_Λ^ω sur $\mathbb{N}^\Lambda$ associée à un système de spins discrets sur $\mathbb{N}$ avec condition au bord $\omega \in \mathbb{N}^{\mathbb{Z}^d \setminus \Lambda}$ est définie par sa densité par rapport à la mesure de référence géométrique multidimensionnelle $\pi_\Lambda := \pi^{\otimes \Lambda}$:

$$\frac{d\pi_\Lambda^\omega}{d\pi_\Lambda}(\sigma) := \frac{1}{Z_\Lambda^\omega} e^{-H_\Lambda^\omega(\sigma)}, \quad \sigma \in \mathbb{N}^\Lambda,$$

où Z_Λ^ω est une constante de normalisation et $H_\Lambda^\omega(\eta) := H_\Lambda(\eta_\Lambda \omega_{\Lambda^c})$, avec la notation $\eta_A \omega_B$ désignant la concaténation de deux configurations η et ω définies respectivement sur $\mathbb{N}^A$ et $\mathbb{N}^B$ avec A et B disjoints. On note Π la mesure de Gibbs sur l'espace $\mathbb{N}^{\mathbb{Z}^d}$ tout entier, et on définit les gradients

$$\mathrm{d}_k^+ f(\eta) = f(\eta + e_k) - f(\eta), \quad \text{et} \quad \mathrm{d}_k^- f(\eta) = 1_{\{\eta_k > 0\}} \left(f(\eta - e_k) - f(\eta) \right),$$

pour toute fonction $f : \mathbb{N}^{\mathbb{Z}^d} \to \mathbb{R}$, où $(e_k)_{k \in \mathbb{Z}^d}$ désigne la base canonique $\{e_k = 1_{\{k\}} : k \in \mathbb{Z}^d\}$. La dynamique markovienne que l'on considère est

$$\mathcal{L}_\Lambda^\omega f(\eta) = \sum_{k \in \Lambda} c_\Lambda^\omega(k, \eta, +) \mathrm{d}_k^+ f(\eta) + c_\Lambda^\omega(k, \eta, -) \mathrm{d}_k^- f(\eta),$$

où les fonctions de transition $c_\Lambda^\omega(k, \eta, \pm)$ sont supposées assurer la symétrie du générateur infinitésimal $\mathcal{L}_\Lambda^\omega$ dans $L^2(\pi_\Lambda^\omega)$, et sont uniformément bornées : il existe une constante $C > 0$ dépendant seulement de $\|\Phi\|$ telle que

$$\frac{1}{C} \leq c_\Lambda^\omega(k, \eta, +) \leq C, \quad \eta \in \mathbb{N}^\Lambda, \quad k \in \Lambda. \tag{0.3.3}$$

Enfin, on désigne par $\mathscr{R}_L$ l'ensemble des rectangles dans $\mathbb{Z}^d$ de taille plus petite que L. Introduisons maintenant une condition de mélange sur la mesure de Gibbs en volume fini qui permet de contrôler la décroissance des covariances.

Définition 0.3.6. *La mesure de Gibbs π_Λ^ω satisfait la condition de mélange s'il existe deux constantes positives C_1 et C_2, dépendant seulement de la dimension d et de $\|\Phi\|$, telles que*

$$\sup_{\sigma,\omega}\left|\frac{\pi_\Lambda^\omega(\{\eta:\ \eta_A=\sigma_A\})\pi_\Lambda^\omega(\{\eta:\ \eta_B=\sigma_B\})}{\pi_\Lambda^\omega(\{\eta\ :\ \eta_{A\cup B}=\sigma_{A\cup B}\})}-1\right|\le C_1e^{-C_2d(A,B)}, \tag{0.3.4}$$

pour tout $\Lambda\in\mathscr{R}_L$, $L\ge 1$, et tous ensembles disjoints $A,B\subset\Lambda$.

Sous les conditions de bornitude (0.3.3) et de mélange (0.3.4), on peut montrer en adaptant la méthode utilisée dans [31] que la mesure de Gibbs π_Λ^ω vérifie une inégalité de Sobolev logarithmique modifiée dont la constante est indépendante de Λ et de la condition au bord ω. Ce résultat entraîne par la méthode de Herbst le

Corollaire 0.3.7. ***(Corollary 2.2.3)*** *Sous la condition de mélange (0.3.4), si $c\in(0,-\log p)$, alors pour toute fonction f vérifiant les conditions $\|\mathrm{d}^+f\|_{l^\infty(\Lambda)}\le\beta$ et*

$$\sum_{k\in\Lambda}c_\Lambda^\omega(k,\eta,+)|\mathrm{d}_k^+f(\eta)|^2\le\alpha^2,\quad \pi_\Lambda(d\eta)-p.p.,$$

où $\alpha,\beta>0$, on obtient l'inégalité de déviation

$$\pi_\Lambda^\omega\left(f-\mathbb{E}_{\pi_\Lambda^\omega}[f]\ge r\right)\le\exp\left(-\min\left(\frac{c^2r^2}{4\gamma_c\alpha^2\beta^2},\frac{rc}{2\beta}\right)\right),\quad r>0,$$

où γ_c est la constante de l'inégalité de Sobolev logarithmique modifiée satisfaite par π_Λ^ω.

Chapitre 3 : concentration locale de type Poisson et courbures discrètes

Le troisième chapitre de cette thèse est constitué d'une version détaillée de l'article soumis [52]. Nous analysons le comportement de la concentration locale des processus de naissance et de mort en fonction des différentes notions de courbures discrètes, et nous donnons des critères sur le générateur infinitésimal afin que cette concentration soit similaire à celle de type Poisson satisfaite par la file d'attente $M/M/\infty$. Dans ce contexte, ce travail généralise les résultats de Ané et de Ledoux obtenus dans l'article [3] pour des marches aléatoires en courbure nulle.

Contrairement aux processus de diffusion où l'approche par les inégalités fonctionnelles locales est souvent utilisée, il est en général difficile dans le cadre discret d'établir une inégalité de type Sobolev logarithmique. Par exemple, une inégalité de Sobolev logarithmique modifiée locale avec le gradient forward d^+ n'est pas encore démontrée à ce jour pour la file d'attente $M/M/\infty$, alors que l'on connaît explicitement sa distribution. Afin d'établir des résultats de concentration locale pour des processus de naissance et de mort, l'approche que nous proposons dans le chapitre 3 est plutôt basée sur des techniques de

martingales et utilise systématiquement les courbures discrètes de ces processus. Ainsi, il est naturel de formuler des critères sur les fonctions de transition du générateur infinitésimal associé pour que les courbures discrètes soit minorées. En particulier, nous sommes amenés dans le chapitre 3 tout comme dans le chapitre 4 à étudier le rôle fondamental joué dans la concentration de la mesure par les différentes métriques sur l'espace d'état du processus, induisant des gradients discrets différents.
Mentionnons enfin que nous n'allons considérer dans la suite que le cas des processus de naissance et de mort, bien que le chapitre 3 contienne aussi des résultats pour des chaînes de Markov à temps continu générales.

Critères de courbure

La métrique utilisée dans le chapitre 3 est la distance classique d, et l'on suppose que la distribution stationnaire vérifie la condition de moment (0.1.7) avec cette distance. Donnons tout d'abord quelques critères sur le générateur infinitésimal du processus de naissance et de mort $(X_t)_{t\geq 0}$ pour que ses différentes courbures discrètes soient minorées. Commençons par la d-courbure de Wasserstein. La preuve de la proposition 0.3.8 ci-dessous est fondée sur une méthode de couplage utilisant la distance d.

Proposition 0.3.8. ***(Proposition 3.5.2)*** *S'il existe une constante $K \in \mathbb{R}$ telle que les fonctions de transition λ et ν vérifient l'inégalité*

$$\inf_{x\in E\setminus\{0\}} \lambda_{x-1} - \lambda_x + \nu_x - \nu_{x-1} \geq K, \tag{0.3.5}$$

alors la d-courbure de Wasserstein est minorée par K.

Commentons le critère précédent. Comme nous l'avons déjà mentionné dans la partie 0.1, la file d'attente $M/M/\infty$ dans le cas $E = \mathbb{N}$ (ou la chaîne d'Ehrenfest à temps continu dans le cas fini $E = \{0, \ldots, n\}$) est l'analogue discret d'un processus d'Ornstein-Uhlenbeck. Sachant que pour établir des inégalités locales, on cherche à comparer en termes de convexité de potentiel un processus de Kolmogorov avec le processus d'Ornstein-Uhlenbeck le plus proche, il est naturel de chercher des critères permettant de comparer un processus de naissance et de mort général avec la file d'attente $M/M/\infty$. Remarquons que pour ce processus de queue, nous avons l'égalité pour tout $x \in \mathbb{N} \setminus \{0\}$:

$$\lambda_{x-1} - \lambda_x + \nu_x - \nu_{x-1} = \nu,$$

qui n'est autre que sa courbure exacte. Ainsi, la relation (0.3.5) traduit une comparaison avec la file d'attente $M/M/\infty$, comparaison qui peut être assimilée à l'analogue discret de la condition spectrale (i') intervenant dans le théorème 0.1.2 pour les processus de Kolmogorov. En particulier, plus la force de rappel vers 0 est importante, plus la d-courbure de Wasserstein du processus est grande.

En adaptant au cadre discret le calcul Γ_2 de Bakry, nous établissons l'équivalence entre les versions discrètes des assertions (iv) et (v) du théorème 0.1.2. Nous obtenons alors la

Proposition 0.3.9. ***(Proposition 3.5.6)*** *S'il existe une constante $\rho \geq 0$ telle que*

$$\inf_{x \in E \setminus \{0, \sup E\}} \min\{\lambda_{x-1} - \lambda_x, \nu_{x+1} - \nu_x\} \geq \rho, \tag{0.3.6}$$

alors la Γ-courbure est minorée par ρ.

D'après l'homogénéité en les fonctions de transition λ et ν des opérateurs markoviens Γ et Γ_2, il n'est pas surprenant de faire apparaître ici une symétrie entre λ et ν. Ainsi, ce critère de comparaison paraît plus restrictif que le précédent dans le cas $E = \mathbb{N}$ car le taux de naissance λ doit être décroissant (donc convergent), mais reste très intéressant dans le cas fini.

À présent, nous sommes en mesure de donner des inégalités de déviation locales de type Poisson pour des processus de naissance et de mort. Bien que les preuves soient identiques dans le cas où l'espace d'état E est infini ou non, nous devons différencier les résultats. En effet, lorsque $E = \mathbb{N}$, les hypothèses des théorèmes 0.3.10 et 0.3.11 ci-dessous entraînent que les bornes inférieures sur les courbures discrètes sont au mieux minorées par 0, tandis qu'elles peuvent être minorées par un nombre strictement positif dans le cas fini $E = \{0, \ldots, n\}$. Par soucis de simplification des formules, nous ne donnons dans la suite que le comportement en $x \log(1 + x)$ de la déviation, alors que des estimées plus précises faisant intervenir la fonction $g(x) = (1 + x) \log(1 + x) - x$ sont disponibles dans le chapitre 3.

Concentration dans le cas infini $E := \mathbb{N}$

La preuve du résultat suivant repose sur des techniques de calcul stochastique utilisant la théorie des martingales.

Théorème 0.3.10. ***(Theorem 3.5.8)*** *On suppose que les fonctions de transition λ et ν sont bornées sur $\mathbb{N}$ et qu'elles vérifient l'inégalité (0.3.5) avec $K \leq 0$. Si $f \in \mathrm{Lip}$ est une fonction réelle d-Lipschitzienne sur $\mathbb{N}$, alors pour tout niveau de déviation $y > 0$ et tout $t > 0$, on obtient l'inégalité de déviation locale de type Poisson*

$$\mathbb{P}_x \left(f(X_t) - \mathbb{E}_x \left[f(X_t) \right] \geq y \right) \quad \leq \quad \exp \left(- \frac{y e^{tK}}{2 \|f\|_{\mathrm{Lip}}} \log \left(1 + \frac{yK}{\sinh(tK) \|\lambda + \nu\|_\infty \|f\|_{\mathrm{Lip}}} \right) \right).$$

Fondé sur une minoration par 0 de la Γ-courbure du processus, le prochain résultat ne requiert pas la bornitude des fonctions de transition du générateur infinitésimal, contrairement au théorème 0.3.10, mais une hypothèse additionnelle de type Lipschitz sur les fonctions considérées. En effet, si ν n'est pas bornée, l'hypothèse $\|\Gamma f\|_\infty < +\infty$ entraîne que le gradient de f introduit en (0.1.10) par rapport à la métrique d décroît très rapidement. La preuve du théorème 0.3.11 ci-dessous est une adaptation au cas markovien de la méthode des covariances utilisée par Houdré dans [42] afin d'établir des inégalités de déviation pour des vecteurs indéfiniment divisibles.

Théorème 0.3.11. ***(Theorem 3.5.9)*** *On suppose que les fonctions de transition λ et ν sont respectivement décroissante et croissante sur $E = \mathbb{N}$. Soit $f \in \mathrm{Lip}$ une fonction réelle d-Lipschitzienne sur $\mathbb{N}$ avec de plus $\|\Gamma f\|_\infty < +\infty$. Alors pour tout $y > 0$ et tout $t > 0$, on obtient l'inégalité de déviation locale de type Poisson*

$$\mathbb{P}_x \left(f(X_t) - \mathbb{E}_x \left[f(X_t) \right] \geq y \right) \quad \leq \quad \exp \left(- \frac{y}{2\|f\|_{\mathrm{Lip}}} \log \left(1 + \frac{y\|f\|_{\mathrm{Lip}}}{2t\|\Gamma f\|_\infty} \right) \right).$$

Ainsi, les inégalités de déviation des théorèmes 0.3.10 et 0.3.11 sont intéressantes lorsque t est proche de 0, mais elles ne sont pas satisfaisantes en temps grand car on ne peut pas les étendre à la distribution stationnaire. Ce point sera corrigé dans le chapitre 4 par l'utilisation d'une autre distance que la métrique usuelle d, nous permettant d'obtenir des bornes inférieures strictement positives sur la courbure de Wasserstein associée à cette nouvelle distance.

Concentration dans le cas fini $E := \{0, 1, \ldots, n\}$

Considérons maintenant le cas fini, $E := \{0, 1, \ldots, n\}$. Nous améliorons les théorèmes 0.3.10 et 0.3.11, afin d'établir par ergodicité des estimées sous la distribution stationnaire π. Ainsi, le point déterminant est d'obtenir des bornes inférieures strictement positives sur les courbures discrètes.

Théorème 0.3.12. ***(Theorem 3.5.11)*** *Supposons que les fonctions de transition λ et ν vérifient l'inégalité (0.3.5) de constante $K \in \mathbb{R}$ sur $\{0, 1, \ldots, n\}$. Si $f \in \mathrm{Lip}$ est une fonction réelle d-Lipschitzienne sur $\{0, 1, \ldots, n\}$, alors pour tout niveau de déviation $y > 0$ et tout $t > 0$,*

$$\mathbb{P}_x \left(f(X_t) - \mathbb{E}_x \left[f(X_t) \right] \geq y \right) \quad \leq \quad \exp \left(- \frac{y}{2\|f\|_{\mathrm{Lip}}} \log \left(1 + \frac{2Ky}{(1 - e^{-2Kt})\|\lambda + \nu\|_\infty \|f\|_{\mathrm{Lip}}} \right) \right). \tag{0.3.7}$$

En particulier, si $K > 0$, alors par ergodicité lorsque t tend vers l'infini dans l'inégalité locale (0.3.7), on obtient l'estimée de déviation sous la distribution stationnaire π :

$$\pi \left(f - \mathbb{E}_\pi[f] \geq y \right) \leq \exp \left(- \frac{y}{2\|f\|_{\mathrm{Lip}}} \log \left(1 + \frac{2Ky}{\|\lambda + \nu\|_\infty \|f\|_{\mathrm{Lip}}} \right) \right).$$

Une version similaire faisant intervenir les hypothèses de la proposition 0.3.9 est aussi disponible dans le cas fini.

Introduisons à présent la chaîne d'Ehrenfest à temps continu. Soit n un entier naturel strictement positif et considérons le processus de naissance et de mort $(X_t^n)_{t \geq 0}$ à valeurs dans $E := \{0, 1, \ldots, n\}$, dont le générateur infinitésimal est donné par

$$\mathcal{L}_n f(x) = \lambda(n - x) \left(f(x+1) - f(x) \right) + \nu x \left(f(x-1) - f(x) \right), \quad x \in \{0, 1, \ldots, n\},$$

où les nombres $0 < \lambda \leq \nu < 1$ satisfont $\lambda + \nu = 1$. Ce processus est nommé chaîne d'Ehrenfest à temps continu. Comme nous l'avons déjà mentionné, ce processus est l'analogue discret d'un processus d'Ornstein-Uhlenbeck réel. En effet, supposons que $\lim_{n\to+\infty} X_0^n/n = \lambda$, et définissons pour tout entier n strictement positif le processus renormalisé $U_t^n := (X_t^n - \lambda n)/\sqrt{n}$, $t > 0$. Si la suite d'états initiaux $(U_0^n)_{n\geq 1}$ tend vers un nombre quelconque u_0, alors la suite de processus $(U_t^n)_{t\geq 0}$ converge lorsque n tend vers l'infini vers le processus d'Ornstein-Uhlenbeck réel

$$U_t = u_0 e^{-t} + \sqrt{2\lambda\nu} \int_0^t e^{-(t-s)} dB_s, \quad t > 0,$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien standard.
Ainsi, en appliquant le théorème 0.3.12 à la chaîne d'Ehrenfest $(X_t^n)_{t\geq 0}$ avec $K = \lambda + \nu = 1$ et la fonction d-lipschitzienne $h_n(x) := f((x - n\lambda)/\sqrt{n})$ sur $\{0, 1, \ldots, n\}$, où f est une fonction lipschitzienne au sens classique sur $\mathbb{R}$ de constante $\|f\|_{\text{Lip}}$, puis en passant à la limite lorsque n tend vers l'infini, on obtient pour tout $y > 0$ et tout $t > 0$ l'inégalité de déviation gaussienne locale

$$\mathbb{P}_{u_0}\left(f(U_t) - \mathbb{E}_{u_0}\left[f(U_t)\right] \geq y\right) \leq \exp\left(-\frac{y^2}{(1 - e^{-2t})\nu\|f\|_{\text{Lip}}^2}\right).$$

On remarque que cette dernière inégalité est optimale en temps car à t fixé, le facteur $(1 - e^{-2t})\nu$ au dénominateur est la variance de la variable aléatoire gaussienne U_t (à la constante λ près).

Enfin, précisons que le chapitre 3 se termine par la donnée d'une inégalité de concentration multidimensionnelle locale pour la file d'attente $M/M/1$. La preuve repose sur la tensorisation de la transformée de Laplace locale d'une fonction lipschitzienne, via une formule d'intégration par parties satisfaite par le semigroupe associé. Ce résultat est un cas particulier de la méthodologie que l'on va utiliser dans le chapitre 4.

Chapitre 4 : concentration de la distribution empirique vers l'équilibre

Ce chapitre fait l'objet de l'article soumis [50] et est consacré à l'estimation non-asymptotique de la vitesse de convergence de la distribution empirique vers l'équilibre pour des processus de naissance et de mort ergodiques. Ce résultat étend ceux de Lezaud [63, 64] du cas des fonctions bornées au cas des fonctions lipschitziennes par rapport à une distance sur $\mathbb{N}$ bien choisie, mais aussi aux processus dont le générateur infinitésimal n'est pas nécessairement borné.

Rappelons que si le processus de naissance et de mort $(X_t)_{t\geq 0}$ est ergodique, le théorème ergodique assure que pour toute fonction $g \in L^1(\pi)$, la probabilité

$$\Lambda(t) := \mathbb{P}_x\left(\left|\frac{1}{t}\int_0^t g(X_s)ds - \mathbb{E}_\pi[g]\right| \geq y\right), \quad y > 0, \tag{0.3.8}$$

converge vers 0 lorsque t tend vers l'infini. Il est bien connu que la théorie des grandes déviations donne une vitesse de convergence asymptotique de la quantité $t^{-1}\log \Lambda(t)$, cf. par exemple le chapitre 8 de [29]. Cependant, un tel résultat n'est pas satisfaisant lorsque l'on cherche à exploiter numériquement des bornes non-asymptotiques sur la probabilité (0.3.8). Récemment, Lezaud a établi dans les articles [63, 64] ce type de résultats pour des chaînes de Markov à temps continu à générateur borné sur un espace d'état fini ou dénombrable. En supposant l'existence d'un trou spectral, les résultats obtenus par Lezaud sont des vitesses de convergence de type Poisson pour des fonctions g bornées, dont les preuves s'appuient essentiellement sur la théorie de la perturbation des opérateurs linéaires développée par Kato. Ainsi, il est intéressant de regarder si l'on peut améliorer cette vitesse de convergence via l'approche par les courbures discrètes, tout en relaxant les diverses hypothèses de bornitude imposées par Lezaud.

Changement de métrique

D'après les parties précédentes, il est fondamental obtenir des bornes inférieures strictement positives sur les courbures discrètes afin d'établir des inégalités de concentration locales fournissant de l'information en temps grand. Cependant, comme nous l'avons déjà vu dans la présentation du chapitre 3, la distance usuelle d ne permet pas d'obtenir ce type de résultats dans le cas $E = \mathbb{N}$. C'est pourquoi l'idée est de considérer la courbure de Wasserstein du processus par rapport à une autre distance sur $\mathbb{N}$ bien choisie.

Définition 0.3.13. *Étant donnée une fonction réelle u strictement positive sur $\mathbb{N}$, on définit la distance $\delta : \mathbb{N} \times \mathbb{N} \to [0, +\infty)$ par*

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

avec la convention $\sum_{k=0}^{-1} u(k) := 0$.

Notons que la distance δ est utilisée par Chen dans [28] afin d'établir des formules variationnelles sur le trou spectral du générateur infinitésimal. On remarque que si la fonction test u est identiquement égale à 1, alors $\delta \equiv d$ et la δ-courbure de Wasserstein est réduite à la courbure de Wasserstein par rapport à la distance d.

Avant de donner le résultat principal, faisons un ensemble d'hypothèses sur les fonctions de transition du générateur infinitésimal du processus. On note dans la suite $a \wedge b := \min\{a, b\}$ et $a \vee b := \max\{a, b\}$, $a, b \in \mathbb{R}$.

Hypothèse (A). *Il existe deux constantes $K > 0$ et $C > 0$ telles que*

$$\left(\inf_{x \geq 0} \lambda_x \right) \wedge \left(\inf_{x \geq 1} \nu_x \right) \geq K \quad \textit{et} \quad \delta(x, x+1) := u(x) \leq C \left(\frac{1}{\sqrt{\nu_{x+1}}} \wedge \frac{1}{\sqrt{\lambda_x}} \right), \quad x \in \mathbb{N}.$$

Hypothèse (B). *La distribution stationnaire π vérifie la condition de moment (0.1.7) avec la métrique δ, et il existe une constante strictement positive α telle que*

$$\inf_{x\in\mathbb{N}} \left\{ \nu_{x+1} + \lambda_x - \nu_x \frac{u(x-1)}{u(x)} - \lambda_{x+1} \frac{u(x+1)}{u(x)} \right\} \geq \alpha. \tag{0.3.9}$$

L'intérêt d'introduire l'hypothèse (A) est qu'elle nous permet de comparer les deux distances δ et d : elles ne sont pas équivalentes, et la métrique δ est strictement plus faible que la métrique usuelle d lorsqu'au moins une des fonctions de transition du générateur infinitésimal n'est pas bornée.

Lorsque la fonction u est identiquement égale à 1, on observe que l'hypothèse (B) devient l'inégalité (0.3.5) associée à la métrique usuelle d. Ainsi, il est naturel d'envisager que l'inégalité (0.3.9) avec le poids u est fortement liée à la distance δ. Cette remarque est justifiée par le résultat suivant, dont la preuve est fondée sur une méthode de couplage.

Proposition 0.3.14. ***(Proposition 4.2.6)*** *Sous l'hypothèse (B), la δ-courbure de Wasserstein du processus est minorée par $\alpha > 0$.*

Le résultat principal

Énonçons à présent le résultat principal du chapitre 4.

Théorème 0.3.15. ***(Theorem 4.2.7)*** *Sous les hypothèses (A) et (B), si $\phi \in \mathrm{Lip}_\delta$ est une fonction réelle δ-Lipschitzienne sur $\mathbb{N}$, alors pour tout $t > 0$, tout niveau de déviation $y > 0$ et toute condition initiale $x \in \mathbb{N}$, on a la vitesse de convergence suivante :*

$$\begin{aligned}\mathbb{P}_x \left(\left| \frac{1}{t} \int_0^t \left(\phi(X_s) - \mathbb{E}_x[\phi(X_s)] \right) ds \right| \geq y \right) &\leq 2e^{-2Ktg\left(\frac{y\alpha}{2\sqrt{K}C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}} \right)} \\ &\leq 2e^{-\frac{ty\alpha\sqrt{K}}{2C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}} \log\left(1+\frac{y\alpha}{2\sqrt{K}C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}}\right)},\end{aligned} \tag{0.3.10}$$

où la fonction g est donnée par $g(u) = (1+u)\log(1+u) - u$, $u > 0$.

Commentons l'estimée ci-dessus. Tout d'abord, par invariance de la mesure stationnaire π ainsi que par l'hypothèse (B),

$$\begin{aligned}\left| \frac{1}{t} \int_0^t \mathbb{E}_x[\phi(X_s)] ds - \pi(\phi) \right| &= \left| \frac{1}{t} \int_0^t \sum_{z\in\mathbb{N}} \left(P_s\phi(x) - P_s\phi(z) \right) \pi(z) ds \right| \\ &\leq \|\phi\|_{\mathrm{Lip}_\delta} \sum_{z\in\mathbb{N}} \delta(x,z)\pi(z) \, \frac{1}{t} \int_0^t e^{-\alpha s} ds \\ &= \|\phi\|_{\mathrm{Lip}_\delta} \sum_{z\in\mathbb{N}} \delta(x,z)\pi(z) \, \frac{1-e^{-\alpha t}}{t\alpha} \\ &=: M_t.\end{aligned}$$

Donc pour y suffisamment grand,

$$\mathbb{P}_x\left(\left|\frac{1}{t}\int_0^t \phi(X_s)ds - \pi(\phi)\right| \geq y\right) \leq \mathbb{P}_x\left(\left|\frac{1}{t}\int_0^t (\phi(X_s) - \mathbb{E}_x[\phi(X_s)])\, ds\right| \geq y - M_t\right).$$

Par conséquent, l'inégalité de déviation (0.3.10) donne la vitesse de convergence de la distribution empirique vers l'équilibre, au prix d'une légère restriction (car $\lim_{t\to+\infty} M_t = 0$) sur le niveau de déviation y.

D'autre part, la fonction $u \mapsto g(u)$ dans le théorème 0.3.15 est équivalente à $u^2/2$ lorsque u est proche de 0 et à $u\log(u)$ lorsque u tend vers l'infini. Ainsi, l'inégalité de type Bennett (0.3.10) exhibe une queue gaussienne pour les petites valeurs de y, en accord avec le TCL pour les processus de Markov, et une décroissance de type Poisson pour les grandes valeurs de y. Par conséquent, on améliore dans le cas des processus de naissance et de mort l'inégalité de Chernoff du théorème 1.1 appliqué à l'exemple 1.7 de [64], ou encore celle de la remarque 2.6 dans [64], car aucune hypothèse de bornitude n'est requise sur la fonction ϕ et sur le générateur infinitésimal.
De plus, mentionnons que l'on retrouve aussi le théorème 3.4 dans [63] en adaptant la preuve du théorème 0.3.15 au cas d'un espace d'état fini.
Cependant, le prix à payer pour ces améliorations est de supposer que l'hypothèse (B) est satisfaite, ce qui est plus fort que l'existence d'un trou spectral supposée par Lezaud dans ses articles [63, 64].

L'idée principale de la preuve du théorème 0.3.15 repose essentiellement sur la tensorisation de la transformée de Laplace locale, en utilisant la proposition 0.3.14. Une telle propriété de tensorisation peut être vue comme une généralisation de la méthode des martingales utilisée notamment par Rio dans [71] et par Djellout, Guillin et Wu dans [35], afin d'établir de la concentration gaussienne pour des suites de variables aléatoires faiblement dépendantes.
Donnons un bref aperçu de cette tensorisation. En adaptant pour la distance δ la preuve du théorème 0.3.10 de la partie précédente, nous avons tout d'abord le

Lemme 0.3.16. ***(Lemma 4.3.1)*** *Sous les hypothèses (A) et (B), si $f \in \mathrm{Lip}_\delta$ est une fonction réelle δ-Lipschitzienne sur $\mathbb{N}$, alors pour tout $t > 0$ et tout $\tau > 0$, nous avons l'estimée suivante sur la transformée de Laplace :*

$$\mathbb{E}_x\left[e^{\tau(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] \leq \exp\left\{h\left(\tau, t, \frac{C\|f\|_{\mathrm{Lip}_\delta}}{\sqrt{K}}\right)\right\}, \quad x \in \mathbb{N}, \tag{0.3.11}$$

où h est la fonction définie sur $(\mathbb{R}_+)^3$ par

$$h(\tau, t, z) := \frac{K\left(1 - e^{-2\alpha t}\right)}{\alpha}\left(e^{\tau z} - \tau z - 1\right).$$

Maintenant, voyons comment l'utilisation d'une borne inférieure strictement positive sur la δ-courbure de Wasserstein du processus va nous permettre de tensoriser via l'inégalité (0.3.11) la tranformée de Laplace en dimension 2. Soit $f : \mathbb{N}\times\mathbb{N} \to \mathbb{R}$ une fonction ℓ^1-lipschitzienne sur l'espace produit par rapport à la distance δ, c'est-à-dire

$$\|f\|_{\mathrm{Lip}_\delta(2)} := \sup_{(x_1,x_2),(y_1,y_2)\in\mathbb{N}^2} \frac{|f(x_1,x_2)-f(y_1,y_2)|}{\delta(x_1,y_1)+\delta(x_2,y_2)} < +\infty.$$

Pour tout $0<s<t$, on note les fonctions unidimensionnelles $f_y(z) := f(y,z)$ et $f_1(y) := \sum_{z\in\mathbb{N}} f(y,z)P_{t-s}(y,z)$. Il est clair que l'on a $\|f_y\|_{\mathrm{Lip}_\delta} \le \|f\|_{\mathrm{Lip}_\delta(2)}$. Vérifions que la fonction f_1 est elle-aussi δ-Lipschitzienne, de constante

$$\|f_1\|_{\mathrm{Lip}_\delta} \le \left(1+e^{-\alpha(t-s)}\right)\|f\|_{\mathrm{Lip}_\delta(2)}. \tag{0.3.12}$$

Pour tous $y,w\in\mathbb{N}$, on a :

$$\begin{aligned}
&|f_1(w)-f_1(y)|\\
&\quad= \left|\sum_{z\in\mathbb{N}}\left(f(w,z)P_{t-s}(w,z)-f(y,z)P_{t-s}(y,z)\right)\right|\\
&\quad\le \left|\sum_{z\in\mathbb{N}} f(w,z)\left(P_{t-s}(w,z)-P_{t-s}(y,z)\right)\right| + \left|\sum_{z\in\mathbb{N}}\left(f(w,z)-f(y,z)\right)P_{t-s}(y,z)\right|\\
&\quad\le e^{-\alpha(t-s)}\|f_w\|_{\mathrm{Lip}_\delta}\delta(y,w) + \|f\|_{\mathrm{Lip}_\delta(2)}\delta(y,w)\\
&\quad\le \left(1+e^{-\alpha(t-s)}\right)\|f\|_{\mathrm{Lip}_\delta(2)}\,\delta(y,w),
\end{aligned}$$

où dans la deuxième inégalité est utilisée la proposition 0.3.14, i.e. $\|P_t f\|_{\mathrm{Lip}_\delta} \le e^{-\alpha t}\|f\|_{\mathrm{Lip}_\delta}$. Par conséquent, on a bien que $f_1 \in \mathrm{Lip}_\delta$ et la borne (0.3.12) est vérifiée.
Maintenant, par la propriété de Markov puis par l'inégalité exponentielle (0.3.11) appliquée successivement aux fonctions f_y et f_1, on obtient

$$\begin{aligned}
&\mathbb{E}_x\left[\exp\left(\tau f(X_s,X_t)\right)\right]\\
&\quad= \sum_{y,z\in\mathbb{N}} \exp\left(\tau f_y(z)\right)P_{t-s}(y,z)P_s(x,y)\\
&\quad\le \sum_{y\in\mathbb{N}} \exp\left\{\tau\sum_{z\in\mathbb{N}} f_y(z)P_{t-s}(y,z) + h\left(\tau, t-s, \frac{C\|f_y\|_{\mathrm{Lip}_\delta}}{\sqrt{K}}\right)\right\}P_s(x,y)\\
&\quad\le \exp\left\{h\left(\tau,t-s,\frac{C\|f\|_{\mathrm{Lip}_\delta(2)}}{\sqrt{K}}\right)\right\}\sum_{y\in\mathbb{N}}\exp\left(\tau f_1(y)\right)P_s(x,y)\\
&\quad\le \exp\left\{h\left(\tau,t-s,\frac{C\|f\|_{\mathrm{Lip}_\delta(2)}}{\sqrt{K}}\right) + h\left(\tau,s,\frac{C\|f_1\|_{\mathrm{Lip}_\delta}}{\sqrt{K}}\right) + \tau\mathbb{E}_x[f(X_s,X_t)]\right\},\\
&\quad\le \exp\left\{h\left(\tau,t-s,\frac{C\|f\|_{\mathrm{Lip}_\delta(2)}}{\sqrt{K}}\right) + h\left(\tau,s,\frac{C(1+e^{-\alpha(t-s)})\|f\|_{\mathrm{Lip}_\delta(2)}}{\sqrt{K}}\right)\right.\\
&\qquad\left.+\tau\mathbb{E}_x[f(X_s,X_t)]\right\},
\end{aligned}$$

car la fonction h est croissante en sa dernière variable. Ainsi, il en résulte la borne suivante sur la transformée de Laplace locale en dimension 2 :

$$\mathbb{E}_x\left[e^{\tau(f(X_s,X_t)-\mathbb{E}_x[f(X_s,X_t)])}\right] \\ \leq \exp\left\{h\left(\tau, t-s, \frac{C\|f\|_{\mathrm{Lip}_\delta(2)}}{\sqrt{K}}\right) + h\left(\tau, s, \frac{C(1+e^{-\alpha(t-s)})\|f\|_{\mathrm{Lip}_\delta(2)}}{\sqrt{K}}\right)\right\}.$$

Le cas général pour les fonctions cylindriques $F = f(X_{t_1}, \dots, X_{t_d})$, $0 < t_1 < \cdots < t_d$, où f est une fonction ℓ^1-lipschitzienne sur l'espace produit $\mathbb{N}^d$, est similaire.

Enfin, le chapitre 4 se termine par un exemple classique satisfaisant les hypothèses du théorème 0.3.15, à savoir la file d'attente $M/M/\infty$.

Chapitre 5 : inégalités maximales pour des intégrales stables stochastiques

Le chapitre 5 est consacré à l'étude trajectorielle des intégrales stables stochastiques, et est constitué de l'article à paraître [51]. On y établit pour ce type de processus de nouvelles inégalités maximales pour des grands et des petits niveaux de déviation. Afin de contrôler les petits sauts de la partie martingale d'un processus stable, l'idée est d'utiliser la méthode de troncature de la mesure de Lévy stable introduite précédemment. Le résultat principal de cette partie est la donnée d'une borne supérieure qui n'explose pas lorsque le paramètre de stabilité tend vers 2. Ces résultats sont appliqués à l'estimation de temps de passage du module d'un processus symétrique stable au dessus de fonctions continues. Bien que l'on ait également obtenu des résultats dans le cas où le processus stable est à variation finie, c'est-à-dire lorsque le paramètre de stabilité α est dans $(0,1)$, nous présentons dans ce qui suit seulement le cas $\alpha \in (1,2)$.

Inégalités maximales pour des grands niveaux de déviation

Le premier résultat que nous donnons améliore significativement l'inégalité (0.2.2) établie par Giné et Marcus. D'après l'estimée asymptotique (0.2.3), l'optimalité en termes du niveau de déviation x ainsi que de la norme du processus $(H_t)_{t\geq 0}$ est conservée, et la vitesse d'explosion lorsque α tend vers 2 devient seulement linéaire.

Proposition 0.3.17. ***(Proposition 5.2.4)*** *On suppose que $(Z_t)_{t\geq 0}$ est symétrique d'indice $\alpha \in (1,2)$. Alors il existe une constante $K_\alpha > 0$, bornée en α, telle que*

$$\sup_{x>0} x^\alpha \, \mathbb{P}\left(\sup_{0\leq s\leq t}\left|\int_0^s H_\tau dZ_\tau\right| \geq x\right) \leq \frac{K_\alpha}{2-\alpha}\int_0^t \mathbb{E}\left[|H_s|^\alpha\right] ds, \quad t \geq 0.$$

La preuve de ce résultat repose sur une adaptation dans le cas des intégrales stables stochastiques de la méthode utilisée par Bass et Levin dans [9] afin d'établir des inégalités de Harnack pour des processus de Markov dont l'intensité de sauts est de type stable.

Dans la construction de l'intégrale stable (0.2.8), le choix du niveau de troncature est $R := x$.

Le théorème 0.3.18 ci-dessous est le résultat principal de cette partie. Il nous permet de contrôler le comportement du processus supremum lorsque α tend vers 2. Le prix à payer est de supposer que le processus prévisible $(H_t)_{t\geq 0}$ satisfait des hypothèses d'intégrabilité plus fortes, et de restreindre l'intervalle de validité du niveau de déviation x. La preuve de ce résultat est délicate et repose sur une estimation fine de moments d'intégrales stochastiques avec le choix du niveau de troncature $R := x/\|H\|_{(\alpha+p,t)}$.

Théorème 0.3.18. ***(Theorem 5.3.2)*** *On suppose que $\alpha \in (1,2)$. Soit $(H_t)_{t\geq 0} \in \mathscr{P}_{\alpha+p}$ avec $p > 2-\alpha$. Alors pour tout $t \geq 0$, il existe deux constantes $K > 0$ et $L_{\alpha+p} > 1$, dont la première est indépendante de α, telles que pour tout niveau de déviation x vérifiant*

$$x^\alpha > \|H\|^\alpha_{(\alpha+p,t)}/(2-\alpha)^{L_{\alpha+p}},$$

l'inégalité maximale suivante est satisfaite :

$$\mathbb{P}\left(\sup_{0\leq s\leq t}\left|\int_0^s H_\tau dZ_\tau\right| \geq x\right) \leq \frac{K}{x^\alpha}\left(\int_0^t \mathbb{E}\left[|H_s|^{\alpha+p}\right] ds\right)^{\frac{\alpha}{\alpha+p}}. \qquad (0.3.13)$$

Remarquons que lorsque p s'approche de 0 par valeurs supérieures, α est alors proche de 2 et le membre de droite dans l'inégalité (0.3.13) devient comparable à l'estimée (0.2.3) donnant comme norme optimale la norme L^α du processus prévisible $(H_t)_{t\geq 0}$.

Petits niveaux de déviation

Récemment, Breton et Houdré ont étudié dans [21] le phénomène de concentration de la mesure pour des vecteurs aléatoires stables. En particulier, ils établissent des inégalités pour des petits niveaux de déviation x, dont l'ordre de grandeur est approximativement $\exp\left(-x^{\alpha/(\alpha-1)}\right)$. Nous établissons maintenant un résultat similaire pour des intégrales stables stochastiques, dont la preuve est fondée sur des techniques de martingales avec le choix du niveau de troncature donné par la relation $\lambda R^{1-\alpha} = x/\|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}$.

Théorème 0.3.19. ***(Theorem 5.4.2)*** *On suppose que $(Z_t)_{t\geq 0}$ est symétrique d'indice $\alpha \in (1,2)$ et que le processus borné $(H_t)_{t\geq 0} \in \mathscr{B}_\alpha$ vérifie p.s. $\lim_{t\to+\infty}\int_0^t |H_s|^\alpha ds = +\infty$. Alors pour tout paramètre $\lambda > \lambda_0(\alpha)$, où $\lambda_0(\alpha)$ est l'unique solution de l'équation*

$$\lambda \log\left(1 + \frac{(2-\alpha)\lambda}{2c}\right) = \frac{4c}{\alpha},$$

il existe $x_1(\alpha,\lambda) > 0$ tel que pour tout $0 \leq x \leq x_1(\alpha,\lambda)$ et tout $t \geq 0$,

$$\mathbb{P}\left(\sup_{0\leq s\leq t}\int_0^s H_\tau dZ_\tau \geq x\right) \qquad (0.3.14)$$

$$\leq \frac{2c}{\alpha}\left(\frac{x}{\lambda\|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}}\right)^{\frac{\alpha}{\alpha-1}} + \exp\left(-\frac{\lambda\log\left(1+\frac{(2-\alpha)\lambda}{2c}\right)}{2}\left(\frac{x}{\lambda\|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}}\right)^{\frac{\alpha}{\alpha-1}}\right).$$

Ainsi, on observe que le comportement du processus supremum est différent selon l'ordre de grandeur du niveau de déviation x. Dans le cas où x est grand, la queue de distribution du processus supremum est de l'ordre de $x^{-\alpha}$, ce qui signifie que les grands sauts sont prépondérants, à l'inverse du comportement en $\exp\left(-x^{\alpha/(\alpha-1)}\right)$ établi pour des petits niveaux de déviation. Notons qu'il serait intéressant d'étudier le cas intermédiaire comme le font Breton et Houdré dans [21], où une estimée du type $\exp(-x^{\alpha})$ est donnée pour des valeurs de x variant dans un intervalle compact.

En choisissant les poids de la mesure de Lévy stable de manière appropriée, le processus symétrique stable converge vers un mouvement brownien standard. D'autre part, lorsque α tend vers 2, on a $\lambda_0(\alpha) \to 0$ et $x_1(\alpha, \lambda) \to +\infty$, et le résultat précédent nous permet par un passage à la limite dans (0.3.14) de retrouver l'inégalité maximale gaussienne classique, valable pour tout niveau de déviation $x > 0$:

Corollaire 0.3.20. ***(Corollary 5.4.6)*** *Si $(B_t)_{t\geq 0}$ désigne un mouvement brownien standard, alors on a :*

$$\mathbb{P}\left(\sup_{0\leq s\leq t} B_s \geq x\right) \leq \exp\left(-\frac{x^2}{2t}\right), \quad x > 0, \quad t \geq 0.$$

Temps de passage de processus symétriques stables

Dans l'article [1], Alili et Patie étudient des transformations fonctionnelles liées à des problèmes de temps de passage de processus de diffusion. Nous adaptons cette méthode afin d'estimer le premier instant de passage du module d'un processus symétrique stable au dessus de certaines fonctions continues, en utilisant les inégalités maximales précédentes. Soit $(X_t^{\phi})_{t\geq 0}$ un processus de type Ornstein-Uhlenbeck stable d'indice $\alpha \in (1, 2)$ et de paramètre ϕ. Autrement dit, le processus $(X_t^{\phi})_{t\geq 0}$ a la représentation en intégrale stable

$$X_t^{\phi} := \phi(t) \int_0^t \frac{dZ_s}{\phi(s)}, \quad t \in [0, T), \quad T \in [0, +\infty],$$

où $(Z_t)_{t\geq 0}$ est un processus symétrique α-stable et ϕ est une fonction positive de classe $C^{\infty}([0, T))$. Deux exemples standard de ce type de processus sont le processus d'Ornstein-Uhlenbeck stable ainsi que le pont stable partant de 0. En effet, la fonction ϕ est donnée par $\phi(t) := \exp(-\lambda t)$ pour un $\lambda > 0$ et $T = +\infty$ dans le premier cas, tandis que pour le pont stable, $\phi(t) := T - t$, où $T > 0$ est un horizon fini. Définissons

$$T_x^{\phi} := \inf\{t \in [0, T) : |X_t^{\phi}| \geq x\}$$

le premier instant de sortie du processus $(X_t^{\phi})_{t\geq 0}$ de la boule centrée en 0 de rayon x. Étant donnée une fonction continue positive f telle que $f(0) \neq 0$, on introduit aussi

$$T^{(f)} := \inf\{t \geq 0 : |Z_t| \geq f(t)\}$$

le premier temps de passage du processus positif $(|Z_t|)_{t\geq 0}$ au-dessus de la fonction f. Afin d'utiliser des techniques de changement de temps, on va supposer dans la suite que la fonction τ donnée par

$$\tau_t := \int_0^t \frac{ds}{\phi(s)^\alpha},$$

est bien définie sur $[0, T)$ et qu'elle tend vers l'infini lorsque t tend vers T. Enfin, on désigne par τ^{-1} l'inverse de τ et par $h_{\phi,\tau}$ la transformation fonctionnelle

$$h_{\phi,\tau}(t) := \frac{1}{\phi \circ \tau^{-1}(t)}, \quad t > 0.$$

Ainsi, le choix de la fonction ϕ détermine le changement de temps τ et donc la fonction $h_{\phi,\tau}$. Par exemple, cette fonction est donnée par $t \mapsto (1+\lambda\alpha t)^{1/\alpha}$ dans le cas du processus d'Ornstein-Uhlenbeck stable, tandis que pour le pont stable, $h_{\phi,\tau}$ est la fonction $t \mapsto (T^{1-\alpha} + t(\alpha-1))^{1/(\alpha-1)}$.
On a le lemme fondamental suivant :

Lemme 0.3.21. *Pour tout $x > 0$, on a l'égalité*

$$\mathbb{P}\left(T_x^\phi \in dr\right) = \mathbb{P}\left(\tau^{-1}\left(T^{(xh_{\phi,\tau})}\right) \in dr\right), \quad r \in [0, T). \qquad (0.3.15)$$

Par conséquent, l'identité en loi (0.3.15) fait le lien entre les distributions des premiers temps de passage définis ci-dessus, à travers la transformation fonctionnelle $h_{\phi,\tau}$. Il en résulte que l'analyse du supremum du processus $(X_t^\phi)_{t\geq 0}$, qui caractérise la distribution de la variable aléatoire T_x^ϕ, fournit aussi des informations intéressantes sur celle de $T^{(xh_{\phi,\tau})}$. Bien que les inégalités maximales établies précédemment pour des intégrales stables stochastiques ne puissent pas directement s'appliquer pour le processus $(X_t^\phi)_{t\geq 0}$, l'utilisation d'une formule d'intégration par parties nous permet de ramener l'étude d'inégalités maximales pour $(X_t^\phi)_{t\geq 0}$ à celle concernant le processus symétrique stable $(Z_t)_{t\geq 0}$. Par nos résultats précédents ainsi que par un résultat de support, on obtient alors le

Théorème 0.3.22. ***(Theorem 5.5.5)*** *Supposons $\alpha \in (1, 2)$. Alors pour tout $x > 0$,*

$$\mathbb{P}\left(T^{(xh_{\phi,\tau})} > r\right) \leq \exp\left(-\frac{2c\tau_r^{-1}}{\alpha(2x)^\alpha}\right), \quad r > 0.$$

De plus, il existe une constante $K > 0$, indépendante de α, telle que pour tout $x > 0$ et tout $0 \leq r < r_0(\alpha, x)$,

$$\mathbb{P}\left(T^{(xh_{\phi,\tau})} \leq r\right) \leq \frac{K\tau_r^{-1}}{x^\alpha}\left(1 + \left\|\phi(\cdot)\int_0^\cdot \frac{\phi'(t)}{\phi(t)^2}\,dt\right\|_{L^\infty([0,\tau_r^{-1}])}\right)^\alpha,$$

où $r_0(\alpha, x)$ est l'unique solution de l'équation

$$(2-\alpha)^{\frac{\alpha+1}{\alpha-1}} x^\alpha = c\tau_r^{-1}\left(1 + \left\|\phi(\cdot)\int_0^\cdot \frac{\phi'(t)}{\phi(t)^2}\,dt\right\|_{L^\infty([0,\tau_r^{-1}])}\right)^\alpha.$$

Ainsi, on observe que les estimées sont différentes selon que l'on considère la queue de distribution ou la fonction de répartition de la variable aléatoire $T^{(xh_{\phi,\tau})}$.
Par exemple, pour le processus d'Ornstein-Uhlenbeck stable, nous avons les estimées suivantes : si l'on définit sur $\mathbb{R}_+$ la fonction $f_{\alpha,x,\lambda}(t) := x(1+\lambda\alpha t)^{1/\alpha}$, on a pour tout $x>0$ la borne suivante sur la queue de distribution de la variable aléatoire $T^{(f_{\alpha,x,\lambda})}$:

$$\mathbb{P}\left(\inf\{t\geq 0 : |Z_t| \geq f_{\alpha,x,\lambda}(t)\} > r\right) \leq \frac{1}{(1+\lambda\alpha r)^{\frac{c}{\lambda\,\alpha^2 2^{\alpha-1} x^\alpha}}}, \quad r>0.$$

Par ailleurs, sa fonction de répartition est majorée pour tout $0\leq r<r_0(\alpha,x,\lambda)$ par

$$\mathbb{P}\left(\inf\{t\geq 0 : |Z_t| \geq f_{\alpha,x,\lambda}(t)\} \leq r\right) \leq \frac{(2-(1+\lambda\alpha r)^{-\frac{1}{\alpha}})^\alpha \log(1+\lambda\alpha r) K_c}{\lambda\alpha x^\alpha}.$$

Dans le cas du pont stable, en notant $g_{\alpha,x,T}(t) := x(T^{1-\alpha}+(\alpha-1)t)^{1/(\alpha-1)}$, $t\geq 0$, $x>0$, on obtient l'estimée pour tout $r>0$

$$\mathbb{P}\left(\inf\{t\geq 0 : |Z_t| \geq g_{\alpha,x,\lambda}(t)\} > r\right) \leq \exp\left(-\frac{c}{\alpha 2^{\alpha-1} x^\alpha}\left(T-(T^{1-\alpha}+(\alpha-1)r)^{\frac{1}{1-\alpha}}\right)\right),$$

alors que pour tout $0\leq r<r_0(\alpha,x,T)$,

$$\mathbb{P}\left(\inf\{t\geq 0 : |Z_t| \geq g_{\alpha,x,T}(t)\} \leq r\right) \leq \frac{K_c(Tg_{\alpha,x,T}(r)-x)(2Tg_{\alpha,x,T}(r)-x)^\alpha}{T^\alpha g_{\alpha,x,T}(r)^{\alpha+1}x^\alpha}.$$

Chapitre 6 : domination convexe pour des intégrales brownienne et stable dépendantes

Le chapitre 6 fait l'objet de l'article en préparation [53], co-écrit avec Yutao Ma, dont l'objectif est d'établir un principe de domination convexe pour des intégrales stochastiques dirigées par un mouvement brownien et un processus stable symétrique non nécessairement indépendants. Nous démontrons que sous certaines conditions de bornitude, une variable aléatoire admettant une représentation en termes d'intégrales stochastiques brownienne et stable corrélées est dominée au sens convexe par la somme indépendante de variables aléatoires gaussienne et stable symétrique. L'approche utilisée repose sur le calcul stochastique *forward-backward* récemment développé dans [59] et nous permet de découpler la paire d'intégrales stochastiques dépendantes.

