# Contributions to functional inequalities and limit theorems on the configuration space

Ronan Herry

Université du Luxembourg
Faculté des sciences, des techniques et de la communication
Thèse n°PhD-FSTC-2018-76
Université Paris-Est Marne-la-Vallée
École doctorale de mathématiques et STIC

THÈSE

présentée le 03.12.2018 au Luxembourg en vue d'obtenir le grade de

DOCTEUR EN MATHÉMATIQUES

par

Ronan HERRY

# Contributions to functional inequalities and limit theorems on configuration spaces

**Jury**

Professeur Vlad BALLY, Vice-président
Université Paris-Est Marne-la-Vallée

Professeur Nathael GOZLAN, Directeur de thèse
Université Paris 5 Descartes

Professeur Michel LEDOUX, Rapporteur
Université Toulouse 3 Paul Sabatier

Professeur Ivan NOURDIN, Président
Université du Luxembourg

Professeur Giovanni PECCATI, Directeur de thèse
Université du Luxembourg

Professeur Matthias REITZNER, Rapporteur
Université d'Osnabrück

*À Bouba et Zeïde.*

Perdu. A quelques pas de la maison, cependant, à guère plus que trois jets de pierre.
Là où retombe la flèche qui fût lancée au hasard.

---

Yves Bonnefoy

Freedom is the freedom to say that two plus two makes four. If that is granted, all else follows.

---

George Orwell

I have developed the habit of studying finite-dimensional random phenomenons from an infinite dimensional point of view.

---

Kiyosi Itô

# Contributions to functional inequalities and limit theorems on configurations space

**Abstract.** We present functional inequalities and limit theorems for point processes. We prove a modified logarithmic Sobolev inequalities, a Stein inequality and an exact fourth moment theorem for a large class of point processes including mixed binomial processes and Poisson point processes. The proofs of these inequalities are inspired by the Malliavin-Stein approach and the Γ-calculus of Bakry-Emery. The implementation of these techniques requires a development of a stochastic analysis for point processes. As point processes are essentially discrete, we design a theory to study non-diffusive random objects. For Poisson point processes, we extend the Stein inequality to study stable convergence with respect to limits that are conditionally Gaussian. Applications to Poisson approximations of Gaussian processes and random geometry are given. We discuss transport inequalities for mixed binomial processes and their consequences in terms of concentration of measure.

On a generic metric measured space, we present a refinement of the notion of concentration of measure that takes into account the parallel enlargement of distinct sets. We link this notion of improved concentration with the eigenvalues of the metric Laplacian and with a version of the Ricci curvature based on multi-marginal optimal transport.

**Keywords.** Limit theorems; functional inequalities; stochastic analysis; optimal transport; point processes; Malliavin calculus; Γ-calculus.

# Inégalités fonctionnelles et théorèmes limites sur les espaces de configurations

**Résumé.** Nous présentons des inégalités fonctionnelles pour les processus ponctuels. Nous prouvons une inégalité de Sobolev logarithmique modifiée, une inégalité de Stein et un théorème du moment quatrième sans terme de reste pour une classe de processus ponctuels qui contient les processus binomiaux et les processus de Poisson. Les preuves reposent sur des techniques inspirées de l'approche de Malliavin-Stein et du calcul avec l'opérateur carré du champ de Bakry-Émery. Pour mettre en œuvre ces techniques nous développons une analyse stochastique pour les processus ponctuels. Plus généralement, nous mettons au point une théorie d'analyse stochastique sans hypothèse de diffusion. Dans le cadre des processus de Poisson ponctuels, l'inégalité de Stein est généralisée pour étudier la convergence stable vers des limites conditionnellement gaussiennes. Nous appliquons ces résultats pour approcher des processus Gaussiens par des processus de Poisson composés et pour étudier des graphes aléatoires. Nous discutons d'inégalités de transport et de leur conséquence en terme de concentration de la mesure pour les processus binomiaux dont la taille de l'échantillon est aléatoire.

Sur un espace métrique mesuré quelconque, nous présentons un développement de la concentration de la mesure qui prend en compte l'agrandissement parallèle d'ensembles disjoints. Cette concentration améliorée donne un contrôle de toutes les valeurs propres du Laplacien métrique. Nous discutons des liens de cette nouvelle notion avec une version de la courbure de Ricci qui fait intervenir le transport à plusieurs marginales.

**Mots clefs.** Théorèmes limites ; inégalités fonctionnelles ; analyse stochastique ; transport optimal ; processus ponctuels ; calcul de Malliavin ; opérateur carré du champ.

# Contents

# PRÉAMBULE

Cette thèse réunit la majorité des résultats obtenus sous la direction de N. GOZLAN et G. PECCATI durant les trois dernières années. De manière informelle, j'ai cherché à comprendre les propriétés structurelles de certains objets aléatoires. En mathématiques, les objets aléatoires ont été introduits pour décrire certains phénomènes dont nous avons une trop faible compréhension pour en avoir un suivi détaillé. Par exemple, si je joue à pile ou face, pour peu que la pièce ne soit pas truquée, il me sera impossible de prédire le résultat du prochain jet. Par contre je sais que, après "un grand nombre" de lancers, il devrait y avoir "a peu près" autant de piles que de faces. Un des problèmes fondamentaux en probabilités est de donner un sens précis à cette intuition. On peut lire à ce sujet le traité de L. BACHELIER (1901) [11].

De manière moins anecdotique, le mouvement brownien décrit la trajectoire de petites particules qui flottent à l'intérieur d'une graine de pollen (observé par R. BROWN en 1827). La petite particule est heurtée dans tous les sens par les molécules du fluide de pollen, donnant lieu à un mouvement saccadé impossible à suivre en détail. L'approximation continue (c'est-à-dire quand on fait tendre le temps entre chaque choc vers $0$, ou de manière équivalente quand on fait tendre le nombre de molécules dans le pollen vers l'infini) de ce mouvement saccadé est le mouvement brownien. Il s'agit du prototype de déplacement pour la physique des milieux continus mais une étude classique (déterministe) de ces propriétés est impossible : on ne peut pas suivre la trajectoire de la particule. Une étude probabiliste du mouvement brownien, dont la première a été conduite par L. BACHELIER (1900) [10], permet de dégager d'excellentes propriétés structurelles du mouvement brownien. A cet égard, on cite également les travaux précurseurs de N. WIENER (1930) [149] ainsi que de K. ITÔ (1944) [69] et (1951) [67]. En fait, le mouvement brownien (et ses généralisations, les mesures aléatoires gaussiennes) constitue une des briques fondamentales du calcul des probabilités moderne et l'école française dite "de Strasbourg" sous l'impulsion de P.-A. MEYER a eu une influence majeure sur le développement du calcul par rapport au mouvement brownien. A ce sujet, on pourra consulter, entre autres, les volumes du Séminaire de probabilités [99].

Le pendant discret ou discontinu du mouvement brownien est le processus de Poisson. Il a été introduit à différentes époques par différentes personnes de manière indépendante pour modéliser des phénomènes variés (télécommunication, radioactivité, cosmologie). L'exemple historique provient de S.-D. POISSON (1837) [130] et concerne l'évolution du nombre d'actes criminels au cours du temps. Par exemple, on veut modéliser le nombre d'individus qui seront condamnés durant l'année à Paris. On peut supposer que les condamnations passées n'influent pas sur les condamnations futures (on parle d'indépendance entre les condamnations). De plus, on suppose que le temps entre chaque condamnation suit une loi exponentielle (il est peu probable que deux condamnations se suivent de manière très proche mais après un certain temps il est quasiment certain qu'il y aura une nouvelle condamnation). Sous ces seules deux hypothèses, [130] montre que le nombre de condamnations à chaque instant suit une loi explicite qui ne dépend que d'un paramètre : le nombre moyen de condamnations. Cette loi porte le nom de loi de Poisson et le processus obtenu quand on fait varier le temps tout au long de l'année s'appelle le processus de Poisson. C'est un processus qui à un temps associe un nombre entier et n'augmente que par saut de taille $1$ (on peut montrer sous les hypothèses du modèle qu'on ne peut pas avoir deux condamnations exactement en même temps).

Le mouvement brownien et le processus de Poisson présentés précédemment sont des processus dits temporels : à un temps $t$, l'un associe un nombre réel aléatoire et l'autre un nombre entier aléatoire. Il est assez naturel de considérer les versions dites spatiales de ces objets aléatoires : au lieu d'être indexés par un temps $t$ les processus sont indexés par une position $x$. On parle aussi de *champ aléatoire*. Le passage de la version temporelle à la version spatiale nécessite un saut conceptuel : on perd l'interprétation dynamique du processus ; ce n'est plus une particule qui se déplace ou une quantité qui évolue mais plutôt un champ statique défini en tout point dont l'intensité (aléatoire) varie avec la position. Typiquement, un champ brownien servira à représenter des champs invariants par rotation tandis qu'un processus de Poisson spatial servira à compter le nombre d'évènements indépendants arrivés en chaque point de l'espace (par exemple, les chutes de météorites sur la Lune). La définition formelle de ces objets spatiaux nécessite une attention particulière. En effet, la seule définition raisonnable d'un champ de Poisson spatial $\eta$, satisfait $\eta_x = 0$ pour toute position $x$. De même, la seule définition du champ Brownien spatial (plus couramment appelé dans le jargon *champ libre gaussien*) $G$, vérifie $G_x = \infty$ pour toute position $x$.

La description de nos deux processus temporels comme des éléments aléatoires de l'espace des fonctions de la demie droite (qui représente le temps) vers les nombres réels (pour le mouvement brownien on peut même choisir cet élément dans l'espace des fonctions continues) est suffisante pour leur étude et en ce sens satisfaisante. Cette description comme une fonction aléatoire cesse d'être satisfaisante lorsque l'on passe aux processus spatiaux. Comme je l'ai expliqué, en tant que fonctions de l'espace, $\eta$ est constamment nul et que $G$ est constamment infini ; en tant que fonctions ces deux objets ne présentent aucun intérêt.

C'est à ce niveau que le saut conceptuel intervient. On doit regarder $G$ et $\eta$ non pas comme des fonctions (éventuellement très compliquées) de l'espace des positions (typiquement représenté par $\mathbb{R}^2$ ou $\mathbb{R}^3$ qui sont des espaces vectoriels de dimension respective $2$ et $3$ et donc en particulier finie) mais comme des fonctions (linéaires, donc très simples) d'un certain "gros" espace $\mathscr{H}$ qui contient lui même des fonctions (qui satisfont elles même certaines propriétés de régularité). L'espace $\mathscr{H}$ est choisi de manière à ce que $G$ et $\eta$ ont les propriétés que l'on attend d'eux (et ce n'est pas nécessairement le même $\mathscr{H}$ pour $G$ et pour $\eta$). Cette construction présente un premier avantage : il suffit de faire varier $\mathscr{H}$ parmi tous les ensembles acceptables et on obtient une grande famille de processus. En fait, on peut construire des champs browniens et de Poisson avec comme espace de paramètre (presque) n'importe quel ensemble (et plus seulement les temps ou les positions). Un second avantage, qui est plus de l'ordre de l'esthétisme intellectuel : cette construction permet d'étudier avec le même formalisme les processus temporels et les processus spatiaux. On cherche ensuite à établir de propriétés générales, c'est-à-dire qui ne dépendent pas du $\mathscr{H}$ choisis.

N. WIENER (1938) [150] a adopté ce point de vue et a montré que le mouvement brownien et le processus de Poisson (et leur généralisation en terme de mesure aléatoires) sont au cœur de la compréhension de nombreux phénomènes physiques. Combinés au travaux de N. WIENER & A. WINTNER (1943) [151], cela a permis de commencer le développement d'une vraie théorie de l'analyse par rapport à $G$ et $\eta$. Ce sont avec les travaux de K. ITÔ (1951) [67] et (1956) [68] que cette analyse a pris ses lettres de noblesse : Itô montrent que les fonctions de $G$ et de $\eta$ (sous certaines hypothèses de régularité) admettent une décomposition suivant des fréquences (comme pour les modes de Fourier).

Dans cette thèse je me suis intéressé aux inégalités fonctionnelles et aux théorèmes limites relativement à $G$ et $\eta$ et surtout par rapport à des processus ponctuels (qui sont des généralisations du processus de Poisson et que l'on va présenter plus loin). Pour faciliter la compréhension, prenons le cas le plus simple. Pour cela on prend $G$ le champ brownien qui est paramétré par un ensemble qui ne contient qu'un seul élément. Dans ce cas on peut voir le champ comme non dépendent d'un paramètre et le considérer comme un seul nombre réel aléatoire. Dans ce cas, l'objet obtenu est bien connu il s'agit d'une variable aléatoire gaussienne qui prend ses valeurs suivants la fameuse courbe de Gauss en forme de cloche. On peut positionner un point sur la droite réelle à distance $G$ de l'origine. On obtient alors une position aléatoire sur les réels et on dit qu'elle est *distribuée selon la loi de $G$* que l'on note $\gamma$. Étant donné n'importe quel intervalle, la probabilité que cette position aléatoire soit dans cet intervalle est égale à l'aire sous la courbe en cloche délimitée par cet intervalle. Bien sûr on pourrait considérer d'autres lois qui correspondent à d'autres nombres aléatoires. De façon simplifiée, pour tout intervalle la loi d'un nombre aléatoire donne la probabilité que ce nombre appartienne à l'intervalle en question. Il est important de noté que deux objets aléatoires peuvent avoir la même loi sans pour autant qu'ils soient égaux. Par exemple, du à la symétrie de la courbe en cloche $\gamma$ par rapport à l'axe des abscisses $G$ et $-G$ ont la même loi mais sont différents.

La loi $\gamma$ en forme de cloche a une très bonne propriété. Regardons le mouvement brownien d'une petite particule à l'intérieur du fluide de pollen et considérons uniquement les déplacements le long d'un axe fixe (disons gauche droite pour l'observateur, identifié à la droite réelle). Ajoutons un potentiel qui va empêcher la particule de trop s'éloigner de son point de départ, l'évolution résultante est appelée *dynamique d'Ornstein-Uhlenbeck*. La loi $\gamma$ est invariante pour cette dynamique : si la position initiale de la particule le long de l'axe est distribuée suivant $\gamma$, alors la position de la particule à tout temps est encore distribuée suivant $\gamma$. De plus on peut vérifier que la loi gaussienne est la seule loi invariante. On en déduit le corollaire immédiat : le seul équilibre possible pour la dynamique d'Orstein-Uhlenbeck est la loi gaussienne. De là, on peut se demander pour quelles loi initiales $\mu$ l'équilibre $\gamma$ sera atteint. Si on note $\{P_t^*,\, t \geq 0\}$ la dynamique d'Ornstein-Uhlenbeck, c'est à dire que étant donnée une loi initiale $\mu$ le symbole $P_t^*\mu$ représente la loi de la particule dans le fluide au temps $t$, on a vu que $P_t^*\gamma = \gamma$ pour tout temps $t$ et on se demande étant donné une loi initiale $\mu$ combien la loi $P_t^*\mu$ est proche de $\gamma$. Pour cela on a besoin de considérer une façon de mesurer à quel point la loi $\mu$ est proche de $\gamma$. On se contentera de n'importe quelle fonction $\mathcal{F}$ qui prend en entrée une loi et qui donne en sortie un nombre réel positif tel que $\mathcal{F}(\mu) = 0$ si et seulement si $\mu = \gamma$. Il existe de nombreuses fonctions $\mathcal{F}$ qui satisfont cette contrainte :

(i) L'*entropie relativement à $\gamma$*, notée $\mathcal{H}(\mu|\gamma)$ qui peut être vue comme la différence d'énergie entre le système à l'état $\mu$ et l'énergie minimale du système à l'état d'équilibre $\gamma$.

(ii) La *distance de Wasserstein relativement à* $\gamma$, $\mathcal{W}_2(\mu|\gamma)$ qui peut être vue comme la distance moyenne parcourue quand on déplace $\mu$ vers $\gamma$ de manière optimale.

L'inégalité de Sobolev logarithmique prouvée par L. GROSS (1975) [61] nous dit que quelque soit la mesure $\mu$, $\mathcal{H}(\mu|\gamma) \leq \frac{1}{2}\mathcal{I}(\mu|\gamma)$. Ici $\mathcal{I}(\mu|\gamma)$ est l'*information relative à* $\gamma$ que l'on peut penser comme la dérivée de l'entropie. Ainsi on peut voir, l'inégalité de Sobolev logarithmique comme une inégalité différentielle qui donne de la convergence exponentielle vers l'équilibre de la dynamique de Ornstein-Uhlenbeck pour toute condition initiale (d'énergie finie), de manière symbolique on écrit : $\mathcal{H}(P_t^*\mu|\gamma) \leq \mathrm{e}^{-2t}\,\mathcal{H}(\mu|\gamma)$. D'autres inégalités fonctionnelles existent pour la gaussienne ; la plupart se déduisent directement de l'inégalité de Sobolev logarithmique. Comme par exemple certaines inégalités de transport (voir F. OTTO & C. VILLANI (2000) [118]). L'inégalité de Sobolev logarithmique comme les inégalités de transport que l'on en déduit sont intrinsèquement continues. N. GOZLAN, C. ROBERTO, P.-M. SAMSON & P. TETALI (2017) [58] ont introduit des inégalités de transport dites "généralisées" qui sont a priori mieux adaptés pour les objets discrets.

Les théorèmes limites eux cherchent à comprendre combien la distribution d'une statistique est éloignée d'une distribution cible (typiquement la gaussienne $\gamma$). Une statistique est une valeur réelle aléatoire qui résume les propriétés d'un objet plus complexe. Par exemple on peut regarder la statistique qui à un mouvement brownien associe son aire sous la courbe jusqu'au temps $1$. Dans ce cas la statistique est linéaire et le calcul exact de sa distribution ne présente aucune difficulté. La compréhension de statistiques non linéaires dans le cas gaussien et Poissonnien a été grandement facilité par l'introduction par I. NOURDIN & G. PECCATI (2009) [112] de la méthode de Malliavin-Stein. Cette méthode exploite les propriétés de symétrie de $P_t$ afin d'obtenir des bornes explicites sur la distance de la distribution d'un statistique à $\gamma$. G. PECCATI, J. L. SOLÉ, M. S. TAQQU & F. UTZET (2010) [121] et C. DÖBLER & G. PECCATI (2018) [40] ont ensuite adapté cette méthode au cadre de Poisson.

Dans cette thèse je me suis donc intéressé à des résultats similaires pour processus ponctuels. Un processus ponctuel peut être vu comme une manière (aléatoire) de repartir une population suivant des classes. Un processus ponctuel peut donc être vu comme une famille dénombrable de points aléatoire $X_i$, $X_i$ représentant la classe de l'individu $i$. Le cardinal de cette famille noté $N$ est lui aussi aléatoire. En général, on préfère noter le processus ponctuel comme une mesure aléatoire $\mu = \sum_{i=1}^{N} \delta_{X_i}$. Comme les processus de Poisson, les processus ponctuels sont des objets de nature discrète (c'est-à-dire avec des sauts, pas continus, non-diffusifs).

J'ai obtenu les résultats suivants.

- J'ai proposé une méthode systématique pour obtenir des inégalités fonctionnelles et des théorèmes limites pour certains objets non-diffusifs sous une hypothèse de décomposition en fréquence.
- J'ai montré que les processus ponctuels satisfont les hypothèses du cadre d'étude précèdent, en particulier j'ai construit la décomposition en fréquence.
- J'ai utilisé mon résultat pour unifier l'analyse des processus de Poisson, de l'hypercube et des processus ponctuels.

- J'ai étudié des inégalités de transport (classiques et généralisées) pour certains processus ponctuels par des méthodes directes.

- J'ai généralisé les théorèmes limites pour les processus de Poisson afin de considérer des limites gaussiennes dont la variance est aléatoire.

Dans un domaine plus géométrique je me suis aussi intéressé à la manière dont certaines quantités géométriques canoniques (les valeurs propres de l'opérateur de Laplace-Beltrami) contrôlent la vitesse à laquelle des ensembles disjoints grossissent.

# Acknowledgements

This thesis was shaped in light of the knowledge and experience of two exceptional researchers: Nathael Gozlan and Giovanni Peccati. I received their advice and their trust in the freedom they gave me in the preparation and writing of my thesis. This dissertation, being at the same time faithful and original, expresses my gratitude towards what they gave me better than any words could ever express it. Arriving at the situation that gave birth to this document is more or less a random trajectory along which I crossed paths with individuals who slowly but steadily attracted me to being an apprentice researcher. Djalil Chafaï, back in Dauphine, gave me a leg up by supervising my master's project and helping me to get it published. He, Guillaume Carlier, Jacques Féjoz, Bénedicte Hass, Jimmy Lambolley and Joseph Lehec make up the numerous outstanding professors who made me realize that mathematics is much more than computing. Looking back, I am grateful to my parents that raised me with "le démon des maths" and to my grandparents who made draw geometric figures and play logic games.

After I started my PhD, I arrived in Luxembourg in a small but cohesive department where I felt quickly at ease, learning and producing. Christian Döbler was nice enough to share with me, in addition to his deadly precision regarding mathematics, a true friendship that helped me navigate the development of my thesis. Discussing with Robert Baumgarth, Simon Campese, Gauthier Dierickx, Ivan Nourdin, Maurizia Rossi, James Thompson, Anna Vidotto, Guangqu Zheng and other members of the Probability group, I learned to listen, to read, to present my and other's works and, most importantly, to not always understand. I overused the time and patience of my friend François Petit learning from him about subjects apparently unrelated to probability but that turned out quite helpful. I am also thankful to him and Yannick Voglaire for helping me to prepare for life after the thesis, and for numerous dinners full of anecdotes and more serious discussion. This thesis was also facilitated by the work of the two secretaries in Luxembourg: Katharina Heil and Marie Leblanc. They simplified most of the paperwork and allowed me to focus on research. Finally, the academic and non-academic life in Luxembourg is not as boring as one could expect and I would like to warmly thank everybody in the department for the time spent here.

In Paris, I benefited from discussions that took place in and around the monthly CTOP seminar organized by Nathael Gozlan and Christian Leonard at the IHP whose pillars are: Matthieu Fradelizi, Olivier Guédon and Paul-Marie Samson. Scientific life in Paris (and in the rest of the world) is very intense and, of course, not limited to this seminar. I am thankful to the MAP5 in Paris Descartes University that hosts me and for the interesting dicussions and workshops with: Nathalie Eisenbaum, Anne Estrade, Eric Luçon, Raphaël Lachièze-Rey and Maurizia Rossi. I would like also to thank informally the organizers of seminars, schools and conferences I attended in Paris and all around the world during my PhD which helped me widen my mathematical horizons, gave me the opportunity to present my own work and to meet mathematicians of my generation. I am particularly happy to have met and befriended: Alex Dunlap, Erik Bates, Clayton Barnes, Guillaume Baverez Giovanni Conforti, Lorenzo Dello Schiavo, Jason Ledwige, Ilaria Mondello, Raquel Perales, Chiara Rigoni, Luca Tamanini, David Tewodrose.

Finally, I am honoured to be able to present my work in front of a jury of such exceptional researchers. The works of Vlad Bally, Michel Ledoux, Ivan Nourdin and Matthias Reitzner represent to me the most rigorous and cutting-edge research. Having such professors assess my thesis is not only a recognition of my work but it also testifies to the openness of the scientific world towards new research. I hope this dissertation will meet their expectations and contribute, to some extent, to the edifice of mathematics always in renewal.

# General introduction

## 1.1. Preliminaries and notations

All random variables will be defined on a sufficiently big probability space $(\Omega, \mathscr{W}, \mathbb{P})$ and the symbol $\mathbb{E}$ will be systematically used to denote the expectation, *i. e.* integration with respect to $\mathbb{P}$, and $\mathbb{V}\mathrm{ar}$ the variance. Unless otherwise specified, relations involving random variables are always understood in an almost-sure sense. Given a symmetric positive $d \times d$ matrix $C$, the symbol $\mathbf{N}(0, C)$ will designate the law of a centered Gaussian vector with covariance $C$. The symbols $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{N}$ and $\mathbb{N}_{>0}$ will always designate the set of real numbers, non-negative real numbers, non-negative integers and positive integers respectively. For $q \in \mathbb{N}_{>0}$, we write $[q] = \{1, \dots, q\}$ and $[0] = \emptyset$.

1.1.1. **Norms and differential calculus.** For $x, y \in \mathbb{R}^d$, we write $\langle x, y\rangle_{\ell^2}$ for the standard scalar product of $x$ and $y$ and, for $p \in [1, \infty]$, we write $|x|_{\ell^p}$ for $(\sum |x_i|^p)^{1/p}$ (or $\max |x_i|$ if $p = \infty$). We will always regard the space of $p$-linear functionals of $\mathbb{R}^d$ as the linear space $\mathbb{R}^{d^p}$. In particular, given two matrices $A$ and $B$ of size $d \times d$, we write $\langle A, B\rangle_{\ell^2}$ for $\mathrm{tr}(A^T B)$, $|A|_{\ell^2}$ for ${\langle A, A\rangle_{\ell^2}}^{1/2}$, $|A|_{\ell^\infty}$ for $\max |a_{ij}|$ and $|A|_{\ell^1}$ for $\sum |a_{ij}|$. For a matrix $A$, we also let

$$|A|_{op} = \sup_{|x|_{\ell^2}=1} |Ax|_{\ell^2}. \tag{1.1.1.1}$$

For $k \in \mathbb{N} \cup \{\infty\}$, the space of $k$ times continuously differentiable functions from $\mathbb{R}^d$ to $\mathbb{R}$ is denoted $\mathscr{C}^k(\mathbb{R}^d)$. For $\phi \in \mathscr{C}^k(\mathbb{R}^d)$, we will write $\nabla^k \phi$ for the $k$-th derivative that we identify with a $k$-form over $\mathbb{R}^d$. In particular, for all $x \in \mathbb{R}^d$, $\nabla^k \phi(x)$ is a $k$-tensor whose coordinates (in the canonical basis) are written $\{\partial^k_{i_1,\dots,i_k} \phi(x),\ i_1, \dots, i_k \in [d]\}$. We will write $\nabla = \nabla^1$ and $\partial = \partial^1$. We let

$$|\nabla^k \phi|_{\ell^p,\infty} = \sup_{x\in\mathbb{R}^d} |\nabla^k \phi(x)|_{\ell^p} = \sup_{x\in\mathbb{R}^d} \left( \sum_{i_1,\dots,i_k \le d} |\partial^k_{i_1,\dots,i_k} \phi(x)|^p \right)^{1/p}. \tag{1.1.1.2}$$

1.1.2. **Reminders about topology and measure theory.** Let $Z$ be a measurable space with its $\sigma$-algebra $\mathfrak{Z}$. Given a measure $\nu$ and a non-negative (or $\nu$-integrable) function $f$, we write $\nu(f)$ or $\int_Z f(x)\nu(\mathrm{d}x)$ to designate the Lebesgue integral of $f$ with respect to $\nu$. For $p$ and $q \in \mathbb{N}_{>0}$, $p \geq q$, if $\nu$ is a measure on the product space $Z^q$ and $f \colon Z^p \to \mathbb{R}_+$, we write

(1.1.2.1)

$$\int f(x)\nu(\mathrm{d}x_{[L]}) = \int f(x_1, \dots, x_p)\nu\left(\prod_{j\in J} \mathrm{d}x_j\right), \quad \text{for } x = (x_1, \dots, x_p)\, x_i \in Z,\ \forall i \in [p] \setminus J.$$

In other words, this means that we integrate only with respect to coordinates in $L$. For $p \in [1, \infty]$, the space of (equivalence classes of $\nu$-almost everywhere equal) measurable functions $f$ such that $\nu(|f|^p) < \infty$ (or $esssup(|f|) < \infty$ when $p = \infty$) is denoted by $\mathscr{L}^p(Z, \mathfrak{Z}, \nu)$ and is equipped with its standard Banach structure. We will commonly abbreviate this notation to $\mathscr{L}^p(Z)$ or $\mathscr{L}^p(\mathfrak{Z})$ or $\mathscr{L}^p(\nu)$. We will use the shorthand notation $\mathscr{L}^0(\mathfrak{Z})$ for the space of $\mathfrak{Z}$-measurable functions. We will write $F = (F_1, \ldots, F_d) \in \mathscr{L}^2(\Omega)$ rather than the shorter but cumbersome notation $F \in \left(\mathscr{L}^2(\Omega)\right)^d$. We will also use this notation for the other functional spaces we will introduce later. We will always made clear the size of the vector by writing explicitly its components in a way that no confusion is possible. We also write $\mathfrak{Z}_\nu$ for the collection of measurable sets $A$ such that $\nu(A) < \infty$. The measure $\nu$ is *σ-finite* if $Z$ can be written as a countable union of elements of $\mathfrak{Z}_\nu$. The space of probability measures on $Z$ is denoted by $\mathscr{P}(Z)$, the space of finite non-negative measures is denoted by $\mathscr{M}_b(Z)$, the space of non-negative measures is denoted by $\mathscr{M}_+(Z)$ and the space of signed measures is denoted by $\mathscr{M}(Z)$. We also consider $\mathscr{M}_\mathbb{N}(Z)$, the set of measures $\xi$ such that $\xi(B) \in \mathbb{N}$ for all $B \in \mathfrak{Z}$, and the set $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$ of countable sums of elements of $\mathscr{M}_\mathbb{N}(Z)$. An element $\nu \in \mathscr{M}_{\bar{\mathbb{N}}}(Z)$ is called a *point measure* (over $Z$), and $\nu$ is called a *proper point measure* if there exists $n \in \mathbb{N} \cup \{\infty\}$ and $(x_1, \ldots, x_n) \in Z^n$ such that

$$\nu = \sum_{k=1}^{n} \delta_{x_k}. \tag{1.1.2.2}$$

For $n = 0$, the previous sum has to be understood as the $0$ measure. Note that without any topological assumption on $Z$, it is possible to construct point measures that are not proper. A signed measure $\nu$ is called a *compound point measure* if there exists $n \in \mathbb{N} \cup \{\infty\}$, $(x_1, \ldots, x_n) \in Z^n$ and $(y_1, \ldots, y_n) \in \mathbb{R}^n$ such that

$$\nu = \sum_{k=1}^{n} y_k \delta_{x_k}. \tag{1.1.2.3}$$

All the measure spaces $\mathscr{M}(Z)$, $\mathscr{M}_+(Z)$, $\mathscr{M}_b(Z)$, $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$ and $\mathscr{M}_\mathbb{N}(Z)$ are equipped with the $\sigma$-algebra generated by the projection maps

$$\nu \mapsto \nu(B), \quad B \in \mathfrak{Z}. \tag{1.1.2.4}$$

A *random measure* is a random object taking value in one of those spaces. A *random point measure* is usually referred to as a *point process* and it is *proper* if it is almost surely proper. Given a random measure $\mu$, its expectation is a (non-random) measure $\nu$ called the *intensity measure*, verifying

$$\mathbb{E}\mu(A) = \nu(A), \quad \text{for all } A \in \mathfrak{Z}. \tag{1.1.2.5}$$

For two measures $\nu_1$ and $\nu_2$ and $q \in \mathbb{N}$, we write $\nu_1 \otimes \nu_2$ for the *tensor product* of $\nu_1$ and $\nu_2$, and we simply write $\nu^q$ for the tensor product of $\nu$ iterated $q$ times.

1.1.3. **Probabilistic approximations in a nutshell.** Assume $Z$ is a topological space with its collection of open sets $\tau$. Unless otherwise specified, we will always assume that the topological space $Z$ is made measurable by equipping it with its collection of Borel sets $\mathfrak{Z}$, that is, the $\sigma$-algebra generated by $\tau$. The space of real-valued bounded continuous functions on $Z$ is denoted by $\mathscr{C}_b(Z)$. We say that the space $Z$ is *Polish* if it is separable, that is it contains a dense sequence, and completely metrizable, that is: there exists a distance $d$ on $Z$ that generates the same topology as $\tau$, and such that $(Z, d)$ is complete. Assume $Z$ is a Polish space. The spaces $\mathscr{C}_b(Z)$ and $\mathscr{M}_b(Z)$ are in separating duality and the weak-$*$ topology associated with this duality is called the *narrow* topology. In other words, it is the sequential topology such that $\mu_n \xrightarrow[n\to\infty]{narrow} \mu$ if and only if $\mu_n(f) \to \mu(f)$ for all $f \in \mathscr{C}_b(Z)$. Given a sequence of $Z$-valued random variables $(F_n)$ we say that the sequence *converges in law* to a random variable $F$; and we write $F_n \xrightarrow[n\to\infty]{law} F$, if $law(F_n) \xrightarrow[n\to\infty]{narrow} law(F)$. The space $\mathscr{M}_b(Z)$ with the narrow topology is Polish [24, Chapter 5]. Consequently, the space of laws of $Z$-valued random variables with the topology of the convergence in law is Polish.

We let $d$ be a distance that completely metrizes the topology of $Z$. We now introduce different notions of distance on the space $\mathscr{P}(Z)$, that is regarded as the set of all laws of $Z$-valued random variables. We say that a function $\phi\colon Z \to \mathbb{R}$ is *Lipschitz* if

$$Lip(\phi) := \sup_{x,y} \frac{|\phi(x) - \phi(y)|}{d(x, y)} < \infty. \tag{1.1.3.1}$$

The space of Lipschitz functions is denoted by $Lip(Z)$. Note that, if $Z = \mathbb{R}^d$ with the $\ell^2$-topology and $\phi \in \mathscr{C}^1(\mathbb{R}^d) \cap Lip(\mathbb{R}^d)$, we have that $|\nabla\phi|_{op} = Lip(\phi)$. In particular, by comparison of norms we have that $|\nabla\phi|_{\ell^2,\infty} \leq \sqrt{d}Lip(\phi)$. The space of bounded Lipschitz functions on the metric space $Z$ is denoted by $\mathscr{W}^{1,\infty}(Z)$ and it is a Banach space for the norm

$$|\phi|_{\mathscr{W}^{1,\infty}(Z)} = |\phi|_\infty + Lip(\phi), \tag{1.1.3.2}$$

where $|\phi|_\infty$ is the supremum norm of $\phi$. The *Monge-Kantorovich-Rubinstein distance* between the laws of two $Z$-valued random variables $X$ and $Y$ is defined by

$$d_1(law(X), law(Y)) = \inf \mathbb{E}d(\tilde{X}, \tilde{Y}). \tag{1.1.3.3}$$

where the infimum is running over all random vectors $(\tilde{X}, \tilde{Y})$ such that $law(\tilde{X}) = law(X)$ and $law(\tilde{Y}) = law(Y)$. Due to the Kantorovich duality [55, Thm 2.1], if $X$ and $Y$ are such that $\mathbb{E}d(X, z_0) + \mathbb{E}d(Y, z_0) < \infty$ for some $z_0 \in Z$, $d_1(law(X), law(Y))$ can be rewritten as

$$d_1(law(X), law(Y)) = \sup\left\{\mathbb{E}\phi(X) - \mathbb{E}\phi(Y), \phi \in \mathscr{W}^{1,\infty}(\mathbb{R}^d), |\nabla\phi|_\infty \leq 1\right\}. \tag{1.1.3.4}$$

The Monge-Kantorovich-Rubinstein distance induces a topology on the space of probability measures that corresponds to the convergence in law together with the convergence of the first moment [146, Thm 6.9]. In general, we will want to compute the Monge-Kantorovich-Rubinstein between the law of a vector of real-valued square-integrable random variables and a $\mathbf{N}(0, C)$. To that extent, we will often refer to the following bound that (see [111, Thm 4.4.1]).

**Theorem 1.1.3.1.** *Let $F = (F_1, \ldots, F_d) \in \mathscr{L}^2(\Omega)$ and $C$ be a symmetric positive $d \times d$ matrix, then*

$$d_1(law(F), \mathbf{N}(0, C)) \leq \sup_{\mathscr{F}_1} |\mathbb{E}\langle C, \nabla^2\phi(F)\rangle_{\ell^2} - \mathbb{E}\langle F, \nabla\phi(F)\rangle_{\ell^2}|, \tag{1.1.3.5}$$

*where $\mathscr{F}_1$ is the collection of functions $\phi \in \mathscr{C}^2(\mathbb{R})$ such that $|\nabla^2\phi|_{op} \leq |C^{-1}|_{op}|C|_{op}^{1/2}$.*

In the non-diffusive setting, the Monge-Kantorovich-Rubinstein distance is not really practicable and we will often use the following distance:
(1.1.3.6)
$$d_2(law(X), law(Y)) = \sup\left\{\mathbb{E}\phi(X) - \mathbb{E}\phi(Y), \phi \in \mathscr{C}^2(\mathbb{R}^d), Lip(\phi) \le 1,\ |\nabla^2\phi|_{op} \le 1\right\}.$$

This distance was introduced by [127]. They observed that it induces a topology stronger than the topology of the convergence in law and they also proved [127, Lemma 2.17] that

**Theorem 1.1.3.2.** *Let $F = (F_1, \dots, F_d) \in \mathscr{L}^2(\Omega)$ and $C$ be a symmetric positive $d \times d$ matrix, then*

$$d_2(law(F), \mathbf{N}(0, C)) \le \sup_{\mathscr{F}_2} |\mathbb{E}\langle C, \nabla^2\phi(F)\rangle_{\ell^2} - \mathbb{E}\langle F, \nabla\phi(F)\rangle_{\ell^2}|, \tag{1.1.3.7}$$

*where $\mathscr{F}_2$ is the collection of functions $\phi \in \mathscr{C}^3(\mathbb{R})$ such that $|\nabla^2\phi|_{op} \le |C^{-1}|_{op}|C|_{op}^{1/2}$ and $|\nabla^3\phi|_{op} \le \frac{\sqrt{2\pi}}{4}|C^{-3/2}|_{op}|C|_{op}^{1/2}$*

The commonly used *Kolmogorov distance* between the distributions of two real-valued random variables $X$ and $Y$ is defined by

$$Kol(X, Y) = \sup_{t\in\mathbb{R}} |\mathbb{P}(X \le t) - \mathbb{P}(Y \le t)|. \tag{1.1.3.8}$$

1.1.4. **Tensor notations.** We let $\mathscr{H}$ be an Hilbert space of (equivalence classes of) functions over some space $Z$. Given $f$ and $g \in \mathscr{H}$ we write $f \otimes g$ for the *tensor product* of $f$ and $g$, that is for $z$ and $z' \in Z$, $f \otimes g(z, z') = f(z)g(z')$. We fix $q \in \mathbb{N}_{>0}$. We write $\mathscr{H}^{\otimes q}$ for the Hilbert space obtained by completion of the space of functions of the form

$$f_1 \otimes \cdots \otimes f_q, \quad f_1, \dots, f_q \in \mathscr{H}, \tag{1.1.4.1}$$

under the scalar product

$$\langle f_1 \otimes \cdots \otimes f_q, g_1 \otimes \cdots \otimes g_q\rangle_{\mathscr{H}^{\otimes q}} = \sum_{i=1}^{q} \langle f_i, g_i\rangle_{\mathscr{H}}. \tag{1.1.4.2}$$

In particular, if $\mathscr{H} = \mathscr{L}^2(Z, \mathfrak{Z}, \nu)$ for some measured space $(Z, \mathfrak{Z}, \nu)$, we find that $\mathscr{H}^{\otimes q} = \mathscr{L}^2(Z^q, \mathfrak{Z}^{\otimes q}, \nu^q)$. Given a permutation $\sigma \in \Sigma_q$ and $f \in \mathscr{H}^{\otimes q}$, we write $f \circ \sigma$ for the function $f$ whose entries are permuted by $\sigma$, namely

$$f \circ \sigma(z_1, \dots, z_q) = f(z_{\sigma(1)}, \dots, z_{\sigma(q)}), \quad z_1, \dots, z_q \in Z. \tag{1.1.4.3}$$

We say that an element $f \in \mathscr{H}^{\otimes q}$ is *symmetric* and we write $f \in \mathscr{H}^{\odot q}$ if it invariant under the permutations, that is $f \circ \sigma = f$ for all $\sigma \in \Sigma_q$. Similarly $f \in \mathscr{H}^{\otimes q}$ is said to be *anti-symmetric* and we write $\mathscr{H}^{\oslash q}$ if $f \circ \sigma = |\sigma| f$ where $|\sigma| \in \{-1, 1\}$ is the signature of the permutation. We have the orthogonal decomposition

$$\mathscr{H}^{\otimes q} = \mathscr{H}^{\odot q} \oplus \mathscr{H}^{\oslash q}. \tag{1.1.4.4}$$

Given $f \in \mathscr{H}^{\otimes q}$, we write $f^{\odot}$ for its symmetrization, that is its projection on $\mathscr{H}^{\odot q}$ and $f^{\oslash}$ for its anti-symmetrization, that is its projection on $\mathscr{H}^{\oslash q}$. We also write $f \odot g$ for $(f \otimes g)^{\odot}$ and similarly with $\oslash$ and $f^{\otimes q}$ for $f \otimes \cdots \otimes f$ repeated $q$ times and similarly for $f^{\odot q}$ and $f^{\oslash q}$ We also write $\mathscr{H}^{\otimes 0} = \mathscr{H}^{\odot 0} = \mathscr{H}^{\oslash 0} = \mathbb{R}$. For $p \in \mathbb{N}$, we write $\mathscr{L}^p(\Omega) \otimes \mathscr{H}$ for the Banach space (Hilbert space for $p = 2$) of $\mathscr{H}$-valued random variables $u$ such that $|u|_{\mathscr{H}} \in \mathscr{L}^p(\Omega)$.

## 1.2. What is this thesis about?

This thesis discusses the development and the intertwining of two fields of mathematics: functional inequalities and limit theorems, with a particular focus on (but not limited to) the setting of configuration spaces.