Soit $(W_t)_{t\geq 0}$ un mouvement brownien standard réel non nécessairement indépendant d'un processus symétrique stable $(Z_t)_{t\geq 0}$ d'indice $\alpha\in(1,2)$ et de mesure de Lévy stable définie sur $\mathbb{R}\setminus\{0\}$ par $\sigma(dx) := c|x|^{-\alpha-1}dx$, $c>0$. En notant la filtration

$$\mathscr{F}_t^{W,Z} := \sigma\left(W_s, Z_s : 0\leq s\leq t\right), \quad t\geq 0,$$

on suppose dans la suite que ces deux processus sont des $(\mathscr{F}_t^{W,Z})_{t\geq 0}$-martingales. En d'autres termes, les accroissements du premier processus sont indépendants du passé du second, et réciproquement. On considère une variable aléatoire F ayant la représentation

$$F - \mathbb{E}[F] = \int_0^{+\infty} H_t dW_t + \int_0^{+\infty} K_t dZ_t, \tag{0.3.16}$$

où les processus bornés $(H_t)_{t\geq 0} \in \mathscr{P}_2$ et $(K_t)_{t\geq 0} \in \mathscr{P}_\alpha$ sont supposés $(\mathscr{F}_t^{W,Z})_{t\geq 0}$-prévisibles. On suppose de plus que $(K_t)_{t\geq 0}$ est de carré integrable afin que la variable aléatoire $\int_0^{+\infty} K_t dZ_t$ soit bien définie comme intégrale stochastique par rapport à la décomposition de Lévy-Itô du processus stable $(Z_t)_{t\geq 0}$. Nous allons établir un principe de domination convexe pour la variable aléatoire centrée $F - \mathbb{E}[F]$.

Théorème 0.3.23. ***(Theorem 6.2.1)*** *Il existe une variable aléatoire gaussienne $\tilde{W}(\beta_1)$ de variance*

$$\beta_1 := \left\| \int_0^{+\infty} |H_t|^2 dt \right\|_\infty,$$

indépendante d'une variable aléatoire réelle symétrique stable $\tilde{Z}(\beta_2)$ d'index α et de mesure de Lévy sur $\mathbb{R} \setminus \{0\}$ donnée par

$$\tilde{\sigma}(dx) = \frac{\beta_2 \tilde{c} dx}{|x|^{\alpha+1}}, \quad \textit{avec} \quad 0 < c \leq \tilde{c} \quad \textit{et} \quad \beta_2 := \left\| \int_0^{+\infty} |K_t|^\alpha dt \right\|_\infty,$$

telles que pour toute fonction convexe ϕ à croissance au plus polynomiale d'ordre $p \in (0, \alpha)$ à l'infini, on ait la relation de domination convexe

$$\mathbb{E}\left[\phi(F - \mathbb{E}F)\right] \leq \mathbb{E}\left[\phi\left(\tilde{W}(\beta_1) + \tilde{Z}(\beta_2)\right)\right]. \tag{0.3.17}$$

La preuve du théorème 0.3.23, que nous allons décrire brièvement, est divisée en plusieurs étapes. Tout d'abord, nous conditionnons la variable aléatoire centrée $F - \mathbb{E}[F]$ afin d'obtenir une martingale $(X_t)_{t\geq 0}$, à laquelle on associe une filtration $(\mathscr{F}_t)_{t\geq 0}$. Ensuite, nous construisons de manière appropriée une martingale rétrograde $(X_t^*)_{t\geq 0}$ par rapport à une filtration décroissante $(\mathscr{F}_t^*)_{t\geq 0}$, qui ne dépend de $(X_t)_{t\geq 0}$ qu'à travers des temps aléatoires donnés par les intégrales $\int_0^t |H_s|^2 ds$ et $\int_0^t |K_s|^\alpha ds$. En particulier, la valeur finale X_0^* étant la somme des variables indépendantes $\tilde{W}(\beta_1)$ et $\tilde{Z}(\beta_2)$, on remarque que la somme corrélée $X_t + X_t^*$, $t > 0$, est découplée à l'instant $t = 0$. Après avoir identifié les caractéristiques locales de ces martingales, la formule d'Itô pour des martingales forward-backward, cf. [59, théorème 8.1], appliquée à la somme $(X_t + X_t^*)_{t\geq 0}$, donne l'identité

$$\begin{aligned} &\mathbb{E}\left[\phi(X_t + X_t^*)\right] - \mathbb{E}\left[\phi(\tilde{W}(\beta_1) + \tilde{Z}(\beta_2))\right] \\ &\quad = \mathbb{E}\left[\int_0^t \int_{-\infty}^{+\infty} \int_0^1 (1-\tau) x^2 \phi''(X_u + X_u^* + \tau x) |K_u|^\alpha \frac{(c - \tilde{c})}{|x|^{\alpha+1}} d\tau dx du\right], \quad t \geq 0, \end{aligned}$$

quantité qui est négative car la fonction ϕ est convexe et $0 < c \leq \tilde{c}$. Enfin, comme la projection de la variable aléatoire X_t^* sur la tribu $\mathscr{F}_t^{W,Z}$ est nulle à chaque instant $t > 0$ fixé, un argument de type Jensen complète la démonstration.

Partie I

Concentration des processus de naissance et de mort

Chapitre 1

Functional inequalities for discrete gradients and applications to the geometric distribution

Ce chapitre fait l'objet d'un article écrit en collaboration avec Nicolas Privault et publié dans le journal *ESAIM, Probability and Statistics*, volume 8, pages 87-101, 2004.

Abstract

We present several functional inequalities for finite difference gradients, such as a Cheeger inequality, a proof that Poincaré inequalities imply modified logarithmic Sobolev inequalities and associated deviation estimates, and an exponential integrability property. In the particular case of the geometric distribution on $\mathbb{N}$ we use an integration by parts formula to compute the optimal isoperimetric and Poincaré constants, and to obtain an improvement of our general logarithmic Sobolev inequality. By a limiting procedure we recover the corresponding inequalities for the exponential distribution.

1.1 Introduction

Isoperimetry consists in finding sets of minimal surface among sets of a given volume, i.e. to search for optimal constants in inequalities of the form

$$cI\left(\mu(A)\right) \leq \mu_s(A), \tag{1.1.1}$$

where μ_s and μ are respectively surface and volume measures and I is a non-negative function on $[0, 1]$. Isoperimetric constants are linked via co-area formulas to functional inequalities such as Poincaré or logarithmic Sobolev inequalities. Discrete isoperimetry has been studied in various contexts, such as reversible Markov chains [34], [46], graph theory [17, 78], statistical mechanics, cf. e.g. [31].

In this paper we consider the general discrete setting of a probability space $(E, \mathcal{E}, \mu)$, and a finite difference gradient d^+ defined as $\mathrm{d}^+ f = f \circ \tau - f$, where $\tau : E \to E$ is an absolutely continuous mapping. Typically $E = \mathbb{N}$ and $\mathrm{d}^+ f(k) = f(k+1) - f(k)$, in this case d^+ can be used to express the surface measure of a set as the expectation of a discrete gradient norm. However, E can be a more general, even uncountable, space. The abstract case of a metric space has been considered in [16], [18] for a gradient having the derivation property.

In Section 1.2 we prove a discrete generalization of Cheeger's inequality [26], i.e. a lower bound on the spectral gap λ_μ in terms of the isoperimetry constant h_μ, using the arguments of [17] and [78]. When μ is the geometric distribution π on $\mathbb{N}$ with parameter $p \in (0,1)$ we show in Section 2.1 that $h_\pi = (1-p)/p$ and $\lambda_\pi = (1-\sqrt{p})^2/p$. The lower bound for λ_π obtained from Cheeger's inequality turns out to be optimal for the geometric distribution.

A measure μ is said to satisfy a logarithmic Sobolev inequality [41] with gradient d when

$$\mathrm{Ent}_\mu[f^2] \le C \mathbb{E}_\mu \left[|\mathrm{d}f|^2 \right], \tag{1.1.2}$$

where $\mathrm{Ent}_\mu[f] = \mathbb{E}_\mu[f \log f] - \mathbb{E}_\mu[f] \log \mathbb{E}_\mu[f]$ denotes the entropy of f under μ. If the gradient $\mathrm{d}f$ has the derivation property, (1.1.2) is equivalent to the following modified logarithmic Sobolev inequality

$$\mathrm{Ent}_\mu[e^f] \le \frac{C}{4} \mathbb{E}_\mu \left[|\mathrm{d}f|^2 e^f \right]. \tag{1.1.3}$$

Such modified inequalities have been proved for Poisson and Bernoulli measures on $\mathbb{N}$ in [19], using the finite difference gradient d^+. On the other hand, modified logarithmic Sobolev inequalities for the exponential distribution have been obtained in [18] under the additional hypothesis that f is $c-$Lipschitz, i.e. $|\mathrm{d}f| \le c$, when d has the derivation property.

In Section 1.3.2 we adapt the method of [18] to the geometric distribution, which can be viewed as a discrete analog of the exponential distribution, since the inter-jump times of the Poisson, resp. binomial, process have exponential, resp. geometric, distributions. For this we use an integration by parts formula and replace the derivation rule used in [18] with bounds on the finite difference gradient d^+, deduced from the mean value theorem. As noted in [31], (1.1.3) does not hold as stated for d^+ under the geometric distribution with parameter p (take $f_a(n) = n \log a$ and let $a \nearrow 1/p$). We will show that (1.1.3) does hold for the geometric distribution under the further assumption $|\mathrm{d}^+ f| \le c$, with a constant depending on c. In Section 1.3.3, using the Herbst method we obtain a deviation result for the geometric distribution, which differs from the deviation inequality recently obtained in [42] from the covariance representation method for infinitely divisible distributions. Although the integral part $[X]$ of an exponential random variable with parameter λ has a geometric law of parameter $e^{-\lambda}$, it does not seem possible to apply existing results on the exponential distribution [18] in our setting. For example, $f([X])$ is not the composition of a Lipschitz function with X. However, exponential random variables can be approximated in distribution by geometric random variables, and in this way we recover the functional inequalities proved in [18] for the exponential distribution.

In Section 1.4 we obtain a more general result, stating that any distribution μ that satisfies a Poincaré inequality with constant λ_μ for a finite difference gradient also satisfies a logarithmic Sobolev inequality of modified type for all function f such that $|\mathrm{d}^+ f| \leq c$, which implies deviation bounds. Finally, we present in Section 1.4.2 an exponential integrability criterion.

Notation

Given a probability space $(E, \mathcal{E}, \mu)$, let $\tau : E \to E$ denote a map absolutely continuous with respect to μ. We denote by d^+ the finite difference gradient operator defined as

$$\mathrm{d}^+ f = f \circ \tau - f = \tau f - f,$$

where $f \circ \tau$ will be denoted by τf for shortness of notation. If $x \in \mathbb{R}$ is such that $\mu(f \geq x) \geq 1/2$ and $\mu(f \leq x) \geq 1/2$, we say that x is a median of f under μ, and write $m(f) = x$. We recall that for every median of f we have

$$\mathbb{E}_\mu\left[|f - m(f)|\right] = \inf_{a \in \mathbb{R}} \mathbb{E}_\mu\left[|f - a|\right].$$

We will need the co-area formula

$$\mathbb{E}_\mu\left[|\mathrm{d}^+ f|\right] = \int_{-\infty}^{+\infty} \mathbb{E}_\mu\left[|\mathrm{d}^+ 1_{\{f>t\}}|\right] dt, \tag{1.1.4}$$

which follows easily from the relations

$$(b-a)^{\pm} = \int_{-\infty}^{\infty} (1_{\{a>t\}} - 1_{\{b>t\}})^{\pm} dt, \quad a, b \in \mathbb{R}.$$

If $E = \mathbb{N}$ then it is natural to consider the shift $\tau f(k) = f(k+1)$, $k \in \mathbb{N}$, and the associated gradient

$$\mathrm{d}^+ f(k) = f(k+1) - f(k), \quad k \in \mathbb{N}. \tag{1.1.5}$$

Note that given $A \subset \mathbb{N}$ we have

$$\{|\mathrm{d}^+ 1_A| > 0\} = \{|\mathrm{d}^+ 1_A| \geq 1\} = \{k \in A \ : \ k+1 \in A^c\} \cup \{k \in A^c \ : \ k+1 \in A\},$$

i.e. $\{|\mathrm{d}^+ 1_A| > 0\}$ represents a frontier ∂A of A, and $\mathbb{E}_\mu[|\mathrm{d}^+ 1_A|]$ represents the measure of ∂A. Throughout this paper, μ denotes an arbitrary probability measure on E, while π denotes the geometric distribution with parameter $p \in (0,1)$ on $E = \mathbb{N}$.

1.2 Isoperimetric and Poincaré inequalities

Given a measure μ on E, let h_μ denote the optimal constant in the inequality

$$h_\mu \mathbb{E}_\mu\left[|f - m(f)|\right] \leq \mathbb{E}_\mu\left[|\mathrm{d}^+ f|\right], \tag{1.2.1}$$

i.e.

$$h_\mu = \inf_{f \neq \mathrm{const}} \frac{\mathbb{E}_\mu\left[|\mathrm{d}^+ f|\right]}{\mathbb{E}_\mu\left[|f - m(f)|\right]}.$$

Several analogs of Proposition 1.2.1 and Proposition 1.2.2 below have already been proved in [17], [46], [78], for connected graphs and for Markov chains, under reversibility or ergodicity assumptions. The gradient used in our setting is different but the proofs are similar and stated for completeness.

Proposition 1.2.1. *We have*

$$h_\mu = \inf_{0 < \mu(A) \leq \frac{1}{2}} \frac{\mathbb{E}_\mu\left[|\mathrm{d}^+ 1_A|\right]}{\mu(A)}. \tag{1.2.2}$$

Proof. We will prove the equality

$$h_\mu = \inf_{\mu(A) > 0} \frac{\mathbb{E}_\mu\left[|\mathrm{d}^+ 1_A|\right]}{\min(\mu(A), 1 - \mu(A))},$$

which clearly implies (1.2.2). Recall that $m(1_A) = 0$ if $\mu(A) \leq 1/2$, and $m(1_A) = 1$ if $\mu(A) \geq 1/2$, and $\mathbb{E}_\mu\left[|1_A - m(1_A)|\right] = \min(\mu(A), 1 - \mu(A))$. Assume that for some $h > 0$,

$$h\mathbb{E}_\mu\left[|1_A - m(1_A)|\right] \leq \mathbb{E}_\mu\left[|\mathrm{d}^+ 1_A|\right], \quad A \in \mathcal{E}.$$

From the co-area formula (1.1.4) we have, since $m(1_{\{f>t\}}) = 0$, $t \geq m(f)$, and $m(1_{\{f \leq t\}}) = 0$, $t \leq m(f)$:

$$\begin{aligned}
\mathbb{E}_\mu\left[|\mathrm{d}^+ f|\right] &= \int_{-\infty}^{+\infty} \mathbb{E}_\mu\left[|\mathrm{d}^+ 1_{\{f>t\}}|\right] dt \\
&\geq \int_{m(f)}^{+\infty} \mathbb{E}_\mu\left[|\mathrm{d}^+ 1_{\{f>t\}}|\right] dt + \int_{-\infty}^{m(f)} \mathbb{E}_\mu\left[|\mathrm{d}^+ 1_{\{f \leq t\}}|\right] dt \\
&\geq h \int_{m(f)}^{+\infty} \mathbb{E}_\mu\left[1_{\{f>t\}}\right] dt + h \int_{-\infty}^{m(f)} \mathbb{E}_\mu\left[1_{\{f \leq t\}}\right] dt \\
&= h\mathbb{E}_\mu\left[(f - m(f))^+\right] + h\mathbb{E}_\mu\left[(m(f) - f)^+\right] \\
&= h\mathbb{E}_\mu[|f - m(f)|],
\end{aligned}$$

hence $h \leq h_\mu$. This concludes the proof, since the converse inequality is obvious. ∎

Let λ_μ denote the optimal constant in the Poincaré inequality

$$\lambda_\mu \mathrm{Var}_\mu[f] \leq \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right], \tag{1.2.3}$$

under μ, i.e.

$$\lambda_\mu = \inf_{f\neq \mathrm{const}} \frac{\mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right]}{\mathrm{Var}_\mu\left[f\right]}.$$

We now prove a Cheeger type inequality, i.e. a lower bound on λ_μ which shows that the strict positivity of h_μ implies a Poincaré inequality.

Proposition 1.2.2. *We have*

$$\left(\sqrt{1+h_\mu}-1\right)^2 \leq \lambda_\mu \leq 2h_\mu. \tag{1.2.4}$$

Proof. Given a function f, let $g = f - m(f)$. We have $m(g) = 0$, which implies $m({g^+}^2) = m({g^-}^2) = 0$. Applying (1.2.1) to ${g^+}^2$ and ${g^-}^2$ we get

$$\begin{aligned}
h_\mu \mathbb{E}_\mu\left[g^2\right] &= h_\mu \mathbb{E}_\mu\left[{g^+}^2\right] + h_\mu \mathbb{E}_\mu\left[{g^-}^2\right] \\
&\leq \mathbb{E}_\mu\left[|\mathrm{d}^+ {g^+}^2| + |\mathrm{d}^+ {g^-}^2|\right] \\
&= \mathbb{E}_\mu\left[|2g^+\mathrm{d}^+g^+ + |\mathrm{d}^+g^+|^2| + |2g^-\mathrm{d}^+g^- + |\mathrm{d}^+g^-|^2|\right] \\
&\leq 2\mathbb{E}_\mu\left[g^+|\mathrm{d}^+g^+| + g^-|\mathrm{d}^+g^-|\right] + \mathbb{E}_\mu\left[|\mathrm{d}^+g^+|^2 + |\mathrm{d}^+g^-|^2\right] \\
&\leq 2\mathbb{E}_\mu\left[|g|\left(|\mathrm{d}^+g^+| + |\mathrm{d}^+g^-|\right)\right] + \mathbb{E}_\mu\left[|\mathrm{d}^+g^+|^2 + |\mathrm{d}^+g^-|^2\right] \\
&\leq 2\mathbb{E}_\mu\left[|g||\mathrm{d}^+g|\right] + \mathbb{E}_\mu\left[|\mathrm{d}^+g|^2\right] \\
&\leq 2\|g\|_2\|\mathrm{d}^+g\|_2 + \|\mathrm{d}^+g\|_2^2,
\end{aligned}$$

where we used the relations $|\mathrm{d}^+g^+| + |\mathrm{d}^+g^-| = |\mathrm{d}^+g|$ and $|\mathrm{d}^+g^+|^2 + |\mathrm{d}^+g^-|^2 \leq |\mathrm{d}^+g|^2$. This implies

$$(\sqrt{1+h_\mu}-1)\|g\|_2 \leq \|\mathrm{d}^+g\|_2.$$

In the general case we have

$$\begin{aligned}
(\sqrt{1+h_\mu}-1)^2\mathrm{Var}_\mu[f] &= (\sqrt{1+h_\mu}-1)^2\mathrm{Var}_\mu[g] \\
&\leq (\sqrt{1+h_\mu}-1)^2\|g\|_2^2 \\
&\leq \mathbb{E}_\mu[|\mathrm{d}^+g|^2] = \mathbb{E}_\mu[|\mathrm{d}^+f|^2],
\end{aligned}$$

therefore $\lambda_\mu \geq \left(\sqrt{1+h_\mu}-1\right)^2$. Moreover we have

$$\lambda_\mu = \inf_{f\neq \mathrm{const}} \frac{\mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right]}{\mathrm{Var}_\mu\left[f\right]} \leq \inf_{\substack{\emptyset\neq A\in\mathcal{E} \\ \mu(A)\leq 1/2}} \frac{\mathbb{E}_\mu\left[|\mathrm{d}^+ 1_A|\right]}{\mu(A)(1-\mu(A))} \leq 2h_\mu.$$

∎

Note that (1.2.4) also yields an upper bound on h_μ in terms of λ_μ:

$$h_\mu \leq \lambda_\mu + 2\sqrt{\lambda_\mu}. \tag{1.2.5}$$

1.3 The geometric distribution

1.3.1 Optimal isoperimetric and Poincaré constants

We consider $E = \mathbb{N}$ and the gradient

$$\mathrm{d}^+ f(k) = f(k+1) - f(k), \quad k \in \mathbb{N}.$$

Under π the Laplacian $\mathcal{L} = -\mathrm{d}_\pi^{+*}\mathrm{d}^+$ is given by

$$-\mathrm{d}_\pi^{+*}\mathrm{d}^+ f(k) = f(k+1) - f(k) + \frac{1}{p} 1_{\{k\geq 1\}}(f(k-1) - f(k)),$$

i.e. $\mathcal{L} = \mathrm{d}^+ + \frac{1}{p}\mathrm{d}^-$ with

$$\mathrm{d}^- f(k) = 1_{\{k\geq 1\}}(f(k-1) - f(k)), \quad k \in \mathbb{N}.$$

Poincaré inequalities for general discrete distributions have been proved in [15], [23], [27], [67]. Theorem 1.3 in [15] shows in particular that a discrete distribution μ on $\mathbb{N}$ satisfies (1.2.3) if and only if

$$\mu(\{n\}) \geq c\mu([0,n])(1 - \mu([0,n])), \quad n \geq 0,$$

for some constant $c > 0$. It is easily seen that the geometric distribution π with parameter $p \in (0,1)$ given by:

$$\pi(\{k\}) = p^k(1-p), \quad k \in \mathbb{N},$$

does satisfy this hypothesis. In this section we prove an isoperimetric inequality for the geometric distribution, which will imply a Poincaré inequality from Cheeger's inequality (1.2.4). The proof relies as in [18] on an integration by parts formula under π.

Lemma 1.3.1. *Let $f : \mathbb{N} \to \mathbb{R}$. We have*

$$\mathbb{E}_\pi[f] = f(0) + \frac{p}{1-p}\mathbb{E}_\pi[\mathrm{d}^+ f]. \tag{1.3.1}$$

Proof. Letting $g = f - f(0)$ we have the Radon-Nikodym type relation

$$\mathbb{E}_\pi[\tau g] = \frac{1}{p}\mathbb{E}_\pi[g], \tag{1.3.2}$$

since $g(0) = 0$, and

$$\mathbb{E}_\pi[\mathrm{d}^+ f] = \mathbb{E}_\pi[\mathrm{d}^+ g] = \mathbb{E}_\pi[\tau g] - \mathbb{E}_\pi[g] = \left(\frac{1}{p} - 1\right)\mathbb{E}_\pi[g] = \frac{1-p}{p}(\mathbb{E}_\pi[f] - f(0)).$$

■

Note that to some extent, (1.3.2) characterizes the values $\pi(k)$ of the geometric distribution, $k \geq 1$, except for $\pi(0)$. Instead of d^+ we may use the gradient d^-, since similarly to the integration by parts formula we have the isometry

$$\mathbb{E}_\pi[N(\mathrm{d}^+ f)] = \frac{1}{p}\mathbb{E}_\pi[N(-\mathrm{d}^- f)],$$

for e.g. $N(x) = x$, $N(x) = |x|$, $N(x) = |x|^2$, which is equivalent to the reversibility of the birth and death process with generator $\mathcal{L} = -\mathrm{d}_\pi^{+*}\mathrm{d}^+$. In particular the gradient norm expectations generally used in the context of graphs and Markov chains [17], [46], [78], are here of the form

$$\mathbb{E}_\pi[N(\mathrm{d}^+ f)] + \frac{1}{p}\mathbb{E}_\pi[N(\mathrm{d}^- f)] = 2\mathbb{E}_\pi[N(\mathrm{d}^+ f)]$$

for $N(x) = |x|$, $N(x) = |x|^2$, and coincide with $\mathbb{E}_\pi[N(\mathrm{d}^+ f)]$ up to a constant factor.

Proposition 1.3.2. *Under the geometric distribution π we have*

$$h_\pi = \frac{1-p}{p}. \tag{1.3.3}$$

Proof. From the integration by parts formula (1.3.1) we have

$$\mathbb{E}_\pi\left[|f - m(f)|\right] \leq \mathbb{E}_\pi\left[|f - f(0)|\right] = \frac{p}{1-p}\mathbb{E}_\pi\left[\mathrm{d}^+|f - f(0)|\right] \leq \frac{p}{1-p}\mathbb{E}_\pi\left[|\mathrm{d}^+ f|\right],$$

which shows $h_\pi \geq (1-p)/p$. On the other hand, letting $f_n = 1_{[n+1,\infty)}$, $n \in \mathbb{N}$, we have for any $n \in \mathbb{N}$ such that $\pi([n+1,\infty)) \leq 1/2$:

$$h_\pi \leq \frac{\mathbb{E}_\pi\left[|\mathrm{d}^+ f|\right]}{\mathbb{E}_\pi\left[|f - m(f)|\right]} = \frac{\pi(\{n\})}{\pi([n+1,\infty))} = \frac{1-p}{p}.$$

■

In particular, the isoperimetric inequality becomes an inequality for functions of the form $f_n = 1_{[n+1,\infty)}$, with $n \geq -\log 2/\log p$.

Proposition 1.3.3. *Under the geometric distribution π we have*

$$\lambda_\pi = \frac{(1-\sqrt{p})^2}{p}. \tag{1.3.4}$$

Proof. Using Cheeger's inequality (1.2.4) and Relation (1.3.3) we get $(1-\sqrt{p})^2/p \leq \lambda_\pi$. On the other hand, with $f_a(k) = a^k$ we have:

$$\lambda_\pi \leq \frac{\mathbb{E}_\pi\left[|\mathrm{d}^+ f_a|^2\right]}{\mathrm{Var}_\pi[f_a]} = (a-1)^2\frac{a^2p^2 + 1 - 2ap}{a^2p + p - 2ap}, \quad a < 1/\sqrt{p},$$

and taking the limit as $a \to 1/\sqrt{p}$ we get $\lambda_\pi \leq (1-\sqrt{p})^2/p$. ■

The Poincaré inequality is not an equality in the linear case $f(k) = a + bk$:

$$\mathrm{Var}_\pi[f] = \frac{p}{(1-p)^2}\mathbb{E}_\pi\left[|\mathrm{d}^+ f|^2\right].$$

Here, the lower bound on λ_π from Cheeger's inequality coincides with the optimal Poincaré constant.

Remark 1.3.4. *The lower bound of λ_π can be directly obtained from the integration by parts formula* (1.3.1) *under π.*

Proof. Letting $g = f - f(0)$ we have $g(0) = 0$ and from (1.3.1) applied to g^2 we obtain:

$$\begin{aligned}
\|g\|_2^2 &= \frac{p}{1-p}\mathbb{E}_\pi\left[\mathrm{d}^+(g^2)\right] \\
&= \frac{p}{1-p}\mathbb{E}_\pi\left[g\mathrm{d}^+ g + \tau g\mathrm{d}^+ g\right] \\
&\leq \frac{p}{1-p}\left(\|g\|_2\|\mathrm{d}^+ g\|_2 + \|\tau g\|_2\|\mathrm{d}^+ g\|_2\right) \\
&= \frac{p}{1-p}\left(\|g\|_2\|\mathrm{d}^+ f\|_2 + \frac{1}{\sqrt{p}}\|g\|_2\|\mathrm{d}^+ f\|_2\right) \\
&= \frac{\sqrt{p}}{1-\sqrt{p}}\|g\|_2\|\mathrm{d}^+ f\|_2,
\end{aligned}$$

hence

$$\|g\|_2 \leq \frac{\sqrt{p}}{1-\sqrt{p}}\|\mathrm{d}^+ f\|_2,$$

and

$$\frac{(1-\sqrt{p})^2}{p}\mathrm{Var}_\pi[f] = \frac{(1-\sqrt{p})^2}{p}\mathrm{Var}_\pi[g] \leq \frac{(1-\sqrt{p})^2}{p}\|g\|_2^2 \leq \mathbb{E}_\pi\left[|\mathrm{d}^+ f|^2\right].$$

■

If X_ε is geometric with parameter p^ε, then $\varepsilon X_\varepsilon$ converges in distribution to an exponential random variable Y with parameter $-\log p$. Given f a c-Lipschitz function on $\mathbb{R}$ we have

$$\mathrm{Var}\left[f(\varepsilon X_\varepsilon)\right] \leq \frac{\varepsilon^2 p^\varepsilon}{(1-\sqrt{p^\varepsilon})^2}\mathbb{E}\left[\left(\frac{f(\varepsilon X_\varepsilon + \varepsilon) - f(\varepsilon X_\varepsilon)}{\varepsilon}\right)^2\right],$$

which as ε goes to 0 yields the Poincaré inequality of Lemma 2.1 in [18] for an exponentially distributed random variable Y with parameter $-\log p$:

$$\mathrm{Var}\left[f(Y)\right] \leq \frac{4}{(\log p)^2}\mathbb{E}\left[|f'(Y)|^2\right],$$

the constant $4/(\log p)^2$ being also optimal, cf. [38]. In a similar way, we obtain from Proposition 1.3.2 an isoperimetric inequality under the exponential distribution with parameter $-\log p$:

$$\mathbb{E}[|f(Y) - m(f(Y))|] \leq -\frac{1}{\log p}\mathbb{E}\left[|f'(Y)|\right].$$

The above constants $h = -\log p$ and $\lambda = (\log p)^2/4$ also satisfy the classical Cheeger inequality $\lambda \geq h^2/4$ which holds in the continuous case, cf. [26].

1.3.2 Modified logarithmic Sobolev inequality

In this section we obtain a modified logarithmic Sobolev inequality for the geometric distribution π on $E = \mathbb{N}$, with $\mathrm{d}^+ f(k) = f(k+1) - f(k)$, $k \in \mathbb{N}$.

Lemma 1.3.5. *Let $c < -\log p$ and let $f : \mathbb{N} \to \mathbb{R}$ be such that $\mathrm{d}^+ f \leq c$ and $f(0) = 0$. We have*

$$\mathbb{E}_\pi \left[f^2 e^f \right] \leq \frac{pe^c}{(1-\sqrt{pe^c})^2} \mathbb{E}_\pi \left[e^f |\mathrm{d}^+ f|^2 \right]. \tag{1.3.5}$$

Proof. From the integration by parts formula (1.3.1) we have

$$\begin{aligned}
\mathbb{E}_\pi \left[f^2 e^f \right] &= \frac{p}{1-p} \mathbb{E}_\pi \left[\mathrm{d}^+ (f^2 e^f) \right] \\
&= \frac{p}{1-p} \mathbb{E}_\pi \left[e^f \left(e^{\mathrm{d}^+ f} \left(|\mathrm{d}^+ f|^2 + 2f \mathrm{d}^+ f \right) + f^2 (e^{\mathrm{d}^+ f} - 1) \right) \right] \\
&\leq \frac{pe^c}{1-p} \mathbb{E}_\pi \left[e^f \left(|\mathrm{d}^+ f|^2 + 2|f||\mathrm{d}^+ f| \right) \right] + \frac{p(e^c - 1)}{1-p} \mathbb{E}_\pi \left[f^2 e^f \right],
\end{aligned}$$

hence

$$\begin{aligned}
\mathbb{E}_\pi \left[f^2 e^f \right] &\leq \frac{pe^c}{1-pe^c} \mathbb{E}_\pi \left[e^f \left(|\mathrm{d}^+ f|^2 + 2|f||\mathrm{d}^+ f| \right) \right] \\
&\leq \frac{pe^c}{1-pe^c} \mathbb{E}_\pi \left[|\mathrm{d}^+ f|^2 e^f \right] + 2 \frac{pe^c}{1-pe^c} \mathbb{E}_\pi \left[f^2 e^f \right]^{1/2} \mathbb{E}_\pi \left[e^f |\mathrm{d}^+ f|^2 \right]^{1/2},
\end{aligned}$$

which implies (1.3.5). ■

Theorem 1.3.6. *Let $0 < c < -\log p$ and let $f : \mathbb{N} \to \mathbb{R}$ such that $|\mathrm{d}^+ f| \leq c$. We have*

$$\mathrm{Ent}_\pi \left[e^f \right] \leq \frac{pe^c}{(1-p)(1-\sqrt{pe^c})} \mathbb{E}_\pi \left[|\mathrm{d}^+ f|^2 e^f \right]. \tag{1.3.6}$$

Proof. From the inequality $-u \ln u \leq 1 - u$, $u > 0$, we have:

$$\mathrm{Ent}_\pi \left[e^f \right] = \mathbb{E}_\pi \left[f e^f \right] - \mathbb{E}_\pi \left[e^f \right] \ln \mathbb{E}_\pi \left[e^f \right] \leq \mathbb{E}_\pi \left[f e^f - e^f + 1 \right]. \tag{1.3.7}$$

Let again $g = f - f(0)$, and let $h(v) = ve^v - e^v + 1$. We have $h \circ g(0) = 0$, and applying (1.3.1) to $h \circ g$ we get :

$$\begin{aligned}
\mathrm{Ent}_\pi \left[e^g \right] &\leq \mathbb{E}_\pi [h \circ g] \\
&= \frac{p}{1-p} \mathbb{E}_\pi [\mathrm{d}^+ (h \circ g)] \\
&= \frac{p}{1-p} \mathbb{E}_\pi [h \circ (g + \mathrm{d}^+ g) - h \circ g]
\end{aligned}$$

$$\leq \frac{p}{1-p}\mathbb{E}_\pi\left[\left(|\mathrm{d}^+g|^2 + |g||\mathrm{d}^+g|\right)e^{g+|\mathrm{d}^+g|}\right],$$

where the inequality

$$h(a+b) - h(a) \leq (b^2 + |ab|)e^{a+|b|}, \qquad a, b \in \mathbb{R},$$

follows from the mean value theorem. From Lemma 1.3.5 and the Schwarz inequality we obtain:

$$\begin{aligned}
\mathrm{Ent}_\pi\left[e^f\right] &= e^{f(0)}\mathrm{Ent}_\pi\left[e^g\right] \\
&\leq \frac{pe^{f(0)}}{1-p}\mathbb{E}_\pi\left[\left(|\mathrm{d}^+g|^2 + |g||\mathrm{d}^+g|\right)e^{g+|\mathrm{d}^+g|}\right]. \\
&\leq \frac{pe^{c+f(0)}}{1-p}\left(\mathbb{E}_\pi\left[|\mathrm{d}^+g|^2e^g\right] + \mathbb{E}_\pi\left[g^2e^g\right]^{1/2}\mathbb{E}_\pi\left[e^g|\mathrm{d}^+g|^2\right]^{1/2}\right) \\
&\leq \frac{pe^{c+f(0)}}{1-p}\left(1 + \frac{\sqrt{pe^c}}{1-\sqrt{pe^c}}\right)\mathbb{E}_\pi\left[|\mathrm{d}^+g|^2e^g\right] \\
&= \frac{pe^c}{(1-p)(1-\sqrt{pe^c})}\mathbb{E}_\pi\left[|\mathrm{d}^+f|^2e^f\right].
\end{aligned}$$

■

In higher dimension we consider the multi-dimensional gradient defined as

$$\mathrm{d}_i^+f(k) = f(k+e_i) - f(k), \quad i = 1, \ldots, n,$$

where f is a function on $\mathbb{N}^n$, $k = (k_1, \ldots, k_n) \in \mathbb{N}^n$, $(e_1, \ldots, e_n)$ is the canonical basis of $\mathbb{R}^n$, and the gradient norm

$$\|\mathrm{d}^+f(k)\|^2 = \sum_{i=1}^n |\mathrm{d}_i^+f(k)|^2 = \sum_{i=1}^n |f(k+e_i) - f(k)|^2. \tag{1.3.8}$$

From the tensorization property of entropy, (1.3.6) still holds with respect to $\pi^{\otimes n}$ in any finite dimension n:

$$\mathrm{Ent}_{\pi^{\otimes n}}\left[e^f\right] \leq \frac{pe^c}{(1-p)(1-\sqrt{pe^c})}\mathbb{E}_{\pi^{\otimes n}}\left[\|\mathrm{d}^+f\|^2e^f\right], \tag{1.3.9}$$

provided $|\mathrm{d}_i^+f| \leq c$, $i = 1, \ldots, n$ (we may also take $(1-p)^{-1}(1-\sqrt{pe^c})^{-1}$ as logarithmic Sobolev constant). Applying again (1.3.6) to X_ε we get for every c-Lipschitz function f:

$$\mathrm{Ent}_\pi\left[e^{f(\varepsilon X_\varepsilon)}\right] \leq \frac{\varepsilon^2 p^\varepsilon e^{\varepsilon c}}{(1-p^\varepsilon)(1-\sqrt{p^\varepsilon e^{\varepsilon c}})}\mathbb{E}_\pi\left[e^{f(\varepsilon X_\varepsilon)}\left(\frac{f(\varepsilon X_\varepsilon + \varepsilon) - f(X_\varepsilon)}{\varepsilon^2}\right)^2\right],$$

which in the limit as ε goes to 0 yields the logarithmic Sobolev inequality of Proposition 2.2 in [18] for the exponential distribution with parameter $-\log p$:

$$\mathrm{Ent}\left[e^{f(Y)}\right] \leq \frac{2}{(\log p)(\log(p) + c)}\mathbb{E}\left[e^{f(Y)}|f'(Y)|^2\right].$$

1.3.3 Deviation inequality

In this section we prove a deviation inequality for functions of several variables under $\pi^{\otimes n}$ using the Herbst method and the above modified logarithmic Sobolev inequality.

Corollary 1.3.7. *Let $0 < c < -\log p$ and let f such that $|\mathrm{d}_i^+ f| \leq \beta$, $i = 1, \ldots, n$, and $\|\mathrm{d}^+ f\|^2 \leq \alpha^2$ for some $\alpha, \beta > 0$. Then for all $r > 0$,*

$$\pi^{\otimes n}(f - \mathbb{E}_{\pi^{\otimes n}}[f] \geq r) \leq \exp\left(-\min\left(\frac{c^2 r^2}{4a_{p,c}\alpha^2\beta^2}, \frac{rc}{\beta} - \alpha^2 a_{p,c}\right)\right), \tag{1.3.10}$$

where

$$a_{p,c} = \frac{pe^c}{(1-p)(1-\sqrt{pe^c})}$$

denotes the logarithmic Sobolev constant in (1.3.6).

Proof. Assume that $|\mathrm{d}_i^+ f| \leq c$, $i = 1, \ldots, n$. For $0 < t \leq 1$, let

$$H(t) = \frac{1}{t}\log \mathbb{E}_{\pi^{\otimes n}}[e^{tf}]$$

with $H(0^+) = \mathbb{E}_{\pi^{\otimes n}}[f]$. In order for $H(t)$ to be finite we may first assume that f is bounded, and then remove this assumption via a limiting argument once (1.3.10) is obtained. From (1.3.6) we have:

$$H'(t) = \frac{1}{t^2}\frac{\mathrm{Ent}_{\pi^{\otimes n}}[e^{tf}]}{\mathbb{E}_{\pi^{\otimes n}}[e^{tf}]} \leq \alpha^2 a_{p,c},$$

so that

$$H(t) \leq \mathbb{E}_{\pi^{\otimes n}}[f] + t\alpha^2 a_{p,c},$$

hence

$$\mathbb{E}_{\pi^{\otimes n}}[e^{tf}] \leq \exp\left(t\mathbb{E}_{\pi^{\otimes n}}[f] + t^2\alpha^2 a_{p,c}\right), \quad 0 < t \leq 1. \tag{1.3.11}$$

Finally, using Chebychev's inequality we obtain from (1.3.11):

$$\begin{aligned}
\pi^{\otimes n}(f - \mathbb{E}_{\pi^{\otimes n}}[f] \geq r) &\leq \inf_{t\in(0,1]} e^{-tr}\mathbb{E}_{\pi^{\otimes n}}[\exp(t(f - \mathbb{E}_{\pi^{\otimes n}}[f]))] \\
&\leq \exp\left(\inf_{t\in(0,1]} -tr + t^2\alpha^2 a_{p,c}\right) \\
&= \exp\left(-\min\left(\frac{r^2}{4\alpha^2 a_{p,c}}, r - \alpha^2 a_{p,c}\right)\right), \quad r > 0,
\end{aligned}$$

where we used the fact (see e.g. Corollary 2.11 in [61]) that the above minimum is attained at $t = \min(1, \frac{r}{2\alpha^2 a_{p,c}})$. Assume now that f satisfies $|\mathrm{d}_i^+ f| \leq \beta$, $i = 1, \ldots, n$, for some $\beta > 0$. Then cf/β satisfies the above hypothesis and we get

$$\pi^{\otimes n}(f - \mathbb{E}_{\pi^{\otimes n}}[f] \geq r) \leq \exp\left(-\min\left(\frac{c^2 r^2}{4a_{p,c}\alpha^2\beta^2}, \frac{rc}{\beta} - \alpha^2 a_{p,c}\right)\right).$$

■

Corollary 1.3.7 implies in particular $\mathbb{E}_\pi[e^{\alpha f}] < \infty$ for all $\alpha < c/\beta$ and $|\mathrm{d}^+ f| < c$. The condition $c < -\log p$ in Corollary 1.3.7 is necessary, since $f(k) = ck$ is not exponentially integrable under the geometric distribution π when $c \geq -\log p$. When $n = 1$, $\alpha = \beta$ and $r \geq 2c\beta a_{p,c}$, we have

$$\pi(f - \mathbb{E}_\pi[f] \geq r) \leq \exp\left(-\frac{rc}{\beta} + c^2 a_{p,c}\right) \leq \exp\left(-\frac{rc}{2\beta}\right), \tag{1.3.12}$$

and if $r \leq 2c\beta a_{p,c}$:

$$\pi(f - \mathbb{E}_\pi[f] \geq r) \leq \exp\left(-\frac{r^2}{4a_{p,c}\beta^2}\right).$$

These bounds can be compared to the result of [42] :

$$\pi(f - \mathbb{E}_\pi[f] \geq r) \leq \left(1 + (1-p)\frac{r}{\beta}\right) \exp\left(-\left(\frac{r}{\beta} + \frac{p}{1-p}\right) \log \frac{p + p(1-p)r/\beta}{p + (1-p)r/\beta}\right),$$

$r > 0$, and to the exact deviation

$$\pi(X - \mathbb{E}_\pi[X] \geq r) = \exp\left(\left(\left[r + \frac{1}{1-p}\right] - 1\right) \log p\right),$$

for X a geometric random variable with parameter p, where $[x]$ denotes the integral part of $x \in \mathbb{R}$. Applying the inequality (1.3.10) to $-f$, we obtain the concentration inequality

$$\pi^{\otimes n}(|f - \mathbb{E}_\pi[f]| \geq r) \leq 2\exp\left(-\min\left(\frac{c^2 r^2}{4a_{p,c}\alpha^2\beta^2}, \frac{rc}{\beta} - \alpha^2 a_{p,c}\right)\right). \tag{1.3.13}$$

Consider the negative binomial distribution ν with parameters $n \geq 1$ and $p \in (0,1)$, defined as

$$\nu(\{k\}) = \binom{n+k-1}{n-1}(1-p)^n p^k, \quad k \in \mathbb{N}.$$

Negative binomial random variables can be constructed as sums of n independent and identically distributed geometric variables with parameter p. Therefore, if we apply (1.3.10) to

$$f(k_1, \ldots, k_n) = \phi(k_1 + \cdots + k_n), \quad (k_1, \ldots, k_n) \in \mathbb{N}^n,$$

we obtain the modified logarithmic Sobolev inequality

$$\mathrm{Ent}_\nu\left[e^\phi\right] \leq n a_{p,c} \mathbb{E}_\nu\left[|\mathrm{d}^+\phi|^2 e^\phi\right],$$

and the deviation inequality

$$\nu(\phi - \mathbb{E}_\nu[\phi] \geq r) \leq \exp\left(-\min\left(\frac{r^2}{4na_{p,c}\beta^2}, \frac{rc}{\beta} - c^2 n a_{p,c}\right)\right),$$

where $\phi : \mathbb{N} \to \mathbb{R}$ satisfies $|d\phi| \leq \beta$, for the negative binomial distribution μ. Similar results can be obtained for the product of negative binomial laws with parameters

$n_1, \ldots, n_d$, namely by replacing $a_{p,c} = \frac{pe^c}{(1-p)(1-\sqrt{pe^c})}$ with $(n_1 + \cdots + n_d)a_{p,c}$ in (1.3.6) and (1.3.10).

Geometric and negative binomial random variables can be constructed as hitting times of the binomial process, thus they can be viewed as random variables on Bernoulli space. However, applying to them the Poincaré and logarithmic Sobolev inequalities on Bernoulli space, see e.g. [44], yields results that are weaker than the above inequalities.

1.4 The abstract case

In this section, we turn again to the general case of a probability space $(E, \mathcal{E}, \mu)$ with an absolutely continuous mapping $\tau : E \to E$. We show that modified logarithmic Sobolev and deviation inequalities hold for every measure μ on E which satisfies a Poincaré inequality

$$\lambda_\mu \mathrm{Var}_\mu[f] \leq \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right] \tag{1.4.1}$$

with respect to d^+, i.e. for every measure μ such that $\lambda_\mu > 0$. The application of the general results of this section to the geometric distribution using the spectral gap value (1.3.4) of λ_π allow to recover the results of Section 1.3.2. However, explicit calculations show that the results are recovered with worse constants for all $p \in (0, e^{-c})$, especially as p approaches e^{-c}.

1.4.1 Logarithmic Sobolev inequality and deviation inequality

Before turning to the main result of this section, we need the two following propositions whose proofs are adapted from [18], replacing the chain rule of derivation by the mean value theorem, and postponed to the end of this section. The next proposition is a generalization of Lemma 1.3.5.

Proposition 1.4.1. *Let $c > 0$. For any f on E such that $|\mathrm{d}^+ f| \leq c$ with $c^2 e^c \leq 4\lambda_\mu$ and $\mathbb{E}_\mu[f] = 0$,*

$$\mathbb{E}_\mu\left[f^2 e^f\right] \leq \alpha_{\mu,c} \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2 e^f\right], \tag{1.4.2}$$

where $\alpha_{\mu,c} = \frac{e^c\left((2+c)\sqrt{\lambda_\mu}+c\right)^2}{\lambda_\mu\left(2\sqrt{\lambda_\mu}-ce^{c/2}\right)^2}$.

The next statement is a modification of Proposition 3.4 in [18].

Proposition 1.4.2. *For any $f : E \to \mathbb{R}$ such that $\mathbb{E}_\mu[f] = 0$ and $|\mathrm{d}^+ f| \leq c$ we have*

$$\mathbb{E}_\mu\left[f^2 + \tau f^2\right] \leq e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \mathbb{E}_\mu\left[\left(f^2 + \tau f^2\right) e^{-|f|}\right]. \tag{1.4.3}$$

The following is a modified logarithmic Sobolev inequalities which holds whenever $\lambda_\mu > 0$.

Theorem 1.4.3. *Assume that $f : E \to \mathbb{R}$ satisfies $|\mathrm{d}^+ f| \leq c$ with $c^2 e^c \leq 4\lambda_\mu$,*

$$\mathrm{Ent}_\mu\left[e^f\right] \leq \frac{1}{2} e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \mathbb{E}_\mu\left[(\alpha_{\mu,c}|\mathrm{d}^+ f|^2 + 2e^{2c}\alpha_{\mu,c}|\mathrm{d}^+\tau f|^2 + 2e^{2c}\|\mathrm{d}^+ f\|^2_{L^2(\mu)})e^f\right]. \quad (1.4.4)$$

Proof. It suffices to suppose $\mathbb{E}_\mu[f] = 0$. By Taylor's formula,

$$\mathrm{Ent}_\mu\left[e^f\right] \leq \mathbb{E}_\mu\left[f e^f - e^f + 1\right] \leq \int_0^1 t\varphi(t)dt,$$

where $\varphi(t) = \mathbb{E}_\mu\left[\left(f^2 + \tau f^2\right) e^{tf}\right]$, $0 \leq t \leq 1$, is a convex function with $\varphi(t) \leq \max(\varphi(0), \varphi(1))$, $0 \leq t \leq 1$. By Proposition 1.4.2,

$$\varphi(0) \leq e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \varphi(1),$$

hence

$$\mathrm{Ent}_\mu\left[e^f\right] \leq \int_0^1 t e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \varphi(1) dt = \frac{1}{2} e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \mathbb{E}_\mu\left[\left(f^2 + \tau f^2\right) e^f\right].$$

Since $|\mathrm{d}^+\left(\tau f - \mathbb{E}_\mu[\tau f]\right)| = |\mathrm{d}^+ \tau f| \leq c$, Proposition 1.4.1 applied to $\tau f - \mathbb{E}_\mu[\tau f]$ implies:

$$\begin{aligned}
\mathbb{E}_\mu\left[\tau f^2 e^f\right] &\leq e^c \mathbb{E}_\mu\left[\tau f^2 e^{\tau f}\right] \\
&\leq 2e^{c+\mathbb{E}_\mu[\tau f]} \mathbb{E}_\mu\left[\left(\tau f - \mathbb{E}_\mu[\tau f]\right)^2 e^{\tau f - \mathbb{E}_\mu[\tau f]}\right] + 2e^c (\mathbb{E}_\mu[\tau f])^2 \mathbb{E}_\mu\left[e^{\tau f}\right] \\
&\leq 2e^c \alpha_{\mu,c} \mathbb{E}_\mu\left[|\mathrm{d}^+\tau f|^2 e^{\tau f}\right] + 2e^{2c} (\mathbb{E}_\mu[\tau f])^2 \mathbb{E}_\mu\left[e^f\right] \\
&= 2e^{2c} \alpha_{\mu,c} \mathbb{E}_\mu\left[|\mathrm{d}^+\tau f|^2 e^f\right] + 2e^{2c} (\mathbb{E}_\mu[\mathrm{d}^+ f])^2 \mathbb{E}_\mu\left[e^f\right].
\end{aligned}$$

Hence

$$\begin{aligned}
\mathrm{Ent}_\mu\left[e^f\right] &\leq \frac{1}{2} e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \mathbb{E}_\mu\left[\left(f^2 + \tau f^2\right) e^f\right] \\
&\leq \frac{1}{2} e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \mathbb{E}_\mu\left[(\alpha_{\mu,c}|\mathrm{d}^+ f|^2 + 2e^{2c}\alpha_{\mu,c}|\mathrm{d}^+\tau f|^2 + 2e^{2c}\mathbb{E}_\mu[|\mathrm{d}^+ f|^2])e^f\right].
\end{aligned}$$

■

We also have

$$\mathrm{Ent}_\mu\left[e^f\right] \leq \frac{1}{2} e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} (\alpha_{\mu,c} + 2e^{2c}\alpha_{\mu,c} + 2e^{2c}) |\mathrm{d}^+ f|_\infty \mathbb{E}_\mu[e^f], \qquad |\mathrm{d}^+ f| \leq c.$$

By tensorization, Theorem 1.4.3 implies in higher dimension

$$\begin{aligned}
\mathrm{Ent}_{\mu^{\otimes n}}\left[e^f\right] &\leq \frac{1}{2} e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \mathbb{E}_\mu\left[(\alpha_{\mu,c}\|\mathrm{d}^+ f\|^2 + 2e^{2c}\alpha_{\mu,c}\|\mathrm{d}^+\tau f\|^2 + 2e^{2c}\|\mathrm{d}^+ f\|^2_{L^2(E;\mathbb{R}^n)})e^f\right] \\
&\leq m_{\mu,c} \|\mathrm{d}^+ f\|^2_{L^\infty(E^n,\mathbb{R}^n)} \mathbb{E}_\mu\left[e^f\right],
\end{aligned}$$

where

$$m_{\mu,c} = \frac{1}{2} e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \left(\alpha_{\mu,c} + 2e^{2c}\alpha_{\mu,c} + 2e^{2c}\right)$$

and

$$\begin{aligned}
\mathrm{d}_i^+ f(x_1,\ldots,x_n) &= \tau_i f(x_1,\ldots,x_n) - f(x_1,\ldots,x_n),\\
&= f(x_1,\ldots,x_{i-1},\tau_i(x_i),x_{i+1},\ldots,x_n) - f(x_1,\ldots,x_n),
\end{aligned}$$

provided $|\mathrm{d}_i^+ f| \le c$, $i = 1,\ldots,n$. We obtain as a corollary as in Section 1.3.2 a deviation inequality for the product measure $\mu^{\otimes n}$ on E^n:

Corollary 1.4.4. *Assume that μ satisfies a Poincaré inequality (1.4.1). Let $c > 0$ such that $c^2 e^c \le 4\lambda_\mu$, and let f such that $|\mathrm{d}_i^+ f| \le \beta$, $i = 1,\ldots,d$, and $\|\mathrm{d}^+ f\|^2 \le \alpha^2$, for some $\alpha,\beta > 0$. Then for all $r > 0$,*

$$\mu^{\otimes n}\left(f - \mathbb{E}_\mu^{\otimes n}[f] \ge r\right) \le \exp\left(-\min\left(\frac{c^2 r^2}{4 m_{\mu,c}\alpha^2\beta^2}, \frac{rc}{\beta} - \alpha^2 m_{\mu,c}\right)\right). \tag{1.4.5}$$

Next we provide the proofs of Proposition 1.4.1 and Proposition 1.4.2.

Proof of Proposition 1.4.1. Set $a^2 = \mathbb{E}_\mu\left[f^2 e^f\right]$ and $b^2 = \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2 e^f\right]$. Since $\mathbb{E}_\mu[f] = 0$, the Poincaré inequality (1.4.1) implies

$$\begin{aligned}
\lambda_\mu^2 \mathbb{E}_\mu\left[f e^{f/2}\right]^2 &\le \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right]\mathbb{E}_\mu\left[|\mathrm{d}^+\left(e^{f/2}\right)|^2\right]\\
&\le \frac{1}{4}\mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right]\mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2 e^{f+|\mathrm{d}^+ f|}\right]\\
&\le \frac{1}{4} e^c\, c^2\, b^2.
\end{aligned} \tag{1.4.6}$$

Applying again the Poincaré inequality to $f e^{f/2}$ and using the mean value theorem we have

$$\begin{aligned}
\lambda_\mu \mathrm{Var}_\mu\left[f e^{f/2}\right] &\le \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\left(1 + \frac{|f| + |\mathrm{d}^+ f|}{2}\right)^2 e^{f+|\mathrm{d}^+ f|}\right]\\
&\le e^c\,\mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\left(1 + \frac{|f| + c}{2}\right)^2 e^f\right]\\
&\le \left(1 + \frac{c}{2}\right)^2 e^c b^2 + \frac{c^2 e^c a^2}{4} + \left(1 + \frac{c}{2}\right) e^c \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2 |f| e^f\right]\\
&\le \left(1 + \frac{c}{2}\right)^2 e^c b^2 + \frac{c^2 e^c a^2}{4} + \left(1 + \frac{c}{2}\right) e^c abc\\
&\le e^c\left(\left(1 + \frac{c}{2}\right) b + \frac{ac}{2}\right)^2.
\end{aligned}$$

Hence

$$a^2 = \mathbb{E}_\mu\left[f e^{f/2}\right]^2 + \mathrm{Var}_\mu\left[f e^{f/2}\right] \le \frac{e^c\, c^2\, b^2}{4\lambda_\mu^2} + \frac{e^c}{\lambda_\mu}\left(\left(1 + \frac{c}{2}\right) b + \frac{ac}{2}\right)^2,$$

which leads to

$$a \le \frac{e^{c/2}\left((2+c)\sqrt{\lambda_\mu} + c\right)}{\sqrt{\lambda_\mu}\left(2\sqrt{\lambda_\mu} - c e^{c/2}\right)},$$

from which the conclusion follows. □

With $\lambda_\pi = (1-\sqrt{p})^2/p$, the condition $c^2e^c \le 4\lambda_\pi$ implies $c < -\log p$, $p \in (0,1)$, hence Theorem 1.4.3 and Corollary 1.4.4 are weaker than Theorem 1.3.6 and Corollary 1.3.7 respectively, when $\mu = \pi$ is the geometric distribution.

Proof of Proposition 1.4.2. We have from the Poincaré inequality (1.4.1):

$$\begin{aligned}
\lambda_\mu \mathbb{E}_\mu\left[f^4\right] &= \lambda_\mu \mathrm{Var}_\mu\left[f^2\right] + \lambda_\mu(\mathbb{E}_\mu\left[f^2\right])^2 \\
&\le \mathbb{E}_\mu\left[|\mathrm{d}^+ f^2|^2\right] + \lambda_\mu(\mathbb{E}_\mu\left[f^2\right])^2 \\
&= \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2 \left(f + \tau f\right)^2\right] + \lambda_\mu(\mathbb{E}_\mu\left[f^2\right])^2 \\
&\le 2c^2\mathbb{E}_\mu\left[f^2 + \tau f^2\right] + \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right]\mathbb{E}_\mu\left[f^2\right] \\
&\le 3c^2\mathbb{E}_\mu\left[f^2\right] + 2c^2\mathbb{E}_\mu\left[\tau f^2\right].
\end{aligned}$$

Hence for all $u > 0$,

$$\begin{aligned}
\mathbb{E}_\mu\left[|f|^3\right] &\le \frac{u}{2}\mathbb{E}_\mu\left[f^2\right] + \frac{1}{2u}\mathbb{E}_\mu\left[f^4\right] \\
&\le c_1\mathbb{E}_\mu\left[f^2\right] + c_2\mathbb{E}_\mu\left[\tau f^2\right] \qquad (1.4.7)
\end{aligned}$$

with $c_1 = \frac{3c^2}{2u\lambda_\mu} + \frac{u}{2}$ and $c_2 = \frac{c^2}{u\lambda_\mu}$. Let us consider the probability measure

$$d\rho = \frac{c_1 f^2 + c_2 \tau f^2}{c_1\mathbb{E}_\mu\left[f^2\right] + c_2\mathbb{E}_\mu\left[\tau f^2\right]} d\mu.$$

By Jensen's inequality,

$$\begin{aligned}
\mathbb{E}_\mu\left[\left(c_1 f^2 + c_2\tau f^2\right)e^{-|f|}\right] &= \mathbb{E}_\rho\left[e^{-|f|}\right]\mathbb{E}_\mu\left[c_1 f^2 + c_2\tau f^2\right] \\
&\ge e^{-\mathbb{E}_\rho[|f|]}\mathbb{E}_\mu\left[c_1 f^2 + c_2\tau f^2\right].
\end{aligned}$$

From the inequality $ab^2 \le a^3 + |b-a|(a^2+b^2)$, $a, b \ge 0$, we have

$$|f|\tau f^2 \le |f|^3 + c\left(f^2 + \tau f^2\right),$$

hence

$$\begin{aligned}
\mathbb{E}_\rho\left[|f|\right]\mathbb{E}_\mu\left[c_1 f^2 + c_2\tau f^2\right] &= \mathbb{E}_\mu\left[c_1|f|^3 + c_2|f|\tau f^2\right] \\
&\le \mathbb{E}_\mu\left[(c_1+c_2)|f|^3 + c_2 c\left(f^2 + \tau f^2\right)\right] \\
&\le (c_1+c_2)\mathbb{E}_\mu\left[c_1 f^2 + c_2\tau f^2\right] + c_2 c\mathbb{E}_\mu\left[f^2 + \tau f^2\right] \\
&\le (c_1+c_2+c)\mathbb{E}_\mu\left[c_1 f^2 + c_2\tau f^2\right]
\end{aligned}$$

where we used the fact that $c_2 \le c_1$. Therefore,

$$\mathbb{E}_\rho\left[|f|\right] \le c_1 + c_2 + c = \frac{5c^2 + u^2\lambda_\mu}{2u\lambda_\mu} + c.$$

Optimizing in u we obtain for $u = c\sqrt{\frac{5}{\lambda_\mu}}$:

$$\mathbb{E}_\rho\left[|f|\right] \le c\left(1 + \sqrt{\frac{5}{\lambda_\mu}}\right).$$

$\square$

As in [18] and references therein, we can obtain the following bound.

Proposition 1.4.5. *Let A, B be disjoint subsets of E. We have*

$$\mu(A)\mu(B) \le 3\exp(-\sqrt{\lambda_\mu}e^{-\gamma_1/2}d(A,B)), \tag{1.4.8}$$

with $\gamma_1^2 e^{\gamma_1} = 2\lambda_\mu$.

Proof. From the Poincaré inequality on $(E^2, \mu^{\otimes 2})$ we have:

$$\lambda_\mu \mathbb{E}_{\mu^{\otimes 2}}[f^2] \le \mathbb{E}_{\mu^{\otimes 2}}[|\mathrm{d}_1^+ f|^2 + |\mathrm{d}_2^+ f|^2],$$

provided $\mathbb{E}_{\mu^{\otimes 2}}[f] = 0$. Applying this inequality to $f(x,y) = \sinh(tg(x,y)/2)$, $0 \le t < \gamma_1$ with $g(x,y) = h(x) - h(y)$ and $|\mathrm{d}h| \le 1$, and using the bound

$$|\sinh(x+y) - \sinh(x)| \le |y|e^{|y|}\cosh x, \quad x, y \in \mathbb{R},$$

we have:

$$\begin{aligned}
\lambda_\mu \mathbb{E}_{\mu^{\otimes 2}}[\sinh^2(tg/2)] &\le \frac{t^2}{4}\mathbb{E}_{\mu^{\otimes 2}}\left[\left(|\mathrm{d}_1^+ g|^2 e^{t|\mathrm{d}_1^+ g|} + |\mathrm{d}_2^+ g|^2 e^{t|\mathrm{d}_2^+ g|}\right)\cosh^2(tg/2)\right] \\
&\le \frac{t^2}{2}e^{\gamma_1}\mathbb{E}_{\mu^{\otimes 2}}\left[\cosh^2(tg/2)\right].
\end{aligned}$$

Hence

$$\mathbb{E}_{\mu^{\otimes 2}}[\cosh^2(tg/2)] = \frac{1}{2}\left(\mathbb{E}_{\mu^{\otimes 2}}[e^{tg}] + 1\right) \le \frac{2\lambda_\mu}{2\lambda_\mu - t^2 e^{\gamma_1}}.$$

and for all $t < \gamma_1$, if $h(x) = d(x, B)$ then

$$e^{td(A,B)}\mu(A)\mu(B) \le \mathbb{E}_{\mu^{\otimes 2}}[e^{tg}] \le \frac{2\lambda_\mu + t^2 e^{\gamma_1}}{2\lambda_\mu - t^2 e^{\gamma_1}}.$$

and it remains to take $t = \sqrt{\lambda_\mu}e^{-\gamma_1/2}$. ∎

1.4.2 Exponential integrability

The Herbst method used in the preceding sections relies on exponential integrability. Following [18], we obtain a bound of the Laplace transform with respect to any measure μ on E, provided it follows a Poincaré inequality (1.4.1).

Proposition 1.4.6. *Let $f : E \to \mathbb{R}$ such that $\mathbb{E}_\mu[f] = 0$, with $|\mathrm{d}^+ f| \le \beta$ for some $\beta > 0$, and let c such that $c^2 e^c \le 4\lambda_\mu$. Then, for every $0 \le t < c/\beta$ we have*

$$\mathbb{E}_\mu[e^{tf}] \le \frac{2\sqrt{\lambda_\mu} + t\beta e^{c/2}}{2\sqrt{\lambda_\mu} - t\beta e^{c/2}}. \tag{1.4.9}$$

Proof. We adapt the proof of Proposition 4.1 in [18]. It is sufficient to assume $\beta = 1$. We have

$$\begin{aligned}
|\mathrm{d}^+ e^{\frac{t}{2}f}(x)| &= |e^{\frac{t}{2}\tau f(x)} - e^{\frac{t}{2}f(x)}| \\
&= \frac{t}{2}\left|\int_{f(x)}^{f(\tau(x))} e^{\frac{t}{2}} dt\right| \\
&\le \frac{t}{2} e^{\frac{t}{2}(f(x)+|\mathrm{d}^+ f(x)|)} |\mathrm{d}^+ f(x)| \\
&\le \frac{t}{2} e^{\frac{c}{2}+\frac{t}{2}f(x)} |\mathrm{d}^+ f(x)|, \qquad x \in E,
\end{aligned}$$

and applying (1.4.1) to $e^{\frac{t}{2}f}$ we get, with $u(t) = \mathbb{E}_\mu[e^{tf}]$:

$$\lambda_\mu \left(u(t) - u(t/2)^2\right) \le e^c \frac{t^2}{4} u(t),$$

i.e.

$$u(t) \le \frac{4\lambda_\mu}{4\lambda_\mu - t^2 e^c} u(t/2)^2.$$

Applying the same inequality for $t/2$ and iterating, we have

$$u(t) \le \prod_{k=0}^{\infty} \left(\frac{4\lambda_\mu}{4\lambda_\mu - e^c t^2/4^k}\right)^{2^k} \le \frac{4\lambda_\mu}{4\lambda_\mu - e^c t^2} V(t),$$

with

$$V(t) = \prod_{k=1}^{\infty} \left(\frac{4\lambda_\mu}{4\lambda_\mu - e^c t^2/4^k}\right)^{2^k},$$

where the product converges whenever $t < c$. It can be shown as in [18] that $\sqrt{V}$ is convex. Moreover $V(0) = 1$ and $V\left(\frac{2\sqrt{\lambda_\mu}}{e^{c/2}}\right) \le 4$, hence

$$\sqrt{V(t)} \le \frac{2\sqrt{\lambda_\mu} + te^{c/2}}{2\sqrt{\lambda_\mu}}.$$

■

It is easily checked that the assumption of Corollary 1.4.4 is consistent with that of Proposition 1.4.6.

Chapitre 2

A logarithmic Sobolev inequality for an interacting spin system under a geometric reference measure

Ce chapitre fait l'objet d'une courte note, écrite en collaboration avec Nicolas Privault, à paraître dans les actes de la conférence *Quantum Probability and Infinite Dimensional Analysis*, Levico (Italie), février 2005.

Abstract

Logarithmic Sobolev inequalities are an essential tool in the study of interacting particle systems, cf. e.g. [65], [85]. In this note we show that the logarithmic Sobolev inequality proved on the configuration space $\mathbb{N}^{\mathbb{Z}^d}$ under Poisson reference measures in [31] can be extended to geometric reference measures using the results of [55]. As a corollary we obtain a deviation estimate for an interacting particle system.

2.1 Logarithmic Sobolev inequality for the geometric distribution

Consider the forward and backward gradient operators

$$\mathrm{d}^+ f(k) = f(k+1) - f(k), \qquad \mathrm{d}^- f(k) = 1_{\{k\geq 1\}}(f(k-1) - f(k)), \quad k \in \mathbb{N},$$

and the Laplacian

$$\mathcal{L} = -\mathrm{d}_\pi^{+*}\mathrm{d}^+ = \mathrm{d}^+ + \frac{1}{p}\mathrm{d}^-$$

which generates a Markov process on $\mathbb{N}$ whose invariant measure is the geometric distribution π on $\mathbb{N}$ with parameter $p \in (0,1)$, i.e.

$$\pi(\{k\}) = (1-p)p^k, \qquad k \in \mathbb{N}.$$

Denote by $\mathbb{E}_\pi$ the expectation under π and by Ent_π the entropy under π, defined as

$$\mathrm{Ent}_\pi[f] = \mathbb{E}_\pi[f \log f] - \mathbb{E}_\pi[f] \log \mathbb{E}_\pi[f].$$

We recall the modified logarithmic Sobolev inequality proved in [55] for the geometric distribution π.

Theorem 2.1.1. *Let $0 < c < -\log p$ and let $f : \mathbb{N} \to \mathbb{R}$ such that $|\mathrm{d}^+ f| \leq c$. We have*

$$\mathrm{Ent}_\pi\left[e^f\right] \leq \frac{pe^c}{(1-p)(1-\sqrt{pe^c})} \mathbb{E}_\pi\left[|\mathrm{d}^+ f|^2 e^f\right]. \tag{2.1.1}$$

In higher dimensions the multi-dimensional gradient is defined as

$$\mathrm{d}_i^+ f(k) = f(k+e_i) - f(k), \quad i = 1, \ldots, n,$$

where f is a function on $\mathbb{N}^n$, $k = (k_1, \ldots, k_n) \in \mathbb{N}^n$, $(e_1, \ldots, e_n)$ is the canonical basis of $\mathbb{R}^n$, and the gradient norm is

$$\|\mathrm{d}^+ f(k)\|^2 = \sum_{i=1}^n |\mathrm{d}_i^+ f(k)|^2 = \sum_{i=1}^n |f(k+e_i) - f(k)|^2. \tag{2.1.2}$$

From the tensorization property of entropy, (2.1.1) still holds with respect to $\pi^{\otimes n}$ in any finite dimension n:

$$\mathrm{Ent}_{\pi^{\otimes n}}\left[e^f\right] \leq \frac{pe^c}{(1-p)(1-\sqrt{pe^c})} \mathbb{E}_{\pi^{\otimes n}}\left[\|\mathrm{d}^+ f\|^2 e^f\right], \tag{2.1.3}$$

provided $|\mathrm{d}_i^+ f| \leq c$, $i = 1, \ldots, n$. As a consequence the following deviation inequality for functions of several variables under $\pi^{\otimes n}$ has been proved in [55] using (2.1.1) and the Herbst method.

Corollary 2.1.2. *Let $0 < c < -\log p$ and let f such that $|\mathrm{d}_i^+ f| \leq \beta$, $i = 1, \ldots, n$, and $\|\mathrm{d}^+ f\|^2 \leq \alpha^2$ for some $\alpha, \beta > 0$. Then for all $r > 0$,*

$$\pi^{\otimes n}(f - \mathbb{E}_{\pi^{\otimes n}}[f] \geq r) \leq \exp\left(-\min\left(\frac{c^2 r^2}{4a_{p,c}\alpha^2\beta^2}, \frac{rc}{\beta} - \alpha^2 a_{p,c}\right)\right), \tag{2.1.4}$$

where

$$a_{p,c} = \frac{pe^c}{(1-p)(1-\sqrt{pe^c})}$$

denotes the logarithmic Sobolev constant in (2.1.1).

Our goal in the next section will be to extend these results to interacting spin systems under a geometric reference measure.

2.2 Logarithmic Sobolev inequality for an interacting spin system

Given a bounded finite range interaction potential $\Phi = \{\Phi_R \ : \ R \subset \mathbb{Z}^d\}$, i.e.

$$\|\Phi\| = \sup_{k\in\mathbb{Z}^d} \sum_{R\ni k} \|\Phi_R\|_\infty < \infty,$$

let the Hamiltonian H_Λ be defined as

$$H_\Lambda(\eta) = \sum_{R\bigcap\Lambda\neq\emptyset} \Phi_R(\eta_R),$$

where η_R denotes the restriction of η to $\mathbb{N}^R$, $R \subset \mathbb{Z}^d$. The Gibbs measure π_Λ^ω on $\mathbb{N}^\Lambda$ associated to a $\mathbb{N}$-valued spin system on a finite lattice $\Lambda \subset \mathbb{Z}^d$ with boundary condition $\omega \in \mathbb{N}^{\mathbb{Z}^d\setminus\Lambda}$ is defined by its density with respect to $\pi_\Lambda := \pi^{\otimes\Lambda}$ as:

$$\frac{d\pi_\Lambda^\omega}{d\pi_\Lambda}(\sigma) = \frac{1}{Z_\Lambda^\omega} e^{-H_\Lambda^\omega(\sigma)}, \quad \sigma \in \mathbb{N}^\Lambda,$$

where π is the geometric reference distribution on $\mathbb{N}$, Z_Λ^ω is a normalization factor, and

$$H_\Lambda^\omega(\eta) = H_\Lambda(\eta_\Lambda\omega_{\Lambda^c}), \quad \eta \in \mathbb{N}^{\mathbb{Z}^d},$$

where $\eta_A\omega_B$ is defined as

$$(\eta_A\omega_B)_k = \eta_k 1_A(k) + \omega_k 1_B(k), \qquad k \in \mathbb{Z}^d,$$

whenever $\eta \in \mathbb{N}^A$, $\omega \in \mathbb{N}^B$, and $A, B \subset \mathbb{Z}^d$ are such that $A \cap B = \emptyset$. Let again

$$\mathrm{d}_k^+ f(\eta) = f(\eta + e_k) - f(\eta), \quad \text{and} \quad \mathrm{d}_k^- f(\eta) = 1_{\{\eta_k>0\}} \left(f(\eta - e_k) - f(\eta) \right),$$

for every function $f : \mathbb{N}^{\mathbb{Z}^d} \to \mathbb{R}$, where $(e_k)_{k\in\mathbb{Z}^d}$, denotes the canonical basis $\{e_k = 1_{\{k\}} \ : \ k \in \mathbb{Z}^d\}$. Consider the Markov generator

$$\mathcal{L}_\Lambda^\omega f(\eta) = \sum_{k\in\Lambda} c_\Lambda^\omega(k,\eta,+)\mathrm{d}_k^+ f(\eta) + c_\Lambda^\omega(k,\eta,-)\mathrm{d}_k^- f(\eta),$$

where $c_\Lambda^\omega(k,\eta,\pm)$ are rate functions such that $\mathcal{L}_\Lambda^\omega$ is self-adjoint in $L^2(\pi_\Lambda^\omega)$, i.e.

$$c_\Lambda^\omega(k,\eta,+)\pi_\Lambda^\omega(\{\eta\}) = c_\Lambda^\omega(k,\eta+e_k,-)\pi_\Lambda^\omega(\{\eta+e_k\}), \quad k \in \Lambda, \ \eta \in \mathbb{N}^\Lambda,$$

$$c_\Lambda^\omega(k,\eta,-)\pi_\Lambda^\omega(\{\eta\}) = c_\Lambda^\omega(k,\eta-e_k,+)\pi_\Lambda^\omega(\{\eta-e_k\}), \quad k \in \Lambda, \ \eta \in \mathbb{N}^\Lambda.$$

We assume that there exists a constant $C > 0$ depending on $\|\Phi\|$ only, with

$$\frac{1}{C} \le c_\Lambda^\omega(k,\eta,+) \le C, \qquad \eta \in \mathbb{N}^\Lambda, \ \Lambda \subset \mathbb{Z}^d, \ k \in \Lambda. \tag{2.2.1}$$

For $f : \mathbb{N}^{\mathbb{Z}^d} \to \mathbb{R}$ we let:

$$\mathscr{E}^{\omega}_{\Lambda}(e^f) = \sum_{k\in\Lambda} \int_{\mathbb{N}^{\mathbb{Z}^d}} c^{\omega}_{\Lambda}(k,\sigma,+) e^{f(\sigma)} |\mathrm{d}^+_k f(\sigma)|^2 d\pi^{\omega}_{\Lambda}(\sigma),$$

and

$$\mathscr{E}_{\Lambda}(e^f) = \sum_{k\in\Lambda} \int_{\mathbb{N}^{\mathbb{Z}^d}} e^{f(\sigma)} |\mathrm{d}^+_k f(\sigma)|^2 d\pi_{\Lambda}(\sigma).$$

Next we consider the family of rectangles of the form

$$R = R(k, l_1, ..., l_d) = k + ([1, l_1] \times \cdots \times [1, l_d]) \cap \mathbb{Z}^d,$$

where $k \in \mathbb{Z}^d$ and $l_1, \ldots, l_d \in \mathbb{N}$, with

$$\mathrm{size}(R) = \max_{k=1,\ldots,d} l_k.$$

Let $\mathscr{R}_L$ denote the set of rectangles such that

$$\mathrm{size}(R) \leq L \quad \text{and} \quad \mathrm{size}(R) \leq 10 \min_{k=1,\ldots,d} l_k.$$

Definition 2.2.1. *We say that π^{ω}_{Λ} satisfies the mixing condition if there exists constants C_1 and C_2, depending on d and $\|\Phi\|$ only, such that:*

$$\sup_{\sigma,\omega} \left| \frac{\pi^{\omega}_{\Lambda}(\{\eta :\ \eta_A = \sigma_A\})\pi^{\omega}_{\Lambda}(\{\eta :\ \eta_B = \sigma_B\})}{\pi^{\omega}_{\Lambda}(\{\eta\ :\ \eta_{A\cup B} = \sigma_{A\cup B}\})} - 1 \right| \leq C_1 e^{-C_2 d(A,B)}, \tag{2.2.2}$$

for all $L \geq 1$, $\Lambda \in \mathscr{R}_L$ and $A, B \subset \Lambda$ such that $A, B \in \mathscr{R}_L$ with $A \cap B = \emptyset$.

We refer to [31] and [65] for conditions on Φ under which (2.2.2) holds under a geometric reference measure.

Our goal is to prove the following logarithmic Sobolev inequality under the Gibbs measure π^{ω}_{Λ}.

Theorem 2.2.2. *Assume that the mixing condition (2.2.2) holds, and let $0 < c < -\log p$. Then there exists a constant $\gamma_c > 0$, independent of Λ and ω, such that*

$$\mathrm{Ent}_{\pi^{\omega}_{\Lambda}}\left[e^f\right] \leq \gamma_c \mathscr{E}^{\omega}_{\Lambda}(e^f), \tag{2.2.3}$$

for every $f : \mathbb{N}^{\mathbb{Z}^d} \to \mathbb{R}$ such that $\|\mathrm{d}^+ f\|_{l^\infty(\Lambda)} \leq c$, π_{Λ}-a.e.

In particular we have

$$\mathrm{Ent}_{\pi^{\omega}_{\Lambda}}\left[e^f\right] \leq \gamma_c \left\| \sum_{k\in\Lambda} c^{\omega}_{\Lambda}(k,\cdot,+)|\mathrm{d}^+_k f(\cdot)|^2 \right\|_{L^\infty(\pi_\Lambda)} \times \mathbb{E}_{\pi^{\omega}_{\Lambda}}\left[e^f\right],$$

which implies, as in Corollary 2.1.2, a deviation inequality under Gibbs measures.

Corollary 2.2.3. *Assume that the mixing condition* (2.2.2) *holds, and let* $0 < c < -\log p$. *Let* f *be such that* $\|\mathrm{d}^+ f\|_{l^\infty(\Lambda)} \leq \beta$ *and*

$$\sum_{k\in\Lambda} c_\Lambda^\omega(k,\eta,+)|\mathrm{d}_k^+ f(\eta)|^2 \leq \alpha^2, \quad \pi_\Lambda(d\eta) - a.e., \tag{2.2.4}$$

for some $\alpha, \beta > 0$. *Then for all* $r > 0$,

$$\pi_\Lambda^\omega\left(f - \mathbb{E}_{\pi_\Lambda^\omega}[f] \geq r\right) \leq \exp\left(-\min\left(\frac{c^2 r^2}{4\gamma_c \alpha^2 \beta^2}, \frac{rc}{\beta} - \alpha^2\gamma_c\right)\right). \tag{2.2.5}$$

Due to Hypothesis (2.2.1), condition (2.2.4) can be replaced by

$$\|\mathrm{d}^+ f(\eta)\|^2_{l^2(\Lambda)} \leq C^{-1}\alpha^2, \quad \pi_\Lambda(d\eta) - a.e.$$

Denoting by Π denote the infinite volume Gibbs measure associated to π_Λ^ω, for some $r_0 > 0$ we get the Ruelle type bound:

$$\Pi(\{\eta \in \mathbb{N}^{\mathbb{Z}^d} : |\eta_\Lambda| \geq r|\Lambda|\}) \leq \exp\left(-(cr - C\gamma_c)|\Lambda| + c\mathbb{E}_\Pi[|\eta_\Lambda|]\right), \quad r > r_0,$$

for all finite subset Λ of $\mathbb{Z}^d$, under the mixing condition (2.2.2). Indeed, it suffices to apply the uniform bound (2.2.5) with $f(\eta) = |\eta_\Lambda|$, $\alpha^2 = C|\Lambda|$, $\beta = 1$, and the compatibility condition

$$\Pi(E) = \int_{\mathbb{N}^{\mathbb{Z}^d}} \pi_\Lambda^\omega(E_\Lambda)\Pi(d\omega),$$

to $E = \{\eta \in \mathbb{N}^{\mathbb{Z}^d} : |\eta_\Lambda| \geq r|\Lambda|\}$, with

$$E_\Lambda := \{\eta \in \mathbb{N}^\Lambda : \eta_\Lambda\omega_{\Lambda^c} \in E\} = \{\eta \in \mathbb{N}^\Lambda : |\eta_\Lambda| \geq r|\Lambda|\}.$$

This shows in particular that Π satisfies the $(RPB)^1$ condition in [60].

2.3 Proof of Theorem 2.2.2

Recall that for $0 < c < -\log p$, by tensorization, Theorem 2.1.1 yields as in (2.1.3) the logarithmic Sobolev inequality

$$\mathrm{Ent}_{\pi_\Lambda}[e^f] \leq s_c \mathscr{E}_\Lambda(e^f), \tag{2.3.1}$$

for all $f : \mathbb{N}^{\mathbb{Z}^d} \to \mathbb{R}$ such that $\|\mathrm{d}^+ f\|_{l^\infty(\Lambda)} \leq c$, π_Λ-a.e., with an optimal constant $s_c \leq a_{p,c}$ which is independent of $\Lambda \subset \mathbb{Z}^d$. Let now $s_{\Lambda,\omega,c}$ denote the optimal constant in the inequality

$$\mathrm{Ent}_{\pi_\Lambda^\omega}[e^f] \leq s_{\Lambda,\omega,c}\mathscr{E}_\Lambda^\omega(e^f), \qquad \|\mathrm{d}^+ f\|_{l^\infty(\Lambda)} \leq c.$$

Lemma 2.3.1. *For every* $\Lambda \subset \mathbb{Z}^d$, *there exists a constant* $A := Ce^{4|\Lambda|\|\Phi\|} > 0$ *depending only on* $|\Lambda|$, c *and independent of* $\omega \in \mathbb{N}^{\mathbb{Z}^d\setminus\Lambda}$, *such that*

$$\frac{s_c}{A} \leq s_{\Lambda,\omega,c} \leq A\, s_c.$$

Proof. We follow the proof of Proposition 3.1 in [31]. From (2.2.1) we obtain:

$$C^{-1}e^{-2|\Lambda|\|\Phi\|}\,\mathscr{E}_\Lambda(e^f) \le \mathscr{E}^\omega_\Lambda(e^f) \le Ce^{2|\Lambda|\|\Phi\|}\,\mathscr{E}_\Lambda(e^f). \tag{2.3.2}$$

From the relation

$$\mathrm{Ent}_\mu[f] = \min_{t>0}\mathbb{E}_\mu[f\log f - f\log t - f + t]$$

and the bound

$$e^{-2|\Lambda|\|\Phi\|} \le \frac{d\pi^\omega_\Lambda}{d\pi_\Lambda} \le e^{2|\Lambda|\|\Phi\|},$$

we have

$$e^{-2|\Lambda|\|\Phi\|}\mathrm{Ent}_{\pi_\Lambda}\left[e^f\right] \le \mathrm{Ent}_{\pi^\omega_\Lambda}\left[e^f\right] \le e^{2|\Lambda|\|\Phi\|}\mathrm{Ent}_{\pi_\Lambda}\left[e^f\right],$$

from which the conclusion follows using (2.3.1) and (2.3.2). ∎

Let for $L \ge 1$:

$$S_{L,c} := \sup_{R\in\mathscr{R}_L}\sup_{\omega\in E} s_{R,\omega,c} \le Cs_c e^{4|\Lambda|\|\Phi\|} < \infty,$$

which is finite by Lemma 2.3.1.

Proposition 2.3.2. *Assume the mixing condition* (2.2.2) *is satisfied. Then there exists a constant* κ *depending on* $\|\Phi\|$*, such that*

$$S_{2L,c} \le \left(1 - \frac{\kappa}{\sqrt{L}}\right)^{-1} S_{L,c} \tag{2.3.3}$$

for L *large enough.*

Proof. The proof of this proposition is identical to that of Proposition 4.1, pp. 1970-1972 and Proposition 5.1, p. 1975 in [31], replacing the Dirichlet form used in [31] with $\mathscr{E}^\omega_\Lambda$. ∎

Finally, Theorem 2.2.2 is proved by taking $\gamma_c = \sup_L S_{L,c}$, which is finite from Proposition 2.3.2.

Chapitre 3

Poisson-type deviation inequalities for curved continuous time Markov chains

Ce chapitre fait l'objet d'un article soumis pour publication.

Abstract

In this paper, we present new local Poisson-type deviation inequalities for continuous time Markov chains whose Wasserstein curvature or Γ-curvature is bounded below. Although these two curvatures are equivalent for Brownian motion on Riemannian manifolds, they are not comparable in discrete settings and yield deviation bounds involving different Lipschitz seminorms. In the case of birth-death process, we provide some conditions on the rates of the associated generator for such discrete curvatures to be bounded below, and we extend to this framework the local deviation inequalities of [3] established for continuous time random walks on graphs, seen as models in null curvature. By a limiting argument, deviation bounds are derived for the stationary distribution of birth-death process in the finite state space case and we recover the optimal Gaussian deviation for Ornstein-Uhlenbeck processes constructed as fluid limits of rescaled continuous time Ehrenfest chains. Finally, an extension of these local deviation inequalities to sample vectors of the $M/M/1$ queueing process completes this work.

3.1 Introduction

In recent years, the area of concentration of measure has been deeply investigated in the context of discrete time Markov chains, using mass transportation and functional inequalities related to the convergence to stationarity. For instance, in the contracting case, Gaussian concentration of measure was put forward by K. Marton [66], via Pinsker-type inequalities derived from information theory. It has been then extended by P.M. Samson [75] to a large class of Markov chains, among them Doeblin recurrent Markov chains, whereas H. Djellout, A. Guillin and L. Wu [35], and lately G. Blower and F. Bolley [13], established similar deviation bounds under assumptions of transportation inequalities. On the other hand, C. Houdré and P. Tetali [45] in the case of reversible Markov chains, and C. Ané and M. Ledoux [3] for continuous time random walks on graphs corresponding to null curvature models, obtained Poisson-type tail estimates using modified logarithmic Sobolev inequalities and the Herbst method.

The purpose of the present paper is to give new local Poisson-type deviation bounds for continuous time Markov chains, which extend and sharpen in the case of curved birth-death processes the tail inequalities of [3] mentioned above. Our approach is based on semigroup analysis and uses the notion of curvature for Markov processes on general metric measure spaces recently investigated in [79], in the context of continuous time Markov chains: the Wasserstein curvature involving the Lipschitz seminorm of the Markov semigroup, and the Γ-curvature related to a commutation relation between the semigroup and the "carré du champ" operator Γ.

In the case of Brownian motion on smooth Riemannian manifolds, Theorem 2 together with Corollary 1 in [79] state that the following assertions are equivalent for any $K \in \mathbb{R}$:

(i) the Brownian Wasserstein curvature is bounded below by K,

(ii) the Brownian Γ-curvature is bounded below by K,

(iii) the Ricci curvature of the manifold is bounded below by K.

Therefore, such an equivalence gives a characterization of uniform lower bounds of the Ricci curvature of the manifold in terms of gradient estimates of heat kernels. However, the equivalence between (i) and (ii) does not hold in the framework of continuous time Markov chains since discrete gradients do not satisfy the chain rule formula. Thus, it is natural to study the role played by each type of discrete curvature in the concentration of measure phenomenon. Actually, the constants in the deviation inequalities we establish in this paper turn out to be different when one or the other discrete curvature above is bounded below. For instance, let $(X_t)_{t\geq 0}$ be a regular continuous time Markov chain on a discrete metric space E, with jumps bounded by some positive b. Let $f : E \to \mathbb{R}$ be a Lipschitz function and denote $g(u) = (1+u)\log(1+u) - u$, $u > 0$. If $(X_t)_{t\geq 0}$ has Wasserstein curvature bounded below by $\rho > 0$ and angle bracket bounded by some positive V^2, we show via Theorem 3.3.1 the tail probability:

$$\begin{aligned}\sup_{x\in E} \mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right) &\leq \exp\left(-\frac{(1-e^{-2\rho t})V^2}{2\rho b^2} g\left(\frac{2\rho b y}{(1-e^{-2\rho t})V^2\|f\|_{\mathrm{Lip}}}\right)\right) \\ &\leq \exp\left(-\frac{y}{2b\|f\|_{\mathrm{Lip}}}\log\left(1+\frac{2\rho b y}{(1-e^{-2\rho t})V^2\|f\|_{\mathrm{Lip}}}\right)\right),\end{aligned}$$

whereas if the Γ-curvature is bounded below by the same ρ and if $\|\Gamma f\|_\infty < +\infty$, we get the estimate:

$$\begin{aligned}\sup_{x\in E}\mathbb{P}_x\left(f(X_t)-\mathbb{E}_x\left[f(X_t)\right]\geq y\right) &\leq \exp\left(-\frac{(1-e^{-2\rho t})\|\Gamma f\|_\infty}{\rho b^2\|f\|_{\mathrm{Lip}}^2}g\left(\frac{\rho b y\|f\|_{\mathrm{Lip}}}{(1-e^{-2\rho t})\|\Gamma f\|_\infty}\right)\right)\\ &\leq \exp\left(-\frac{y}{2b\|f\|_{\mathrm{Lip}}}\log\left(1+\frac{\rho b y\|f\|_{\mathrm{Lip}}}{(1-e^{-2\rho t})\|\Gamma f\|_\infty}\right)\right),\end{aligned}$$

cf. Corollary 3.4.4. Although the exponential decays above are somewhat similar, we note that a lower bound on the Γ-curvature entails more general inequalities involving the mixed Lipschitz seminorms $\|\cdot\|_{\mathrm{Lip}}$ and $f\mapsto\|\Gamma f\|_\infty^{1/2}$, whereas a lower bound on the Wasserstein curvature leads to deviation results including the sole $\|\cdot\|_{\mathrm{Lip}}$ and enforces the angle bracket of the chain to be bounded.

The paper is organized as follows. In Section 3.2, some basic material on continuous time Markov chains is recalled and we introduce two notions of curvatures of Markov chains, namely the Wasserstein curvature and the Γ-curvature. In Section 3.3, Theorem 3.3.1, a local Poisson-type deviation inequality is established for continuous time Markov chain with Wasserstein curvature bounded below, and we analyze the influence of the sign of such a lower bound in large deviation inequalities. In Section 3.4, a general estimate is derived in Theorem 3.4.2 under the hypothesis of a lower bound on the Γ-curvature, and with further assumptions on the chain, these upper bounds are computed to yield local Poisson tail probabilities involving the mixed Lipschitz seminorms $\|\cdot\|_{\mathrm{Lip}}$ and $f\mapsto\|\Gamma f\|_\infty^{1/2}$. The case of birth-death process on $\mathbb{N}$ or $\{0,1,\ldots,n\}$ is investigated in Section 3.5, in which we give some conditions on the transition rates of the associated generator for such discrete curvatures to be bounded below. As a result, we extend to birth-death processes the deviation inequalities of [3] established for continuous time random walks on graphs, seen as models in null curvature. By a limiting argument, deviation bounds are derived for the stationary distribution of birth-death process on the finite state space $\{0,1,\ldots,n\}$ and we recover the optimal Gaussian concentration for Ornstein-Uhlenbeck processes constructed as fluid limits of rescaled continuous time Ehrenfest chains. Finally, these local Poisson-type deviation inequalities are extended to sample vectors of the $M/M/1$ queueing process by using a tensorization procedure of the Laplace transform together with an integration by parts formula satisfied by the underlying semigroup.

3.2 Notation and preliminaries

Throughout the paper, E is a countable set endowed with a non-trivial metric d, $\mathscr{F}(E)$ is the collection of all real-valued functions on E, $\mathscr{B}(E)\subset\mathscr{F}(E)$ is the space of all real-valued bounded functions on E equipped with the supremum norm $\|f\|_\infty=\sup_{x\in E}|f(x)|$, and $\mathrm{Lip}(E)$ is the subspace of $\mathscr{F}(E)$ consisting of Lipschitz functions on E, i.e.

$$\|f\|_{\mathrm{Lip}}:=\sup_{x\neq y}\frac{|f(x)-f(y)|}{d(x,y)}<+\infty.$$

3.2.1 Basic material on continuous time Markov chains

On a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, consider an E-valued continuous time Markov chain $(X_t)_{t\geq 0}$ with its natural filtration $(\mathscr{F}_t)_{t\geq 0}$ and homogeneous semigroup $(P_t)_{t\geq 0}$ acting on $\mathscr{B}(E)$ as follows:

$$P_t f(x) := \mathbb{E}_x[f(X_t)] = \sum_{y\in E} f(y) P_t(x,y), \quad x \in E.$$

We assume through the paper that the càdlàg chain $(X_t)_{t\geq 0}$ is regular, i.e. the number of its discontinuities is finite on each compact time interval. Define

$$Q_x = \lim_{t\downarrow 0} \frac{1 - P_t(x,x)}{t} \in [0, +\infty], \qquad Q(x,y) = \lim_{t\downarrow 0} \frac{P_t(x,y)}{t} \in [0,+\infty), \quad y \neq x,$$

and denote $Q(x,x) = -Q_x$, $x \in E$. By Theorem 2.2 page 337 in [20], the regularity assumption implies that $(X_t)_{t\geq 0}$ is stable and conservative, i.e. for any $x \in E$, $Q_x < +\infty$, and $\sum_{y\in E} Q(x,y) = 0$, respectively. The generator $\mathcal{L}$ of the chain is given by

$$\mathcal{L}f(x) = \sum_{y\in E} (f(y) - f(x))\, Q(x,y), \quad x \in E,$$

where f belongs to an algebra $\mathcal{A}$ (say) containing the constant functions and which is stable by the action of $\mathcal{L}$, P_t and the C^∞-functions. See for instance [6] for a discussion on the existence of this algebra.

In the remainder of the paper, the chains we consider are implicitly assumed to be non-explosive. In other words, if $(T_n)_{n\in\mathbb{N}}$ denotes the sequence of jump times of the chain $(X_t)_{t\geq 0}$, i.e. $T_0 = 0$ and $T_{n+1} = \inf\{t > T_n : X_t \neq X_{T_n}\}$, $n \in \mathbb{N}$, then for any initial state $x \in E$, we have $\mathbb{P}_x\left(\lim_{n\to+\infty} T_n = +\infty\right) = 1$.

Given $f \in \mathscr{B}(E)$, the process $M^f = (M_t^f)_{t\geq 0}$ defined by

$$M_t^f = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)ds, \quad t \geq 0,$$

is a $(\mathbb{P}_x, \mathscr{F}_t)$-martingale for any $x \in E$, which has the representation:

$$M_t^f = \sum_{y,z\in E} \int_0^t (f(y) - f(z))\, 1_{\{X_{s-}=z\}} (N_{z,y} - \sigma_{z,y})(ds),$$

where $(N_{z,y})_{z,y\in E}$ is a family of independent Poisson processes on $\mathbb{R}_+$ with respective intensity $\sigma_{z,y}(dt) = Q(z,y)dt$.

If $(X_t)_{t\geq 0}$ is square-integrable, then the angle bracket process exists and is given by

$$\langle X, X\rangle_t = \sum_{y,z\in E} d(z,y)^2 \int_0^t 1_{\{X_s = z\}} Q(z,y) ds, \quad t \geq 0.$$

If there exists $V > 0$ such that $\left\| \sum_{y\in E} d(\cdot, y)^2 Q(\cdot, y) \right\|_\infty \leq V^2$, then $\langle X, X\rangle_t \leq V^2 t$ and we say that $(X_t)_{t\geq 0}$ has angle bracket bounded by V^2.

Finally, we say that the chain $(X_t)_{t\geq 0}$ has jumps bounded by some positive b if $\sup_{t>0} d(X_{t-}, X_t) \leq b$.

3.2.2 Curved continuous time Markov chains

Wasserstein curvature of regular Markov chains

Let us introduce the notion of curved Markov chain in the Wasserstein sense.

Definition 3.2.1. *The Wasserstein curvature at time $t > 0$ of a regular Markov chain with semigroup $(P_t)_{t\geq 0}$ is defined by*

$$K_t := -\frac{1}{t} \sup \left\{ \log \left(\frac{\|P_t f\|_{\mathrm{Lip}}}{\|f\|_{\mathrm{Lip}}} \right) : f \in \mathcal{A} \cap \mathrm{Lip}(E),\ f \neq \text{constant} \right\} \quad \in [-\infty, +\infty).$$

It is said to be bounded below by $K \in \mathbb{R}$ if $\inf_{t>0} K_t \geq K$.

Remark 3.2.2. We call this curvature the Wasserstein curvature since it is connected with the so-called Wasserstein distances. Indeed, if $\mathscr{P}(E)$ denotes the space of probability measures on the subsets of E equipped with the weak topology and $\mathscr{P}_1(E)$ is the subset of $\mathscr{P}(E)$ consisting of all μ such that $\sum_{y\in E} d(x,y)\mu(y) < +\infty$ for some (or equivalently for any) $x \in E$, then given $\mu, \nu \in \mathscr{P}_1(E)$, define the Wasserstein distance W between μ and ν by

$$W(\mu,\nu) := \inf_{\pi} \sum_{x,y\in E} d(x,y)\pi(x,y),$$

where the infimum runs over all $\pi \in \mathscr{P}_1(E \times E)$ with marginals μ and ν, making $\mathscr{P}_1(E)$ a Polish space, see for instance [81]. The Kantorovich-Rubinstein duality theorem states that the Wasserstein distance rewrites as

$$W(\mu,\nu) = \sup \left\{ \left| \sum_{x\in E} f(x)(\mu(x) - \nu(x)) \right| : \|f\|_{\mathrm{Lip}} \leq 1 \right\}.$$

Thus, if a Markov kernel $P_t(x,\cdot) \in \mathscr{P}_1(E)$ for some $x \in E$ and any positive t, then the following assertions are equivalent:

(i) $\inf_{t>0} K_t \geq K$;

(ii) $\|P_t f\|_{\mathrm{Lip}} \leq e^{-Kt}\|f\|_{\mathrm{Lip}}$, for any $f \in \mathcal{A} \cap \mathrm{Lip}(E)$ and any $t > 0$;

(iii) $W(P_t(x,\cdot), P_t(y,\cdot)) \leq e^{-Kt} d(x,y)$ for any $x, y \in E$ and any $t > 0$.

Hence, these assertions characterize lower bounds on the Wasserstein curvature in terms of contraction properties of the semigroup in the Wasserstein metric W. Note that a version of (iii) above has been introduced by K. Marton in [66] with the trivial metric $d(x,y) = 1_{x\neq y}$, and also by H. Djellout, A. Guillin and L. Wu through the condition (C1) in [35], in order to study transportation cost inequalities for weakly dependent sequences.

Remark 3.2.3. By the Kantorovich-Rubinstein duality theorem together with [29, Theorem 5.23], any chain with Wasserstein curvature bounded below by some positive constant

K is positive recurrent and thus has a unique stationary distribution $\pi \in \mathscr{P}_1(E)$. Therefore, according to the Kantorovich-Rubinstein duality theorem, we have:

$$\begin{aligned} W(P_t(x,\cdot),\pi) &= \sup_{\|f\|_{\mathrm{Lip}}\le 1} \left| \sum_{y\in E} f(y)(P_t(x,y)-\pi(y)) \right| \\ &= \sup_{\|f\|_{\mathrm{Lip}}\le 1} \left| \sum_{y,z\in E} f(y)(P_t(x,y)-P_t(z,y))\pi(z) \right| \\ &\le \sup_{\|f\|_{\mathrm{Lip}}\le 1} \sum_{z\in E} |P_t f(x) - P_t f(z)|\pi(z) \\ &\le e^{-Kt} \sum_{z\in E} d(x,z)\pi(z), \end{aligned}$$

which goes to 0 as t tends to infinity. Hence, the positive number K describes the speed of convergence of the Markov chain to stationarity with respect to the Wasserstein metric W.

Γ-curvature of regular Markov chains

Recall that the "carré du champ" operator Γ is the symmetric bilinear mapping defined on $\mathcal{A}\times\mathcal{A}$ by

$$\begin{aligned} \Gamma(f,g)(x) &:= \frac{1}{2}\left(\mathcal{L}(fg)(x) - f(x)\mathcal{L}g(x) - g(x)\mathcal{L}f(x)\right) \\ &= \frac{1}{2}\sum_{y\in E} \left(f(y)-f(x)\right)\left(g(y)-g(x)\right) Q(x,y). \end{aligned}$$

We let $\Gamma f = \Gamma(f,f)$ and introduce the notion of curved Markov chains in the Γ-sense:

Definition 3.2.4. *The Γ-curvature at time $t>0$ of a regular Markov chain with semigroup $(P_t)_{t\ge 0}$ is defined by*

$$\rho_t := -\frac{1}{t}\sup\left\{ \log \left\| \frac{(\Gamma P_t f)^{1/2}}{P_t(\Gamma f)^{1/2}} \right\|_\infty : f\in\mathcal{A},\ f\ne \text{constant} \right\} \quad \in [-\infty,+\infty).$$

It is said to be bounded below by $\rho\in\mathbb{R}$ if $\inf_{t>0}\rho_t \ge \rho$.

Remark 3.2.5. By definition, the Γ-curvature is bounded below by $\rho\in\mathbb{R}$ if and only if for any $f\in\mathcal{A}$,

$$(\Gamma P_t f)^{1/2}(x) \le e^{-\rho t} P_t (\Gamma f)^{1/2}(x), \quad x\in E, \quad t>0, \tag{3.2.1}$$

which is the analogue in discrete settings of the classical commutation relation between local gradient and heat kernel on Riemannian manifolds with Ricci curvature bounded below, see [7].

Main differences between discrete curvatures

As already mentioned in the introduction, both curvatures are essentially equivalent for Brownian motions on Riemannian manifolds, see [79, Theorem 2]. This is no longer the case in discrete settings since discrete gradients do not satisfy the chain rule formula, and the discrete curvatures defined above are not directly comparable. However, note that the inequality (3.2.1) is a pointwise commutation relation between the semigroup and a discrete gradient induced by the operator Γ, whereas a lower bound on the Wasserstein curvature entails via the item (ii) of Remark 3.2.2 an inequality between Lipschitz semi-norms and where the semigroup is dropped in its right-hand-side. Hence, the assumption (ii) is weaker than (3.2.1) in some sense and we expect that a lower bound K on the Wasserstein curvature entails weaker deviation results than that established under the assumption of the same lower bound K on the Γ-curvature.

Preliminary comments on tail estimates

Let us make some comments on the deviation inequalities we will establish in the remainder of this paper:

- Our estimates are said to be local since they are given with respect to the probability measures $P_t(x,\cdot)$, $t>0$, uniformly in the initial state $x \in E$. Moreover, we give in general two estimates for each result, to emphasize the good order of magnitude of the exponential decays in the deviation bounds. The second one is easily deduced from the first one by using the elementary inequality $(1+u)\log(1+u) - u \geq u\log(1+u)/2$, $u \geq 0$.

- For the sake of simplicity, our results are concerned with right tail estimates of type $\mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right)$, where the level of deviation y is positive. However, replacing in the corresponding inequalities f by $-f$, two-side tail estimates of type $\mathbb{P}_x\left(\left|f(X_t) - \mathbb{E}_x\left[f(X_t)\right]\right| \geq y\right)$ can be obtained.

- Similarly to the paper [42] for infinitely divisible random vectors with compactly supported Lévy measures, the boundedness assumption on the jumps of the chain allows us to derive explicit Poisson like inequalities, see for instance Theorem 3.3.1 or Corollary 3.4.4, whereas the general case yields the formal tail estimate (3.4.2) of Theorem 3.4.2. Moreover, all our results are still available when replacing the upper bound on the jumps $b \geq \sup_{t>0} d(X_{t-}, X_t)$ by the deterministic time-dependent upper bound $b_t \geq \sup_{0<s\leq t} d(X_{s-}, X_s)$, $t>0$.

- We do not investigate in this paper the case of independent product Markov chains, since our results would be sub-optimal with respect to the dimension. Indeed, our proofs are based on the tensorization of the Laplace transform with respect to the ℓ^1-metric, which is not well-adapted to handle dimension-free concentration results, see for instance the discussion in [62, Section 1.6].

- Denote $\log_+(x) = \max(\log(x), 0)$, $x>0$. A classical consequence of our Poisson-type deviation inequalities is the following exponential integrability property: for any

$f \in \mathrm{Lip}(E)$, any positive t and sufficiently small $\lambda > 0$, we have:

$$\sup_{x\in E} \mathbb{E}_x \left[e^{\lambda |f(X_t)-\mathbb{E}_x[f(X_t)]| \log_+ |f(X_t)-\mathbb{E}_x[f(X_t)]|} \right] < +\infty.$$

3.3 Deviation bounds for curved Markov chains in the Wasserstein sense

In this part, we present Poisson-type deviation results under the assumption of a lower bound on the Wasserstein curvature.