1.2.1. **Functional inequalities.** Functional inequalities can be seen as an analytic and synthetic manifestation of structural properties in mathematics. As such, they are prevalent in many areas of current research. For instance, the *logarithmic Sobolev inequality* of L. Gross (1975) [61] compares for a function $p$, say a smooth probability density on $\mathbb{R}$, its *entropy* $\mathcal{H}(p|\gamma)$ relatively to the normal law $\gamma = \mathbf{N}(0,1)$ defined by

$$\mathcal{H}(p|\gamma) = \gamma(p \log p), \tag{1.2.1.1}$$

with its *information* $\mathcal{I}(p|\gamma)$ (again relatively to the normal law $\gamma$) defined by

$$\mathcal{I}(p|\gamma) = \gamma\left(\frac{(p')^2}{p}\right) = \int \rho(u)^2 p(u)\gamma(\mathrm{d}u), \tag{1.2.1.2}$$

where $\rho(u) = \frac{\mathrm{d}}{\mathrm{d}u}\log p(u)$ is the *score function* of $p$. Namely, the logarithmic Sobolev inequality states that, for all such densities $p$,

$$\mathcal{H}(p|\gamma) \leq \frac{1}{2}\mathcal{I}(p|\gamma). \tag{1.2.1.3}$$

In information theory, this inequality is a consequence of an inequality of A. J. Stam (1959) [140] quantifying an uncertainty principle for the information (relatively to the Lebesgue measure though) and an exponential functional of the entropy (also with respect to the Lebesgue measure). In the theory of dynamical systems, the logarithmic Sobolev inequality is an instance of a regularizing property known as *hypercontractivity* as well as a control on the speed of convergence to the equilibrium of the Ornstein-Uhlenbeck dynamic on the line

$$\mathrm{d}X_t = \mathrm{d}W_t - X_t \mathrm{d}t, \tag{1.2.1.4}$$

where $W$ is a Brownian motion. Equivalently, the logarithmic Sobolev inequality accounts for the same properties for the solutions of the associated Fokker-Planck equation

$$\partial_t u = \partial^2_{xx} u - x\partial_x u. \tag{1.2.1.5}$$

For differential geometers, the logarithmic Sobolev inequality shows that the weighted Riemannian manifold $(\mathbb{R}, |\cdot|, \gamma)$ has the same geometry as a sphere of big dimension and in fact the logarithmic Sobolev inequality plays a pivotal role in controlling the behaviour of the Hamilton-Ricci flow in Perelman's proof of Poincaré conjecture (see G. Perelman (2002) [128] notably sections 3 & 5). The logarithmic Sobolev inequality also has consequences in terms of concentration of measure via the Herbst argument presented in the book of M. Ledoux (2001) [85, Chapter 5] that gives back the well-known fact that there exists $c > 0$ such that:

$$\gamma(A)(1 - \gamma(A + [-r, r])) \leq \mathrm{e}^{-cr^2}, \tag{1.2.1.6}$$

for all Borel sets $A \subset \mathbb{R}$, where $A + [-r, r] = \{u + t, \quad u \in A,\, t \in [-r, r]\}$.

Remark that, by a change of variables and the chain rule, the logarithmic Sobolev inequality (1.2.1.2) is equivalent to

$$\mathcal{H}(F^2) \leq 2\mathbb{E}|DF|^2, \quad F \in \mathscr{D}\mathrm{om}\, D, \tag{1.2.1.7}$$

where $\mathscr{D}\mathrm{om}\, D$ is the space of $F \in \mathscr{L}^2(\Omega)$ of the form $F = f(N)$ with $N \sim \mathbf{N}(0,1)$ and $f$ Lipschitz such that $f'(N) \in \mathscr{L}^2(\Omega)$, the operator $D$ is defined by $DF = f'(N)$ and $\mathcal{H}(F^2) = \mathbb{E}(F^2 \log F^2) - (\mathbb{E}F^2) \log(\mathbb{E}F^2)$. Hence, we can either say that the metric measured space $(\mathbb{R}, |\cdot|, \mathbf{N}(0,1))$ supports a logarithmic Sobolev inequality or that the probability space $(\Omega, \sigma(N), \mathbb{P})$ supports a logarithmic Sobolev inequality when equipped with the operator $D$.

With these perspectives, the logarithmic Sobolev inequality (as many other functional inequalities) can by extended in two directions:

(i) one is rather geometrical and use the formalism of *metric measured space*;

(ii) the other is rather probabilistic and exploits the idea of integration parts via the $\Gamma$-calculus of D. Bakry & M. Émery (1985) [12].

Remark that generalizing the notion of entropy is a standard task. Given a probability measure $\nu$ on a measured space $(E, \mathfrak{E}, \lambda)$, we define its *entropy* relatively to $\lambda$ by

$$\mathcal{H}(\nu|\lambda) = \lambda(p \log p), \tag{1.2.1.8}$$

if $\nu$ is absolutely continuous with respect to $\lambda$, with Radon-Nikodym derivative $p$; and we set $\mathcal{H}(\nu|\lambda) = \infty$ otherwise. In the setting of a metric measured space $(E, d, \lambda)$, we can extend the notion of length of the gradient by

$$|\nabla\phi|(x) = \limsup_{y \to x} \frac{|\phi(x) - \phi(y)|}{d(x,y)}. \tag{1.2.1.9}$$

We can use this generalized notion of derivative to generalize the notion of information (relative to $\lambda$) by

$$\mathcal{I}(\nu|\lambda) = \lambda\left(\frac{|\nabla p|^2}{p}\right), \tag{1.2.1.10}$$

where $p$ is the density of $\nu$ with respect to $\lambda$, and we set $I(\nu|\lambda) = \infty$ if $\nu$ is not absolutely continuous with respect to $\lambda$. We say that the metric measured space $(E, d, \lambda)$ supports a *logarithmic Sobolev inequality* (with constant $c > 0$) if for all probability measure $\nu$

$$\mathcal{H}(\nu|\lambda) \leq \frac{c}{2}\mathcal{I}(\nu|\lambda). \tag{1.2.1.11}$$

For probability spaces, the generalization is slightly more intricate and we do not explain it here in full generality. Let us simply point out some facts for the case of the Gaussian measure on the line. For every $F = f(N)$ where $N \sim \mathbf{N}(0,1)$ with $f \in \mathscr{C}^2(\mathbb{R})$ such that $f''(N) \in \mathscr{L}^2(\Omega)$, we let

$$LF = D^2F - XDF. \tag{1.2.1.12}$$

The attentive reader would have recognized in the definition of $L$ the probabilistic equivalent of the differential operator appearing on the right-hand side of the Fokker-Planck equation (1.2.1.5). The following facts are rather straightforward:

$$\mathbb{E}LF = 0; \tag{1.2.1.13}$$

$$|DF|^2 = \frac{1}{2}LF^2 - FLF. \tag{1.2.1.14}$$

Consequently, we find that

$$\mathbb{E}|DF|^2 = -\mathbb{E}FLF. \tag{1.2.1.15}$$

In view of this integration by parts formula, in order to generalize the logarithmic Sobolev inequality (1.2.1.7) to a more general probability space $(\Omega, \mathfrak{W}, \mathbb{P})$, we need either to define a derivative operator $D$ or to find a substitute for the differential operator $L$. These generalizations are given in Chapter 2 when $\mathfrak{W}$ is the $\sigma$-algebra generated by a random element typically enjoying some independence properties.

We will refer the study of a random object via an associated derivative operator $D$ as the one mentioned before as the *Malliavin calculus* and we will refer a similar study carried via the operator $L$ as the *Bakry-Emery calculus* or the $\Gamma$-*calculus*. The Malliavin calculus and the Bakry-Emery calculus intertwines beautifully and yields striking results from the point of view of functional inequalities as well as limit theorems.

1.2.2. **Limit theorems.** Limit theorems are concerned with the asymptotic behaviour of a sequence of random variables. The most well-known result in that field is the *central limit theorem* due to A. DE MOIVRE (1733) [104] (see also the recent reprint of his work [105]) for the case of Bernoulli random variables and P.-S. DE LAPLACE (1812) [77] (reedited in the two volumes [78, 79]) for the general case. This theorem says that the distribution of the empirical mean $S_n$ of $n$ centered and normalized independent random variables is well approximated by $\mathbf{N}(0, 1/n)$. This qualitative result has been extended in various directions, in particular with regard to how we measure the distance between the distribution of $S_n$ and the one of the Gaussian. For instance, the two independent results of A. C. BERRY (1941) [15] and C.-G. ESSEEN (1942) [45] show that the Kolmogorov distance between the law of $\sqrt{n}S_n$ and $\mathbf{N}(0, 1)$ is of order $n^{-1/2}$. More recently, S. ARTSTEIN, K. M. BALL, F. BARTHE & A. NAOR (2004) [7] and A. BARRON & O. JOHNSON (2004) [14] (also independently) generalized this result when the Kolmogorov distance is replaced by the *entropic distance* (that is, the difference of the entropies relatively to the Lebesgue measure). Conditions are also known in order to allow dependencies in the random variables in consideration. Such conditions are often referred under the name of *Lyapunov conditions* (see [16, Thm 27.3]). In any case, the philosophical meaning of the central limit theorem is that if $e$ is a random object with a lot of intrinsic independence and $f$ is linear, then we can more or less precisely measure how far the distribution of $f(e)$ is from $\mathbf{N}(\mathbb{E}f(e), \mathbb{V}\mathrm{ar}(f(e)))$. Note that in that case, we will only require to know the mean and variance of $f(e)$. Obtaining results similar to the central limit theorem for non-linear functionals has been a prominent problem in the modern theory of limit theorems. The main difficulty is to understand a good notion of non-linearity. While linearity is an ubiquitous phenomenon in all areas of mathematics, defining what non-linear means is quite difficult and has to be handle case by case. The simplest form we can hope for is the one of polynomials. To that extent, the striking characterization of D. NUALART & G. PECCATI (2005) [116], referred as the *fourth moment theorem*, says that if we consider a family $P = (P_n)$ of homogeneous Gaussian polynomials of degree $d$

$$P_n = \sum_{i_1 < \cdots < i_d} a_n(i_1, \ldots, i_d) X_{i_1} \ldots X_{i_d}, \tag{1.2.2.1}$$

where $(X_i)$ is a sequence of independent and normally distributed random variables, then, $P_n \xrightarrow[n\to\infty]{law} \mathbf{N}(0,1)$ if and only if $\mathbb{E}P_n^2 \to 1$ and $\mathbb{E}P_n^4 \to 3$. Remark that the fourth moment of a $\mathbf{N}(0,1)$ is 3.

In fact, this result was not only showed for homogeneous Gaussian polynomials, but also for Gaussian multiple stochastic integrals and more details will be given about them in Chapter 2. This result is quite natural as the stochastic integrals can be regarded as a probabilistic analogous of orthogonal polynomials in the theory of smooth functions. Stochastic integrals behaves particularly well and are intrinsically intertwined with the Malliavin calculus. The work of D. NUALART & S. ORTIZ-LATORRE (2008) [114] has shown that the derivation of the fourth moment theorem can be done solely with Malliavin calculus based techniques. Whence, trying to generalize such results for other random objects with some independence appears quite appealing. Thus, we could construct a unified approach in order to deal both with functional inequalities and limit theorems based on the study of the operator $D$. The Malliavin calculus is available in three cases: Gaussian processes [115, Chapter 2]; Poisson point processes [80] and independent random variables on the cube [38]. In these three cases, from the operator $D$, we can obtain functional inequalities and limit theorems. Apart from those three cases, such construction has remained quite elusive and we do not know any other form of satisfactory results concerning functional inequalities and limit theorems via Malliavin calculus.

Regarding the Bakry-Emery calculus, the breakthrough contribution of M. LEDOUX (2012) [84] showed that the techniques of D. NUALART & S. ORTIZ-LATORRE (2008) [114] can be understood in the context of the Bakry-Emery calculus. E. AZMOODEH, S. CAMPESE & G. POLY (2014) [9] used this perspective, in order to extend the fourth moment theorem for Gaussian polynomials to more general random polynomials but the theory of stochastic integrals or Malliavin calculus in their extended setting is, to my knowledge, not known.

1.2.3. **Configuration spaces.** In this dissertation, we will apply the techniques described above to derive functional inequalities and limit theorems especially on configurations spaces. A *configuration* can be thought of as a way of partitioning a countable population into (possibly) uncountable many classes. Each individual $i \in \mathbb{N}$ is sorted in one (and only one class) but one class can contain several individuals. If we denote by $Z$ the space of all possible classes, a configuration $C$ can intuitively be thought as a random set (with multiplicities) of $Z$. It is convenient to represent this random set as the random measure on $Z$ given by

$$\mu = \sum_{X \in C} m(X, C) \delta_X, \tag{1.2.3.1}$$

where we write $m(X, C)$ for the multiplicity of $X$ in $C$, and the sum over the empty set or when $|C| = 0$ is understood as the zero measure. Note that configurations generalize the concept of discrete random variables as every discrete random variable is represented by a configuration over a sole class. The formal definition of a configuration involves the notion of *point process* that is a random element of $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$. Point processes appear in many different fields of mathematics. In stochastic geometry, point processes over a metric space give random points of this space from which we can construct random geometrical objects. In matrix theory, the spectral measure (that is, a sum of Dirac masses at the eigenvalues) is a point process. In number theory, the function that counts all prime numbers smaller than a fixed integer can also be seen as a point process. Despite their pervasive nature, no general study of point processes from the point of view of the Malliavin calculus or the Bakry-Emery calculus has been proposed apart for the canonical example of the Poisson point process. Following the previously introduced intuition, if the random measure has a lot of independence then a Malliavin calculus and a Bakry-Emery calculus should be accessible and functional inequalities and limit theorems could be deduced from them. The easiest stronger form of independence for the example of point process given in (1.2.3.1) is when the $X_i$'s are independent and identically distributed and independent of $\mu(Z)$. We show that this case, known as the *mixed binomial*, indeed supports a Malliavin calculus and a Bakry-Emery calculus.

## 1.3. Lay summary of the original results

This dissertation collects various results I obtained during the last three years in the field of stochastic analysis applied to functional inequalities and limit theorems. I was particularly interested in the geometry induced by the Malliavin calculus and the use of geometric methods for functional inequalities and limit theorems. This section briefly summarizes the original results obtained and how they were obtained.

1.3.1. **Functional inequalities without diffusion.** Developed in Section 2.4.

The field of functional inequalities makes use of a large array of probabilistic, geometric and analytical methods. For instance, the Bakry-Emery theory, based on a systematic study of a semi-group from the point view of the convexity of the entropy along the semi-group itself, yielded striking results in these three fields. For details on the theory we recommend the seminal paper of D. Bakry & M. Émery (1985) [12], the lecture notes of M. Ledoux (2000) [86] and the comprehensive monograph of D. Bakry, I. Gentil & M. Ledoux (2014) [13]. A crucial assumption is that the semi-group is diffusive. For a probabilist, this means essentially that we look at a stochastic process (possibly on a infinite-dimensional space) driven by a Brownian motion. This assumption is rather restrictive and the problem of studying non-diffusive infinite-dimensional semi-groups did not receive, to our knowledge, a lot of attention. The exception is the Poisson point process that is the canonical example of a discrete probabilistic object. To that extent, let us mention the two references by G. Peccati & M. Reitzner (2016) [122] and G. Last & M. Penrose (2018) [82]. We propose to study such discrete random objects in a abstract and systematic way, under the additional assumption that they support stochastic integrals. We stress that this assumption does not seem as restrictive as it is in the diffusive setting and we provide two examples (outside the Poisson setting) where our analysis applies. Also note that G. Peccati & M. S. Taqqu (2011) [124] have carried out an analysis for random measures supporting stochastic integrals, but they were more interested in the combinatorial properties of such objects and were not concerned about functional inequalities nor limit theorems. The key idea in our analysis is to use the Malliavin gradient that is provided by the stochastic integrals (those abstract constructions are recalled in Section 2.3) to quantify the lack of diffusiveness of the semi-group. Namely, we will ask that the Malliavin derivative $D$ takes values in a space of function over some measurable space $(Z, \mathfrak{Z})$ and that $D$ is representable as follows

$$D_z f(e) = C_z(f(T_z e) - f(e)), \quad z \in Z, \tag{1.3.1.1}$$

where $f$ is a measurable function, $e$ is the random object supporting stochastic integrals, $T_z$ is an injective maps and $C_z$ is a random variable (see Section 2.4.2 for details). We show that, under that representability assumption, random variables always satisfy:

(i) a modified logarithmic Sobolev inequality;

(ii) a non-exact fourth moment theorem.

The latter says that the law of a stochastic integrals $F$ is close to the one of a centered Gaussian if and only if $\mathbb{E}F^4$ is close to $3(\mathbb{E}F^2)^2$ and that a polynomial reminder involving the Malliavin derivative of the functional is close to $0$. The modified logarithmic Sobolev inequality is just an adaptation of the celebrated logarithmic Sobolev inequality of L. GROSS (1975) [61] that bounds the entropy of a random variable. We can also compute explicitly the carré du champ of Bakry-Emery in terms of the square of the Malliavin gradient plus a randomized term. In the diffusive case, the randomized case never appears, and being able to explicitly compute and control this extra term is at the core of our argument for the non-exact fourth moment theorem. This was already observed in a Poisson setting by C. DÖBLER & G. PECCATI (2018) [40].

1.3.2. **Stochastic analysis for point processes.** Developed in Section 2.7.

Point processes can be thought of as the generalization of the notion of discrete random variables. Not only do they count a total population but the population can be partitioned into several (possibly uncountably many) classes. These random objects intersect many fields such as computational geometry, stochastic geometry, mathematical biology, renewal theory. The formalism of random measures gives a particularly nice framework to study point processes. This was already the perspective adopted by K. ITÔ (1956) [68] in his seminal paper on stochastic analysis for Poisson point processes, that form the most well-known example of point processes. Apart from Poisson point processes, the subject of stochastic analysis for generic point processes has received very little attention. Based on the abstract consideration of Section 2.4, we show that it is possible to develop a stochastic calculus à la Itô for a class of point processes much larger than Poisson point processes (for instance, it includes all Poisson point processes and all mixed binomial processes). Indeed for a functional $F = f(\mu)$ of a point process $\mu$ on a measurable space $(Z, \mathfrak{Z})$, a very natural discrete analogue of the derivation exists via the difference operator given, for every $z \in Z$, by

$$D_z^+ F = f(\mu + \delta_z) - f(\mu), \quad \text{for all } z \in Z. \tag{1.3.2.1}$$

This form looks very similar to (1.3.1.1) and we show that, under some compatibility conditions for the underlying point process $\mu$, the abstract framework of Section 2.4

applies. In particular, we construct stochastic integrals with respect to such point processes. We show that every mixed binomial process and every Poisson point process with $\sigma$-finite intensity measure satisfies those compatibility conditions. We show that, under the compatibility condition, the fourth moment theorem with remainder turns to an exact fourth moment theorem, that is: the law of a stochastic integral $F$ is close to a Gaussian law if $\mathbb{E}F^4$ is close to $3(\mathbb{E}F^2)^2$, when the correlations of stochastic integrals satisfy some algebraic condition.

1.3.3. **Stable approximations of Poisson functionals.** Developed in Chapter 3.

For Gaussian functionals, I. NOURDIN & D. NUALART (2010) [109] observed that the quantitative method proposed above to measure the distance between the law of the functional and a Gaussian law can be applied to the estimation of the distance from the law of a conditionally Gaussian random variable. They also considered stable convergence, that is a qualitative refinement of the convergence in law. We obtained similar results in a Poisson setting. The main difficulty is to control the remainder. In the diffusive Gaussian setting, thanks to the chain rule, the rest can be controlled via a iteration of the chain rule obtained by F. FAÀ DI BRUNO (1855) [46]. Due to the absence of a chain rule, we cannot derive a workable expression for the remainder but by a careful use of the Cauchy-Schwarz inequality (or the Hölder inequality), we can bound this remainder by geometrically meaningful quantities. In particular, the quartic remainder that appears in the non-exact fourth moment theorem of Section 2.4 will appear again. Using this result, I was able to study finely the asymptotic behaviour of models arising from the theory of stochastic processes (in particular, Volterra processes with respect to an independently scattered random measure). In this case, I can also obtain a stable version of the fourth moment theorem on the Poisson space of C. DÖBLER & G. PECCATI (2018) [40].

1.3.4. **Transport inequalities for random point measures.** *Based on a ongoing work with* N. GOZLAN & G. PECCATI. Developed in Chapter 4. Transport inequalities form a cornerstone in functional inequalities. They generally compare a cost of displacing a probability distribution $\nu_1$ to a distribution $\nu_2$ to a information theoretical quantity (such as the relative entropy) that measures how far $\nu_1$ is to $\nu_2$ in terms of fluctuations. The interaction between the geometric displacement and the relative information is particularly well adapted for obtaining concentration of measure results, as noticed by K. MARTON (1986) [93], and motivates the study of such transport-entropy inequalities. In a diffusive setting, following the work of S. G. BOBKOV, I. GENTIL & M. LEDOUX (2001) [20], such inequalities can be derived from the logarithmic Sobolev inequality (2.2.2.22). For discrete models, such as point processes, there is no general theory. M. REITZNER (2013) [132] showed that binomial process and Poisson point processes with finite intensity measure satisfy concentration of measure with respect to Talagrand convex distance. This result encouraged us to conduct further investigations about transport-entropy inequalities on the Poisson space. For the time being, we can prove new transport-entropy inequalities for mixed binomial processes with respect to various costs. From them, we can recover the result of [132] as well as one of the results of M. ERBAR & M. HUESMANN (2015) [44], that is a Talagrand inequality on the configurations spaces.

1.3.5. **Multiple sets exponential concentration, higher order eigenvalues and multi--marginal optimal transport.** *Based on the paper N. Gozlan & R. Herry [53], to appear in Potential analysis, and an ongoing work with* N. GOZLAN & P.-M. SAMSON*.* Developed in Chapter 5.

On a compact Riemannian manifold, it is well-known [85, Theorem 3.1] that the first eigenvalue of the Laplace-Beltrami operator gives exponential concentration of measure. This means that the volume of a set grows at an exponential rate given by the square root of the eigenvalue as the set enlarges. More precisely, denoting by $\mathrm{vol}$ the Riemannian volume and $\lambda_1$ the first eigenvalue, we have that, for a Borel set $A$ with $\mathrm{vol}(A) \geq 1/2$:

$$\mathrm{vol}(A_r) \geq 1 - \mathrm{e}^{-\frac{r}{3}\sqrt{\lambda_1}}, \tag{1.3.5.1}$$

where $A_r$ is the set of points $x$ of the manifold such that there exists $y \in A$ such that $x$ and $y$ are at distance less than $r$. Such a phenomenon known as *concentration of measure* has been extensively used in many fields, such as geometry of metric measured spaces (see the book of M. GROMOV (2007) [60, Section 3½]) and probability theory (see the monograph of M. TALAGRAND (1995) [145]), and initiated an independent field of research (see M. LEDOUX (2001) [85]). Concentration inequalities are connected to many other functional inequalities such as the Poincaré inequality (S. BOBKOV & M. LEDOUX (1997) [19]), the logarithmic Sobolev inequality (S. G. BOBKOV & F. GÖTZE (1999) [17]), the Talagrand inequality (M. TALAGRAND (1995) [145]) or to the criteria of Bakry-Emery (D. BAKRY, I. GENTIL & M. LEDOUX (2014) [13, Section 4.6]) or Lott-Sturm-Villani (F. OTTO & C. VILLANI (2000) [118]) and thus, concentration inequalities often serve as a guideline in formulating new inequalities as in the recent work of N. GOZLAN, C. ROBERTO, P.-M. SAMSON & P. TETALI (2017) [58] or in obtaining equivalent formulation of them, as E. MILMAN (2009) [102], who obtained equivalence between concentration of measure and isoperimetry under a curvature assumption.

Together with N. GOZLAN, we introduced in [53] a notion of improved concentration of measure that accounts for the parallel enlargement of $k$ distinct sets and showed that $\lambda_k$, the $k$-th eigenvalue of the Laplacian, gives exponential improved concentration. Namely, given Borel sets $A_1, \ldots, A_k$ satisfying the geometrical conditions that $\mu(A_i) \geq 1/(k+1)$ for all $i$ and such that their enlargements do not overlap, we proved, in particular, that

$$\mathrm{vol}(A_r) \geq 1 - \frac{1}{k+1} \exp\left(-c \min\left(r^2\lambda_k, r\sqrt{\lambda_k}\right)\right), \tag{1.3.5.2}$$

where $A = \cup_i A_i$ and $c > 0$ is some universal constant. This bound generalizes a famous result obtained by M. GROMOV & V. D. A. MILMAN (1983) [59]. The method of proof works in a general framework that encompasses compact Riemannian manifolds but graphs as well: the one of metric measured spaces. Our result is reminiscent of a result of F. R. K. CHUNG, A. GRIGOR'YAN & S.-T. YAU (1996) [34]. It is an open question to know whether, as for classical concentration of measure, multiple sets concentration of measure can be obtained from certain functional inequalities. We will discuss a partial result we obtained in that direction, jointly with N. GOZLAN & P.-M. SAMSON, by using the notion of displacement convexity along the Wasserstein barycenters introduced by M. AGUEH & G. CARLIER (2011) [1] generalizing notions of curvature based on optimal transport developed independently by K.-T. STURM (2006) [141] & [142] and J. LOTT & C. VILLANI (2009) [90].

# Functional inequalities and limit theorems for probabilistic models

## 2.1. Motivations and context

This part of the thesis is about stochastic analysis and its use for deriving functional inequalities and limit theorems. Roughly speaking, stochastic analysis is a collection of probabilistic tools for analyzing some specific infinite-dimensional spaces of functionals of a probabilistic object $e$, living on a measurable space $(E, \mathfrak{E})$. This object $e$ can be, for instance:

(G) a Gaussian field on some space $\mathscr{H}$ that lives on $\mathbb{R}^{\mathscr{H}}$;

(IID) an independent and identically distributed sequence of random variables, that lives on the space of sequences;

(MC) a Markov chain, that lives on the space of sequences;

(PP) a Poisson point process or a binomial process, that live on the space of point measures;

(X) an exchangeable sequence, that also lives on the space of sequences.

Note that, in these examples, the space $E$ does not always come with a natural distance or even with a natural topology. In what follows, the $\sigma$-algebra generated by $e$ is denoted by the symbol $\mathfrak{W}$. Tools of stochastic analysis are typically used in order to study infinite-dimensional spaces of the form $\mathscr{L}^p(\mathfrak{W})$ for some $p \geq 1$. The starting point of our approach towards stochastic analysis is to define three operators, that we will denote $L$, $D$ and $\delta$ in all this thesis, that play the role of the Laplace-Beltrami operator, the Riemannian gradient and the divergence (that is the adjoint of gradient) in the geometric study of compact Riemannian manifolds.

In the setting of Gaussian fields and point processes, the keystone of stochastic analysis is based on the existence of *multiple stochastic integrals* first developed by K. Itô (1951) [67] for functionals of a Gaussian random field (Section 2.8 contains an historical and bibliographical survey of the development of the theory so in the rest of this chapter we will keep the bibliographical references light). Multiple stochastic integrals can be regarded as a family of isometric linear mappings $I_q\colon \mathscr{H}^{\circ q} \to \mathscr{L}^2(\mathfrak{W})$ $(q \in \mathbb{N})$, where $\mathscr{H}^{\circ q}$ is a Hilbert space that will be defined in Section 2.3. The stochastic integrals give an orthogonal decomposition, that is, denoting by $\mathscr{C}_q$ the range of $I_q$, we have the Hilbert space decomposition $\mathscr{L}^2(\mathfrak{W}) = \oplus_{q\in\mathbb{N}}\mathscr{C}_q$. This means that such mappings give an orthogonal decomposition (rather referred, in this dissertation, as a *chaotic decomposition* by analogy with the Gaussian case) of $\mathscr{L}^2(\mathfrak{W})$ and a representation of every $F \in \mathscr{L}^2(\mathfrak{W})$ in terms of a family of elements $h_q \in \mathscr{H}^{\circ q}$ $(q \in \mathbb{N})$. Hence, we will define the operator $D$ as acting on such random variables via their representation in terms of the family $\{h_q; ; q \in \mathbb{N}\}$. The resulting object, denoted by $DF$, will be a $\mathscr{H}$-valued random variable. The two other operators $\delta$ and $L$ can be constructed from $D$.

These operators thus generalize tools from Riemannian geometry to our infinite-dimensional non-geometric setting and give us a solid groundwork in order to recover inequalities that generalize well-known geometrical inequalities from the finite-dimensional case. We will also see that we can derive some new inequalities that are intrinsically infinite-dimensional. In order to conveniently mirror the Riemannian setting, it is common to assume that we work with a *diffusion*, stating essentially that $L$ acts as a second-order differential operator without constant term. In that case, we can recover, in the framework of stochastic analysis, the celebrated *logarithmic Sobolev inequality* of L. GROSS (1975) [61]. This inequality generalizes the one presented on $\mathbb{R}$ in Section 1.2 and compares the *entropy* of a functional $F$, that is $\mathcal{H}(F) = \mathbb{E}(F \log F) - (\mathbb{E}F)\log(\mathbb{E}F)$, and its *energy*, that is $\mathcal{E}(F) = \mathbb{E}|DF|^2$. In the diffusion setting, the derivation of the logarithmic Sobolev inequality relies on an interpolation argument initiated by the groundbreaking work of D. BAKRY & M. ÉMERY (1985) [12]. M. LEDOUX (2012) [84] has shown that these techniques can also be used to derive the *Stein inequality*, that bounds the Monge-Kantorovich-Rubinstein distance of the law of a (sufficiently smooth) random variable $F$ to the one of a Gaussian random variable by the variance of the *Stein kernel* $S(F)$ that is expressed in terms of $D$ and $L$. The Stein inequality gives back the celebrated fourth moment theorem of D. NUALART & G. PECCATI (2005) [116], stating that the law of Gaussian multiple stochastic integral $F$ is close to the normal law if $\mathbb{E}F^4$ is close to $3(\mathbb{E}F^2)^2$. Let us also mention that M. LEDOUX, I. NOURDIN & G. PECCATI (2015) [87] used the notion of Stein kernel, initially developed to study limit theorems in a infinite-dimensional framework, to improve the finite-dimensional logarithmic Sobolev inequality: a result that we will not investigate further in this dissertation but that illustrates the flexibility of such methods.

When working with diffusions, we can also construct from the operator $D$ an *intrinsic distance* on $\Omega$ by

$$d(\omega, \omega') = \sup\{F(\omega) - F(\omega'), \text{ such that } |DF| \leq 1\}, \quad \omega, \omega' \in \Omega. \tag{2.1.0.1}$$

Then, it can be shown (see L. AMBROSIO, N. GIGLI & G. SAVARÉ (2015) [4] for details) that the metric measured space $(\Omega, \mathfrak{W}, d, \mathbb{P})$ is similar, from the point of view of functional inequalities, to the metric measured space $(\mathbb{R}, |\cdot|, \mathbf{N}(0,1))$. In particular, let us mention the *Talagrand inequality* that compares the quadratic Wasserstein transport distance (with respect to to the intrinsic distance) between two probabilities absolutely continuous with respect to $\mathbb{P}$ and there relative entropy to $\mathbb{P}$ (that is the entropy of there density). This inequality initially proved on $\mathbb{R}$ endowed with the Gaussian law by M. TALAGRAND (1996) [144] also holds in this setting.

The Bakry-Emery theory, on which all these results are built, is based on the formalism of the $\Gamma$-calculus. For suitable functional $F$ and $G$, the expression of the carré du champ $\Gamma$ is given by

$$2\Gamma(F, G) = L(FG) - FLG - GLF. \tag{2.1.0.2}$$

When $D$ satisfies a chain rule, then $L$ is a diffusion and $\Gamma(F, G) = \langle DF, DG\rangle_{\mathscr{H}}$ (Theorem 2.3.3.3). This fact is particularly convenient: we can use interchangeably nice properties of $D$ (such as the chain rule) or of $\Gamma$ (such as the fact that $\Gamma(F, F)$ is explicitly computable when $F$ is a multiple stochastic integral).

The diffusion assumption is critical, and non-trivial counter-examples can be produced for most facts mentioned above outside the diffusive setting. Studying functional inequalities in a non-diffusive setting is particularly difficult in general as, to my knowledge, without topological assumption, there exists no analytical equation, such as the chain rule for diffusions, that accounts for not being a diffusion. Nonetheless, these non-diffusive objects are essentially discrete and we can expect that the three operators $D$, $\delta$ and $L$ are simpler to define in this case. In thus respect, let us allude to the difference operator $df(x) = f(x+1) - f(x)$ that is well- and easily-defined for all real functions $f$. This difference operator is the discrete counterpart of the classical derivative that is more difficult to define but nicer to work with due to the chain rule. In the framework of non-diffusive stochastic analysis, we propose a definition of *representability* for the operator $D$, that essentially states that a formula analogous to the one for the difference operator $d$ is available for $D$. Under the assumption of representability, we can show several results of interest that are new at this level of generality:

(i) an explicit representation of the carré du champ under minimal hypothesis (Theorem 2.4.2.4);

(ii) a modified logarithmic Sobolev inequality (Theorem 2.4.4.1);

(iii) a modified Stein inequality (Theorem 2.4.4.2);

(iv) a fourth moment theorem with quartic remainder (Theorem 2.4.4.3).

The latter states that the law of a multiple integral $F$ is close to a Gaussian if $\mathbb{E}F^4$ is close to $3(F^2)^2$ and a quartic remainder involving $DF$ is close to $0$. Under an additional assumption we can turn this non-exact fourth moment theorem to an exact fourth moment theorem (Theorem 2.4.4.6), that is not involving the quartic remainder. Let us note that for those models the relation $\Gamma(F, F) = |DF|^2$ never holds.

The framework of representable structures is very flexible and provides an unified scheme in order to recover functional inequalities and quantitative limit theorems that were partially known for different classical probabilistic models. We develop a theory of stochastic analysis for functionals of point processes. We show that, under some regularity and compatibility assumptions, our stochastic analysis for point processes enters the framework of representable structure (Theorem 2.7.2.5). Two examples of interest are covered by this framework: mixed binomial processes and Poisson point processes. For Poisson point processes, the exact fourth moment theorem applies, while for binomial processes of fixed size there is a remainder. In the general case, for mixed binomial process, whether or not the fourth moment theorem with remainder becomes exact depends only on the law of the size of the binomial process under study. We can also study the hypercube with a non-symmetric probability (Theorem 2.6.1). Concerning existing results in the literature we recover:

(i) The fourth moment theorem with remainder on Rademacher chaoses, that was known from the work of C. Döbler & K. Krokowski (2017) [38] (Theorem 2.6.1).

(ii) The modified logarithmic Sobolev inequality for Poisson point processes of L. Wu (2000) [152] and the exact fourth moment theorem for Poisson functionals of C. Döbler & G. Peccati (2018) [40] (Theorem 2.7.3.3).

Regarding new results, we obtain:

(i) A modified logarithmic Sobolev inequality, a Stein inequality and a fourth moment theorem with remainder for functionals of a mixed binomial process (Theorem 2.7.3.1).

(ii) A modified logarithmic Sobolev inequality on the hypercube (Theorem 2.6.2).

The strategy of proof for all the previous theorems is to construct the three operators $D$, $\delta$ and $L$ associated with these random objects and show that they satisfy our assumptions and we are confident that we will be able to study other examples with our method in the future. Also, from Theorem 2.4.2.4, we obtain a representation of the carré du champ for all of those models with the sole assumption that the carré du champ exists. This was not known in any of those models.

The rest of the chapter is divided as follows. In Sections 2.2 and 2.3, we introduce the abstract framework of Itô structures and we show that, in a diffusion setting, the logarithmic Sobolev inequality (Theorem 2.2.2.1), the Stein inequality (Theorem 2.2.3.1) and the exact fourth moment theorem (Theorems 2.2.3.2 and 2.3.3.5) can be recovered. We do not prove the Talagrand inequality, as defining properly the intrinsic distance is rather technical and of no interest for the rest of this dissertation. All of the results of these two sections are well-known but, for completeness, we give proofs that are sufficiently detailed in order to follow the reasoning. In Section 2.4, we develop a non-diffusive framework in order to obtain modified versions of the inequalities presented in the diffusive setting. The essential notion is the one of Malliavin derivative representable by a transitive action as defined in Section 2.4.2. With this notion a modified logarithmic Sobolev inequality (Theorem 2.4.4.1), a modified Stein inequality (Theorem 2.4.4.2), a fourth moment theorem with quartic remainder (Theorem 2.4.4.3), and an exact fourth moment theorem under an additional assumption (Theorem 2.4.4.6) are obtained. In Section 2.5, we construct the first (for the order of this thesis as well as for the chronological order) example of diffusive Itô structure, involving functionals of a Gaussian field. This structure can be thought of as the infinite tensorization of the space $\mathbb{R}$ with the Gaussian measure so it is of no surprise that we recover the celebrated logarithmic Sobolev inequality. Section 2.6 gives a toy model of non-diffusive Itô structure by studying unfair coin tosses on the infinite-dimensional cube (also referred as Rademacher space by some authors). Due to the simplicity of the model, all proofs are straightforward. We purposely chose the point of view of random measures (that might appear unnecessarily heavy) in order to better anticipate the subsequent Section 2.7 on point processes. Finally, in Section 2.7, we construct a stochastic analysis for functionals of point processes. This construction is applied to the two examples of Poisson point processes and mixed binomial processes in Section 2.7.3. Most of the results of Sections 2.4 and 2.7 are, to my knowledge, new. For non-original results, along the text, we usually cite either the seminal paper or a comprehensive reference where this result can be found (sometimes in a slightly less general framework). The bibliographical discussion of Section 2.8 closes this chapter.

## 2.2. Functional inequalities in a diffusive setting

**Outline.** We introduce the notion of chaotic decomposition for the space $\mathscr{L}^2(\mathfrak{W})$, where $\mathfrak{W}$ is a $\sigma$-algebra that we will regard as the $\sigma$-algebra generated by a random object $e$. This decomposition allows us to define the *Ornstein-Uhlenbeck semi-group* $P$ that is a *Markov semi-group*, its *generator* $L$, the associated *Dirichlet energy* $\mathcal{E}$, and the *carré du champ* $\Gamma$. In this section, we introduce and we mostly work under a diffusion assumption that essentially states that the generator is a second-order differential operator. We discuss two inequalities presented in the introduction: the logarithmic Sobolev inequality and the Stein inequality. These inequalities are formulated only in terms of $L$ and $\Gamma$ and we will see in Section 2.3 that, for diffusions, they are equivalent to inequalities involving $D$. For the logarithmic Sobolev inequality Theorem 2.2.2.1, we need to assume that the carré du champ and the semi-group interact in a particular way. This is known as the Bakry-Emery criterion and we will see in Section 2.3 that this criterion is always satisfied for our diffusive probabilistic models. On the other hand, the Stein inequality Theorem 2.2.3.1 is true in a diffusive setting without any further assumptions. Under an additional assumption on the chaoses, we recover in Theorem 2.2.3.2 the fourth moment theorem that states that the law of a functional $F$ living in a fixed chaos is close (for the Monge-Kantorovich-Rubinstein distance) to a Gaussian law if $\mathbb{E}F^4$ is close to $3(\mathbb{E}F^2)^2$. All these results are known (see [12, 84, 9, 30]).

2.2.1. **Chaotic decomposition and spectral theory.** Let $(\mathscr{C}_q)_{q\in\mathbb{N}}$ be a family of orthogonal sub-Hilbert spaces of $\mathscr{L}^2(\mathfrak{W})$ such that $\mathscr{C}_0$ is the linear space of constant functions. We say that this family is a *chaotic decomposition* if

$$\mathscr{L}^2(\mathfrak{W}) = \bigoplus_{q\in\mathbb{N}} \mathscr{C}_q. \tag{2.2.1.1}$$

**Remark 1.** The definition of a chaotic decomposition is simply the one of an orthogonal decomposition. Keeping in mind the probabilistic interpretation of $\mathscr{L}^2(\mathfrak{W})$ and anticipating Section 2.3, we choose to call it a chaotic decomposition at this level of generality in order to have a unique denomination for such decomposition in the entire document.

The space $\mathscr{C}_q$ is referred as the *Wiener chaos of order* $q$ or simply *chaos* and an element $F \in \mathscr{L}^2(\mathfrak{W})$ is said to have *finite chaotic decomposition* if it lives in a finite sum of $\mathscr{C}_q$. The linear space of such random variables is denoted by $\mathscr{C}$. The chaotic decomposition assumption means that $\mathscr{C}$ is dense in $\mathscr{L}^2(\mathfrak{W})$. In a more explicit way, this means that, if for $F \in \mathscr{L}^2(\mathfrak{W})$ we write, for all $q \in \mathbb{N}$, $J_qF$ for the projection of $F$ onto $\mathscr{C}_q$, then $J_qF$ and $J_{q'}F$ are uncorrelated for $q \neq q'$ and we have the formula

$$F = \sum_{q\in\mathbb{N}} J_qF = \mathbb{E}F + \sum_{q\in\mathbb{N}_{>0}} J_qF, \tag{2.2.1.2}$$

where the sum is in a $\mathscr{L}^2(\mathfrak{W})$-sense.

We define the family of linear operators $P = \{P_t,\, t \geq 0\}$ by:

$$P_tF = \sum_{q\in\mathbb{N}} \mathrm{e}^{-qt}\, J_qF, \quad t \geq 0. \tag{2.2.1.3}$$

We recall (see, for instance [13, Section 1.2.1]) that a *Markov semi-group* over $\mathscr{L}^2(\mathfrak{W})$ is a family $Q = \{Q_t,\, t \geq 0\}$ such that:

(i) For every $t \geq 0$, $Q_t$ is a bounded linear operator on $\mathscr{L}^2(\mathfrak{W})$.

(ii) The operator $P_0$ is the identity of $\mathscr{L}^2(\mathfrak{W})$.

(iii) For every $t \geq 0$, $P_t 1 = 1$.

(iv) For every $t \geq 0$ and non-negative $F \in \mathscr{L}^2(\mathfrak{W})$, $P_t F \geq 0$.

(v) For every $s$ and $t \geq 0$, $P_t P_s = P_{t+s}$.

We recall that Markov semi-groups satisfy a *Jensen inequality*. This fact is very well-known for integration with respect to measures (see, for instance [47, Lemma 1 page 76]) but we could not find any reference to a proof of this inequality at our level of generality. Despite the proof is the same as for measures, we give a proof for completeness.