Theorem 3.3.1. *Let $(X_t)_{t\geq 0}$ be a regular Markov chain on E with jumps and angle bracket bounded respectively by $b > 0$ and $V^2 > 0$. Assume moreover that its Wasserstein curvature is bounded below by $K \in \mathbb{R}$. Let $f \in \mathrm{Lip}(E)$ and define for any $t > 0$ the positive numbers $C_{t,K} = \sup_{0\leq s\leq t} e^{-K(t-s)}$ and $M_{t,K} = (1-e^{-2Kt})/(2K)$ ($M_{t,K} = t$ if $K = 0$). Then for any initial state $x \in E$, any $y > 0$ and any $t > 0$, we have the local Poisson-type deviation inequality:*

$$\begin{aligned} \mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right) &\leq \exp\left(-\frac{M_{t,K}V^2}{b^2 C_{t,K}^2} g\left(\frac{bC_{t,K}y}{M_{t,K}V^2\|f\|_{\mathrm{Lip}}}\right)\right) \qquad (3.3.1) \\ &\leq \exp\left(-\frac{y}{2bC_{t,K}\|f\|_{\mathrm{Lip}}}\log\left(1+\frac{bC_{t,K}y}{M_{t,K}V^2\|f\|_{\mathrm{Lip}}}\right)\right), \end{aligned}$$

where $g(u) = (1+u)\log(1+u) - u$, $u > 0$.

Proof. Fix $x \in E$, $t > 0$, and assume first that f is bounded. The process $\left(Z_s^f\right)_{0\leq s\leq t}$ given by

$$Z_s^f := P_{t-s}f(X_s) - P_t f(X_0)$$

is a real $\mathbb{P}_x$-martingale with respect to the truncated filtration $(\mathscr{F}_s)_{0\leq s\leq t}$ and we have by Itô's formula:

$$Z_s^f = \sum_{y,z\in E} \int_0^s \left(P_{t-\tau}f(y) - P_{t-\tau}f(z)\right) 1_{\{X_{\tau-}=z\}}(N_{z,y} - \sigma_{z,y})(d\tau).$$

Since the Wasserstein curvature is bounded below, the jumps of $\left(Z_s^f\right)_{0\leq s\leq t}$ are bounded for any $s \in [0,t]$:

$$\begin{aligned} \left|Z_s^f - Z_{s-}^f\right| &= \left|P_{t-s}f(X_s) - P_{t-s}f(X_{s-})\right| \\ &\leq d(X_s, X_{s-})\|f\|_{\mathrm{Lip}} C_{t,K} \\ &\leq b\|f\|_{\mathrm{Lip}} C_{t,K}, \end{aligned}$$

as its angle bracket:

$$\langle Z^f, Z^f\rangle_s = \sum_{y,z\in E} \int_0^s \left(P_{t-\tau}f(y) - P_{t-\tau}f(z)\right)^2 1_{\{X_{\tau-}=z\}}\,\sigma_{z,y}(d\tau)$$

$$\begin{aligned}
&\leq \|f\|_{\mathrm{Lip}}^2 \sum_{y,z\in E} \int_0^s e^{-2K(t-\tau)} d(z,y)^2 1_{\{X_{\tau-}=z\}} Q(z,y) d\tau \\
&\leq \|f\|_{\mathrm{Lip}}^2 M_{t,K} V^2.
\end{aligned}$$

By [56, Lemma 23.19], for any positive λ, the process $(Y_s^{(\lambda)})_{0\leq s\leq t}$ given by

$$Y_s^{(\lambda)} := \exp\left(\lambda Z_s^f - \lambda^2 \psi(\lambda b\|f\|_{\mathrm{Lip}} C_{t,K}) \langle Z^f, Z^f\rangle_s\right)$$

is a $\mathbb{P}_x$-supermartingale with respect to $(\mathscr{F}_s)_{0\leq s\leq t}$, where $\psi(z) = z^{-2}\left(e^z - z - 1\right)$, $z > 0$. Thus, we get for any $\lambda > 0$:

$$\begin{aligned}
\mathbb{E}_x\left[e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] &= \mathbb{E}_x\left[e^{\lambda Z_t^f}\right] \\
&\leq \exp\left(\lambda^2 \|f\|_{\mathrm{Lip}}^2 M_{t,K} V^2\, \psi(\lambda b\|f\|_{\mathrm{Lip}} C_{t,K})\right) \mathbb{E}_x\left[Y_t^{(\lambda)}\right] \\
&\leq \exp\left(\lambda^2 \|f\|_{\mathrm{Lip}}^2 M_{t,K} V^2\, \psi(\lambda b\|f\|_{\mathrm{Lip}} C_{t,K})\right) \\
&= \exp\left(\frac{M_{t,K}V^2}{b^2 C_{t,K}^2}\left(e^{\lambda b\|f\|_{\mathrm{Lip}} C_{t,K}} - \lambda b\|f\|_{\mathrm{Lip}} C_{t,K} - 1\right)\right).
\end{aligned}$$

Finally, using the exponential Chebychev's inequality and optimizing in $\lambda > 0$ in the exponential estimate above, the deviation inequality (3.3.1) is established in the bounded case.
To remove the boundedness assumption, let $f \in \mathrm{Lip}(E)$ and consider the bounded function $f_n = \max\{-n, \min\{f, n\}\}$, $n \in \mathbb{N}$. We have the pointwise convergence $f_n \uparrow f$ and by a classical argument, see for instance the proof of Proposition 10 in [19], $(f_n)_{n\in\mathbb{N}}$ is uniformly integrable with respect to the probability measure $P_t(x,\cdot)$, which implies the convergence of $\mathbb{E}_x[f_n(X_t)]$ to $\mathbb{E}_x[f(X_t)]$. Since $\|f_n\|_{\mathrm{Lip}} \leq \|f\|_{\mathrm{Lip}}$ and that g is non-decreasing on $\mathbb{R}_+$, we finally have by Fatou's lemma:

$$\begin{aligned}
\mathbb{P}_x\left(f(X_t) - \mathbb{E}_x[f(X_t)] \geq y\right) &\leq \liminf_{n\to+\infty} \mathbb{P}_x\left(f_n(X_t) - \mathbb{E}_x[f_n(X_t)] \geq y\right) \\
&\leq \liminf_{n\to+\infty} \exp\left(-\frac{M_{t,K}V^2}{b^2 C_{t,K}^2} g\left(\frac{bC_{t,K}y}{M_{t,K}V^2\|f_n\|_{\mathrm{Lip}}}\right)\right) \\
&\leq \exp\left(-\frac{M_{t,K}V^2}{b^2 C_{t,K}^2} g\left(\frac{bC_{t,K}y}{M_{t,K}V^2\|f\|_{\mathrm{Lip}}}\right)\right).
\end{aligned}$$

Theorem 3.3.1 is established in full generality. ■

Remark 3.3.2. If $K = 0$, then the estimate in Theorem 3.3.1 is comparable to the tail inequalities of [42, 77] established for Lévy processes and infinitely divisible distributions with compactly supported Lévy measure. If $K < 0$, then the decay in (3.3.1) is slower, due to some exponential factors, whereas if $K > 0$, the chain is positive recurrent and such estimates can be extended to the stationary distribution, as illustrated below and in Section 3.5.2.

Large deviation bounds

Let us now analyze the influence of the sign of the lower bound K of the Wasserstein curvature on some large deviation inequalities which are direct applications of (3.3.1). Fix the initial state $x \in E$ and the deviation level $y > 0$. Under the assumptions of Theorem 3.3.1, we have the following behaviors:

(i) If t tends to 0, then we have the estimate independent of K:

$$\limsup_{t \to 0} -\frac{1}{\log(t)} \log \mathbb{P}_x \left(f(X_t) - \mathbb{E}_x \left[f(X_t) \right] \geq y \right) \leq -\frac{y}{b \| f \|_{\mathrm{Lip}}}.$$

The speed of convergence is $-1/\log(t)$, which is sharp in the case of continuous time Markov chains whose rate functions of the generator are bounded, see [2]. One deduces that the sign of K has no influence in small time in (3.3.1).

(ii) If t tends to infinity, then the sign of K is crucial in (3.3.1). Indeed, if K is positive, then the existence of a unique stationary distribution π is assured by positive recurrence, as noted in Remark 3.2.3. The positivity of K achieves the best deviation inequality and as t tends to infinity, (3.3.1) entails an inequality for the stationary distribution π:

$$\pi \left(f - \mathbb{E}_\pi[f] \geq y \right) \leq \exp \left(\frac{y}{b \| f \|_{\mathrm{Lip}}} - \left(\frac{y}{b \| f \|_{\mathrm{Lip}}} + \frac{V^2}{2b^2 K} \right) \log \left(1 + \frac{2bKy}{V^2 \| f \|_{\mathrm{Lip}}} \right) \right),$$

where $\mathbb{E}_\pi[f]$ denotes the expectation of f with respect to π. See Section 3.5.2 for a more careful analysis of deviation estimates for stationary distributions of curved birth-death processes on finite state spaces. On the other hand, if $K \leq 0$, then the worst deviation inequality is realized and the limiting argument above is no longer available in (3.3.1), since $C_{t,K}$ and $M_{t,K}$ (which strongly depend on K) are not bounded uniformly in time.

To conclude this section, note that Theorem 3.3.1 allows us to consider neither Markov chains with unbounded angle bracket nor another Lipschitz seminorms than $\| \cdot \|_{\mathrm{Lip}}$. To overcome this difficulty, one has to require some assumptions on another curvature of the chain, namely the Γ-curvature.

3.4 Tail estimates for curved Markov chains in the Γ-sense

In this section, we adapt to the Markovian case the covariance method of the papers [42, 44] to derive local deviation inequalities for curved Markov chains in the Γ-sense. Although Wasserstein and Γ-curvatures are not comparable in discrete spaces, the results we give in this part are more general than that in Section 3.3.

3.4.1 A general bound

Given $(X_t)_{t\geq 0}$ a regular Markov chain on E and two functions $f, g \in \mathscr{B}(E)$, we define the local covariance of $f(X_t)$ and $g(X_t)$ by

$$\mathrm{Cov}_x\left[f(X_t), g(X_t)\right] := \mathbb{E}_x\left[\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right]\right)\left(g(X_t) - \mathbb{E}_x\left[g(X_t)\right]\right)\right], \quad x \in E.$$

Before turning to Theorem 3.4.2 below, let us establish the following

Lemma 3.4.1. *Let $(X_t)_{t\geq 0}$ be a regular Markov chain on E with Γ-curvature bounded below by $\rho \in \mathbb{R}$. Let $g_1, g_2 \in \mathscr{B}(E)$ with $\|\Gamma g_1\|_\infty < +\infty$ and define $L_{t,\rho} = (1 - e^{-2\rho t})/(2\rho)$ if $\rho \neq 0$, and $L_{t,\rho} = t$ otherwise. Then for any initial state $x \in E$ and any time $t > 0$, we have the local covariance inequality:*

$$\mathrm{Cov}_x\left[g_1(X_t), g_2(X_t)\right] \leq 2L_{t,\rho}\|\Gamma g_1\|_\infty^{1/2}\,\mathbb{E}_x\left[(\Gamma g_2)^{1/2}(X_t)\right], \quad t > 0.$$

Proof. Fix $x \in E$ and $t > 0$. As in the proof of Theorem 3.3.1, we have for $i = 1, 2$:

$$g_i(X_t) - \mathbb{E}_x\left[g_i(X_t)\right] = \sum_{y,z\in E}\int_0^t \left(P_{t-s}g_i(y) - P_{t-s}g_i(z)\right) 1_{\{X_{s-}=z\}}(N_{z,y} - \sigma_{z,y})(ds).$$

By Cauchy-Schwarz inequality,

$$\begin{aligned}
&\mathrm{Cov}_x\left[g_1(X_t), g_2(X_t)\right] \\
&\quad = \mathbb{E}_x\left[\left(g_1(X_t) - \mathbb{E}_x\left[g_1(X_t)\right]\right)\left(g_2(X_t) - \mathbb{E}_x\left[g_2(X_t)\right]\right)\right] \\
&\quad = 2\,\mathbb{E}_x\left[\int_0^t \Gamma(P_{t-s}g_1, P_{t-s}g_2)(X_s)\,ds\right] \\
&\quad = 2\int_0^t P_s\left(\Gamma(P_{t-s}g_1, P_{t-s}g_2)\right)(x)\,ds \\
&\quad \leq 2\int_0^t P_s\left((\Gamma P_{t-s}g_1)^{1/2}\,(\Gamma P_{t-s}g_2)^{1/2}\right)(x)\,ds \\
&\quad \leq 2\int_0^t e^{-2\rho(t-s)}P_s\left(P_{t-s}(\Gamma g_1)^{1/2}\,P_{t-s}(\Gamma g_2)^{1/2}\right)(x)\,ds,
\end{aligned} \tag{3.4.1}$$

where in (3.4.1) we used the assumption of a lower bound ρ on the Γ-curvature. Since $(P_t)_{t\geq 0}$ is a contraction operator on $\mathscr{B}(E)$, we have:

$$\begin{aligned}
\mathrm{Cov}_x\left[g_1(X_t), g_2(X_t)\right] &\leq 2\,\|\Gamma g_1\|_\infty^{1/2}\int_0^t e^{-2\rho(t-s)}P_s\left(P_{t-s}(\Gamma g_2)^{1/2}\right)(x)\,ds \\
&= 2L_{t,\rho}\|\Gamma g_1\|_\infty^{1/2}\,\mathbb{E}_x\left[(\Gamma g_2)^{1/2}(X_t)\right].
\end{aligned}$$

■

Now, we are able to state Theorem 3.4.2 which presents a general deviation bound for curved Markov chains in the Γ-sense:

Theorem 3.4.2. *Let $(X_t)_{t\geq 0}$ be a regular Markov chain on E with Γ-curvature bounded below by $\rho \in \mathbb{R}$. Let $f \in \mathrm{Lip}(E)$ with $\|\Gamma f\|_\infty < +\infty$, and define the function $\psi_{f,t}$: $(0,+\infty) \to \mathbb{R}_+ \cup \{\infty\}$ by*

$$\psi_{f,t}(\lambda) := \sqrt{2} L_{t,\rho} \|\Gamma f\|_\infty^{1/2} \left\| \sum_{y\in E} (f(y) - f(\cdot))^2 \left(\frac{e^{\lambda \|f\|_{\mathrm{Lip}} d(\cdot,y)} - 1}{\|f\|_{\mathrm{Lip}} d(\cdot,y)} \right)^2 Q(\cdot,y) \right\|_\infty^{1/2}, \quad t > 0,$$

where $L_{t,\rho}$ is defined in Lemma 3.4.1. Then for any initial state $x \in E$, any deviation level $y > 0$ and any $t > 0$, we get the local tail probability:

$$\mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right) \leq \exp \inf_{\lambda\in(0,M_{f,t})} \int_0^\lambda (\psi_{f,t}(\tau) - y)\, d\tau, \tag{3.4.2}$$

where $M_{f,t} = \sup\{\lambda > 0 : \psi_{f,t}(\lambda) < +\infty\}$.

Remark 3.4.3. Note that $\psi_{f,t}$ is bijective from $(0, M_{f,t})$ to $(0,+\infty)$, so that the term in the exponential is negative and the inequality (3.4.2) makes sense.

Proof. Fix $x \in E$ and $t > 0$. Proceeding as in the end of the proof of Theorem 3.3.1, it is sufficient to establish the result for bounded Lipschitz function f. Applying Lemma 3.4.1 with $g_1 = f - \mathbb{E}_x[f(X_t)]$ and $g_2 = \exp\left(\lambda(f - \mathbb{E}_x[f(X_t)])\right)$, $\lambda \in (0, M_{f,t})$, we have:

$$\begin{aligned}
&\mathbb{E}_x\left[(f(X_t) - \mathbb{E}_x[f(X_t)])\, e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] \\
=\;& \mathrm{Cov}_x\left[f(X_t) - \mathbb{E}_x[f(X_t)], e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] \\
\leq\;& 2L_{t,\rho}\|\Gamma f\|_\infty^{1/2}\, e^{-\lambda\mathbb{E}_x[f(X_t)]}\mathbb{E}_x\left[\left(\Gamma e^{\lambda f}\right)^{1/2}(X_t)\right] \\
=\;& \sqrt{2}\, L_{t,\rho}\|\Gamma f\|_\infty^{1/2}\, \mathbb{E}_x\left[e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])} \left(\sum_{y,z\in E} \left(e^{\lambda(f(y)-f(z))} - 1\right)^2 1_{\{X_t=z\}} Q(z,y)\right)^{1/2}\right] \\
\leq\;& \sqrt{2}\, L_{t,\rho}\|\Gamma f\|_\infty^{1/2} \\
\times\;& \mathbb{E}_x\left[e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])} \left(\sum_{y,z\in E} (f(y)-f(z))^2\, 1_{\{X_t=z\}} \left(\frac{e^{\lambda\|f\|_{\mathrm{Lip}} d(y,z)} - 1}{\|f\|_{\mathrm{Lip}} d(y,z)}\right)^2 Q(z,y)\right)^{1/2}\right],
\end{aligned}$$

where in the last inequality we used the elementary $|e^z - 1| \leq e^{|z|} - 1$, $z \in \mathbb{R}$, together with the increase of the function $z \mapsto (e^z - 1)/z$ on $(0,+\infty)$. Thus, we obtain:

$$\mathbb{E}_x\left[(f(X_t) - \mathbb{E}_x[f(X_t)])\, e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] \leq \psi_{f,t}(\lambda)\mathbb{E}_x\left[e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])}\right].$$

Letting $H_{f,t,x}(\lambda) = \mathbb{E}_x\left[e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])}\right]$, the latter inequality rewrites as

$$\frac{dH_{f,t,x}(\lambda)}{d\lambda} \leq \psi_{f,t}(\lambda) H_{f,t,x}(\lambda),$$

and integrating the above differential inequality yields:

$$\mathbb{E}_x\left[e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] \leq e^{\int_0^\lambda \psi_{f,t}(\tau)d\tau}, \quad \lambda \in (0, M_{f,t}).$$

Finally, using the exponential Chebychev inequality, Theorem 3.4.2 is established. ∎

3.4.2 Some explicit tail estimates

Since the estimate (3.4.2) is very general, let us make further assumptions on the chain to get more explicit inequalities. Denote in the sequel $L_{t,\rho} = (1 - e^{-2\rho t})/(2\rho)$ if $\rho \neq 0$, and $L_{t,\rho} = t$ otherwise, and denote the function $g(u) = (1+u)\log(1+u) - u$, $u > 0$. Using the notation of Theorem 3.4.2, we have the

Corollary 3.4.4. *Under the hypothesis of Theorem 3.4.2, suppose moreover that $(X_t)_{t\geq 0}$ has jumps bounded by $b > 0$. Then for any initial state $x \in E$, any $y > 0$ and any $t > 0$, we get the local Poisson-type deviation inequality:*

$$\begin{aligned} \mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right) &\leq \exp\left(-\frac{2L_{t,\rho}\|\Gamma f\|_\infty}{b^2\|f\|_{\mathrm{Lip}}^2} g\left(\frac{by\|f\|_{\mathrm{Lip}}}{2L_{t,\rho}\|\Gamma f\|_\infty}\right)\right) \qquad (3.4.3) \\ &\leq \exp\left(-\frac{y}{2b\|f\|_{\mathrm{Lip}}}\log\left(1 + \frac{by\|f\|_{\mathrm{Lip}}}{2L_{t,\rho}\|\Gamma f\|_\infty}\right)\right). \end{aligned}$$

Proof. Under the notation of Theorem 3.4.2, the boundedness of the jumps implies $M_{f,t} = +\infty$, $t > 0$, and $\psi_{f,t}$ is bounded by

$$\psi_{f,t}(\lambda) \leq 2L_{t,\rho}\|\Gamma f\|_\infty \frac{e^{\lambda b\|f\|_{\mathrm{Lip}}} - 1}{b\|f\|_{\mathrm{Lip}}}, \quad \lambda > 0.$$

Using then Theorem 3.4.2 and optimizing in $\lambda > 0$, the proof is achieved. ∎

Note that (3.4.3) is more general than (3.3.1), since the finiteness assumption on $\|\Gamma f\|_\infty$ allows us to consider Markov chains with non necessarily bounded angle bracket. Thus, when the angle bracket of the process is bounded, the next corollary exhibits an estimate comparable to that of Theorem 3.3.1:

Corollary 3.4.5. *Let $(X_t)_{t\geq 0}$ be a regular Markov chain on E with jumps and angle bracket bounded respectively by $b > 0$ and $V^2 > 0$. Assume moreover that its Γ-curvature is bounded below by $\rho \in \mathbb{R}$, and let $f \in \mathrm{Lip}(E)$. Then for any initial state $x \in E$, any $y > 0$ and any $t > 0$, we get the local Poisson tail probability:*

$$\begin{aligned} \mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right) &\leq \exp\left(-\frac{L_{t,\rho}V^2}{b^2} g\left(\frac{by}{L_{t,\rho}V^2\|f\|_{\mathrm{Lip}}}\right)\right) \qquad (3.4.4) \\ &\leq \exp\left(-\frac{y}{2b\|f\|_{\mathrm{Lip}}}\log\left(1 + \frac{by}{L_{t,\rho}V^2\|f\|_{\mathrm{Lip}}}\right)\right). \end{aligned}$$

Proof. By the boundedness of the jumps and of the angle bracket, the function $\psi_{f,t}$ in Theorem 3.4.2 is bounded by

$$\psi_{f,t}(\lambda) \leq L_{t,\rho}V^2\|f\|_{\mathrm{Lip}} \frac{e^{\lambda b\|f\|_{\mathrm{Lip}}} - 1}{b}, \quad \lambda > 0.$$

Finally, applying Theorem 3.4.2 yields the result. ∎

Remark 3.4.6. As in Section 3.3, a similar discussion about large deviation bounds under the assumption of Γ-curvature bounded below can be derived from the estimates (3.4.3) and (3.4.4), so we omit it.

3.5 Deviation probabilities for birth-death processes

Among the main results of the paper [3], some local deviation inequalities are established for continuous time random walks on graphs. Actually, such processes may be seen as models in null curvature, since the transition rates of the associated generator do not depend on the space-variable.

By using the general results of Sections 3.3 and 3.4, the purpose of this section is to extend and sharpen these local tail estimates to birth-death processes whose discrete curvatures are bounded below. Let us introduce now some basic material about birth-death processes.

Let $(X_t)_{t\geq 0}$ be a birth-death process on the state space $E=\mathbb{N}$ or $E=\{0,1,\ldots,n\}$. It is a regular Markov chain with generator defined on $\mathscr{F}(E)$ (recall that $\mathscr{F}(E)$ is the collection of real-valued functions on E) by

$$\mathcal{L}f(x)=\lambda_x\left(f(x+1)-f(x)\right)+\nu_x\left(f(x-1)-f(x)\right),\quad x\in E, \tag{3.5.1}$$

where the function rates λ and ν are respectively called the birth and death rates of the chain. The chain $(X_t)_{t\geq 0}$ is irreducible on E if and only if the transition rates λ and ν are positive with 0 as reflecting state, i.e. $\nu_0=0$ (if $E=\{0,1,\ldots,n\}$, the state n is also reflecting, i.e. $\lambda_n=0$), and we assume irreducibility in the remainder of the paper.

The transition probabilities of the associated semigroup $(P_t)_{t\geq 0}$ are given for any $x\in E$ by

$$P_t(x,y)=\begin{cases}\lambda_x t+o(t) & \text{if } y=x+1,\\ \nu_x t+o(t) & \text{if } y=x-1,\\ 1-(\lambda_x+\nu_x)t+o(t) & \text{if } y=x,\end{cases}$$

where the function o is defined in a neighborhood of 0 and is such that $o(t)/t$ converges to 0 as t tends to 0. The chain is positive recurrent if and only if

$$\sum_{x\in E\setminus\{0\}}\prod_{y=1}^{x}\frac{\lambda_{y-1}}{\nu_y}<+\infty,$$

and in this case, the unique stationary distribution π is given for any $x\in E$ by

$$\pi(x)=\pi(0)\prod_{y=1}^{x}\frac{\lambda_{y-1}}{\nu_y},\quad \text{with}\quad \pi(0)=\left(1+\sum_{x\in E\setminus\{0\}}\prod_{y=1}^{x}\frac{\lambda_{y-1}}{\nu_y}\right)^{-1}. \tag{3.5.2}$$

By Reuter's criterion, any irreducible birth-death process on $E=\mathbb{N}$ is non-explosive in finite time if and only if

$$\sum_{x=1}^{+\infty}\left(\frac{1}{\lambda_x}+\frac{\nu_x}{\lambda_x\lambda_{x-1}}+\cdots+\frac{\nu_x\cdots\nu_1}{\lambda_x\cdots\lambda_1\lambda_0}\right)=+\infty,$$

see for instance Theorem 4.5 page 352 in [20].

Remark 3.5.1. If the birth rate λ is bounded, then Reuter's criterion immediately applies.

Before stating the main results of this section, let us give some criteria on the rates of the generator of a birth-death process on E which ensure that its discrete curvatures are bounded below.
First, we deal with the Wasserstein curvature.

Proposition 3.5.2. *Let $(X_t)_{t\geq 0}$ be a birth-death process on E with generator $\mathcal{L}$ and stationary distribution π given respectively by (3.5.1) and (3.5.2). Assume that π satisfies the moment condition $\sum_{y\in E} y\pi(y) < \infty$. If there exists a real number K such that*

$$\inf_{x\in E\setminus\{0\}} \lambda_{x-1} - \lambda_x + \nu_x - \nu_{x-1} \geq K, \tag{3.5.3}$$

then the Wasserstein curvature of the chain is bounded below by K.

Remark 3.5.3. If $E = \mathbb{N}$ and the rates of the generator are bounded and satisfy the assumptions of Proposition 3.5.2, then necessarily $K \leq 0$.

Proof. Consider $(X_t^x)_{t\geq 0}$ and $(X_t^y)_{t\geq 0}$ two independent copies of $(X_t)_{t\geq 0}$, starting respectively from x and y. Then the two-dimensional process $(X_t^x, X_t^y)_{t\geq 0}$ has generator $\tilde{\mathcal{L}}$ given by

$$\tilde{\mathcal{L}} f(z,w) = (\mathcal{L} f(\cdot, w))(z) + (\mathcal{L} f(z,\cdot))(w), \quad z, w \in E.$$

Denote d the classical distance on E, i.e. $d(z,w) = |z-w|$, $z, w \in E$. Because the rates of the generator satisfy the inequality (3.5.3), we have immediately the bound $\tilde{\mathcal{L}} d(z, z+1) \leq -K$, $z \in E$, which is equivalent to the inequality

$$\tilde{\mathcal{L}} d(z,w) \leq -K d(z,w), \quad z, w \in E. \tag{3.5.4}$$

Since the stationary distribution π has a finite first moment, the process $(d(X_t^x, X_t^y))_{t\geq 0}$ has finite expectation and using Itô's formula, the drift inequality (3.5.4) and Gronwall's lemma, we obtain

$$\mathbb{E}\left[d(X_t^x, X_t^y)\right] \leq e^{-Kt} d(x,y).$$

Thus, the latter estimate implies the following inequality in terms of Wasserstein distance:

$$W(P_t(x,\cdot), P_t(y,\cdot)) \leq e^{-Kt} d(x,y),$$

and by the equivalent statements of Remark 3.2.2, the Wasserstein curvature of $(X_t)_{t\geq 0}$ is bounded below by K. ∎

In order to establish modified logarithmic Sobolev inequalities for continuous time random walks on $\mathbb{Z}$, the authors in [3] used a suitable Γ_2-calculus to give a criterion under which the Γ-curvature is bounded below by 0. Actually, this criterion can be generalized to any real lower bound on the Γ-curvature via Lemma 3.5.4 below. As in the diffusion case [7], define the Γ_2-operator on $\mathscr{F}(E)$ by

$$\Gamma_2 f(x) = \frac{1}{2}\left(\mathcal{L}\Gamma f(x) - 2\Gamma(f, \mathcal{L} f)(x)\right), \quad x \in E.$$

By adapting the proof in [3] mentioned above, we get the

Lemma 3.5.4. *Let $(X_t)_{t\geq 0}$ be a birth-death process on E with generator $\mathcal{L}$ given by (3.5.1). Assume that there exists $\rho \in \mathbb{R}$ such that the inequality*

$$\Gamma_2 f(x) - \Gamma\left(\Gamma f\right)^{1/2}(x) \geq \rho \Gamma f(x), \quad x \in E, \tag{3.5.5}$$

is satisfied for any $f \in \mathscr{F}(E)$. Then $(X_t)_{t\geq 0}$ has Γ-curvature bounded below by ρ.

Remark 3.5.5. We mention that the equivalence holds in Lemma 3.5.4. Indeed, if the process $(X_t)_{t\geq 0}$ has Γ-curvature bounded below by ρ, then the function α given on $[0,\infty)$ by $\alpha(t) = e^{-\rho t} P_t \sqrt{\Gamma f} - \sqrt{\Gamma P_t f}$ is non-negative and null in 0. Hence we have $\alpha'(0) \geq 0$, which is (3.5.5).

Proposition 3.5.6. *Let $(X_t)_{t\geq 0}$ be a birth-death process on E with generator $\mathcal{L}$ given by (3.5.1). Assume that there exists some non-negative number ρ such that*

$$\inf_{x \in E\setminus\{0, \sup E\}} \min\{\lambda_{x-1} - \lambda_x, \nu_{x+1} - \nu_x\} \geq \rho. \tag{3.5.6}$$

Then the Γ-curvature is bounded below by ρ.

Remark 3.5.7. If $E = \mathbb{N}$ and the rates of the generator satisfy the assumptions of Proposition 3.5.6, then necessarily $\rho = 0$.

Proof. By Lemma 3.5.4, the result holds true if the Γ_2-inequality (3.5.5) above is satisfied, that we prove now. Letting the forward and backward gradients be defined as $\mathrm{d}^+ f = f(\cdot + 1) - f$ and $\mathrm{d}^- f = f(\cdot - 1) - f$, we have for any $x \in E$:

$$\begin{aligned} &2\Gamma_2 f(x) - 2\Gamma\left(\Gamma f\right)^{1/2}(x) = \\ &\qquad \lambda_x(\nu_{x+1} - \nu_x)\left(\mathrm{d}^+ f(x)\right)^2 + \nu_x(\lambda_{x-1} - \lambda_x)\left(\mathrm{d}^- f(x)\right)^2 + I(x) + J(x), \end{aligned}$$

where:

$$\begin{aligned} I(x) &:= \lambda_x \lambda_{x+1} \mathrm{d}^- f(x+1) \mathrm{d}^+ f(x+1) + \lambda_x \nu_x \mathrm{d}^- f(x+1) \mathrm{d}^+ f(x-1) \\ &\quad + \lambda_x \left(\lambda_{x+1}\left(\mathrm{d}^+ f(x+1)\right)^2 + \nu_{x+1}\left(\mathrm{d}^+ f(x)\right)^2\right)^{1/2} \left(\lambda_x\left(\mathrm{d}^+ f(x)\right)^2 + \nu_x\left(\mathrm{d}^- f(x)\right)^2\right)^{1/2}, \end{aligned}$$

and

$$\begin{aligned} J(x) &:= \nu_x \nu_{x-1} \mathrm{d}^+ f(x-1) \mathrm{d}^- f(x-1) + \lambda_x \nu_x \mathrm{d}^- f(x+1) \mathrm{d}^+ f(x-1) \\ &\quad + \nu_x \left(\lambda_{x-1}\left(\mathrm{d}^+ f(x-1)\right)^2 + \nu_{x-1}\left(\mathrm{d}^- f(x-1)\right)^2\right)^{1/2} \left(\lambda_x\left(\mathrm{d}^+ f(x)\right)^2 + \nu_x\left(\mathrm{d}^- f(x)\right)^2\right)^{1/2}. \end{aligned}$$

Since the rates λ and ν satisfy (3.5.6), we get:

$$2\Gamma_2 f(x) - 2\Gamma\left(\Gamma f\right)^{1/2}(x) \geq 2\rho \Gamma f(x) + I(x) + J(x).$$

Proving in the same way that $J \geq 0$, it is sufficient to establish that I is non-negative. Letting $a = \mathrm{d}^- f(x+1)$, $b = \mathrm{d}^+ f(x-1)$ and $c = \mathrm{d}^+ f(x+1)$, we obtain:

$$I(x) \;=\; \lambda_x\left(\lambda_{x+1}c^2 + \nu_{x+1}a^2\right)^{1/2}\left(\lambda_x a^2 + \nu_x b^2\right)^{1/2} + \lambda_x\lambda_{x+1}ac + \lambda_x\nu_x ab$$

$$\begin{aligned}
&\geq \lambda_x \left(\lambda_{x+1} c^2 + \nu_{x+1} a^2\right)^{1/2} \left(\lambda_x a^2 + \nu_x b^2\right)^{1/2} - \lambda_x \lambda_{x+1} |ac| - \lambda_x \nu_x |ab| \\
&= \lambda_x \left(I_1(x) - I_2(x)\right),
\end{aligned}$$

where

$$I_1(x) := \left(\lambda_{x+1} c^2 + \nu_{x+1} a^2\right)^{1/2} \left(\lambda_x a^2 + \nu_x b^2\right)^{1/2} \quad \text{and} \quad I_2(x) := \lambda_{x+1}|ac| + \nu_x |ab|.$$

Using again (3.5.6), we have:

$$\begin{aligned}
&(I_1(x))^2 - (I_2(x))^2 \\
&\quad = \lambda_{x+1}(\lambda_x - \lambda_{x+1}) a^2 c^2 + \nu_x(\nu_{x+1} - \nu_x) a^2 b^2 + \lambda_x \nu_{x+1} a^4 + \lambda_{x+1} \nu_x b^2 c^2 - 2\nu_x \lambda_{x+1} a^2 bc \\
&\quad \geq \nu_x \lambda_{x+1} (a^2 - bc)^2 \geq 0.
\end{aligned}$$

The proof is complete. ■

3.5.1 The case $E = \mathbb{N}$

An estimate for bounded generators

In order to apply the deviation inequalities of Theorem 3.3.1, one has to require that regular Markov chain has Wasserstein curvature bounded below and bounded angle bracket. In the case of birth-death processes on $\mathbb{N}$, the latter assumption follows from the boundedness of the transition rates of the generator.

Theorem 3.5.8. *Let $(X_t)_{t\geq 0}$ be a birth-death process on $\mathbb{N}$ with generator $\mathcal{L}$ and stationary distribution π given respectively by (3.5.1) and (3.5.2). We suppose that the transition rates λ and ν are bounded on $\mathbb{N}$ and π has finite first moment. Assume moreover that there exists $K \leq 0$ such that $\inf_{x\in\mathbb{N}\setminus\{0\}} \lambda_{x-1} - \lambda_x + \nu_x - \nu_{x-1} \geq K$. Then for any $f \in \mathrm{Lip}(\mathbb{N})$, any initial state $x \in \mathbb{N}$, any deviation level $y > 0$ and any $t > 0$, we have the local Poisson-type tail estimate:*

$$\begin{aligned}
&\mathbb{P}_x \left(f(X_t) - \mathbb{E}_x [f(X_t)] \geq y\right) \\
&\quad \leq \exp\left(-\frac{\sinh(tK)\|\lambda+\nu\|_\infty}{Ke^{-tK}} g\left(\frac{yK}{\sinh(tK)\|\lambda+\nu\|_\infty \|f\|_{\mathrm{Lip}}}\right)\right) \\
&\quad \leq \exp\left(-\frac{ye^{tK}}{2\|f\|_{\mathrm{Lip}}} \log\left(1 + \frac{yK}{\sinh(tK)\|\lambda+\nu\|_\infty \|f\|_{\mathrm{Lip}}}\right)\right),
\end{aligned} \tag{3.5.7}$$

where $g(u) = (1+u)\log(1+u) - u$, $u > 0$. If $K = 0$, then replace (3.5.7) by its limit as $K \to 0$.

Proof. By Proposition 3.5.2, the Wasserstein curvature is bounded below by K. Hence using Theorem 3.3.1 achieves the proof. ■

An inequality for non necessarily bounded generators

In this part, no particular boundedness assumption is made on the generator of birth-death process.

Theorem 3.5.9. *Let $(X_t)_{t\geq 0}$ be a birth-death process on $\mathbb{N}$ with generator $\mathcal{L}$ given by (3.5.1). Assume that λ and ν are respectively non-increasing and non-decreasing. Let $f \in \mathrm{Lip}(\mathbb{N})$ with furthermore $\|\Gamma f\|_\infty < +\infty$. Then for any initial state $x \in \mathbb{N}$, any $y > 0$ and any $t > 0$, we have the local deviation estimate:*

$$\begin{aligned}\mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right) &\leq \exp\left(-\frac{2t\|\Gamma f\|_\infty}{\|f\|_{\mathrm{Lip}}^2} g\left(\frac{y\|f\|_{\mathrm{Lip}}}{2t\|\Gamma f\|_\infty}\right)\right) \\ &\leq \exp\left(-\frac{y}{2\|f\|_{\mathrm{Lip}}}\log\left(1 + \frac{y\|f\|_{\mathrm{Lip}}}{2t\|\Gamma f\|_\infty}\right)\right),\end{aligned}$$

where $g(u) = (1+u)\log(1+u) - u$, $u > 0$.

Proof. By Proposition 3.5.6, the Γ-curvature is bounded below by 0. Since the birth rate λ is bounded above by λ_0, the chain is non-explosive by Reuter's criterion, and applying Corollary 3.4.4 with the lower bound $\rho = 0$ yields the result. ∎

Remark 3.5.10. As claimed above, Theorem 3.5.9 is available for birth-death processes with non necessarily bounded generator, in contrast to Theorem 3.5.8. However, the price to pay in the unbounded case is to require that f has a sub-linear growth at infinity.

3.5.2 The case $E = \{0, 1, \ldots, n\}$

If π denotes the stationary distribution of an irreducible Markov chain on a finite state space, then it satisfies a logarithmic Sobolev inequality, see [74], which in turn implies via the Herbst method that Lipschitz functions have Gaussian tails under π. However, it is sometimes interesting to weaken the upper bound in terms of the deviation level to have a better control of the tail with respect to some parameters, see for instance the discussion in [19] about concentration for Bernoulli distributions and penalties.
In this way, the purpose of this part is to refine Theorem 3.5.8 and Theorem 3.5.9 when the state space is finite, in order to establish by a limiting argument Poisson-type deviation estimates for stationary distributions of birth-death processes. To do so, the crucial point is to obtain positive lower bounds on discrete curvatures.
Our estimates below may be compared to that of [45, Proposition 4] established under reversibility assumptions and without notion of discrete curvatures.

Theorem 3.5.11. *Let $(X_t)_{t\geq 0}$ be a birth-death process on $\{0, 1, \ldots, n\}$ with generator $\mathcal{L}$ and stationary distribution π given respectively by (3.5.1) and (3.5.2). Assume that there exists $K > 0$ such that $\min_{x\in\{1,\ldots,n\}} \lambda_{x-1} - \lambda_x + \nu_x - \nu_{x-1} \geq K$, and let $f \in$*

Lip$(\{0,1,\dots,n\})$. *Then for any initial state* $x \in \{0,1,\dots,n\}$*, any deviation level* $y > 0$ *and any* $t > 0$*, we have:*

$$\begin{aligned}&\mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right)\\&\quad\leq \exp\left(-\frac{(1-e^{-2Kt})\|\lambda+\nu\|_\infty}{2K} g\left(\frac{2Ky}{(1-e^{-2Kt})\|\lambda+\nu\|_\infty\|f\|_{\mathrm{Lip}}}\right)\right),\end{aligned}$$

where $g(u) = (1+u)\log(1+u) - u$*,* $u > 0$*.*
In particular, letting t *going to infinity in the above local inequality yields the deviation estimate under* π*:*

$$\pi\left(f - \mathbb{E}_\pi[f] \geq y\right) \leq \exp\left(\frac{y}{\|f\|_{\mathrm{Lip}}} - \left(\frac{y}{\|f\|_{\mathrm{Lip}}} + \frac{\|\lambda+\nu\|_\infty}{2K}\right)\log\left(1 + \frac{2Ky}{\|\lambda+\nu\|_\infty\|f\|_{\mathrm{Lip}}}\right)\right).$$

Proof. By Proposition 3.5.2, the Wasserstein curvature is bounded below by K. Therefore, it remains to apply Theorem 3.3.1 to get the result. ■

Under different assumptions on the transition rates of the generator, we get a somewhat similar estimate:

Theorem 3.5.12. *Let* $(X_t)_{t\geq 0}$ *be a birth-death process on* $\{0,1,\dots,n\}$ *with generator* $\mathcal{L}$ *and stationary distribution* π *given respectively by (3.5.1) and (3.5.2). Assume that there exists* $\rho > 0$ *such that* $\min_{x\in\{1,\dots,n-1\}}\min\{\lambda_{x-1} - \lambda_x, \nu_{x+1} - \nu_x\} \geq \rho$*. Let* $f \in$ Lip$(\{0,1,\dots,n\})$*. Then for any initial state* $x \in \{0,1,\dots,n\}$*, any deviation level* $y > 0$ *and any* $t > 0$*, we have:*

$$\begin{aligned}&\mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right)\\&\quad\leq \exp\left(-\frac{(1-e^{-2\rho t})(\lambda_0+\nu_n)}{2\rho} g\left(\frac{2\rho y}{(1-e^{-2\rho t})(\lambda_0+\nu_n)\|f\|_{\mathrm{Lip}}}\right)\right),\end{aligned}$$

where $g(u) = (1+u)\log(1+u) - u$*,* $u > 0$*.*
In particular, letting t *going to infinity in the above local inequality entails the following tail probability under the stationary distribution* π*:*

$$\pi\left(f - \mathbb{E}_\pi[f] \geq y\right) \leq \exp\left(\frac{y}{\|f\|_{\mathrm{Lip}}} - \left(\frac{y}{\|f\|_{\mathrm{Lip}}} + \frac{\lambda_0+\nu_n}{2\rho}\right)\log\left(1 + \frac{2\rho y}{(\lambda_0+\nu_n)\|f\|_{\mathrm{Lip}}}\right)\right).$$

Proof. By Proposition 3.5.6, the Γ-curvature is bounded below by ρ. Hence, applying Corollary 3.4.5 achieves the proof. ■

Remark 3.5.13. In order to obtain deviation bounds for stationary distributions, the positivity of lower bounds of discrete curvatures is crucial and thus does not allow us to extend such estimates to birth-death processes on the infinite state space $E = \mathbb{N}$, see the Remark 3.5.3 and Remark 3.5.7.
In particular, it excludes the $M/M/\infty$ queueing process recently investigated by D. Chafai in [25] and whose stationary distribution is the Poisson measure on $\mathbb{N}$. Therefore, we expect to recover the classical deviation inequality satisfied by the Poisson distribution

by taking the limit as $t \to +\infty$ in some appropriate local deviation inequalities, and such an interesting problem will be addressed in a forthcoming research. Note also that Theorem 3.5.9 is available for the $M/M/\infty$ queueing process, but such a result does not reflect the positive exact curvature of this queue emphasized in [25].

3.5.3 Ornstein-Uhlenbeck processes as fluid limits of rescaled Ehrenfest chains

In this part, we recover via Theorem 3.5.11 the optimal Gaussian concentration for an Ornstein-Uhlenbeck process constructed as a fluid limit of a rescaled continuous time Ehrenfest chain.
Given $n \in \mathbb{N}$, let $(X_t^n)_{t\geq 0}$ be the continuous time Ehrenfest chain on $\{0, 1, \dots, n\}$ starting from some $x_n \in \{0, 1, \dots, n\}$ and with generator given by:

$$\mathcal{L}_n f(x) = \lambda(n-x)\left(f(x+1) - f(x)\right) + \nu x \left(f(x-1) - f(x)\right), \quad x \in \{0, 1, \dots, n\},$$

where $0 < \lambda \leq \nu < 1$ are such that $\lambda + \nu = 1$.
Let $y(t) = \lambda + (y_0 - \lambda)e^{-t}$, $t > 0$, where $y_0 = \lim_{n\to+\infty} X_0^n/n$, and define for any $n \in \mathbb{N}\backslash\{0\}$ the process $(Z_t^n)_{t\geq 0}$ by $Z_t^n = (X_t^n - ny(t))/\sqrt{n}$, $t > 0$. Assume furthermore that the sequence of initial states $(Z_0^n)_{n\in\mathbb{N}}$ converges to z_0 (say).
By the central limit theorem in [37, Chapter 11], the sequence of processes $(Z_t^n)_{t\geq 0}$ converges as n goes to infinity to the process $(Z_t)_{t\geq 0}$ which is the unique solution of the equation

$$Z_t = z_0 + \int_0^t \sqrt{\lambda + (\nu - \lambda)y(s)}\,dB_s - \int_0^t Z_s ds, \quad t > 0,$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion.
In particular, if $y_0 = \lambda$, then $y(t) = \lambda$ for any $t > 0$ and $(Z_t)_{t\geq 0}$ rewrites as the Ornstein-Uhlenbeck process $(U_t)_{t\geq 0}$:

$$U_t = z_0 e^{-t} + \sqrt{2\lambda\nu}\int_0^t e^{-(t-s)} dB_s, \quad t > 0.$$

Now, fix $n \in \mathbb{N}\backslash\{0\}$ and time $t > 0$, and let $f \in \mathrm{Lip}(\mathbb{R})$. If h_n denotes the function $h_n = f \circ \phi_n$, where ϕ_n is defined on $\{0, 1, \dots, n\}$ by $\phi_n(x) = (x - n\lambda)/\sqrt{n}$, then $h_n \in \mathrm{Lip}(\{0, 1, \dots, n\})$ with constant at most $n^{-1/2}\|f\|_{\mathrm{Lip}}$. Therefore we can apply Theorem 3.5.11 to $(X_t^n)_{t\geq 0}$ and h_n, with $K = 1$, to get for any fixed $n \in \mathbb{N}\backslash\{0\}$, any deviation level $y > 0$ and any $t > 0$, the local deviation estimate:

$$\begin{aligned}
&\mathbb{P}_{x_n}\left(h_n(X_t^n) - \mathbb{E}_{x_n}\left[h_n(X_t^n)\right] \geq y\right) \\
&\leq \ \exp\left(-\frac{(1-e^{-2t})n\nu}{2} g\left(\frac{2\sqrt{n}y}{(1-e^{-2t})n\nu\|f\|_{\mathrm{Lip}}}\right)\right),
\end{aligned}$$

where $g(u) = (1+u)\log(1+u) - u$, $u > 0$. Finally, letting n going to infinity in the above inequality yields for any $y > 0$ and any $t > 0$ the classical Gaussian deviation:

$$\mathbb{P}_{z_0}\left(f(U_t) - \mathbb{E}_{z_0}\left[f(U_t)\right] \geq y\right) \leq \exp\left(-\frac{y^2}{(1-e^{-2t})\nu\|f\|_{\mathrm{Lip}}^2}\right),$$

see for instance Theorems 5.1 and 5.3 in [62].

3.5.4 A local inequality for samples of the $M/M/1$ queue

In this part, we give a local deviation estimate for sample vectors of the $M/M/1$ queueing process. Recall it is an irreducible birth-death process whose generator is given by

$$\mathcal{L}f(x) = \lambda\left(f(x+1) - f(x)\right) + \nu 1_{\{x\neq 0\}}\left(f(x-1) - f(x)\right), \quad x \in \mathbb{N},$$

where the positive numbers λ and ν correspond respectively to the input rate and service rate of the queue: the independent and identically distributed interarrival times and independent and identically distributed service times of the customers follow an exponential law with respective parameters λ and ν. The existence of an integration by parts formula for the associated semigroup together with a tensorization procedure of the Laplace transform allow us to provide with Theorem 3.5.14 below a local inequality for sample vectors of the $M/M/1$ queue.
We say in the sequel that a function $f : \mathbb{N}^d \to \mathbb{R}$ is ℓ^1-Lipschitz if

$$\|f\|_{\mathrm{Lip}(d)} = \sup_{x\neq y} \frac{|f(x) - f(y)|}{\|x-y\|_1} < +\infty,$$

where $\|\cdot\|_1$ denotes the ℓ^1-norm $\|z\|_1 = \sum_{i=1}^d |z_i|$, $z \in \mathbb{N}^d$.
Now, we can state the following

Theorem 3.5.14. *Let $(X_t)_{t\geq 0}$ be the $M/M/1$ queue with input and service rates λ and ν. Let f be ℓ^1-Lipschitz on $\mathbb{N}^d$ and consider the sample $X^d = (X_{t_1}, \dots, X_{t_d})$, $0 = t_0 < t_1 < \cdots < t_d = T$. Then for any initial state $x \in \mathbb{N}$ and any deviation level $y > 0$, we have the mulitdimensional local Poisson like deviation inequality:*

$$\begin{aligned}\mathbb{P}_x\left(f(X^d) - \mathbb{E}_x[f(X^d)] \geq y\right) &\leq \exp\left(-T(\lambda+\nu)g\left(\frac{y}{Td(\lambda+\nu)\|f\|_{\mathrm{Lip}(d)}}\right)\right) \qquad (3.5.8)\\ &\leq \exp\left(-\frac{y}{2d\|f\|_{\mathrm{Lip}(d)}}\log\left(1+\frac{y}{Td(\lambda+\nu)\|f\|_{\mathrm{Lip}(d)}}\right)\right),\end{aligned}$$

where $g(u) = (1+u)\log(1+u) - u$, $u > 0$.

Proof. Fix the initial state $x \in \mathbb{N}$. If u is a one dimensional Lipschitz function on $\mathbb{N}$ and $t > 0$, then rewriting the proof of Theorem 3.4.2 for the $M/M/1$ queue yields for any $\tau > 0$:

$$\mathbb{E}_x\left[e^{\tau u(X_t)}\right] \leq \exp\left\{\tau\mathbb{E}_x[u(X_t)] + h(\tau, t, \|u\|_{\mathrm{Lip}})\right\}, \qquad (3.5.9)$$

where h is the function defined on $(\mathbb{R}_+)^3$ by $h(\tau,t,z) = t(\lambda+\nu)\left(e^{\tau z} - \tau z - 1\right)$ and $\|\cdot\|_{\mathrm{Lip}}$ remains for the classical Lipschitz seminorm on $\mathbb{N}$.
To obtain a multidimensional version of (3.5.9), the idea is to tensorize the Laplace transform via an integration by parts formula satisfied by the semigroup $(P_t)_{t\geq 0}$ of the $M/M/1$

queueing process.
First, observe that we have the commutation relation $\mathcal{L}\mathrm{d}^+ = \mathrm{d}^+\mathcal{L}$, where d^+ is the forward gradient $\mathrm{d}^+ f(x) = f(x+1) - f(x)$, $x \in \mathbb{N}$. It implies $P_t \mathrm{d}^+ = \mathrm{d}^+ P_t$ for any non-negative t, which in turn entails for any bounded function u on $\mathbb{N}$ the integration by parts formula:

$$\sum_{y\in\mathbb{N}} u(y) P_t(x+1,y) = \sum_{y\in\mathbb{N}} u(y+1) P_t(x,y), \quad x \in \mathbb{N}. \tag{3.5.10}$$

Let f be bounded and ℓ^1-Lipschitz on $\mathbb{N}^d$. Set $f_d := f$ and define for any $k = 1, \ldots, d-1$, the function f_k on $\mathbb{N}^k$ by

$$f_k(x_1, \ldots, x_k) := \sum_{x_{k+1},\ldots,x_d \in\mathbb{N}} f(x_1, \ldots, x_k, \ldots, x_d) P_{t_{k+1}-t_k}(x_k, x_{k+1}) \cdots P_{t_d - t_{d-1}}(x_{d-1}, x_d).$$

Let $x_1, \ldots, x_{k-1}, y \in \mathbb{N}$. Using recursively (3.5.10), we have:

$$\begin{aligned}
f_k(x_1, \ldots, x_{k-1}, y+1) &= \sum_{x_d,\ldots,x_{k+1}\in\mathbb{N}} f(x_1, \ldots, x_{k-1}, y+1, x_{k+1}, x_{k+2}, \ldots, x_d) \\
&\quad \times P_{t_{k+1}-t_k}(y+1, x_{k+1}) P_{t_{k+2}-t_{k+1}}(x_{k+1}, x_{k+2}) \cdots P_{t_d-t_{d-1}}(x_{d-1}, x_d) \\
&= \sum_{x_d,\ldots,x_{k+1}\in\mathbb{N}} f(x_1, \ldots, x_{k-1}, y+1, x_{k+1}+1, x_{k+2}, \ldots, x_d) \\
&\quad \times P_{t_{k+1}-t_k}(y, x_{k+1}) P_{t_{k+2}-t_{k+1}}(x_{k+1}+1, x_{k+2}) \cdots P_{t_d-t_{d-1}}(x_{d-1}, x_d) \\
&= \cdots \\
&= \sum_{x_d,\ldots,x_{k+1}\in\mathbb{N}} f(x_1, \ldots, x_{k-1}, y+1, x_{k+1}+1, \ldots, x_{d-1}+1, x_d) \\
&\quad \times P_{t_{k+1}-t_k}(y, x_{k+1}) \cdots P_{t_{d-1}-t_{d-2}}(x_{d-2}, x_{d-1}) P_{t_d-t_{d-1}}(x_{d-1}+1, x_d) \\
&= \sum_{x_d,\ldots,x_{k+1}\in\mathbb{N}} f(x_1, \ldots, x_{k-1}, y+1, x_{k+1}+1, \ldots, x_{d-1}+1, x_d+1) \\
&\quad \times P_{t_{k+1}-t_k}(y, x_{k+1}) \cdots P_{t_{d-1}-t_{d-2}}(x_{d-2}, x_{d-1}) P_{t_d-t_{d-1}}(x_{d-1}, x_d).
\end{aligned}$$

Hence we obtain for any $k = 1, \ldots, d$, and any $x_1, \ldots, x_{k-1} \in \mathbb{N}$,

$$\begin{aligned}
&\|f_k(x_1, \ldots, x_{k-1}, \cdot)\|_{\mathrm{Lip}} \\
&= \sup_{y\in\mathbb{N}} |f_k(x_1, \ldots, x_{k-1}, y+1) - f_k(x_1, \ldots, x_{k-1}, y)| \\
&\leq \sup_{y\in\mathbb{N}} \sum_{x_d,\ldots,x_{k+1}\in\mathbb{N}} |f(x_1, \ldots, x_{k-1}, y+1, \ldots, x_d+1) - f(x_1, \ldots, x_{k-1}, y, \ldots, x_d)| \\
&\quad \times P_{t_{k+1}-t_k}(y, x_{k+1}) \cdots P_{t_{d-1}-t_{d-2}}(x_{d-2}, x_{d-1}) P_{t_d-t_{d-1}}(x_{d-1}, x_d) \\
&\leq (d-k+1)\|f\|_{\mathrm{Lip}(d)} \\
&\leq d\|f\|_{\mathrm{Lip}(d)}.
\end{aligned} \tag{3.5.11}$$

Using successively in the following lines the inequality (3.5.9) with the one-dimensional Lipschitz functions $x_k \mapsto f_k(*, x_k)$, $k = d, d-1, \ldots, 1$, and plugging the upper bound of

(3.5.11) into the right-hand-side of (3.5.9) since the function h is non-decreasing in its last variable, we apply the Markov property and we get

$$\begin{aligned}
&\mathbb{E}_x\left[e^{\tau f(X^d)}\right] \\
&= \sum_{x_1,\dots,x_{d-1}\in\mathbb{N}} \sum_{x_d\in\mathbb{N}} e^{\tau f_d(x_1,\dots,x_d)} P_{t_d-t_{d-1}}(x_{d-1},x_d)\cdots P_{t_1}(x,x_1) \\
&\le \exp\left\{h(\tau, t_d-t_{d-1}, d\|f\|_{\mathrm{Lip}(d)})\right\} \\
&\quad \times \sum_{x_1,\dots,x_{d-2}\in\mathbb{N}} \sum_{x_{d-1}\in\mathbb{N}} e^{\tau f_{d-1}(x_1,\dots,x_{d-1})} P_{t_{d-1}-t_{d-2}}(x_{d-2},x_{d-1})\cdots P_{t_1}(x,x_1) \\
&\le \exp\left\{h(\tau, t_d-t_{d-1}, d\|f\|_{\mathrm{Lip}(d)}) + h(\tau, t_{d-1}-t_{d-2}, d\|f\|_{\mathrm{Lip}(d)})\right\} \\
&\quad \times \sum_{x_1,\dots,x_{d-3}\in\mathbb{N}} \sum_{x_{d-2}\in\mathbb{N}} e^{\tau f_{d-2}(x_1,\dots,x_{d-2})} P_{t_{d-2}-t_{d-3}}(x_{d-3},x_{d-2})\cdots P_{t_1}(x,x_1) \\
&\le \cdots \\
&\le \exp\left(\sum_{k=1}^{d-1} h\left(\tau, t_{d-k+1}-t_{d-k}, d\|f\|_{\mathrm{Lip}(d)}\right)\right) \sum_{x_1\in\mathbb{N}} e^{\tau f_1(x_1)} P_{t_1}(x,x_1) \\
&\le \exp\left(\sum_{k=1}^{d} h\left(\tau, t_{d-k+1}-t_{d-k}, d\|f\|_{\mathrm{Lip}(d)}\right)\right) e^{\tau\sum_{x_1\in\mathbb{N}} f_1(x_1)P_{t_1}(x,x_1)} \\
&= \exp\left(\sum_{k=1}^{d} h\left(\tau, t_k-t_{k-1}, d\|f\|_{\mathrm{Lip}(d)}\right)\right) e^{\tau\mathbb{E}_x[f(X^d)]} \\
&= \exp\left\{\tau\mathbb{E}_x[f(X^d)] + T(\lambda+\nu)\left(e^{\tau d\|f\|_{\mathrm{Lip}(d)}} - \tau d\|f\|_{\mathrm{Lip}(d)} - 1\right)\right\}.
\end{aligned}$$

Dividing in both sides by $e^{\tau\mathbb{E}_x[f(X^d)]}$ and using the exponential Chebychev inequality achieve the proof in the bounded case. Finally, a classical argument allows us to remove the boundedness assumption on the function f. The proof is now complete. ∎

Remark 3.5.15. Note that Theorem 3.5.14 does not allow us to extend (3.5.8) to functionals on path spaces. Thus, it would be an interesting project to refine suitably (3.5.8) in terms of the increments $\Delta_i = t_i - t_{i-1}$, as $\Delta_i \to 0$.

Chapitre 4

A new Poisson-type deviation inequality for the empirical distribution of ergodic birth-death processes

Ce chapitre fait l'objet d'un article soumis pour publication.

Abstract

In this paper, we present a new Poisson-type deviation inequality of the empirical distribution of an ergodic birth-death process on $\mathbb{N}$, that generalizes the results of Lezaud [63, 64]. Our approach relies on the notion recently developed in [52] of Wasserstein curvatures of the process, which characterize contraction properties of the associated Markov semigroup on a suitable space of Lipschitz functions on $\mathbb{N}$.

4.1 Introduction

Let $(X_t)_{t\geq 0}$ be an ergodic Markov process on a Polish state space E, with stationary distribution π. The weak law of large numbers asserts that for any function $\phi \in L^1(\pi)$, the probability

$$\Lambda(t) := \mathbb{P}_x \left(\left| \frac{1}{t} \int_0^t \phi(X_s)ds - \int_E \phi d\pi \right| \geq y \right) \tag{4.1.1}$$

tends to 0 as t goes to infinity. Although large deviations theory gives an upper bound on the quantity $\limsup_{t\to+\infty} t^{-1} \log \Lambda(t)$, cf. [33], such an asymptotic estimate might be unsatisfactory for instance in the determination of confidence intervals, since one wants to control the convergence for fixed parameters. Actually, this problem has been raised and addressed by several authors. Using the Lumer-Philips theorem, Wu derived in

[82] a non-asymptotic estimate which is sharp for symmetric semigroups, but not really tractable. In the diffusion framework, various authors obtained explicit upper bounds on (4.1.1), provided the stationary distribution π satisfies some functional inequalities such as Poincaré, log-Sobolev or transportation type inequalities, see [24, 35, 40]. However, due to the discrete structure of countable state spaces, very few results are known in this area for continuous time Markov chains. To the author's knowledge, such a topic has been investigated only recently by Lezaud in his papers [63, 64]. Namely, using Kato's perturbation theory for linear operators, Lezaud established in [63] Poisson-type deviation bounds involving the spectral gap of the symmetrized generator of a continuous time Markov chain on a finite state space, and generalized in [64] such estimates to the countable state space case under boundedness assumptions on the function ϕ and on the transition rates of the generator.

The purpose of this paper is to present a new Poisson-type deviation inequality for the empirical distribution $t^{-1}\int_0^t \phi(X_s)ds$ when the process $(X_t)_{t\geq 0}$ is an ergodic birth-death process on $\mathbb{N}$. In particular, since our estimate is available for an unbounded Lipschitz function ϕ and for unbounded generators, we extend (and sharpen) in the case of birth-death processes the results obtained by Lezaud in [63, 64]. Our approach relies on the notion of Wasserstein curvature of continuous time Markov chains which characterizes contraction properties of the associated semigroup on a space of Lipschitz functions. For instance in the paper [52], a special emphasis is given to the curvature method in order to obtain deviation inequalities for the random variable $f(X_t)$, where f is a Lipschitz function with respect to the classical distance on $\mathbb{N}$, say d. However, the Wasserstein curvature associated to the metric d does not provide tail estimates involving some information in large time, hence on the stationary distribution. To correct this problem, the idea is to consider in the present paper the Wasserstein curvature related to a suitable distance, so that we are able to establish for the path functional $t^{-1}\int_0^t \phi(X_s)ds$ a convenient deviation bound that gives the correct order of convergence to equilibrium as the time parameter t goes to infinity.

The paper is organized as follows. Given an integer-valued ergodic birth-death process $(X_t)_{t\geq 0}$, we introduce in Section 4.2 its Wasserstein curvature related to a suitable metric on $\mathbb{N}$, say δ, and we provide in Proposition 4.2.6 some conditions on the associated generator under which the Wasserstein curvature above is bounded below by a positive constant. Under these criteria, we state in the second part of Section 4.2 our main contribution of the paper which is contained in Theorem 4.2.7, where a Poisson-type deviation bound is established for the empirical distribution $t^{-1}\int_0^t \phi(X_s)ds$, provided the function ϕ is Lipschitz with respect to the distance δ. In particular, no boundedness assumption is required on the transition rates of the generator. The whole Section 4.3 is devoted to the proof of Theorem 4.2.7, which is rather technical and is divided into several lemma. The key point of the proof corresponds to Lemma 4.3.3 with the tensorization of the Laplace transform in order to obtain a multi-dimensional estimate. Finally, the example of the $M/M/\infty$ queueing process is investigated in Section 4.4.

4.2 Preliminaries and main result

On a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$, we consider throughout the paper an irreducible ergodic birth-death process $\{(X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in\mathbb{N}}\}$ on the infinite state space $\mathbb{N} := \{0, 1, \ldots\}$. Such a process is a stable conservative continuous time Markov chain with generator given for any function $f : \mathbb{N} \to \mathbb{R}$ by

$$\mathcal{L}f(x) = \lambda_x \left(f(x+1) - f(x)\right) + \nu_x \left(f(x-1) - f(x)\right), \quad x \in \mathbb{N},$$

where the transition rates λ and ν are positive with $\nu_0 = 0$. Letting

$$\mu(0) = 1, \quad \mu(x) := \frac{\lambda_0\lambda_1 \cdots \lambda_{x-1}}{\nu_1\nu_2 \cdots \nu_x}, \quad x \geq 1,$$

the reversible stationary distribution π of the process $(X_t)_{t\geq 0}$ is given by

$$\pi(x) = \frac{\mu(x)}{\sum_{y\in\mathbb{N}} \mu(y)}, \quad x \in \mathbb{N}. \tag{4.2.1}$$

Denote $\mathbb{E}_x$ the expectation with respect to $\mathbb{P}_x$. The homogeneous semigroup $(P_t)_{t\geq 0}$ defined by

$$P_t f(x) := \mathbb{E}_x[f(X_t)] = \sum_{y\in\mathbb{N}} f(y)P_t(x, y), \quad t > 0, \quad x \in \mathbb{N},$$

is positivity preserving and contractive on every space $L^p(\pi)$, $p \in [1, +\infty]$.

Let ρ be a distance on $\mathbb{N}$ and define Lip_ρ the space of Lipschitz function $f : \mathbb{N} \to \mathbb{R}$ endowed with the seminorm

$$\|f\|_{\mathrm{Lip}_\rho} := \sup_{x,y\in\mathbb{N}} \frac{|f(x) - f(y)|}{\rho(x, y)} < +\infty.$$

If the stationary distribution satisfies the moment condition

$$\sum_{x\in\mathbb{N}} \rho(x, y)\pi(x) < +\infty, \quad y \in \mathbb{N}, \tag{4.2.2}$$

then the inclusion $\mathrm{Lip}_\rho \subset L^1(\pi)$ holds and the semigroup is well-defined on the space Lip_ρ.

Now, we recall from the paper [52] the definition of the ρ-Wasserstein curvature of the birth-death process $(X_t)_{t\geq 0}$.

Definition 4.2.1. *We assume that the stationary distribution π satisfies the moment condition (4.2.2). The ρ-Wasserstein curvature of the process $(X_t)_{t\geq 0}$ is defined for any $t > 0$ by*

$$\alpha_t := -\frac{1}{t} \sup \left\{ \log \left(\frac{\|P_t f\|_{\mathrm{Lip}_\rho}}{\|f\|_{\mathrm{Lip}_\rho}} \right) : f \in \mathrm{Lip}_\rho,\ f \neq \text{constante} \right\} \quad \in [-\infty, +\infty).$$

The ρ-Wasserstein curvature is said to be bounded below by $\alpha \in \mathbb{R}$ if $\inf_{t>0} \alpha_t \geq \alpha$. In other words, the semigroup $(P_t)_{t\geq 0}$ is contractive in the following sense: for any $f \in \mathrm{Lip}_\rho$ and any $t > 0$,

$$\|P_t f\|_{\mathrm{Lip}_\rho} \leq e^{-\alpha t} \|f\|_{\mathrm{Lip}_\rho}.$$

If the stationary distribution π satisfies the moment condition (4.2.2), then by the Kantorovich-Rubinstein duality theorem, see Theorem 5.10 in [29], the ρ-Wasserstein curvature of the process is bounded below by α if and only if we have

$$W_\rho(P_t(x,\cdot), P_t(y,\cdot)) \leq e^{-\alpha t} \rho(x,y), \quad x,y \in \mathbb{N}, \quad t > 0,$$

where $W_\rho(\cdot,\cdot)$ denotes the Wasserstein distance between probability measures on $\mathbb{N}$, endowed with the cost function ρ, that is

$$W_\rho(\mu,\nu) := \inf_\eta \sum_{x,y\in\mathbb{N}} \rho(x,y)\eta(x,y),$$

where the infimum runs over any probability measure η on $\mathbb{N}\times\mathbb{N}$ having marginals μ and ν. If α is positive, then the semigroup $(P_t)_{t\geq 0}$ converges exponentially fast to the stationary distribution π with respect to the metric W_ρ, cf. Theorem 5.23 in [29], and one deduces that the positivity of the Wasserstein curvatures is of crucial importance to study the ergodic properties of the process $(X_t)_{t\geq 0}$.

Denote d the classical distance on $\mathbb{N}$ given by

$$d(x,y) = |x-y|, \quad x,y \in \mathbb{N}.$$

Among the results of the article [52], a Poisson-type deviation inequality is established for birth-death processes with d-Wasserstein curvature bounded below by a non-positive constant. As noticed above, such a tail estimate is not convenient for large time and thus does not involve any information on the stationary distribution. To provide an estimate of the correct order as the time parameter is large, the idea is to consider the Wasserstein curvature related to a suitable distance δ on $\mathbb{N}$, that we introduce now. We mention that the metric δ defined below has been used by Chen in [28] to obtain variational formulae for the spectral gap of birth-death processes.

Definition 4.2.2. *Given a positive function u on $\mathbb{N}$, define the distance $\delta : \mathbb{N}\times\mathbb{N} \to [0,+\infty)$ as*

$$\delta(x,y) := \left| \sum_{k=0}^{x-1} u(k) - \sum_{k=0}^{y-1} u(k) \right|,$$

with the convention $\sum_{k=0}^{-1} u(k) = 0$.

We denote in the sequel $a \wedge b := \min\{a,b\}$ and $a \vee b := \max\{a,b\}$, $a,b \in \mathbb{R}$. Let us introduce a set of assumptions.

Assumption (A) *There exists two constants $K > 0$ and $C > 0$ such that*

$$\left(\inf_{x\geq 0} \lambda_x\right) \wedge \left(\inf_{x\geq 1} \nu_x\right) \geq K \quad \text{and} \quad u(x) \leq C \left(\frac{1}{\sqrt{\nu_{x+1}}} \wedge \frac{1}{\sqrt{\lambda_x}} \right), \quad x \in \mathbb{N}.$$

Assumption (B) *The stationary distribution π satisfies the moment condition (4.2.2) with the metric δ, and there exists a positive constant α such that*

$$\inf_{x\in\mathbb{N}} \left\{ \nu_{x+1} + \lambda_x - \nu_x \frac{u(x-1)}{u(x)} - \lambda_{x+1} \frac{u(x+1)}{u(x)} \right\} \geq \alpha. \tag{4.2.3}$$

Remark 4.2.3. Note that $u(-1)$ does not need to be defined in (4.2.3) since it is multiplied by $\nu_0 = 0$.

Under Assumption (A), we have a control on the distance δ as follows:

Lemma 4.2.4. *Under Assumption (A), the two inequalities below hold:*

(1) $\delta(x,y) \leq \frac{C}{\sqrt{K}}\, d(x,y)$, $x, y \in \mathbb{N}$;

(2) $\sup_{x\in\mathbb{N}} \lambda_x \delta(x,x+1)^2 + \nu_x \delta(x,x-1)^2 \leq 2C^2$.

Proof. Let $x, y \in \mathbb{N}$. If $x = y$, then $\delta(x,y) = 0 = d(x,y)$, so (1) is trivially true. Suppose $x \neq y$. Under Assumption (A), we have

$$\delta(x,y) = \left| \sum_{k=0}^{x-1} u(k) - \sum_{k=0}^{y-1} u(k) \right| = \sum_{k=x\wedge y}^{(x\vee y)-1} u(k) \leq \sum_{k=x\wedge y}^{(x\vee y)-1} \frac{C}{\sqrt{K}} = \frac{C}{\sqrt{K}} d(x,y).$$

Hence (1) is established.
To obtain (2), we use the second inequality of Assumption (A):

$$\sup_{x\in\mathbb{N}} \lambda_x \delta(x,x+1)^2 + \nu_x \delta(x,x-1)^2 = \sup_{x\in\mathbb{N}} \lambda_x u(x)^2 + \nu_x u(x-1)^2 \leq 2C^2.$$

The proof of Lemma 4.2.4 is complete. ∎

Remark 4.2.5. If at least one of the transition rates of the generator is unbounded, then the second inequality of Assumption (A) entails that u vanishes at infinity, so that the proper inclusion $\mathrm{Lip}_\delta \subsetneq \mathrm{Lip}_d$ holds. In particular, the identity function $f(x) = x$ is not Lipschitz on $\mathbb{N}$ with respect to the metric δ.

When the weight u is identically equal to 1, the metric δ is reduced to the classical distance d and by Proposition 5.1 in [52], the inequality (4.2.3) applied with $u \equiv 1$ implies that the d-Wasserstein curvature of the process is bounded below. Therefore, we expect in the general case that the inequality (4.2.3) of Assumption (B) is related to the δ-Wasserstein curvature of the process.

Proposition 4.2.6. *Under Assumption (B), the δ-Wasserstein curvature of the process is bounded below by α.*

Proof. Consider $(X_t^x)_{t\geq 0}$ and $(X_t^y)_{t\geq 0}$ two independent copies of $(X_t)_{t\geq 0}$, starting respectively from two different states $x, y \in \mathbb{N}$. Then the generator $\tilde{\mathcal{L}}$ of the two-dimensional process $(X_t^x, X_t^y)_{t\geq 0}$ is given for any real function f on $\mathbb{N} \times \mathbb{N}$ by

$$\tilde{\mathcal{L}} f(z,w) = (\mathcal{L} f(\cdot, w))(z) + (\mathcal{L} f(z, \cdot))(w), \quad z, w \in \mathbb{N}.$$

Let $z, w \in \mathbb{N}$, $z \neq w$. Using Assumption (B), we have

$$\begin{aligned}
\tilde{\mathcal{L}}\delta(z,w) &= (\mathcal{L}\delta(\cdot, w))(z) + (\mathcal{L}\delta(z,\cdot))(w) \\
&= \lambda_{z\vee w} u(z\vee w) - \lambda_{z\wedge w} u(z\wedge w) - \nu_{z\vee w} u((z\vee w)-1) + \nu_{z\wedge w} u((z\wedge w)-1) \\
&= \sum_{k=z\wedge w}^{(z\vee w)-1} (\lambda_{k+1} u(k+1) - \lambda_k u(k) - \nu_{k+1} u(k) + \nu_k u(k-1)) \\
&\leq -\alpha \sum_{k=z\wedge w}^{(z\vee w)-1} u(k) \\
&= -\alpha \delta(z,w). \qquad (4.2.4)
\end{aligned}$$

If $z = w$, then we have clearly $\tilde{\mathcal{L}}\delta(z,w) = 0 = -\alpha\delta(z,w)$, hence (4.2.4) is verified for any $z, w \in \mathbb{N}$. By Assumption (B), the stationary distribution π satisfies the moment condition (4.2.2) with the distance δ and the process $(\delta(X_t^x, X_t^y))_{t\geq 0}$ has finite expectation. Therefore, using Dynkin's formula together with the drift inequality (4.2.4), we have for any $t > 0$:

$$\begin{aligned}
\mathbb{E}[\delta(X_t^x, X_t^y)] &= \delta(x,y) + \mathbb{E}\int_0^t \tilde{\mathcal{L}}\delta(X_s^x, X_s^y) ds \\
&\leq \delta(x,y) - \alpha \int_0^t \mathbb{E}[\delta(X_s^x, X_s^y)] ds.
\end{aligned}$$

By Gronwall's lemma, we get

$$\mathbb{E}[\delta(X_t^x, X_t^y)] \leq e^{-\alpha t}\delta(x,y),$$

from which we obtain immediately, by the definition of the Wasserstein distance, the inequality

$$W_\delta(P_t(x,\cdot), P_t(y,\cdot)) \leq e^{-\alpha t}\delta(x,y).$$

Finally, as x, y and t are arbitrary, the Kantorovich-Rubinstein duality theorem entails that the δ-Wasserstein curvature of the process is bounded below by α. ∎

Now we are able to state the main result of this paper, whose proof is given in the next section. Denote in the sequel the function $g(u) := (1+u)\log(1+u) - u$, $u > 0$.

Theorem 4.2.7. *Under Assumptions (A) and (B), then for any Lipschitz function $\phi \in \mathrm{Lip}_\delta$, any $t > 0$, any initial state $x \in \mathbb{N}$ and any deviation level $y > 0$, we have the following Poisson-type deviation inequality:*

$$\mathbb{P}_x\left(\left| \frac{1}{t}\int_0^t (\phi(X_s) - \mathbb{E}_x[\phi(X_s)])\, ds \right| \geq y \right) \leq 2e^{-2Ktg\left(\frac{y\alpha}{2\sqrt{K}C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}}\right)} \qquad (4.2.5)$$

$$\leq \quad 2e^{-\frac{ty\alpha\sqrt{K}}{2C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}}\log\left(1+\frac{y\alpha}{2\sqrt{K}C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}}\right)}.$$

Remark 4.2.8. Denote $\pi(\phi) = \sum_{z\in\mathbb{N}} \phi(z)\pi(z)$. By invariance of the stationary distribution π together with Assumption (B), we have

$$\begin{aligned}
\left|\frac{1}{t}\int_0^t \mathbb{E}_x[\phi(X_s)]ds - \pi(\phi)\right| &= \left|\frac{1}{t}\int_0^t \sum_{z\in\mathbb{N}} (P_s\phi(x) - P_s\phi(z))\,\pi(z)ds\right| \\
&\leq \|\phi\|_{\mathrm{Lip}_\delta} \sum_{z\in\mathbb{N}} \delta(x,z)\pi(z)\,\frac{1}{t}\int_0^t e^{-\alpha s}ds \\
&= \|\phi\|_{\mathrm{Lip}_\delta} \sum_{z\in\mathbb{N}} \delta(x,z)\pi(z)\,\frac{1-e^{-\alpha t}}{t\alpha} \\
&=: M_t.
\end{aligned}$$

Hence we obtain for sufficiently large y,

$$\mathbb{P}_x\left(\left|\frac{1}{t}\int_0^t \phi(X_s)ds - \pi(\phi)\right| \geq y\right) \leq \mathbb{P}_x\left(\left|\frac{1}{t}\int_0^t (\phi(X_s) - \mathbb{E}_x[\phi(X_s)])\,ds\right| \geq y - M_t\right).$$

Therefore, the deviation inequality (4.2.5) gives an estimate of the speed of convergence of the empirical distribution to equilibrium, at the price of strengthening - slightly since $\lim_{t\to+\infty} M_t = 0$ - the range of the deviation level y.

Remark 4.2.9. The function $u \mapsto g(u)$ in Theorem 4.2.7 is equivalent to $u^2/2$ as u is close to 0 and to $u\log(u)$ as u tends to infinity. Hence, the Bennett-type inequality (4.2.5) exhibits a Gaussian tail for small values of the deviation level y, in accordance with the central limit theorem for Markov processes, and a Poisson tail for its large values. Therefore, we extend (and sharpen) in the case of birth-death processes the Chernoff inequality of Theorem 1.1 applied to Example 1.7 in [64], or that of Remark 2.6 in [64], since no boundedness assumption is required on the function ϕ and on the transition rates of the generator. Moreover, we point out that we also recover Theorem 3.4 in [63] when adapting the proof of Theorem 4.2.7 to the finite state space case. However, the price to pay here is to suppose Assumption (B), which is stronger than the existence of a spectral gap assumed by Lezaud in [63, 64]. See for instance Theorem 9.25 (1) in [29] for a comparison between these two criteria.

4.3 Proof of Theorem 4.2.7

This section is devoted to the proof of Theorem 4.2.7, which is divided into several lemma. First, we establish a convenient upper bound in large time on the Laplace transform of a Lipschitz function with respect to the distance δ, cf. Lemma 4.3.1. Using then a rather technical method of tensorization, the extension of such an estimate to the multi-dimensional case is considered in Lemma 4.3.3. Finally, with the help of the previous lemma and a suitable approximation of the empirical distribution, we finish the proof of Theorem 4.2.7.

4.3.1 A Laplace transform estimate

Let us start with an upper bound on the Laplace transform of a Lipschitz function with respect to the distance δ.

Lemma 4.3.1. *Suppose that Assumptions (A) and (B) are satisfied. Then for any Lipschitz function $f \in \mathrm{Lip}_\delta$, any $t > 0$, any initial state $x \in \mathbb{N}$ and any $\tau > 0$, we have the following estimate on the Laplace transform:*

$$\mathbb{E}_x\left[e^{\tau(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] \leq \exp\left\{\frac{K(1-e^{-2\alpha t})}{\alpha}\left(e^{\tau CK^{-1/2}\|f\|_{\mathrm{Lip}_\delta}} - \tau CK^{-1/2}\|f\|_{\mathrm{Lip}_\delta} - 1\right)\right\}. \tag{4.3.1}$$

Proof. Fix the initial state $x \in \mathbb{N}$, let $t > 0$, and assume first that f is bounded Lipschitz. The process $\left(Z_s^f\right)_{0\leq s\leq t}$ given by $Z_s^f := P_{t-s}f(X_s) - P_t f(X_0)$ is a real $\mathbb{P}_x$-martingale with respect to the truncated filtration $(\mathscr{F}_s)_{0\leq s\leq t}$ and we have by Itô's formula:

$$\begin{aligned} Z_s^f &= \int_0^s \left(P_{t-\tau}f(z+1) - P_{t-\tau}f(z)\right) 1_{\{X_{\tau-}=z\}} d(N_\tau^{(z,\uparrow)} - \lambda_z \tau) \\ &\quad + \int_0^s \left(P_{t-\tau}f(z-1) - P_{t-\tau}f(z)\right) 1_{\{X_{\tau-}=z\}} d(N_\tau^{(z,\downarrow)} - \nu_z \tau), \end{aligned}$$

where $\{(N_t^{(z,\uparrow)})_{t\geq 0} : z \in \mathbb{N}\}$ and $\{(N_t^{(z,\downarrow)})_{t\geq 0} : z \in \mathbb{N}\}$ are two independent families of independent Poisson processes on $\mathbb{R}_+$ with respective intensities $\lambda_z t$ and $\nu_z t$, $z \in \mathbb{N}$, $t > 0$. By Proposition 4.2.6, the δ-Wasserstein curvature of the process $(X_t)_{t\geq 0}$ is bounded below by $\alpha > 0$. Hence, the jumps of $\left(Z_s^f\right)_{0\leq s\leq t}$ satisfy

$$\begin{aligned} \sup_{0<s\leq t}\left|Z_s^f - Z_{s-}^f\right| &= \sup_{0<s\leq t}|P_{t-s}f(X_s) - P_{t-s}f(X_{s-})| \\ &\leq \|f\|_{\mathrm{Lip}_\delta} \sup_{0<s\leq t} e^{-\alpha(t-s)}\delta(X_s, X_{s-}) \\ &\leq \|f\|_{\mathrm{Lip}_\delta} \sup_{z\in\mathbb{N}} \delta(z, z+1) \\ &\leq \frac{C\|f\|_{\mathrm{Lip}_\delta}}{\sqrt{K}}, \end{aligned}$$

where in the last inequality we used the inequality (1) of Lemma 4.2.4. Moreover, the angle bracket process is bounded for any $s \in [0, t]$:

$$\begin{aligned} &\langle Z^f, Z^f\rangle_s \\ &= \int_0^s \left\{\lambda_{X_{\tau-}}\left(P_{t-\tau}f(X_{\tau-}+1) - P_{t-\tau}f(X_{\tau-})\right)^2 + \nu_{X_{\tau-}}\left(P_{t-\tau}f(X_{\tau-}-1) - P_{t-\tau}f(X_{\tau-})\right)^2\right\} d\tau \\ &\leq \|f\|_{\mathrm{Lip}_\delta}^2 \int_0^s e^{-2\alpha(t-\tau)}\left\{\lambda_{X_{\tau-}}\delta(X_{\tau-}, X_{\tau-}+1)^2 + \nu_{X_{\tau-}}\delta(X_{\tau-}, X_{\tau-}-1)^2\right\} d\tau \\ &\leq \frac{C^2(1-e^{-2\alpha t})\|f\|_{\mathrm{Lip}_\delta}^2}{\alpha}, \end{aligned} \tag{4.3.2}$$

where in the last inequality we used the estimate (2) of Lemma 4.2.4.
Define the function ψ on $\mathbb{R}_+$ by $\psi(z) = z^{-2}(e^z - z - 1)$. By Lemma 23.19 in [56], the process $(Y_s^{(\tau)})_{0\le s\le t}$ given for any $\tau > 0$ by

$$Y_s^{(\tau)} := \exp\left\{\tau Z_s^f - \tau^2 \psi\left(\tau C K^{-1/2}\|f\|_{\mathrm{Lip}_\delta}\right)\langle Z^f, Z^f\rangle_s\right\}$$

is a $\mathbb{P}_x$-supermartingale with respect to $(\mathscr{F}_s)_{0\le s\le t}$. Thus, we get for any $\tau > 0$:

$$\begin{aligned}
\mathbb{E}_x\left[e^{\tau(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] &= \mathbb{E}_x\left[e^{\tau Z_t^f}\right] \\
&\le \exp\left\{\frac{\tau^2 C^2(1-e^{-2\alpha t})\|f\|^2_{\mathrm{Lip}_\delta}}{\alpha}\psi\left(\tau C K^{-1/2}\|f\|_{\mathrm{Lip}_\delta}\right)\right\}\mathbb{E}_x\left[Y_t^{(\tau)}\right] \\
&\le \exp\left\{\frac{\tau^2 C^2(1-e^{-2\alpha t})\|f\|^2_{\mathrm{Lip}_\delta}}{\alpha}\psi\left(\tau C K^{-1/2}\|f\|_{\mathrm{Lip}_\delta}\right)\right\} \\
&= \exp\left\{\frac{K(1-e^{-2\alpha t})}{\alpha}\left(e^{\tau C K^{-1/2}\|f\|_{\mathrm{Lip}_\delta}} - \tau C K^{-1/2}\|f\|_{\mathrm{Lip}_\delta} - 1\right)\right\}.
\end{aligned}$$

The Laplace transform estimate is established in the bounded case.

To remove the boundedness assumption on the function f, the argument is standard and is given for completeness. Let $f \in \mathrm{Lip}_\delta$ and consider the bounded function $f_n = \max\{-n, \min\{f, n\}\}$, $n \in \mathbb{N}$, which converges to f by construction. We have clearly $\|f_n\|_{\mathrm{Lip}_\delta} \le \|f\|_{\mathrm{Lip}_\delta}$. Let us show that the sequence of random variables $(f_n(X_t))_{n\in\mathbb{N}}$ converges to $f(X_t)$ in $L^1(\mathbb{P}_x)$.
Let y_0 be such that $2C^2(1-e^{-2\alpha t})\|f\|^2_{\mathrm{Lip}_\delta} < \alpha y_0^2$. By Chebychev's inequality,

$$\begin{aligned}
\mathbb{P}_x\left(|f_n(X_t) - \mathbb{E}_x[f_n(X_t)]| \ge y_0\right) &\le \frac{\mathbb{E}_x\left[|f_n(X_t) - \mathbb{E}_x[f_n(X_t)]|^2\right]}{y_0^2} \\
&= \frac{\mathbb{E}_x\left[\langle Z^{f_n}, Z^{f_n}\rangle_t\right]}{y_0^2} \\
&\le \frac{C^2(1-e^{-2\alpha t})\|f_n\|^2_{\mathrm{Lip}_\delta}}{y_0^2\alpha} \\
&\le \frac{C^2(1-e^{-2\alpha t})\|f\|^2_{\mathrm{Lip}_\delta}}{y_0^2\alpha} \\
&< \frac{1}{2}, \qquad (4.3.3)
\end{aligned}$$

where in the third line we used the bound (4.3.2) applied to the function f_n.
On the other hand, let z be such that $\mathbb{P}_x(|f(X_t)| \le z) \ge 3/4$ and let $n_0 \in \mathbb{N}$ be such that for any $n \ge n_0$, we have $\mathbb{P}_x(|f_n(X_t) - f(X_t)| \ge 1) \le 1/4$. Hence for any $n \ge n_0$,

$$\begin{aligned}
\mathbb{P}_x\left(|f_n(X_t)| \le z+1\right) &\ge \mathbb{P}_x\left(|f(X_t)| \le z\right) - \mathbb{P}_x\left(|f(X_t)| \le z\,;\, |f_n(X_t)| > z+1\right) \\
&\ge \frac{3}{4} - \mathbb{P}_x\left(|f_n(X_t) - f(X_t)| \ge 1\right)
\end{aligned}$$

$$\geq \frac{1}{2}. \tag{4.3.4}$$

Thus, (4.3.3) and (4.3.4) entail for any $n \geq n_0$ the bound $|\mathbb{E}_x[f_n(X_t)]| \leq y_0 + z + 1$. Now, together with the help of (4.3.2) applied again to f_n, observe that for any $n \geq n_0$,

$$\begin{aligned}
\mathbb{E}_x\left[f_n(X_t)^2\right] &= \mathbb{E}_x\left[|f_n(X_t) - \mathbb{E}_x[f_n(X_t)]|^2\right] + \mathbb{E}_x[f_n(X_t)]^2 \\
&= \mathbb{E}_x\left[\langle Z^{f_n}, Z^{f_n}\rangle_t\right] + \mathbb{E}_x[f_n(X_t)]^2 \\
&\leq \frac{C^2(1-e^{-2\alpha t})\|f_n\|^2_{\mathrm{Lip}_\delta}}{\alpha} + (y_0+z+1)^2 \\
&\leq \frac{C^2(1-e^{-2\alpha t})\|f\|^2_{\mathrm{Lip}_\delta}}{\alpha} + (y_0+z+1)^2,
\end{aligned}$$

which implies $\sup_{n\in\mathbb{N}} \mathbb{E}_x\left[f_n(X_t)^2\right] < +\infty$. Therefore, the uniform integrability is verified and the sequence of random variables $(f_n(X_t))_{n\in\mathbb{N}}$ converges in $L^1(\mathbb{P}_x)$ to $f(X_t)$. Finally, applying (4.3.1) to f_n and using Fatou's lemma, we obtain:

$$\begin{aligned}
&\mathbb{E}_x\left[e^{\tau f(X_t)}\right] \\
&\leq \liminf_{n\to+\infty} \mathbb{E}_x\left[e^{\tau f_n(X_t)}\right] \\
&\leq \liminf_{n\to+\infty} e^{\tau \mathbb{E}_x[f_n(X_t)]} \exp\left\{\frac{K(1-e^{-2\alpha t})}{\alpha}\left(e^{\tau C K^{-1/2}\|f_n\|_{\mathrm{Lip}_\delta}} - \tau C K^{-1/2}\|f_n\|_{\mathrm{Lip}_\delta} - 1\right)\right\} \\
&\leq \liminf_{n\to+\infty} e^{\tau \mathbb{E}_x[f_n(X_t)]} \exp\left\{\frac{K(1-e^{-2\alpha t})}{\alpha}\left(e^{\tau C K^{-1/2}\|f\|_{\mathrm{Lip}_\delta}} - \tau C K^{-1/2}\|f\|_{\mathrm{Lip}_\delta} - 1\right)\right\} \\
&= e^{\tau \mathbb{E}_x[f(X_t)]} \exp\left\{\frac{K(1-e^{-2\alpha t})}{\alpha}\left(e^{\tau C K^{-1/2}\|f\|_{\mathrm{Lip}_\delta}} - \tau C K^{-1/2}\|f\|_{\mathrm{Lip}_\delta} - 1\right)\right\},
\end{aligned}$$

where we used in the last inequality that the function $z \mapsto e^z - z - 1$ is non-decreasing on $\mathbb{R}_+$. ■

Remark 4.3.2. The Laplace transform estimate (4.3.1) allows us to sharpen in large time the deviation inequalities given in [52] for birth-death processes on $\mathbb{N}$. Indeed, we get easily from the Chebychev inequality and (4.3.1):

$$\begin{aligned}
\mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right) &\leq \inf_{\tau>0} e^{-\tau y}\mathbb{E}_x\left[e^{\tau(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] \\
&\leq e^{-\frac{K(1-e^{-2\alpha t})}{\alpha} g\left(\frac{\alpha y}{C\sqrt{K}(1-e^{-2\alpha t})\|f\|_{\mathrm{Lip}_\delta}}\right)}, \quad y > 0,
\end{aligned}$$

where $g(u) := (1+u)\log(1+u) - u$, $u > 0$. In particular, letting t tend to infinity in the latter inequality entails the estimate under the stationary distribution π:

$$\pi\left(f - \pi(f) \geq y\right) \leq e^{\frac{y\sqrt{K}}{C\|f\|_{\mathrm{Lip}_\delta}} - \left(\frac{K}{\alpha} + \frac{y\sqrt{K}}{C\|f\|_{\mathrm{Lip}_\delta}}\right)\log\left(1+\frac{\alpha y}{C\sqrt{K}\|f\|_{\mathrm{Lip}_\delta}}\right)}.$$

However, in contrast to the deviation inequalities given in [52], the price to pay here is to require a stronger assumption on the function f, namely $f \in \mathrm{Lip}_\delta$.

4.3.2 Tensorization of the Laplace transform

This part is devoted to the extension to the multi-dimensional case of the Laplace transform estimate (4.3.1), by using the method of tensorization. Such an approach may be seen as the continuous time analogous of the argument used in the articles [71], [35], to establish Gaussian concentration inequalities for weakly dependent sequences.
Given $n \in \mathbb{N} \setminus \{0,1\}$, define $\mathrm{Lip}_\delta(n)$ the space of real Lipschitz functions on the product space $\mathbb{N}^n$, endowed with the seminorm

$$\|f\|_{\mathrm{Lip}_\delta(n)} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{\delta_n(x,y)} < +\infty,$$

where δ_n is the ℓ^1-distance on $\mathbb{N}^n$ with respect to the metric δ, i.e. $\delta_n(y,z) := \sum_{i=1}^n \delta(y_i, z_i)$, $y, z \in \mathbb{N}^n$.

Lemma 4.3.3. *Suppose that Assumptions (A) and (B) are satisfied. Denote the sample $X^n = (X_{t_1}, \dots, X_{t_n})$, $0 = t_0 < t_1 < \cdots < t_n$, and let $f \in \mathrm{Lip}_\delta(n)$. Then for any initial state $x \in \mathbb{N}$ and any $\tau > 0$, we have the Laplace transform estimate:*

$$\mathbb{E}_x \left[e^{\tau(f(X^n) - \mathbb{E}_x[f(X^n)])} \right] \leq \exp \left\{ \sum_{k=1}^n h\left(\tau, t_k - t_{k-1}, M_k C K^{-1/2} \|f\|_{\mathrm{Lip}_\delta(n)}\right) \right\}, \tag{4.3.5}$$

where $M_k := \sum_{l=k}^n e^{-\alpha(t_l - t_k)}$ and h is the function defined on $(\mathbb{R}_+)^3$ by

$$h(\tau, t, z) := \frac{K\left(1 - e^{-2\alpha t}\right)}{\alpha} \left(e^{\tau z} - \tau z - 1\right).$$

Proof. Let $f_n := f$ and define for any $k = 1, \dots, n-1$, the function f_k on $\mathbb{N}^k$ by

$$\begin{aligned} f_k(x_1, \dots, x_k) &:= \sum_{x_{k+1}, \dots, x_n \in \mathbb{N}} f(x_1, \dots, x_n) P_{t_n - t_{n-1}}(x_{n-1}, x_n) \cdots P_{t_{k+1} - t_k}(x_k, x_{k+1}) \\ &= \sum_{x_{k+1} \in \mathbb{N}} f_{k+1}(x_1, \dots, x_k, x_{k+1}) P_{t_{k+1} - t_k}(x_k, x_{k+1}). \end{aligned}$$

We divide the proof of Lemma 4.3.3 into two parts.

• Step 1 : By a downward recursive argument on k, let us show first that the one-dimensional function $x_k \mapsto f_k(*, x_k)$ is Lipschitz with respect to the distance δ, with furthermore the inequality

$$\sup_{x_1, \dots, x_{k-1} \in \mathbb{N}} \|f_k(x_1, \dots, x_{k-1}, \cdot)\|_{\mathrm{Lip}_\delta} \leq M_k \|f\|_{\mathrm{Lip}_\delta(n)}. \tag{4.3.6}$$

Since $M_n = 1$, the property (4.3.6) is trivially true for $k = n$.
Suppose now that (4.3.6) is satisfied for some $k \in \{2, \dots, n-1\}$.
First, letting $x_1, \dots, x_{k-2}, y, z, x_k \in \mathbb{N}$, we have:

$|f_k(x_1, \dots, x_{k-2}, y, x_k) - f_k(x_1, \dots, x_{k-2}, z, x_k)|$

$$\begin{aligned}
&= \Bigg| \sum_{x_{k+1},\ldots,x_n\in\mathbb{N}} f(x_1,\ldots,x_{k-2},y,x_k,x_{k+1},\ldots,x_n) P_{t_n-t_{n-1}}(x_{n-1},x_n)\cdots P_{t_{k+1}-t_k}(x_k,x_{k+1}) \\
&\quad - \sum_{x_{k+1},\ldots,x_n\in\mathbb{N}} f(x_1,\ldots,x_{k-2},z,x_k,x_{k+1},\ldots,x_n) P_{t_n-t_{n-1}}(x_{n-1},x_n)\cdots P_{t_{k+1}-t_k}(x_k,x_{k+1}) \Bigg| \\
&\le \|f\|_{\mathrm{Lip}_\delta(n)} \delta(y,z) \sum_{x_{k+1},\ldots,x_n\in\mathbb{N}} P_{t_n-t_{n-1}}(x_{n-1},x_n)\cdots P_{t_{k+1}-t_k}(x_k,x_{k+1}) \\
&= \|f\|_{\mathrm{Lip}_\delta(n)} \delta(y,z),
\end{aligned}$$

from which follows the inequality

$$\sup_{x_1,\ldots,x_{k-2},x_k\in\mathbb{N}} \|f_k(x_1,\ldots,x_{k-2},\cdot,x_k)\|_{\mathrm{Lip}_\delta} \le \|f\|_{\mathrm{Lip}_\delta(n)}. \tag{4.3.7}$$

Now, let us show that the property (4.3.6) is satisfied at the step $k-1$ with the help of (4.3.7).
Let $x_1,\ldots,x_{k-2},y,z\in\mathbb{N}$. By Proposition 4.2.6, the δ-curvature of the process is bounded below by α. Using this argument in the second inequality below, we have:

$$\begin{aligned}
&|f_{k-1}(x_1,\ldots,x_{k-2},y) - f_{k-1}(x_1,\ldots,x_{k-2},z)| \\
&= \left| \sum_{x_k\in\mathbb{N}} f_k(x_1,\ldots,x_{k-2},y,x_k) P_{t_k-t_{k-1}}(y,x_k) - \sum_{x_k\in\mathbb{N}} f_k(x_1,\ldots,x_{k-2},z,x_k) P_{t_k-t_{k-1}}(z,x_k) \right| \\
&\le \left| \sum_{x_k\in\mathbb{N}} f_k(x_1,\ldots,x_{k-2},y,x_k) (P_{t_k-t_{k-1}}(y,x_k) - P_{t_k-t_{k-1}}(z,x_k)) \right| \\
&\quad + \sum_{x_k\in\mathbb{N}} |f_k(x_1,\ldots,x_{k-2},y,x_k) - f_k(x_1,\ldots,x_{k-2},z,x_k)| \, P_{t_k-t_{k-1}}(z,x_k) \\
&\le e^{-\alpha(t_k-t_{k-1})} \|f_k(x_1,\ldots,x_{k-2},y,\cdot)\|_{\mathrm{Lip}_\delta} \delta(y,z) \\
&\quad + \sum_{x_k\in\mathbb{N}} \|f_k(x_1,\ldots,x_{k-2},\cdot,x_k)\|_{\mathrm{Lip}_\delta} \delta(y,z) P_{t_k-t_{k-1}}(z,x_k) \\
&\le \left(1 + M_k e^{-\alpha(t_k-t_{k-1})}\right) \|f\|_{\mathrm{Lip}_\delta(n)} \delta(y,z) \\
&= M_{k-1} \|f\|_{\mathrm{Lip}_\delta(n)} \delta(y,z),
\end{aligned}$$

where in the last inequality we used the assumption (4.3.6) at the step k together with (4.3.7). Therefore, we obtain the inequality

$$\|f_{k-1}(x_1,\ldots,x_{k-2},\cdot)\|_{\mathrm{Lip}_\delta} \le M_{k-1} \|f\|_{\mathrm{Lip}_\delta(n)},$$

and the parameters $x_1,\ldots,x_{k-2}$ being arbitrary, the property (4.3.6) is established at the step $k-1$, hence in full generality.

• Step 2 : Proof of the Laplace transform estimate (4.3.5).

Using successively in the following lines Lemma 4.3.1 with the one-dimensional Lipschitz functions $x_k \mapsto f_k(*, x_k)$, $k = n, n-1, \dots, 1$, and plugging the upper bound of (4.3.6) into the right-hand-side of (4.3.1) since the function h is non-decreasing in its last variable, we apply the Markov property and we get

$$
\begin{aligned}
&\mathbb{E}_x\left[e^{\tau f(X^n)}\right] \\
&= \sum_{x_1,\dots,x_{n-1}\in\mathbb{N}} \sum_{x_n\in\mathbb{N}} e^{\tau f_n(x_1,\dots,x_n)} P_{t_n-t_{n-1}}(x_{n-1},x_n)\cdots P_{t_1}(x,x_1) \\
&\le \exp\left\{h(\tau, t_n - t_{n-1}, CK^{-1/2}\|f\|_{\mathrm{Lip}_\delta(n)})\right\} \\
&\quad \times \sum_{x_1,\dots,x_{n-2}\in\mathbb{N}} \sum_{x_{n-1}\in\mathbb{N}} e^{\tau f_{n-1}(x_1,\dots,x_{n-1})} P_{t_{n-1}-t_{n-2}}(x_{n-2},x_{n-1})\cdots P_{t_1}(x,x_1) \\
&\le \exp\left\{h(\tau, t_n - t_{n-1}, CK^{-1/2}\|f\|_{\mathrm{Lip}_\delta(n)}) + h(\tau, t_{n-1} - t_{n-2}, CK^{-1/2}M_{n-1}\|f\|_{\mathrm{Lip}_\delta(n)})\right\} \\
&\quad \times \sum_{x_1,\dots,x_{n-3}\in\mathbb{N}} \sum_{x_{n-2}\in\mathbb{N}} e^{\tau f_{n-2}(x_1,\dots,x_{n-2})} P_{t_{n-2}-t_{n-3}}(x_{n-3},x_{n-2})\cdots P_{t_1}(x,x_1) \\
&\le \cdots \\
&\le \exp\left(\sum_{k=1}^{n-1} h\left(\tau, t_{n-k+1} - t_{n-k}, CK^{-1/2}M_{n-k+1}\|f\|_{\mathrm{Lip}_\delta(n)}\right)\right) \sum_{x_1\in\mathbb{N}} e^{\tau f_1(x_1)} P_{t_1}(x,x_1) \\
&\le \exp\left(\sum_{k=1}^{n} h\left(\tau, t_{n-k+1} - t_{n-k}, CK^{-1/2}M_{n-k+1}\|f\|_{\mathrm{Lip}_\delta(n)}\right)\right) e^{\tau \sum_{x_1\in\mathbb{N}} f_1(x_1) P_{t_1}(x,x_1)} \\
&= \exp\left(\sum_{k=1}^{n} h\left(\tau, t_k - t_{k-1}, CK^{-1/2}M_k\|f\|_{\mathrm{Lip}_\delta(n)}\right)\right) e^{\tau \mathbb{E}_x[f(X^n)]}.
\end{aligned}
$$

The proof of Lemma 4.3.3 is complete. ■

4.3.3 Proof of Theorem 4.2.7

Now we are able to prove Theorem 4.2.7.