**Lemma 2.2.1.1** (Jensen inequality)**.** *Let $F = (F_1, \dots, F_d) \in \mathscr{L}^2(\mathfrak{W})$. Let $\psi \colon \mathbb{R}^d \to \mathbb{R}$ be convex such that $\psi(F) \in \mathscr{L}^2(\mathfrak{W})$. Then,*

$$\psi(P_t F) \leq P_t \psi(F), \quad \text{for all } t \geq 0. \tag{2.2.1.4}$$

*Proof.* Let $t \geq 0$ and let $x = P_t F \in \mathbb{R}^d$. Without loss of generality, we can assume that the domain of $\psi$ is $\mathbb{R}^d$. Since $\psi$ is convex, there exists $y \in \mathbb{R}^d$ such that

$$\langle y, x' - x \rangle_{\ell^2} + \psi(x) \leq \psi(x'), \quad \text{for all } x' \in \mathbb{R}^d. \tag{2.2.1.5}$$

In technical terms, $y$ is a sub-gradient of $\psi$ at $x$. Hence, since $P$ preserves the positivity, is linear and has total mass $1$

$$P_t \psi(F) \geq \langle y, P_t F - x \rangle_{\ell^2} + \psi(x) = \psi(P_t F), \quad t \geq 0. \tag{2.2.1.6}$$

This concludes the proof. □

From the property of the exponential function, it is clear that the family $P$ defined in (2.2.1.3) forms a Markov semi-group on $\mathscr{L}^2(\mathfrak{W})$ whose generator in $\mathscr{L}^2(\mathfrak{W})$ is the unbounded operator $L \colon \mathscr{L}^2(\mathfrak{W}) \to \mathscr{L}^2(\mathfrak{W})$ given by

$$LF = \lim_{t \to 0^+} \frac{1}{t}(P_t - 1)F = -\sum_{q \in \mathbb{N}} q J_q F. \tag{2.2.1.7}$$

The domain of $L$ is

$$\mathscr{D}\mathrm{om}\, L = \left\{ F \in \mathscr{L}^2(\mathfrak{W}), \text{ such that } \sum_{q \in \mathbb{N}} q^2 \mathbb{E}\left( (J_q F)^2 \right) < \infty \right\}. \tag{2.2.1.8}$$

We call $P$ (resp. $L$) the *Ornstein-Uhlenbeck semi-group* (resp. *Ornstein-Uhlenbeck generator*) associated with the chaos decomposition $\oplus \mathscr{C}_q$. The following theorem summarizes the properties of $L$ and $P$. Along the document, we will invoke these properties without always referring explicitly to this theorem.

**Theorem 2.2.1.2.** *The operator $L$ is a densely defined unbounded closed self-adjoint operator with domain $\mathscr{D}\mathrm{om}\, L$ given in (2.2.1.8). When equipped with the inner product*

$$\langle F, G\rangle_{\mathscr{D}\mathrm{om}\, L} = \mathbb{E}FG + \mathbb{E}LFLG, \tag{2.2.1.9}$$

*the space $\mathscr{D}\mathrm{om}\, L$ is Hilbert. The space $\mathscr{C}$ of random variables with finite chaotic decomposition is dense in $\mathscr{D}\mathrm{om}\, L$. The spectrum of $L$ is only pure-point and given by $-\mathbb{N}$ and for $q \in \mathbb{N}$,*

$$\ker(L+q) = \mathscr{C}_q. \tag{2.2.1.10}$$

*The Ornstein-Uhlenbeck semi-group and its generator satisfy an* invariance *property:*

$$\mathbb{E}LF = 0, \quad \text{for } F \in \mathscr{D}\mathrm{om}\, L; \tag{2.2.1.11}$$

$$\mathbb{E}P_tF = \mathbb{E}F, \quad \text{for } t \geq 0,\ F \in \mathscr{L}^2(\mathfrak{W}). \tag{2.2.1.12}$$

*The generator commutes with the action of the semi group:*

$$P_tF \in \mathscr{D}\mathrm{om}\, L \text{ and } LP_tF = P_tLF, \quad \text{for all } t \geq 0,\ F \in \mathscr{D}\mathrm{om}\, L. \tag{2.2.1.13}$$

*Proof.* Since $\mathscr{D}\mathrm{om}\, L$ contains $\mathscr{C}$, the space $\mathscr{D}\mathrm{om}\, L$ is dense in $\mathscr{L}^2(\mathfrak{W})$. Let $(F_n)_{n\in\mathbb{N}} \subset \mathscr{D}\mathrm{om}\, L$, $F \in \mathscr{L}^2(\mathfrak{W})$ and $G \in \mathscr{L}^2(\mathfrak{W})$. Assume that the sequence $(F_n)$ converges to $F$ in $\mathscr{L}^2(\mathfrak{W})$ and that $(LF_n)$ converges to $G \in \mathscr{L}^2(\mathfrak{W})$. Then, for all $q \in \mathbb{N}$, $(J_qF_n)$ converges to $J_qF$ in $\mathscr{L}^2(\mathfrak{W})$ and so

$$LF_n \xrightarrow[n\to\infty]{\mathscr{L}^2(\mathfrak{W})} = -\sum_{q\in\mathbb{N}} qJ_qF = G. \tag{2.2.1.14}$$

Since, by assumption, $G \in \mathscr{L}^2(\mathfrak{W})$, this shows that $F \in \mathscr{D}\mathrm{om}\, L$ and $LF = G$ and so $L$ is closed. From the orthogonality of the chaotic decomposition, we find for $F \in \mathscr{L}^2(\mathfrak{W})$ and $G \in \mathscr{D}\mathrm{om}\, L$,

$$\mathbb{E}FLG = \mathbb{E}\sum_{q\in\mathbb{N}}(-qJ_qF)J_qG = \mathbb{E}GLF. \tag{2.2.1.15}$$

This shows that $L$ is symmetric and, by the Cauchy-Schwarz inequality, this shows that $\mathscr{D}\mathrm{om}\, L^* \subset \mathscr{D}\mathrm{om}\, L$ (where $L^*$ is the adjoint of $L$). Thus, $L$ is in fact self-adjoint. It is clear that $\langle\cdot,\cdot\rangle_{\mathscr{D}\mathrm{om}\, L}$ is an inner product. The space $\mathscr{D}\mathrm{om}\, L$ is Hilbert with respect to this inner product since $L$ is closed and densely defined. Let $F$ be an element of the orthogonal of $\mathscr{C}$ in $\mathscr{D}\mathrm{om}\, L$. By definition of the orthogonality

$$(1+q^2)\mathbb{E}(J_qF)^2 = 0, \quad \text{for all } q \in \mathbb{N}. \tag{2.2.1.16}$$

This shows that $F = 0$ and, thus, the orthogonal of $\mathscr{C}$ contains only $0$ and we obtain the density of $\mathscr{C}$ in $\mathscr{D}\mathrm{om}\, L$. By the fact that $\oplus\mathscr{C}_q$ is an orthogonal decomposition and that $\mathscr{C}_0$ contains the constant functions, we have that

$$J_0F = \mathbb{E}F; \tag{2.2.1.17}$$

$$\mathbb{E}J_qF = 0, \quad \text{for all } q \in \mathbb{N}_{>0}. \tag{2.2.1.18}$$

This yields the announced invariance property. The commutation property is a consequence of the trivial identity

$$q\,\mathrm{e}^{-tq} = \mathrm{e}^{-tq}\,q, \quad \text{for all } q \in \mathbb{N},\ t \geq 0. \tag{2.2.1.19}$$

This completes the proof. □

We give the construction of the *pseudo-inverse* that will be used later. For every function $\psi\colon -\mathbb{N} \to \mathbb{R}$, we define the (possibly) unbounded operator

$$\psi(L)F = \sum \psi(-q) J_q F, \tag{2.2.1.20}$$

with domain

$$\mathscr{D}\mathrm{om}\,\psi(L) = \left\{ F \in \mathscr{L}^2(\mathfrak{W}), \text{ such that } \sum_{q\in\mathbb{N}} \psi(-q)^2 \mu\left((J_q F)^2\right) < \infty \right\}. \tag{2.2.1.21}$$

It follows that, the operator $L^{-1}$ is defined on the class of centered $F \in \mathscr{L}^2(\mathfrak{W})$ by

$$L^{-1}F = -\sum_{q\in\mathbb{N}_{>0}} q^{-1} J_q F. \tag{2.2.1.22}$$

For $F \in \mathscr{L}^2(\mathfrak{W})$ centered, we have that $L^{-1}F \in \mathscr{D}\mathrm{om}\, L$ and $LL^{-1}F = F$. Also with this notation $P_t = \mathrm{e}^{tL}$.

Very generally, we obtain a *spectral gap* inequality.

**Theorem 2.2.1.3** (Spectral gap inequality)**.** *Assume that $\mathscr{L}^2(\mathfrak{W})$ has a chaotic decomposition.*

$$-\mathbb{E}FL^{-1}F \leq \mathbb{V}\mathrm{ar}(F), \quad \text{for all} F \in \mathscr{L}^2(\mathfrak{W}),\ \mathbb{E}F = 0; \tag{2.2.1.23}$$

$$\mathbb{V}\mathrm{ar}(F) \leq -\mathbb{E}FLF, \quad \text{for all} F \in \mathscr{D}\mathrm{om}\, L. \tag{2.2.1.24}$$

**Remark 2.** Anticipating (2.2.2.15), this inequality is equivalent to a Poincaré inequality in the sense of [13, Def 4.2.1]. Anticipating (2.3.2.18), on the Gaussian and Poisson spaces the spectral gap inequality is thus equivalent to the Poincaré inequality on these spaces (see resp. [108, Exercice 2.11.1] and [80, Cor 1]).

*Proof.* We prove only the second inequality, the first being proved in the same way. Let $F \in \mathscr{D}\mathrm{om}\, L$. By isometry, we find that

$$\mathbb{V}\mathrm{ar}(F) = \sum_{q\in\mathbb{N}_{>0}} \mathbb{E}(J_q F)^2; \tag{2.2.1.25}$$

$$-\mathbb{E}FLF = \sum_{q\in\mathbb{N}_{>0}} q\mathbb{E}(J_q F)^2. \tag{2.2.1.26}$$

Since the summation is over $q \geq 1$, this proves the claim. □

2.2.2. **Bakry-Emery condition and the logarithmic Sobolev inequality.** We now give the definition of the carré du champ operator $\Gamma$. We adopt the point of view of Dirichlet forms as it yields a carré du champ that has a linear domain. We follow N. BOULEAU & F. HIRSCH (1991) [23, Chapter I Section 2] and L. AMBROSIO, N. GIGLI & G. SAVARÉ (2015) [4, Section 2]. We define the *energy* associated with the chaotic decomposition $\oplus\mathscr{C}_q$ as the unbounded bilinear form on $\mathscr{L}^2(\mathfrak{W}) \times \mathscr{L}^2(\mathfrak{W})$ defined by

$$\mathcal{E}(F, G) = \sum_{q\in\mathbb{N}} q\mathbb{E}J_q F J_q G, \quad F, G \in \mathscr{L}^2(\mathfrak{W}). \tag{2.2.2.1}$$

The domain of $\mathcal{E}$ is given by

$$\mathscr{D}\mathrm{om}\,\mathcal{E} = \left\{ F \in \mathscr{L}^2(\mathfrak{W}), \text{ such that } \sum_{q\in\mathbb{N}} q\mathbb{E}\left((J_qF)^2\right) < \infty \right\}. \tag{2.2.2.2}$$

The space $\mathscr{D}\mathrm{om}\,\mathcal{E}$ is Hilbert for the scalar product given by

$$\langle F, G\rangle_{\mathscr{D}\mathrm{om}\,\mathcal{E}} = \mathbb{E}FG + \mathcal{E}(F,G), \quad F, G \in \mathscr{D}\mathrm{om}\,\mathcal{E}. \tag{2.2.2.3}$$

Since $\mathscr{D}\mathrm{om}\,\mathcal{E}$ contains $\mathscr{C}$, the space $\mathscr{D}\mathrm{om}\,\mathcal{E}$ is dense in $\mathscr{L}^2(\mathfrak{W})$. We write $\mathcal{E}(F) = \mathcal{E}(F,F)$. Observe that, for every $t \geq 0$, the space $\mathscr{D}\mathrm{om}\,\mathcal{E}$ is stable under the action of $P_t$ and that

$$\mathcal{E}(P_tF) \leq e^{-2t}\mathcal{E}(F), \quad \text{for all } F \in \mathscr{D}\mathrm{om}\,\mathcal{E}. \tag{2.2.2.4}$$

Note that $\mathscr{D}\mathrm{om}\,L \subsetneq \mathscr{D}\mathrm{om}\,\mathcal{E}$ and that

$$\mathcal{E}(F,G) = -\mathbb{E}FLG, \quad \text{for all } F, G \in \mathscr{D}\mathrm{om}\,L. \tag{2.2.2.5}$$

Thus, $\mathcal{E}$ can be seen as an extension of the bilinear form

$$\mathscr{L}^2(\mathfrak{W}) \times \mathscr{L}^2(\mathfrak{W}) \ni (F,G) \mapsto -\mathbb{E}FLG. \tag{2.2.2.6}$$

Since

$$\mathcal{E}(F) = \lim_{t\to 0} \frac{1}{t}\mathbb{E}F(F - P_tF), \tag{2.2.2.7}$$

and that $P$ is a Markov semi-group, [23, I Prop 3.2.1] ensures that the energy $\mathcal{E}$ is a *Dirichlet form*. In particular, in view of [23, I Prop 3.3.1, I Rmk 2.2.4], we have that $\mathscr{D}\mathrm{om}\,\mathcal{E}$ is stable by composition with Lipschitz functions. That is $F \in \mathscr{D}\mathrm{om}\,\mathcal{E}$ and $\phi$ is Lipschitz then $\phi(F) \in \mathscr{D}\mathrm{om}\,\mathcal{E}$. It is a consequence of [23, I Cor 3.3.2] that the space $\mathscr{A} = \mathscr{D}\mathrm{om}\,\mathcal{E} \cap \mathscr{L}^\infty(\mathfrak{W})$ is an algebra with respect to the pointwise multiplication of functions. In view of the growth property of Lipschitz functions, $\mathscr{A}$ is also stable by composition with Lipschitz functions and, since the elements of $\mathscr{C}^1(\mathbb{R})$ are locally Lipschitz, $\mathscr{A}$ is in particular stable by composition with $\mathscr{C}^k(\mathbb{R})$ functions for any $k \in \mathbb{N} \cup \{\infty\}$. Hence, for every $F \in \mathscr{A}$, we define the *functional carré du champ* of $F$ as the linear form $\mathbf{\Gamma}(F)$ on $\mathscr{A}$, defined by

$$\mathbf{\Gamma}(F)[\Phi] = \mathcal{E}(F, F\Phi) - \frac{1}{2}\mathcal{E}(F^2, \Phi), \quad \text{for all } \Phi \in \mathscr{A}. \tag{2.2.2.8}$$

From [23, I Prop 4.1.1],

$$0 \leq \mathbf{\Gamma}(F)[\Phi] \leq |\Phi|_{\mathscr{L}^\infty(\mathfrak{W})}\mathcal{E}(F), \quad \text{for all } F, \Phi \in \mathscr{A}. \tag{2.2.2.9}$$

That allows us to extend the linear form $\mathbf{\Gamma}(F)$ for every $F \in \mathscr{D}\mathrm{om}\,\mathcal{E}$. For $F \in \mathscr{D}\mathrm{om}\,\mathcal{E}$, we write that $F \in \mathscr{D}\mathrm{om}\,\Gamma$ if the linear form $\mathbf{\Gamma}(F)$ can be represented by a measure absolutely continuous with respect to $\mathbb{P}$ whose density is denoted by $\Gamma(F)$. In other words, $F \in \mathscr{D}\mathrm{om}\,\Gamma$ if and only if there exists a non-negative $\Gamma(F) \in \mathscr{L}^1(\mathfrak{W})$ such that

$$\mathbf{\Gamma}(F)[\Phi] = \mathbb{E}\Gamma(F)\Phi, \quad \text{for all } \Phi \in \mathscr{A}. \tag{2.2.2.10}$$

The space $\mathscr{D}\mathrm{om}\,\Gamma$ is a closed sub-linear space of $\mathscr{D}\mathrm{om}\,\mathcal{E}$. In particular, it is Hilbert for the induced topology. By polarization, we extend the definition of $\Gamma$ to a symmetric bilinear continuous mapping called *carré du champ* by

$$\Gamma(F,G)=\frac{1}{4}(\Gamma(F+G)-\Gamma(F-G))\in\mathscr{L}^1(\mathfrak{W}),\quad \text{for all } F,G\in\mathscr{D}\mathrm{om}\,\Gamma. \tag{2.2.2.11}$$

Since $L$ is invariant and $1\in\mathscr{D}\mathrm{om}\,L$, we have that

$$\mathcal{E}(F,1)=0,\quad \text{for all } F\in\mathscr{D}\mathrm{om}\,\mathcal{E}. \tag{2.2.2.12}$$

Hence, we have that

$$\mathcal{E}(F,G)=\mathbb{E}\Gamma(F,G),\quad \text{for all } F,G\in\mathscr{D}\mathrm{om}\,\Gamma. \tag{2.2.2.13}$$

Observe that if $F\in\mathscr{D}\mathrm{om}\,L$ is such that $F^2\in\mathscr{D}\mathrm{om}\,L$ then $F\in\mathscr{D}\mathrm{om}\,\Gamma$ and

$$\Gamma(F)=\frac{1}{2}LF^2-FLF. \tag{2.2.2.14}$$

Note that, for $F\in\mathscr{D}\mathrm{om}\,\Gamma\cap\mathscr{D}\mathrm{om}\,L$, we have that

$$\mathbb{E}\Gamma(F)=-\mathbb{E}FLF=\mathcal{E}(F). \tag{2.2.2.15}$$

We say that the semi-group satisfies the *Bakry-Emery condition* [4, Cor 2.3-(vi)] if:

(i) The space $\mathscr{D}\mathrm{om}\,\Gamma$ is dense in $\mathscr{L}^2(\mathfrak{W})$.

(ii) For every $t\geq 0$ and $F\in\mathscr{D}\mathrm{om}\,\Gamma$, $P_tF\in\mathscr{D}\mathrm{om}\,\Gamma$ and:

$$\Gamma(P_tF)\leq \mathrm{e}^{-2t}\,P_t\Gamma(F). \tag{2.2.2.16}$$

To obtain more precise results under the Bakry-Emery condition a classical assumption is that $L$ is a diffusion. We say that the Ornstein-Uhlenbeck generator $L$ is a *diffusion* if the associated Dirichlet form $\mathcal{E}$ is *strongly local* in the sense that

$$\mathcal{E}(F,G)=0,\quad \text{for all } F,G\in\mathscr{D}\mathrm{om}\,\mathcal{E},\ \text{such that } FG=0. \tag{2.2.2.17}$$

Remark that, since $\mathbb{P}$ is a probability measure and that $\mathcal{E}(1)=0$ this definition is indeed equivalent to the classical definition of strong locality by [23, I Cor 5.1.4]. In that case, the carré du champ $\Gamma$ satisfies a chain rule [23, I Cor 7.1.2]:

$$\Gamma(\phi(F),G)=\phi'(F)\Gamma(F,G),\quad \text{for all } F,G\in\mathscr{D}\mathrm{om}\,\Gamma,\ \phi\in Lip(\mathbb{R}). \tag{2.2.2.18}$$

Remark that, by the Rademacher theorem, the derivative of the Lipschitz function $\phi$ is defined up to a negligible set. It is part of the corollary that, on this negligible set, the two sides of the expression can be taken to be $0$. If $L$ is a diffusion and satisfies a Bakry-Emery condition then by [4, Cor 2.3], we have that $\mathscr{D}\mathrm{om}\,\Gamma=\mathscr{D}\mathrm{om}\,\mathcal{E}$ and by [23, I Cor 6.1.4] $L$ satisfies a chain rule: for every $\phi\in\mathscr{C}^2(\mathbb{R})$ and bounded first and second derivatives, for every $F\in\mathscr{D}\mathrm{om}\,L$ such that $\Gamma(F)\in\mathscr{L}^2(\mathfrak{W})$, we have that $\phi(F)\in\mathscr{D}\mathrm{om}\,L$ and

$$L\phi(F)=\phi'(F)LF+\phi''(F)\Gamma(F). \tag{2.2.2.19}$$

Remark that, in the previous expression, we have that $F\in\mathscr{D}\mathrm{om}\,L\subset\mathscr{D}\mathrm{om}\,\mathcal{E}=\mathscr{D}\mathrm{om}\,\Gamma$ so that $\Gamma(F)$ is well-defined.

In the diffusion setting, under the Bakry-Emery criterion, we obtain the stronger logarithmic Sobolev inequality (see [13, Section 5.7]). Consider the convex function $\phi(x)=x\log x$ for $x\geq 0$ (with $\phi(0)=0$). Recall that the entropy functional is defined for a non-negative random variable $F$ as $\mathcal{H}(F)=\mathbb{E}\phi(F)-\phi(\mathbb{E}F)$.

**Theorem 2.2.2.1** (Logarithmic Sobolev inequality). *Assume that $\mathscr{L}^2(\mathfrak{W})$ has a chaotic decomposition and that $L$ is diffusive and satisfies the Bakry-Emery criterion, then for all $F \in \mathscr{D}\mathrm{om}\,\mathcal{E}$,*

$$\mathcal{H}(F^2) \le 2\mathcal{E}(F) = 2\mathbb{E}\Gamma(F). \tag{2.2.2.20}$$

*Proof.* See [31, Thm 2.1]. We fix $G \in \mathscr{A}$ with $G \ge 0$, $\epsilon > 0$ and we let $G_\epsilon = G + \epsilon$. By construction, we have that $\lim_{t\to\infty} P_t G_\epsilon = J_0 G_\epsilon = \mathbb{E}G_\epsilon$. Writing $P_\infty G_\epsilon = \mathbb{E}G_\epsilon$, we have that

$$\begin{aligned}\mathcal{H}(G_\epsilon) = \mathbb{E}\phi(G_\epsilon) - \phi(\mathbb{E}G_\epsilon) &= -\mathbb{E}\int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t}\phi(P_t G_\epsilon)\mathrm{d}t \\ &= -\int_0^\infty \mathbb{E}\phi'(P_t G_\epsilon) L P_t G_\epsilon.\end{aligned} \tag{2.2.2.21}$$

Note that $\phi'(x) = 1 + \log x$ is Lipschitz on $(\epsilon, |G_\epsilon|_{\mathscr{L}^\infty(\mathfrak{W})})$. By the fact that $\mathscr{A}$ is stable under the composition with Lipschitz function and the fact that $P_t$ preserves the positivity, the random variable $\phi'(P_t G_\epsilon) \in \mathscr{A}$. Consequently, we obtain that

$$\mathcal{H}(G_\epsilon) = \int_0^\infty \mathbb{E}\Gamma(\phi'(P_t G_\epsilon), P_t G_\epsilon). \tag{2.2.2.22}$$

By the chain rule for $\Gamma$ (2.2.2.18) we have that

$$\Gamma(\phi'(P_t G_\epsilon), P_t G_\epsilon) = \phi''(P_t G_\epsilon)\Gamma(P_t G_\epsilon). \tag{2.2.2.23}$$

In the diffusion setting, by [13, Eq 5.5.1], the Bakry-Emery relation (2.2.2.16) is equivalent to

$$\sqrt{\Gamma(P_t H)} \le \mathrm{e}^{-t}\, P_t\sqrt{(\Gamma(H))}, \quad H \in \mathscr{D}\mathrm{om}\,\Gamma. \tag{2.2.2.24}$$

Consequently,

$$\phi''(P_t G_\epsilon)\Gamma(P_t G_\epsilon) \le \mathrm{e}^{-2t}\,\frac{\left(P_t\sqrt{\Gamma(G_\epsilon)}\right)^2}{P_t G_\epsilon}. \tag{2.2.2.25}$$

The function $\psi\colon (u, v) \mapsto u^2 v^{-1}$ is convex. By Jensen's inequality for $P_t$, we have that

$$\psi(P_t\sqrt{\Gamma(G_\epsilon)}, P_t G_\epsilon) \le P_t\psi(\sqrt{\Gamma(G_\epsilon)}, G_\epsilon). \tag{2.2.2.26}$$

By invariance, we arrive at

$$\mathcal{H}(G_\epsilon) \le \int_0^\infty \mathrm{e}^{-2t}\,\mathbb{E}P_t\frac{\Gamma(G_\epsilon)}{G_\epsilon} = \frac{1}{2}\mathbb{E}\frac{\Gamma(G_\epsilon)}{G_\epsilon}. \tag{2.2.2.27}$$

From the strong locality property of $\mathcal{E}$, we have that for all $H \in \mathscr{D}\mathrm{om}\,\Gamma$, $\Gamma(H) = 0$ on $\{H = 0\}$. Hence, we can let $\epsilon \to 0$ in (2.2.2.27) and replace $G_\epsilon$ by $G$. Since $\mathscr{A}$ is an algebra, $F^2 \in \mathscr{A}$ and by (2.2.2.27) with $G = F^2$, by the chain rule for $\Gamma$ (2.2.2.18), we find that

$$\mathcal{H}(F^2) \le \frac{1}{2}\mathbb{E}\frac{\Gamma(F^2)}{F^2} = 2\mathbb{E}\Gamma(F) = 2\mathcal{E}(F). \tag{2.2.2.28}$$

Thus, we proved the claim for $F \in \mathscr{A}$. The quadratic form $\mathcal{E}$ is continuous on $\mathscr{D}\mathrm{om}\,\mathcal{E}$ and $\mathcal{H}$ is continuous for the $\mathscr{L}^2(\mathfrak{W})$-topology and hence for the topology of $\mathscr{D}\mathrm{om}\,\mathcal{E}$ that is finer. The claim concerned continuous functionals on $\mathscr{D}\mathrm{om}\,\mathcal{E}$ and is proved on a dense subset of $\mathscr{D}\mathrm{om}\,\mathcal{E}$. We conclude by density. □

2.2.3. **Polynomial chaoses and the fourth moment theorem.** We now turn to functional inequalities involving the law of random variables. These theorems are interesting from the point of view of limit theorems as they give quantitative estimates. We still work under the diffusion assumption and we recall that in this case $\mathscr{D}\mathrm{om}\,\mathcal{E} = \mathscr{D}\mathrm{om}\,\Gamma$. Whenever $F = (F_1, \dots, F_{d_1}) \in \mathscr{D}\mathrm{om}\,\Gamma$ and $G = (G_1, \dots, G_{d_2}) \in \mathscr{D}\mathrm{om}\,\Gamma$, we will write $\Gamma(F, G)$ for the random symmetric matrix whose coefficient $(i, j) \in [d_1] \times [d_2]$ is given by

$$\frac{1}{2}(\Gamma(F_i, G_j) + \Gamma(F_j, G_i)). \tag{2.2.3.1}$$

We introduce the *Stein kernel* for a centered random vector $F = (F_1, \dots, F_{d_1}) \in \mathscr{L}^2(\mathfrak{W})$ and $G = (G_1, \dots, G_{d_2}) \in \mathscr{D}\mathrm{om}\,\mathcal{E}$

$$S(F, G) = -\Gamma(L^{-1}F, G). \tag{2.2.3.2}$$

Remark that since $L^{-1}F \in \mathscr{D}\mathrm{om}\,L \subset \mathscr{D}\mathrm{om}\,\mathcal{E}$, no further assumptions are needed on $F$. For such $F$ and $G$, by (2.2.2.15), we have the following integration by parts

$$\mathbb{E}F^T G = \mathbb{E}LL^{-1}F^T G = \mathbb{E}S(F, G). \tag{2.2.3.3}$$

As usual, we write $S(F) = S(F, F)$.

**Theorem 2.2.3.1** (Stein inequality). *Suppose $\mathscr{L}^2(\mathfrak{W})$ has a chaotic decomposition and that the associated Ornstein-Uhlenbeck generator $L$ is diffusive, then for all $F = (F_1, \dots, F_d) \in \mathscr{D}\mathrm{om}\,\Gamma$ with $\mathbb{E}F = 0$ such that $S(F) \in \mathscr{L}^2(\mathfrak{W})$, we have that, with $C = \mathbb{E}F^T F$*

$$d_1(law(F), \mathbf{N}(0, C)) \leq \sqrt{d}|C^{-1}|_{op}|C|_{op}^{1/2}\sqrt{\mathbb{E}|S(F) - C|_{\ell^2}^2}, \tag{2.2.3.4}$$

*where $d_1$ is the Monge-Kantorovich-Rubinstein distance on $\mathbb{R}$.*

*Proof.* The proof is adapted from the Gaussian case [111, Thm 6.1.1]. First, we assume that $F \in \mathscr{A}$. Let $\phi \in \mathscr{C}^2(\mathbb{R})$. By the integration by parts (2.2.3.3) and the chain rule for $\Gamma$ (2.2.2.18) (that are justified since $F \in \mathscr{A}$ and $\phi$ is $\mathscr{C}^2(\mathbb{R})$), we have that

$$\begin{aligned}
\mathbb{E}\langle F, \nabla\phi(F)\rangle_{\ell^2} &= \sum_{i=1}^{d} \mathbb{E}\Gamma(L^{-1}F_i, \partial_i\phi(F)) \\
&= \sum_{ij} \mathbb{E}\partial_{ij}\phi(F)\Gamma(L^{-1}F_i, F_j) \\
&= \mathbb{E}\langle\nabla^2\phi(F), S(F)\rangle_{\ell^2}.
\end{aligned} \tag{2.2.3.5}$$

By Theorem 1.1.3.1 and the Cauchy-Schwarz inequality, this proves the bound for $F \in \mathscr{A}$. To conclude in the case $F \in \mathscr{D}\mathrm{om}\,\Gamma$, we use that $\mathscr{A}$ is dense in $\mathscr{D}\mathrm{om}\,\Gamma$ and that both side of the bound are continuous with respect to the topology of $\mathscr{D}\mathrm{om}\,\Gamma$. □

We would like to bound the quantity $\mathbb{E}|S(F) - \mathbb{E}F^2|_{\ell^2}^2$ in a more explicit way. Since by the integration by parts (2.2.3.3), $\mathbb{E}S(F) = \mathbb{E}F^T F$, in the setting of the previous theorem, $\mathbb{E}|S(F) - \mathbb{E}F^2|_{\ell^2}^2$ is the covariance matrix of $S(F)$. In the case $d = 1$ and in a Gaussian setting, M. LEDOUX (2012) [84] observed that, if $F \in \mathscr{C}_q$ for a given $q$, then the expression of $\mathbb{V}\mathrm{ar}(S(F))$ simplifies. At our level of generality, this can be done under an additional assumption introduced by E. AZMOODEH, S. CAMPESE & G. POLY (2014) [9]. We say that the chaotic decomposition has *polynomial chaoses* if for all $p$ and $q \in \mathbb{N}$, $F \in \mathscr{C}_p$ and $G \in \mathscr{C}_q$ such that $FG \in \mathscr{L}^2(\mathfrak{W})$, we have that

$$FG \in \bigoplus_{r \leq p+q} \mathscr{C}_r. \tag{2.2.3.6}$$

When $F$ is a vector of random variables each living in a fixed chaos, under the polynomial chaoses assumption and the diffusion assumption the covariance of the Stein kernel simplifies as follows. The non-quantitative version of the following theorem appeared in [30]. For short given $F \in \mathscr{L}^4(\mathfrak{W})$, we write $\mathbb{M}(F) = \mathbb{E}F^4 - 3(\mathbb{E}F^2)^2$.

**Theorem 2.2.3.2** (Fourth moment theorem). *Assume $\mathscr{L}^2(\mathfrak{W})$ has a polynomial chaotic structure with diffusive Ornstein-Uhlenbeck generator $L$. Let $(p_1, \dots, p_d) \in \mathbb{N}^d$. There exists $c > 0$, such that for all $F = (F_1, \dots, F_d)$ with $F_i \in \mathscr{C}_{p_i} \cap \mathscr{L}^4(\mathfrak{W})$, we have that*

(2.2.3.7)

$$\begin{aligned}\mathbb{E}|S(F) - \mathbb{E}F^T F|^2_{HS} \leq c \sum_{i}^{d} \mathbb{M}(F_i) + c \sum_{\substack{i,j=1 \\ p_j < p_i}}^{d} \left(\mathbb{E}F_i^4\right)^{\frac{1}{2}} \mathbb{M}(F_j)^{\frac{1}{2}} \\ + c \sum_{\substack{i,j=1 \\ i \neq j \\ p_i = p_j}}^{d} \left( \mathbb{M}(F_i)^{\frac{1}{2}} \mathbb{M}(F_j)^{\frac{1}{2}} + \left[\mathbb{E}J_{2p_i} F_j^2 J_{2p_i} F_i^2 - 2(\mathbb{E}F_i F_j)^2\right]_+ \right).\end{aligned}$$

**Remark 3.** Remark that, at this level of generality, we cannot recover the celebrated quantitative fourth moment theorem that generalizes the fourth moment theorem for vectors of multiple Gaussian integrals [111, Thm 6.2.6]. However, in dimension $1$ or if the components of the vector all live in a different chaos, we obtain an exact fourth moment theorem: the law of the vector is close to the normal law each for each of its component $F$, $\mathbb{M}(F)$ is close to $0$. See Theorem 2.3.3.5 for an another statement.

*Proof.* The proof follows along the lines of the proof of [9, Thm 3.2]. Before proving the theorem let us state and prove several lemmas. The first claim is

**Lemma 2.2.3.3.** *For all $p \in \mathbb{N}$, there exists $c_p > 0$ such that,*

$$\mathbb{V}\mathrm{ar}\, S(G) \leq c_p \left(\mathbb{E}G^4 - 3\left(\mathbb{E}G^2\right)^2\right), \quad G \in \mathscr{C}_p \cap \mathscr{L}^4(\mathfrak{W}). \tag{2.2.3.8}$$

Let $G \in \mathscr{C}_p \cap \mathscr{L}^4(\mathfrak{W})$. Since $G$ is a stochastic integral, by the assumption of polynomial chaoses, polynomials in $G$ have finite chaotic decomposition and, thus, belongs to $\mathscr{D}\mathrm{om}\, L \subset \mathscr{D}\mathrm{om}\, \mathcal{E} = \mathscr{D}\mathrm{om}\, \Gamma$. This justifies the computations we carry below. Since $LG = -pG$, by the chain rule (2.2.2.18), we find that

$$\mathbb{E}G^4 = -\frac{1}{p}\mathbb{E}G^3 LG = \frac{1}{p}\mathbb{E}\Gamma(G^3, G) = \frac{3}{p}\mathbb{E}G^2\Gamma(G). \tag{2.2.3.9}$$

Observe that

$$\Gamma(G) = \left(\frac{1}{2}L + p\right) G^2 = p\mathbb{E}G^2 + \sum_{q=1}^{2p} \left(p - \frac{q}{2}\right) J_q G^2. \tag{2.2.3.10}$$

By orthogonality of the chaotic decomposition, we obtain that

$$\mathbb{E}G^2\Gamma(G) = p(\mathbb{E}G)^2 + \sum_{q=1}^{2p}\left(p-\frac{q}{2}\right)\mathbb{E}\left[\left(J_qG^2\right)^2\right]. \tag{2.2.3.11}$$

Finally, we have that

$$\mathbb{E}G^4 - 3\left(\mathbb{E}G^2\right)^2 = \frac{3}{p}\sum_{q=1}^{2p}\left(p-\frac{q}{2}\right)\mathbb{E}\left[\left(J_qG^2\right)^2\right]. \tag{2.2.3.12}$$

By the definition of $L$ and $\Gamma$, we find that

$$\mathbb{V}\mathrm{ar}(pS(G)) = \mathbb{E}\left(\Gamma(G) - p\mathbb{E}G^2\right)^2 = \sum_{q=1}^{2p}\left(p-\frac{q}{2}\right)^2\mathbb{E}\left[\left(J_qG^2\right)^2\right]. \tag{2.2.3.13}$$

Combining the two previous expressions yields

$$\mathbb{V}\mathrm{ar}(S(G)) \leq \frac{p-\frac{1}{2}}{3p}\left(\mathbb{E}G^4 - 3\left(\mathbb{E}G^2\right)^2\right). \tag{2.2.3.14}$$

This concludes the proof of Lemma 2.2.3.3.

We now prove the following:

**Lemma 2.2.3.4.** *Let $q$ and $p \in \mathbb{N}$, $F \in \mathscr{C}_p \cap \mathscr{L}^4(\mathfrak{W})$ and $G \in \mathscr{C}_q \cap \mathscr{L}^4(\mathfrak{W})$, we have that*

$$\mathbb{V}\mathrm{ar}(\Gamma(F,G)) \leq \frac{p+q-1}{4}\left(\mathbb{E}\left[F^2\left(qG^2 - \Gamma(G)\right)\right] - 2q(\mathbb{E}FG)^2\right). \tag{2.2.3.15}$$

Since $F$ and $G$ have finite chaotic decomposition, the following computations are justified by the chain rule (2.2.2.18) and the integration by parts (2.2.2.15):

$$\mathbb{E}F^2G^2 = -\frac{1}{q}\mathbb{E}F^2GLG = \frac{1}{q}\mathbb{E}\Gamma(F^2G,G) = \frac{2}{q}\mathbb{E}FG\Gamma(F,G) + \frac{1}{q}\mathbb{E}F^2\Gamma(G). \tag{2.2.3.16}$$

By definition of $\Gamma$, we have that

$$\Gamma(F,G) = q\mathbb{E}FG + \frac{1}{2}\sum_{k=1}^{p+q}(p+q-k)J_k(FG). \tag{2.2.3.17}$$

Consequently we obtain that,

$$\mathbb{E}FG\Gamma(F,G) = q(\mathbb{E}FG)^2 + \frac{1}{2}\sum_{k=1}^{p+q}(p+q-k)\mathbb{E}\left[(J_kFG)^2\right]. \tag{2.2.3.18}$$

Using the previous relations, we find that

$$\begin{aligned}
\mathbb{V}\mathrm{ar}(\Gamma(F,G)) &= \sum_{k=1}^{p+q}\left(\frac{p+q-k}{2}\right)^2\mathbb{E}\left[(J_kFG)^2\right] \\
&\leq \frac{p+q-1}{2}\left(\mathbb{E}FG\Gamma(F,G) - q(\mathbb{E}FG)^2\right) \\
&= \frac{p+q-1}{2}\left(\frac{q}{2}\mathbb{E}F^2G^2 - \frac{1}{2}\mathbb{E}F^2\Gamma(G) - q(\mathbb{E}FG)^2\right).
\end{aligned} \tag{2.2.3.19}$$

This concludes the proof of Lemma 2.2.3.4.

Now we prove the theorem. Let $C = \mathbb{E}F^T F$. Observe that

$$\mathbb{E}|S(F) - C|_{\ell^2}^2 = \frac{1}{4}\sum_{i,j=1}^{d}\left(\frac{1}{p_i} + \frac{1}{p_j}\right)\mathbb{V}\mathrm{ar}\,\Gamma(F_i, F_j). \tag{2.2.3.20}$$

We will prove the theorem by bounding each of the terms appearing in the sum. Let $i$ and $j \in [d]$ such that $p_i \neq p_j$ and we can assume that $p_i > p_j$. By orthogonality of the stochastic integrals, we have that $\mathbb{E}F_iF_j = 0$. By Lemma 2.2.3.4, we find that

$$\mathbb{V}\mathrm{ar}(\Gamma(F_i, F_j)) \leq \frac{p + q - 1}{4}\mathbb{E}F_j^2(p_iF_i^2 - \Gamma(F_i)). \tag{2.2.3.21}$$

Since

$$p_iF_i^2 - \Gamma(F_i) = \sum_{k=1}^{2p_i}\frac{k}{2}J_kF_i^2, \tag{2.2.3.22}$$

we find that

$$\begin{aligned}
\mathbb{V}\mathrm{ar}(\Gamma(F_i, F_j)) &\leq \frac{p_i + p_j - 1}{4}p_i^2\sum_{k=1}^{2p_j}\mathbb{E}\left[J_kF_j^2J_kF_i^2\right] \\
&\leq \frac{p + q - 1}{4}q^2\mathbb{E}F_j^2\sum_{k=1}^{2p_i-1}J_kF_j^2 \\
&\leq \frac{p_i + p_j - 1}{2}p_i^2\sqrt{\mathbb{E}F_i^4}\sqrt{\sum_{k=1}^{2p_i-1}\left(p_i - \frac{k}{2}\right)^2\mathbb{E}\left[\left(J_kF_j^2\right)^2\right]} \\
&= \frac{p_i + p_j - 1}{2}p_i^2\sqrt{\mathbb{E}F_i^4}\sqrt{\mathbb{V}\mathrm{ar}(p_iS(F_j))} \\
&\leq c\sqrt{\mathbb{E}F_i^4}\sqrt{\mathbb{E}F_j^4 - 3\left(\mathbb{E}F_j^2\right)^2}.
\end{aligned} \tag{2.2.3.23}$$

The last inequality is true by Lemma 2.2.3.3.

Now we take $i$ and $j \in [d]$ such that $p_i = p_j$. By Lemma 2.2.3.4, we find that

$$\mathbb{V}\mathrm{ar}(\Gamma(F_i, F_j)) = \frac{2p_j - 1}{4}(\mathbb{E}F_i^2(p_iF_j^2 - \Gamma(F_j)) - 2p_i(\mathbb{E}F_iF_j)^2). \tag{2.2.3.24}$$

Writing the chaotic decomposition, we find that

(2.2.3.25)

$$\begin{aligned}
\mathbb{V}\mathrm{ar}(\Gamma(F_i, F_j)) &= \sum_{k=1}^{2p_i-1}\frac{k}{2}\mathbb{E}\left[J_kF^2J_kG^2\right] + p_i\left(\mathbb{E}J_{2p_i}F^2J_{2p_i}G^2 - 2(\mathbb{E}FG)^2\right) \\
&\leq p_i\sqrt{\sum_{k=1}^{2p_i-1}\mathbb{E}J_kF_i^2}\sqrt{\sum_{k=1}^{2p_i-1}\mathbb{E}J_kF_j^2} + q\left[\mathbb{E}J_{2p_i}F^2J_{2p_i}G^2 - 2(\mathbb{E}FG)^2\right]_+.
\end{aligned}$$

Since

$$\sum_{k=1}^{2p_i-1}\mathbb{E}J_kF_i^2 \leq c\,\mathbb{V}\mathrm{ar}(p_i\Gamma(F_i)), \tag{2.2.3.26}$$

we conclude, by Lemma 2.2.3.3, that
(2.2.3.27)

$$\mathbb{V}\mathrm{ar}(\Gamma(F_i, F_j)) \leq c\sqrt{\mathbb{E}F_i^4 - 3(\mathbb{E}F_i^2)^2}\sqrt{\mathbb{E}F_j^4 - 3\big(\mathbb{E}F_j^2\big)^2} + c\big[\mathbb{E}J_{2p_i}F^2 J_{2p_i}G^2 - 2(\mathbb{E}FG)^2\big]_+.$$

This concludes the proof. $\square$

## 2.3. Fock space structure and Malliavin calculus

**Outline.** In the previous section, we recovered workable conditions for obtaining functional inequalities and quantitative limit theorems in a diffusive setting. We would like to improve them in three ways:

(i) we would like to check the Bakry-Emery condition (2.2.2.16);

(ii) in a diffusive setting, we would like to improve the non-exact fourth moment theorem Theorem 2.2.3.2 to recover the exact fourth moment theorem on the Gaussian space;

(iii) we would like to extend those results to the non-diffusive framework.