Proof of Theorem 4.2.7. We use the notation of Section 4.3.2. Fix the initial state $x \in \mathbb{N}$ of the birth-death process $(X_t)_{t\ge 0}$ and a finite time horizon $t > 0$. Define $t_k = kt/n$, $k = 0, \dots, n$, a regular subdivision of the time interval $[0, t]$ and let the sample $X^n = (X_{t_1}, \dots, X_{t_n})$. The function f given by

$$
f(z) := \frac{1}{n}\sum_{k=1}^{n} \phi(z_k), \quad z = (z_1, \dots, z_n),
$$

is Lipschitz on the product space $\mathbb{N}^n$ with respect to the ℓ^1-metric δ_n and its Lipschitz seminorm satisfies the bound

$$
\|f\|_{\mathrm{Lip}_\delta(n)} \le \frac{1}{n}\|\phi\|_{\mathrm{Lip}_\delta}.
$$

Hence, since we have

$$\sup_{k=1,\dots,n} M_k = \sup_{k=1,\dots,n} \sum_{l=k}^{n} e^{-\alpha t(l-k)/n} = \frac{1-e^{-\alpha t}}{1-e^{-\alpha t/n}},$$

and that the function h is non-decreasing in its last variable, Lemma 4.3.3 entails for any $\tau > 0$:

$$\mathbb{E}_x \left[e^{\tau(f(X^n)-\mathbb{E}_x[f(X^n)])} \right] \leq \exp \left\{ nh \left(\tau, \frac{t}{n}, \frac{C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}}{n\sqrt{K}(1-e^{-\alpha t/n})} \right) \right\}.$$

By Chebychev's inequality, we get for any $y > 0$:

$$\begin{aligned} \mathbb{P}_x \left(f(X^n) - \mathbb{E}_x[f(X^n)] \geq y \right) &\leq \inf_{\tau>0} e^{-\tau y}\, \mathbb{E}_x \left[e^{\tau(f(X^n)-\mathbb{E}_x[f(X^n)])} \right] \\ &\leq e^{-\frac{nK}{\alpha}(1-e^{-2\alpha t/n}) g\left(\frac{y\alpha(1-e^{-\alpha t/n})}{C\sqrt{K}(1-e^{-2\alpha t/n})(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}} \right)}, \end{aligned}$$

where we recall $g(u) := (1+u)\log(1+u) - u$, $u > 0$. Applying also the same reasoning to the function $-f$ yields

$$\begin{aligned} \mathbb{P}_x \left(|f(X^n) - \mathbb{E}_x[f(X^n)]| \geq y \right) &\leq 2e^{-\frac{nK}{\alpha}(1-e^{-2\alpha t/n}) g\left(\frac{y\alpha(1-e^{-\alpha t/n})}{C\sqrt{K}(1-e^{-2\alpha t/n})(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}} \right)} \\ &=: 2e^{-A_n}. \end{aligned} \tag{4.3.8}$$

Now, the Riemann sum $f(X^n) = n^{-1}\sum_{k=1}^{n} \phi(X_{kt/n})$ converges $\mathbb{P}_x$-a.s. to the empirical distribution $t^{-1}\int_0^t \phi(X_s)ds$, and up to a slight change in the end of the proof of Lemma 4.3.1, we can show that the convergence also holds in $L^1(\mathbb{P}_x)$. Finally, using Fatou's lemma and the estimate (4.3.8), we obtain

$$\begin{aligned} \mathbb{P}_x \left(\left| \frac{1}{t}\int_0^t \left(\phi(X_s) - \mathbb{E}_x[\phi(X_s)]\right) ds \right| \geq y \right) &\leq \liminf_{n\to+\infty} \mathbb{P}_x \left(|f(X^n) - \mathbb{E}_x[f(X^n)]| \geq y \right) \\ &\leq \liminf_{n\to+\infty} 2e^{-A_n} \\ &= 2e^{-2Ktg\left(\frac{y\alpha}{2C\sqrt{K}(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}} \right)}. \end{aligned}$$

The proof of Theorem 4.2.7 is established. ∎

Remark 4.3.4. The problem to find a similar rate of convergence as (4.2.5) for the identity function $\phi_0(x) = x$ is unsolved when the generator of the birth-death process is unbounded. Indeed, our estimate is only available for Lipschitz functions in the space Lip_δ, which excludes ϕ_0 as noticed in Remark 4.2.5.

4.4 Application to the $M/M/\infty$ queueing process

Consider a ticket booth system where each customer arriving in the queue is immediately served. Denoting X_t the number of busy servers - the length of the queue - at time $t > 0$, we assume that the customers' arrival process is a Poisson process of intensity $\lambda > 0$ and that conditionally on the event $\{X_s = x\}$, the service time $T := \inf\{t > s : X_t \neq X_s\} - s$ follows an exponential distribution with parameter $\lambda + \nu x$, $\nu > 0$. The stochastic process $(X_t)_{t\geq 0}$ is called a $M/M/\infty$ queueing process. This is an ergodic birth-death process with generator given for any $f : \mathbb{N} \to \mathbb{R}$ by

$$\mathcal{L}f(x) = \lambda(f(x+1) - f(x)) + \nu x(f(x-1) - f(x)), \quad x \in \mathbb{N}.$$

By (4.2.1), the stationary distribution of $(X_t)_{t\geq 0}$ is the Poisson measure $\mathscr{P}(\sigma)$ on $\mathbb{N}$ with parameter $\sigma := \lambda/\nu$, i.e.

$$\mathscr{P}(\sigma)(x) = e^{-\sigma}\frac{\sigma^x}{x!}, \quad x \in \mathbb{N}.$$

For the sake of simplicity, we assume in the sequel that the process is normalized, i.e. $\lambda = \nu$. Denote $\mathscr{B}(n,p)$ the binomial distribution with parameters $n \in \mathbb{N}$ and $p \in (0,1)$. The knowledge of the distribution at time $t > 0$ of the $M/M/\infty$ queueing process allows us to make explicit computations. Indeed, by the Mehler-type convolution formula given for instance by Chafaï in [25]:

$$\mathcal{L}(X_t | X_0 = x) = \mathscr{B}\left(x, e^{-\nu t}\right) * \mathscr{P}\left(1 - e^{-\nu t}\right), \quad t > 0,$$

we get for any $\tau > 0$,

$$\begin{aligned} \mathbb{E}_x\left[e^{\tau(X_t - \mathbb{E}_x[X_t])}\right] &= \exp\left\{x \log\left(1 + e^{-\nu t}(e^\tau - 1)\right) - \tau x e^{-\nu t} + (1 - e^{-\nu t})(e^\tau - \tau - 1)\right\} \\ &\leq \exp\left\{\left(xe^{-\nu t} + 1 - e^{-\nu t}\right)(e^\tau - \tau - 1)\right\} \\ &= \exp\left\{\mathbb{E}_x[X_t]\,(e^\tau - \tau - 1)\right\}, \end{aligned} \tag{4.4.1}$$

where we used the inequality $\log(1+x) \leq x$, $x > 0$. Thus, by Chebychev's inequality, we obtain for any $y > 0$ the deviation inequality

$$\begin{aligned} \mathbb{P}_x\left(X_t - \mathbb{E}_x[X_t] \geq y\right) &\leq \inf_{\tau > 0} e^{-\tau y}\, \mathbb{E}_x\left[e^{\tau(X_t - \mathbb{E}_x[X_t])}\right] \\ &\leq \exp\left\{y - (\mathbb{E}_x[X_t] + y)\log\left(1 + \frac{y}{\mathbb{E}_x[X_t]}\right)\right\}, \end{aligned} \tag{4.4.2}$$

which entails by ergodicity as $t \to +\infty$ the classical tail estimate for a Poisson random variable X with intensity 1, that is

$$\mathbb{P}\left(X - \mathbb{E}[X] \geq y\right) \leq \exp\left\{y - (1+y)\log(1+y)\right\}.$$

Therefore, we expect that the Poisson-type deviation inequality (4.4.2) may be extended to the empirical distribution of the $M/M/\infty$ queueing process, a question to which we turn now.

Choosing the positive function $u(x) := (x+1)^{-1/2}$, $x \in \mathbb{N}$, in the definition of the distance δ, the transition rates of the generator satisfy Assumption (A) with the constants $C = \sqrt{K} = \sqrt{\nu}$. Moreover, a short computation shows that Assumption (B) is also verified with $\alpha = \nu/2 > 0$, which is the half of the exact curvature of the $M/M/\infty$ queueing process, see [25]. Hence, Theorem 4.2.7 entails for any Lipschitz function $\phi \in \mathrm{Lip}_\delta$, any $t > 0$, any initial state $x \in \mathbb{N}$ and any $y > 0$, the Poisson-type deviation inequality

$$\begin{aligned}\mathbb{P}_x\left(\left|\frac{1}{t}\int_0^t (\phi(X_s) - \mathbb{E}_x[\phi(X_s)])\, ds\right| \geq y\right) &\leq 2e^{-2\nu t g\left(\frac{y}{4(1-e^{-\nu t/2})\|\phi\|_{\mathrm{Lip}_\delta}}\right)} \\ &\leq 2e^{-\frac{ty\nu}{4(1-e^{-\nu t/2})\|\phi\|_{\mathrm{Lip}_\delta}}\log\left(1+\frac{y}{4(1-e^{-\nu t/2})\|\phi\|_{\mathrm{Lip}_\delta}}\right)},\end{aligned}$$

where $g(u) = (1+u)\log(1+u) - u$, $u > 0$.

Remark 4.4.1. We mention that the Laplace transform estimate (4.4.1) cannot be extended to the multi-dimensional case by using the general method of Section 4.3, since the upper bound in (4.4.1) depends on the initial condition $x \in \mathbb{N}$. Hence, although its distribution at time t is known, the problem raised in Remark 4.3.4 still subsists in the case of the $M/M/\infty$ queueing process.

Partie II

Fluctuations des intégrales stables stochastiques

Chapitre 5

On maximal inequalities for stable stochastic integrals

Ce chapitre fait l'objet d'un article à paraître dans le journal *Potential Analysis*.

Abstract

Sharp maximal inequalities in large and small range are derived for stable stochastic integrals. In order to control the tail of a stable process, we introduce a truncation level in the support of its Lévy measure: we show that the contribution of the compound Poisson stochastic integral is negligible as the truncation level is large, so that the study is reduced to establish maximal inequalities for the martingale part with a suitable choice of truncation level. The main problem addressed in this paper is to give upper bounds which remain bounded as the parameter of stability of the underlying stable process goes to 2. Applications to estimates of first passage times of symmetric stable processes above positive continuous curves complete this work.

5.1 Introduction

Given a filtered probability space $\mathbf{\Omega} = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$, consider on $\mathbf{\Omega}$ a càdlàg real stable process $Z = (Z_t)_{t\geq 0}$ of index $\alpha \in (0,2)$ without Gaussian component and let $H = (H_t)_{t\geq 0}$ be a sufficiently integrable predictable càdlàg process. The purpose of this paper is to give maximal inequalities for stable stochastic integrals $H\cdot Z = (\int_0^t H_s dZ_s)_{t\geq 0}$. We show that their decay in the bilateral case is

$$\mathbb{P}\left(\sup_{0\leq s\leq t}\left|\int_0^s H_\tau dZ_\tau\right| \geq x\right) \leq \frac{K}{\alpha x^\alpha}\,\|H\|^\alpha_{L^{\alpha+p}(\Omega\times[0,t])}, \quad x \geq x_\alpha, \quad p > 2-\alpha, \qquad (5.1.1)$$

whereas in the unilateral case, if Z is symmetric and $\alpha \in (1,2)$, it is

$$\mathbb{P}\left(\sup_{0\leq s\leq t}\int_0^s H_\tau dZ_\tau \geq x\right) \leq L_\alpha \exp\left(-M_\alpha\left(\frac{x}{\|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}}\right)^{\alpha/(\alpha-1)}\right), \quad x \leq \widetilde{x}_\alpha. \tag{5.1.2}$$

Here $L_\alpha, M_\alpha, x_\alpha$ and $\widetilde{x}_\alpha$ stand for positive numbers depending explicitly on α, whereas K is a positive constant independent of α.
It is known since the early 80's that stable stochastic integrals inherit regularly varying tails from the underlying stable process. For example, in order to prove the central limit theorem for stable stochastic integrals in the Skorohod space, Giné and Marcus established in [39] the maximal inequality

$$\sup_{x>0} x^\alpha\,\mathbb{P}\left(\sup_{0\leq t\leq 1}\left|\int_0^t H_s dZ_s\right| \geq x\right) \leq \frac{D}{\alpha(2-\alpha)^2}\,\|H\|^\alpha_{L^\alpha(\Omega\times[0,t])}, \tag{5.1.3}$$

where D is a universal constant independent of α. However, as α tends to 2, the upper bound in their maximal inequality (5.1.3) goes to infinity. On the other hand, the extremal behavior of stochastic integrals driven by multivariate Lévy processes with regularly varying tails have been studied recently in [47] by Hult and Lindskog, and by Applebaum, see [5]. In particular, if Z is symmetric and H is square-integrable and satisfies further the uniform integrability condition $\mathbb{E}\left[\sup_{t\in[0,1]}|H_t|^{\alpha+p}\right] < +\infty$ for some $p>0$, then Example 3.2 in [47] yields the extremal behavior

$$\lim_{x\to+\infty} x^\alpha\,\mathbb{P}\left(\sup_{0\leq s\leq t}\left|\int_0^s H_\tau dZ_\tau\right| \geq x\right) = C_\alpha\,\|H\|^\alpha_{L^\alpha(\Omega\times[0,t])}, \quad t\in[0,1], \tag{5.1.4}$$

where C_α depends on α and remains bounded as $\alpha \in (0,2]$. Therefore, as α gets close to 2, the maximal inequality (5.1.3) of Giné and Marcus does not recover the non-explosive asymptotic estimate (5.1.4).
Our approach to establish maximal inequalities for stable stochastic integrals is based on stochastic calculus for jump processes and allows us to avoid the limiting explosion of the upper bound described above. Following Pruitt in [69] for Lévy processes and more recently Houdré and Marchal in [43] in the specific case of stable random vectors, the method relies on the use of the Lévy-Itô decomposition of Z with a truncation level R in the support of its Lévy measure, in order to control the jump size of the martingale part: Z is split into the sum of a square-integrable martingale with infinitely many jumps bounded by R on each compact time interval, and a compound Poisson process which represents the large jumps of Z, plus a drift part. Constructing then the stable stochastic integral $H\cdot Z$ with respect to the above semimartingale decomposition, we show that the contribution of the compound Poisson stochastic integral in both bilateral and unilateral cases is negligible as the truncation level is large, reducing the study to the proof of maximal inequalities for the martingale part of $H\cdot Z$. Using stochastic calculus for Poisson random measures, sharp estimates follow by choosing suitably the truncation level R.

Let us describe the content of the paper. In Section 5.2, some notation and basic properties of stable processes are introduced. Then we apply a truncation method somewhat similar to that of Pruitt to derive maximal inequalities for stable stochastic integrals, and compare them with the corresponding results of Giné and Marcus, and Hult and Lindskog, see [39] and [47]. In particular, Proposition 5.2.4 slightly improves the estimate in [39, Theorem 3.5] when the index of stability α of the underlying stable process lies in $(1, 2)$ and under some integrability conditions. The main contribution of this paper is contained in Section 5.3, Theorem 5.3.2, where large range inequalities are given in the bilateral case (5.1.1), freeing us from the explosion of the upper bound as α goes to 2. Section 5.4 is devoted to small range tail estimates in the unilateral case (5.1.2). As a result, we recover the classical maximal Gaussian inequality via Theorem 5.4.2 and a limiting procedure in the Skorohod space. Finally, we apply in Section 5.5 the results of Section 5.2 and 5.3 to estimate first passage times of a symmetric stable process above several positive continuous curves. The method relies on an extension to the stable case of the results of [1, 68] established for Brownian motions.

5.2 Notation and preliminaries

Let $\mathbf{\Omega} = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space and let Z be a càdlàg real stable process on $\mathbf{\Omega}$ of index $\alpha \in (0, 2)$ without Gaussian component. For the sake of briefness, by a *stable process* we will implicitly mean an $(\mathscr{F}_t)_{t\geq 0}$*-adapted real càdlàg stable process* in the remainder of this paper. Recall that its characteristic function is defined by

$$\varphi_{Z_t}(u) = \exp t \left(iub + \int_{-\infty}^{+\infty} \left(e^{iuy} - 1 - iuy\, 1_{\{|y|\leq 1\}} \right) \nu(dy) \right), \tag{5.2.1}$$

where ν stands for the stable Lévy measure on $\mathbb{R}$:

$$\nu(dy) = \left(c_-\, 1_{\{y<0\}} + c_+\, 1_{\{y>0\}} \right) \frac{dy}{|y|^{\alpha+1}}, \quad c_-, c_+ \geq 0, \quad c_- + c_+ > 0. \tag{5.2.2}$$

As a Lévy process, Z is a semimartingale whose Lévy-Itô decomposition is given by

$$Z_t = bt + \int_0^t \int_{|y|\leq 1} y\, (\mu - \sigma)(dy, ds) + \int_0^t \int_{|y|>1} y\, \mu(dy, ds), \quad t \geq 0, \tag{5.2.3}$$

where μ is a Poisson random measure on $\mathbb{R} \times [0, +\infty)$ with intensity $\sigma(dy, dt) = \nu(dy) \otimes dt$ and b is the drift. In particular, if $\alpha < 1$, then Z is a finite variation process whereas when $\alpha \geq 1$, we have a.s.

$$\sum_{s\leq t} |\Delta Z_s| = +\infty, \quad t > 0,$$

where ΔZ_s denotes the jump size of Z at time $s > 0$.
Z is said to be strictly stable if we have the self-similarity property

$$(Z_{kt})_{t\geq 0} \stackrel{(d)}{=} (k^{\frac{1}{\alpha}} Z_t)_{t\geq 0},$$

where $k > 0$ and the equality $\overset{(d)}{=}$ is in the sense of finite dimensional distributions. If moreover $c := c_+ = c_-$, then Z is symmetric and its characteristic function (5.2.1) is computed to be

$$\varphi_{Z_t}(u) = e^{-t\rho_\alpha |u|^\alpha}, \tag{5.2.4}$$

where

$$\rho_\alpha := \frac{\sqrt{\pi}\Gamma((2-\alpha)/2)}{\alpha 2^\alpha \Gamma((1+\alpha)/2)} 2c.$$

5.2.1 The truncation method

In order to control the jump size of the martingale part of the stable stochastic integral, let us introduce the truncation method of the stable Lévy measure (5.2.2). For some truncation level $R > 1$, let $Z^{(R+)}$ and $Z^{(R-)}$ be the independent Lévy processes defined by

$$Z_t^{(R-)} := \int_0^t \int_{|y|\le R} y\,(\mu - \sigma)(dy, ds), \quad Z_t^{(R+)} := \int_0^t \int_{|y|>R} y\,\mu(dy, ds), \quad t \ge 0.$$

The first one has a compactly supported Lévy measure and is a square-integrable martingale with infinitely many jumps bounded by R on each compact time interval, whereas the second one is a compound Poisson process. The Lévy-Itô decomposition (5.2.3) rewrites as

$$Z_t = b_R t + Z_t^{(R-)} + Z_t^{(R+)}, \quad t \ge 0, \tag{5.2.5}$$

where $b_R := b + \int_{1<|y|\le R} y\,\nu(dy)$ is a drift depending on R.
Given a predictable càdlàg process H, let

$$\|H\|_{(p,t)} := \|H\|_{L^p(\Omega\times[0,t])} = \left(\int_0^t \mathbb{E}\left[|H_s|^p\right] ds\right)^{\frac{1}{p}}, \quad t \ge 0, \quad p > 0,$$

and define $\mathscr{P}_p$ (resp. $\mathscr{B}_p$) as the space of predictable càdlàg process H such that for all $t \ge 0$, $\|H\|_{(p,t)} < +\infty$ (resp. $\|H\|_{L^\infty(\Omega, L^p([0,t]))} < +\infty$). In particular, H is said integrable if $H \in \mathscr{P}_1$ and square-integrable if $H \in \mathscr{P}_2$.
Following [4, Chapter 4], we construct the stable stochastic integral of a square-integrable predictable process H as the sum of L^2-type and Lebesgue-Stieltjes stochastic integrals: letting

$$X_t^{(R-)} := \int_0^t H_s dZ_s^{(R-)}, \quad X_t^{(R+)} := \int_0^t H_s dZ_s^{(R+)}, \quad A_t^R := b_R \int_0^t H_s ds, \quad t \ge 0,$$

the first integral $X^{(R-)} = H\cdot Z^{(R-)}$ is a square-integrable martingale, whereas the integrals $X^{(R+)} = H \cdot Z^{(R+)}$ and A^R are constructed in the Lebesgue-Stieltjes sense, and we define the stable stochastic integral as

$$X_t := \int_0^t H_s dZ_s = A_t^R + X_t^{(R-)} + X_t^{(R+)}, \quad t \ge 0. \tag{5.2.6}$$

We denote respectively by $a \vee b$ and $a \wedge b$ the maximum and the minimum between two real numbers a and b.
We finish by making two remarks on the maximal inequalities of type (5.1.1) or (5.1.2) we will establish in the remainder of this paper:

Remark 5.2.1. The truncation level R is related to the deviation level x and to some L^p-norm of the process H, and is chosen each time equal to its optimal value.

Remark 5.2.2. Although they can be computed, the constants appearing in the upper bounds are not given explicitly in general, since their numerical value is not of crucial importance in our study.

5.2.2 A first maximal inequality

In order to study the rates of growth of Lévy processes, Pruitt established in [69] some maximal inequalities whose proofs are based on a truncation method for general Lévy measures, with a particular choice of truncation level.
Inspired by this work, we derive in this part a first maximal inequality for stable stochastic integrals by using the semimartingale decomposition (5.2.6).
Fix $t \geq 0$ and $x > \|H\|_{(2,t)}$. Using the above notation, we have by (5.2.6):

$$
\begin{aligned}
&\mathbb{P}\left(\sup_{0\leq s\leq t} |X_s| \geq x\right) \\
&\quad \leq \mathbb{P}\left(\sup_{0\leq s\leq t} \left|A_s^R\right| + \sup_{0\leq s\leq t} \left|X_s^{(R-)}\right| + \sup_{0\leq s\leq t} \left|X_s^{(R+)}\right| \geq x\right) \\
&\quad \leq \mathbb{P}\left(\sup_{0\leq s\leq t} \left|A_s^R\right| \geq \frac{x}{2}\right) + \mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(R-)}| \geq \frac{x}{2}\right) + \mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(R+)}| > 0\right).
\end{aligned} \tag{5.2.7}
$$

First, we investigate the absolutely continuous part A^R. By Chebychev's inequality,

$$
\begin{aligned}
\mathbb{P}\left(\sup_{0\leq s\leq t} \left|A_s^R\right| \geq \frac{x}{2}\right) &\leq \mathbb{P}\left(\int_0^t |H_\tau|\, d\tau \geq \frac{x}{2|b_R|}\right) \\
&\leq \frac{4{b_R}^2}{x^2}\,\mathbb{E}\left[\left(\int_0^t |H_\tau|\, d\tau\right)^2\right].
\end{aligned}
$$

Using the elementary inequality $(a+b)^2 \leq 2\left(a^2+b^2\right)$, $a, b \in \mathbb{R}$, and then Cauchy-Schwarz' inequality,

$$
\begin{aligned}
{b_R}^2 &= \left(b + \int_{1<|y|\leq R} y\, \nu(dy)\right)^2 \\
&\leq 2b^2 + 2\left(\int_{1<|y|\leq R} y\, \nu(dy)\right)^2
\end{aligned}
$$

$$\begin{aligned}
&\leq 2b^2 + 2\nu\left(\{y \in \mathbb{R} : 1 < |y| \leq R\}\right) \int_{1<|y|\leq R} y^2\,\nu(dy) \\
&\leq 2b^2 + 2\nu\left(\{y \in \mathbb{R} : |y| > 1\}\right) \int_{|y|\leq R} y^2\,\nu(dy) \\
&= 2\left(b^2 + \frac{(c_- + c_+)^2}{\alpha(2-\alpha)} R^{2-\alpha}\right).
\end{aligned}$$

By Cauchy-Schwarz' inequality again and since $x > \|H\|_{(2,t)}$, we have

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\leq s\leq t} |A_s^R| \geq \frac{x}{2}\right) &\leq \frac{8t}{x^2}\left(b^2 + \frac{(c_- + c_+)^2}{\alpha(2-\alpha)} R^{2-\alpha}\right) \|H\|_{(2,t)}^2 \\
&< \frac{8tb^2\|H\|_{(2,t)}^\alpha}{x^\alpha} + \frac{8t(c_- + c_+)^2 R^{2-\alpha}\|H\|_{(2,t)}^2}{\alpha(2-\alpha)x^2}. \qquad (5.2.8)
\end{aligned}$$

Now, we show that the contribution of the compound Poisson stochastic integral $X^{(R+)}$ is negligible as the truncation level R is sufficiently large. Recall that the integral $X^{(R+)}$, and so its supremum process $(\sup_{0\leq s\leq t} |X_s^{(R+)}|)_{t\geq 0}$, has piecewise constant sample paths and its distribution at any time has an atom at 0. Now, denote by T_1^R the first jump time of the Poisson process $(\mu\left(\{y \in \mathbb{R} : |y| > R\} \times [0,t]\right))_{t\geq 0}$ on the set $\{y \in \mathbb{R} : |y| > R\}$. If a.s. T_1^R occurs after time t, then the compound Poisson stochastic integral $X^{(R+)}$ (and so its supremum process) is identically 0 on the interval $[0,t]$. Thus we have

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(R+)}| > 0\right) &= 1 - \mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(R+)}| = 0\right) \\
&\leq 1 - \mathbb{P}\left(T_1^R > t\right) \\
&= 1 - \exp\left(-t\nu\left(\{y \in \mathbb{R} : |y| > R\}\right)\right) \\
&\leq t\nu\left(\{y \in \mathbb{R} : |y| > R\}\right) \\
&= \frac{(c_- + c_+)t}{\alpha R^\alpha}, \qquad (5.2.9)
\end{aligned}$$

where we used in the second equality above that T_1^R is exponentially distributed with parameter $\nu\left(\{y \in \mathbb{R} : |y| > R\}\right)$, see e.g. [76, Theorem 21.3].
Recall now that $X^{(R-)}$ is a square-integrable martingale involving the small jumps of Z. By Doob's inequality together with the isometry formula for Poisson stochastic integrals,

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(R-)}| \geq \frac{x}{2}\right) &\leq \frac{4}{x^2}\,\mathbb{E}\left[\left|\int_0^t \int_{|y|\leq R} H_\tau y\,(\mu - \sigma)(dy, d\tau)\right|^2\right] \\
&= \frac{4}{x^2}\,\mathbb{E}\left[\int_0^t \int_{|y|\leq R} H_\tau^2 y^2\,\nu(dy) d\tau\right] \\
&= \frac{4}{x^2} \int_{|y|\leq R} y^2\,\nu(dy) \int_0^t \mathbb{E}\left[H_\tau^2\right] d\tau,
\end{aligned}$$

that is to say

$$\mathbb{P}\left(\sup_{0\le s\le t}|X_s^{(R-)}|\ge\frac{x}{2}\right)\le\frac{4(c_-+c_+)\|H\|^2_{(2,t)}R^{2-\alpha}}{(2-\alpha)x^2}. \tag{5.2.10}$$

Finally, using (5.2.7) and choosing the truncation level

$$R=\frac{x}{\|H\|_{(2,t)}}>1$$

in (5.2.8), (5.2.9) and (5.2.10) show that there exists $K:=K(b,c_-,c_+,t)>0$, independent of α, such that

$$\mathbb{P}\left(\sup_{0\le s\le t}\left|\int_0^s H_\tau dZ_\tau\right|\ge x\right)\le\frac{K\|H\|^\alpha_{(2,t)}}{\alpha(2-\alpha)x^\alpha},\quad x>\|H\|_{(2,t)}. \tag{5.2.11}$$

Let us comment the estimate (5.2.11).
If Z is symmetric with Lévy measure $\nu(dy)=c|y|^{-\alpha-1}dy$, $c>0$, and H satisfies further the uniform integrability condition $\mathbb{E}\left[\sup_{0\le t\le 1}|H_t|^{\alpha+p}\right]<+\infty$, $p>0$, then Example 3.2 in [47] entails the asymptotic estimate

$$\lim_{x\to+\infty}x^\alpha\,\mathbb{P}\left(\sup_{0\le s\le t}\left|\int_0^s H_\tau dZ_\tau\right|\ge x\right)=\frac{K_{\alpha,c}}{\alpha}\int_0^t|H_\tau|^\alpha\,d\tau,\quad t\in[0,1], \tag{5.2.12}$$

where

$$K_{\alpha,c}:=\begin{cases}\frac{2c\sqrt{\pi}(1-\alpha)\Gamma((2-\alpha)/2)}{2^{\alpha+1}\Gamma(2-\alpha)\Gamma((1+\alpha)/2)\cos(\pi\alpha/2)} & \text{if } \alpha\ne1,\\ c & \text{if } \alpha=1,\end{cases}$$

which remains bounded as $\alpha\in[0,2]$. It shows that (5.2.11) is sharp for $\alpha\in(0,2)$ and also as α converges to 0, but goes to infinity as α tends to 2, in contrast to (5.2.12). On the other hand, assuming $H\in\mathscr{P}_\alpha$, then a combination of Theorem 3.5 and Example 3.7 in [39] implies the maximal inequality

$$\sup_{x>0}x^\alpha\,\mathbb{P}\left(\sup_{0\le t\le 1}\left|\int_0^t H_s dZ_s\right|\ge x\right)\le\frac{D}{\alpha(2-\alpha)^2}\int_0^1\mathbb{E}\left[|H_t|^\alpha\right]dt, \tag{5.2.13}$$

where D is a universal constant independent of α. Therefore, the speed of explosion in (5.2.11) is better than in (5.2.13) since it is linear in α and not quadratic, but it is worse in terms of L^p-norm of H, since the L^2-norm is involved instead of the optimal L^α-norm. Before avoiding in Section 5.3 the explosion of its upper bound as α gets close to 2, let us now improve (5.2.11) in terms of L^p-norm of H.

5.2.3 A maximal inequality in optimal L^α-norm

First, we quote [8, Proposition 2.1], up to a minor modification related to the integrability property of H:

Lemma 5.2.3. *Consider a stable stochastic integral $X := H \cdot Z$, where Z is a symmetric stable process of index $\alpha \in (1,2)$ with generator $\mathcal{L}$, and H is square-integrable. Let f be a $C^2(\mathbb{R})$-function with bounded first and second derivatives. Then the process M^f given by*

$$M_t^f := f(X_t) - f(X_0) - \int_0^t |H_s|^\alpha \mathcal{L} f(X_{s-}) ds, \quad t \geq 0,$$

is a martingale.

Now, we improve the upper bound in (5.2.11) in terms of L^p-norm of H. Actually, the estimate in Proposition 5.2.4 below recovers via a different proof the inequality (5.2.13) of Giné and Marcus, and slightly improves it as α tends to 2, since the speed of the explosion of the upper bound is not quadratic but only linear in α:

Proposition 5.2.4. *Let Z be a symmetric stable process of index $\alpha \in (1,2)$ and Lévy measure $\nu(dz) = c|z|^{-\alpha-1}dz$, $c > 0$, and let H be square-integrable. Then there exists $K_{\alpha,c} > 0$, finite as α tends to 2, such that*

$$\sup_{x>0} x^\alpha \, \mathbb{P}\left(\sup_{0 \leq t \leq 1} \left| \int_0^t H_s dZ_s \right| \geq x \right) \leq \frac{K_{\alpha,c}}{2-\alpha} \int_0^1 \mathbb{E}\left[|H_t|^\alpha \right] dt.$$

Proof. The present proof is an adaptation to the case of stable stochastic integrals of that of Bass in [9, Proposition 3.1]. Denote by $\mathcal{L}$ the infinitesimal generator of Z. Let f be a non-negative $C^2(\mathbb{R})$-function such that $f(0) = 0$, $f(y) = 1$ if $|y| \geq 1$ and whose first and second derivatives are bounded above in absolute value respectively by $c_1 > 0$ and $c_2 > 0$. Let $x > 0$, $f_x(y) := f(y/x)$ and let

$$\tau_x := \inf\{t \geq 0 : |X_t| \geq x\}$$

be the first exit time of the stable stochastic integral $X = H \cdot Z$ of the centered ball of radius x. If the process exits the ball before time 1, then $f_x(X_{1 \wedge \tau_x}) = 1$ and by Lemma 5.2.3 and a conditioning argument,

$$\begin{aligned}
\mathbb{P}\left(\sup_{0 \leq t \leq 1} |X_t| \geq x \right) &= \mathbb{P}(\tau_x \leq 1) \\
&\leq \mathbb{E}\left[f_x(X_{1 \wedge \tau_x}) \right] \\
&= \mathbb{E}\left[\int_0^{1 \wedge \tau_x} |H_t|^\alpha \mathcal{L} f_x(X_{t-}) dt \right] \\
&\leq \int_0^1 \mathbb{E}\left[|H_t|^\alpha \left| \mathcal{L} f_x(X_{t-}) \right| \right] dt.
\end{aligned}$$

Therefore,

$$\mathbb{P}\left(\sup_{0 \leq t \leq 1} |X_t| \geq x \right) \leq \|\mathcal{L} f_x\|_{L^\infty(\mathbb{R})} \int_0^1 \mathbb{E}\left[|H_t|^\alpha \right] dt. \tag{5.2.14}$$

By the symmetry of ν,

$$\begin{aligned}
\mathcal{L}f_x(y) &= \int_{\mathbb{R}} (f_x(y+z) - f_x(y) - zf_x'(y))\,\nu(dz) \\
&\leq \int_{|z|\leq R} \frac{c_2 z^2}{2x^2}\nu(dz) + \int_{|z|>R} \frac{2c_1|z|}{x}\nu(dz) \\
&= \frac{c_2 c R^{2-\alpha}}{(2-\alpha)x^2} + \frac{4c_1 c R^{1-\alpha}}{(\alpha-1)x}.
\end{aligned}$$

If we choose the truncation level $R = x$, then denoting $K_{\alpha,c} := c_2 c + 4c_1 c(2-\alpha)/(\alpha-1)$, the calculus above implies the bound

$$\|\mathcal{L}f_x\|_{L^\infty(\mathbb{R})} \leq \frac{K_{\alpha,c}}{(2-\alpha)x^\alpha}.$$

Finally, plugging this into (5.2.14), the proof is complete. ∎

5.3 Large range estimates for α close to 2

The purpose of the present part is to control the upper bound in (5.2.11), freeing us from its explosion as α tends to 2. The price to pay is to require stronger integrability conditions on the process H and to reduce the range interval of the deviation level x.
First, we recall Bihari's inequality, which is a Gronwall-type inequality. See e.g. [36, Chapter 1] for a proof of such an inequality.

Lemma 5.3.1. *Let T be a positive time horizon and let ρ, ψ and g be positive measurable functions such that ρ is monotone-increasing, $s \mapsto \psi(s)\rho\,(g(s))$ is integrable on $[0,T]$ and*

$$g(s) \leq K_T + \int_0^s \psi(\tau)\,\rho(g(\tau))\,d\tau, \quad s \in [0,T], \tag{5.3.1}$$

where $K_T \geq 0$. Then the Bihari inequality

$$g(T) \leq \phi^{-1}\left(\phi(K_T) + \int_0^T \psi(s)ds\right)$$

holds, where $\phi(x) := \int_0^x \frac{dy}{\rho(y)}$.

We can now state the main result of this paper:

Theorem 5.3.2. *Let Z be a stable process of index $\alpha \in (1,2)$ and Lévy measure ν given by (5.2.2). Let $p > 2-\alpha$, $\epsilon > 0$ and let $H \in \mathscr{P}_{\alpha+p}$. Then for all $t \geq 0$, there exists $K := K(b, c_-, c_+, t, p, \epsilon) > 0$, independent of α, such that for all*

$$x^\alpha \;>\; \|H\|^\alpha_{(\alpha+p,t)} \max\left\{1, \left(\frac{(2p)^{\frac{2}{\alpha+p}}}{\epsilon(2-\alpha)(\alpha+p)^{\frac{2}{\alpha+p}}}\right)^{\frac{\alpha+p}{\alpha+p-2}} \left(2^{\frac{\alpha+p-4}{2}} \vee 1\right)(c_- + c_+)t\right\},$$

we have the maximal inequality

$$\mathbb{P}\left(\sup_{0\le s\le t}\left|\int_0^s H_\tau dZ_\tau\right| \ge x\right) \le \frac{K\|H\|^{\alpha}_{(\alpha+p,t)}}{x^\alpha}. \tag{5.3.2}$$

Proof. We proceed as in the proof of inequality (5.2.11) and investigate first the absolutely continuous part A^R analogously to (5.2.8). Fix $t \ge 0$ and $x > \|H\|_{(\alpha+p,t)}$. By the elementary inequality $(a+b)^q \le 2^{q-1}\left(|a|^q + |b|^q\right)$, $a, b \in \mathbb{R}$, $q \ge 1$, applied with $q = \alpha + p$, together with Hölder's inequality, we get

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\le s\le t}|A_s^R| \ge \frac{x}{2}\right)\\
&\le \frac{2^{2\alpha+2p-1}t^{\alpha+p-1}\|H\|^{\alpha+p}_{(\alpha+p,t)}}{x^{\alpha+p}}\left(|b|^{\alpha+p} + \nu\left(\{y\in\mathbb{R} : |y|>1\}\right)^{\alpha+p-1}\int_{|y|\le R} y^{\alpha+p}\nu(dy)\right)\\
&= \frac{2^{2\alpha+2p-1}t^{\alpha+p-1}\|H\|^{\alpha+p}_{(\alpha+p,t)}}{x^{\alpha+p}}\left(|b|^{\alpha+p} + \frac{(c_-+c_+)^{\alpha+p}R^p}{p\,\alpha^{\alpha+p-1}}\right)\\
&\le \frac{2^{2\alpha+2p-1}t^{\alpha+p-1}\|H\|^{\alpha}_{(\alpha+p,t)}}{x^{\alpha}}\left(|b|^{\alpha+p} + \frac{(c_-+c_+)^{\alpha+p}R^p\|H\|^p_{(\alpha+p,t)}}{p\,\alpha^{\alpha+p-1}x^p}\right),
\end{aligned} \tag{5.3.3}$$

where we used in the last inequality $x > \|H\|_{(\alpha+p,t)}$.
Now, let us control the martingale part $X^{(R-)} = H \cdot Z^{(R-)}$. By Doob's and Burkholder's inequalities for martingales with jumps, see e.g. pp. 303-4 in [32], we have

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\le s\le t}|X_s^{(R-)}| \ge \frac{x}{2}\right) &\le \frac{2^{\alpha+p}}{x^{\alpha+p}}\mathbb{E}\left[\left|\int_0^t H_s dZ_s^{(R-)}\right|^{\alpha+p}\right]\\
&\le \frac{2^{\alpha+p}C_{\alpha+p}}{x^{\alpha+p}}\mathbb{E}\left[\left[\int_0^{\cdot} H_s dZ_s^{(R-)}, \int_0^{\cdot} H_s dZ_s^{(R-)}\right]_t^{\frac{\alpha+p}{2}}\right]\\
&= \frac{2^{\alpha+p}C_{\alpha+p}}{x^{\alpha+p}}\mathbb{E}\left[\left(\int_0^t\int_{|y|\le R} H_s^2\, y^2\,\mu(dy,ds)\right)^{\frac{\alpha+p}{2}}\right].
\end{aligned} \tag{5.3.4}$$

Let $(Y_s)_{s\in[0,t]}$ be the finite variation process defined by

$$Y_s := \int_0^s\int_{|y|\le R} H_\tau^2\, y^2\,\mu(dy,d\tau), \quad 0 \le s \le t.$$

By Itô's formula for jump processes and the inequality $(a+b)^q - a^q \le qb(a+b)^{q-1}$, $0 \le a \le b$, $q \ge 1$, applied with $q = (\alpha+p)/2$, we have

$$\begin{aligned}
Y_s^{\frac{\alpha+p}{2}} &= \int_0^s\int_{|y|\le R}\left(\left(Y_{\tau-} + H_\tau^2\, y^2\right)^{\frac{\alpha+p}{2}} - Y_{\tau-}^{\frac{\alpha+p}{2}}\right)\mu(dy,d\tau)\\
&\le \frac{\alpha+p}{2}\int_0^s\int_{|y|\le R} H_\tau^2\, y^2\left(Y_{\tau-} + H_\tau^2\, y^2\right)^{\frac{\alpha+p-2}{2}}\mu(dy,d\tau)
\end{aligned}$$

$$\leq \quad \frac{\alpha+p}{2}\,(2^{\frac{\alpha+p-4}{2}} \vee 1) \int_0^s \int_{|y|\leq R} H_\tau^2\, y^2 \left(Y_{\tau-}^{\frac{\alpha+p-2}{2}} + |H_\tau|^{\alpha+p-2}\, |y|^{\alpha+p-2} \right) \, \mu(dy, d\tau),$$

where we used in the last inequality the elementary bound $(a+b)^q \leq (2^{q-1} \vee 1)\; (a^q + b^q)$, $a, b \geq 0$, $q \geq 0$, applied with $q = (\alpha + p - 2)/2$. Denote $D_{\alpha,p} = 2^{\frac{\alpha+p-4}{2}} \vee 1$. Taking expectations and using Hölder's inequality, we get

$$\begin{aligned}
&\mathbb{E}\left[Y_s^{\frac{\alpha+p}{2}}\right] \\
&\quad \leq D_{\alpha,p}\frac{(\alpha+p)(c_-+c_+)}{2} \left(\frac{R^{2-\alpha}}{2-\alpha} \int_0^s \mathbb{E}\left[H_\tau^2\, Y_\tau^{\frac{\alpha+p-2}{2}}\right]\, d\tau + \frac{R^p}{p}\, \|H\|_{(\alpha+p,s)}^{\alpha+p} \right) \\
&\quad \leq D_{\alpha,p}\frac{(\alpha+p)(c_-+c_+)}{2} \left(\frac{R^{2-\alpha}}{2-\alpha} \int_0^s \mathbb{E}\left[|H_\tau|^{\alpha+p}\right]^{\frac{2}{\alpha+p}}\, \mathbb{E}\left[Y_\tau^{\frac{\alpha+p}{2}}\right]^{\frac{\alpha+p-2}{\alpha+p}} d\tau + \frac{R^p}{p}\, \|H\|_{(\alpha+p,t)}^{\alpha+p} \right).
\end{aligned}$$

Applying Lemma 5.3.1 with $T = t$,

$$g(s) := \mathbb{E}\left[Y_s^{\frac{\alpha+p}{2}}\right], \quad \psi(\tau) := \mathbb{E}\left[|H_\tau|^{\alpha+p}\right]^{\frac{2}{\alpha+p}}, \quad K_t := D_{\alpha,p}\frac{(\alpha+p)(c_-+c_+)R^p}{2p}\, \|H\|_{(\alpha+p,t)}^{\alpha+p}$$

and

$$\rho(x) := D_{\alpha,p}\frac{(\alpha+p)(c_-+c_+)R^{2-\alpha}}{2(2-\alpha)}\, x^{\frac{\alpha+p-2}{\alpha+p}},$$

and by using Hölder's inequality to estimate $\int_0^t \psi(\tau)d\tau$, we obtain

$$g(t) \leq \Phi^{-1}\left(\Phi\left(D_{\alpha,p}\frac{(\alpha+p)(c_-+c_+)R^p}{2p}\, \|H\|_{(\alpha+p,t)}^{\alpha+p} \right) + t^{\frac{\alpha+p-2}{\alpha+p}}\, \|H\|_{(\alpha+p,t)}^2 \right),$$

where

$$\begin{aligned}
\Phi(x) \;&:=\; \int_0^x \frac{dy}{\rho(y)} \\
&=\; \frac{2-\alpha}{D_{\alpha,p}(c_-+c_+)R^{2-\alpha}}\, x^{\frac{2}{\alpha+p}}.
\end{aligned}$$

Hence we have

$$\begin{aligned}
&\mathbb{E}\left[Y_t^{\frac{\alpha+p}{2}}\right] \leq \\
&\quad \frac{D_{\alpha,p}^{\frac{\alpha+p}{2}}\,(c_-+c_+)^{\frac{\alpha+p}{2}} R^{\frac{(2-\alpha)(\alpha+p)}{2}}}{(2-\alpha)^{\frac{\alpha+p}{2}}} \left(\frac{(2-\alpha)(\alpha+p)^{\frac{2}{\alpha+p}} R^{\frac{\alpha(\alpha+p-2)}{\alpha+p}}}{D_{\alpha,p}^{\frac{\alpha+p-2}{\alpha+p}}\,(c_-+c_+)^{\frac{\alpha+p-2}{\alpha+p}}\,(2p)^{\frac{2}{\alpha+p}}} + t^{\frac{\alpha+p-2}{\alpha+p}} \right)^{\frac{\alpha+p}{2}} \|H\|_{(\alpha+p,t)}^{\alpha+p}.
\end{aligned}$$

Now, choose the truncation level

$$R = \frac{x}{\|H\|_{(\alpha+p,t)}} > 1.$$

Since the assumption on x claims that

$$x^{\frac{\alpha(\alpha+p-2)}{\alpha+p}} > \frac{\|H\|_{(\alpha+p,t)}^{\frac{\alpha(\alpha+p-2)}{\alpha+p}} \, t^{\frac{\alpha+p-2}{\alpha+p}} \, D_{\alpha,p}^{\frac{\alpha+p-2}{\alpha+p}} \, (c_- + c_+)^{\frac{\alpha+p-2}{\alpha+p}} (2p)^{\frac{2}{\alpha+p}}}{\epsilon(2-\alpha)(\alpha+p)^{\frac{2}{\alpha+p}}},$$

we establish the following bound on moments

$$\mathbb{E}\left[Y_t^{\frac{\alpha+p}{2}}\right] \;\leq\; \frac{D_{\alpha,p}(c_- + c_+)(\alpha+p)(1+\epsilon)^{\frac{\alpha+p}{2}} \|H\|_{(\alpha+p,t)}^{\alpha}}{2p} \, x^p.$$

Finally, plugging the latter inequality into (5.3.4) yields

$$\mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(R-)}| \geq \frac{x}{2}\right) \leq \frac{2^{\alpha+p} C_{\alpha+p} D_{\alpha,p}(c_- + c_+)(\alpha+p)(1+\epsilon)^{\frac{\alpha+p}{2}} \|H\|_{(\alpha+p,t)}^{\alpha}}{2p x^{\alpha}},$$

and together with (5.2.7) and the choice of truncation level $R = x/\|H\|_{(\alpha+p,t)}$ in (5.2.9) and (5.3.3), Theorem 5.3.2 is proved. ■

Under further assumptions on Z and H, the process $H \cdot Z$ is a time-changed stable process and we get the following maximal inequality, which is asymptotically optimal in terms of L^{α}-norm when $\|H\|_{L^{\alpha}([0,t])}$ is bounded on $\mathbf{\Omega}$ for all $t \geq 0$:

Corollary 5.3.3. *Let Z be a symmetric stable process of index $\alpha \in (1,2)$ and Lévy measure $\nu(dy) = c|y|^{-\alpha-1}dy$, $c > 0$. Let $H \in \mathscr{B}_{\alpha}$ with a.s. $\lim_{t\to+\infty} \int_0^t |H_s|^{\alpha} ds = +\infty$. Let $p > 2-\alpha$ and $\epsilon > 0$. Then there exists $K := K(c,p,\epsilon) > 0$, independent of α, such that for all $t \geq 0$ and for all*

$$x^{\alpha} \;>\; \left\|\int_0^t |H_s|^{\alpha}\, ds\right\|_{L^{\infty}(\Omega)} \left(\frac{2p^{\frac{2}{\alpha+p}}}{\epsilon(2-\alpha)(\alpha+p)^{\frac{2}{\alpha+p}}}\right)^{\frac{\alpha+p}{\alpha+p-2}} \left(2^{\frac{\alpha+p-4}{2}} \vee 1\right) c,$$

we have the estimate

$$\mathbb{P}\left(\sup_{0\leq s\leq t}\left|\int_0^s H_{\tau} dZ_{\tau}\right| \geq x\right) \leq \frac{K}{x^{\alpha}} \left\|\int_0^t |H_s|^{\alpha}\, ds\right\|_{L^{\infty}(\Omega)}. \tag{5.3.5}$$

Proof. By [73, Theorem 3.1], the process $H \cdot Z$ is a time-changed process of Z, i.e. we have the identity a.s.

$$\int_0^t H_s dZ_s = \hat{Z}_{\tau_t}, \quad t \geq 0,$$

where $\tau = (\tau_t)_{t\geq 0}$ given by $\tau_t := \int_0^t |H_s|^{\alpha} ds$ is a time change process, and $\hat{Z}$ is a symmetric stable process defined on $\mathbf{\Omega}$ and having the same distribution as Z. Since the symmetry of $\hat{Z}$ implies it is self-similar of index α, then so is the supremum process:

$$\left(\sup_{0\leq s\leq kt} \hat{Z}_s\right)_{t\geq 0} \overset{(d)}{=} \left(k^{\frac{1}{\alpha}} \sup_{0\leq s\leq t} \hat{Z}_s\right)_{t\geq 0}, \quad k > 0.$$

Thus, denoting $\beta(t) := \|\tau_t\|_{L^\infty(\Omega)}^{1/\alpha}$, we have

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\le s\le t}\left|\int_0^s H_\tau dZ_\tau\right| \ge x\right) &= \mathbb{P}\left(\sup_{0\le s\le t}|\hat{Z}_{\tau_s}| \ge x\right) \\
&\le \mathbb{P}\left(\sup_{0\le s\le \tau_t}|\hat{Z}_s| \ge x\right) \\
&\le \mathbb{P}\left(\sup_{0\le s\le \beta(t)^\alpha}|\hat{Z}_s| \ge x\right) \\
&= \mathbb{P}\left(\sup_{0\le s\le 1}|\hat{Z}_s| \ge \frac{x}{\beta(t)}\right).
\end{aligned}$$

Finally, applying Theorem 5.3.2 with $H_s = 1$ for all $0 \le s \le t = 1$, the proof is complete. ■

Remark 5.3.4. If Z is a non-symmetric strictly stable process and H is positive and satisfies further the hypothesis of Corollary 5.3.3 (resp. that of Theorem 5.4.2 below), then the stable stochastic integral $H \cdot Z$ is still a time-changed process of Z. Thus, applying in the proof above (resp. in the proof of Theorem 5.4.2) Theorem 3 in [57] instead of Theorem 3.1 in [73], an estimate somewhat similar to that of Corollary 5.3.3 (resp. Theorem 5.4.2) can be established.

5.4 Small range maximal inequalities

In this part, we derive small range estimates in the unilateral case (5.1.2). Recently, Breton and Houdré investigated in [21] small and intermediate range concentration for stable random vectors. In particular, the small range behavior is covered by their Theorem 1, whose small deviation rate is of order $\exp\left(-c_\alpha x^{\alpha/(\alpha-1)}\right)$ for some positive c_α depending on α. Before proving a similar rate for suprema of stable stochastic integrals, let us establish first the result for symmetric stable processes via Proposition 5.4.1 below. We point out that using the scaling property, it is sufficient to get the result on the time interval $[0, 1]$.