These three tasks can be accomplished through the introduction of a Fock space structure and Malliavin operators. As we will see, this structure arises in all the examples of probabilistic models we will study. The idea behind the Fock space formalism is to represent each chaos $\mathscr{C}_q$ ($q \in \mathbb{N}$) of the chaotic decomposition of $\mathscr{L}^2(\mathfrak{W})$ with a Hilbert space $\mathscr{H}^{\circ q}$ and a *multiple stochastic integral* mapping $I_q \colon \mathscr{H}^{\circ q} \to \mathscr{C}_q$. Provided some compatibility conditions on the family $\{\mathscr{H}^{\circ q},\, q \in \mathbb{N}\}$, we can look at $\oplus\mathscr{H}^{\circ q}$ as a graded structure of Hilbert spaces and then we define the Malliavin derivative $DF$ as the resulting object obtained when shifting the representation of $F$ in this graded structure; and we define $\delta$ as the adjoint of $D$. The operator $D$ is then an unbounded operator and for every $F \in \mathscr{D}\mathrm{om}\, D$, the object $DF$ is a $\mathscr{H} = \mathscr{H}^{\circ 1}$-valued square-integrable random variable. The two Malliavin operators $D$ and $\delta$ are linked to the Ornstein-Uhlenbeck generator: as we will see in Theorem 2.3.2.5, we have that $L = -\delta D$. We will obtain the following commutation relation between the operator $D$ and the Ornstein-Uhlenbeck semi-group, namely:

(2.3.0.1)
$$DP_tF = \mathrm{e}^{-t}\, P_tDF, \quad \text{for all } F \in \mathscr{D}\mathrm{om}\, D,\ t \geq 0.$$

In particular, the bilinear map

(2.3.0.2)
$$\Gamma_0 \colon\ \mathscr{D}\mathrm{om}\, D \times \mathscr{D}\mathrm{om}\, D \ni (F, G) \mapsto \langle DF, DG\rangle_{\mathscr{H}},$$

satisfies a relation similar to the Bakry-Emery condition (2.2.2.16). Thus, if $\Gamma = \Gamma_0$ then the operator $L$ also satisfies the Bakry-Emery relation. The equality $\Gamma = \Gamma_0$ is not true in general (we will see it when computing explicitly the carré du champ for discrete models in Section 2.4), and anticipating (2.3.2.4), one can already observe that, we have

(2.3.0.3)
$$\mathscr{D}\mathrm{om}\, D = \mathscr{D}\mathrm{om}\,\Gamma_0 = \mathscr{D}\mathrm{om}\,\mathcal{E},$$

while the equality $\mathscr{D}\mathrm{om}\,\Gamma = \mathscr{D}\mathrm{om}\,\mathcal{E}$ under a Bakry-Emery condition is, to my knowledge, known in general only for diffusions (see [4, Cor 2.3]). If $D$ is a derivation, essentially meaning that it acts as the usual derivative, then we show that $L$ is a diffusion and that the carré du champ is the square of the gradient (Theorem 2.3.3.3). Also we show that if $D$ is a derivation, we obtain an abstract product formula (Lemma 2.3.3.4) that can be used to improve the fourth moment theorem with remainder to a fourth moment theorem without remainder (Theorem 2.3.3.5). The question of using the Malliavin derivative outside of the diffusive setting will be discussed in Section 2.4.2.

2.3.1. **Abstract Fock space.** We fix a measured space $(Z, \mathfrak{Z}, \nu)$. We let $\mathscr{H} = \mathscr{L}^2(\nu)$. Recall that, for $q \in \mathbb{N}$, we denote the space of symmetric functions of $\mathscr{L}^2(\nu^q)$ by $\mathscr{H}^{\odot q}$. The space $\mathscr{H}^{\odot q}$ is endowed with the scalar product

$$\langle h, \tilde{h} \rangle_{\mathscr{H}^{\odot q}} = q! \nu^q(h\tilde{h}), \quad \text{for all } h, \tilde{h} \in \mathscr{H}^{\odot q}. \tag{2.3.1.1}$$

Let $\{\mathscr{H}^{\circ q}, q \in \mathbb{N}\}$ be a family of sets such that for all $q \in \mathbb{N}$, $\mathscr{H}^{\circ q}$ is a sub-Hilbert space of $\mathscr{H}^{\odot q}$ (we take $\mathscr{H}^{\circ 0} = \mathbb{R}$). Of course, for every $q \in \mathbb{N}_{>0}$, the space $\mathscr{H}^{\circ q}$ is equipped with the topology induced by $\mathscr{H}^{\odot q}$ but, a priori, it can be equipped with an inner product $\langle \cdot, \cdot \rangle_{\mathscr{H}^{\circ q}}$ that is equivalent but different from the one of $\mathscr{H}^{\odot q}$. See the examples below. We say that this family is a *compatible with a Fock space structure* if it is compatible with the restriction, that is, for all $q \in \mathbb{N}$,

$$h(z, \cdot) \in \mathscr{H}^{\circ q}, \quad \text{for } \nu\text{-almost every } z \in Z. \tag{2.3.1.2}$$

Note that this implies the stronger property that, for all $q$ and $p \in \mathbb{N}$ and $h \in \mathscr{H}^{\circ(p+q)}$,

$$h(z_1, \ldots, z_p, \cdot) \in \mathscr{H}^{\circ q}, \quad \text{for } \nu^p\text{-almost every } (z_1, \ldots, z_p) \in Z^p. \tag{2.3.1.3}$$

A *Fock space* is any space of the form $\mathscr{H}^{\circ} = \oplus_{q \in \mathbb{N}} \mathscr{H}^{\circ q}$, where $\{\mathscr{H}^{\circ q}, q \in \mathbb{N}\}$ is compatible with a Fock space structure.

Before investigating further the Malliavin operators, we give several examples of Fock space that we will encounter in this document. Note that, the last two definitions are new and are well-adapted to the study of point processes.

**Example 2.3.1.1** (The bosonic Fock space)**.** We present the historical example of Fock space taken from [48], in the context of quantum mechanics. Given a measured space $(Z, \mathfrak{Z}, \nu)$. The associated *bosonic Fock space* corresponds to the choice of $\mathscr{H}^{\circ q} = \mathscr{H}^{\odot q}$ (as Hilbert spaces), for all $q \in \mathbb{N}$. When $\mathscr{H} = \mathscr{L}^2(\nu)$, we will use the symbol $\mathscr{H}^{\odot}$ to designate the bosonic Fock space over $\mathscr{H}$, that is $\mathscr{H}^{\odot} = \oplus_q \mathscr{H}^{\odot q}$.

**Example 2.3.1.2** (The mixed bosonic Fock space)**.** Let $(\lambda_q)$ be a sequence of positive real numbers and $(Z, \mathfrak{Z}, \nu)$ be a measured space. The associated *mixed bosonic Fock space* corresponds to the choice of $\mathscr{H}^{\circ q} = \mathscr{H}^{\odot q}$ (as sets) and $\langle \cdot, \cdot \rangle_{\mathscr{H}^{\circ q}} = \lambda_q \langle \cdot, \cdot \rangle_{\mathscr{H}^{\odot q}}$, for all $q \in \mathbb{N}$.

**Example 2.3.1.3** (The vanishing Fock space)**.** As before, we take a measured space $(Z, \mathfrak{Z}, \nu)$ and $\mathscr{H} = \mathscr{L}^2(\nu)$. We take $\mathscr{H}^{\circ q}$ be the sub-Hilbert space (equipped with $\langle \cdot, \cdot \rangle_{\mathscr{H}^{\odot q}}$) of functions $h \in \mathscr{H}^{\odot q}$ vanishing on the diagonal. By this, mean that

$$h(x_1, \ldots, x_q) = 0, \quad \forall x = (x_1, \ldots, x_q) \in Z^q, \text{ such that } \exists i, j \in [q],\ x_i = x_j. \tag{2.3.1.4}$$

The *vanishing Fock space* associated with $\mathscr{H}$ is $\mathscr{H}^{\circ} = \oplus_q \mathscr{H}^{\circ q}$.

Let $\mathscr{H}^\circ$ be a Fock space. If the space $\mathscr{L}^2(\mathfrak{W})$ admits a chaotic decomposition $\oplus\mathscr{C}_q$ (defined in Section 2.2), we say that $\mathscr{H}^\circ$ is the *Fock space associated with the chaotic decomposition*, or that $\mathscr{L}^2(\mathfrak{W})$ *supports the Fock space* $\mathscr{H}^\circ$, if the graded structure $\oplus\mathscr{C}_q$ is isomorphic to the graded structure $\mathscr{H}^\circ$. By this, we mean that, for all $q \in \mathbb{N}$, there exists linear bijective maps $I_q \colon \mathscr{H}^{\circ q} \to \mathscr{C}_q$ such that

$$\mathbb{E} I_q(h_q) I_{q'}(\tilde{h}_{q'}) = 1_{q=q'} q! \nu^q(h_q \tilde{h}_{q'}), \quad q, q' \in \mathbb{N}_{>0},\ h_q \in \mathscr{H}^{\circ q},\ \tilde{h}_{q'} \in \mathscr{H}^{\circ q'}, \tag{2.3.1.5}$$

and we require that $I_0$ is the identity map from $\mathbb{R}$ to $\mathbb{R}$. For $q \in \mathbb{N}$, the map $I_q$ is called the *multiple stochastic integral map of order* $q$. We will say that $F$ is a *stochastic integral* if there exist $q \in \mathbb{N}$ and $h \in \mathscr{H}^{\circ q}$ such that $F = I_q(h)$. If $\mathscr{L}^2(\mathfrak{W})$ supports a Fock space and has polynomial chaoses, we say that $\mathscr{L}^2(\mathfrak{W})$ has an *Itô structure*. For $p$ and $q \in \mathbb{N}$ and $h_{p+q} \in \mathscr{H}^{\circ(p+q)}$, the symbol $I_q(h_{p+q})$ is used to designate the $\mathscr{H}^{\circ p}$-valued random variable $u$ such that

$$u_{z_1,\dots,z_p} = I_q(h(z_1, \dots, z_p, \cdot)). \tag{2.3.1.6}$$

We denote by $\mathscr{L}^2(\mathfrak{W}) \circ \mathscr{H}^{\circ p}$ the space of all such random variables $u$. Remark that this space is a priori smaller than $\mathscr{L}^2(\mathfrak{W}) \otimes \mathscr{H}^{\circ p}$.

2.3.2. **Malliavin calculus.** In the following, we fix a Fock space $\mathscr{H}^\circ = \oplus\mathscr{H}^{\circ q}$ and we assume that $\mathscr{L}^2(\mathfrak{W})$ supports the Fock space $\mathscr{H}^\circ$. From the Fock space structure every element $F \in \mathscr{L}^2(\mathfrak{W})$ can be represented by a sequence $(h_0, h_1, \dots) \in \mathscr{H}^\circ$. Namely, the representation (2.2.1.2) becomes

$$F = \sum_{q \in \mathbb{N}} I_q(h_q), \tag{2.3.2.1}$$

for a unique element $h = (h_0, h_1, \dots) \in \mathscr{H}^\circ$.

A very natural operation is to shift that sequence that is to look at the operation

$$(h_0, h_1, \dots) \mapsto (h_1, h_2, \dots). \tag{2.3.2.2}$$

This operation can be translated at the level of $\mathscr{L}^2(\mathfrak{W})$ and yields the formalism of *Malliavin gradient*. For all $p \in \mathbb{N}$, we define the unbounded operator $D^p \colon \mathscr{L}^2(\mathfrak{W}) \to \mathscr{L}^2(\mathfrak{W}) \circ \mathscr{H}^{\circ p}$ called the *Malliavin derivative of order* $p$ via

$$D^p F = \sum_{q \geq p} \frac{p!}{(q-p)!} I_{q-p}(h_q), \tag{2.3.2.3}$$

with

$$\mathscr{D}\mathrm{om}\, D^p = \left\{ F \in \mathscr{L}^2(\mathfrak{W}),\ \text{such that} \sum_{q \geq p} \left( \frac{p!}{(q-p)!} \right)^2 |h_q|^2_{\mathscr{H}^{\odot q}} < \infty \right\}. \tag{2.3.2.4}$$

We also consider the unbounded operator $\delta^p \colon \mathscr{L}^2(\mathfrak{W}) \circ \mathscr{H}^{\circ p} \to \mathscr{L}^2(\mathfrak{W})$ with

$$\mathscr{D}\mathrm{om}\, \delta^p = \left\{ u \in \mathscr{L}^2(\mathfrak{W}) \circ \mathscr{H}^p,\ \text{such that},\ \forall\, F \in \mathscr{D}\mathrm{om}\, D^p,\ |\mathbb{E}\langle u, D^p F \rangle_{\mathscr{H}^{\odot q}}| \leq c(\mathbb{E}F^2)^{1/2} \right\}, \tag{2.3.2.5}$$

and for $u \in \mathscr{D}\mathrm{om}\,\delta^p$, the value of $\delta^p u$ is characterized by the following duality relation:

$$\mathbb{E}\delta^p u F = \mathbb{E}\langle u, D^p F\rangle_{\mathscr{H}^{\odot q}}, \quad \text{for all } F \in \mathscr{D}\mathrm{om}\, D^p. \tag{2.3.2.6}$$

We write $D$ for $D^1$ and $\delta$ for $\delta^1$. With our notation, we also have that $D^0$ and $\delta^0$ both are the identity operator of $\mathscr{L}^2(\mathfrak{W})$. For $F \in \mathscr{D}\mathrm{om}\, D^p$, the quantity $D^p F$ is a random element of $\mathscr{H}^{\circ p}$ and for $z_1, \dots, z_p \in Z$, we write

$$D^p_{z_1 \dots z_p} F \in \mathscr{L}^2(\mathfrak{W}), \tag{2.3.2.7}$$

for the random element obtained by evaluating the function $D^p F$ at $(z_1, \dots, z_p)$.

The following three theorems summarize the main properties of $D^p$ and $\delta^p$.

**Theorem 2.3.2.1.** *Let $p \in \mathbb{N}$. We have that $D^{p+1} = DD^p = D^p D$ and $\delta^{p+1} = \delta\delta^p = \delta^p\delta$. The operator $D^p$ is closed and densely defined; $\delta^p$ is its adjoint and is also closed and densely defined. Moreover, $\operatorname{im} D^p$ has full range in $\mathscr{L}^2(\mathfrak{W}) \circ \mathscr{H}^{\circ p}$ and $\ker \delta^p = \{0\}$.*

**Remark 4.** Our definition of $\delta^p$ does not coincide with the classical definition of the divergence operator in a Gaussian setting (see [115, Chapter 1] or [111, Chapter 2]). Indeed, in this case the authors work with the bosonic Fock space and they define the operators $D^p$ in the same way that we do but they define $D$ as a map from $\mathscr{L}^2(\mathfrak{W})$ to $\mathscr{L}^2(\mathfrak{W}) \otimes \mathscr{H}^{\odot p}$ and consider $\delta^p$ as the adjoint of $D^p$ with this extended range. We choose to work with our definition rather than the usual one since it provides us nicer properties for the operators $D^p$ and $\delta^p$. The counterpart being that $\mathscr{D}\mathrm{om}\,\delta^p$ is much smaller. This is not a drawback in our case as we are not really interested in the properties of $\delta$. By analogy with a case that might be more familiar to the reader, every smooth vector field $u$ on a Riemannian manifold $M$ can be written uniquely as $u = \nabla F + u_0$ where $F$ is a smooth function, $\nabla$ is the Riemannian gradient and $u_0$ is a smooth vector field with vanishing divergence. Our choice would consist, in this case, to only look at the subspace $\mathscr{V}_0$ of vector fields for which $u_0 = 0$. In that case, clearly, $\nabla\colon \mathscr{C}^\infty(M) \to \mathscr{V}_0$ is surjective. If we want to study the geometry of vector fields, this restriction is impairing. However, if, as in our case, we are jut interested in properties of functions $F$ and their gradients this definition is rather convenient.

*Proof.* Observe that $D^p$ simply corresponds to iterating $p$ times the shift operation and, thus, we verify that $D^p$ is indeed $D$ iterated $p$ times. As $\mathscr{D}\mathrm{om}\, D^p$ contains $\mathscr{C}$ it is densely defined. The proof that $D^p$ is closed is the same as the proof that $L$ is closed: this consists in considering a sequence $(F_n) \subset \mathscr{D}\mathrm{om}\, D^p$ that converges to some $F \in \mathscr{L}^2(\mathfrak{W})$ and such that $(D^p F_n)$ converges to some $G \in \mathscr{L}^2(\mathfrak{W})$ and by writing explicitly the chaotic decomposition of $F_n$ and $D^p F_n$. Since $D^p$ is densely defined, it admits a closed adjoint; by definition, $\delta^p$ is such an adjoint. Since $D^p$ is closed its adjoint $\delta^p$ is densely defined. Let $u = \sum_{q\in\mathbb{N}} I_q(g_{q+p}) \in \mathscr{L}^2(\mathfrak{W}) \circ \mathscr{H}^p$. We let $h_q = (q-p)!q!^{-1} g_q$ if $q \geq p$ and $h_q = 0$ otherwise. Then $F = \sum_{q\in\mathbb{N}} I_q(h_q) \in \mathscr{D}\mathrm{om}\, D^p$ and $D^p F = u$. This shows that $\operatorname{im} D^p = \mathscr{L}^2(\mathfrak{W}) \circ \mathscr{H}^{\circ p}$. Since $D$ and $\delta$ are adjoint of each other, we have that $\ker \delta^p$ is the orthogonal of $\operatorname{im} D^p$, that is $\{0\}$. $\square$

**Theorem 2.3.2.2.** *Let $p \in \mathbb{N}$ and $F \in \mathscr{D}\mathrm{om}\, D^p$. Then for all $q \leq p$, the kernels $h_q \in \mathscr{H}^{\circ q}$ in the representation (2.3.2.1) are given by $h_q = q!^{-1}\mathbb{E}D^q F$.*

*Proof.* Immediate from the definition of $D^p$ and the fact that for $q \leq p$, $\mathscr{D}\mathrm{om}\, D^p \subset \mathscr{D}\mathrm{om}\, D^q$. □

**Theorem 2.3.2.3.** *The linear space $\mathscr{D}\mathrm{om}\, D^p$ is a Hilbert space for the inner product*

$$\langle F, G\rangle_{\mathscr{D}\mathrm{om}\, D^p} = \sum_{j=0}^{p} \mathbb{E}\langle D^j F, D^j G\rangle_{\mathscr{H}^{\odot j}}. \tag{2.3.2.8}$$

*The space $\mathscr{C}$ is dense in $\mathscr{D}\mathrm{om}\, D^p$ for every $p \in \mathbb{N}$. In particular, $q \leq p$, $\mathscr{D}\mathrm{om}\, D^p$ is dense in $\mathscr{D}\mathrm{om}\, D^q$. The linear space $\mathscr{D}\mathrm{om}\, \delta^p$ is a Hilbert space for the inner product*

$$\langle u, v\rangle_{\mathscr{D}\mathrm{om}\, \delta^p} = \mathbb{E}\delta^p u \delta^p v. \tag{2.3.2.9}$$

*Proof.* It is clear that $\langle \cdot, \cdot\rangle_{\mathscr{D}\mathrm{om}\, D^p}$ defines an inner product and $\langle \cdot, \cdot\rangle_{\mathscr{D}\mathrm{om}\, \delta^p}$ defines an inner product since $\ker \delta^p = \{0\}$. The fact that the spaces are complete for the induced topology comes from the fact that those operators are closed and densely defined. The fact that $\mathscr{C}$ is dense in every $\mathscr{D}\mathrm{om}\, D^p$ comes from the density of $\mathscr{C}$ in $\mathscr{L}^2(\mathfrak{W})$ and the fact that $D^p$ is closed. □

By duality, we obtain the following representation of stochastic integrals as iterated divergences.

**Proposition 2.3.2.4.** *We have that $\mathscr{H}^{\odot p} \subset \mathscr{D}\mathrm{om}\, \delta^p$ and $\delta^p h = I_p(h)$ for $h \in \mathscr{H}^{\odot p}$.*

*Proof.* Let $h \in \mathscr{H}^{\odot p}$. Let $G = \sum_{q \in \mathbb{N}} I_q(g_q) \in \mathscr{D}\mathrm{om}\, D^p$, then

$$\mathbb{E}\langle h, D^p G\rangle_{\mathscr{H}^{\odot p}} = p!\langle h, g_p\rangle_{\mathscr{H}^{\odot p}}. \tag{2.3.2.10}$$

So, by the Cauchy-Schwarz inequality, $h \in \mathscr{D}\mathrm{om}\, \delta^p$. By duality, we have that

$$\mathbb{E}\delta^p h G = \mathbb{E}\langle h, D^p G\rangle_{\mathscr{H}^{\odot p}} = p!\langle h, g_p\rangle_{\mathscr{H}^{\odot p}} = \mathbb{E}I_p(h)G. \tag{2.3.2.11}$$

In other words, $\delta^p h = I_p(h)$. □

2.3.2.1. *Combining $D$ and $L$.* Comparing (2.2.1.8), (2.2.2.2) and (2.3.2.4), we see that

$$\mathscr{D}\mathrm{om}\, L = \mathscr{D}\mathrm{om}\, D^2; \quad \text{and} \quad \mathscr{D}\mathrm{om}\, D = \mathscr{D}\mathrm{om}\, \mathcal{E}. \tag{2.3.2.12}$$

The next series of statements links $L$ and $\mathcal{E}$ more precisely with the operators $D$ and $\delta$. For short, we write

$$\Gamma_0(F, G) = \langle DF, DG\rangle_{\mathscr{H}}, \quad \text{for all } F, G \in \mathscr{D}\mathrm{om}\, D. \tag{2.3.2.13}$$

Remark that, since $\mathscr{D}\mathrm{om}\, D$ is a linear space, $\Gamma_0$ is a bilinear form. The first theorem is a representation of $L$.

**Theorem 2.3.2.5.** *The self-adjoint operator $L$ coincides with $-\delta D$. Namely, $F \in \mathscr{D}\mathrm{om}\, L$ if and only if $F \in \mathscr{D}\mathrm{om}\, D$ and $DF \in \mathscr{D}\mathrm{om}\, \delta$ and in that case $LF = -\delta D$.*

*Proof.* Since $D$ and $\delta$ are closed and densely-defined and adjoint of each other. The operator $\tilde{L} = -\delta D$ is self-adjoint closed and densely defined. Observe that

$$\mathscr{D}\text{om}\, \tilde{L} = \{F \in \mathscr{D}\text{om}\, D, \text{ such that } DF \in \mathscr{D}\text{om}\, \delta\}. \tag{2.3.2.14}$$

For $F \in \mathscr{D}\text{om}\, \tilde{L}$, we have, by Proposition 2.3.2.4,

$$\tilde{L}F = \delta\left(\sum_{q \in \mathbb{N}} q I_{q-1}(h_q)\right) = LF. \tag{2.3.2.15}$$

This shows that $L$ is an extension of $\tilde{L}$. As both $L$ and $\tilde{L}$ are self-adjoint, this implies that $L = \tilde{L}$. □

From this representation, we deduce various integration by parts formulae that are summarized in the following theorem.

**Theorem 2.3.2.6** (Integration by parts)**.** *The following relations hold:*

$$\mathbb{E}\Gamma_0(F, G) = \mathcal{E}(F, G), \quad \textit{for all } F, G \in \mathscr{D}\text{om}\, D; \tag{2.3.2.16}$$

$$\mathbb{E}\Gamma_0(F, G) = -\mathbb{E}GLF, \quad \textit{for all } F \in \mathscr{D}\text{om}\, L,\ G \in \mathscr{D}\text{om}\, D; \tag{2.3.2.17}$$

$$\mathbb{E}\Gamma_0(F, G) = \mathbb{E}\Gamma(F, G), \quad \textit{for all } F, G \in \mathscr{D}\text{om}\, \Gamma. \tag{2.3.2.18}$$

*Proof.* In view of Theorem 2.3.2.5, we obtain (2.3.2.17). Since $\mathscr{D}\text{om}\, \mathcal{E} = \mathscr{D}\text{om}\, \Gamma$, the relation (2.3.2.16) is deduced from (2.3.2.17). In view of (2.2.2.13), the relation (2.3.2.18) is also a consequence of (2.3.2.16). □

By direct computations on the chaotic decomposition, we obtain, the following commutation relations.

**Theorem 2.3.2.7.** *For $F \in \mathscr{D}\text{om}\, D$,*

$$\mathrm{e}^{-t}\, P_t DF = DP_t F. \tag{2.3.2.19}$$

*For $F \in \mathscr{D}\text{om}\, D^3$,*

$$DLF = LDF - DF. \tag{2.3.2.20}$$

In particular, we see that the $\Gamma_0$ satisfies a Bakry-Emery type condition

$$\Gamma_0(P_t F) = \mathrm{e}^{-2t}\, P_t \Gamma_0(F). \tag{2.3.2.21}$$

However, it is a priori not clear how to compare $\Gamma_0$ and $\Gamma$. An important part of the rest of this chapter is about partially solving this question.

2.3.2.2. *More properties of $\delta$.* We now present several lemmas regarding the operator $\delta$. We start with an approximation for $\mathscr{D}\text{om}\, \delta$.

**Lemma 2.3.2.8.** *Let $\mathscr{V}$ be the image by $D$ of $\mathscr{D}\text{om}\, L$. Then, $\mathscr{V}$ is dense in $\mathscr{D}\text{om}\, \delta$.*

*Proof.* Let $u$ be an element of the orthogonal of $\mathscr{V}$ in $\mathscr{D}\mathrm{om}\,\delta$. Then, by duality, for all $v = DF \in \mathscr{V}$ ($F \in \mathscr{D}\mathrm{om}\,L$), we have that

$$0 = \mathbb{E}\delta u\delta v = -\mathbb{E}\delta u LF. \tag{2.3.2.22}$$

Since for all $q \in \mathbb{N}$, $\mathscr{C}_q \subset \mathscr{D}\mathrm{om}\,L$, we see that necessarily $\delta u = 0$. Since $\delta$ is injective, we infer that $u = 0$. This proves that the orthogonal of $\mathscr{V}$ in $\mathscr{L}^2(\mathfrak{W}) \circ \mathscr{H}$ is $\{0\}$ and, hence, the announced density. $\square$

**Lemma 2.3.2.9.** *Let $F$ and $G \in \mathscr{D}\mathrm{om}\,L$. Then,*

$$\mathbb{E}LFLG = \mathbb{E}\langle D^2F, D^2G\rangle_{\mathscr{H}^{\odot 2}} + \mathbb{E}\langle DF, DG\rangle_{\mathscr{H}}. \tag{2.3.2.23}$$

*Consequently, for $u$ and $v \in \mathscr{D}\mathrm{om}\,\delta$,*

$$\mathbb{E}\delta u\delta v = \mathbb{E}\langle u, v\rangle_{\mathscr{H}} + \mathbb{E}\langle Du, Dv\rangle_{\mathscr{H}^{\odot 2}}. \tag{2.3.2.24}$$

*Proof.* We assume first that $F$ and $G \in \mathscr{D}\mathrm{om}\,D^3$. By integration by parts Theorem 2.3.2.6 and the commutation (2.3.2.20), we have that

$$\begin{aligned} \mathbb{E}LFLG &= -\mathbb{E}\langle DF, DLG\rangle_{\mathscr{H}} = -\mathbb{E}\langle DF, LDG\rangle_{\mathscr{H}} + \mathbb{E}\langle DF, DG\rangle_{\mathscr{H}} \\ &= \mathbb{E}\langle D^2F, D^2G\rangle_{\mathscr{H}^{\odot 2}} + \mathbb{E}\langle DF, DG\rangle_{\mathscr{H}}. \end{aligned} \tag{2.3.2.25}$$

The previous relation is between continuous bilinear forms of $\mathscr{D}\mathrm{om}\,L$ and hold on the dense subset $\mathscr{D}\mathrm{om}\,D^3$ of $\mathscr{D}\mathrm{om}\,L = \mathscr{D}\mathrm{om}\,D^2$. Hence, it holds on all $\mathscr{D}\mathrm{om}\,L$.

By Lemma 2.3.2.8, we prove the second relation on the subset $\mathscr{V}$ dense in $\mathscr{D}\mathrm{om}\,\delta$. Let let $u$ and $v \in \mathscr{V}$ and let $F$ and $G \in \mathscr{D}\mathrm{om}\,L$ such that $u = DF$ and $v = DG$. Since by Theorem 2.3.2.5, $L = -\delta D$, by the previous relation, we have that

$$\begin{aligned} \mathbb{E}\delta u\delta v &= \mathbb{E}LFLG \\ &= \mathbb{E}\langle DF, DG\rangle_{\mathscr{H}} + \mathbb{E}\langle D^2F, D^2G\rangle_{\mathscr{H}^{\odot 2}} \\ &= \mathbb{E}\langle u, v\rangle_{\mathscr{H}} + \mathbb{E}\langle Du, Dv\rangle_{\mathscr{H}^{\odot 2}}. \end{aligned} \tag{2.3.2.26}$$

This proves the claim. $\square$

2.3.3. **Derivation and diffusion.** Recall that we have set $\mathscr{A} = \mathscr{D}\mathrm{om}\,\mathcal{E} \cap \mathscr{L}^\infty(\mathfrak{W})$ and that $\mathscr{A}$ is an algebra stable by composition with Lipschitz functions. We say that the Malliavin gradient $D$ is a *derivation* if:

$$D(FG) = FDG + GDF, \quad \text{for all } F, G \in \mathscr{A}. \tag{2.3.3.1}$$

The fact that $D$ is a derivation is equivalent to the following *chain rule*.

**Theorem 2.3.3.1.** *The operator $D$ is a derivation if and only if for all $\phi \in \mathscr{C}_1(\mathbb{R})$:*

$$D\phi(F) = \sum_{i=1}^{d} \partial_i\phi(F)DF_i, \quad F = (F_1, \dots, F_d) \in \mathscr{A}. \tag{2.3.3.2}$$

*Proof.* By selecting $\phi(x, y) = xy$ in (2.3.3.2), we see that the chain rule is sufficient for $D$ to be a derivation. Let us show that this is also necessary. Since $D$ is a derivation, by [25, III Cor of Prop 10.4.2], for all multivariate polynomials $P$,

$$DP(F) = \sum_{i=1}^{d} \partial_i P(F) DF_i, \quad F = (F_1, \dots, F_d) \in \mathscr{A}. \tag{2.3.3.3}$$

Observe that, since $\mathscr{A}$ is an algebra, in the previous formula $P(F) \in \mathscr{A}$. We let $M = \max_{i=1,\dots,d} |F_i|_{\mathscr{L}^\infty(\mathfrak{W})}$ and we let $K = [-2M, 2M]^d$. Without loss of generality, we can assume that $\phi$ is compactly supported in $K$. By the Stone-Weierstrass approximation theorem, we can find a sequence $(p_n)$ that converges uniformly on $K$ to $\phi$ such that $(p'_n)$ converges uniformly on $K$ to $\phi'$. Consequently, we have that

$$p_n(F) \xrightarrow[n\to\infty]{\mathscr{L}^2(\mathfrak{W})} \phi(F); \quad \text{and} \quad p'_n(F) \xrightarrow[n\to\infty]{\mathscr{L}^2(\mathfrak{W})} \phi'(F). \tag{2.3.3.4}$$

Since $D$ is a closed operator, this proves the claim. □

By duality we also obtain the following *Leibniz rule* for the divergence.

**Theorem 2.3.3.2.** *Let $u \in \mathscr{D}\mathrm{om}\,\delta$ and $F \in \mathscr{A}$ such that $Fu \in \mathscr{D}\mathrm{om}\,\delta$, we have*

$$\delta(Fu) = F\delta u - \langle u, DF\rangle_{\mathscr{H}}. \tag{2.3.3.5}$$

*Proof.* For short, we write $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_{\mathscr{H}}$. Let $G \in \mathscr{A}$. Then,

$$\mathbb{E}G\delta(Fu) = \mathbb{E}\langle DG, Fu\rangle = \mathbb{E}\langle FDG, u\rangle. \tag{2.3.3.6}$$

Since $D$ is a derivation over $\mathscr{A}$ and $F$ and $G \in \mathscr{A}$, we have that $FDG = D(FG) - GDF$. This shows that, for all $G \in \mathscr{A}$,

$$\mathbb{E}G\delta(Fu) = \mathbb{E}GF\delta u - \mathbb{E}G\langle DF, u\rangle. \tag{2.3.3.7}$$

By density of $\mathscr{A}$ in $\mathscr{D}\mathrm{om}\,D$ and definition of $\delta$, this proves the claim. □

We have the following theorem connecting $D$ and $L$ in a diffusive case.

**Theorem 2.3.3.3.** *The Malliavin gradient $D$ is a derivation if and only if the Ornstein-Uhlenbeck generator $L$ is a diffusion and*

$$\Gamma_0(F, G) = \Gamma(F, G), \quad \text{for all } F, G \in \mathscr{D}\mathrm{om}\,\Gamma. \tag{2.3.3.8}$$

*Proof.* Assume $D$ is derivation. By (2.3.2.18), we have that

$$\mathbf{\Gamma}(F)[\Phi] = \mathbb{E}\langle DF, D(F\Phi)\rangle - \frac{1}{2}\mathbb{E}\langle DF^2, D\Phi\rangle = \mathbb{E}\Phi|DF|^2, \quad \text{for all} F, \Phi \in \mathscr{A}. \tag{2.3.3.9}$$

By density of $\mathscr{A}$ in $\mathscr{D}\mathrm{om}\,\mathcal{E}$ we find that $\mathscr{D}\mathrm{om}\,\Gamma = \mathscr{D}\mathrm{om}\,\mathcal{E}$ and that (2.3.3.8) holds. Let $F$ and $G \in \mathscr{D}\mathrm{om}\,\mathcal{E}$ such that $FG = 0$. Then $D(FG) = FDG + GDF = 0$. Hence, we find that $1_{F\neq 0}DG = 1_{G\neq 0}DF = 0$. By (2.3.2.18), we have that

$$\mathcal{E}(F, G) = \mathbb{E}\langle DF, DG\rangle = \mathbb{E}1_{F=0}1_{G=0}\langle DF, DG\rangle = 0. \tag{2.3.3.10}$$

This shows that $L$ is a diffusion.

Conversely, assume that $L$ is a diffusion and that (2.3.3.8) holds. Let $F$ and $G \in \mathscr{A}$. By Lemma 2.3.2.8, to show that $D$ is a derivation, it is sufficient to show that

$$\mathbb{E}\langle D(FG), DH\rangle = \mathbb{E}F\langle DG, DH\rangle + \mathbb{E}G\langle DF, DH\rangle, \quad \text{for all } H \in \mathscr{D}\mathrm{om}\, L. \tag{2.3.3.11}$$

By (2.3.3.8), we have that this expression is equivalent to

$$\mathcal{E}(FG, H) = \mathbb{E}F\Gamma(G, H) + \mathbb{E}G\Gamma(F, H). \tag{2.3.3.12}$$

This expression holds true by the chain rule (2.2.2.18). This concludes the proof. □

The following lemma provides an abstract product formula when $D$ is a derivation. Recall that we say that $\mathscr{L}^2(\mathfrak{W})$ supports the mixed bosonic Fock space (Example 2.3.1.2), when the Fock space $\mathscr{H}^\circ = \oplus \mathscr{H}^{\odot q}$ (as sets) and there exists positive constants $(\lambda_q)$ such that $\langle \cdot, \cdot \rangle_{\mathscr{H}^{\circ q}} = \lambda_q = \langle \cdot, \cdot \rangle_{\mathscr{H}^{\odot q}}$. The $(\lambda_q)$ will be referred as the *bosonic constants*. In a more explicit way, this means that for every $h \in \mathscr{H}^{\odot q}$ and $\tilde{h} \in \mathscr{H}^{\odot p}$ ($q$ and $p \in \mathbb{N}$) the quantities $I_q(h)$ and $I_p(\tilde{h})$ are well-defined and

$$\mathbb{E}I_q(h)I_p(\tilde{h}) = 1_{p=q}\lambda_q\langle h, \tilde{h}\rangle_{\mathscr{H}^{\odot q}} = 1_{p=q}\lambda_q q!\nu^q(h\tilde{h}). \tag{2.3.3.13}$$

**Lemma 2.3.3.4.** *Assume the Malliavin derivative is a derivation then $\mathscr{L}^2(\mathfrak{W})$ supports the mixed bosonic Fock space (meaning that $\mathscr{H}^{\circ q} = \mathscr{H}^{\odot q}$ as sets, for all $q \in \mathbb{N}$), $\mathscr{L}^2(\mathfrak{W})$ has polynomial chaoses and for all $q$ and $p \in \mathbb{N}$, $h \in \mathscr{H}^{\circ p}$ and $\tilde{h} \in \mathscr{H}^{\circ q}$ such that $I_p(h)I_q(\tilde{h}) \in \mathscr{L}^2(\mathfrak{W})$, we have that*

$$J_{p+q}(I_p(h)I_q(\tilde{h})) = I_{p+q}(h \odot \tilde{h}). \tag{2.3.3.14}$$

*Proof.* The proof is the same as the one of Lemma 2.4.2.1 that is a similar statement in a non-diffusive setting. We do not reproduce it here. □

The product formula will help us controlling the remainder that appeared in Theorem 2.2.3.2. The same behaviour will be observed in a non-diffusive setting (see Theorem 2.4.4.6). Recall that for $F \in \mathscr{L}^4(\mathfrak{W})$, we write $\mathbb{M}(F) = \mathbb{E}F^4 - 3(\mathbb{E}F^2)^2$. Given $c > 0$, we also write $\mathbb{M}_c(F) = \mathbb{E}F^4 - c(\mathbb{E}F^2)^2$.

**Theorem 2.3.3.5.** *Assume $\mathscr{L}^2(\mathfrak{W})$ supports a Fock and that $D$ is a derivation. Let $(\lambda_q)$ be the bosonic constant associated with $\mathscr{L}^2(\mathfrak{W})$ (they exist by Lemma 2.3.3.4). Let $(p_1, \ldots, p_d) \in \mathbb{N}^d$. For all $q \in \mathbb{N}$, we let $c_q = 1 + 2\frac{\lambda_{2q}}{\lambda_q^2}$. There exists $c > 0$, such that for all $F = (F_1, \ldots, F_d)$ with $F_i \in \mathscr{C}_{p_i} \cap \mathscr{L}^4(\mathfrak{W})$, we have that*

(2.3.3.15)

$$\begin{aligned}\mathbb{E}|S(F) - \mathbb{E}F^TF|_{HS}^2 \leq c\sum_{i}^{d} \mathbb{M}(F_i) + c\sum_{\substack{i,j=1 \\ p_j < p_i}}^{d} (\mathbb{E}F_i^4)^{\frac{1}{2}}\mathbb{M}(F_j)^{\frac{1}{2}} \\ + c\sum_{\substack{i,j=1 \\ i \neq j \\ p_i = p_j}}^{d} \left(\mathbb{M}(F_i)^{\frac{1}{2}}\mathbb{M}(F_j)^{\frac{1}{2}} + \mathbb{M}_{c_{p_i}}(F_i)^{\frac{1}{2}}\mathbb{M}_{c_{p_i}}(F_j)^{\frac{1}{2}} + \frac{\lambda_{2p_i} - \lambda_{p_i}^2}{\lambda_{p_i}}\mathbb{E}F_i\mathbb{E}F_j\right).\end{aligned}$$

**Remark 5.** This theorem is useful if and only if $c_q \geq 3$, that is $\lambda_{2q} \geq \lambda_q^2$. We however state it in full generality in order for the reader to observe the mechanic preventing a fourth moment theorem to be recovered.