Proposition 5.4.1. *Let Z be a symmetric stable process of index $\alpha \in (1, 2)$ and Lévy measure $\nu(dy) = c|y|^{-\alpha-1}dy$, $c > 0$. Then for all $\lambda > \lambda_0(\alpha)$, where $\lambda_0(\alpha)$ is the unique solution of the equation*

$$\lambda \log\left(1 + \frac{(2-\alpha)\lambda}{2c}\right) = \frac{4c}{\alpha},$$

there exists $x_0(\alpha, \lambda) > 0$ such that for all $0 \le x \le x_0(\alpha, \lambda)$,

$$\mathbb{P}\left(\sup_{0\le t\le 1} Z_s \ge x\right) \le \frac{2c}{\alpha}\left(\frac{x}{\lambda}\right)^{\frac{\alpha}{\alpha-1}} + \exp\left(-\frac{\lambda\log\left(1+\frac{(2-\alpha)\lambda}{2c}\right)}{2}\left(\frac{x}{\lambda}\right)^{\frac{\alpha}{\alpha-1}}\right). \tag{5.4.1}$$

Proof. As in the proof of inequality (5.2.9), we have

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le t\le 1} Z_s \ge x\right) &\le \mathbb{P}\left(\sup_{0\le t\le 1} Z_s^{(R+)} > 0\right) + \mathbb{P}\left(\sup_{0\le t\le 1} Z_s^{(R-)} \ge x\right)\\ &\le \frac{2c}{\alpha R^\alpha} + \mathbb{P}\left(\sup_{0\le t\le 1} Z_s^{(R-)} \ge x\right). \end{aligned} \tag{5.4.2}$$

The Lévy process $Z^{(R-)}$ is a martingale with jumps bounded by R, hence has exponential moments, see e.g. [30, Proposition 3.14]. Moreover, the angle bracket process $< Z^{(R-)}, Z^{(R-)} >$ is computed to be

$$\begin{aligned}< Z^{(R-)}, Z^{(R-)} >_t &= \int_0^t \int_{|y|\le R} y^2\, \nu(dy) ds\\ &= \frac{2ct}{2-\alpha} R^{2-\alpha}\\ &= v_t(R)^2.\end{aligned}$$

Let $\phi(z) := z^{-2}\left(e^z - z - 1\right)$, $z > 0$, and define for all $\beta > 0$ the process $S^{(\beta,R)}$ by

$$S_t^{(\beta,R)} = \exp\left(\beta Z_t^{(R-)} - \beta^2 \phi(\beta R) < Z^{(R-)}, Z^{(R-)} >_t\right), \quad t \ge 0.$$

By [56, Lemma 23.19], $S^{(\beta,R)}$ is a supermartingale for all $\beta > 0$. Thus, the exponential Markov's inequality yields

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le t\le 1} Z_s^{(R-)} \ge x\right) &\le \inf_{\beta>0} \mathbb{P}\left(\sup_{0\le t\le 1} S_t^{(\beta,R)} \ge \exp\left(\beta x - \beta^2 v_1(R)^2 \phi(\beta R)\right)\right)\\ &\le \inf_{\beta>0} \exp\left(-\beta x + \beta^2 v_1(R)^2 \phi(\beta R)\right)\\ &= \exp\left(\frac{x}{R} - \left(\frac{x}{R} + \frac{v_1(R)^2}{R^2}\right) \log\left(1 + \frac{Rx}{v_1(R)^2}\right)\right)\\ &\le \exp\left(-\frac{x}{2R} \log\left(1 + \frac{Rx}{v_1(R)^2}\right)\right)\\ &= \exp\left(-\frac{x}{2R} \log\left(1 + \frac{(2-\alpha)R^{\alpha-1}x}{2c}\right)\right),\end{aligned}$$

where in the latter inequality we used $(1+u)\log(1+u) - u \ge \frac{u}{2}\log(1+u)$, $u \ge 0$, which is equivalent to $(1+u/2)\log(1+u) \ge u$, $u \ge 0$, established by a standard convexity argument. Now, let the truncation level R be such that $x = \lambda R^{1-\alpha}$ for some $\lambda > 0$. Plugging the last inequality into (5.4.2), we get

$$\mathbb{P}\left(\sup_{0\le t\le 1} Z_s \ge x\right) \le \frac{2c}{\alpha}\left(\frac{x}{\lambda}\right)^{\frac{\alpha}{\alpha-1}} + \exp\left(-\frac{\lambda \log\left(1 + \frac{(2-\alpha)\lambda}{2c}\right)}{2}\left(\frac{x}{\lambda}\right)^{\frac{\alpha}{\alpha-1}}\right)$$

$$=: \quad F\left(\left(\frac{x}{\lambda}\right)^{\frac{\alpha}{\alpha-1}}\right). \tag{5.4.3}$$

A necessary condition for the upper bound in (5.4.3) to make sense is that the real number $F\left((x/\lambda)^{\alpha/(\alpha-1)}\right)$ has to be smaller than 1, which is the case in a neighborhood of 0_+ if $\lambda > \lambda_0(\alpha)$. Finally, choose $x_0(\alpha,\lambda) > 0$ such that $F\left((x_0(\alpha,\lambda)/\lambda)^{\alpha/(\alpha-1)}\right) = 1$ to obtain the maximum range of validity for the result. ■

Now, we can establish a small range maximal inequality for stable stochastic integrals:

Theorem 5.4.2. *Let Z be a symmetric stable process of index $\alpha \in (1,2)$ and Lévy measure $\nu(dy) = c|y|^{-\alpha-1}dy$, $c > 0$, and let $H \in \mathscr{B}_\alpha$ with a.s. $\lim_{t\to+\infty}\int_0^t |H_s|^\alpha\, ds = +\infty$. Then for all $\lambda > \lambda_0(\alpha)$, where $\lambda_0(\alpha)$ is the unique solution of the equation*

$$\lambda \log\left(1 + \frac{(2-\alpha)\lambda}{2c}\right) = \frac{4c}{\alpha},$$

there exists $x_1(\alpha,\lambda) > 0$ such that for all $0 \le x \le x_1(\alpha,\lambda)$ and all $t \ge 0$,

$$\mathbb{P}\left(\sup_{0\le s\le t}\int_0^s H_\tau dZ_\tau \ge x\right) \le \frac{2c}{\alpha}\left(\frac{x}{\lambda\|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}}\right)^{\frac{\alpha}{\alpha-1}} + \exp\left(-\frac{\lambda\log\left(1+\frac{(2-\alpha)\lambda}{2c}\right)}{2}\left(\frac{x}{\lambda\|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}}\right)^{\frac{\alpha}{\alpha-1}}\right). \tag{5.4.4}$$

Proof. Following the proof of Corollary 5.3.3, we have by time change and scaling

$$\mathbb{P}\left(\sup_{0\le s\le t}\int_0^s H_\tau dZ_\tau \ge x\right) \le \mathbb{P}\left(\sup_{0\le s\le 1}\hat{Z}_s \ge \frac{x}{\|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}}\right),$$

where $\hat{Z}$ is a symmetric stable process defined on $\mathbf{\Omega}$ and having the same law as Z. Finally, Proposition 5.4.1 applied to $\hat{Z}$ achieves the proof. ■

Remark 5.4.3. For all $\epsilon > 0$, let x_ϵ be the unique solution of the equation

$$\frac{2c}{\alpha}\left(\frac{x}{\lambda\|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}}\right)^{\frac{\alpha}{\alpha-1}} = \epsilon\exp\left(-\frac{\lambda\log\left(1+\frac{(2-\alpha)\lambda}{2c}\right)}{2}\left(\frac{x}{\lambda\|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}}\right)^{\frac{\alpha}{\alpha-1}}\right).$$

Then for all $0 \le x \le x_\epsilon$, the inequality (5.4.4) implies

$$\mathbb{P}\left(\sup_{0\le s\le t}\int_0^s H_\tau dZ_\tau \ge x\right) \le (1+\epsilon)\exp\left(-\frac{\lambda\log\left(1+\frac{(2-\alpha)\lambda}{2c}\right)}{2}\left(\frac{x}{\lambda\|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}}\right)^{\frac{\alpha}{\alpha-1}}\right).$$

Thus, the order of the upper bound in (5.4.4) is $\exp\left(-c_\alpha \left(x/\|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}\right)^{\alpha/(\alpha-1)}\right)$, and is comparable to that in [21, Theorem 1] for Lipschitz functions of stable random vectors.

Remark 5.4.4. The quantity $x_1(\alpha,\lambda)$ in Theorem 5.4.2 can be given explicitly. Indeed, let $x_0^*(\alpha,\lambda) > 0$ be the real number where the function F in (5.4.3) reaches its unique minimum, i.e.

$$x_0^*(\alpha,\lambda)^{\frac{\alpha}{\alpha-1}} = \frac{2\lambda^{\frac{1}{\alpha-1}}}{\log\left(1+\frac{(2-\alpha)\lambda}{2c}\right)} \log\left(\frac{\alpha\lambda\log\left(1+\frac{(2-\alpha)\lambda}{2c}\right)}{4c}\right) < x_0(\alpha,\lambda)^{\frac{\alpha}{\alpha-1}},$$

then choose $x_1(\alpha,\lambda) = \|H\|_{L^\infty(\Omega,L^\alpha([0,t]))}\, x_0^*(\alpha,\lambda)$.

Remark 5.4.5. There is no optimal choice for the parameter λ in Theorem 5.4.2: on the one hand, $\lambda = \lambda_0(\alpha)$ achieves the best maximal inequality (5.4.4) but in this case the domain for the deviation level x is empty; on the other hand, as λ increases, the domain expands but in this case the maximal inequality (5.4.4) is the worst.

As an application of Theorem 5.4.2, let us recover the classical maximal inequality in the Gaussian case, cf. Proposition 1.8 p.55 in [70].

Corollary 5.4.6. *Let $(B_t)_{t\geq 0}$ be a standard Brownian motion. Then the following maximal inequality holds*

$$\mathbb{P}\left(\sup_{0\leq s\leq t} B_s \geq x\right) \leq \exp\left(-\frac{x^2}{2t}\right), \quad x>0, \quad t\geq 0.$$

Proof. Let $(X^n)_{n\geq 2}$ be a sequence of symmetric stable processes of index $\alpha_n = 2-1/n$ and Lévy measure $\nu_n(dy) = (2n)^{-1}dy/|y|^{\alpha_n+1}$. Applying Theorem 5.4.2 to X^n, $n\geq 2$, the inequality (5.4.4) becomes for all $0\leq x\leq x_1(\alpha_n,\lambda)$, all $\lambda > \lambda_0(\alpha_n)$ and all $t\geq 0$

$$\mathbb{P}\left(\sup_{0\leq s\leq t} X_s^n \geq x\right) \leq \frac{1}{2n-1}\left(\frac{x}{\lambda t^{\frac{n}{2n-1}}}\right)^{\frac{2n-1}{n-1}} + \exp\left(-\frac{\lambda\log(1+\lambda)}{2}\left(\frac{x}{\lambda t^{\frac{n}{2n-1}}}\right)^{\frac{2n-1}{n-1}}\right), \tag{5.4.5}$$

where

$$x_1(\alpha,\lambda)^{\frac{2n-1}{n-1}} = \frac{2(t\lambda)^{\frac{n}{n-1}}}{\log(1+\lambda)} \log\left((n-\frac{1}{2})\lambda\log(1+\lambda)\right),$$

and $\lambda_0(\alpha_n)$ is the unique solution of the equation

$$\lambda\log(1+\lambda) = \frac{2}{2n-1}.$$

Note that $\lambda_0(\alpha_n)$ converges to 0 and $x_1(\alpha_n,\lambda)$ to infinity as n goes to infinity. Denoting $D[0,+\infty)$ the Skorohod space of real-valued càdlàg functions on $[0,+\infty)$ equipped with the Skorohod topology, the sequence of processes $(X^n)_{n\geq 2}$ converges weakly in $D[0,+\infty)$ as $n \to +\infty$ to a standard Brownian motion $(B_t)_{t\geq 0}$ (say), see e.g. Section 3 of Chapter VII in [49]. Since the supremum functional is continuous on $D[0,+\infty)$, then the Continuous Mapping Theorem p.20 in [12] implies

$$\lim_{n\to+\infty} \mathbb{P}\left(\sup_{0\leq s\leq t} X_s^n \geq x\right) = \mathbb{P}\left(\sup_{0\leq s\leq t} B_s \geq x\right), \quad x > 0, \quad t \geq 0.$$

Finally, letting n going to infinity and then λ to 0 in the right-hand-side of (5.4.5) yield the result. ∎

5.5 Estimates of first passage times of symmetric stable processes above positive continuous curves

In [1, 68], the authors investigate functional transformations related to first crossing problems for self-similar diffusions. More precisely, they show via a time change transformation how the distribution of the first passage time of a Gauss-Markov process of Ornstein-Uhlenbeck type can be deduced from the law of the first crossing time of a continuous curve by a Brownian motion. In this part, we adapt this method in order to estimate the first passage time of a symmetric stable process above several positive continuous curves, by using the maximal inequalities of Section 5.2 and 5.3.
To do so, let X^ϕ be a stable-Markov process of Ornstein-Uhlenbeck type of index $\alpha \in (0,2)$ and parameter ϕ, i.e. X^ϕ has the integral representation

$$X_t^\phi := \phi(t)\int_0^t \frac{dZ_s}{\phi(s)}, \quad t \in [0,T), \quad T \in (0,+\infty],$$

where Z is a symmetric stable process of index α and Lévy measure $\nu(dy) = c|y|^{-\alpha-1}dy$, $c > 0$, and ϕ is a positive $C^\infty([0,T))$-function. Let also

$$T_x^\phi := \inf\{t \in [0,T) : |X_t^\phi| \geq x\}$$

be its first exit time of the centered ball of radius x. Given a positive continuous function f such that $f(0) \neq 0$, define

$$T^{(f)} := \inf\{t \geq 0 : |Z_t| \geq f(t)\}$$

as the first passage time of $|Z|$ above f. Let us give a first lemma which states an identity in law between first passage times:

Lemma 5.5.1. *Let X^ϕ be a stable-Markov process of Ornstein-Uhlenbeck type of index $\alpha \in (0,2)$ and parameter ϕ. Assume that $\tau_t := \int_0^t \frac{ds}{\phi(s)^\alpha} < +\infty$ for all $t \in [0,T)$ and that $\lim_{t\to T} \tau_t = +\infty$. Denote by τ^{-1} the inverse of τ and let $h_{\phi,\tau}$ be the function defined on $(0,+\infty)$ by $h_{\phi,\tau}(t) = 1/(\phi \circ \tau^{-1}(t))$. Then for all $x > 0$, we have the identity in distribution*

$$\mathbb{P}\left(T_x^\phi \in dr\right) = \mathbb{P}\left(\tau^{-1}\left(T^{(xh_{\phi,\tau})}\right) \in dr\right), \quad r \in [0,T).$$

Proof. By [73, Theorem 3.1], the process X^ϕ rewrites as a time-changed symmetric stable process, i.e. we have a.s.

$$X_t^\phi = \phi(t)\hat{Z}_{\tau_t}, \quad t \in [0,T),$$

where $\hat{Z}$ is a symmetric stable process defined on the same probability space as Z and having the same distribution. Thus, we have for all $r \in [0,T)$

$$\begin{aligned}
\mathbb{P}\left(T_x^\phi \le r\right) &= \mathbb{P}\left(\inf\{t \in [0,T) : |X_t^\phi| \ge x\} \le r\right) \\
&= \mathbb{P}\left(\inf\left\{t \in [0,T) : |\hat{Z}_{\tau_t}| \ge \frac{x}{\phi(t)}\right\} \le r\right) \\
&= \mathbb{P}\left(\tau^{-1}\left(T^{(xh_{\phi,\tau})}\right) \le r\right).
\end{aligned}$$

■

Now, we establish via an integration by parts formula several maximal inequalities for stable-Markov processes of Ornstein-Uhlenbeck type:

Lemma 5.5.2. *Let X^ϕ be a stable-Markov process of Ornstein-Uhlenbeck type of index $\alpha \in (0,2)$ and parameter ϕ. Let $t \in [0,T)$. Then we have the support estimate*

$$\mathbb{P}\left(\sup_{0\le s\le t} |X_s^\phi| < y\right) \le \exp\left(-\frac{ct}{\alpha 2^{\alpha-1} y^\alpha}\right), \quad y > 0. \tag{5.5.1}$$

If $\alpha \in (0,1]$, then we have the maximal inequality

$$\mathbb{P}\left(\sup_{0\le s\le t} |X_s^\phi| \ge x\right) \le \frac{4ct}{\alpha x^\alpha}\left(1 + \left\|\phi(\cdot)\int_0^\cdot \frac{\phi'(\tau)}{\phi(\tau)^2}\,d\tau\right\|_{L^\infty([0,t])}\right)^\alpha, \quad x > 0, \tag{5.5.2}$$

whereas if $\alpha \in (1,2)$, then for all

$$x^\alpha > \frac{tc}{(2-\alpha)^{\frac{\alpha+1}{\alpha-1}}}\left(1 + \left\|\phi(\cdot)\int_0^\cdot \frac{\phi'(\tau)}{\phi(\tau)^2}\,d\tau\right\|_{L^\infty([0,t])}\right),$$

we have

$$\mathbb{P}\left(\sup_{0\le s\le t} |X_s^\phi| \ge x\right) \le \frac{K_c t}{x^\alpha}\left(1 + \left\|\phi(\cdot)\int_0^\cdot \frac{\phi'(\tau)}{\phi(\tau)^2}\,d\tau\right\|_{L^\infty([0,t])}\right)^\alpha, \tag{5.5.3}$$

where $K_c > 0$ only depends on c.

Proof. Fix $t \in [0,T)$ and $y > 0$. If a.s. the path of the process X^ϕ lies in the interval $(-y,y)$ up to time t, then there are no jumps of magnitude larger than $2y$ before time t, so that we have the set inclusion $\left\{\sup_{0\le s\le t} |X_s^\phi| < y\right\} \subset \left\{\sup_{0\le s\le t} |\Delta X_s^\phi| < 2y\right\}$. Moreover, the process X^ϕ has the same jumps as the process Z by definition. Thus, if

T_1^{2y} denotes the first jump time on the set $\{z \in \mathbb{R} : |z| > 2y\}$ of the Poisson process $(\mu(\{z \in \mathbb{R} : |z| > 2y\} \times [0,t]))_{t\in[0,T)}$, then we have

$$\begin{aligned} \mathbb{P}\left(\sup_{0\le s\le t} |X_s^\phi| < y\right) &\le \mathbb{P}\left(\sup_{0\le s\le t} |\Delta X_s^\phi| < 2y\right) \\ &= \mathbb{P}\left(\sup_{0\le s\le t} |\Delta Z_s| < 2y\right) \\ &\le \mathbb{P}\left(T_1^{2y} > t\right) \\ &= \exp\left(-t\nu\left(\{z \in \mathbb{R} : |z| \ge 2y\}\right)\right) \\ &= \exp\left(-\frac{2ct}{\alpha(2y)^\alpha}\right), \end{aligned}$$

where in the second equality we used that T_1^{2y} is exponentially distributed with parameter $\nu(\{z \in \mathbb{R} : |z| > 2y\})$. The support estimate (5.5.1) is proved.
Now, we establish (5.5.2) and (5.5.3). By the classical integration by parts formula for semimartingales, cf. [30, Proposition 8.11], we have

$$\begin{aligned} \int_0^t \frac{dZ_s}{\phi(s)} &= \frac{Z_t}{\phi(t)} - \int_0^t Z_{s-} d\left(\frac{1}{\phi}\right)(s) \\ &= \frac{Z_t}{\phi(t)} + \int_0^t \frac{\phi'(s) Z_s}{\phi(s)^2} ds. \end{aligned}$$

Hence, the process X^ϕ rewrites as

$$X_t^\phi = Z_t + \phi(t) \int_0^t \frac{\phi'(s)}{\phi(s)^2} Z_s \, ds, \quad t \in [0,T). \tag{5.5.4}$$

Denote $A_t := \left\| \phi(\cdot) \int_0^{\cdot} \frac{\phi'(\tau)}{\phi(\tau)^2} \, d\tau \right\|_{L^\infty([0,t])}$ and let us distinguish two cases:

• if $\alpha \in (0,1]$, then following the proof of inequality (5.2.11) but restricted to the symmetric stable process Z yields the inequality

$$\mathbb{P}\left(\sup_{0\le s\le t} |Z_s| \ge x\right) \le \frac{4ct}{\alpha x^\alpha}.$$

Thus, together with (5.5.4), we have

$$\begin{aligned} \mathbb{P}\left(\sup_{0\le s\le t} |X_s^\phi| \ge x\right) &\le \mathbb{P}\left(\sup_{0\le s\le t} |Z_s| \ge \frac{x}{1+A_t}\right) \\ &\le \frac{4ct(1+A_t)^\alpha}{\alpha x^\alpha}; \end{aligned}$$

• if $\alpha \in (1,2)$, then Corollary 5.3.3 applied with e.g. $p = 1$ and $\epsilon = 2^{(\alpha-1)/(\alpha+1)}$, together with (5.5.4) show that there exists $K_c > 0$, which only depends of c, such that

$$\mathbb{P}\left(\sup_{0\le s\le t} |X_s^\phi| \ge x\right) \le \mathbb{P}\left(\sup_{0\le s\le t} |Z_s| \ge \frac{x}{1+A_t}\right)$$

$$\leq \quad \frac{K_c t(1+A_t)^\alpha}{x^\alpha}$$

for all $x^\alpha > (tc(1+A_t)^\alpha)/((2-\alpha)^{(\alpha+1)/(\alpha-1)})$. ∎

Remark 5.5.3. The support estimate (5.5.1) is independent of ϕ and thus is similar to that of a symmetric stable process.

Remark 5.5.4. No time change techniques are required in the proof of Lemma 5.5.2 but just the integration by parts formula which entails (5.5.4). However, if we assume $\tau_t := \int_0^t \frac{ds}{\phi(s)^\alpha} < +\infty$, $t \in [0,T)$, with $\tau_t \to +\infty$ as $t \to T$ and that ϕ is non-decreasing on $[0,T)$, then time change, scaling and Corollary 5.3.3 entail for sufficiently large x

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^\phi| \geq x\right) &\leq \mathbb{P}\left(\sup_{0\leq s\leq t} |\hat{Z}_{\tau_s}| \geq \frac{x}{\phi(t)}\right) \\
&\leq \mathbb{P}\left(\sup_{0\leq s\leq 1} |\hat{Z}_s| \geq \frac{x}{\phi(t)\tau_t^{\frac{1}{\alpha}}}\right) \\
&\leq \frac{K_c}{x^\alpha}\phi(t)^\alpha \int_0^t \frac{ds}{\phi(s)^\alpha}.
\end{aligned}$$

Now, we are able to state the main result of this part:

Theorem 5.5.5. *Let Z be a symmetric stable process of index $\alpha \in (0,2)$ and Lévy measure $\nu(dy) = c|y|^{-\alpha-1}dy$, $c > 0$. Let ϕ be a positive $C^\infty([0,T))$-function such that $\tau_t := \int_0^t \frac{ds}{\phi(s)^\alpha} < +\infty$ for all $t \in [0,T)$ and that $\lim_{t\to T}\tau_t = +\infty$. Denote by τ^{-1} the inverse of τ and by $h_{\phi,\tau}$ the function defined on $(0,+\infty)$ by $h_{\phi,\tau}(t) := 1/(\phi\circ\tau^{-1}(t))$. Then for all $x > 0$,*

$$\mathbb{P}\left(T^{(xh_{\phi,\tau})} > r\right) \leq \exp\left(-\frac{2c\tau_r^{-1}}{\alpha(2x)^\alpha}\right), \quad r > 0. \tag{5.5.5}$$

If $\alpha \in (0,1]$, then for all $x > 0$, we have

$$\mathbb{P}\left(T^{(xh_{\phi,\tau})} \leq r\right) \leq \frac{4c\tau_r^{-1}}{\alpha x^\alpha}\left(1 + \left\|\phi(\cdot)\int_0^\cdot \frac{\phi'(t)}{\phi(t)^2}\,dt\right\|_{L^\infty([0,\tau_r^{-1}])}\right)^\alpha, \quad r > 0, \tag{5.5.6}$$

whereas if $\alpha \in (1,2)$, then there exists $K_c > 0$, which only depends of c, such that for all $x > 0$ and for all $0 \leq r < r_0(\alpha,x)$, we have

$$\mathbb{P}\left(T^{(xh_{\phi,\tau})} \leq r\right) \leq \frac{K_c\tau_r^{-1}}{x^\alpha}\left(1 + \left\|\phi(\cdot)\int_0^\cdot \frac{\phi'(t)}{\phi(t)^2}\,dt\right\|_{L^\infty([0,\tau_r^{-1}])}\right)^\alpha, \tag{5.5.7}$$

where $r_0(\alpha,x)$ is the unique solution of the equation

$$(2-\alpha)^{\frac{\alpha+1}{\alpha-1}}x^\alpha = c\tau_r^{-1}\left(1 + \left\|\phi(\cdot)\int_0^\cdot \frac{\phi'(t)}{\phi(t)^2}\,dt\right\|_{L^\infty([0,\tau_r^{-1}])}\right)^\alpha.$$

Proof. It is sufficient to apply Lemma 5.5.1 and Lemma 5.5.2. ∎

Thus, given ϕ, the quantity in the right-hand-side of the inequalities (5.5.5), (5.5.6) and (5.5.7) can be computed explicitly. Let us give two applications of Theorem 5.5.5. If $\phi(t) := e^{-\lambda t}$ for $\lambda > 0$ and $T = +\infty$, then X^ϕ is the stable Ornstein-Uhlenbeck process of index α. Therefore, a direct computation in Theorem 5.5.5 implies the

Corollary 5.5.6. *Let Z be a symmetric stable process of index $\alpha \in (0,2)$ and Lévy measure $\nu(dy) = c|y|^{-\alpha-1}dy$, $c > 0$. Letting $f_{\alpha,x,\lambda}(t) := x(1+\lambda\alpha t)^{1/\alpha}$, $t \geq 0$, $\lambda > 0$, we have for all $x > 0$*

$$\mathbb{P}\left(\inf\{t \geq 0 : |Z_t| \geq f_{\alpha,x,\lambda}(t)\} > r\right) \leq \frac{1}{(1+\lambda\alpha r)^{\frac{c}{\lambda\,\alpha^2 2^{\alpha-1} x^\alpha}}}, \qquad r > 0.$$

If $\alpha \in (0,1]$, then for all $x > 0$ and all $r > 0$,

$$\begin{aligned}\mathbb{P}\left(\inf\{t \geq 0 : |Z_t| \geq f_{\alpha,x,\lambda}(t)\} \leq r\right) &\leq \frac{4c(2-(1+\lambda\alpha r)^{-\frac{1}{\alpha}})^\alpha \log(1+\lambda\alpha r)}{\lambda\alpha^2 x^\alpha} \\ &\leq \frac{16cr}{\alpha x^\alpha}.\end{aligned}$$

Finally, if $\alpha \in (1,2)$, then for all $x > 0$ and for all $0 \leq r < r_0(\alpha,x,\lambda)$, we have the estimate

$$\begin{aligned}\mathbb{P}\left(\inf\{t \geq 0 : |Z_t| \geq f_{\alpha,x,\lambda}(t)\} \leq r\right) &\leq \frac{(2-(1+\lambda\alpha r)^{-\frac{1}{\alpha}})^\alpha \log(1+\lambda\alpha r) K_c}{\lambda\alpha x^\alpha} \\ &\leq \frac{4rK_c}{x^\alpha},\end{aligned}$$

where K_c is the constant of Theorem 5.5.5 and $r_0(\alpha,x,\lambda)$ is the unique solution of the equation

$$\lambda\alpha x^\alpha = \frac{c(2-(1+\lambda\alpha r)^{-\frac{1}{\alpha}})^\alpha (\log(1+\lambda\alpha r))}{(2-\alpha)^{\frac{\alpha+1}{\alpha-1}}}.$$

Now, we present the case of the stable bridge. Given a symmetric stable process $Z = (Z_t)_{t\geq 0}$ of index $\alpha \in (0,2)$, there exists a Markov process $X^{(\mathrm{br})} = (X_t^{(\mathrm{br})})_{0\leq t\leq T}$ starting from 0 and ending in 0 at a finite time horizon T, such that its distribution $\mathbb{Q}$ is given by

$$d\mathbb{Q}_{|\mathscr{F}_t} = \frac{p_{T-t}(-X_t)}{p_T(0)} d\mathbb{P}_{|\mathscr{F}_t}, \quad t \in (0,T),$$

where p_t is a version everywhere positive of the distribution of the stable random variable Z_t, see [10, Chapter VIII]. The process $X^{(\mathrm{br})}$ is called a stable bridge. By e.g. Exercise 12.2 in [84], $X^{(\mathrm{br})}$ is the unique solution of the linear equation

$$X_t^{(\mathrm{br})} = Z_t - \int_0^t \frac{X_s^{(\mathrm{br})}}{T-s} ds, \quad t \in (0,T),$$

which rewrites by the integration by parts formula of Proposition 8.11 in [30] as

$$X_t^{(\mathrm{br})} = (T-t)\int_0^t \frac{dZ_s}{T-s}ds, \quad t \in (0,T).$$

Hence, the stable bridge $X^{(\mathrm{br})}$ is a stable-Markov process of Ornstein-Uhlenbeck type with parameter ϕ given by $\phi(t) = T-t$, $t \in [0,T]$. Thus, using Theorem 5.5.5, we get the

Corollary 5.5.7. *Let Z be a symmetric stable process of index $\alpha \in (1,2)$ and Lévy measure $\nu(dy) = c|y|^{-\alpha-1}dy$, $c > 0$. Letting $g_{\alpha,x,T}(t) := x(T^{1-\alpha} + (\alpha-1)t)^{1/(\alpha-1)}$, $t \geq 0$, we have for all $x > 0$ and all $r > 0$*

$$\begin{aligned}
\mathbb{P}\left(\inf\{t \geq 0 : |Z_t| \geq g_{\alpha,x,\lambda}(t)\} > r\right) &\leq \exp\left(-\frac{c}{\alpha 2^{\alpha-1}x^\alpha}\left(T - (T^{1-\alpha} + (\alpha-1)r)^{\frac{1}{1-\alpha}}\right)\right) \\
&= \exp\left(-\frac{c(Tg_{\alpha,x,T}(r) - x)}{\alpha 2^{\alpha-1} g_{\alpha,x,T}(r)x^\alpha}\right),
\end{aligned}$$

whereas for all $x > 0$ and for all $0 \leq r < r_0(\alpha,x,T)$, we have

$$\begin{aligned}
\mathbb{P}\left(\inf\{t \geq 0 : |Z_t| \geq g_{\alpha,x,T}(t)\} \leq r\right) &\leq \frac{K_c(Tg_{\alpha,x,T}(r) - x)(2Tg_{\alpha,x,T}(r) - x)^\alpha}{T^\alpha g_{\alpha,x,T}(r)^{\alpha+1}x^\alpha} \\
&\leq \frac{4TK_c}{x^\alpha},
\end{aligned}$$

where K_c is the constant of Theorem 5.5.5 and $r_0(\alpha,x,\lambda)$ is the unique solution of the equation

$$(2-\alpha)^{\frac{\alpha+1}{\alpha-1}}T^\alpha g_{\alpha,x,T}(r)^{\alpha+1}x^\alpha = c(Tg_{\alpha,x,T}(r) - x)(2Tg_{\alpha,x,T}(r) - x)^\alpha.$$

Remark 5.5.8. In the latter corollary, only the case $\alpha \in (1,2)$ is considered, since the time change techniques we use in the proof of Theorem 5.5.5 are not satisfied when $\alpha \in (0,1)$.

Chapitre 6

A convex domination principle for dependent Brownian and stable stochastic integrals

Ce chapitre fait l'objet d'un article en préparation, écrit avec Yutao Ma.

Abstract

Based on the forward-backward stochastic calculus developed in [59], we show in this note that under some boundedness assumptions, the sum of correlated Brownian and stable stochastic integrals is convex dominated by the independent sum of Gaussian and symmetric stable random variables.

6.1 Introduction

A real random variable X is said to be convex dominated by another real random variable Y if we have

$$\mathbb{E}\left[\phi(X)\right] \leq \mathbb{E}\left[\phi(Y)\right], \tag{6.1.1}$$

for any integrable convex function ϕ. Such a domination principle between random variables may be seen as a generalization of the classical moment inequalities developed in the theory of stochastic processes and entail many interesting results, among them Doob's and Burkholder's inequalities for martingales, Orlicz embeddings and deviation inequalities, see for instance [32, 59]. Recently, such a problem has been considered by Klein in his PhD thesis [58]. More precisely, using some stochastic calculus techniques, he showed that under several boundedness assumptions, a stochastic integral driven by a point process is convex dominated by a centered Poisson random variable whose intensity depends on the characteristics of the previous stochastic integral. This result has been extended in [59] to the general framework of martingales with jumps. In particular, they considered the case of Poisson random measures and derived a convex domination principle contained in their Theorem 5.1, which is stated as follows: a centered random variable admitting a representation in terms of dependent Brownian and compensated Poisson stochastic integrals, is convex dominated by the independent sum of centered Gaussian and Poisson random variables, provided some boundedness and moment assumptions are made on the integrated processes. However, the analysis does not concern for instance the case of martingales with unbounded jumps and infinite variance. Hence, the purpose of this note is to extend such a result by replacing the driving compensated Poisson process by a symmetric stable process whose sample paths are of infinite variation. Our approach relies on the Itô's formula for forward-backward martingales introduced in [59] and allows us to decouple the pair of Brownian and stable stochastic integrals.

6.2 Main result

Consider on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ a real standard Brownian motion $(W_t)_{t\geq 0}$ which is non-necessarily independent of a symmetric stable process $(Z_t)_{t\geq 0}$ of index $\alpha \in (1,2)$ and with stable Lévy measure defined on $\mathbb{R} \setminus \{0\}$ by

$$\sigma(dx) := \frac{cdx}{|x|^{\alpha+1}}, \quad c > 0. \tag{6.2.1}$$

Denoting the filtration $\mathcal{F}_t^{W,Z} := \sigma\left(W_s, Z_s : 0 \leq s \leq t\right)$, $t \geq 0$, we assume in the remainder of the paper that the Lévy processes $(W_t)_{t\geq 0}$ and $(Z_t)_{t\geq 0}$ are $(\mathcal{F}_t^{W,Z})_{t\geq 0}$-martingales. In other words, the latter condition states that the increments of $(W_t)_{t\geq 0}$ are independent of the past of $(Z_t)_{t\geq 0}$, and reciprocally. Let F be a random variable having the representation

$$F - \mathbb{E}[F] = \int_0^{+\infty} H_t dW_t + \int_0^{+\infty} K_t dZ_t, \tag{6.2.2}$$

where the bounded processes $(H_t)_{t\geq 0} \in L^\infty(\Omega, L^2(0,+\infty))$ and $(K_t)_{t\geq 0} \in L^\infty(\Omega, L^\alpha(0,+\infty))$ are $(\mathcal{F}_t^{W,Z})_{t\geq 0}$-predictable. We assume moreover that $(K_t)_{t\geq 0} \in L^2(\Omega \times (0,+\infty))$ so that the random variable $\int_0^{+\infty} K_t dZ_t$ is well-defined as a stochastic integral with respect to the Lévy-Itô decomposition of the stable process $(Z_t)_{t\geq 0}$. Let $S_{p,\alpha}$ be the set of convex functions on $\mathbb{R}$ with at most polynomial growth of order $p \in (0,\alpha)$ at infinity, i.e. $\lim_{|x|\to+\infty} |x|^{-p} |\phi(x)| < +\infty$. The main result of this note, whose proof is given in the next section, is the following

Theorem 6.2.1. *There exists a Gaussian random variable $\tilde{W}_{\beta_1}$ with variance*

$$\beta_1 := \left\| \int_0^{+\infty} |H_t|^2 dt \right\|_\infty,$$

independent of a symmetric stable random variable $\tilde{Z}_{\beta_2}$ of index α and with Lévy measure given on $\mathbb{R} \setminus \{0\}$ by

$$\tilde{\sigma}(dx) := \frac{\beta_2 \tilde{c} dx}{|x|^{\alpha+1}}, \quad \textit{with} \quad 0 < c \leq \tilde{c} \quad \textit{and} \quad \beta_2 := \left\| \int_0^{+\infty} |K_t|^\alpha dt \right\|_\infty,$$

such that for any $\phi \in S_{p,\alpha}$, we have the convex domination inequality

$$\mathbb{E}\left[\phi(F - \mathbb{E}F)\right] \leq \mathbb{E}\left[\phi\left(\tilde{W}_{\beta_1} + \tilde{Z}_{\beta_2}\right)\right]. \tag{6.2.3}$$

Remark 6.2.2. Note that the processes $(W_t)_{t\geq 0}$ and $(Z_t)_{t\geq 0}$ in the predictable representation of the random variable F are correlated, whereas the Gaussian and the symmetric stable random variables in the right-hand-side of (6.2.3) are independent. Hence the convex domination inequality (6.2.3) allows us to decouple the dependent pair $(W_t, Z_t)_{t\geq 0}$.

6.3 Proof of Theorem 6.2.1

Before proceeding to the proof of Theorem 6.2.1, let us introduce some notation and preliminary results on the forward-backward stochastic calculus recently developed in [59].

6.3.1 Forward-backward stochastic calculus

We endow the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with an increasing filtration $(\mathcal{F}_t)_{t\geq 0}$ and a decreasing filtration $(\mathcal{F}_t^*)_{t\geq 0}$, and we consider a $(\mathcal{F}_t)_{t\geq 0}$-martingale $(X_t)_{t\geq 0}$ and a backward $(\mathcal{F}_t^*)_{t\geq 0}$-martingale $(X_t^*)_{t\geq 0}$. For the sake of briefness, the (forward) processes considered in the remainder of this section are supposed to be right-continuous with left limits, whereas the backward martingales are naturally supposed to be left-continuous with right limits. We denote in the sequel by $(X_t^c)_{t\geq 0}$, $(X_t^{*c})_{t\geq 0}$, the continuous parts of the processes $(X_t)_{t\geq 0}$, $(X_t^*)_{t\geq 0}$, respectively. Let

$$\Delta X_s := X_s - X_{s-}, \text{and} \quad \Delta^* X_s^* := X_s^* - X_{s+}^*,$$

be the respective forward and backward jump sizes at time $s > 0$ and we denote by ν_X, ν_{X^*}, the $(\mathcal{F}_t)_{t\geq 0}$-, $(\mathcal{F}_t^*)_{t\geq 0}$-, dual predictable projections of the jumping measures μ_X, μ_{X^*}, of the martingales $(X_t)_{t\geq 0}$, $(X_t^*)_{t\geq 0}$, respectively. Define as the limits in probability the quadratic variations

$$[X, X]_t := \lim_{n\to+\infty} \sum_{i=1}^{n} |X_{t_i^n} - X_{t_{i-1}^n}|^2 \quad \text{and} \quad [X^*, X^*]_t := \lim_{n\to+\infty} \sum_{i=0}^{n-1} |X_{t_i^n}^* - X_{t_{i+1}^n}^*|^2,$$

for all refining subdivisions $(t_i^n)_{i=0,\ldots,n}$ of the time interval $[0, t]$, and denote the continuous brackets

$$\langle X^c, X^c\rangle_t := [X, X]_t - \sum_{0<s\leq t} |\Delta X_s|^2 \quad \text{and} \quad \langle X^{*c}, X^{*c}\rangle_t := [X^{*c}, X^{*c}]_t - \sum_{0\leq s<t} |\Delta^* X_s^*|^2.$$

According to the terminology of [48, 59], the pairs

$$(\nu_X(dt, dx), \langle X^c, X^c\rangle_t) \quad \text{and} \quad (\nu_{X^*}(dt, dx), \langle X^{*c}, X^{*c}\rangle_t)$$

are called the local characteristics of the processes $(X_t)_{t\geq 0}$ and $(X_t^*)_{t\geq 0}$, respectively.

Remark 6.3.1. Note that the bracket processes $([X,X]_t)_{t\geq 0}$, $(\langle X^c, X^c\rangle_t)_{t\geq 0}$, $([X^*,X^*]_t)_{t\geq 0}$ and $(\langle X^{*c}, X^{*c}\rangle_t)_{t\geq 0}$ are $(\mathcal{F}_t)_{t\geq 0}$-adapted but not $(\mathcal{F}_t^*)_{t\geq 0}$-adapted.

Now, let us quote Theorem 8.1 in [59], which is an Itô's type formula for forward-backward martingales:

Lemma 6.3.2. *Let $(X_t)_{t\geq 0}$, $(X_t^*)_{t\geq 0}$, be a $(\mathcal{F}_t^*)_{t\geq 0}$-adapted $(\mathcal{F}_t)_{t\geq 0}$-martingale and a $(\mathcal{F}_t)_{t\geq 0}$-adapted $(\mathcal{F}_t^*)_{t\geq 0}$-backward martingale, respectively. Then for any real-valued function $f \in C^2(\mathbb{R}^2)$, we have the Itô's formula:*

$$\begin{aligned} f(X_t, X_t^*) &= f(X_0, X_0^*) + \int_{0+}^t \frac{\partial f}{\partial x_1}(X_{s-}, X_s^*) dX_s + \frac{1}{2}\int_{0+}^t \frac{\partial^2 f}{\partial x_1^2}(X_{s-}, X_s^*) d\langle X^c, X^c\rangle_s \\ &\quad - \int_0^{t-} \frac{\partial f}{\partial x_2}(X_s, X_{s+}^*) d^* X_s^* - \frac{1}{2}\int_0^{t-} \frac{\partial^2 f}{\partial x_2^2}(X_s, X_{s+}^*) d\langle X^{*c}, X^{*c}\rangle_s \\ &\quad + \sum_{0<s\leq t} \left(f(X_{s-} + \Delta X_s, X_s^*) - f(X_{s-}, X_s^*) - \Delta X_s \frac{\partial f}{\partial x_1}(X_{s-}, X_s^*) \right) \\ &\quad - \sum_{0\leq s<t} \left(f(X_s, X_{s+}^* + \Delta^* X_s^*) - f(X_s, X_{s+}^*) - \Delta^* X_s^* \frac{\partial f}{\partial x_2}(X_s, X_{s+}^*) \right), \end{aligned}$$

where d and d^* are the forward and backward Itô's differential, respectively, and the integral with respect to $(\langle X^{*c}, X^{*c}\rangle_t)_{t\geq 0}$ is defined as a Stieltjes integral with respect to an (non-necessarily $(\mathcal{F}_t^*)_{t\geq 0}$-adapted) increasing process.

Remark 6.3.3. As noticed in [59], The crossing adaptedness assumption of Lemma 6.3.2 allows us to define properly the forward and the backward Itô's stochastic integrals.

6.3.2 Integrability of convex functions

In order for the convex domination principle of Theorem 6.2.1 to make sense, we have first to establish the following integrability property of convex functions in $S_{p,\alpha}$.

Lemma 6.3.4. *If the random variable F has the representation (6.2.2), then for any $\phi \in S_{p,\alpha}$, the random variable $\phi(F - \mathbb{E}F)$ is integrable.*

Proof. Let $\phi \in S_{p,\alpha}$. By the continuity of the convex function ϕ, we only have to check its integrability at infinity. Thus, it is sufficient to show that the centered random variable $F - \mathbb{E}[F]$ has a finite moment of order $p \in (0, \alpha)$. Since the process $(H_t)_{t\geq 0} \in L^2(\Omega \times (0, +\infty))$, we only have to verify the latter condition for the random variable $\int_0^{+\infty} K_t dZ_t$, and up to a conditioning argument, for the random variable $\int_0^T K_t dZ_t$, where $T > 0$ is a fixed time horizon. We have

$$\begin{aligned} \mathbb{E}\left[\left|\int_0^T K_t dZ_t\right|^p\right] &= \int_0^{+\infty} \mathbb{P}\left(\left|\int_0^T K_t dZ_t\right| \geq x^{1/p}\right) dx \\ &\leq \int_0^{+\infty} \mathbb{P}\left(\sup_{0\leq s\leq T}\left|\int_0^s K_t dZ_t\right| \geq x^{1/p}\right) dx. \end{aligned}$$

By the maximal inequality (2.11) in [51], there exists a constant $D_\alpha > 0$ depending on α such that for any $x > \mathbb{E}\left[\int_0^T |K_t|^2 dt\right]^{p/2}$, we have

$$\mathbb{P}\left(\sup_{0\le s\le T}\left|\int_0^s K_t dZ_t\right| \ge x^{1/p}\right) \;\le\; \frac{D_\alpha}{x^{\alpha/p}}\,\mathbb{E}\left[\int_0^T |K_t|^2 dt\right]^{\alpha/2}.$$

Hence, denoting $x_0 := \mathbb{E}\left[\int_0^T |K_t|^2 dt\right]^{p/2}$, we obtain

$$\mathbb{E}\left[\left|\int_0^T K_t dZ_t\right|^p\right] \;\le\; x_0 + D_\alpha \mathbb{E}\left[\int_0^T |K_t|^2 dt\right]^{\alpha/2} \int_{x_0}^{+\infty} \frac{dx}{x^{\alpha/p}},$$

which is finite provided $0 < p < \alpha$. The proof is complete. ■

6.3.3 Proof of Theorem 6.2.1

We are able to start the proof of the main Theorem 6.2.1, which is divided into several steps. First, we have to introduce a forward and a backward martingales (relying on the representation of the random variable F) with respect to an increasing and a decreasing filtration, respectively. Then we identify their local characteristics and take expectation in the Itô's formula for forward-backward martingales. Finally, we get the result by a limiting argument.

Let $(\tilde{W}_t)_{t\ge0}$ be a standard Brownian motion independent of a symmetric stable process $(\tilde{Z}_t)_{t\ge0}$ of index α and with Lévy measure given on $\mathbb{R}\setminus\{0\}$ by $\tilde{\sigma}(dx) := \tilde{c}|x|^{-\alpha-1}dx$. We assume that both processes are independent of the filtration $(\mathcal{F}_t^{W,Z})_{t\ge0}$. Consider the $(\mathcal{F}_t^{W,Z})_{t\ge0}$-martingale $(X_t)_{t\ge0}$ given by

$$X_t := \mathbb{E}[F - \mathbb{E}F|\mathcal{F}_t^{W,Z}], \quad t \ge 0.$$

By assumption, the symmetric stable process $(Z_t)_{t\ge0}$ is a martingale with respect to the filtration $(\mathcal{F}_t^{W,Z})_{t\ge0}$, and so is the stochastic integral $(\int_0^t K_s dZ_s)_{t\ge0}$, since it is constructed as the L^1-limit of square-integrable martingales. Using a similar argument for the Brownian part, one deduces that the martingale $(X_t)_{t\ge0}$ is identified as

$$X_t = \int_0^t H_s dW_s + \int_0^t K_s dZ_s, \quad t \ge 0.$$

Define the enlarged filtration $(\mathcal{F}_t)_{t\ge0}$ as

$$\mathcal{F}_t := \mathcal{F}_t^{W,Z} \vee \sigma(\tilde{W}_s, \tilde{Z}_s : s \ge 0), \quad t \ge 0,$$

then the process $(X_t)_{t\ge0}$ is still a martingale with respect to $(\mathcal{F}_t)_{t\ge0}$. Denote the continuous increasing processes

$$\gamma_t := \int_0^t |H_s|^2 ds \quad \text{and} \quad \tau_t := \int_0^t |K_s|^\alpha ds, \quad t \ge 0,$$

and consider the time-changed process

$$X_t^* = \tilde{W}_{\beta_1} - \tilde{W}_{\gamma_t} + \tilde{Z}_{\beta_2} - \tilde{Z}_{\tau_t}, \quad t \geq 0.$$

Finally, we endow the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the decreasing filtration given by

$$\mathcal{F}_t^* := \sigma(\tilde{W}_{\beta_1} - \tilde{W}_{\gamma_s}, \tilde{Z}_{\beta_2} - \tilde{Z}_{\tau_s} : s \geq t) \vee \mathcal{F}_\infty^{W,Z}, \quad t \geq 0.$$

Note that the correlated processes $(X_t)_{t\geq 0}$ and $(X_t^*)_{t\geq 0}$ are $(\mathcal{F}_t^*)_{t\geq 0}$- and $(\mathcal{F}_t)_{t\geq 0}$-adapted, respectively, and their dependence is given through $(\gamma_t)_{t\geq 0}$ and $(\tau_t)_{t\geq 0}$.
On the one hand, the independent processes $(\tilde{W}_t)_{t\geq 0}$ and $(\tilde{Z}_t)_{t\geq 0}$ are both independent of the $(\mathcal{F}_t^{W,Z})_{t\geq 0}$-measurable time-change processes $(\gamma_t)_{t\geq 0}$ and $(\tau_t)_{t\geq 0}$. On the other hand, they are centered and have independent increments. Hence, the process $(X_t^*)_{t\geq 0}$ is a $(\mathcal{F}_t^*)_{t\geq 0}$-backward martingale.
Now, let us identify the local characteristics of the martingales $(X_t)_{t\geq 0}$ and $(X_t^*)_{t\geq 0}$. First, the continuous brackets are the same, i.e.

$$d\langle X^c, X^c\rangle_t = |H_t|^2 d\langle W, W\rangle_t = |H_t|^2 dt, \quad t \geq 0,$$

and

$$d\langle X^{*c}, X^{*c}\rangle_t = d\langle \tilde{W}, \tilde{W}\rangle_{\gamma_t} = d\gamma_t = |H_t|^2 dt,$$

cf. Proposition 1.15 p.173 in [70].
If we set $Y_t := \int_0^t K_s dZ_s$, $t \geq 0$, then the following change of variables formula

$$\begin{aligned}
&\int_0^t \int_{-\infty}^{+\infty} (f(Y_s + K_s x) - f(Y_s) - K_s x f'(Y_s)) \frac{dxds}{|x|^{\alpha+1}} \\
&\quad = \int_0^t \int_{-\infty}^{+\infty} (f(Y_s + y) - f(Y_s) - y f'(Y_s)) |K_s|^\alpha \frac{dyds}{|y|^{\alpha+1}}, \quad t \geq 0,
\end{aligned}$$

available for any real-valued function $f \in C^1(\mathbb{R})$, allows us to identify the local characteristics of the jump parts of $(X_t)_{t\geq 0}$ as

$$\nu_X(dt, dx) = \nu_Y(dt, dx) = |K_t|^\alpha \, dt \, \sigma(dx),$$

whereas we have by Theorem 10.27 (b),(e) in [48]:

$$\nu_{X^*}(dt, dx) = d\tau_t \, \tilde{\sigma}(dx) = |K_t|^\alpha \, dt \, \tilde{\sigma}(dx).$$

Assume without loss of generality that the convex function $\phi \in C^2(\mathbb{R})$, since any convex function can be approximated by an increasing sequence of $C^2(\mathbb{R})$ convex functions. Using the Itô's formula of Lemma 6.3.2 and taking then expectation, which is allowed by Lemma 6.3.4 since $\phi \in S_{p,\alpha}$, we get for any $0 \leq s \leq t$:

$$\begin{aligned}
&\mathbb{E}\left[\phi(X_t + X_t^*)\right] - \mathbb{E}\left[\phi(X_s + X_s^*)\right] \\
&\quad = \mathbb{E}\left[\int_s^t \phi'(X_u + X_u^*) d\langle X^c, X^c\rangle_u\right] - \mathbb{E}\left[\int_s^t \phi'(X_u + X_u^*) d\langle X^{*c}, X^{*c}\rangle_u\right]
\end{aligned}$$

$$
\begin{aligned}
&+\mathbb{E}\left[\int_s^t \int_{-\infty}^{+\infty} \left(\phi(X_u + X_u^* + x) - \phi(X_u + X_u^*) - x\phi'(X_u + X_u^*)\right) \nu_X(du, dx)\right] \\
&-\mathbb{E}\left[\int_s^t \int_{-\infty}^{+\infty} \left(\phi(X_u + X_u^* + x) - \phi(X_u + X_u^*) - x\phi'(X_u + X_u^*)\right) \nu_{X^*}(du, dx)\right] \\
&= \mathbb{E}\left[\int_s^t \int_{-\infty}^{+\infty} \int_0^1 (1-\tau) x^2 \phi''(X_u + X_u^* + \tau x) |K_u|^\alpha \frac{(c - \tilde{c})}{|x|^{\alpha+1}}\, d\tau dx du\right].
\end{aligned}
$$

By the convexity of ϕ, the second derivative ϕ'' is non-negative and according to the comparison assumption on the weights $0 < c \leq \tilde{c}$, one deduces that the function $t \mapsto \mathbb{E}\left[\phi(X_t + X_t^*)\right]$ is non-increasing on $\mathbb{R}_+$. Since we have the null projection $\mathbb{E}\left[X_t^* | \mathcal{F}_t^{W,Z}\right] = 0$ for any $t \geq 0$, we obtain by Jensen's inequality

$$
\begin{aligned}
\mathbb{E}\left[\phi\left(\mathbb{E}\left[F - \mathbb{E}[F] | \mathcal{F}_t^{W,Z}\right]\right)\right] &= \mathbb{E}[\phi(X_t)] \\
&\leq \mathbb{E}[\phi(X_t + X_t^*)] \\
&\leq \mathbb{E}[\phi(X_0^*)] \\
&= \mathbb{E}\left[\phi\left(\tilde{W}_{\beta_1} + \tilde{Z}_{\beta_2}\right)\right].
\end{aligned}
$$

Letting t going to infinity in the left-hand-side above achieves the proof of the Theorem 6.2.1.