*Proof.* See [41, Lem 2.3] Let $q \in \mathbb{N}$, $h$ and $\tilde{h} \in \mathscr{H}^{\odot q}$. We write $F = I_q(h)$ and $G = I_q(\tilde{h})$. In view of Theorem 2.2.3.2, it is enough to show that

$$|\mathbb{E}J_{2q}F^2 J_{2q}G^2 - 2(\mathbb{E}FG)^2| \leq c\mathbb{M}_{c_q}(F_i)^{\frac{1}{2}}\mathbb{M}_{c_q}(F_j)^{\frac{1}{2}} + \frac{\lambda_q - \lambda_q^2}{\lambda_q}\mathbb{E}F_i\mathbb{E}F_j. \tag{2.3.3.16}$$

By Lemma 2.3.3.4,

$$\mathbb{E}F^4 = (\mathbb{E}F^2)^2 + \sum_{k=1}^{2q-1} \mathbb{E}\left[\left(J_k F^2\right)^2\right] + \lambda_{2q}|h \odot \tilde{h}|^2_{\mathscr{H}^{\odot 2q}}. \tag{2.3.3.17}$$

It is an algebraic fact of tensor calculus [113, Lem 2.2(2)] that

$$|h \odot \tilde{h}|^2_{\mathscr{H}^{\odot 2q}} = 2\langle h, \tilde{h}\rangle^2_{\mathscr{H}^{\odot q}} + \sum_{r=1}^{q-1} q!^2 \binom{q}{r} \langle h \otimes_r \tilde{h}, \tilde{h} \otimes_r h\rangle_{\mathscr{H}^{\otimes 2(q-r)}}, \tag{2.3.3.18}$$

where the exact expression of $h \otimes_r \tilde{h} \in \mathscr{H}^{\otimes 2(q-r)}$ is irrelevant (see [111, Appendix B.4]). In particular, we infer that
(2.3.3.19)

$$\mathbb{E}F^4 = \left(1 + 2\frac{\lambda_{2q}}{\lambda_q^2}\right)(\mathbb{E}F^2)^2 + \sum_{k=1}^{2q-1} \mathbb{E}\left[\left(J_k F^2\right)^2\right] + \lambda_{2q}\sum_{l=1}^{q-1} q!^2 \binom{q}{l} |h \otimes_r h|^2_{\mathscr{H}^{\otimes 2(q-l)}}.$$

By Lemma 2.3.3.4 and (2.3.3.18), we find that
(2.3.3.20)

$$\mathbb{E}J_{2q}F^2 J_{2q}G^2 - 2(\mathbb{E}FG)^2 = 2(\lambda_{2q} - \lambda_q^2)\langle h, \tilde{h}\rangle^2_{\mathscr{H}^{\odot q}} + \lambda_{2q}\sum_{r=1}^{q-1} q!^2 \binom{q}{r} \langle h \otimes_r \tilde{h}, \tilde{h} \otimes_r h\rangle_{\mathscr{H}^{\otimes 2(q-r)}}.$$

Applying the Cauchy-Schwarz inequality several times and some algebraic identities similar to the one in the proof of [41, Lem 2.3], we find that

$$\begin{aligned} &\sum_{r=1}^{q-1} q!^2 \binom{q}{r} \langle h \otimes_r \tilde{h}, \tilde{h} \otimes_r h\rangle_{\mathscr{H}^{\otimes 2(q-r)}} \\ &\leq \sqrt{\sum_{r=1}^{q-1} (q!)^2 \binom{q}{r}^2 |h \otimes_r h|_{\mathscr{H}^{\otimes 2(q-r)}}} \sqrt{\sum_{r=1}^{q-1} (q!)^2 \binom{q}{r}^2 |\tilde{h} \otimes_r \tilde{h}|_{\mathscr{H}^{\otimes 2(q-r)}}} \end{aligned} \tag{2.3.3.21}$$

Combining the previous expression with (2.3.3.19) yields

(2.3.3.22)

$$\begin{aligned} |\mathbb{E}J_{2q}F^2 J_{2q}G^2 - 2(\mathbb{E}FG)^2| &\leq (\lambda_{2q} - \lambda_q^2)\langle h, \tilde{h}\rangle^2_{\mathscr{H}^{\odot q}} \\ &+ \sqrt{\mathbb{E}F^4 - \left(1 + 2\frac{\lambda_{2q}}{\lambda_q}\right)(\mathbb{E}F^2)^2}\sqrt{\mathbb{E}G^4 - \left(1 + 2\frac{\lambda_{2q}}{\lambda_q}\right)(\mathbb{E}G^2)^2}. \end{aligned}$$

This concludes the proof. □

**Outline.** We recall that, we want to study $\mathfrak{W} = \sigma(e)$, where $e$ is a random object of interest living on a measurable space $(E, \mathfrak{E})$. In this section, we assume that $\mathscr{L}^2(\mathfrak{W})$ has a chaotic decomposition with Ornstein-Uhlenbeck generator $L$ but we will drop the assumption that $L$ is a diffusion. However, getting practical results under the only assumption that $L$ is not diffusive seems unreachable and we will introduce and study a particular class of such non-diffusive $L$. We recall that, by definition, every $\mathfrak{W}$-random variable $F$ has a *representative* $f$ such that $F = f(e)$. The representative is unique up to $law(e)$-negligible sets. In the rest of the section, whenever $F$ and $G \in \mathscr{L}^2(\mathfrak{W})$, the symbols $f$ and $g$ will designate one of their respective representatives. In Section 2.4.1, we will suggest a class of non-diffusive generators from which we expect positive results. These will be generators for which the carré du champ has the following representation

$$\Gamma(F, G) = \int (f(y) - f(e))(g(y) - g(e))q(e, \mathrm{d}y), \quad \text{for all } F, G \in \mathscr{D}\mathrm{om}\,\Gamma, \tag{2.4.0.1}$$

where $q = \{q(x, A),\ x \in E,\ A \in \mathfrak{E}\}$ is a collection of measures satisfying some conditions, ensuring that this representation does not depend on the choice of $f$ and $g$. We will call such generators *pure-jump type* generators. However we were not able to develop a technology in order to obtain functional inequalities and limit theorems in the setting of those operators in full generality. For that reason, in Section 2.4.2, we will study a particular class of pure-jump generators for which the measure $q$ can be computed explicitly. We will assume the existence of a Fock space structure and we will consider a condition on the Malliavin derivative in order to quantify the non-diffusiveness. Recall that for $F \in \mathscr{D}\mathrm{om}\, F$, $DF$ is a random element of $\mathscr{H} \subset \mathscr{L}^2(Z, \mathfrak{Z}, \nu)$ where $\nu$ is a $\sigma$-finite measure on the measurable space $(Z, \mathfrak{Z})$. Roughly speaking, we will require that for all $z \in Z$, there exists a random variable $C_z$, and a map $T_z$ satisfying some conditions such that

$$D_z F = C_z(f(T_z e) - f(e)), \quad \text{for all } F \in \mathscr{D}\mathrm{om}\, D,\ z \in Z. \tag{2.4.0.2}$$

Under this sole condition, we obtain a representation of the carré du champ $\Gamma$ and the Ornstein-Uhlenbeck generator $L$ in Theorems 2.4.2.3 and 2.4.2.4. In particular, we always have that $\mathscr{D}\mathrm{om}\,\Gamma = \mathscr{D}\mathrm{om}\,\mathcal{E}$ and $2\Gamma(F) = \Gamma_0(F) + T(F)$ where $T(F)$ is an explicit term that can be thought as a randomized derivative. In this setting, we can show a pseudo-chain rule Lemma 2.4.3.1, a modified logarithmic Sobolev inequality Theorem 2.4.4.1 and a modified Stein inequality Theorem 2.4.4.2. As in the diffusive case, under an additional polynomial chaoses assumption, we deduce a fourth moment theorem with quartic remainder Theorem 2.4.4.3. In Sections 2.6 and 2.7, we will see, that this framework is particularly well-adapted to study point processes. We give conditions on the covariance of multiple stochastic integrals, in order for this fourth moment theorem with remainder to simplify to an exact fourth moment theorem Theorem 2.4.4.6, that is the theorem is as good (up to numerical constants) as the fourth moment theorem for diffusions (Theorem 2.2.3.2). The proof of this exact fourth moment theorem is based on an abstract product formula for stochastic integrals (Lemma 2.4.2.1) of independent interest. These results are new at this level of generality and give a unified framework to understand several recent results on limit theorems in a discrete setting ([40, 38, 41]). We will study in Section 2.7.3 the two important cases of Poisson point processes and mixed binomial processes.

2.4.1. **Pure-jump generators.** This short section is only aimed at providing the reader with an intuition concerning a possibly "nice" form of the generator $L$. Unfortunately, contrary to the diffusive setting, we are not yet able to obtain positive results in the setting presented here that solely concerns the operator $L$ and this is why we present the more advanced technology of transitive operators based on the Malliavin derivative in Section 2.4.3. An *absolutely continuous* kernel on $E$ is a collection of non-negative real numbers $q = \{q(x, A)\}$ for $x \in E$ and $A \in \mathfrak{E}$ such that: for all $A \in \mathfrak{E}$, the map $x \mapsto q(x, A)$ is measurable, and, for all $x \in E$, $q(x, \cdot)$ is a non-negative measure on $(E, \mathfrak{E})$ absolutely continuous with respect to the law of $e$. We say that the semi-group is of *pure-jump* type if, there exists a kernel $q$ on $E$ such that:

$$\Gamma(F, G) = \frac{1}{2} \int (f(y) - f(e))(g(y) - g(e))q(e, \mathrm{d}y), \quad \text{for all } F, G \in \mathscr{D}\mathrm{om}\,\Gamma. \tag{2.4.1.1}$$

Remark that the assumption of absolute continuity makes this definition independent of the choice of the representative. The symmetry of the semigroup implies that the measure $J(\mathrm{d}x\mathrm{d}y) = q(x, \mathrm{d}y)law(e)(\mathrm{d}x)$ is symmetric.

**Remark 6.** By defining the *covariant derivative* as the unbounded operator $\nabla_x \colon \mathscr{L}^2(\nu) \to \mathscr{L}^2(q(x, \mathrm{d}y))$ with $(\nabla_x f)(y) = f(y) - f(x)$ and the *connection* $\nabla \colon \mathscr{L}^2(\nu) \to \mathscr{L}^2(J)$ as $\nabla f = f \ominus f$, that is $\nabla f(x, y) = f(y) - f(x)$. We think of $\mathscr{L}^2(J)$ as a vector bundle and $\nabla$ as a connection (hence the name). We have that $F \in \mathscr{D}\mathrm{om}\,\Gamma$ if and only if $f \in \mathscr{D}\mathrm{om}\,\nabla_e$, and

$$\Gamma(F, G) = \nabla_e f \cdot \nabla_e g, \qquad \text{for all } F, G \in \mathscr{D}\mathrm{om}\,\Gamma; \tag{2.4.1.2}$$

$$\mathbb{E}\Gamma(F, G) = \nabla f \cdot \nabla g, \qquad \text{for all } F, G \in \mathscr{D}\mathrm{om}\,\mathcal{E}. \tag{2.4.1.3}$$

With $\nabla^*$ the adjoint of $\nabla$, we find that

$$LF = -\nabla^* \nabla f(e). \tag{2.4.1.4}$$

This expression has to be compared with the one obtained in Theorem 2.3.2.5, stating that in the setting of Malliavin operators $L = -\delta D$. Thus, we see that we have two competing geometrical structures: the *flat* structure provided by the Malliavin calculus directly on $\mathscr{L}^2(\mathfrak{W})$, and the *curved* structure provided by the connection on $\mathscr{L}^2(\mathfrak{E})$. It is not yet clear how to use the operator $\nabla$ and in particular how it interacts with stochastic integrals or the Ornstein-Uhlenbeck semi-group. For instance, we can obtained a modified Stein inequality similar to the one we will obtain in Theorem 2.4.4.2 in the setting of pure-jump operators but the remainder is not tractable as it is expressed in terms of $\nabla$. Better understanding the operator $\nabla$ is of independent interest and this work will be completed elsewhere.

2.4.2. **Transitive operators.** In this section, we fix a measured space $(Z, \mathfrak{Z}, \nu)$, we let $\mathscr{H} = \mathscr{L}^2(\nu)$ and we assume that $\mathscr{L}^2(\mathfrak{W})$ has an abstract Fock space $\mathscr{H}^\circ = \oplus_{q \in \mathbb{N}} \mathscr{H}^{\circ q}$. For $F \in \mathscr{D}\mathrm{om}\, D$ and $z \in Z$, we write $\partial_z f$ for a representative of $D_z F$. We say that the Malliavin derivative $D$ is *representable with a transitive action* if there exists a measurable mapping $T \colon Z \times E \to E$ and a measurable mapping $c : E \to \mathbb{R}$, such that, for all $z \in Z$, the map $T_z$ is injective, the map $c_z$ is bounded and never takes the value $0$ and

$$\partial_z f(e) = c_z(e)(f(T_z e) - f(e)), \quad \text{for all } F = f(e) \in \mathscr{D}\mathrm{om}\, D,\ z \in Z, \tag{2.4.2.1}$$

where we write $T_z x$ for $T(z,x)$ and $c_z(e)$ for $c(z,e)$. The map $T$ is called the *action map* or the *transitive action* and $c_z$ the *mixing map*. We also need to assume that this formula is independent of the choice of the representatives $f$ and $\partial_z f$ (this will be easily checked on examples).

**Remark 7.** We could take $T_z$ and $c_z$ with an extra randomness in which case we would ask that almost surely, for all $z \in Z$, the random map $T_z$ is injective, $c_z$ is bounded and never vanishes and

$$\partial_z f(e) = \mathbb{E}[c_z(e)(f(T_z e) - f(e))|e], \tag{2.4.2.2}$$

where we write $T_z x$ for the random variable $T(z,x)$ and $c_z(e)$ for the random variable $c(z,e)$. Most of the definitions and theorems would adapt straightforwardly. However, since we do not know applications for this possible generalization, we choose to keep the definitions as simple as possible.

For the rest of the section, we assume that $D$ is representable with action map $T$ and mixing map $c$. For $z \in Z$, we set $C_z = c_z(e)$ and $\tilde{C}_z = c_z(T_z^{-1}e)1_{\{e \in \operatorname{im} T_z\}}$. We say moreover that the Malliavin derivative is *pure* if $C_z = 1$ for all $z \in Z$. We define the map $D^+$ by

$$D_z^+ F = f(T_z e) - f(e), \quad \text{for all } z \in Z,\ F \in \mathscr{L}^0(\mathfrak{W}). \tag{2.4.2.3}$$

Then, we obtain that for all $F \in \mathscr{D}\mathrm{om}\, D$, $DF = CD^+F$. Conversely, a $\mathfrak{W}$-random variable $F$ belongs to $\mathscr{D}\mathrm{om}\, D$ if and only if $D^+F \in \mathscr{L}^2(\mathfrak{W}) \otimes \mathscr{H}$ and in that case $DF = CD^+F$. Note that, $D^+$ is everywhere defined, while $D$ has, in general, a smaller domain. We also define,

$$D_z^- F = (f(e) - f(T_z^{-1}e))1_{\{e \in \operatorname{im} T_z\}}, \quad \text{for all } z \in Z,\ F \in \mathscr{L}^0(\mathfrak{W}). \tag{2.4.2.4}$$

Observe that the operator $D^+$ and $D^-$ are not derivations but satisfy the combinatorial properties

$$D_z^+(FG) = FD_z^+G + GD_z^+F + D^+FD_z^+G \tag{2.4.2.5}$$
$$D_z^-(FG) = FD_z^-G + GD_z^-F - D_z^-FD_z^-G. \tag{2.4.2.6}$$

In particular, this indicates that Malliavin gradients representable with transitive action are unlikely to be derivations. In particular, if the action is pure, then $D = D^+$ and $D$ is not a derivation, unless in the trivial case where $T_z$ is the identity map for all $z \in Z$. Note that we can of course iterate the definition of $D^+$ and $D^-$. Given $l \in \mathbb{N}$ and $z_1, \dots, z_l \in Z$, we write

$$D^{+l}_{z_1,\dots,z_l} F = D^+_{z_1} \dots D^+_{z_l} F, \tag{2.4.2.7}$$

with the convention that $D^{+0} = 1$. It is immediate to check that, with $f$ a representative of $F$,

$$D^{+l}_{z_1,\dots,z_l} F = \sum_{\{j_1,\dots,j_k\} = J \subset [l]} \left( (-1)^{l-|J|} f\left( T_{z_{j_1}} \circ \dots \circ T_{z_{j_k}} e \right) \right). \tag{2.4.2.8}$$

Recall that the mixed bosonic Fock space (associated to $\mathscr{H} = \mathscr{L}^2(\nu)$) is given by $\mathscr{H}^\circ = \oplus_q \mathscr{H}^{\circ q}$ with, for all $q \in \mathbb{N}$, $\mathscr{H}^{\circ q} = \mathscr{L}^2_\sigma(\nu^q)$ (as sets) such that there exists a sequence of positive constants $(\lambda_q)$ such that $\langle\cdot,\cdot\rangle_{\mathscr{H}^{\circ q}} = \langle\cdot,\cdot\rangle_{\mathscr{H}^{\odot q}}$, and that we call the sequence $(\lambda_q)$ the bosonic constants. In a more explicit way, this means that for every $h \in \mathscr{H}^{\odot q}$ and $\tilde{h} \in \mathscr{H}^{\odot p}$ ($q$ and $p \in \mathbb{N}$) the quantities $I_q(h)$ and $I_p(\tilde{h})$ are well-defined and

$$\mathbb{E}I_q(h)I_p(\tilde{h}) = 1_{p=q}\lambda_q\langle h,\tilde{h}\rangle_{\mathscr{H}^{\odot q}} = 1_{p=q}\lambda_q q!\nu^q(h\tilde{h}). \tag{2.4.2.9}$$

We have the following lemma giving an abstract product formula when the Malliavin derivative is pure (compare to [39] and Lemma 2.3.3.4).

**Lemma 2.4.2.1.** *Assume that the Malliavin derivative is pure. Then $\mathscr{L}^2(\mathfrak{W})$ supports the mixed bosonic Fock space, has polynomial chaoses and for $F = I_p(h_p)$ and $G = I_q(h_q)$, where $p$ and $q \in \mathbb{N}$ and $h_p \in \mathscr{H}^{\odot p}$ and $h_q \in \mathscr{H}^{\odot q}$ such that $FG \in \mathscr{L}^2(\mathfrak{W})$ we have that*

$$J_{p+q}(I_p(h_p)I_q(h_q)) = I_{p+q}(h_q \odot h_p). \tag{2.4.2.10}$$

*Proof.* By assumption, we have that for $F \in \mathscr{D}\mathrm{om}\, D$, $DF = D^+F$. We start by proving that $\mathscr{L}^2(\mathfrak{W})$ has polynomial chaoses. From Theorem 2.3.2.2, we know that

$$FG = \sum_{k=0}^{\infty} I_k(\tilde{h}_k), \tag{2.4.2.11}$$

where

$$\tilde{h}_k(z_1,\dots,z_k) = \frac{1}{k!}\mathbb{E}D^k_{z_1,\dots,z_k}(FG), \quad k \in \mathbb{N}. \tag{2.4.2.12}$$

Hence, the property of polynomial chaoses will follow from the claim:

$$FG \in \mathscr{D}\mathrm{om}\, D^m \text{ and } D^m(FG) = 0, \quad \text{for all } m > p+q. \tag{2.4.2.13}$$

We prove the claim by induction on $p+q$. If $p+q = 2$ then $p = q = 1$. Since $F$ and $G \in \mathscr{C}_1$, we find that both $D^{+2}F = D^2F$ and $D^{+2}G = D^2G$ vanish. By (2.4.2.5), we find that $D^{+3}(FG) = 0$. This shows that $FG \in \mathscr{D}\mathrm{om}\, D^3$ and $D^3(FG) = 0$. This proves the claim for $p+q = 2$. We assume that $p+q > 2$ and we let $m > p+q$. By (2.4.2.5), we also have that

$$\begin{aligned} D^{+m}_{z_1\dots z_m}FG = {} & pD^{+(m-1)}_{z_1\dots z_{m-1}}(I_{p-1}(h_p(z_m,\cdot))I_q(h_q)) \\ & + qD^{+(m-1)}_{z_1\dots z_{m-1}}(I_{q-1}(h_q(z_m,\cdot))I_p(h_p)) \\ & + qpD^{+(m-1)}_{z_1\dots z_{m-1}}(I_{q-1}(h_q(z_m,\cdot))I_{p-1}(h_p(z_m,\cdot))). \end{aligned} \tag{2.4.2.14}$$

By the induction hypothesis, all the terms in the right-hand side vanish and this proves the claim (2.4.2.13). Let us prove at once that $\mathscr{L}^2(\mathfrak{W})$ supports the mixed bosonic Fock space and that $J_{p+q}(FG) = I_{p+q}(h_p \odot h_q)$. Let us prove by induction on $p+q$ that

$$D^{p+q}(FG) = (p+q)!h_p \odot h_q. \tag{2.4.2.15}$$

Regarding the previous expression, by assumption, $FG \in \mathscr{L}^2(\mathfrak{W})$ and, since we proved that the chaoses are polynomials, we know that $FG \in \mathscr{C}$ and hence $FG \in \mathscr{D}\mathrm{om}\, D^{p+q}$. If $p+q = 2$ then $p = q = 1$, by (2.4.2.5), we have that $D^2F$ and $D^2G$ vanish and that

$$D^2_{z_1,z_2}(FG) = D_{z_1}FD_{z_2}G + D_{z_1}GD_{z_2}F = h_p(z_1)h_q(z_2) + h_p(z_2)h_q(z_1) = 2h_p \odot h_q. \tag{2.4.2.16}$$

Now if $p+q>2$, we obtain that

$$\begin{aligned} D^{p+q}_{z_1\dots z_{p+q}} FG &= pD^{(p+q-1)}_{z_1\dots z_{p+q-1}}(I_{p-1}(h_p(z_{p+q},\cdot))I_q(h_q)) \\ &\quad + qD^{(p+q-1)}_{z_1\dots z_{p+q-1}}(I_{q-1}(h_q(z_{p+q},\cdot))I_p(h_p)) \\ &\quad + qpD^{(p+q-1)}_{z_1\dots z_{p+q-1}}(I_{q-1}(h_q(z_{p+q},\cdot))I_{p-1}(h_p(z_{p+q},\cdot))). \end{aligned} \tag{2.4.2.17}$$

By the fact that the chaoses are polynomial the last line vanishes, and, by the induction hypothesis, we find that

$$\begin{aligned} D^{p+q}_{z_1\dots z_{p+q}} FG &= p(p-1+q)!(h_p(z_{p+q},\cdot)\odot h_q)(z_1,\dots,z_{p+q-1}) \\ &\quad + q(q-1+p)!(h_q(z_{p+q},\cdot)\odot h_p)(z_1,\dots,z_{p+q-1}). \end{aligned} \tag{2.4.2.18}$$

In view of [40, Eq 6.3], this proves the claim (2.4.2.15). Let us conclude the proof. By Theorem 2.3.2.2, (2.4.2.15) shows that for all $p$ and $q\in\mathbb{N}$,

$$\mathscr{H}^{\circ p}\odot\mathscr{H}^{\circ q}\subset\mathscr{H}^{\circ(p+q)}. \tag{2.4.2.19}$$

This proves by an immediate induction that $\mathscr{H}^{\circ r}\supset\mathscr{H}^{\odot r}$ for all $r\in\mathbb{N}$. This shows that $\mathscr{L}^2(\mathfrak{W})$ supports the mixed bosonic Fock space. Also (2.4.2.10) is a consequence of Theorem 2.3.2.2 and (2.4.2.15). The proof is completed. □

A random measure $\eta$ measurable with respect to $\mathfrak{W}$ is the *Campbell measure* associated with the transitive action $T$ if, for all non-negative bi-measurable functions

$$h\colon Z\times E\ni(z,e)\mapsto h_z(e)\in\mathbb{R}_+, \tag{2.4.2.20}$$

we have that

$$\mathbb{E}\int h_z(e)\eta(\mathrm{d}z)=\mathbb{E}\int h_z(T_z e)\nu(\mathrm{d}z). \tag{2.4.2.21}$$

Note that, if it exists, the intensity of a Campbell measure $\eta$ is given by $\nu$, that is for all $A\in\mathfrak{Z}$,

$$\mathbb{E}\eta(A)=\nu(A). \tag{2.4.2.22}$$

It is classical, that if $h$ is as before without the positivity constraint and

$$\mathbb{E}\int|h_z(e)|\nu(\mathrm{d}z)<\infty, \tag{2.4.2.23}$$

then, we can extend (2.4.2.21) to this $h$. We present the following lemma.

**Lemma 2.4.2.2.** *Assume that the transitive action $T$ admits the Campbell measure $\eta$. Almost surely $\eta(\mathrm{d}z)$-almost everywhere $e\in\operatorname{im}T_z$.*

*Proof.* Let $A\in\mathfrak{Z}$, $z\in Z$ and $h_z(x)=1_A 1_{\{x\in\operatorname{im}T_z\}}$ for $x\in Z$. By definition, we have that

$$h_z(T_z e)=1_A 1_{\{T_z e\in\operatorname{im}T_z\}}=1_A,\quad\text{for all } z\in Z. \tag{2.4.2.24}$$

Applying (2.4.2.21), we thus find

$$\mathbb{E}\eta(A)=\nu(A)=\mathbb{E}\int_A 1_{\{e\in\operatorname{im}T_z\}}\eta(\mathrm{d}z). \tag{2.4.2.25}$$

Hence, the desired conclusion. □

Finally, we arrive at the following representation of $\delta$ and hence $L$ for transitive actions.

**Theorem 2.4.2.3.** *Assume that the Malliavin gradient $D$ is representable by the transitive map $T$, mixing map $C$ and admits the Campbell measure $\eta$. For $u \in \mathscr{D}\mathrm{om}\,\delta$ such that $u \in \mathscr{L}^1(\mathfrak{W} \otimes \mathfrak{E}, \mathbb{P} \otimes \nu)$, we have that*

$$\delta u = \int c_z(T_z^{-1}e)h_z(T_z^{-1}e)\eta(\mathrm{d}z) - \int c_z(e)h_z(e)\nu(\mathrm{d}z), \tag{2.4.2.26}$$

*where $h$ is a representative of $u$. Let*

$$\mathscr{D}\mathrm{om}_0\, L = \{F \in \mathscr{D}\mathrm{om}\, L, \text{ such that } DF \in \mathscr{L}^1(\mathfrak{W} \otimes \mathfrak{E}, \mathbb{P} \otimes \nu)\}. \tag{2.4.2.27}$$

*For all $F \in \mathscr{D}\mathrm{om}_0\, L$, we have that*

$$LF = \int C_z^2 D_z^+ F\nu(\mathrm{d}z) - \int \tilde{C}_z^2 D_z^- F\eta(\mathrm{d}z). \tag{2.4.2.28}$$

*In particular, for pure actions, for such $F$:*

$$LF = \int D_z^+ F\nu(\mathrm{d}z) - \int D_z^- F\eta(\mathrm{d}z). \tag{2.4.2.29}$$

*Proof.* Recall that we defined the algebra $\mathscr{A} = \mathscr{D}\mathrm{om}\,\mathcal{E} \cap \mathscr{L}^\infty(\mathfrak{W})$. We start by proving (2.4.2.26). Observe that if $u \in \mathscr{L}^1(\mathfrak{W} \otimes \mathfrak{Z}, \mathbb{P} \otimes \nu)$ by (2.4.2.21),

$$\mathbb{E}\int |(1 - D_z^-)u_z|\eta(\mathrm{d}z) < \infty. \tag{2.4.2.30}$$

The formula is proved by duality. Let $u$ be as in the theorem and $F \in \mathscr{A}$, then we have

$$\mathbb{E}\langle u, DF\rangle_{\mathscr{H}} = \mathbb{E}\int c_z(e)h_z(e)(f(T_z e) - f(e))\nu(\mathrm{d}z). \tag{2.4.2.31}$$

Since $F$ is bounded, by (2.4.2.21), we find that

$$\mathbb{E}\int c_z(e)h_z(e)f(T_z e)\nu(\mathrm{d}z) = \mathbb{E}\int c_z(T_z^{-1}e)h_z(T_z^{-1}e)\eta(\mathrm{d}z). \tag{2.4.2.32}$$

This proves (2.4.2.26) by density of $\mathscr{A}$ in $\mathscr{D}\mathrm{om}\, D$. To prove (2.4.2.28), we simply write $L = -\delta D$ and apply (2.4.2.26). □

Finally, we arrive at the announced representation of the carré du champ.

**Theorem 2.4.2.4.** *Let the assumptions of Theorem 2.4.2.3 prevail. Then, $\mathscr{D}\mathrm{om}\,\Gamma = \mathscr{D}\mathrm{om}\, D$ and*

$$\Gamma(F, G) = \frac{1}{2}\int C_z D_z^+ F D_z^+ G\nu(\mathrm{d}z) + \frac{1}{2}\int \tilde{C}_z D_z^- F D_z^- G\eta(\mathrm{d}z), \quad \text{for all } F, G \in \mathscr{D}\mathrm{om}\, D. \tag{2.4.2.33}$$

*Proof.* First of all remark that that by (2.4.2.21), the right-hand side of (2.4.2.33) is well-defined whenever $F$ and $G \in \mathscr{D}\mathrm{om}\, D$. We write $T(F, G)$ for the right-hand side of (2.4.2.33) and $T(F)$ when $F = G$. By polarization it is enough to show the theorem for $F = G$ in (2.4.2.33) and by density of $\mathscr{A}$ in $\mathscr{D}\mathrm{om}\,\mathcal{E}$, we can restrict the proof to the case $F \in \mathscr{A}$. Let $F \in \mathscr{A}$. From Theorem 2.4.2.3, we easily compute that

$$\mathcal{E}(F) = \mathbb{E}T(F). \tag{2.4.2.34}$$

Let $\Phi \in \mathscr{A}$. In view of (2.4.2.5) and (2.4.2.6), we find that

$$\begin{aligned}\mathcal{E}(F, F\Phi) &= \mathbb{E}\Phi T(F) + \mathbb{E}FT(\Phi, F)\\ &\quad + \frac{1}{2}\mathbb{E}\int C_z(D_z^+F)^2 D_z^+\Phi\nu(\mathrm{d}z) - \frac{1}{2}\mathbb{E}\int \tilde{C}_z(D_z^-F)^2 D_z^-\Phi\eta(\mathrm{d}z).\end{aligned} \tag{2.4.2.35}$$

Since $\Phi$ is, by assumption, bounded, it admits a ($law(e)$-almost everywhere) bounded representative and thus $D_z^+\Phi$ and $D_z^-\Phi$ are also bounded. Thus we can apply (2.4.2.21) to the last term of the right-hand side of the previous expression and we see that the last line vanishes. We can obtain a similar expression for $\mathcal{E}(\Phi, F^2)$ and we obtain that

$$\mathcal{E}(F, F\Phi) - \frac{1}{2}\mathcal{E}(F^2, \Phi) = \mathbb{E}\Phi T(F),\ \forall \Phi \in \mathscr{A}. \tag{2.4.2.36}$$

Hence $\Gamma(F) = T(F)$. This completes the proof. □

In view of the previous theorem and (2.4.2.21), the following integration by parts holds:

$$\mathcal{E}(F, G) = \mathbb{E}\Gamma_0(F, G) = \mathbb{E}\Gamma(F, G) \quad \text{for all } F, G \in \mathscr{D}\mathrm{om}\, D. \tag{2.4.2.37}$$

However, the equality $\Gamma_0 = \Gamma$ does not hold in general.

2.4.3. **Difference operators.** We say that the Malliavin derivative $D$ is a *difference operator* if, for all $F \in \mathscr{D}\mathrm{om}\, D$, $\phi$ such that $\phi(F) \in \mathscr{D}\mathrm{om}\, D$ and $z \in Z$, we have that

$$D_z\phi(F) = \phi(F + D_zF) - \phi(F). \tag{2.4.3.1}$$

We now observe that transitive operators act as difference operators and, hence enjoy an ersatz of the chain rule. Indeed, let $D$ have a transitive action, with the notations of the previous section, we have that, for all $F \in \mathscr{L}^2(\mathfrak{W})$, $\phi$ such that $\phi(F) \in \mathscr{D}\mathrm{om}\, D$ and $z \in Z$,

$$D_z\phi(F) = C_z(\phi(F + D_z^+F) - \phi(F)). \tag{2.4.3.2}$$

We have the following pseudo chain rule.

**Lemma 2.4.3.1.** *Assume the Malliavin gradient is representable by a transitive action and let the previous notations prevail. Let $F = (F_1, \dots, F_d)$ and $G \in \mathscr{D}\mathrm{om}\, D$ such that*

$$DGDF_iDF_j \in \mathscr{L}^1(\mathfrak{W} \otimes \mathfrak{E}, \mathbb{P} \otimes \nu), \quad i, j \in [d]. \tag{2.4.3.3}$$

*Let $\phi \in \mathscr{C}^1(\mathbb{R}^d)$ such that $\phi(F) \in \mathscr{D}\mathrm{om}\, D$. Then,*

$$\Gamma_0(\phi(F), G) = \langle \nabla\phi(F), \Gamma_0(F, G)\rangle_{\ell^2} + R^+_\phi(F, G); \tag{2.4.3.4}$$

$$\Gamma(\phi(F), G) = \langle \nabla\phi(F), \Gamma(F, G)\rangle_{\ell^2} + R_\phi(F, G), \tag{2.4.3.5}$$

*where*

$$R_\phi(F, G) = \frac{1}{2}\left(R^+_\phi(F, G) - R^-_\phi(F, G)\right);$$

$$R^+_\phi(F, G) = \sum_{i,j=1}^{d} \int C_z^2 D_z^+ G D_z^+ F_i D_z^+ F_j R^+_{ij}(z)\nu(\mathrm{d}z);$$

$$R^-_\phi(F, G) = \int \tilde{C}_z^2 D_z^- G D_z^- F_i D_z^- F_j R^-_{ij}(z)\eta(\mathrm{d}z),$$

*with*

$$R^+_{ij}(z) = \int_0^1 \int_0^1 \alpha \partial_{ij}\phi(F + \alpha\beta D_z^+ F)\mathrm{d}\alpha\mathrm{d}\beta;$$

$$R^-_{ij}(z) = \int_0^1 \int_0^1 \alpha \partial_{ij}\phi(F - \alpha\beta D_z^- F)\mathrm{d}\alpha\mathrm{d}\beta.$$

*Moreover, if $D$ is a difference operator, the formula (2.4.3.4) holds with $D^+ = D$ and $C = 1$.*

*Proof.* We give the proof only for $d = 1$, the generalization to higher dimension being straightforward. By the fundamental theorem of calculus, we have

$$\phi(x + h) - \phi(x) = h \int_0^1 \phi'(x + \alpha h)\mathrm{d}\alpha. \tag{2.4.3.6}$$

Therefore, applying this formula once more, we find that

$$\phi(x + h) - \phi(x) - h\phi'(x) = h^2 \int_0^1 \int_0^1 \alpha\phi''(x + \alpha\beta h)\mathrm{d}\alpha\mathrm{d}\beta. \tag{2.4.3.7}$$

Applying the previous relation to $x = F$, $h = D_z^+ F$ and $x = F$, $h = -D_z^- F$, we have that

$$D_z^+\phi(F) = \phi'(F)D_z^+ F + (D_z^+ F)^2 \int_0^1 \int_0^1 \alpha\phi''(F + \alpha\beta D_z^+ F)\mathrm{d}\alpha\mathrm{d}\beta; \tag{2.4.3.8}$$

$$D_z^-\phi(F) = \phi'(F)D_z^- F - (D_z^- F)^2 \int_0^1 \int_0^1 \alpha\phi''(F - \alpha\beta D_z^- F)\mathrm{d}\alpha\mathrm{d}\beta. \tag{2.4.3.9}$$

Multiplying the two previous equation by $D_z^+ G$ (resp. $D_z^- G$), and integrating against $\nu$ (resp. $\eta$), proves the claim for the case of transitive actions. The generalization to difference operators is immediate. □

2.4.4. **Non-diffusive inequalities.** We will use the previous formalism to deduce functional inequalities and quantitative limit theorems. Recall that the spectral gap inequality Theorem 2.2.1.3 was obtained from the chaotic decomposition only and thus also holds here. We obtain the following modified logarithmic Sobolev inequality (see [152, Thm 1.1]). Consider the convex function $\phi(x) = x \log x$ for $x \geq 0$. Recall that the entropy functional is defined for a positive random variable as $\mathcal{H}(F) = \mathbb{E}\phi(F) - \phi(\mathbb{E}F)$.

**Theorem 2.4.4.1** (Modified logarithmic Sobolev inequality). *Assume that $\mathscr{L}^2(\mathfrak{W})$ has a chaotic decomposition and a Fock space. If the Malliavin gradient $D$ has a transitive action, then, for all $F \in \mathscr{D}\mathrm{om}\, D$ with $F > 0$, we have*

$$\mathcal{H}(F) \leq \mathbb{E} \int C_z^2 (D_z^+ \phi(F) - \phi'(F) D_z^+ F) \nu(\mathrm{d}z). \tag{2.4.4.1}$$

*Moreover, if $D$ is a difference operator the conclusion of the theorem holds with $D^+ = D$ and $C = 1$.*

*Proof.* The proof follows the lines of [31, Thm 5.1]. Since, by Theorem 2.4.2.4, $\mathscr{D}\mathrm{om}\, \Gamma = \mathscr{D}\mathrm{om}\, \mathcal{E}$, the computations carried out at the beginning of the proof of Theorem 2.2.2.1 are valid and we can start from (2.2.2.22), that is:

$$\mathcal{H}(F) = \int_0^\infty \mathbb{E}\Gamma_0(\phi'(P_t F), P_t F). \tag{2.4.4.2}$$

Since $D$ has a transitive action, we obtain that, for all $t \geq 0$,

$$\begin{aligned} \Gamma_0(\phi'(P_t F), P_t F) &= \int C_z^2 (\phi'(P_t F + D_z^+ P_t F) - \phi'(P_t F)) D_z^+ P_t F \nu(\mathrm{d}z) \\ &= \int C_z^2 \psi(P_t F, \mathrm{e}^{-t} P_t D_z F) \nu(\mathrm{d}z), \end{aligned} \tag{2.4.4.3}$$

where $\psi(u, v) = v(\phi'(u+v) - \phi'(u))$ is convex. By Jensen's inequality, we find that

$$\Gamma_0(\phi'(P_t F), P_t F) \leq P_t \int C_z^2 \psi(F, \mathrm{e}^{-t} D_z^+ F) \nu(\mathrm{d}z), \quad \forall t \geq 0. \tag{2.4.4.4}$$

Observe that

$$\int_0^\infty \mathrm{e}^{-t} \phi'(F + \mathrm{e}^{-t} t D_z^+ F) \mathrm{d}t = \frac{D_z^+ \phi(F)}{D_z^+ F}. \tag{2.4.4.5}$$

By invariance of the semi-group, this yields the desired conclusion for transitive actions by integrating (2.4.4.4) with respect to $\mathbb{E}$. Again, the generalization to difference operator is easily obtained. □

We now turn to inequalities involving the law of random variables. Namely, we want to establish a Stein inequality in this non-diffusive setting. As we will see, we obtain a result similar to Theorem 2.2.3.1 up to a remainder term. Recall that we work with the symmetrized matrix-valued carré du champ, that is whenever $F = (F_1, \ldots, F_{d_1}) \in \mathscr{D}\mathrm{om}\, \Gamma$ and $G = (G_1, \ldots, G_{d_2}) \in \mathscr{D}\mathrm{om}\, \Gamma$, we write $\Gamma(F, G)$ for the random symmetric matrix whose coefficient $(i, j) \in [d_1] \times [d_2]$ is given by

$$\frac{1}{2}(\Gamma(F_i, G_j) + \Gamma(F_j, G_i)). \tag{2.4.4.6}$$

We adopt a similar convention for $\Gamma_0$. We define for $F = (F_1, \dots, F_{d_1}) \in \mathscr{L}^2(\mathfrak{W})$ centered and $G = (G_1, \dots, G_{d_2}) \in \mathscr{D}\mathrm{om}\, D$:

$$S_0(F, G) = -\Gamma_0(L^{-1}F, G). \tag{2.4.4.7}$$

In the setting of transitive actions, recall that we have defined the Stein kernel, for such $F$ and $G$, by

$$S(F, G) = -\Gamma(L^{-1}F, G). \tag{2.4.4.8}$$

Remark that since $L^{-1}F \in \mathscr{D}\mathrm{om}\, D^2 \subset \mathscr{D}\mathrm{om}\, D$, no further assumptions are needed on $F$. For such $F$ and $G$, we have the following integration by parts

$$\mathbb{E}F^T G = \mathbb{E}S_0(F, G) = \mathbb{E}S(F, G). \tag{2.4.4.9}$$

As usual, we write $S(F) = S(F, F)$ and $S_0(F) = S_0(F, F)$. In the setting of transitive actions, we write,

$$\epsilon(F) = \sum_{i,j,k=1}^{d} \mathbb{E} \int C_z^2 |D_z^+ L^{-1} F_i||D_z^+ F_j||D_z^+ F_k| \nu(\mathrm{d}z), \quad F \in \mathscr{L}^0(\mathfrak{W}). \tag{2.4.4.10}$$

If $D$ is a difference operator, we use the same notation with $C = 1$ and $D^+ = D$. Remark that this quantity is well-defined, though potentially infinite, for all random variables $F$. We recall that $d_2$ designates a distance introduced in Section 1.1 and that $d_2$ induces a topology stronger than the one of the convergence in law. We obtain the following Stein inequality in a discrete setting. To my knowledge, this is the first quantitative bounds measuring the distance between the law of a functional of a generic non-diffusive probabilistic object and a multivariate Gaussian law in any dimension. The only two other references, we are aware of, where convergence of multivariate functionals in a non-diffusive setting is considered are [41] for the Poisson space and G. ZHENG (2017) [153] for the Rademacher space and the authors do not provide bounds for general functionals.

**Theorem 2.4.4.2** (Stein inequality)**.** *Suppose that $\mathscr{L}^2(\mathfrak{W})$ has a chaotic decomposition and a Fock space. Assume that the Malliavin derivative $D$ is a difference operator or has a transitive action, then there exists $c > 0$, such that, for all $F \in \mathscr{D}\mathrm{om}\, D$ such that $\mathbb{E}F = 0$ and $S(F) \in \mathscr{L}^2(\mathfrak{W})$, with $C = \mathbb{E}F^T F$,*

$$d_2(law(F), \mathbf{N}(0, C)) \leq c\sqrt{d}(|C^{-1}|_{op}|C|_{op}^{\frac{1}{2}}\sqrt{\mathbb{E}|S_0(F) - \sigma^2|_{\ell^2}^2} + |C^{-\frac{3}{2}}|_{op}|C|_{op}^{\frac{1}{2}}\epsilon(F)). \tag{2.4.4.11}$$

*If $D$ has a transitive action, then*

$$d_2(law(F), \mathbf{N}(0, C)) \leq c\sqrt{d}(|C^{-1}|_{op}|C|_{op}^{\frac{1}{2}}\sqrt{\mathbb{E}|S(F) - \sigma^2|_{\ell^2}^2} + |C^{-\frac{3}{2}}|_{op}|C|_{op}^{\frac{1}{2}}\epsilon(F)). \tag{2.4.4.12}$$

*Proof.* The proof is very similar to Theorem 2.3.3.5. As the content is essentially new at this level of generality, we give a complete proof. We start by proving the claim for $S_0$. Assume $F \in \mathscr{A}$. Let $\phi \in \mathscr{C}^2(\mathbb{R})$. By the integration by parts (2.4.4.9) and the pseudo-chain rule Lemma 2.4.3.1 (these operations are justified in view of the properties of

$\mathscr{A}$), we have that

$$\begin{aligned}\text{(2.4.4.13)}\qquad \mathbb{E}\langle F, \nabla\phi(F)\rangle_{\ell^2} &= \sum_{i=1}^{d} \mathbb{E}\Gamma(L^{-1}F_i, \partial_i\phi(F)) \\ &= \langle \nabla^2\phi(F), S_0(F)\rangle_{\ell^2} + \sum_{i=1}^{d} R^{+}_{\partial_i\phi}(F, L^{-1}F_i).\end{aligned}$$

It follows that

(2.4.4.14)

$$\mathbb{E}\langle F, \nabla\phi(F)\rangle_{\ell^2} - \mathbb{E}\langle C, \nabla^2\phi(F)\rangle_{\ell^2} = \mathbb{E}\langle \nabla^2\phi(F), (S_0(F) - C)\rangle_{\ell^2} + \sum_{i=1}^{d} R^{+}_{\partial_i\phi}(F, L^{-1}F_i).$$

Observe that, by the Cauchy-Schwarz inequality, on the one hand we have that

$$\text{(2.4.4.15)}\qquad |\mathbb{E}\langle \nabla^2\phi(F), S_0(F) - C\rangle_{\ell^2}| \leq |\nabla^2\phi|_{\ell^2,\infty}\mathbb{E}|S_0(F) - C|_{\ell^2},$$

and on the other hand we have that

$$\text{(2.4.4.16)}\qquad \left|\sum_{i=1}^{d} R^{+}_{\partial_i\phi}(F, L^{-1}F_i)\right| \leq \sum_{i,j,k=1}^{d} |\partial_{ijk}\phi|_\infty \int |D_z^+ L^{-1}F_i||D_z^+F_j||D_z^+F_k|\nu(\mathrm{d}z).$$

We use Theorem 1.1.3.2 to conclude.

In the case of transitive actions, if we work with $S$ rather than $S_0$, the quantity $R^+$ has to be replaced with $R$. Let $\eta$ the Campbell measure associated to the action, by (2.4.2.21), we obtain that

$$\text{(2.4.4.17)}\qquad \left|\mathbb{E}\sum_{i=1}^{d} R_{\partial_i\phi}(F, L^{-1}F)\right| \leq |\nabla^3\phi|_{\ell^2,\infty}\epsilon(F).$$

Then, we conclude as before. □

The main interest of working with $S$ rather than with $S_0$ is that it behaves well with respect to the stochastic decomposition and hence we can hope for a simplification of the $\mathbb{E}|S(F) - C|^2_{\ell^2}$ as in Theorem 2.2.3.2. We say that $\mathscr{L}^2(\mathfrak{W})$ has a *transitive discrete Itô structure* if it has an Itô structure (that is, chaotic decomposition with Fock space and polynomial chaoses) and if the associated Malliavin derivative is representable by a transitive action, and we say that it has a *pure Itô structure* if, moreover, the Malliavin derivative is pure. For short, let us write

$$\text{(2.4.4.18)}\qquad \Delta(F) = \mathbb{E}\int C_z^2 \left(D_z^+F\right)^4 \nu(\mathrm{d}z), \quad F \in \mathscr{L}^0(\mathfrak{W}).$$

The quantity $\Delta(F)$ is well-defined (though potentially infinite) for all random variables $F$. Also recall that for $F \in \mathscr{L}^4(\mathfrak{W})$, we write $\mathbb{M}(F) = \mathbb{E}F^4 - 3(\mathbb{E}F^2)^2$ and we write $\tilde{\mathbb{M}}(F) = \mathbb{M}(F) + \Delta(F)$. Generally, we can obtain the following fourth moment theorem with remainder (compare to Theorem 2.2.3.2). This theorem, was essentially obtained, in a Poisson setting, in [41] via exchangeable pairs techniques rather than the Stein method.

**Theorem 2.4.4.3.** *Assume $\mathscr{L}^2(\mathfrak{W})$ has a transitive discrete Itô structure. Let $(p_1, \dots, p_d) \in \mathbb{N}^d_{>0}$. There exists $c > 0$, such that for all $F = (F_1, \dots, F_d)$ with $F_i \in \mathscr{C}_{p_i} \cap \mathscr{L}^4(\mathfrak{W})$, we have that*

(2.4.4.19)

$$\mathbb{E}|S(F) - \mathbb{E}F^T F|^2_{HS} \leq c \sum_{i}^{d} \tilde{\mathbb{M}}(F_i) + c \sum_{\substack{i,j=1 \\ p_j < p_i}}^{d} \left(\mathbb{E}F_i^4\right)^{\frac{1}{2}} \tilde{\mathbb{M}}(F_j)^{\frac{1}{2}}$$

$$+ c \sum_{\substack{i,j=1 \\ i \neq j \\ p_i = p_j}}^{d} \left( \tilde{\mathbb{M}}(F_i)^{\frac{1}{2}} \tilde{\mathbb{M}}(F_j)^{\frac{1}{2}} + \left[\mathbb{E}J_{2p_i}F_j^2 J_{2p_i}F_i^2 - 2(\mathbb{E}F_iF_j)^2\right]_+ \right).$$

$$\epsilon(F) \leq \left(\sum_{i=1}^{d} \left(\frac{\mathbb{E}F_i^2}{p_i}\right)\right) \left(\sum_{i=1}^{d} \Delta(F_i)^{\frac{1}{4}}\right)^2 \tag{2.4.4.20}$$

**Remark 8.** Observe that, contrary to Theorem 2.2.3.2 (diffusive fourth moment theorem), even in dimension $1$, the convergence of the fourth moment alone does not allow to recover convergence in law to a Gaussian. We also have to ensure that the quartic remainder expressed by the quantity involving $\Delta$ vanishes. The fact that $\Delta(F)$ vanishes when $F$ is a stochastic integral and $\mathbb{M}(F) = 0$ depends on the covariance structure of stochastic integrals, that is it depends on the bosonic constants (see Theorem 2.4.4.6).

*Proof.* By the Cauchy-Schwarz inequality,

$$\epsilon(F) \leq \sum_{i=1}^{d} \left(\frac{1}{p_i}\mathbb{E}S(F_i)\right)^2 \left(\sum_{j=1}^{d} \left(\mathbb{E}\int C_z^2 (D_z^+ F_j)^4 \nu(\mathrm{d}z)\right)^{1/4}\right)^2. \tag{2.4.4.21}$$

As before we have that

$$\mathbb{E}|S(F) - \mathbb{E}F^T F|^2_{\ell^2} = \sum_{i,j=1}^{d} \left(\frac{1}{p_i} + \frac{1}{p_j}\right) \mathbb{V}\mathrm{ar}(\Gamma(F_i, F_j)). \tag{2.4.4.22}$$

On the account of the two previous relations, we will prove the theorem by bounding the quantity $\mathbb{V}\mathrm{ar}(\Gamma(F_i, F_j))$ for all $i$ and $j \in [d]$. Let us first prove the following lemma (compare to Lemma 2.2.3.3 obtained in a diffusive setting). Note that obtaining such lemma is at the heart of the proof of the proof of the fourth moment theorem on the Poisson space (see, for instance [40, Lemmas 3.1 & 3.2]).

**Lemma 2.4.4.4.** *For all $p \in \mathbb{N}$ and $G \in \mathscr{C}_p \cap \mathscr{L}^4(\mathfrak{W})$, we have that*

$$\mathbb{V}\mathrm{ar}(S(G)) \leq c \left(\mathbb{E}G^4 - 3\left(\mathbb{E}G^2\right)^2 + \Delta(G)\right). \tag{2.4.4.23}$$

*Proof.* Let $p \in \mathbb{N}$ and $G \in \mathscr{C}_p \cap \mathscr{L}^4(\mathfrak{W})$. Since polynomials in $G$ will have finite chaotic decomposition all the following are justified. By integration by parts (2.2.2.15) and the fact that $LG = -pG$, we have that

$$\mathbb{E}G^4 = -\frac{1}{p}\mathbb{E}G^3 LG = \frac{1}{p}\mathbb{E}\Gamma(G^3, G). \tag{2.4.4.24}$$

Using (2.4.3.2) with $\phi(x) = x^3$ and using the binomial formula, we see that

$$\begin{aligned} 2\mathbb{E}\Gamma(G^3, G) &= \mathbb{E}\int C_z^2 \left( \left(D_z^+ G\right)^4 + 3G^2\left(D_z^+ G\right)^2 + 3G\left(D_z^+ G\right)^3 \right) \nu(\mathrm{d}z) \\ &\quad + \mathbb{E}\int \tilde{C}_z^2 \left( \left(D_z^- G\right)^4 + 3G^2\left(D_z^- G\right)^2 - 3G\left(D_z^- G\right)^3 \right) \eta(\mathrm{d}z). \end{aligned} \tag{2.4.4.25}$$

Applying (2.4.2.21) to the last term in the second integral makes that the last term of each integral will sum up to

$$-3\mathbb{E}\int C_z^2 \left(D_z^+ G\right)^4 \nu(\mathrm{d}z). \tag{2.4.4.26}$$

Thus, we find

$$\mathbb{E}\Gamma(G^3, G) = 3\mathbb{E}G^2\Gamma(G) - \frac{1}{2}\mathbb{E}\int C_z^2 \left(D_z^+ G\right)^4 \nu(\mathrm{d}z). \tag{2.4.4.27}$$

Observe that

$$\Gamma(G) = \left(\frac{1}{2}L + p\right) G^2 = p\mathbb{E}G^2 + \sum_{q=1}^{2p} \left(p - \frac{q}{2}\right) J_q G^2. \tag{2.4.4.28}$$

By orthogonality of the chaotic decomposition, we obtain that

$$\mathbb{E}G^2\Gamma(G) = p\left(\mathbb{E}G^2\right)^2 + \sum_{q=1}^{2p} \left(p - \frac{q}{2}\right) \mathbb{E}\left[\left(J_q G^2\right)^2\right]. \tag{2.4.4.29}$$

By definition of $L$ and $\Gamma$, we find that

$$\mathbb{V}\mathrm{ar}(pS(G)) = \mathbb{E}\left(\Gamma(G) - p\mathbb{E}G^2\right)^2 = \sum_{q=1}^{2p} \left(p - \frac{q}{2}\right)^2 \mathbb{E}\left[\left(J_q G^2\right)^2\right]. \tag{2.4.4.30}$$

Finally, we have that

$$\begin{aligned} \mathbb{V}\mathrm{ar}(pS(G)) &\leq \left(p - \frac{1}{2}\right) \left(\mathbb{E}G^2\Gamma(G) - p(\mathbb{E}G^2)^2\right) \\ &= \left(p - \frac{1}{2}\right) \left(\frac{p}{3}\mathbb{E}G^4 - p(\mathbb{E}G^2)^2 + \frac{1}{6}\Delta(G)\right). \end{aligned} \tag{2.4.4.31}$$

This concludes the proof of Lemma 2.4.4.4. □

The following lemma is the discrete counterpart of Lemma 2.2.3.4.

**Lemma 2.4.4.5.** *Let $q$ and $p \in \mathbb{N}$, $F \in \mathscr{C}_p \cap \mathscr{L}^4(\mathfrak{W})$ and $G \in \mathscr{C}_q \cap \mathscr{L}^4(\mathfrak{W})$, we have that*

$$\mathbb{V}\mathrm{ar}(\Gamma(F, G)) \leq c\left(\mathbb{E}\left[F^2\left(qG^2 - \Gamma(G)\right)\right] - 2q(\mathbb{E}FG)^2 + \mathbb{E}\int C_z^2 (D_z^+ F)^2 (D_z^+ G)^2 \nu(\mathrm{d}z)\right), \tag{2.4.4.32}$$

*with $c = \frac{p+q-1}{4}$.*

*Proof.* We assume moreover that $F$ and $G \in \mathscr{A}$. As in the diffusive case, we write

$$\mathbb{E}F^2G^2 = \frac{1}{q}\mathbb{E}\Gamma(F^2G, G) = \frac{1}{q}\mathbb{E}F^2\Gamma(G) + \frac{1}{q}\mathbb{E}G\Gamma(F^2, G). \tag{2.4.4.33}$$

The second inequality follows from the following observation (already made in the proof of Theorem 2.4.2.4) that the Dirichlet energy acts as a derivation. Indeed, by (2.4.2.5) and (2.4.2.6)

$$\begin{aligned}\mathbb{E}\Gamma(AB, C) &= \mathbb{E}A\Gamma(B, C) + \mathbb{E}B\Gamma(A, C)\\ &\quad + \frac{1}{2}\int C_z^2 D_z^+ A D_z^+ B D_z^+ C \nu(\mathrm{d}z) - \frac{1}{2}\int \tilde{C}_z^2 D_z^- A D_z^- B D_z^- C \eta(\mathrm{d}z).\end{aligned} \tag{2.4.4.34}$$

And the two last terms cancel out provided we can apply (2.4.2.21) (this is the case if $A$, $B$ and $C \in \mathscr{A}$). A similar argument yields

$$\begin{aligned}\mathbb{E}G\Gamma(F^2, G) &= 2\mathbb{E}FG\Gamma(F, G)\\ &\quad + \frac{1}{2}\mathbb{E}\int G(D_z^+F)^2 D_z^+ G\nu(\mathrm{d}z) - \frac{1}{2}\mathbb{E}\int G(D_z^-F)^2 D_z^- G\eta(\mathrm{d}z)\\ &= 2\mathbb{E}FG\Gamma(F, G) - \frac{1}{2}\mathbb{E}\int (D_z^+F)^2 (D_z^+G)^2 \nu(\mathrm{d}z).\end{aligned} \tag{2.4.4.35}$$

Eventually, we proved that

$$\mathbb{E}F^2G^2 = \frac{2}{q}\mathbb{E}FG\Gamma(F, G) + \frac{1}{q}\mathbb{E}F^2\Gamma(G) - \frac{1}{2q}\mathbb{E}\int (D_z^+F)^2(D_z^+G)^2\nu(\mathrm{d}z). \tag{2.4.4.36}$$

By writing the chaotic decomposition and the previous relation, we obtain

(2.4.4.37)

$$\begin{aligned}\mathbb{V}\mathrm{ar}(\Gamma(F, G)) &= \sum_{k=1}^{p+q}\left(\frac{p+q-k}{2}\right)^2 \mathbb{E}\left[(J_k FG)^2\right]\\ &\leq \frac{p+q-1}{2}\left(\mathbb{E}FG\Gamma(F, G) - q(\mathbb{E}FG)^2\right)\\ &= \frac{p+q-1}{2}\left(\frac{q}{2}\mathbb{E}F^2G^2 - \frac{1}{2}\mathbb{E}F^2\Gamma(G) - q(\mathbb{E}FG)^2 + \mathbb{E}\int (D_z^+F)^2(D_z^+G)^2\right).\end{aligned}$$

This proves Lemma 2.4.4.5. □

It suffices to bound, by the Cauchy-Schwarz inequality,

$$\mathbb{E}\int (D_z^+F)^2(D_z^+G)^2\nu(\mathrm{d}z) \leq \sqrt{\Delta(F)}\sqrt{\Delta(G)}, \tag{2.4.4.38}$$

and the end of the proof is the same as the one of Theorem 2.2.3.2. □

The following theorem asserts that if $\mathscr{L}^2(\mathfrak{W})$ has a pure Itô structure, then we can simplify the previous expressions thanks to the product formula. Again the bosonic constants $(\lambda_q)$ that appear when computing the covariance of stochastic integrals will play a role. As before, we write, for $F \in \mathscr{L}^4$ and $c \in \mathbb{R}$, $\mathbb{M}_c(F) = \mathbb{E}F^4 - c(\mathbb{E}F^2)^2$. This is the first multivariate quantitative bound in the Monge-Kantorovich-Rubinstein distance in a non-diffusive setting (compare to [41, Thm 1.7]).

**Theorem 2.4.4.6.** *Assume $\mathscr{L}^2(\mathfrak{W})$ has a pure Itô structure. Let $(\lambda_q)$ be the bosonic constants associated with $\mathscr{L}^2(\mathfrak{W})$ (that exists thanks to Lemma 2.4.2.1). Let $(p_1, \dots, p_d) \in \mathbb{N}^d_{>0}$. There exists $c > 0$, such that for all $F = (F_1, \dots, F_d)$ with $F_i \in \mathscr{C}_{p_i} \cap \mathscr{L}^4(\mathfrak{W})$, we have that, with $c_q = 1 + 2\frac{\lambda_{2q}}{\lambda_q^2}$, for all $q \in \mathbb{N}$*

(2.4.4.39)
$$\begin{aligned}\mathbb{E}|S(F) - \mathbb{E}F^T F|^2_{HS} \leq c \sum_{i}^{d} \mathbb{M}(F_i) + \mathbb{M}_{c_i}(F_i) + c \sum_{\substack{i,j=1 \\ p_j < p_i}}^{d} \left(\mathbb{E}F_i^4\right)^{\frac{1}{2}} \left(\mathbb{M}(F_j) + \mathbb{M}_{c_j}(F_j)\right)^{\frac{1}{2}} \\ + c \sum_{\substack{i,j=1 \\ i \neq j \\ p_i = p_j}}^{d} \left( \left(\mathbb{M}(F_i) + \mathbb{M}_{c_i}(F_i)\right)^{\frac{1}{2}} \left(\mathbb{M}(F_j) + \mathbb{M}_{c_j}(F_j)\right)^{\frac{1}{2}} \right) \\ + c \sum_{\substack{i,j=1 \\ i \neq j \\ p_i = p_j}}^{d} \left( \mathbb{M}_{c_{p_i}}(F_i)^{\frac{1}{2}} \mathbb{M}_{c_{p_i}}(F_j)^{\frac{1}{2}} + \frac{\lambda_{2p_i} - \lambda_{p_i}^2}{\lambda_{p_i}} \mathbb{E}F_i \mathbb{E}F_j \right).\end{aligned}$$

**Remark 9.** Contrary to the diffusive case (Theorem 2.3.3.5) the bosonic constants play a role even in dimension $d = 1$.

*Proof.* We start from the bound obtained in Theorem 2.4.4.3. The bound on

$$\left[\mathbb{E}J_{2p_i}F_j^2 J_{2p_i}F_i^2 - 2(\mathbb{E}F_iF_j)^2\right]_+, \quad i \neq j,\ p_i = p_j, \tag{2.4.4.40}$$

is obtained as in the proof of Theorem 2.3.3.5.

Hence, it is sufficient to show that

(2.4.4.41)
$$\Delta(F) \leq c\left(\mathbb{E}F^4 - c_q\left(\mathbb{E}F^2\right)^2\right) + c\left(\mathbb{E}F^4 - 3\left(\mathbb{E}F^2\right)^2\right), \quad \text{for all } p \in \mathbb{N},\ F \in \mathscr{C}_p \cap \mathscr{L}^4(\mathfrak{W}).$$

Let $p \in \mathbb{N}$ and $F = I_p(h) \in \mathscr{C}_p \cap \mathscr{L}^4(\mathfrak{W})$. From Lemma 2.4.2.1, we know that $J_{2p}F^2 = I_{2p}(h \odot h)$. Hence, we conclude that

$$\mathbb{E}F^4 - \mathbb{E}F^2 = \mathbb{V}\mathrm{ar}(F^2) = \sum_{q=1}^{2p-1} \mathbb{E}\left[\left(J_qF^2\right)^2\right] + \lambda_{2q}|h \odot h|^2_{\mathscr{H}^{\odot 2p}}. \tag{2.4.4.42}$$

It is a algebraic fact of tensor calculus (see for instance [111, Eq 5.2.12]) that

$$|h \odot h|^2_{\mathscr{H}^{\odot 2p}} \geq 2|h|^4_{\mathscr{H}^{\odot p}} = 2\left(\mathbb{E}F^2\right)^2. \tag{2.4.4.43}$$

In particular, we find that

$$\sum_{q=1}^{2p-1} \mathbb{E}\left[\left(J_qF^2\right)^2\right] \leq \mathbb{E}F^4 - c_q\left(\mathbb{E}F^2\right)^2. \tag{2.4.4.44}$$

From the proof of Theorem 2.4.4.3, we know that

$$\frac{1}{2}\mathbb{E}\int (D_z^+F)^4\nu(\mathrm{d}z) = 3\mathbb{E}F^2\Gamma(F) - p\mathbb{E}F^4$$

$$\leq p\left(3(\mathbb{E}F^2)^2 - \mathbb{E}F^4\right) + 3\sum_{q=1}^{2p-1}\left(p-\frac{q}{2}\right)\mathbb{E}\left[(J_qF^2)^2\right] \tag{2.4.4.45}$$

$$\leq p\left(3(\mathbb{E}F^2)^2 - \mathbb{E}F^4\right) + 3\left(p-\frac{1}{2}\right)\sum_{q=1}^{2p-1}\mathbb{E}\left[(J_qF^2)^2\right]$$

Combining the two last equations, we obtain (2.4.4.41) and this concludes the proof. □

## 2.5. AN HISTORICAL INTERLUDE

**Outline.** We first present the historical and canonical setting of the diffusive Itô structure yielded by isonormal Gaussian processes. This section is only present in order for the reader to get acquainted with the abstract notions introduced in the previous sections. None of the results presented here are original.

### 2.5.1. Isonormal Gaussian processes.

Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of independent and identically distributed (real) normal random variables. The *Gauss space* $\mathscr{G}$ is the $\mathscr{L}^2(\Omega)$-closure of the linear span of the $G_i$'s. The space $\mathscr{G}$ is a separable Hilbert sub-space of $\mathscr{L}^2(\Omega)$ that only contains Gaussian random variables. Let $\mathscr{H}$ be a separable Hilbert space. An *isonormal Gaussian process over* $\mathscr{H}$ is any isometric embedding $W\colon \mathscr{H}\to\mathscr{G}$. We write $\mathfrak{W}$ for the $\sigma$-algebra generated by $W$ and $\mathscr{L}^2(\mathfrak{W}) = \mathscr{L}^2(\Omega,\mathfrak{W},\mathbb{P})$. The goal of this section is to show that, with the notions introduce before, $\mathscr{L}^2(\mathfrak{W})$ has a diffusive Itô structure, thus yielding immediately all the results mentioned in Section 2.2.

**Example 2.5.1.1** (Gaussian vector)**.** If $\mathscr{H}$ is $\mathbb{R}^d$ with the scalar product induced by a symmetric positive definite matrix $A$, we let $X$ be the centered Gaussian vector with covariance $A$. Then for all $h\in\mathscr{H}$, $W(h) = \langle h, X\rangle_{\ell^2}$.

**Example 2.5.1.2** (Reproducing kernel)**.** We consider a continuous centered Gaussian process $X = (X_t)_{t\in[0,1]}$ such that $\mathbb{E}X_1^2 = 1$. For $s$ and $t\in[0,1]$, we write $R(s,t) = \mathbb{E}X_sX_t$ for the covariance function and we assume that $R$ is continuous. We let

$$R_t = R(t,\cdot) = [0,1]\ni s\mapsto R(s,t),\quad \text{for all } t\in[0,1]. \tag{2.5.1.1}$$

By the property of the covariance kernel, the operator $T\colon \mathscr{L}^2(0,1)\to\mathscr{L}^2(0,1)$ defined by

$$(T\phi)(s) = \int R(s,t)\phi(t)\mathrm{d}t,\quad \text{for all } s\in[0,1], \tag{2.5.1.2}$$

is Hilbert-Schmidt. By Mercer's theorem [43, XI.8.50.E.58 p. 1008], it admits an orthogonal basis of continuous eigenfunctions $(e_i)_{i\in\mathbb{N}}$ associated with a sequence of non-negative eigenvalues $(\lambda_i)_{i\in\mathbb{N}}$ and we have that

$$R(t,s) = \sum_{i\in\mathbb{N}}\lambda_i e_i(t)e_i(s),\quad t,s\in[0,1], \tag{2.5.1.3}$$

where the series converges uniformly. So that,

$$R_t = \sum_{i\in\mathbb{M}} \lambda_i e_i(t) e_i, \quad t \in [0,1]. \tag{2.5.1.4}$$

We consider

$$\mathscr{H} = \left\{ f \in \mathscr{L}^2(0,1), \text{ such that } \sum_{i\in\mathbb{N}} \frac{(f\cdot e_i)^2}{\lambda_i} < \infty \right\}, \tag{2.5.1.5}$$

where $\cdot$ is the standard scalar product on $\mathscr{L}^2(0,1)$. The space $\mathscr{H}$ is a Hilbert space when endowed with the scalar product

$$\langle f, g\rangle_{\mathscr{H}} = \sum_{i\in\mathbb{N}} \frac{(f\cdot e_i)(g\cdot e_i)}{\lambda_i}. \tag{2.5.1.6}$$

Observe that we have

$$\langle R_t, R_s\rangle_{\mathscr{H}} = R(s,t), \quad t,s \in [0,1]. \tag{2.5.1.7}$$

and that for $h$ continuous

$$\langle R_t, h\rangle = h(t), \quad t \in [0,1]. \tag{2.5.1.8}$$

We consider $W$ an isonormal Gaussian process over $\mathscr{H}$, then we have that $\{W(R_t);\ t \in [0,1]\}$ is a Gaussian process with covariance function $R$. This isonormal Gaussian process is called the *reproducing kernel representation* of $X$.

**Example 2.5.1.3** (Covariance kernel)**.** Consider a centered Gaussian process $X$ with covariance function $R(t,s) = \mathbb{E}X_t X_s$. We consider $\mathscr{H}$ the closure of the linear span of the functions of the form $R_t = 1_{[0,t]}$ for $t \in [0,1]$ under the scalar product

$$\langle R_s, R_t\rangle_{\mathscr{H}} = R(s,t), \quad s,t \in [0,1]. \tag{2.5.1.9}$$

We let $W$ be an isonormal Gaussian process over the Hilbert space $\mathscr{H}$. Then, the process $\{W(R_t);\ t \in [0,1]\}$ is a Gaussian process with covariance $R$ called the *covariance representation* of $X$. Contrary to the previous example, this method works without any assumptions on the continuity of $X$ or of its covariance kernel. Nonetheless, it is a priori not clear that the Hilbert space $\mathscr{H}$ is contained in a space of functions (and it is, in general, not see [129]).

**Example 2.5.1.4** (Stationary Gaussian field)**.** We consider a centered stationary real-valued Gaussian field $X = (X_x)_{x\in\mathbb{R}^d}$ with $\mathbb{E}X_0^2 = 1$. For $x$ and $y \in \mathbb{R}^d$, we consider its covariance kernel $R(x,y) = \mathbb{E}X_x X_y$ and assume it is continuous. Recall that stationary means that there exists function $\tilde{R}$ such that $R(x,y) = \tilde{R}(x-y)$. Being a covariance function, the function $\tilde{R}$ is non-negative definite and, since the process $X$ is real-valued, $\tilde{R}$ is symmetric. It is then a consequence of a theorem of S. Bochner (1933) [22] that there exists a symmetric probability measure $\bar{\nu}$ on $\mathbb{R}^d$ such that

$$\tilde{R}(x) = \int e^{i2\pi\lambda\cdot x}\, \bar{\nu}(\mathrm{d}\lambda), \quad x \in \mathbb{R}^d. \tag{2.5.1.10}$$

From the symmetry, we can find a probability measure $\nu$ on $\mathbb{R}_+ \times \mathbb{R}^{d-1}$ such that

$$\tilde{R}(x) = \int \cos(2\pi\lambda \cdot x)\nu(\mathrm{d}\lambda), \quad x \in \mathbb{R}^d. \tag{2.5.1.11}$$

If we consider two independent isonormal Gaussian processes $W_{re}$ and $W_{im}$ over the separable Hilbert space $\mathscr{L}^2(\nu)$, we find that the random function

$$F\colon \mathbb{R}^d \ni x \mapsto W_{re}(\cos(2\pi\langle\cdot, x\rangle)) + W_{im}(\sin(2\pi\langle\cdot, x\rangle)), \tag{2.5.1.12}$$

is a centered Gaussian field with same covariance as $X$. The function $F$ is called the *spectral representation of the stationary field* $X$.

**Example 2.5.1.5** (Gaussian processes with stationary increments)**.** We consider a centered continuous Gaussian process $X = (X_t)_{t\in\mathbb{R}_+}$ with stationary increments such that $X_0 = 0$ and the covariance function $R$ is continuous. Recall that the stationary increments means that:

$$R(s,t) = \mathbb{E}X_sX_t = \mathbb{E}(X_{s+\tau} - X_\tau)(X_{t+\tau} - X_\tau), \quad \text{for all } s,t,\tau \in \mathbb{R}_+. \tag{2.5.1.13}$$

A. N. KOLMOGOROV (1940) [73] (see [74, Paper 42]) gave a spectral representation of the covariance function $R$ in this case. Namely, we have that there exists a symmetric measure $\bar{\nu}$ on $\mathbb{R}$ such that

$$\int (u^2 \wedge 1)\bar{\nu}(\mathrm{d}u) < \infty, \tag{2.5.1.14}$$

and

$$R(s,t) = \int_{-\infty}^{\infty} (\mathrm{e}^{ius} - 1)(\mathrm{e}^{-iut} - 1)\bar{\nu}(\mathrm{d}u), \quad s,t \in \mathbb{R}_+. \tag{2.5.1.15}$$

Again from the symmetry properties of $\bar{\nu}$ we can find a measure $\nu$ on $\mathbb{R}_+$ satisfying

$$\int_{\mathbb{R}_+} (u^2 \wedge 1)\nu(\mathrm{d}u) < \infty \tag{2.5.1.16}$$

and such that

$$R(s,t) = \int_0^{\infty} (\cos(u(s-t)) - \cos(ut) - \cos(us) + 1)\nu(\mathrm{d}u), \quad s,t \in \mathbb{R}_+. \tag{2.5.1.17}$$

Hence taking, as before, two independent isonormal Gaussian processes $W_{re}$ and $W_{im}$ over the separable Hilbert space $\mathscr{L}^2(\nu)$ yields to the *spectral representation of the process with stationary increments* $X$ given by

$$[0,1] \ni t \mapsto W_{re}(\cos(t\cdot) - 1) + W_{im}(\sin(t\cdot)). \tag{2.5.1.18}$$

For the rest of the section, we fix a separable real Hilbert space $\mathscr{H}$ and an isonormal Gaussian process $W$ over $\mathscr{H}$ and $\{h_i\}_{i\in\mathbb{N}}$ is an Hilbert basis. We let $\mathfrak{W} = \sigma(W)$. The goal of this section is to show that $\mathscr{L}^2(\mathfrak{W})$ has a diffusive Itô structure. Isonormal Gaussian processes form the canonical example of such Itô structures and we give a rather detailed construction.

2.5.2. **Hermite chaos.** *Hermite polynomials* play an important role in many areas of mathematics and among the many ways of introducing them, they can be defined recursively by

$$H_0 = 1 \quad \text{and} \quad H_{q+1} = XH_q - H_q'. \tag{2.5.2.1}$$

From this recursion, we deduce two crucial properties of the Hermite polynomials:

(i) For $q \in \mathbb{N}$, $H_q$ solves the two linear differential equations

$$H_q'' - XH_q' = -qH_q; \tag{2.5.2.2}$$
$$H_q' = qH_{q-1}. \tag{2.5.2.3}$$

(ii) For $p$ and $q \in \mathbb{N}$,

$$H_pH_q = \sum_{r=0}^{p\wedge q} \binom{q}{r}\binom{p}{r} r! H_{p+q-2r}. \tag{2.5.2.4}$$

Let $d \in \mathbb{N}$ and $\mathbb{R}_d[X]$ the space of real polynomials of degree $d$. The operator

$$\mathbb{R}_d[X] \ni P \mapsto P'' - XP' \in \mathbb{R}_d[X], \tag{2.5.2.5}$$

is symmetric positive and definite in $\mathscr{L}^2(\gamma)$, where $\gamma$ is the normal law. From (2.5.2.2), we deduce that $\{H_q\}_{q\le d}$ is complete orthogonal system of $\mathbb{R}_d[X]$ in $\mathscr{L}^2(\gamma)$. From (2.5.2.4), we deduce that, with $\mathrm{e}_q = q!^{-1/2}H_q$, $\{e_q\}_{q\le d}$ is in fact an Hilbert basis. By approximation, the renormalized Hermite polynomials $\{e_q\}_{q\in\mathbb{N}}$ form an Hilbert basis of $\mathscr{L}^2(\gamma)$. By a direct computation, we obtain the decomposition

$$\mathrm{e}^{xt-t^2/2} = \sum_{q\in\mathbb{N}} \frac{t^q}{q!} H_q(x). \tag{2.5.2.6}$$

From (2.5.2.3) and a Taylor expansion, we have the following binomial formula for Hermite polynomials

$$H_q(x+y) = 2^{-q/2} \sum_{k=0}^{q} H_k(\sqrt{2}x)H_{q-k}(\sqrt{2}y). \tag{2.5.2.7}$$

We define the associated *Wiener chaos* $\mathscr{C}_q$ as the closure of the linear span of $\{H_q(W(h))\}$ for $h \in \mathscr{H}$ and $|h| = 1$. Plainly, $\mathscr{C}_1 = \mathscr{G}$ contains only Gaussian random variables and $\mathscr{C}_0$ is the space of constants. A *multivariate Hermite polynomial* of degree $q$ is a polynomial $H$ in several variable of the form $H = H_{q_1} \otimes \cdots \otimes H_{q_l}$ with $\sum q_i = q$. Thanks to (2.5.2.7), we have an alternative definition of $\mathscr{C}_q$ as the closure of the space of functions of the form $H(W(h_{i_1}), \dots, W(h_{i_l}))$ for $H$ a multi-variate Hermite polynomial of degree $q$ and $i_1, \dots, i_l \in \mathbb{N}$. We also have this well-known result. For $h \in \mathscr{H}$, we write

$$E(h) = \exp\left(W(h) - \frac{|h|^2}{2}\right) \in \mathscr{L}^2(\mathfrak{W}). \tag{2.5.2.8}$$

**Lemma 2.5.2.1.** *The system $\mathscr{E} = \{E(h)\}_{h\in\mathscr{H}}$ is total in $\mathscr{L}^2(\mathfrak{W})$.*

*Proof.* Take $F_m = \phi(W(h_1), \dots, W(h_m))$ in the orthogonal of $\mathscr{E}$. Then with $h = \sum \lambda_i h_i$, computing $\mathbb{E}F \exp(h)$, we see that the Laplace transform of $\phi$ is $0$. As a consequence of the martingale convergence theorem, every $F \in \mathscr{L}^2(\mathfrak{W})$ can be approximated by a sequence of the form $F_m$ with $m \to \infty$. Finally, the orthogonal of $\mathscr{E}$ in $\mathscr{L}^2(\mathfrak{W})$ is $\{0\}$ and this shows the announced density. □

**Theorem 2.5.2.2.** *We have the chaotic decomposition*

(2.5.2.9)
$$\mathscr{L}^2(\mathfrak{W}) = \bigoplus_{q \in \mathbb{N}} \mathscr{C}_q.$$

*Proof.* The fact the $\mathscr{C}_q$ are mutually orthogonal comes from the fact that the Hermite polynomials form an orthogonal family of $\mathscr{L}^2(\gamma)$ and each of the $W(h)$, $h \in \mathscr{H}$, is indeed Gaussian. Let $\mathscr{C} = \oplus \mathscr{C}_q$. By (2.5.2.6), the closure $\mathscr{C}$ contains the system $\mathscr{E}$. Thus, $\mathscr{C}$ is dense by Lemma 2.5.2.1. □

2.5.3. **Stochastic integrals and Malliavin gradient.** We now introduce *Itô stochastic integrals*. Roughly speaking, they provide an orthogonal basis of $\mathscr{C}_q$. Thinking of $\mathscr{C}_q$ as the space of multi-variate Hermite polynomials of degree $q$ evaluated at a Gaussian vector, it is then very natural to use multi-variate Hermite polynomials. The formal construction is carried out below. Recall that $\{h_i\}_{i\in\mathbb{N}}$ is a Hilbert basis of $\mathscr{H}$. By *multi-index*, we mean a sequence of integers that have only finitely many non-zero terms. For a multi-index $q$, we write $|q| = \sum q_i$ and $q! = \prod q_i!$. Let $q$ be a multi-index and $h_q$ be the element of $\mathscr{H}^{\odot|q|}$ obtained by symmetrization of $\otimes_{i=1}^{\infty} h_i^{\otimes q_i}$. For $p = |q|$, we define

(2.5.3.1)
$$I_p(h_q) = \prod_{i=1}^{\infty} H_{q_i}(W(h_i)).$$

**Theorem 2.5.3.1.** *The map $I_p$ can be extended to a Hilbert isomorphism $I_p \colon \mathscr{H}^{\odot p} \to \mathscr{C}_p$ and the maps $(I_p)_{p\in\mathbb{N}}$ give the Fock space representation for $\mathscr{L}^2(\mathfrak{W})$ over the bosonic Fock space $\mathscr{H}^{\odot}$.*

*Proof.* We set $p \in \mathbb{N}$. The family $\{p!^{-\frac{1}{2}} h_q;\ q \in \mathbb{N}^{\mathbb{N}}, |q| = p\}$ is a Hilbert basis of $\mathscr{H}^{\odot p}$. Since the linear space $\mathscr{C}_p$ contains $I_p(h_q)$, we can extend $I_p$ by linearity. We have to show that $\mathscr{I}_p = \{I_p(p!^{-\frac{1}{2}} h_q);\ q \in \mathbb{N}^{\mathbb{N}}, |q| = p\}$ is a Hilbert basis of $\mathscr{C}_p$. From the product formula (2.5.2.4), we get that

(2.5.3.2)
$$\mathbb{E}(H_r(N)H_r(N')) = r!(\mathbb{E}NN')^r 1_{r=r'}, \quad r, r' \in \mathbb{N}.$$

for $N$ and $N'$ jointly Gaussian, centered and of unit variance. This yields

(2.5.3.3)
$$\mathbb{E} I_p\left(p^{-\frac{1}{2}} h_q\right) I_p\left(p^{-\frac{1}{2}} h_{q'}\right) = 1_{q=q'}, \quad q, q' \in \mathbb{N}^{\mathbb{N}},\ |q| = |q'| = p.$$

To show that $\mathscr{I}_p$ is total, we pick $F$ in the orthogonal of $\mathscr{I}_p$ in $\mathscr{C}_p$. By definition of $\mathscr{C}_p$, $F = H_p(W(h))$ for some $h \in \mathscr{H}$. Writing the coordinates of $h$ in the Hilbert basis $\{h_i\}_{i\in\mathbb{N}}$, we have that $h = \tilde{h} + t_1 h_1$, with $\tilde{h} = \sum_{i\geq 2} t_i h_i$. By definition of the orthogonal and (2.5.3.2), we have that

(2.5.3.4)
$$0 = \mathbb{E} F I_p(h_1^{\otimes p}) = t_1.$$

By recursion, we show that $h = 0$. □

As a consequence of (2.5.2.4), we obtain that if $p$ and $q \in \mathbb{N}$, $h \in \mathscr{H}^{\odot p}$ and $\tilde{h} \in \mathscr{H}^{\odot q}$, then,

$$I_p(h)I_q(\tilde{h}) = \sum_{r=0}^{p\wedge q} \binom{p}{r}\binom{q}{r} r! I_{p+q-2r}(h \otimes_r \tilde{h}), \tag{2.5.3.5}$$

where $h \otimes_r \tilde{h} \in \mathscr{H}^{\otimes p+q-2r}$ is defined in [111, Appendix B.4] but whose explicit expression is irrelevant here. Remark that in the previous formula, we do not need to assume that $I_p(h)I_q(\tilde{h}) \in \mathscr{L}^4(\mathfrak{W})$. Indeed, it is automatically the case due to Meyer's inequalities (see [111, Thm 2.5.5]), that are a consequence of hypercontractivity. Thus, the space $\mathscr{L}^2(\mathfrak{W})$ also have polynomial chaoses. Also, as already observed in the previous proof,

$$\mathbb{E}I_p(h)I_q(\tilde{h}) = q!\langle h, \tilde{h}\rangle_{\mathscr{H}^{\otimes q}} 1_{p=q}. \tag{2.5.3.6}$$

From the relation (2.5.2.3) and (2.5.3.1), we deduce that

$$DI_p(h_q) = \sum_{j=1}^{\infty}\prod_{i=1}^{\infty} H^j_{q_i}(W(h_i))h_j, \quad p \in \mathbb{N},\ q \in \mathbb{N}^{\mathbb{N}},\ |q| = p, \tag{2.5.3.7}$$

where $H^j_{q_i} = H'_{q_i}1_{j=i} + H_{q_i}1_{i\neq j}$. Hence, by approximation we deduce that if $\phi$ is smooth with bounded derivatives and $F = (F_1, \dots, F_d)$ where $F_i \in \mathscr{D}\mathrm{om}\, D$, then $\phi(F) \in \mathscr{D}\mathrm{om}\, D$ and

$$D\phi(F) = \sum_{i=1}^{d} \partial_i\phi(F)DF_i. \tag{2.5.3.8}$$

Thus, the Malliavin derivative $D$ satisfies the chain rule or, equivalently it is a derivation (see Section 2.3.3). Ultimately, we proved that $\mathscr{L}^2(\mathfrak{W})$ has a diffusive Itô structure and that it supports the bosonic Fock space (in particular we have $c_i = 3$ for all $i \in [d]$ in Theorem 2.3.3.5).