Annexe A

Concentration markovienne et calcul chaotique

Nous nous intéressons dans cette partie annexe à la généralisation en dimension infinie du phénomène de concentration de processus de naissance et de mort. Plus précisément, nous établissons une inégalité de déviation de type Poisson pour des fonctionnelles de la trajectoire d'un processus de naissance et de mort en utilisant un calcul chaotique. Récemment, quelques auteurs ont employé ce type de techniques afin d'établir des résultats de concentration pour des fonctionnelles de martingales normales à sauts satisfaisant la Propriété de Représentation Chaotique (en bref PRC) et dont le gradient de Malliavin, défini en abaissant le degré des intégrales stochastiques multiples engendrées par ces martingales normales, est borné dans un sens à préciser. Citons par exemple l'approche par des identités de covariance dans les articles [21, 22, 44], ou encore celle par des inégalités de Sobolev logarithmiques et la méthode de Herbst [83]. Néanmoins, on ne connaît pas l'interprétation probabiliste du gradient de Malliavin en dehors du cas à accroissements indépendants (cas brownien, cas Poisson et cas mixte brownien-Poisson), c'est-à-dire que l'on ne dispose pas en général d'expression explicite de ce gradient agissant sur des fonctionnelles cylindriques. La méthode que nous privilégions dans cette partie repose sur un calcul chaotique que l'on développe pour des processus de naissance et de mort, d'après les travaux de Biane [11] à propos de la PRC pour des chaînes de Markov à temps continu générales. En utilisant une formule de Clark-Ocone valable pour toute fonctionnelle de carré intégrable, ainsi que des techniques de martingales similaires à celles utilisées dans les chapitres 3 et 4 de la thèse, l'inégalité de déviation de type Poisson que nous établissons est comparable à celles démontrées par exemple dans les travaux [44, 83] sur l'espace de Poisson. Cependant, les hypothèses de type Lipschitz que nous imposons sont différentes car nous supposons seulement que la projection prévisible de l'opérateur de gradient, qui peut être calculée explicitement dans certains cas, est bornée, alors que l'expression du gradient reste inconnue.

Introduisons à présent le contexte de notre étude. Considérons un processus de naissance et de mort stable et conservatif $(X_t)_{t\geq 0}$ à valeurs entières et muni de sa filtration naturelle $(\mathscr{F}_t)_{t\geq 0}$. Son générateur infinitésimal $\mathcal{L}$ est donné par

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

et l'on suppose dans la suite que les fonctions de taux λ et ν sont strictement positives avec $\nu_0 = 0$, et minorées par un nombre strictement positif :

$$\lambda_* := \inf_{x\in\mathbb{N}} \lambda_x > 0, \quad \nu_* := \inf_{x\geq 1} \nu_x > 0.$$

Définissons les processus

$$M_t^{(1)} := \sum_{s\leq t} \lambda_{X_{s-}}^{-1/2} 1_{\{\Delta X_s = 1\}} - \int_0^t \lambda_{X_s}^{1/2} ds, \quad M_t^{(-1)} := \sum_{s\leq t} \nu_{X_{s-}}^{-1/2} 1_{\{\Delta X_s = -1\}} - \int_0^t \nu_{X_s}^{1/2} ds,$$

correspondant respectivement aux sauts positifs et négatifs de la chaîne $(X_t)_{t\geq 0}$. Biane a démontré le

Lemme A.0.1. *([11, lemme 1]) Les processus $(M_t^{(1)})_{t\geq 0}$ et $(M_t^{(-1)})_{t\geq 0}$ sont des martingales normales au sens de Meyer, c'est-à-dire*

$$\langle M^{(i)}, M^{(j)} \rangle_t = \delta_{ij} t, \quad i,j = 1, -1.$$

Lorsque la variable initiale X_0 est déterministe, ce que l'on suppose dans la suite de cette partie, les deux martingales $(M_t^{(1)})_{t\geq 0}$ et $(M_t^{(-1)})_{t\geq 0}$ engendrent aussi la filtration $(\mathscr{F}_t)_{t\geq 0}$. Ainsi, toute variable aléatoire qui est une fonction des trajectoires de ces deux martingales est aussi une fonction de celle du processus de naissance et de mort originel $(X_t)_{t\geq 0}$. L'idée est donc d'introduire un calcul chaotique pour les martingales normales précédentes afin d'en déduire un résultat de concentration pour des fonctionnelles de la trajectoire du processus $(X_t)_{t\geq 0}$.

Quelques éléments de calcul chaotique

Notons l'espace à deux points $E = \{-1, 1\}$ et considérons l'espace de Hilbert $L^2(E \times \mathbb{R}_+)$ muni du produit scalaire

$$\langle f, g \rangle := \sum_{z\in E} \int_0^{+\infty} f(z,t) g(z,t) dt.$$

On désigne par $L_s^2((E \times \mathbb{R}_+)^n)$ le sous-espace de $L^2((E \times \mathbb{R}_+)^n)$ constitué des fonctions symétriques f_n en leurs n variables couples, i.e. pour toute permutation σ dans le groupe symétrique $\mathbb{S}_n$ d'ordre n, on a

$$f_n((z_1,t_1), \ldots, (z_n,t_n)) = f_n\left((z_{\sigma(1)}, t_{\sigma(1)}), \ldots, (z_{\sigma(n)}, t_{\sigma(n)})\right).$$

Pour $t > 0$, on note $\Delta_n^t := \{((z_1,t_1), \ldots, (z_n,t_n)) \in (E \times \mathbb{R}_+)^n : t_i \in [0,t],\ i = 1, \ldots, n\}$ et pour toute fonction $f_n \in L_s^2((E \times \mathbb{R}_+)^n)$, on définit l'intégrale stochastique multiple par

$$I_n(f_n) \;:=\; n! \sum_{(z_1,\ldots,z_n)\in E^n} \int_0^{+\infty} \int_0^{t_n} \cdots \int_0^{t_2} f_n((z_1,t_1), \ldots, (z_n,t_n)) dM_{t_1}^{(z_1)} \cdots dM_{t_n}^{(z_n)}$$

$$= \quad n \sum_{z\in E} \int_0^{+\infty} I_{n-1}\left(f_n(*1_{\Delta_{n-1}^t}(*),(z,t)) \right) dM_t^{(z)}.$$

Si $f_n \in L^2((E\times\mathbb{R}_+)^n)$, on note $I_n(f_n) := I_n(\hat{f}_n)$, où $\hat{f}_n$ désigne la fonction symétrisée de f_n, c'est-à-dire

$$\hat{f}_n((z_1,t_1),\ldots,(z_n,t_n)) := \frac{1}{n!}\sum_{\sigma\in\mathbb{S}_n} f_n\left((z_{\sigma(1)},t_{\sigma(1)}),\ldots,(z_{\sigma(n)},t_{\sigma(n)})\right).$$

Comme cas particulier du résultat principal de l'article de Biane [11], valable pour des chaînes de Markov à temps continu plus générales, les martingales normales $(M_t^{(1)})_{t\geq 0}$ et $(M_t^{(-1)})_{t\geq 0}$ satisfont la PRC suivante :

Théorème A.0.2. *Toute variable aléatoire de carré intégrable $F \in L^2(\Omega) := L^2(\Omega,\mathscr{F}_\infty,\mathbb{P})$ peut s'écrire comme la somme infinie d'intégrales stochastiques multiples*

$$F = \mathbb{E}[F] + \sum_{n\geq 1} I_n(f_n),$$

où les fonctions $f_n \in L^2((E\times\mathbb{R}_+)^n)$ dépendent éventuellement de la condition initiale déterministe X_0.

Parallèlement aux cas des espaces de Wiener et de Poisson, l'opérateur I_n définit une isométrie de $L^2((E\times\mathbb{R}_+)^n)$ dans l'espace vectoriel engendré par les intégrales multiples $I_n(f_n)$, $f_n \in L^2((E\times\mathbb{R}_+)^n)$, dit chaos d'ordre n.

Définissons à présent des opérateurs de gradient et de divergence, objets usuels du calcul chaotique. Tout d'abord, on désigne par $D : \mathrm{Dom}\, D \subset L^2(\Omega) \to L^2(E\times\mathbb{R}_+\times\Omega)$ le gradient agissant sur les intégrales stochastiques multiples de la manière suivante :

$$D_{z,t} I_n(f_n) := n I_{n-1}(f_n(*,(z,t))), \quad z\in E, \quad t\geq 0,$$

où le domaine du gradient est donné par

$$\mathrm{Dom}\, D := \left\{ F = \mathbb{E}[F] + \sum_{n\geq 1} I_n(f_n) : \|DF\|^2_{L^2(E\times\mathbb{R}_+\times\Omega)} = \sum_{n\in\mathbb{N}} n n! \|f_n\|^2_{L^2((E\times\mathbb{R}_+)^n)} < +\infty \right\}.$$

Rappelons que la PRC du théorème A.0.2 entraîne trivialement la densité de cet espace dans $L^2(\Omega)$. L'opérateur de divergence $\delta : \mathrm{Dom}\,\delta \subset L^2(E\times\mathbb{R}_+\times\Omega) \to L^2(\Omega)$ est défini par

$$\delta(u) := I_{n+1}(\tilde{f}_{n+1}), \quad u(z,t) := I_n(f_{n+1}(*,(z,t))), \quad t\geq 0,$$

de domaine $\mathrm{Dom}\,\delta$, dense dans l'espace $L^2(E\times\mathbb{R}_+\times\Omega)$, donné par

$$\mathrm{Dom}\,\delta := \left\{ u(z,t) = \sum_{n\geq 0} I_n(f_{n+1}(*,(z,t))) : \sum_{n\geq 0}(n+1)! \|f_{n+1}\|^2_{L^2((E\times\mathbb{R}_+)^{n+1})} < +\infty \right\}.$$

Ici, la fonction $\tilde{f}_{n+1}$ désigne la symétrisation de f_{n+1} en ses $n+1$ variables couples.
Il n'est pas difficile de montrer d'après leur définition que ces deux opérateurs sont adjoints, i.e. que la formule d'intégration par partie suivante est vérifiée :

$$\mathbb{E}\left[\langle DF, u\rangle_{L^2(E\times\mathbb{R}_+)}\right] = \mathbb{E}\left[F\delta(u)\right], \quad F \in \operatorname{Dom} D, \quad u \in \operatorname{Dom}\delta. \tag{A.0.1}$$

À présent, soit $T : \mathbb{A} \to \mathbb{B}$ un opérateur linéaire, où $\mathbb{A}$ et $\mathbb{B}$ sont deux espaces vectoriels normés. On dit que T est fermable si pour toute suite $(F_n)_{n\in\mathbb{N}}$ convergeant vers 0 dans $\mathbb{A}$ et telle que $(TF_n)_{n\in\mathbb{N}}$ tend vers un élément G de $\mathbb{B}$, alors $G = 0$.
Il est classique dans le calcul chaotique qu'une formule d'intégration par parties de type (A.0.1) entraîne la fermabilité des opérateurs de gradient et de divergence :

Proposition A.0.3. *Les opérateurs de gradient D et de divergence δ sont fermables.*

Preuve. Soit $(F_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\operatorname{Dom} D$ convergeant vers 0 dans $L^2(\Omega)$ et telle que $(DF_n)_{n\in\mathbb{N}}$ tend vers un élément G de $L^2(E\times\mathbb{R}_+\times\Omega)$. Montrons alors que $G = 0$. Soit $u \in \operatorname{Dom}\delta$. Par la relation de dualité (A.0.1) et l'inégalité de Cauchy-Schwarz, on a

$$\begin{aligned}
\left|\mathbb{E}\left[\langle G, u\rangle_{L^2(E\times\mathbb{R}_+)}\right]\right| &\leq \left|\mathbb{E}\left[F_n\delta(u)\right] - \mathbb{E}\left[\langle G, u\rangle_{L^2(E\times\mathbb{R}_+)}\right]\right| + \left|\mathbb{E}\left[F_n\delta(u)\right]\right| \\
&= \left|\mathbb{E}\left[\langle DF_n - G, u\rangle_{L^2(E\times\mathbb{R}_+)}\right]\right| + \left|\mathbb{E}\left[F_n\delta(u)\right]\right| \\
&\leq \|DF_n - G\|_{L^2(E\times\mathbb{R}_+\times\Omega)}\|u\|_{L^2(E\times\mathbb{R}_+\times\Omega)} + \|F_n\|_{L^2(\Omega)}\|\delta(u)\|_{L^2(\Omega)}.
\end{aligned}$$

La quantité de droite tendant vers 0 lorsque n tend vers l'infini, on a ainsi

$$\mathbb{E}\left[\langle G, u\rangle_{L^2(E\times\mathbb{R}_+)}\right] = 0, \quad u \in \operatorname{Dom}\delta,$$

et par densité dans $L^2(E\times\mathbb{R}_+\times\Omega)$ du domaine $\operatorname{Dom}\delta$, on obtient finalement que la variable G est identiquement égale à 0, ce qui entraîne la fermabilité de l'opérateur D.
La preuve de la fermabilité de l'opérateur de divergence δ s'effectue de la même manière. ■

De plus, l'opérateur de divergence δ coïncide pour les processus adaptés avec l'intégrale d'Itô par rapport aux martingales normales $(M_t^{(1)})_{t\geq 0}$ et $(M_t^{(-1)})_{t\geq 0}$:

Proposition A.0.4. *Soit $u \in L^2(E\times\mathbb{R}_+\times\Omega)$ un processus $(\mathscr{F}_t)_{t\geq 0}$-adapté de carré intégrable. Alors $u \in \operatorname{Dom}\delta$ et on a l'identité*

$$\delta(u) = \sum_{z\in E}\int_0^{+\infty} u(z,t)\,dM_t^{(z)}.$$

Preuve. Supposons tout d'abord que le processus adapté u appartient au chaos d'ordre n. En d'autres termes, il s'écrit sous forme d'une intégrale stochastique multiple $u(z,t) = I_n(f_{n+1}(*, (z,t)))$. Par la formule de conditionnement triviale,

$$\mathbb{E}\left[I_n(f_n)|\mathscr{F}_t\right] = I_n\left(f_n(*1_{\Delta_n^t}(*))\right), \tag{A.0.2}$$

le processus adapté u satisfait la relation

$$u(z,t) = \mathbb{E}\left[I_n(f_{n+1}(*, (z,t)))|\mathscr{F}_t\right] = I_n\left(f_{n+1}(*1_{\Delta_n^t}(*), (z,t))\right).$$

Ainsi, on obtient

$$\begin{aligned}\delta(u) &= I_{n+1}(\tilde{f}_{n+1}) \\ &= \sum_{z\in E}\int_0^{+\infty} I_n\left(f_{n+1}(*1_{\Delta_n^t}(*),(z,t))\right) dM_t^{(z)} \\ &= \sum_{z\in E}\int_0^{+\infty} u(z,t)dM_t^{(z)}.\end{aligned}$$

Le cas général est établi par fermabilité de l'opérateur de divergence δ. ■

La formule de Clark-Ocone suivante, satisfaite par les martingales normales $(M_t^{(1)})_{t\geq 0}$, $(M_t^{(-1)})_{t\geq 0}$, est valable pour toute fonctionnelle de carré intégrable.

Théorème A.0.5. *Toute fonctionnelle de carré intégrable $F \in L^2(\Omega)$ se représente comme une intégrale stochastique du type*

$$F = \mathbb{E}[F] + \sum_{z\in E}\int_0^{+\infty} \mathbb{E}\left[D_{z,t}F|\mathscr{F}_t\right] dM_t^{(z)}. \tag{A.0.3}$$

Preuve. Soit $F \in \operatorname{Dom} D$. Par la PRC du théorème A.0.2, il existe des fonctions $f_n \in L^2((E\times\mathbb{R}_+)^n)$ telles que

$$\begin{aligned}F &= \mathbb{E}[F] + \sum_{n\geq 1} I_n(f_n) \\ &= \mathbb{E}[F] + \sum_{n\geq 1}\sum_{z\in E} n\int_0^{+\infty} I_{n-1}\left(f_n(*1_{\Delta_{n-1}^t}(*),(z,t))\right) dM_t^{(z)} \\ &= \mathbb{E}[F] + \sum_{n\geq 1}\sum_{z\in E}\int_0^{+\infty} \mathbb{E}\left[D_{z,t}I_n(f_n)|\mathscr{F}_t\right] dM_t^{(z)},\end{aligned}$$

où l'on a utilisé la formule de conditionnement (A.0.2) dans la dernière expression. Ainsi, la formule de Clark-Ocone (A.0.3) est démontrée dans le cas des fonctionnelles appartenant au domaine $\operatorname{Dom} D$ du gradient D. Il reste à montrer qu'elle peut être étendue à toute fonctionnelle de l'espace $L^2(\Omega)$. Supposons $F \in L^2(\Omega)$. Par densité, il existe une suite $(F_n)_{n\in\mathbb{N}} \subset \operatorname{Dom} D$ qui converge vers F dans $L^2(\Omega)$. En utilisant alors la formule de Clark-Ocone précédente ainsi que la proposition A.0.4, on obtient

$$F_n - \mathbb{E}[F_n] = \sum_{z\in E}\int_0^{+\infty} \mathbb{E}\left[D_{z,t}F_n|\mathscr{F}_t\right] dM_t^{(z)} = \delta(u_n),$$

où l'on a noté u_n le processus adapté $u_n(z,t) := \mathbb{E}\left[D_{z,t}F_n|\mathscr{F}_t\right]$. Comme $(F_n)_{n\in\mathbb{N}}$ converge vers F dans $L^2(\Omega)$ et que les martingales $(M_t^{(1)})_{t\geq 0}$ et $(M_t^{(-1)})_{t\geq 0}$ sont normales, la suite $(u_n)_{n\in\mathbb{N}}$ tend dans $L^2(E\times\mathbb{R}_+\times\Omega)$ vers un processus adapté, noté v, qui appartient donc au domaine $\operatorname{Dom}\delta$ de l'opérateur de divergence. De plus, $(\delta(u_n))_{n\in\mathbb{N}}$ converge dans $L^2(\Omega)$ vers la variable aléatoire centrée $F - \mathbb{E}[F]$. Ainsi, par fermabilité de l'opérateur de divergence δ, on a l'identité $F - \mathbb{E}[F] = \delta(v)$, et l'on termine la preuve de la formule de Clark-Ocone (A.0.3) en définissant l'élément $\mathbb{E}\left[D_{z,t}F|\mathscr{F}_t\right] := v(z,t)$. ■

Un résultat de concentration

Énonçons à présent le résultat principal de cette partie annexe, à propos de la concentration de fonctionnelles de processus de naissance et de mort. On note $a \wedge b := \min\{a, b\}$, $a, b \in \mathbb{R}$.

Théorème A.0.6. *Soit $F \in L^2(\Omega)$ une fonctionnelle de carré intégrable telle que pour $z \in E$, on ait la borne*

$$|\mathbb{E}\left[D_{z,t}F|\mathscr{F}_t\right]| \leq b, \quad d\mathbb{P}dt - p.s.,$$

avec de plus

$$\sum_{z\in E}\int_0^{+\infty} \mathbb{E}\left[D_{z,t}F|\mathscr{F}_t\right]^2 dt \leq v^2 \quad d\mathbb{P} - p.s.$$

Alors l'inégalité de déviation suivante est vérifiée :

$$\begin{aligned}&\mathbb{P}\left(F - \mathbb{E}[F] \geq x\right)\\ &\quad \leq \exp\left(\frac{x(\lambda_* \wedge \nu_*)^{1/2}}{b} - \left(\frac{x(\lambda_* \wedge \nu_*)^{1/2}}{b} + \frac{v^2(\lambda_* \wedge \nu_*)}{b^2}\right)\log\left(1 + \frac{bx}{v^2(\lambda_* \wedge \nu_*)^{1/2}}\right)\right).\end{aligned}$$

Preuve. Par la formule de Clark-Ocone du théorème A.0.5, on a

$$\begin{aligned}F - \mathbb{E}[F] &\leq \sup_{t\geq 0} \mathbb{E}\left[F|\mathscr{F}_t\right] - \mathbb{E}[F]\\ &= \sup_{t\geq 0}\sum_{z\in E}\int_0^t \mathbb{E}\left[D_{z,s}F|\mathscr{F}_s\right] dM_s^{(z)}\\ &= \sup_{t\geq 0} Z_t,\end{aligned}$$

où le processus $(Z_t)_{t\geq 0}$ est la $(\mathscr{F}_t)_{t\geq 0}$-martingale de carré intégrable donnée par

$$Z_t := \sum_{z\in E}\int_0^t \mathbb{E}\left[D_{z,s}F|\mathscr{F}_s\right] dM_s^{(z)}, \quad t \geq 0, \quad Z_0 = 0.$$

Les sauts de cette martingale vérifient pour tout $t > 0$ la majoration

$$\begin{aligned}|\Delta Z_t| &= |Z_t - Z_{t-}|\\ &= \left|\sum_{z\in E} \mathbb{E}\left[D_{z,t}F|\mathscr{F}_t\right]\Delta M_t^{(z)}\right|\\ &= \left|\mathbb{E}\left[D_{1,t}F|\mathscr{F}_t\right]\lambda_{X_{t-}}^{-1/2}1_{\{\Delta X_t = 1\}} + \mathbb{E}\left[D_{-1,t}F|\mathscr{F}_t\right]\nu_{X_{t-}}^{-1/2}1_{\{\Delta X_t=-1\}}\right|\\ &\leq \frac{b}{(\lambda_* \wedge \nu_*)^{1/2}}.\end{aligned}$$

De plus, les martingales $(M_t^{(1)})_{t\geq 0}$ et $(M_t^{(-1)})_{t\geq 0}$ étant normales, le crochet oblique du processus $(Z_t)_{t\geq 0}$ satisfait la borne

$$\langle Z, Z\rangle_t = \sum_{z\in E}\int_0^t \mathbb{E}\left[D_{z,s}F|\mathscr{F}_s\right]^2 d\langle M^{(z)}, M^{(z)}\rangle_s$$

$$= \sum_{z\in E}\int_0^t \mathbb{E}\left[D_{z,s}F|\mathscr{F}_s\right]^2 ds \leq v^2, \quad t\geq 0.$$

Par le lemme 23.19 de [56], le processus $(Y_t^{(\tau)})_{t\geq 0}$ défini pour tout $\tau > 0$ par

$$Y_t^{(\tau)} := \exp\left\{\tau Z_t - \tau^2\psi\left(\frac{\tau b}{(\lambda_*\wedge\nu_*)^{1/2}}\right)\langle Z,Z\rangle_t\right\}, \quad t\geq 0,$$

est une $(\mathscr{F}_t)_{t\geq 0}$-surmartingale, où ψ désigne la fonction $\psi(z) := z^{-2}(e^z - z - 1)$, $z > 0$. Ainsi, on obtient

$$\begin{aligned}
\mathbb{P}(F-\mathbb{E}[F]\geq x) &\leq \mathbb{P}\left(\sup_{t\geq 0} Z_t \geq x\right)\\
&\leq \inf_{\tau>0}\mathbb{P}\left(\sup_{t\geq 0} Y_t^{(\tau)} \geq \exp\left\{\tau x - \tau^2 v^2\psi\left(\frac{\tau b}{(\lambda_*\wedge\nu_*)^{1/2}}\right)\right\}\right)\\
&\leq \inf_{\tau>0}\exp\left\{-\tau x + \tau^2 v^2\psi\left(\frac{\tau b}{(\lambda_*\wedge\nu_*)^{1/2}}\right)\right\},
\end{aligned}$$

optimisation qui donne l'inégalité de déviation désirée. ■

Le résultat de déviation dans le théorème A.0.6 est comparable à ceux établis sur l'espace de Poisson dans les articles [44, 83], alors que les hypothèses de type Lipschitz imposées sur les fonctionnelles sont différentes. En effet, nous supposons seulement la bornitude de la projection prévisible sur la filtration $(\mathscr{F}_t)_{t\geq 0}$ de l'opérateur de gradient D et non la bornitude du gradient lui-même comme dans les travaux cités ci-dessus. L'intérêt d'introduire ce type d'hypothèses dans le cas des processus de naissance et de mort réside dans le fait que cette projection prévisible peut être calculée explicitement dans certains cas, alors que l'interprétation probabiliste du gradient reste inconnue. Pour illustrer notre propos, considérons par exemple la fonctionnelle simple $F = f(X_T) \in L^2(\Omega)$, où T est un horizon fini. La formule de Clark-Ocone du théorème A.0.5 entraîne

$$f(X_T) = \mathbb{E}[f(X_T)] + \sum_{z\in E}\int_0^{+\infty}\mathbb{E}\left[D_{z,t}f(X_T)|\mathscr{F}_t\right] dM_t^{(z)}. \tag{A.0.4}$$

Bien que l'expression explicite de DF soit inconnue, nous allons calculer sa projection prévisible sur la filtration $(\mathscr{F}_t)_{t\geq 0}$. Par la formule d'Itô appliquée à la $(\mathscr{F}_t)_{t\geq 0}$-martingale $(P_{T-t}f(X_t))_{0\leq t\leq T}$, où $P_{T-t}f(X_t) := \mathbb{E}\left[f(X_T)|\mathscr{F}_t\right]$ est le semigroupe du processus $(X_t)_{t\geq 0}$, on a

$$\begin{aligned}
P_{T-t}f(X_t) &= P_T f(X_0) + \int_0^t \lambda_{X_{s-}}^{1/2}\left(P_{T-s}f(X_{s-}+1) - P_{T-s}f(X_{s-})\right) dM_s^{(1)}\\
&\quad + \int_0^t \nu_{X_{s-}}^{1/2}\left(P_{T-s}f(X_{s-}-1) - P_{T-s}f(X_{s-})\right) dM_s^{(-1)}.
\end{aligned}$$

En choisissant alors $t = T$, on obtient par unicité de la représentation (A.0.4) les identités suivantes :

$$\mathbb{E}\left[D_{1,t}f(X_T)|\mathscr{F}_t\right] = \lambda_{X_{t-}}^{1/2}\left(P_{T-t}f(X_{t-}+1) - P_{T-t}f(X_{t-})\right),$$

$$\mathbb{E}\left[D_{-1,t}f(X_T)|\mathscr{F}_t\right] = \nu_{X_{t-}}^{1/2}\left(P_{T-t}f(X_{t-}-1) - P_{T-t}f(X_{t-})\right).$$

Notons enfin que les hypothèses de bornitude du théorème A.0.6 sont satisfaites si le semi-groupe est lipschitzien et les fonctions de transition du générateur infinitésimal bornées, ce qui nous ramène à l'étude des courbures discrètes du processus $(X_t)_{t\geq 0}$ définies dans les chapitres 3 et 4 de la thèse.
Un autre exemple de fonctionnelle, dont l'expression du gradient est inconnue alors que sa projection prévisible sur la filtration $(\mathscr{F}_t)_{t\geq 0}$ est calculable, est le suivant :

$$F = \mathbb{E}[F] + \sum_{z\in E}\int_0^{+\infty} u_t^{(z)} dM_t^{(z)}.$$

En effet, toujours par unicité de la formule de Clark-Ocone (A.0.3), on peut identifier les projections prévisibles du gradient sur la filtration $(\mathscr{F}_t)_{t\geq 0}$ comme

$$\mathbb{E}\left[D_{1,t}F|\mathscr{F}_t\right] = u_t^{(1)}, \quad \mathbb{E}\left[D_{-1,t}F|\mathscr{F}_t\right] = u_t^{(-1)},$$

alors que l'on ignore l'expression du gradient DF. Ainsi, en supposant le processus $(u_t)_{t\geq 0}$ dans l'espace $L^\infty(\mathbb{R}_+\times\Omega)\cap L^\infty(\Omega, L^2(0,+\infty))$, le théorème A.0.6 peut s'appliquer pour ce type de fonctionnelles.

Annexe B

Convergence d'un processus stable vers un mouvement brownien standard

L'objectif de cette courte partie est d'utiliser très simplement les techniques employées dans le chapitre 5 pour l'obtention d'inégalités maximales, afin d'établir la convergence d'un processus symétrique stable renormalisé vers un mouvement brownien standard.
Soit $\big(\Omega, (\mathscr{F}_t)_{t\in[0,1]}, \mathscr{F}, \mathbb{P}\big)$ un espace de probabilité filtré et soit $(X_t)_{t\in[0,1]}$ un processus réel symétrique $\alpha-$stable $(\mathscr{F}_t)_{t\in[0,1]}$-adapté (càdlàg) d'index $\alpha \in (1,2)$ et sans partie gaussienne (en bref $S\alpha S$). Sa fonction caractéristique est donnée par

$$\varphi_{X_t}(u) = \exp\left(t\int_{\mathbb{R}} \left(e^{iuy} - 1\right)\nu(dy)\right),$$

où $\nu(dy) := c|y|^{-\alpha-1}dy$, $c > 0$, désigne la mesure de Lévy stable sur $\mathbb{R}$. Sa fonction caractéristique se réécrit alors comme

$$\varphi_{X_t}(u) = e^{-t\rho|u|^\alpha},$$

où ρ est la quantité

$$\rho = \frac{\pi c}{\alpha\Gamma(\alpha)\sin(\pi\alpha/2)},$$

et Γ est la fonction Gamma usuelle.
Avant d'énoncer notre théorème de convergence, introduisons tout d'abord quelques résultats intermédiaires. Le lemme suivant est une adaptation de la méthode utilisée dans le chapitre 5 de la thèse afin d'établir une inégalité maximale pour le processus symétrique stable $(X_t)_{t\in[0,1]}$.

Lemme B.0.7. *L'inégalité maximale suivante est satisfaite :*

$$\mathbb{P}\left(\sup_{0\le s\le t} |X_s| \ge x\right) \le \frac{4ct}{\alpha(2-\alpha)x^\alpha}, \quad x > 0. \tag{B.0.1}$$

Preuve. Le processus stable $(X_t)_{t\in[0,1]}$ étant symétrique, sa décomposition de Lévy-Itô est donnée par $X_t = X_t^{(x-)} + X_t^{(x+)}$, où

$$X_t^{(x-)} := \int_0^t \int_{|y|\leq x} y\,(\mu - \sigma)(dy, ds), \quad X_t^{(x+)} := \int_0^t \int_{|y|>x} y\,\mu(dy, ds), \quad t \in [0,1],$$

et μ est une mesure de Poisson aléatoire sur $\mathbb{R} \times [0,1]$ d'intensité $\sigma(dy, dt) = \nu(dy) \otimes dt$. Ainsi,

$$\begin{aligned} \mathbb{P}\left(\sup_{0\leq s\leq t} |X_s| \geq x\right) &\leq \mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(x-)}| + \sup_{0\leq s\leq t} |X_s^{(x+)}| \geq x\right) \\ &\leq \mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(x-)}| \geq x\right) + \mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(x+)}| > 0\right). \end{aligned} \tag{B.0.2}$$

Notons que le processus de Poisson composé $(X_t^{(x+)})_{t\in[0,1]}$ est constant par morceaux tout comme son processus supremum $(\sup_{0\leq s\leq t} |X_s^{(x+)}|)_{t\in[0,1]}$, et que sa loi possède un atome en 0 à chaque instant $t > 0$. Désignons par T_1^x le premier instant de saut du processus de Poisson $(\mu(\{y \in \mathbb{R} : |y| > x\} \times [0,t]))_{t\in[0,1]}$ sur l'ensemble $\{y \in \mathbb{R} : |y| > x\}$. Si presque sûrement T_1^x survient après l'instant t, alors le processus de Poisson composé $(X_t^{(x+)})_{t\in[0,1]}$ (et donc son processus supremum) est identiquement égal à 0 sur l'intervalle $[0,t]$. Par conséquent, on a

$$\begin{aligned} \mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(x+)}| > 0\right) &= 1 - \mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(x+)}| = 0\right) \\ &\leq 1 - \mathbb{P}(T_1^x > t) \\ &= 1 - \exp\left(-t\nu\left(\{y \in \mathbb{R} : |y| > x\}\right)\right) \\ &\leq t\nu\left(\{y \in \mathbb{R} : |y| > x\}\right) \\ &= \frac{2ct}{\alpha x^\alpha}, \end{aligned} \tag{B.0.3}$$

où l'on a utilisé dans la deuxième égalité le fait que la variable aléatoire T_1^x suit une loi exponentielle de paramètre $\nu(\{y \in \mathbb{R} : |y| > x\})$, cf. par exemple [76, théorème 21.3].

Considérons à présent la martingale de carré intégrable $(X_t^{(x-)})_{t\in[0,1]}$. Par l'inégalité de Doob et la formule d'isométrie pour les intégrales stochastiques dirigées par une mesure de Poisson aléatoire,

$$\begin{aligned} \mathbb{P}\left(\sup_{0\leq s\leq t} |X_s^{(x-)}| \geq x\right) &\leq \frac{1}{x^2}\,\mathbb{E}\left[\left|\int_0^t \int_{|y|\leq x} y\,(\mu - \sigma)(dy, d\tau)\right|^2\right] \\ &= \frac{1}{x^2}\,\mathbb{E}\left[\int_0^t \int_{|y|\leq x} y^2\,\nu(dy)d\tau\right] \\ &= \frac{t}{x^2} \int_{|y|\leq x} y^2\,\nu(dy), \end{aligned}$$

c'est-à-dire

$$\mathbb{P}\left(\sup_{0\leq s\leq t}|X_s^{(x-)}|\geq x\right)\leq\frac{2ct}{(2-\alpha)x^\alpha}. \tag{B.0.4}$$

Finalement, l'utilisation des inégalités (B.0.3) et (B.0.4) dans l'inégalité (B.0.2) achève la preuve. ■

Considérons l'espace de Skorohod D constitué des fonctions càdlàg sur l'intervalle $[0,1]$ et muni de la distance de Skorohod d, que l'on définit de la façon suivante, cf. [12] :

$$d(x,y):=\inf_{\lambda\in\Lambda}\left\{\|\lambda-I\|\vee\|x-y\lambda\|\right\},$$

où l'ensemble Λ est donné par

$$\Lambda:=\left\{\lambda:[0,1]\to[0,1]\,\text{continue, strictement croissante, avec}\,\lambda(0)=0,\ \lambda(1)=1\right\},$$

$\|x\|:=\sup_{t\in[0,1]}|x(t)|$, et I est la fonction identité sur $[0,1]$. Le théorème 2.3 de Giné et de Marcus dans l'article [39] fournit des critères permettant d'établir des résultats de convergence dans l'espace de Skorohod (D,d).

Théorème B.0.8. *(Giné-Marcus) Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires à valeurs dans (D,d) telle que*

(i) les distributions finidimensionnelles sont faiblement convergentes.

(ii) il existe $\beta>1/2$, $\gamma>0$ et une fonction F sur $[0,1]$ croissante et continue à droite telle que pour tout $0\leq t_1\leq t\leq t_2\leq 1$,

$$\mathbb{P}\left(|X_n(t)-X_n(t_1)|\geq x,|X_n(t_2)-X_n(t)|\geq x\right)\quad\leq\quad x^{-\gamma}\left(F(t)-F(t_1)\right)^\beta\left(F(t_2)-F(t)\right)^\beta.$$

(iii) pour tout $\epsilon>0$, on a les limites

$$\lim_{\delta\downarrow 0}\sup_{n\geq 1}\mathbb{P}\left(\sup_{s,t\in[1-\delta,1)}|X_n(t)-X_n(s)|>\epsilon\right)=0, \tag{B.0.5}$$

$$\lim_{\delta\downarrow 0}\sup_{n\geq 1}\mathbb{P}\left(|X_n(\delta)-X_n(0)|>\epsilon\right)=0. \tag{B.0.6}$$

Alors la suite de distributions des variables aléatoires X_n dans (D,d) converge faiblement et la limite est déterminée par les limites des lois finidimensionnelles.

À présent, nous sommes en mesure d'énoncer le résultat principal de cette partie.

Théorème B.0.9. *Soit $(X_t^{(\alpha)})_{t\in[0,1]}$ un processus $S\alpha S$ d'index $\alpha\in(1,2)$ et de mesure de Lévy stable $\nu(dx):=2^{-1}(2-\alpha)|x|^{-\alpha-1}dx$. Alors la famille de lois des variables aléatoires $(\sup_{0\leq s\leq t}X_s^{(\alpha)})_{t\in[0,1]}$ paramétrées par α et à valeurs dans l'espace (D,d) converge faiblement lorsque α tend vers 2 vers la distribution du processus supremum d'un mouvement brownien standard.*

Preuve. Afin d'établir le résultat, nous allons démontrer la convergence du processus $(X_t^{(\alpha)})_{t\in[0,1]}$ vers un mouvement brownien standard en vérifiant les conditions (i), (ii) et (iii) du théorème B.0.8, puis établir la continuité de la fonction supremum sur l'espace de Skorohod (D, d), et enfin nous allons utiliser un résultat de transfert de convergence faible par une fonction continue.

Tout d'abord, les accroissements du processus $(X_t^{(\alpha)})_{t\in[0,1]}$ étant indépendants, la convergence des lois finidimensionnelles vers celles d'un mouvement brownien standard $(W_t)_{t\in[0,1]}$ provient du cas unidimensionnel, qui est immédiat. Ainsi, la condition (i) est satisfaite. Par l'indépendance et la stationnarité des accroissements du processus $(X_t^{(\alpha)})_{t\in[0,1]}$, ainsi que le lemme B.0.7, on a

$$\begin{aligned}\mathbb{P}\left(|X_t^{(\alpha)} - X_{t_1}^{(\alpha)}| \geq x, |X_{t_2}^{(\alpha)} - X_t^{(\alpha)}| \geq x\right) &= \mathbb{P}\left(|X_{t-t_1}^{(\alpha)}| \geq x\right)\mathbb{P}\left(|X_{t_2-t}^{(\alpha)}| \geq x\right) \\ &\leq \frac{4(t-t_1)(t_2-t)}{\alpha^2 x^{2\alpha}}.\end{aligned}$$

Par conséquent, en choisissant $\beta = 1$, $\gamma = 2\alpha$ et $F(t) = t$, $t \in [0,1]$, la condition (ii) est vérifiée.
Établissons à présent les identités (B.0.5) et (B.0.6) de la condition (iii). Pour tout $\epsilon > 0$,

$$\mathbb{P}\left(\sup_{s,t\in[1-\delta,1)} |X_t^{(\alpha)} - X_s^{(\alpha)}| > \epsilon\right) \leq \mathbb{P}\left(\sup_{t\in[1-\delta,1)} |X_t^{(\alpha)} - X_{1-\delta}^{(\alpha)}| > \epsilon/2\right),$$

et un léger changement dans la preuve du lemme B.0.7 montre que l'on a

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in[1-\delta,1)} |X_t^{(\alpha)} - X_{1-\delta}^{(\alpha)}| > \epsilon/2\right) &\leq \frac{2^{\alpha+1}\delta}{\alpha\epsilon^\alpha} \\ &\leq \frac{8\delta}{\epsilon \wedge \epsilon^2}.\end{aligned}$$

Ainsi,

$$\lim_{\delta\downarrow 0} \sup_{\alpha\in(1,2)} \mathbb{P}\left(\sup_{s,t\in[1-\delta,1)} |X_t^{(\alpha)} - X_s^{(\alpha)}| > \epsilon\right) = 0.$$

Sachant que $X_0^{(\alpha)} = 0$, le lemme B.0.7 entraîne

$$\begin{aligned}\mathbb{P}\left(|X_\delta^{(\alpha)}| > \epsilon\right) &\leq \frac{2\delta}{\alpha\epsilon^\alpha} \\ &\leq \frac{2\delta}{\epsilon \wedge \epsilon^2},\end{aligned}$$

et donc on obtient la limite

$$\lim_{\delta\downarrow 0} \sup_{\alpha\in(1,2)} \mathbb{P}\left(|X_\delta^{(\alpha)}| > \epsilon\right) = 0.$$

Ainsi, la convergence du processus $(X_t^{(\alpha)})_{t\in[0,1]}$ vers le mouvement brownien standard $(W_t)_{t\in[0,1]}$ est démontrée.

Établissons à présent la continuité de la fonction supremum sur l'espace de Skorohod (D,d), c'est-à-dire que la fonction

$$h : D \to D;\, x \mapsto \sup_{s\in[0,\cdot]} x(s)$$

est continue sur (D,d). Par définition de la distance de Skorohod d, une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de (D,d) tend vers $x \in (D,d)$ si et seulement s'il existe une suite $(\lambda_n)_{n\in\mathbb{N}} \subset \Lambda$ telle que λ_n converge uniformément vers la fonction identité I, et que $x_n \circ \lambda_n$ converge uniformément vers x. Afin de montrer la continuité de h, il est suffisant de prouver que $h(x_n)$ converge dans (D,d) vers $h(x)$, ce qui est immédiat car $\sup_{s\in[0,\lambda_n(\cdot)]} x_n(s)$ converge uniformément vers $\sup_{s\in[0,\cdot]} x(s)$ lorsque n tend vers l'infini.

Finalement, un résultat de transfert de convergence faible par la fonction continue h sur (D,d), nommé en anglais le *Continuous Mapping Theorem* que l'on peut trouver à la page 20 de l'ouvrage de Billingsley [12], entraîne la convergence désirée. ■

Bibliographie

[1] L. Alili and P. Patie. On the first crossing times of a Brownian motion and a family of continuous curves. *C. R. Math. Acad. Sci. Paris*, 340(3):225–228, 2005.

[2] C. Ané. Grandes déviations en temps petit pour un processus de Markov à espace d'états dénombrable. *C. R. Acad. Sci. Paris Sér. I Math.*, 325(9):1025–1028, 1997.

[3] C. Ané and M. Ledoux. On logarithmic Sobolev inequalities for continuous time random walks on graphs. *Probab. Theory Related Fields*, 116(4):573–602, 2000.

[4] D. Applebaum. *Lévy processes and stochastic calculus*, volume 93 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2004.

[5] D. Applebaum. Lévy-type stochastic integrals with regularly varying tails. *Stoch. Anal. Appl.*, 23(3):595–611, 2005.

[6] D. Bakry. On Sobolev and logarithmic Sobolev inequalities for Markov semigroups. In *New trends in stochastic analysis (Charingworth, 1994)*, pages 43–75. World Sci. Publishing, River Edge, NJ, 1997.

[7] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[8] R.F. Bass. Stochastic differential equations driven by symmetric stable processes. In *Séminaire de Probabilités, XXXVI*, volume 1801 of *Lecture Notes in Math.*, pages 302–313. Springer, Berlin, 2003.

[9] R.F. Bass and D.A. Levin. Harnack inequalities for jump processes. *Potential Anal.*, 17(4):375–388, 2002.

[10] J. Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

[11] P. Biane. Chaotic representations for finite Markov chains. *Stochastics and Stochastics Reports*, 30:61–68, 1989.

[12] P. Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999.

[13] G. Blower and F. Bolley. Concentration of measure on product spaces with applications to markov processes. *Studia Math.*, 175:47–72, 2006.

[14] R.M. Blumenthal and R.K. Getoor. Sample functions of stochastic processes with stationary independent increments. *J. Math. Mech.*, 10:493–516, 1961.

[15] S.G. Bobkov and F. Götze. Discrete isoperimetric and Poincaré-type inequalities. *Probab. Theory Related Fields*, 114(2):245–277, 1999.

[16] S.G. Bobkov and C. Houdré. Isoperimetric constants for product probability measures. *Ann. Probab.*, 25(1):184–205, 1997.

[17] S.G. Bobkov, C. Houdré, and P. Tetali. λ_∞, vertex isoperimetry and concentration. *Combinatorica*, 20(2):153–172, 2000.

[18] S.G. Bobkov and M. Ledoux. Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, 107(3):383–400, 1997.

[19] S.G. Bobkov and M. Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.*, 156(2):347–365, 1998.

[20] P. Brémaud. *Markov chains, Gibbs fields, Monte Carlo simulation, and queues*, volume 31 of *Texts in Applied Mathematics.* Springer-Verlag, New York, 1999.

[21] J.C. Breton and C. Houdré. On finite range stable-type concentration. To appear in *Theory Probab. Appl.*, 2006.

[22] J.C. Breton, C. Houdré, and N. Privault. Dimension-free and infinite variance tail estimates on poisson space. Preprint, 2004.

[23] T. Cacoullos and V. Papathanasiou. Characterizations of distributions by generalizations of variance bounds and simple proofs of the CLT. *J. Statist. Plann. Inference*, 63(2):157–171, 1997.

[24] P. Cattiaux and A. Guillin. Deviation bounds for additive functionals of markov processes. To appear in *ESAIM Probab. Stat.*, 2006.

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

[26] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in analysis*, pages 195–199. Princeton Univ. Press, Princeton, N. J., 1970.

[27] L. H. Y. Chen and J. H. Lou. Characterization of probability distributions by Poincaré-type inequalities. *Ann. Inst. H. Poincaré Probab. Statist.*, 23(1):91–110, 1987.

[28] M.F. Chen. Estimation of spectral gap for Markov chains. *Acta Math. Sinica*, 12(4):337–360, 1996.

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

[30] R. Cont and P. Tankov. *Financial modelling with jump processes.* Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[31] P. DaiPra, A.M. Paganoni, and G. Posta. Entropy inequalities for unbounded spin systems. *Ann. Probab.*, 30(4):1959–1976, 2002.

[32] C. Dellacherie and P.A. Meyer. *Probabilités et potentiel. Chapitres V à VIII*, volume 1385 of *Actualités Scientifiques et Industrielles.* Hermann, Paris, revised edition, 1980.

[33] J.D. Deuschel and D.W. Stroock. *Large deviations*, volume 137 of *Pure and Applied Mathematics.* Academic Press Inc., Boston, MA, 1989.

[34] P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.*, 1(1):36–61, 1991.

[35] H. Djellout, A. Guillin, and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.*, 32(3B):2702–2732, 2004.

[36] S.S. Dragomir. *Some Gronwall type inequalities and applications.* Nova Science Publishers Inc., Hauppauge, NY, 2003.

[37] S.N. Ethier and T.G. Kurtz. *Markov processes, Characterization and convergence.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.

[38] 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, Berlin, 2005.

[39] E. Giné and M.B. Marcus. The central limit theorem for stochastic integrals with respect to Lévy processes. *Ann. Probab.*, 11(1):58–77, 1983.

[40] M. Gourcy and L. Wu. Logarithmic Sobolev inequalities of diffusions for the L2-metric. *Potential Anal.*, 25(1):77–102, 2006.

[41] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.

[42] C. Houdré. Remarks on deviation inequalities for functions of infinitely divisible random vectors. *Ann. Probab.*, 30(3):1223–1237, 2002.

[43] C. Houdré and P. Marchal. On the concentration of measure phenomenon for stable and related random vectors. *Ann. Probab.*, 32(2):1496–1508, 2004.

[44] C. Houdré and N. Privault. Concentration and deviation inequalities in infinite dimensions via covariance representations. *Bernoulli*, 8(6):697–720, 2002.

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

[46] C. Houdré and P. Tetali. Isoperimetric invariants for product Markov chains and graph products. *Combinatorica*, 24(3):359–388, 2004.

[47] H. Hult and F. Lindskog. Extremal behavior of stochastic integrals driven by regularly varying Lévy processes. To appear in *Ann. Probab.*, 2006.

[48] J. Jacod. *Calcul stochastique et problèmes de martingales*, volume 714 of *Lecture Notes in Mathematics*. Springer, Berlin, 1979.

[49] J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, second edition, 2003.

[50] A. Joulin. A new Poisson-type deviation inequality for the empirical distribution of ergodic birth-death processes. Submitted.

[51] A. Joulin. On maximal inequalities for stable stochastic integrals. To appear in *Potential Analysis*, 2006.

[52] A. Joulin. Poisson-type deviation inequalities for curved continuous time Markov chains. Submitted, 2006.

[53] A. Joulin and Y. Ma. A convex domination principle for dependent Brownian and stable stochastic integrals. In preparation.

[54] A. Joulin and N. Privault. A logarithmic Sobolev inequality for an interacting spin system under a geometric reference measure. To appear in the Proceedings of the conference *Quantum Probability and White Noise Analysis* of Levico (Italy), february 2005.

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

[56] O. Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, 1997.

[57] J. Kallsen and A.N. Shiryaev. Time change representation of stochastic integrals. *Teor. Veroyatnost. i Primenen.*, 46(3):579–585, 2001.

[58] T. Klein. Inégalités de concentration, martingales et arbres aléatoires. PhD Thesis, Université de Versailles-Saint-Quentin, 2003.

[59] T. Klein, Y. Ma, and N. Privault. Convex concentration inequalities and forward-backward stochastic calculus. *Electron. J. Probab.*, 11(20):486–512, 2006.

[60] Y. Kondratiev, T. Kuna, and O. Kutoviy. On relations between a priori bounds for measures on configuration spaces. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 7(2):195–213, 2004.

[61] 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.

[62] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2001.

[63] P. Lezaud. Chernoff-type bound for finite Markov chains. *Ann. Appl. Probab.*, 8(3):849–867, 1998.

[64] P. Lezaud. Chernoff and Berry-Esséen inequalities for Markov processes. *ESAIM Probab. Statist.*, 5:183–201 (electronic), 2001.

[65] F. Martinelli. Lectures on Glauber dynamics for discrete spin models. In *Lectures on probability theory and statistics (Saint-Flour, 1997)*, volume 1717 of *Lecture Notes in Math.*, pages 93–191. Springer, Berlin, 1999.

[66] K. Marton. A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.*, 6(3):556–571, 1996.

[67] L. Miclo. An example of application of discrete Hardy's inequalities. *Markov Process. Related Fields*, 5(3):319–330, 1999.

[68] P. Patie. On some first passage time problems, motivated by financial applications. Ph.D. Thesis, ETH Zurich, 2004.

[69] W.E. Pruitt. The growth of random walks and Lévy processes. *Ann. Probab.*, 9(6):948–956, 1981.

[70] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, Berlin, third edition, 1999.

[71] E. Rio. Inégalités de Hoeffding pour les fonctions lipschitziennes de suites dépendantes. *C. R. Acad. Sci. Paris Sér. I Math.*, 330(10):905–908, 2000.

[72] P. Robert. *Stochastic networks and queues*, volume 52 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, french edition, 2003. Stochastic Modelling and Applied Probability.

[73] J. Rosiński and W.A. Woyczyński. On Itô stochastic integration with respect to p-stable motion: inner clock, integrability of sample paths, double and multiple integrals. *Ann. Probab.*, 14(1):271–286, 1986.

[74] L. Saloff-Coste. Lectures on finite Markov chains. In *Lectures on probability theory and statistics (Saint-Flour, 1996)*, volume 1665 of *Lecture Notes in Math.*, pages 301–413. Springer, Berlin, 1997.

[75] P.M. Samson. Concentration of measure inequalities for Markov chains and Φ-mixing processes. *Ann. Probab.*, 28(1):416–461, 2000.

[76] K. Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[77] M. Schmuckenschläger. Martingales, Poincaré type inequalities, and deviation inequalities. *J. Funct. Anal.*, 155(2):303–323, 1998.

[78] T. Stoyanov. Isoperimetric and related constants for graphs and Markov chains. PhD thesis, Georgia Institute of Technology, 2001.

[79] K.T. Sturm and M.K. Von Renesse. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.

[80] M. Talagrand. A new isoperimetric inequality and the concentration of measure phenomenon. In *Geometric aspects of functional analysis (1989–90)*, volume 1469 of *Lecture Notes in Math.*, pages 94–124. Springer, Berlin, 1991.

[81] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[82] L. Wu. A deviation inequality for non-reversible Markov processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 36(4):435–445, 2000.

[83] L. Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Related Fields*, 118(3):427–438, 2000.

[84] M. Yor. *Some aspects of Brownian motion. Part II.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1997.

[85] B. Zegarliński. Analysis of classical and quantum interacting particle systems. In *Quantum interacting particle systems (Trento, 2000)*, volume 14 of *QP–PQ: Quantum Probab. White Noise Anal.*, pages 241–336. World Sci. Publishing, River Edge, NJ, 2002.