## 2.6. A TOY EXAMPLE: THE HYPERCUBE

Let us explore further the concepts introduced before with a simple non-diffusive example. Note that we purposely choose to present the stochastic analysis on the hypercube from the point of view of the heavier formalism of random measures. In this way, the presentation of this section is closer to the one of Section 2.7 about stochastic analysis for point processes and helps us to put the two approaches in a common framework. We consider a sequence $(p_k)_{k\in\mathbb{N}}$ of real number in $(0,1)$ and a family of independent random variables $e = (e_k)_{k\in\mathbb{N}}$ such that, for $k \in \mathbb{N}$,

$$e_k = \begin{cases} e_k^+ = \left(\frac{1-p_k}{p_k}\right)^{1/2}, & \text{with probability } p_k; \\ e_k^- = -\left(\frac{p_k}{1-p_k}\right)^{1/2}, & \text{with probability } 1-p_k. \end{cases} \tag{2.6.1}$$

Note that, for all $k \in \mathbb{N}$, $\mathbb{E}e_k = 0$ and $\mathbb{E}e_k^2 = 1$. The law of $e$ can be regarded as a normalized version of the probability measure on the cube $\{-1,+1\}^\infty$ given by

$$m = \bigotimes_{k\in\mathbb{N}}(p_k\delta_1 + (1-p_k)\delta_{-1}). \tag{2.6.2}$$

We let $\mathfrak{W} = \sigma(e)$ and we will show that $\mathscr{L}^2(\mathfrak{W})$ has a transitive discrete Itô structure. We will make this statement more precise soon but we first prepare some notation. We consider the *normalized hypercube*:

$$E = \bigtimes_k \{e_k^-, e_k^+\}. \tag{2.6.3}$$

Let $k \in \mathbb{N}$. For $x \in E$, we write $T_k x$ for the vector $x$ whose $k$-th coordinate is shifted (that is, $(T_k x)_k = e_k^-$ if $x_k = e_k^+$ and conversely) and whose other coordinates are left unchanged, and we also write $c_k(x) = -(p_k(1-p_k))^{1/2} sign(x_k)$. Note that, for all $k \in \mathbb{N}$, the map $T_k \colon E \to E$ is a bijection and is its own inverse. We also consider the two random measures $\mu$ and $\eta$ on $\mathbb{N}$ defined by

$$\mu = \sum_{k \in \mathbb{N}} e_k \delta_k; \tag{2.6.4}$$

$$\eta = \sum_{k \in \mathbb{N}} e_k^2 \delta_k. \tag{2.6.5}$$

The measure $\mu$ is signed with vanishing expectation, while $\eta$ is non-negative with expectation given by $\nu$, the counting measure on $\mathbb{N}$. For $q \in \mathbb{N}$, the *factorial measure* of $\mu$ is the measure on $\mathbb{N}^q$ defined by

$$\mu^{(q)} = \sum_{k_1,\dots,k_q}^{\neq} e_{k_1} \dots e_{k_q} \delta_{k_1 \dots k_q}, \tag{2.6.6}$$

where the superscript $\neq$ indicates that the summation is over $q$-tuples of pairwise different indices. We write $\mathscr{H} = l^2(\mathbb{N}) = \mathscr{L}^2(\mathbb{N}, 2^{\mathbb{N}}, \nu)$ where $\nu$ is the counting measure. For $q \in \mathbb{N}$, we let $\mathscr{H}^{\circ q}$ be the sub Hilbert space of functions $h \in \mathscr{H}^{\odot q}$ that vanish on the diagonal, that is $h_{i_1 \dots i_q} = 0$, whenever there exists $l$ and $l'$ distinct elements of $[q]$ such that $i_l = i_{l'}$. It is clear that $\mathscr{H}^{\circ} = \oplus_{q \in \mathbb{N}} \mathscr{H}^{\circ q}$ is an abstract Fock space in the sense of Section 2.3. In fact, $\mathscr{H}^{\circ}$ is the vanishing Fock space associated with $\mathscr{H}$ presented in Example 2.3.1.3. We let

$$I_q(h) = \mu^{(q)}(h) = \sum_{k_1 \dots k_q}^{\neq} h_{k_1 \dots k_q} e_{k_1} \dots e_{k_q}, \quad q \in \mathbb{N},\, h \in \mathscr{H}^{\circ q}. \tag{2.6.7}$$

With this notation we can state the main result of this section.

**Theorem 2.6.1.** *The space $\mathscr{L}^2(\mathfrak{W})$ has a transitive discrete Itô structure with a Fock space based on $\mathscr{H}^{\circ}$, stochastic integral maps given by $(I_q)_{q \in \mathbb{N}}$ and a Malliavin derivative $D$ with transitive action map $T$ and with mixing map $c$, that is, for all $k \in \mathbb{N}$ and $F \in \mathscr{D}\mathrm{om}\, D$ with representative $f$, we have that*

$$D_k F = -(p_k(1-p_k))^{1/2} sign(e_k)(f(T_k e) - f(e)). \tag{2.6.8}$$

*The Campbell measure associated with this action is given by $\eta$, as defined in (2.6.5).*

**Remark 10.** As $\mathbb{N}$ is, as usual, endowed with the its discrete $\sigma$-algebra the representative is in fact unique and so the definition of the derivative obviously does not depend on the representative.

In view of the analysis of Section 2.4.3, we obtain as corollaries the following two results. We start with a modified logarithmic Sobolev inequality for the weighted hypercube. As the logarithmic Sobolev inequality is rather geometrical, we state our inequality on $\{-1,+1\}^{\infty}$ rather than $E$. The inequality seems new, and when all the $p_k$ are taken equal this we recover an inequality of S. G. BOBKOV & M. LEDOUX (1998) [18]. Let us first introduce some notations. For $f\colon \{-1,+1\}^{\infty} \to \mathbb{R}$ and $k \in \mathbb{N}$, we write $d_k f(x) = f(-x) - f(x)$. Recall that we have defined the measure $m$ in (2.6.2). We let $\phi(x) = x\log x$ ($x \in R_+$), and if $f \geq 0$ we write,

$$\mathbf{Ent}_m(f) = m(\phi(f)) - \phi(m(f)). \tag{2.6.9}$$

Finally, we have that.

**Theorem 2.6.2.** *Let $f\colon \{-1,+1\}^{\infty} \to \mathbb{R}_+$ and let $m$ be the measure defined in (2.6.2). Then,*

$$\mathbf{Ent}_m(f) \leq \sum_{k\in\mathbb{N}} p_k(1-p_k) m(d_k\phi(f) - \phi'(f) d_k f). \tag{2.6.10}$$

*In particular, we have that*

$$\mathbf{Ent}_m(f) \leq \sum_{k\in\mathbb{N}} p_k(1-p_k) m\left(\frac{|d_k f|^2}{f}\right). \tag{2.6.11}$$

**Remark 11.** When there exists $p \in (0,1)$ such that $p_k = p$ for all $k \in \mathbb{N}$, (2.6.11) was derived in [18, Theorem 1].

*Proof.* Let $S\colon \{-1,1\}^{\infty} \to E$ be the bijection that pairs, for all $k \in \mathbb{N}$, $1$ and $e_k^+$. There exists a function $g\colon E \to \mathbb{R}_+$ such that $f = g \circ S$. Let $G = g(e)$. Since for $k \in \mathbb{N}$, we have that $C_k^2 = p_k(1-p_k)$, by Theorem 2.6.1, we can invoke Theorem 2.4.4.1 that asserts that

$$\mathbf{Ent}_m(f) = \mathcal{H}(G) \leq \sum_{k\in\mathbb{N}} p_k(1-p_k)\mathbb{E}\left[\phi(g(T_k e)) - \phi(g(e)) - \phi'(g(e))(g(T_k e) - g(e))\right]. \tag{2.6.12}$$

In view of the independence of $e$ and since $S^{-1}$ maps $law(e)$ to $m$, we obtain (2.6.10). The inequality (2.6.11) is a consequence of the following observation (see [152]). Let $\psi(u,v) = \phi(u+v) - \phi(u) - \phi'(u)v$ for $u > 0$ and $u+v > 0$, then (2.6.10) can be rewritten as

$$\mathbf{Ent}_m(f) = \sum_{k\in\mathbb{N}} p_k(1-p_k)\psi(f, d_k f). \tag{2.6.13}$$

Also we have that $\psi(u,v) \leq uv^{-2}$. This shows (2.6.11) and concludes the proof. $\square$

We now turn to the fourth moment theorem. Remark that the fourth moment theorem with remainder on the cube was obtained by [38] for univariate functionals. The multivariate version was developed in [153] via exchangeable pair techniques and the author of [153] does not provide explicitly a quantitative bound but a priori a bound similar to ours could be deduced from his work. Note that the method proof are different.

**Theorem 2.6.3.** *Let $(p_1, \dots, p_d) \in \mathbb{N}^d_{>0}$. There exists $c > 0$, such that for all $F = (F_1, \dots, F_d)$ with $F_i \in \mathscr{C}_{p_i} \cap \mathscr{L}^4(\mathfrak{W})$, we have that, with $C = \mathbb{E}F^T F$ Then,*

$$d_1(law(F), \mathbf{N}(0,C)) \leq c \sum_{i=1}^{d} \sqrt{\left(\mathbb{E}F^4 - 3(\mathbb{E}F^2)^2 + \sum_{k\in\mathbb{N}} p_k(1-p_k)\mathbb{E}(D_k^+ F)^4\right) + \left(\sum_{i=1}^{d} \frac{C_{ii}}{p_i}\right)\left(\sum_{i=1}^{d}\left(\sum_{k\in\mathbb{N}} p_k(1-p_k)\mathbb{E}(D_k^+F)^4\right)^{\frac{1}{4}}\right)^2}. \tag{2.6.14}$$

**Remark 12.** The authors of [38] showed that the quartic remainders cannot be removed and related it to the influence of boolean functions.

*Proof.* Direct application of Theorem 2.4.4.3 that we can use thanks to Theorem 2.6.1. □

Most of the rest of the section is devoted to the proof of Theorem 2.6.1.

*Proof of Theorem 2.6.1.* We start by proving the part about stochastic integrals. First of all, for $p$ and $q \in \mathbb{N}$, $g \in \mathscr{H}^{\circ p}$ and $h \in \mathscr{H}^{\circ q}$, we have that

$$I_p(g)I_q(h) = \sum_{j_1\dots j_p}^{\neq} \sum_{k_1\dots k_q}^{\neq} g_{j_1\dots j_p} h_{k_1\dots k_q} e_{j_1}\dots e_{j_p} e_{k_1}\dots e_{k_q}. \tag{2.6.15}$$

Since $(e_k)_{k\in\mathbb{N}}$ is a sequence of centered independent random variables, we see that the quantity $\mathbb{E}I_p(g)I_q(h)$ does not vanish if and only if the indices of $e$ appearing in the first integral are exactly the same as in the second integral. This can happen if and only if $p = q$ and in this case,

$$\mathbb{E}I_p(g)I_p(h) = q! \sum_{j_1\dots j_p}^{\neq} g_{j_1\dots j_p} h_{j_1\dots j_p} \mathbb{E}e_{j_1}^2 \dots \mathbb{E}e_{j_p}^2. \tag{2.6.16}$$

Since $\mathbb{E}e_k^2 = 1$, for all $k \in \mathbb{N}$, this shows the isometry. We now define the chaotic decomposition from the stochastic integrals, namely we set $\mathscr{C}_q$ as the set of all $\mathscr{L}^2(\mathfrak{W})$ random variables of the form $I_q(h)$ for some $h \in \mathscr{H}^{\circ q}$. From the previous computations, we know that the chaos are orthogonal. Let us prove that they form a complete family. For all $q \in \mathbb{N}$, $\mathscr{C}_q$ contains the random variables of the form $e_{i_1}\dots e_{i_q}$ for all $i_1, \dots, i_q \in \mathbb{N}$ pairwise different. We take $F$ in the orthogonal of $\oplus_q \mathscr{C}_q$. We have that

$$0 = \mathbb{E}Fe_1 = \mathbb{E}(\mathbb{E}[F|e_1]e_1) = (p_1(1-p_1))^{1/2}(f(e_1^+) - f(e_1^-)), \tag{2.6.17}$$

where $f$ is the representative of $\mathbb{E}[F|e_1]$. Since $\mathscr{C}_0$ contains the constant functions, we also find that

$$0 = \mathbb{E}F = p_1 f(e_1^+) + (1-p_1)f(e_1^-). \tag{2.6.18}$$

Eventually, we find that $\mathbb{E}[F|e_1] = 0$. Observe that for all $k \in \mathbb{N}$, $Fe_{k+1}$ is orthogonal to $e_1 \dots e_k$. Proceeding recursively, we obtain that $\mathbb{E}[F|e_1, \dots, e_k] = 0$ for all $k \in \mathbb{N}$. Thus, $F = 0$. Therefore, $\mathscr{L}^2(\mathfrak{W})$ indeed has the announced Fock space structure. Let us show the associated Malliavin derivative has transitive action map given by $T_k$ and mixing map $c_k$. Let $h \in \mathscr{H}^{\circ q}$. We fix $F = h_{kk_1\dots k_{q-1}} e_k e_{k_1} \dots e_{k_q}$ with representative $f$. We let $\tilde{e}_k$ be the shift of $e_k$. Observe that

$$-sign(e_k)(p_k(1-p_k))^{1/2}(\tilde{e}_k - e_k) = 1. \tag{2.6.19}$$

This implies that

$$c_k(e)(f(T_k e) - f(e)) = h_{kk_1\dots k_{q-1}} e_{k_1} \dots e_{k_{q-1}}. \tag{2.6.20}$$

So we immediately see that if $I_q(h)$ has representative $f$, then

$$c_k(e)(f(T_k e) - f(e)) = qI_q(h(k,\cdot)) = D_k I_q(h). \tag{2.6.21}$$

This shows the announced representation of the Malliavin derivative $D$. This family is clearly dense in every Sobolev space. We show that the Campbell measure associated to the action map is indeed $\eta$. Let $(u_k)_{k\in\mathbb{N}}$ be a family of functions from $E$ to $\mathbb{R}$. By definition of $\eta$, we have that

$$\mathbb{E}\sum_{k\in\mathbb{N}} u_k(e)\eta(k) = \sum_{k\in\mathbb{N}} \mathbb{E}\left[p_k u_k(e_k^+)\frac{1-p_k}{p_k} + (1-p_k)u_k(e_k^-)\frac{p_k}{1-p_k}\right], \tag{2.6.22}$$

where with a slight abuse of notation we write $e_k^+$ for the vector obtained from $e$ where the $k$-th coordinate is given by the number $e_k^+$ (and respectively for $e_k^-$). Finally,

$$\mathbb{E}\sum_{k\in\mathbb{N}} u_k(e)\eta_k = \mathbb{E}\sum_{k\in\mathbb{N}} u_k(T_k e). \tag{2.6.23}$$

We now show that this structure enjoys the property of polynomial chaos. Let $F = \sum_{q\in\mathbb{N}} I_q(h_q) \in \mathscr{L}^2(\mathfrak{W})$. Fix $q \in \mathbb{N}$ and $k_1, \dots, k_q \in \mathbb{N}$ pairwise different. If $F \in \mathscr{D}\mathrm{om}\, D^q$, by Theorem 2.3.2.2 and by duality, we find that

$$h_q(k_1, \dots, k_q) = \mathbb{E}\left[\prod_{i=1}^{q} e_{k_i} F\right]. \tag{2.6.24}$$

By density of $\mathscr{D}\mathrm{om}\, D^q$ in $\mathscr{L}^2(\mathfrak{W})$, this relation still holds even for $F \in \mathscr{L}^2(\mathfrak{W})$. Let $g \in \mathscr{H}^{\circ p}$ and $h \in \mathscr{H}^{\circ q}$, we write

$$I_p(g)I_q(h) = \sum_{r=0}^{\infty} I_r(h_r). \tag{2.6.25}$$

From the previous identity

$$h_r(k_1, \dots, k_r) = \mathbb{E}\left[I_p(g)I_q(h)\prod_{i=1}^{r} e_{k_i}\right]. \tag{2.6.26}$$

The integral $I_p(g)$ is a linear combination of products of $p$ distinct elements of $(e_k)$ and similarly for $I_q(h)$ in a way that if $r > p+q$, there is necessarily one element of the product that does not appear in either of the two integrals. Hence, by independence and the fact that $e_k$ is centered we find $h_r = 0$. The proof is concluded. □

**Remark 13.** By Tychonoff's theorem the space $\{-1,+1\}^\infty$ is compact. Moreover it is a topological group for the pointwise multiplication known as the *Cantor group*. The Haar measure of this group is given by

$$\frac{1}{2}\bigotimes_{k\in\mathbb{N}}(\delta_1+\delta_{-1}). \tag{2.6.27}$$

For $l\in\mathbb{N}$ and $i_1,\ldots,i_l\in\mathbb{N}$ pairwise disjoint, the mapping,

$$h_{i_i,\ldots,i_k}\colon \{-1,+1\}^\infty \ni e\mapsto e_{i_1}\ldots e_{i_l}, \tag{2.6.28}$$

are the *characters* of this group. Hence, when $p_k=\frac{1}{2}$, for all $k\in\mathbb{N}$, the chaotic representation is a consequence of the Peter-Weyl theorem.

**Remark 14.** The definition of stochastic integrals makes sense for functions that are not necessarily vanishing on the diagonal. If we pick $h$ and $\tilde{h}\in\mathscr{H}^{\odot 2}$ such that, for all $k$ and $j\in\mathbb{N}$, $h_{kj}=1_{k=j=1}$ and $\tilde{h}_{kj}=1_{k=j=2}$, we have that

$$\mathbb{E}I_2(h)I_2(\tilde{h})=\mathbb{E}e_1^2e_2^2=1\neq 0=\nu(h\tilde{h}). \tag{2.6.29}$$

Hence, the isometry property does not hold. This explains why we work with vanishing Fock space $\mathscr{H}^\circ$ rather than the bosonic Fock space $\mathscr{H}^\odot$.

For the reader's convenience, we also give the explicit representation of the Ornstein-Uhlenbeck generator $L$ and the carré du champ obtained from Theorem 2.4.2.3. Note that, with the notations of Section 2.4.3, we have $\tilde{C}_k=-C_k$ and $C_kD_k^+F=\tilde{C}_kD_k^-F=D_kf$, so that

$$LF=\sum_{k\in\mathbb{N}}C_kD_kF(e_k^2+1). \tag{2.6.30}$$

Observe that we have the two relations:

$$e_k^2-1=\frac{1-2p_k}{(p_k(1-p_k))^{1/2}}e_k; \tag{2.6.31}$$

$$sign(e_k)=2(p_k(1-p_k))^{1/2}e_k+2p_k-1. \tag{2.6.32}$$

Combining these two relations, we obtain

(2.6.33)

$$\begin{aligned}-C_k(e_k^2+1)&=(2(p_k(1-p_k))^{1/2}e_k-(1-2p_k))((1-2p_k)e_k+2(p_k(1-p_k))^{1/2})\\&=4p_k(1-p_k)e_k-(1-2p_k)^2e_k+2(p_k(1-p_k))^{1/2}(1-2p_k)(e_k^2-1)\\&=\left[4p_k(1-p_k)-(1-2p_k)^2+2(1-p_k)^2\right]e_k=e_k.\end{aligned}$$

Eventually, we obtain the following closed forms:

$$LF=-\sum_{k\in\mathbb{N}}D_kFe_k=-\mu(DF); \tag{2.6.34}$$

$$\Gamma(F,G)=\frac{1}{2}\sum_{k\in\mathbb{N}}D_kFD_kG(1+e_k^2)=\frac{1}{2}(\nu+\eta)(DFDG). \tag{2.6.35}$$

(Note that the form for $\Gamma$ is obtained directly from Theorem 2.4.2.3 and does not rely on the computations carried out for $L$.)

**Outline.** In this section, we introduce, for all $q \in \mathbb{N}$, $\chi^{(q)}$ the *factorial measure of order* $q$ of the point measure $\chi$. If $\mu$ is a point process, $\mu^{(q)}$ is still a point process. A sufficiently integrable random variable of the form $\mu^{(q)}(h)$ is called a $U$-statistics of order $q$. Under mild assumptions, we prove (Lemma 2.7.2.1) that the linear span of $U$-statistics of all order is dense in $\mathscr{L}^2(\mathfrak{W})$, where $\mathfrak{W} = \sigma(\mu)$. We then study from an abstract and systematic point of view the combinatorial properties of the moments of $U$-statistics in terms of the *factorial moment measure* $\nu_{(q)} = \mathbb{E}\mu^{(q)}$. Hence, we define *stochastic integrals* $I_q \colon \mathscr{L}^2(\nu_{(q)}) \to \mathscr{L}^2(\mathfrak{W})$ as an alternating sums of integrals with respect to the factorial measures and the factorial moment measures as it is traditional for Poisson point processes. Those stochastic integrals can be seen as the orthonormalisation of the $U$-statistics and we show that they satisfy an Itô isometry (Lemma 2.7.2.3). We then discuss the possibility to obtain a Fock space decomposition of the space $\mathscr{L}^2(\mathfrak{W})$. The main difficulty is that a priori $\mathscr{L}^2(\nu_{(q)}) \neq \mathscr{L}^2(\nu^q)$. If it is the case, the family of Hilbert spaces $\mathscr{H}^{\circ q} = \mathscr{L}^2_\sigma(\nu_{(q)})$ gives us the Fock space $\mathscr{H}^\circ = \oplus_q \mathscr{H}^{\circ q}$. In Theorem 2.7.2.5, provided the former space is indeed a Fock space and provided mild regularity assumptions on $\mu$, we show that $\mathscr{L}^2(\mathfrak{W})$ has a pure discrete Itô structure. Examples are then given in the two classical frameworks of Poisson point processes and mixed binomial processes in Section 2.7.3. The construction we described is well-known for Poisson point processes, our contributions consists in adapting this construction to the setting of generic point processes.

### 2.7.1. **Some preliminaries.**

2.7.1.1. *Exponential approximation.* Let us state and prove the following result on approximation of functionals of random measures. It is well-known for Poisson point processes [82, Lem 18.4] but the argument works for generic random measures and is given here. The proof relies on a functional version of the monotone class theorem [35, Thm I.21] that we reproduce here without a proof.

**Theorem 2.7.1.1.** *Let $\mathscr{W}$ be a linear space of bounded random variables that contains* $1$, *that is stable with respect to uniform convergence and stable with respect to the bounded monotone convergence, that is: for all increasing uniformly bounded sequence of positive random variables $(F_n)$ the almost sure limit $F = \lim_n F_n$ belongs to $\mathscr{W}$. Let $\mathscr{G} \subset \mathscr{W}$ be stable under multiplication. Then, the space $\mathscr{W}$ contains all the bounded functions measurable with respect to the $\sigma$-algebra generated by $\mathscr{G}$.*

We will use the following lemma at several places. We first introduce some notations. For a measure $\nu$, we say that a real-valued measurable function $u$ has *$\nu$-bounded support* if $\{|u| > 0\} \in \mathfrak{Z}_\nu$. Given a random measure $\mu$, we define its *Laplace functional*,

$$L_\mu(u) = L(u) = \mathbb{E}\,\mathrm{e}^{-\mu(u)}\,. \tag{2.7.1.1}$$

Remark that this quantity is well-defined for all measurable $u \geq 0$, and that in that case $L(u) \in [0, 1]$ and it is well-defined, though possibly infinite, for all $u \in \mathscr{L}^1(\nu)$, where $\nu = \mathbb{E}\mu$ is the intensity measure of $\mu$. We say that $\mu$ is *locally exponentially integrable* if, for every $B \in \mathfrak{Z}_\nu$, we have that $\mathbb{E}\,\mathrm{e}^{\mu(B)} < \infty$.

**Lemma 2.7.1.2.** *Let $\mu$ be a random measure with $\sigma$-finite intensity measure $\nu$. We define the set $\mathscr{G}$ as the linear span of random variables of the form $\mathrm{e}^{-\mu(u)}$ with $u$ non-negative bounded with $\nu$-bounded support. Let $\mathfrak{W} = \sigma(\mu)$, then $\mathscr{G}$ is dense in $\mathscr{L}^2(\mathfrak{W})$.*

*Proof.* First of all, since, $u \geq 0$ with $\nu$-bounded support, it is clear that $\mathrm{e}^{-\mu(u)}$ is square integrable. Let $\mathscr{W}$ be the bounded functions of the closure of $\mathscr{G}$ in $\mathscr{L}^2(\mathfrak{W})$. If $(F_n) \subset \mathscr{W}$ converges uniformly to some $F \in \mathscr{L}^2(\mathfrak{W})$, then, by dominated convergence $(F_n)$ also converges in $\mathscr{L}^2(\mathfrak{W})$. Thus, $\mathscr{W}$ is stable under uniform convergence. Similarly, by the monotone convergence theorem, the space $\mathscr{W}$ is stable with respect to the bounded monotone convergence. Clearly, $\mathscr{W}$ contains the constant and $\mathscr{G}$ is stable by multiplication. By the previous theorem, the space $\mathscr{W}$ contains all the bounded random variables measurable with respect to the $\sigma$-algebra $\mathfrak{G} = \sigma(\mathscr{G})$. Let us show that $\mathfrak{G} = \mathfrak{W}$. This will conclude the proof as in this case, we will have $\mathscr{W} \supset \mathscr{L}^\infty(\mathfrak{W})$ and $\mathscr{L}^\infty(\mathfrak{W})$ is clearly dense in $\mathscr{L}^2(\mathfrak{W})$. Let $A \in \mathfrak{Z}_\nu$. By definition of $\mathscr{G}$, for all $t > 0$, $t^{-1}(1 - \mathrm{e}^{-t\mu(A)})$ is $\mathfrak{G}$-measurable. By taking the limit, as $t \to 0$, we find that $\mu(A)$ is $\mathfrak{G}$-measurable. Let $A \in \mathfrak{Z}$, since $\nu$ is $\sigma$-finite, we can find an increasing sequence of elements of $\mathfrak{Z}_\nu$ whose union is $A$. Therefore, $\mu(A)$ is $\mathfrak{G}$-measurable. As $\mathfrak{W}$ is generated by random variables of the form $\mu(A)$ for $A \in \mathfrak{Z}$, we obtain that $\mathfrak{G} = \mathfrak{W}$. This concludes the proof. $\square$

2.7.1.2. *Factorial measure.* From this approximation theorem a very natural candidate for the chaos of $\mathscr{L}^2(\mathfrak{W})$ is the space of random variables of the form $\mu^q(h)$ for some $h$ sufficiently integrable. As we will see the choice of the tensor product is combinatorially not wise in case of point processes and has to be replaced with the factorial product that we introduce below. In the following, the random measure $\mu$ is a *point process*, that is a random element of $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$. An element $\chi \in \mathscr{M}_{\bar{\mathbb{N}}}(Z)$ is *proper* if it can be written

$$\chi = \sum_{k=1}^{l} \delta_{x_k}, \quad \text{with } l \in \bar{\mathbb{N}},\ (x_k)_{k=1}^{k=l} \subset Z^l. \tag{2.7.1.2}$$

We say that the point process $\mu$ is *proper* if it is almost surely proper, that is if $\mu$ can be written

$$\mu = \sum_{k=1}^{N} \delta_{X_k}, \tag{2.7.1.3}$$

for some random elements $N \in \bar{\mathbb{N}}$ and $X_k \in Z$. We say that $\mu$ is *distributionally proper* if there exists a proper point process $\mu'$ such that $\mu$ and $\mu'$ have the same distribution. We start by defining a deterministic operation on the elements of $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$ that plays the role of the tensor product.

For $\chi \in \mathscr{M}_{\bar{\mathbb{N}}}(Z)$, we define the *factorial measures* that are defined recursively by $\chi^{(1)} = \chi$ and for $q \in \mathbb{N}$ and $A \in Z^q$ measurable

$$\chi^{(q+1)}(A) = \chi^{(q)} \otimes \chi(A) - \sum_{k=1}^{q} \chi^{(q)}(A_k), \tag{2.7.1.4}$$

where

$$A_k = \{(x_1, \dots, x_q) \in Z^q, \text{ such that } (x_1, \dots, x_q, x_k) \in A\}. \tag{2.7.1.5}$$

This equation has a unique solution and the mapping $\chi \mapsto \chi^{(q)}$ is measurable [82, Prop A.18]. Also if $\chi = \sum_{k=1}^{n} \delta_{x_k}$ is a proper point measure, then we have the closed expression

$$\chi^{(q)} = \sum_{k_1,\dots,k_q \le n}^{\neq} \delta_{(x_{k_1},\dots,x_{k_q})}, \tag{2.7.1.6}$$

where we use the superscript $\neq$ to indicate that the summation is taken over pairwise distinct indices. On the other hand, if $\chi^2$ does not charge the diagonal then $\chi^{(q)} = \chi^q$.

**Lemma 2.7.1.3.** *Let $q \in \mathbb{N}$. Let $\chi$ and $\xi \in \mathscr{M}_{\bar{\mathbb{N}}}(Z)$. Then, for every $f\colon Z^q \to [0,\infty]$,*

$$(\chi+\xi)^{(q)}(f) = \sum_{L \subset [q]} \int_{Z^q} f(z) \chi^{(|L|)}(\mathrm{d}z_L) \xi^{(q-|L|)}(\mathrm{d}z_{[q]\setminus L}), \tag{2.7.1.7}$$

*where both sides of this identity can assume the value $\infty$. If $f$ has an arbitrary sign, the previous equality is still valid, provided both side of the equality are finite, when $f$ is replaced by $|f|$.*

*Proof.* We start by assuming that $\chi$ and $\xi$ are proper. Thus, we let $I, J \subset \mathbb{N}$ and $(x_i)_{i\in I} \subset Z^I$, $(y_j)_{j\in J} \subset Z^J$ such that

$$\chi = \sum_{i\in I} \delta_{x_i} \quad \text{and} \quad \xi = \sum_{j \in J} \delta_{y_j}. \tag{2.7.1.8}$$

Without loss of generality, we can assume that $I = [m]$ and $J = [m+l] \setminus [l]$ for some $m$ and $l \in \mathbb{N}$. Let $K = [m+l]$ and $(z_k)_{k\in K} \subset Z^K$ such that $z_k = x_k$ if $k \in I$ and $z_k = y_k$ if $k \in J$. By definition of the factorial power, we have that

$$(\chi+\xi)^{(q)} = \sum_{k_1,\dots,k_q \le m+l}^{\neq} \delta_{(z_{k_1},\dots,z_{k_q})}. \tag{2.7.1.9}$$

For $k_1, \dots, k_q$ distinct elements of $[m+l]$, we let $L$ be the set of those $l \in [q]$ such that $k_l \in I$. By construction, $(k_l)_{l\in L}$ is a family of distinct elements of $I$ and $(k_l)_{l\notin L}$ is a family of distinct elements of $J$. Reciprocally, a family of distinct elements in $I^p$ and a family of distinct elements in $J^{q-p}$ for $0 \le p \le q$ completely determines a family of distinct element in $K^q$. This proves the claim for $\chi$ and $\xi$ proper.

In general, we can write, by definition of $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$,

$$\chi = \sum_{k=0}^{\infty} \chi_k \quad \text{and} \quad \xi = \sum_{k=0}^{\infty} \xi_k, \tag{2.7.1.10}$$

where for all $k \in \mathbb{N}$, $\chi_k$ and $\xi_k \in \mathscr{M}_{\mathbb{N}}(Z)$. Letting

$$\chi_n = \sum_{k=0}^{n} \chi_k \quad \text{and} \quad \xi_n = \sum_{k=0}^{n} \xi_k, \tag{2.7.1.11}$$

The measure $\chi_n$ and $\xi_n$ are proper and the claim is proved for such measures. Moreover, we have that $\chi_n \uparrow \chi$ and $\xi_n \uparrow \xi$ so that by [82, Proposition A.18], we have that for all $k \in \mathbb{N}$, $\chi_n^{(k)} \uparrow \chi^{(k)}$, $\xi_n^{(k)} \uparrow \xi^{(k)}$ and $(\chi_n+\xi_n)^{(k)} \uparrow (\chi+\xi)^{(k)}$. We thus conclude in the general case. □

We also have the following lemma.

**Lemma 2.7.1.4.** *Let $\chi \in \mathscr{M}_{\bar{\mathbb{N}}}(Z)$. For all $q > \chi(Z)$, we have that $\chi^{(q)} = 0$.*

*Proof.* If $\chi$ is a proper point measure it is immediate. We prove it in the general case. By definition, there exists proper point measures $\chi_k = \sum_{i=1}^{N_k} \delta_{X_i^k} \in \mathscr{M}_{\mathbb{N}}(Z)$ such that

$$\chi = \sum_{k=1}^{\infty} \chi_k. \tag{2.7.1.12}$$

We define $\chi_n = \sum_{k=1}^{n} \chi_k$. Then $\chi_n \uparrow \chi$ and on $\{\chi(Z) < q\}$, we have that $\{\chi_n(Z) < q\}$ and then it is clear from (2.7.1.6) that $\chi_n^{(q)} = 0$. Now we can use [82, Prop A.18], to obtain that $\chi_n^{(q)} \uparrow \chi^{(q)}$ and conclude. $\square$

The *order* of the measure $q$ is defined as

$$|\chi| = \inf \left\{ q \in \mathbb{N}, \text{ such that } \chi^{(q)} = 0 \right\} - 1. \tag{2.7.1.13}$$

Note that from the recursive definition of the factorial measure (2.7.1.4), we see that for all $q > |\chi|$, we have that $\chi^{(q)} = 0$ and from the previous lemma, we have that $|\chi| \leq \chi(Z)$. Also the order of $\chi$ might be infinite and the only measure of order $0$ is the null measure.

2.7.2. **$U$-statistics and stochastic integrals.** We now turn to the case where $\mu$ is a random element of $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$. We define the *factorial moment measure* as $\nu_{(q)} = \mathbb{E}\mu^{(q)}$. Since $\mu^{(1)} = \mu$, we have that $\nu_{(1)} = \nu = \mathbb{E}\mu$, the intensity measure of $\mu$. We will need to further assume that for all $k \in \mathbb{N}$, the measure $\nu_{(k)}$ is $\sigma$-finite and we denote $(A_k^n)_{n\in\mathbb{N}}$ an increasing family of elements of $\mathfrak{Z}_{\nu_{(k)}}$ whose union is $Z^k$. We also assume that $\mu$ is locally exponentially integrable and distributionally proper. The quantity $|\mu|$, that is the order of $\mu$ is a $\bar{\mathbb{N}}$ valued random variable and we denote $|\mu|_\infty$ its essential supremum.

A *$U$-statistics of order $k$* is any random variable of the form $\mu^{(k)}(h) \in \mathscr{L}^2(\mathfrak{W})$ for $h \in \mathscr{L}^1(\nu_{(k)})$. Remark that, a priori, we only have $\mu^{(k)}(h) \in \mathscr{L}^1(\mathfrak{W})$ and we will carry out some moment computations in order to obtain workable condition on $h$ to ensure $\mu^{(k)}(h) \in \mathscr{L}^2(\mathfrak{W})$. However, we can already state the following result. We write $\mathscr{U}_k$ for the linear span of $U$-statistics up to order $k$ and for consistency we write $\mathscr{U}_0$ for the space of constants.

**Lemma 2.7.2.1.** *The linear space $\sum_{k\leq|\mu|_\infty} \mathscr{U}_k$ is dense in $\mathscr{L}^2(\mathfrak{W})$.*

*Proof.* We let $u$ be non-negative and have a $\nu$-bounded support and we let $B = \{u > 0\}$. Assume first that $\mu$ is proper. By the previous lemma, we have that

$$\begin{aligned} e^{-\mu(u)} &= \prod_{X\in\mu\cap B} e^{-u(X)} = \sum_{J\subset[\mu(B)]} \prod_{j\in J} (e^{-u(X_i)} - 1) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \mu^{(k)} \left( (e^{-u} - 1)^{\otimes k} \right). \end{aligned} \tag{2.7.2.1}$$

Note that by Lemma 2.7.1.4, we have in fact that

$$e^{-\mu(u)} = \sum_{k=0}^{|\mu|_\infty} \frac{1}{k!} \mu^{(k)} \left( (e^{-u} - 1)^{\otimes k} \right). \tag{2.7.2.2}$$

Thus, with $h_k = k!^{-1}(\mathrm{e}^{-u} - 1)^{\otimes k}$ the random variable $\mathrm{e}^{-\mu(u)}$ is the almost sure limit, as $n \to \infty$, of the series given by

$$S_n = \sum_{k=0}^{n \wedge |\mu|_\infty} \mu^{(k)}(h_k). \tag{2.7.2.3}$$

We will show at once that, for all $k \in \mathbb{N}$, the random variable $\mu^{(k)}(h_k)$ is a $U$-statistics of order $k$ and that the limit can be taken in $\mathscr{L}^2(\mathfrak{W})$. Note that, for all $n \in \mathbb{N}$,

$$|S_n| \leq \sum_{k=0}^{\mu(B)} \frac{1}{k!} |\mu^{(k)}((\mathrm{e}^{-u} - 1)^{\otimes k})| \leq \sum_{k=0}^{\mu(B)} \frac{1}{k!} \mu^{(k)}(B^k). \tag{2.7.2.4}$$

It is easy to check that

$$\mu^{(k)}(B^k) = \frac{\mu(B)!}{(\mu(B) - k)!}. \tag{2.7.2.5}$$

Hence, we find that

$$|S_n| \leq 2^{\mu(B)}. \tag{2.7.2.6}$$

By the assumption of local exponential integrability the right-hand side belongs to $\mathscr{L}^2(\mathfrak{W})$. This shows that $\mu^{(k)}(h_k)$ is a $U$-statistics and, by Lebesgue dominated convergence theorem, we can take the limit in $\mathscr{L}^2(\mathfrak{W})$. This concludes the proof for proper point processes.

If $\mu$ is only distributionally proper, we consider $\mu'$ a proper point process with same law as $\mu$. We let $\mathfrak{W}' = \sigma(\mu')$ and $\mathscr{U}_k'$ the space of $u$-statistics with respect to $\mu'$ of order $k$. Let $F = f(\mu)$ in the orthogonal of $\sum_{k \leq |\mu|}^{\infty} \mathscr{U}_k$. Then, $F' = f(\mu')$ is in the orthogonal of $\sum_{k \leq |\mu|}^{\infty} \mathscr{U}_k'$. This proves that $f = 0$, $law(\mu')$ almost-everywhere and hence $law(\mu)$ almost-everywhere. Hence $F = 0$. This concludes the proof in the general case. □

2.7.2.1. *Moments computation and isometry.* From this result, the family of spaces $(\mathscr{U}_k)$ forms a natural candidate for the chaotic decomposition of $\mathscr{L}^2(\mathfrak{W})$. As we will see in the following moment computations, these spaces are not orthogonal. These computations are essentially combinatorial and are adapted from [82, Chapter 12] where they are done in a Poisson setting.

Let $p$ and $q \in \mathbb{N}$. A *sub-partition* of $[q]$ is a family of disjoint subsets of $[q]$. Given $\sigma$ a sub-partition of $[q]$, we say that $J \in \sigma$ is a *block* of the sub-partition and we say that $\sigma$ is a partition if the union of the blocks is $[q]$ itself. The number of blocks of the partition (that is the cardinality of $\sigma$) is denoted $|\sigma|$ and the cardinality of the union of the blocks is denoted $||\sigma||$. Remark that a sub-partition $\sigma$ is a partition of $[q]$ if and only if $||\sigma|| = q$. Given $J \subset [p+q]$, we write $J_1 = J \cap [p]$, $J_2 = J \cap ([p+q] \setminus [p])$ and $J^c = [p+q] \setminus J$. We let $\Pi$ (resp. $\Pi^*$) be the set of partitions (resp. sub-partitions) $\sigma$ of $[p+q]$ such that for all $J \in \sigma$, $|J_1| \leq 1$ and $|J_2| \leq 1$. We also denote $\Pi_2$ (resp. $\Pi_2^*$) the set of those partitions (resp. sub-partitions) such that each block contains exactly 2 elements. Observe that if $p \neq q$, $\Pi_2$ is empty. Given a sub-partition $\sigma$ of $[p+q]$ and a functions $h\colon Z^{p+q} \to \mathbb{R}$, we write $h_\sigma$ for the function $Z^{p+q+|\sigma|-||\sigma||} \to \mathbb{R}$ obtained by identifying the arguments whose numbers belong to the same blocks of $\sigma$. Namely, let $\sigma$ be a sub-partition of $[p+q]$ whose blocks are given by $\{I_1, \dots, I_l\}$. The vector $z = (z_1, \dots, z_{p+q-|\sigma|+||\sigma||}) \in Z^{p+q-|\sigma|+||\sigma||}$ is mapped to the vector $z' = (z_{i_1}, z_{i_2}, z_{i_3}, \dots, z_{i_{p+q}})$, where

$$i_l = \inf\{i \in [p+q], \text{ such that } (i,l) \in \sigma\} \wedge l. \tag{2.7.2.7}$$

As usual, we take $\inf \emptyset = \infty$. Then, we define $h_\sigma(z) = h(z')$.

**Example 2.7.2.1.** If $p + q = 5$ and $\sigma = \{\{1,3\},\{2,5\}\}$ then

$$h_\sigma(x,y,z) = h(x,y,x,z,y). \tag{2.7.2.8}$$

We state a lemma that expresses the tensor product of the factorial measures of $\mu$.

**Lemma 2.7.2.2.** *Let the previous notation prevail and let $h_p \in \mathscr{L}^1(\nu_{(p)})$ and $h_q \in \mathscr{L}^1(\nu_{(q)})$. Let $h = h_p \otimes h_q$ and assume that $h_\sigma \in \mathscr{L}^1(\nu_{(p+q+|\sigma|-||\sigma||)})$ for all $\sigma \in \Pi_2^*$. Then, if $\mu$ is proper, we have that*

$$\mu^{(p)}(h_p)\mu^{(q)}(h_q) = \sum_{\sigma \in \Pi_2^*} \mu^{(p+q+|\sigma|-||\sigma||)}(h_\sigma). \tag{2.7.2.9}$$

*If $\mu$ is only distributionally proper, we have that,*

$$\mathbb{E}\mu^{(p)}(h_p)\mu^{(q)}(h_q) = \sum_{\sigma \in \Pi_2^*} \nu_{(p+q+|\sigma|-||\sigma||)}(h_\sigma). \tag{2.7.2.10}$$

*Proof.* By the formula (2.7.1.6) for the factorial measure of a proper point measure, we obtain that

$$\mu^{(p)}(h_p)\mu^{(q)}(h_q) = \sum_{i_1,\dots,i_{p+q} \leq N}^{*} h(X_{i_1}, \dots, X_{i_{p+q}}), \tag{2.7.2.11}$$

where the superscript $*$ indicates that the indices $i_k$ and $i_l$ are distinct whenever $k \neq l$ and $i_k$ and $i_l$ both belong to $[q]$ or both belong to $[p+q] \setminus [q]$. Given a set of such indices $(i_1, \dots, i_{p+q})$, we associate the sub-partition $\sigma$ such that for $j \in [q]$ and $l \in [p+q] \setminus [q]$

$$\{j,l\} \in \sigma \iff i_j = i_l. \tag{2.7.2.12}$$

This proves the first part of the claim. The second part is immediately obtained by taking expectation, if $\mu$ is proper. If $\mu$ is only distributionally proper, as the second part of the claim concerns only the distribution of $\mu$, we can work with a proper version of $\mu$ to conclude. □

We now introduce the fundamental objects of this section: the *stochastic integrals* that can be thought as the orthogonalization of the $U$-statistics. For every $q \leq |\mu|_\infty$, we would like to define the stochastic integral of a kernel $h$ as

$$I_q(h) = \sum_{J \subset [q]} (-1)^{q-|J|} \int h(x)\mu^{(|J|)}(\mathrm{d}x_J)\nu_{(q-|J|)}(\mathrm{d}x_{[q]\setminus J}). \tag{2.7.2.13}$$

Compare this definition with the one we gave on the cube in Section 2.6. Note that, by Fubini theorem and the definition of $\nu_{(k)}$, a sufficient (but cumbersome) condition for the previous quantity to be well-defined is that $h \in \mathscr{L}^1(\nu_{(k)} \otimes \nu_{(q-k)})$ for all $k \in [q]$. We say that kernels verifying such a criterion are *stochastically integrable*.

**Lemma 2.7.2.3.** *Let $p$ and $q \in \mathbb{N}$. Let $h_p \in \mathscr{L}^2(\nu_{(p)})$ and $h_q \in \mathscr{L}^2(\nu_{(q)})$ be stochastically integrable. Then,*

$$\mathbb{E}I_p(h_p)I_q(h_q) = 1_{p=q} \sum_{\sigma \in \Sigma_p} \nu_{(p)}(h_p h_q \circ \sigma), \tag{2.7.2.14}$$

*where $\Sigma_p$ is the set of permutations of $[p]$ and $h \circ \sigma$ is the function obtained by permuting the argument according to $\sigma$.*

*Proof.* From Lemma 2.7.2.2, we find that

$$\mathbb{E}I_p(h_p)I_q(h_q) = \sum_{J \subset [p+q]} (-1)^{|J^c|} \sum_{\sigma \in \Pi_2^*, \sigma \subset J} \nu_{(p+q+|\sigma|-||\sigma||)}(h_\sigma), \tag{2.7.2.15}$$

where the notation $\sigma \subset J$ means that for all $I \in \sigma$, $I \subset J$. If $\sigma$ is a partition then $\sum_{\sigma \subset J} (-1)^{p+q-|J|} = 1$ and the sum vanishes otherwise. Recall that for a partition of $[p+q]$, we have that $||\sigma|| = p+q$. Hence, inverting the order of summation we find that

$$\mathbb{E}I_p(h_p)I_q(h_q) = \sum_{\sigma \in \Pi_2} \nu_{(|\sigma|)}(h_\sigma). \tag{2.7.2.16}$$

As already noticed if $p \neq q$ then $\Pi_2$ is empty and the sum vanishes. If $p = q$ then the blocks of $\sigma \in \Pi_2$ are pairs $(l, k)$ with $l \in [p]$ and $k \in [2p] \setminus [p]$ each of them appearing only once. To such $\sigma$ we can associate a permutation of $[p]$ (still denoted $\sigma$) such that $\sigma(l) = k - p$. It is clear that this identification is one-to-one so that we identify $\Pi_2$ with $\Sigma_p$ the set or permutations of $[p]$. Finally, we find that

$$\mathbb{E}I_p(h_p)I_q(h_q) = 1_{p=q} \sum_{\sigma \in \Sigma_p} \nu_{(p)}(h_p h_q \circ \sigma). \tag{2.7.2.17}$$

The proof is complete. □

2.7.2.2. *Fock space representation.* Let $\mathscr{H}^{\circ q}$ be the symmetric functions of $\mathscr{L}^2(\nu_{(q)})$ and $\mathscr{H}^{\circ 0} = \mathbb{R}$. We will now extend the mapping $I_q$ on $\mathscr{H}^{\circ q}$. We let $\mathscr{H}_q$ be the set of bounded elements $h \in \mathscr{H}^{\circ q}$ such that for all $k \in [q]$, $\nu_{(k)} \otimes \nu_{(q-k)}(h \neq 0) < \infty$. Plainly, the elements of $\mathscr{H}_q$ are stochastically integrable. We recall that for all $k \in \mathbb{N}$, the family $(A_k^n)_{n \in \mathbb{N}}$ denote an increasing family of elements of $\mathfrak{Z}_{\nu_{(k)}}$ whose union is $Z^k$. For $h \in \mathscr{H}^{\circ q}$, letting

$$h_n = h 1_{|h| \leq n} \prod_{k \in [q]} (1_{A_n^k} \odot 1_{A_n^{q-k}}), \tag{2.7.2.18}$$

shows that $\mathscr{H}_q$ is dense in $\mathscr{H}^{\circ q}$. By density, we can extend the stochastic integrals to $\mathscr{H}^{\circ q}$.

**Lemma 2.7.2.4.** *There exists a linear isometry $I_q \colon \mathscr{H}^{\circ q} \to \mathscr{L}^2(\mathfrak{W})$ such that for all $h \in \mathscr{H}^{\circ q}$ and stochastically integrable the formula (2.7.2.13) holds.*

*Proof.* The formula (2.7.2.13) define a linear isometry on the dense subset $\mathscr{H}_q$ of $\mathscr{H}^{\circ q}$. We can extend this mapping by density. □

We say that the factorial moments of $\mu$ are *compatible with a Fock space structure* when, for all $q \in \mathbb{N}$, $\mathscr{L}^2_\sigma(\nu_{(q)})$ is a sub-Hilbert space of $\mathscr{L}^2_\sigma(\nu^q)$ (as explained in Section 2.3, we recall the topologies on the two spaces are the same but the inner products can a priori be different on each space, that is the inner products can differ by positive constants), and when it is compatible with the restriction in the sense of Section 2.3. In other words, writing $\mathscr{H}^{\circ q}$ for $\mathscr{L}^2_\sigma(\nu_{(q)})$ ($q \in \mathbb{N}$) and considering $\mathscr{H}^\circ = \oplus_{q \le |\mu|_\infty} \mathscr{H}^{\circ q}$, we require that $\mathscr{H}^\circ$ is an abstract Fock space. The following theorem is the main result of this section and extend the famous Wiener-Itô decomposition for functionals of a Poisson point process to other point processes.

**Theorem 2.7.2.5.** *Let $\mu$ be a locally exponentially integrable and distributionally proper point process whose factorial moment measures $\{\nu_{(q)};\ q \in \mathbb{N}\}$ are $\sigma$-finite and compatible with a Fock space structure. Let $\mathfrak{W} = \sigma(\mu)$. Then $\mathscr{L}^2(\mathfrak{W})$ has a Fock space structure based on $\mathscr{H}^\circ = \oplus_q \mathscr{H}^{\circ q}$, where $\mathscr{H}^{\circ q} = \mathscr{L}^2_\sigma(\nu_{(q)})$, for all $q \in \mathbb{N}$. The stochastic integral maps are given by the family of maps $\{I_q;\ q \in \mathbb{N}\}$ constructed in Lemma 2.7.2.4. We consider the transitive action such that, for all $z \in Z$, the map $T_z$ acts on $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$ by $T_z\chi = \chi + \delta_z$ and we assume that the Campbell measure $\eta$ defined in (2.4.2.21) exists and has $\sigma$-finite second moment measure. Then, the Malliavin derivative is representable by the transitive action $T$ and the space $\mathscr{L}^2(\mathfrak{W})$ has a pure discrete Itô structure. Also, for all $F \in \mathscr{L}^2(\mathfrak{W})$,*

$$F = \mathbb{E}F + \sum_{q \in \mathbb{N}_{>0}} I_q(h_q), \tag{2.7.2.19}$$

*where*

$$h_q(z_1, \ldots, z_q) = \frac{1}{q!} \mathbb{E} D^{+q}_{z_1 \ldots z_q} F. \tag{2.7.2.20}$$

**Remark 15.** In the theory of point processes, it is customary to define the *reduced Campbell measure* associated with a point process $\mu$ as the measure $C$ on $Z \times \mathscr{M}_{\bar{\mathbb{N}}}(Z)$ defined by

$$C(A \times B) = \mathbb{E} \int_A 1_B(\mu - \delta_x)\mu(\mathrm{d}x). \tag{2.7.2.21}$$

The Campbell measure $\eta$ associated with the transitive action $T_z$ in the sense of Section 2.4.2 exists if and only if there exists a probability measure $\Pi \in \mathscr{P}(\mathscr{M}_{\bar{\mathbb{N}}}(Z))$, such that $C = \nu \otimes \Pi$ and, in this case, the law of $\eta$ is given by $\Pi$. This justifies our terminology.

*Proof.* The fact that the mappings $\{I_q;\ q \in \mathbb{N}\}$ are the stochastic integral maps associated with the Fock space structure $\mathscr{H}^\circ$ comes from Lemmas 2.7.2.1 and 2.7.2.4 and the assumption on the compatibility of the moment measures with a Fock space structure. Let $q \in \mathbb{N}$. Pick $h \in \mathscr{H}_q$, then by Lemma 2.7.1.3, we have that

$$D^+_z I_q(h) = qI_{q-1}(h(z, \cdot)) = D_z I_q(h), \quad z \in Z, \tag{2.7.2.22}$$

where the operator $D^+_z$ is associated with the transitive action $T_z$ according to the notations of Section 2.4.2, that is for $F = f(\mu)$, we have $D^+_z F = f(\mu + \delta_z) - f(\mu)$, and $D$ is the Malliavin derivative associated with the Fock space structure according to the notations of Section 2.3. Let $h \in \mathscr{H}^{\circ q}$ and let $(h_n) \subset \mathscr{H}_q$ converging to $h$ defined by (2.7.2.18). It is clear that $h_n(z,\cdot) \uparrow h(z,\cdot)$. Since by assumption $h(z,\cdot) \in \mathscr{H}^{\circ(q-1)}$, by monotone convergence, we find that $h_n(z,\cdot) \xrightarrow[n\to\infty]{\mathscr{H}^{\circ(q-1)}} h_{(z,\cdot)}$. Hence, by continuity of $I_q$,

$$D_z^+ I_q(h_n) \xrightarrow[n\to\infty]{\mathscr{L}^2(\mathfrak{W})} qI_{q-1}(h(z,\cdot)) = D_z I_q(h). \tag{2.7.2.23}$$

On the other hand, considering $f_n$ a representative of $I_q(h_n)$ and $f$ a representative of $I_q(h)$, for $B \in \mathfrak{Z}_\nu$, we have, by (2.4.2.21)

$$\mathbb{E}\int_B |f_n(\mu+\delta_z) - f(\mu+\delta_z)|\nu(\mathrm{d}z) = \mathbb{E}\eta(B)|f_n(\mu) - f(\mu)|. \tag{2.7.2.24}$$

Since $\nu = \mathbb{E}\nu$ and $\nu_{(2)} = \mathbb{E}\mu^{(2)}$ are assumed to be $\sigma$-finite we have that $\mathbb{E}\eta(B)^2 < \infty$ and by the Cauchy-Schwarz inequality, we obtain that the right-hand side of the previous equation vanishes as $n \to \infty$. In particular this implies that for all $B \in \mathfrak{Z}_\nu$ and for $\nu$ almost every $z \in B$,

$$D_z^+ I_q(h_n)1_B \xrightarrow[n\to\infty]{\mathscr{L}^1(\mathfrak{W})} D_z^+ I_q(h)1_B. \tag{2.7.2.25}$$

Combining this relation with (2.7.2.23) and using that the $\mathscr{L}^1(\mathfrak{W})$ topology is coarser than the one of $\mathscr{L}^2(\mathfrak{W})$ and that this topology is separated, we find that for all $B \in \mathfrak{Z}_\nu$ for almost every $z \in B$,

$$1_B D_z^+ I_q(h) = 1_B D_z I_q(h). \tag{2.7.2.26}$$

As $\nu$ is $\sigma$-finite, this shows that for $\nu$-almost every $z \in Z$, $D_z^+ I_q(h) = D_z I_q(h)$. This proves the announced representation of the Malliavin derivative $D$. The representation of the kernels comes from Theorem 2.3.2.2. The fact that $\mathscr{L}^2(\mathfrak{W})$ has polynomial chaoses comes from Lemma 2.4.2.1. This concludes the proof. □

In the setting of the previous theorem we can readily obtain our modified logarithmic Sobolev inequality Theorem 2.4.4.1; our Stein inequality Theorem 2.4.4.2; and our exact fourth moment theorem Theorem 2.4.4.6. In the following section we explicit those results for Poisson point processes and mixed binomials processes.

### 2.7.3. **Binomial processes and Poisson processes.**

2.7.3.1. *Binomial process.* We consider the proper point process

$$\mu = \sum_{i=1}^{N} \delta_{X_i}, \tag{2.7.3.1}$$

where $(X_i)$ is a sequence of independent and identically distributed random variables with common law $p$ on $Z$ and $N$ is a $\mathbb{N}$-valued random variable independent of $(X_i)$. The random measure $\mu$ is called a *binomial process with sampling distribution $p$ and sample size $N$* or simply *mixed binomial process*. We let

$$\lambda_q = q!\mathbb{E}\binom{N}{q} = \mathbb{E}(N(N-1)\dots(N-q+1))_+, \tag{2.7.3.2}$$

where we used the convention that $\binom{n}{q} = 0$ for $q > n$. We assume $\lambda_q < \infty$ for all $q \in \mathbb{N}$. Since $q!\binom{N}{q}$ counts the number of ordered $q$-tuples of distinct elements in $[N]$ and that $N$ and the point are independent, we find that the factorial moment measure is given by $\nu_{(q)} = \mathbb{E}\mu^{(q)} = \lambda_q p^q = \lambda_q \lambda_1^{-q} \nu^q$.

**Remark 16.** It is a well-known fact that when $N$ follows a Poisson law with mean $\lambda$ then $\lambda_q = \lambda^q$ and we recover that for Poisson point processes $\eta$ with finite intensity $\nu$, $\mathbb{E}\eta^{(q)} = \nu^q$.

**Theorem 2.7.3.1.** *Let the previous notations prevail. Then, $\mu$ satisfies the assumptions of Theorem 2.7.2.5.*

*Proof.* First of all, by the finiteness of the $\lambda_q$ ($q \in \mathbb{N}$), the random variable $\mu(Z) = N$ is exponentially integrable. In particular $\mu$ is locally exponentially integrable. Let $q \in \mathbb{N}$. The measure $\nu_{(q)}$ is finite and in particular it is $\sigma$-finite. Moreover, since the $\nu_{(q)} = \lambda_q \lambda_1^{-q} \nu^q$, it is clear that the factorial moment measures are compatible with a Fock space structure. In fact, we are in the setting of the mixed bosonic Fock space where the bosonic constants are given for all $q \in \mathbb{N}$ by $\lambda_q \lambda_1^{-q}$. Let us compute the Campbell measure associated to the action $T_z \chi = \chi + \delta_z$. Let $A \in \mathfrak{Z}$ and $B \subset \mathscr{M}_{\bar{\mathbb{N}}}(Z)$ measurable. We consider the reduced Campbell $C$ and by independence we can write

$$
\begin{aligned}
C(A \times B) &= \mathbb{E} \int_A 1_B(\mu - \delta_x)\mu(\mathrm{d}x) \\
&= \sum_{k=0}^{\infty} \mathbb{P}(N = k) \sum_{i=1}^{k} \mathbb{E} 1_{\{X_i \in A\}} 1_B \left( \sum_{j \neq i} \delta_{X_j} \right) \\
&= p(A) \sum_{k \in \mathbb{N}} \mathbb{P}(N = k) k \mathbb{P} \left( \sum_{j=1}^{k-1} \delta_{X_j} \in B \right).
\end{aligned} \tag{2.7.3.3}
$$

Thus $C = \nu \otimes \Pi$ where $\Pi$ is the probability measure given by

$$
\Pi(B) = \frac{1}{\lambda_1} \sum_{k \in \mathbb{N}} \mathbb{P}(N = k) k \mathbb{P} \left( \sum_{j=1}^{k-1} \delta_{X_j} \in B \right). \tag{2.7.3.4}
$$

We let $\eta \sim \Pi$, then $\eta$ is the Campbell measure of the action $T_z$. Let $A \in \mathfrak{Z}$. We have that

$$
\begin{aligned}
\mathbb{E}\eta^{(2)}(A) &= \frac{1}{\lambda_1} \sum_{k \in \mathbb{N}} k \mathbb{P}(N = k) \mathbb{E} \sum_{i,j \leq k-1}^{\neq} 1_A(X_i, X_j) \\
&= 3 \frac{\lambda_2}{\lambda_1^2} p^2(A) < \infty.
\end{aligned} \tag{2.7.3.5}
$$

Hence, the second factorial moment of $\eta$ is finite and a fortiori $\sigma$-finite. This concludes the proof. □

In the remainder of the section, we study $U$-statistics by writing their chaotic decomposition with respect to multiple stochastic integrals. When the size of the sampling $N$ is deterministic, we link this chaotic decomposition with the Hoeffding decomposition. Since the binomial process has a finite intensity measure $\nu$, we have that

$\mathscr{L}^2(\nu) \subset \mathscr{L}^1(\nu)$ and the representation (2.7.2.13) always hold. The stochastic integrals were defined as an alternating sum of $U$-statistics. This definition can be inverted. Let $h \in \mathscr{L}^2_\sigma(p^q)$. From Lemma 2.7.1.3, we deduce that

$$D_{z_1,\dots,z_k}\mu^{(q)}(h) = \frac{q!}{(q-k)!}\mu^{(q-k)}(h). \tag{2.7.3.6}$$

Hence by Theorem 2.3.2.2, we find that

$$\mu^{(q)}(h) = \sum_{k=0}^{q} I_k(h_k), \tag{2.7.3.7}$$

where

$$h_k(z_1,\dots,z_k) = \binom{q}{k}\lambda_{q-k}\int h(z_1,\dots,z_k,y_1,\dots,y_{q-k})p^{q-k}(\mathrm{d}y). \tag{2.7.3.8}$$

Finally, we obtain that

$$\mu^{(q)}(h) = \sum_{k=0}^{q}\binom{q}{k}\sum_{l=0}^{k}\binom{k}{l}(-1)^{k-l}\left(\mu^{(l)}\otimes\lambda_{k-l}p^{k-l}\otimes\lambda_{q-k}p^{q-k}\right)(h). \tag{2.7.3.9}$$

When $N = q$ is deterministic, this expression becomes

$$\mu^{(q)}(h) = \sum_{k=0}^{q}\binom{q}{k}\sum_{l=0}^{k}\binom{k}{l}(-1)^{k-l}\frac{q!}{k!}\frac{q!}{(q-k-l)!}\left(\mu^{(l)}\otimes p^{q-l}\right)(h). \tag{2.7.3.10}$$

Since $h$ is symmetric and $(X_i)$ is independent and identically distributed, with $F = h(X_1,\dots,X_q)$, we find that

$$\frac{1}{l!}\mu^{(l)}\otimes p^{q-l}(h) = \sum_{L\subset[q],|L|=l}\mathbb{E}[F|\mathfrak{F}_L], \quad l\in[q], \tag{2.7.3.11}$$

where for $J \subset [q]$, we define $\mathfrak{F}_J = \sigma(X_j, j\in J)$. Simplifying the expression of the binomial coefficients yields

$$\begin{aligned} F = \frac{1}{q!}\mu^{(q)}(h) &= \sum_{k=0}^{q}\binom{q}{k}\sum_{l=0}^{k}(-1)^{k-l}\binom{q}{k-l}\sum_{L\subset[q],|L|=l}\mathbb{E}[F|X_L] \\ &= \sum_{K\subset[q]}\sum_{L\subset K}(-1)^{|K|-|L|}\mathbb{E}[F|\mathfrak{F}_L]. \end{aligned} \tag{2.7.3.12}$$

The previous expression is nothing but an orthogonal decomposition used by W. HOEFFDING (1961) [65] to obtain sufficient conditions for the asymptotic normality of general $U$-statistics. We do not know yet if such conditions can be recovered from our Theorem 2.4.4.3. Also note that in this case the bosonic constants are given by $\tilde{\lambda}_q = \frac{n!}{(n-q)!}n^{-q}$ and we can easily check that $\tilde{\lambda}_{2q} \leq \tilde{q}^2$ and we cannot use Theorem 2.4.4.6. Note that, Poisson $U$-statistics were considered by M. REITZNER & M. SCHULTE (2013) [133] using Stein's inequality for Poisson point processes established by G. PECCATI, J. L. SOLÉ, M. S. TAQQU & F. UTZET (2010) [121]. This inequality corresponds to Theorem 2.4.4.2 in our generalized setting and holds for every binomial processes. It seems natural to conjecture that the analysis of [133] will also apply in this case. We are currently investigating this question.

2.7.3.2. *Poisson process.* The Poisson processes form a family of point processes with a non-empty intersection with mixed binomial processes. We say that $\eta$ is a *Poisson point process with intensity measure* $\nu$ if it is a random variable with value in $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$ such that for all $A_1, \ldots, A_l$ pairwise disjoint elements of $\mathscr{Z}_\nu$ the random vector $(\eta(A_1), \ldots, \eta(A_l))$ is a vector of independent Poisson random variables with mean $(\nu(A_1), \ldots, \nu(A_l))$. If such a process exists, its law will be denoted $\Pi_\nu$. Existence of Poisson point processes with arbitrary reference measure is not known. We list some examples below.

**Example 2.7.3.1** (Poisson random variable)**.** Let $Z = \{0\}$. Every measure on $Z$ can be represented its total mass. Let $\lambda > 0$. Then $\eta = N\delta_0$, where $N$ is a Poisson random variable of mean $\lambda$ is a Poisson point process with intensity $\lambda\delta_0$.

**Example 2.7.3.2** (Finite intensity measure)**.** Let $\nu$ be a measure on $Z$ such that $\nu(Z) = \lambda < \infty$. We let $N$ be a Poisson random variable with $\mathbb{E}N = \lambda$ and $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent and identically distributed random variables distributed according to $\frac{\nu}{\lambda}$ and independent of $N$. We set

$$\eta(f) = \sum_{i=1}^{N} f(X_i), \quad f \in \mathscr{L}^0(Z). \tag{2.7.3.13}$$

In other words, we have

$$\eta = \sum_{i=1}^{N} \delta_{X_i}. \tag{2.7.3.14}$$

Then $\eta$ is a Poisson point process with intensity measure $\nu$.

**Example 2.7.3.3** ($s$-finite intensity measure)**.** We say that $\nu$ is $s$-finite if $\nu = \sum_{n\in\mathbb{N}} \nu_n$, where the $\nu_n$'s are finite measures. Let $\eta_n$ be independent Poisson point processes with respective intensity measure $\nu_n$. We let

$$\eta = \sum_{n\in\mathbb{N}} \eta_n. \tag{2.7.3.15}$$

Then, $\eta$ is a Poisson point process with intensity measure $\nu$. Observe that every $\sigma$-finite is a $s$-finite measure.

We have the important Mecke theorem.

**Theorem 2.7.3.2** ([82, Theorems 4.1 & 4.4])**.** *Let $\nu$ be a $s$-finite measure and $\eta$ be a point process. The following are equivalent:*

*(i) $\eta$ is a Poisson point process with intensity measure $\nu$;*

*(ii) for all $q \in \mathbb{N}$ and for all measurable $\phi\colon \mathscr{M}_{\bar{\mathbb{N}}}(Z) \times Z^q \to [0, \infty]$, we have that*

$$\mathbb{E} \int \phi(\eta, z_1, \ldots, z_q) \eta^{(q)}(\mathrm{d}z) = \mathbb{E} \int \phi\left(\eta + \sum_{j=1}^{q} \delta_{z_j}, z_1, \ldots, z_q\right) \nu^q(\mathrm{d}z). \tag{2.7.3.16}$$

*(iii) for all measurable $\phi\colon \mathscr{M}_{\bar{\mathbb{N}}}(Z) \times Z \to [0, \infty]$, we have that*

$$\mathbb{E} \int \phi(\eta, z)\eta(\mathrm{d}z) = \mathbb{E} \int \phi\left(\eta + \delta_z, z_1\right) \nu(\mathrm{d}z). \tag{2.7.3.17}$$

**Remark 17.** Even though, we can construct Poisson point processes for intensity measures more general than $\sigma$-finite, the construction of stochastic integrals via the approximation argument of Section 2.7 relies critically on the $\sigma$-finiteness of the intensity measure and of the factorial moment measures.

**Remark 18.** The Mecke theorem implies that the Poisson point process is the only point process with $s$-finite intensity measure that is its own Campbell measure. This is why we denote a Poisson point process by the symbol $\eta$ rather than $\mu$.

*Idea of proof.* In the case of finite intensity measure, we have the mixed binomial representation of $\eta$ and we compute the reduced Campbell measure as in the previous section. Then the proof reduces to the well-known fact that, if $N$ is a Poisson random variable with mean $\lambda > 0$, then $\mathbb{E}Nf(N) = \lambda\mathbb{E}f(N+1)$. The general case is done by adding independent Poisson point processes of finite intensity measure. $\square$

**Theorem 2.7.3.3.** *Let $\eta$ be a Poisson point process with $\sigma$-finite intensity measure $\nu$. Then, $\eta$ satisfies the assumptions of Theorem 2.7.2.5. Moreover, with $\mathfrak{W} = \sigma(\eta)$, $\mathscr{L}^2(\mathfrak{W})$ supports the bosonic Fock space and in particular the bosonic constants are all equal to* $1$.

**Remark 19.** Observe that when the intensity measure is finite we also are in the setting of the previous section on binomial processes. So that this theorem is non-trivial only for $\nu(Z) = \infty$.

*Proof.* It is a well-known fact [82, Cor 3.7] of the theory of Poisson processes that every Poisson point process with $s$-finite intensity measure is distributionally proper. As a consequence of Theorem 2.7.3.2 we obtain that $\mathbb{E}\eta^{(q)} = \nu^q$ and that $\eta$ is its own Campbell measure. This proves that the factorial moment measure of any order is $\sigma$-finite, that the $\mathscr{L}^2(\mathfrak{W})$ supports the bosonic Fock space and that the Campbell measure has a $\sigma$-finite second factorial moment measure. Hence it is easy to see that any Poisson point process with $\sigma$-finite intensity measure enters the framework of our Theorem 2.7.2.5. $\square$

## 2.8. Bibliographical and historical comments

2.8.1. **Chaotic decomposition and stochastic integrals.** The notion of chaotic decomposition was originally discussed in the seminal paper of N. Wiener (1938) [150] on homogeneous chaoses. At that time, Wiener only considered decompositions with respect to time-indexed processes such as Brownian motions or Poisson processes on the half-line. Moreover, the proposed decomposition was not orthogonal. The orthogonality condition comes from the work of K. Itô (1951) [67] in a Gaussian setting and also from the work of K. Itô (1956) [68] for Poisson point processes. In his two works, Itô worked in a context of random measures (though, I do not know if he was the first to consider it), invented the concept of multiple stochastic integrals that is still used nowadays and already pointed out the link between Gaussian chaoses and Hermite polynomials. The fact that Poisson chaoses can be related to polynomials comes from an observation of H. Ogura (1972) [117] on Charlier polynomials. Concerning Charlier polynomials and Poisson integrals, see also D. Surgailis (1984) [143]. In Section 2.5, our presentation of Gaussian processes from the point of view of isonormal

Gaussian processes (more general than Gaussian random measures) is due to the fundamental contribution of R. M. DUDLEY (1967) [42]. During the preparation of this work, I constantly used the two recent references by D. NUALART (2009) [115, Chapter 1] and I. NOURDIN & G. PECCATI (2012) [111, Chapter 2]. For the toy model of Section 2.6, I consulted the recent reference of C. DÖBLER & K. KROKOWSKI (2017) [38]. For the construction I gave concerning point processes in Section 2.7, I was inspired by the work G. LAST & M. D. PENROSE (2011) [83] that carries a similar construction for Poisson point processes. I also used the subsequent contribution of G. LAST (2016) [80] and the recent monograph [82]. Observe that for a binomial process with deterministic sampling size $n$, W. HOEFFDING (1961) [65] gave a celebrated orthogonal decomposition for a $U$-statistics. As we explained, this decomposition is linked to stochastic integrals for binomial processes. To my knowledge, this decomposition was never put in the framework of stochastic integrals. G.-C. ROTA & T. C. WALLSTROM (1997) [136] initiated a systematic study of stochastic integrals from the combinatorial aspect was initiatied. This analysis was exploited in the comprehensive monograph of G. PECCATI & M. S. TAQQU (2011) [124].

2.8.2. **Malliavin calculus and Fock space.** Malliavin calculus was developed by P. MALLIAVIN (1978) [92] only with respect to a Brownian motion. Malliavin's goal was to give an alternative proof of a celebrated theorem of L. HÖRMANDER (1967) [66] about the regularizing property of some second-order differential operator. Those operators are the generators of a solution of an explicit stochastic differential equation and Malliavin used his calculus to study solutions to these equations. Reading Malliavin original paper is quite demanding and we recommend the lecture notes of M. HAIRER (2016) [62] for a self-contained presentation of the subject. However, the concept of Fock space was formally introduced much earlier in the work of V. FOCK (1932) [48] following the seminal work of P. DIRAC (1927) [37] on second quantization. In this theory, each $\mathscr{H}^{\odot q}$ represents a system with $q$ bosons, so the Malliavin derivative is called the creation operator while its adjoint is called the annihilation operator and the Ornstein-Uhlenbeck generator the number operator.

The idea of constructing a Markov semi-group out of a chaotic decomposition (or rather a Fock space) is also very classical in quantum field theory and goes back (at least) to the seminal paper of E. NELSON (1973) [106] where he constructed the free Markov field over the Gaussian space, that corresponds, in our terminology, to the Ornstein-Uhlenbeck semi-group in a Gaussian setting. Namely, given any Markov dynamic $p_t$ on the Hilbert space $\mathscr{H}$, we can "lift" this dynamic to $\mathscr{H}^{\odot}$ by

$$p_t^{\odot q}(h_1 \odot \cdots \odot h_q) = p_t h_1 \odot \cdots \odot p_t h_q, \quad q \in \mathbb{N},\, h_1, \dots, h_q \in \mathscr{H},\, t \geq 0. \tag{2.8.2.1}$$

and, hence, to $\mathscr{L}^2(\mathfrak{W})$ by

$$P_t F = \sum_{q \in \mathbb{N}} I_q(p_t^{\odot q} h_q), \quad t \geq, \, F = \sum_{q \in \mathbb{N}} I_q(h_q) \in \mathscr{L}^2(\mathfrak{W}). \tag{2.8.2.2}$$

The choice of the free dynamic, that is a straight line evolution $p_t f = \mathrm{e}^{-t} f$, yields the Ornstein-Uhlenbeck semi-group. It follows that, in a sense, physicists knew the Malliavin derivative and the Ornstein-Uhlenbeck semi-group much before probabilist and for them the Ornstein-Uhlenbeck is completely trivial! In the same paper E. NELSON (1973) [106] also proved the hypercontractivity of this semi-group. The influential work of L. GROSS (1975) [61] showed that hypercontractivity is equivalent to a logarithmic Sobolev inequality and Gross proved this inequality for the Gaussian measure via the central limit theorem.

2.8.3. **Logarithmic Sobolev inequality, hypercontractivity and Bakry-Emery.** Studying logarithmic Sobolev inequalities for other processes motivated the key study of hypercontractive diffusions by D. BAKRY & M. ÉMERY (1985) [12] where they introduced (for every semi-group not just the ones associated to chaotic decomposition) the criterion we presented in Section 2.2 (as well as other equivalent formulations) and already proved the equivalence of their criterion with logarithmic Sobolev inequality. The work of Bakry-Emery gave birth to a far reaching chapter of modern probability and for the diffusion setting we only quote the monograph by D. BAKRY, I. GENTIL & M. LEDOUX (2014) [13]. The proof of the logarithmic Sobolev inequality by convexity we present is a bit different from the original one and we learned it from D. CHAFAÏ (2004) [31]. For more details on logarithmic Sobolev inequalities, one can also read the collective work (in french) of C. ANÉ ET AL. (2000) [6]. Outside of the diffusion setting not much is known in general and in fact the Bakry-Emery criterion is *not* equivalent to a logarithmic Sobolev inequality. In fact, several notions of logarithmic Sobolev inequality coexist in a discrete setting (in a diffusive setting they are equivalent thank to the chain rule). S. G. BOBKOV & P. TETALI (2006) [21] gave a detailed study of such logarithmic Sobolev inequalities in a discrete setting. However, many authors have considered modified logarithmic Sobolev inequalities. By the good properties of the logarithmic Sobolev inequality with respect to tensorization (that is, if $\mu$ satisfies the logarithmic Sobolev inequality then $\mu^{\otimes n}$ satisfies a logarithmic Sobolev inequality with the *same* constant), this was used combined with the central limit theorem in the original article of Gross [61] to prove the logarithmic Sobolev inequality for the Gaussian. The same idea was used by S. G. BOBKOV & M. LEDOUX (1998) [18] to prove a modified logarithmic Sobolev inequality for the Poisson measure on the integers. This inequality was generalized and improved for Poisson point processes by L. WU (2000) [152] via martingale techniques. Our inequality of Theorem 2.4.4.1 is new at this level of generality but is essentially the same as the one obtained by Wu put in our more general setting. The proof by convexity follows the one of Theorem 13 in S. BOURGUIN & G. PECCATI (2016) [27] that itself follows the already mentioned work of D. CHAFAÏ (2004) [31].

2.8.4. **The Malliavin-Stein approach.** The idea to use Malliavin calculus to obtain the fourth moment theorem for Gaussian stochastic integrals of D. NUALART & G. PECCATI (2005) [116] comes from the work of D. NUALART & S. ORTIZ-LATORRE (2008) [114]. The combination of the Malliavin calculus with the Stein bound Theorem 1.1.3.1 was introduced by I. NOURDIN & G. PECCATI (2009) [112]. These techniques known as Malliavin-Stein techniques are referenced in a Gaussian setting in the monograph by I. NOURDIN & G. PECCATI (2012) [111] and the webpage of I. Nourdin [107] gives a comprehensive list of works in that direction in and outside the Gaussian setting. The idea to use the $\Gamma$-calculus to derive the Stein inequality in an abstract diffusive setting is due to M. LEDOUX (2012) [84] and the notion of polynomial chaoses comes from the subsequent work of E. AZMOODEH, S. CAMPESE & G. POLY (2014) [9] that most of our presentation follows. For discrete models, only little is known and no general theory exists. For Poisson point processes, the theory of Malliavin-Stein was initiated by G. PECCATI, J. L. SOLÉ, M. S. TAQQU & F. UTZET (2010) [121] and followed by many important contributions. Note that they already already obtained our Stein inequality Theorem 2.4.4.2 in dimension $1$ for $\Gamma_0$ in a Poisson setting but did not link it to the chaotic decomposition as the technology of $\Gamma$-calculus on the Poisson space was not yet available. Eventually, this technology was developed by C. DÖBLER & G. PECCATI (2018) [40] and that leads to an exact fourth moment theorem on the Poisson space as we presented in Theorem 2.4.4.6 in a more general setting. For other discrete models only fourth moment theorems with remainder can be obtained such as the ones obtained for the twisted hypercube or Rademacher space by C. DÖBLER & K. KROKOWSKI (2017) [38].

# STABLE LIMIT THEOREMS ON THE POISSON SPACE

## 3.1. INTRODUCTION

We have seen that, for limit theorems for functionals of Poisson point processes, we can first derive a Stein inequality with remainder (Theorem 2.4.4.2) and then some algebraic manipulations on tensor products turn the fourth moment theorem with a quartic remainder (Theorem 2.4.4.3) into an exact fourth moment theorem (Theorem 2.4.4.6): the law of a functional $F$ living in a fixed Poisson chaos is close to a normal law if $\mathbb{E}F^4$ is close to $3(\mathbb{E}F^2)^2$. In other words, the non-diffusive Poisson fourth moment theorem is as good as the fourth moment theorem in the diffusive Gaussian setting. In a Gaussian framework, I. NOURDIN & D. NUALART (2010) [109] have extended the argument for the Stein inequality to measure the distance between the law of a Poisson functional (not necessarily living in fixed chaos), and the law of a random variable of the form $SN$ where $S$ is measurable with respect to the underlying Poisson process and $N$ is an normal random variable independent of the Poisson point process. It is therefore very natural to ask if the same analysis can be carried out in a Poisson framework: this will be done in the coming chapter from the point of view of quantitative estimates and of stable convergence.

The definition and study of *stable convergence* is one of the celebrated contribution of A. RÉNYI (1958) [134] and (1963) [135]; it is a refinement of the notion of convergence in law. Stable convergence is tailored for studying conditional limits of sequences of random variables. Thus, stable limits are, typically, *mixtures*, that is: objects of the form $SN$, where $N$ is an independent random element (possibly constructed on an extended probability space) and $S$ is a random variable. In a semi-martingales setting, J. JACOD & A. N. SHIRYAEV (2003) [70] obtained archetypal stable convergence results involving such mixtures. More recently, results by I. NOURDIN & D. NUALART (2010) [109], D. HARNETT & D. NUALART (2013) [63] and I. NOURDIN, D. NUALART & G. PECCATI (2016) [110] gave sufficient conditions and quantitative bounds for the stable convergence of functionals of an isonormal Gaussian processes to a *Gaussian mixture*, that is, a mixture as above, where $N$ is taken to be a Gaussian random variable. The typical application of such results is the study of the asymptotic behaviour of functionals of a fractional Brownian motion. The three references [109, 63, 110] made a pervasive use of the Malliavin techniques to prove such limit theorems. This approach was initiated by D. NUALART & G. PECCATI (2005) [116] to prove central limit theorems for iterated Itô integrals. The contribution of D. Nualart & G. Peccati is a milestone in the theory of limit theorems and has led to an independent field of research, known as the *Malliavin-Stein approach* (see the webpage of I. Nourdin [107] for a comprehensive list of contributions on the subject).

Following the trendsetting work of G. PECCATI, J. L. SOLÉ, M. S. TAQQU & F. UTZET (2010) [121], the Malliavin-Stein approach was extended, beyond the scope of Gaussian fields, to Poisson point processes. Despite being a very active field of research, the considered limit distributions have been until now prevalently Gaussian [76, 75, 81, 133, 122, 127, 138, 40, 41, 26] or, sometimes, Poisson [120] or Gamma [125, 40], and, to the best of our knowledge, Gaussian mixtures were never considered as limit distributions. The aim of this chapter is to tackle this problem, by proving an array of new quantitative and stable limit theorems on the Poisson space, with a target distribution given by a Gaussian mixture. We rely on a standard interpolation technique, known as *smart path*, Malliavin calculus for Poisson point processes as presented in Section 2.7, and the recently found representation of the carré du champ of D. BAKRY & M. ÉMERY (1985) [12] on the Poisson space, due to C. DÖBLER & G. PECCATI (2018) [40] that is extended by our Theorem 2.4.2.4. Our approach allows us to deal with any target distribution of the form $SN$, where $S$ is a matrix valued random variable (measurable with respect to the underlying Poisson point process) and $N$ is a Gaussian vector independent of the underlying Poisson point process, provided mild regularity assumptions on $S$ and on the functional under study (Theorems 3.3.1.4 and 3.3.1.5). Various applications are obtained such as:

- a stable fourth moment theorem on the Poisson space (Proposition 3.4.1.1);
- a first attack at the difficult problem of finding sufficient conditions for the limit of a $U$-statistic of order $2$ to be a Gaussian mixture in terms of the contractions of the kernel (Proposition 3.4.2.2);
- the characterization of the stable asymptotic behaviour of quadratic functionals of the Poisson approximation of Gaussian processes with stationary increments (Theorem 3.5.2.3).

The chapter is organized as follows. Each section starts with its own short introduction that presents its structure in detail and that recalls, if necessary, the context and the definition of the main objects under study. Theorems 3.3.1.4 and 3.3.1.5, that contain bounds and stable limit theorems for Poisson functionals, are the main results of the chapter and are presented in Section 3.3. A detailed comparison of these results with the aforementioned works on the Gaussian space of [109, 63, 110] follows in Section 3.3.2. A special attention to stochastic integrals is paid in Section 3.4. From our main results and these generic computations, we deduce, in Sections 3.4.1 and 3.4.2:

- Proposition 3.4.1.1, that is a stable version of the recently proved fourth moment theorem on the Poisson space of [40, 41].
- Proposition 3.4.2.2, that is a criterion for conditionally normal limit for order $2$ $U$-statistics.

In Section 3.5, given a Gaussian process with stationary increments, we introduce a class of compound Poisson processes obtained from the spectral representation of the Gaussian process. Applying Theorems 3.3.1.4 and 3.3.1.5, we prove Theorem 3.5.2.3, completely characterizing the asymptotic behaviour of a quadratic functional of a such compound Poisson process.

## 3.2. PRELIMINARIES

3.2.1. **Reminders on stochastic analysis for Poisson point processes.** We first recall some results about stochastic analysis for Poisson point processes that we obtained in Sections 2.4 and 2.7. Recall that random variables are defined on a probability space $(\Omega, \mathfrak{A}, \mathbb{P})$. We fix some measurable space $(Z, \mathscr{Z})$. Recall that $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$ is the space of countable sums of $\mathbb{N}$-valued measures on $Z$, and that $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$ can be endowed with the $\sigma$-algebra generated by the cylindrical mappings

(3.2.1.1) $$\xi \in \mathscr{M}_{\bar{\mathbb{N}}}(Z) \mapsto \xi(B) \in \mathbb{N} \cup \{\infty\}, \quad B \in \mathscr{Z}.$$

Let $\nu$ be a $\sigma$-finite measure on $(Z, \mathscr{Z})$. A random variable $\eta = \eta_\nu$ with values in $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$ is a *Poisson random measure* with intensity $\nu$, if the following two properties are satisfied:

(P.1) for all $B_1, \dots, B_n \in \mathscr{Z}$ pairwise disjoint, $\eta(B_1), \dots, \eta(B_n)$ are independent;

(P.2) for $B \in \mathscr{Z}$ with $\nu(B) < \infty$, $\eta(B) \sim \mathit{Poisson}(\nu(B))$.

We recall that such random variable exists. We let $\mathfrak{W} = \sigma(\eta)$. Also, for all $q \in \mathbb{N}$, we have defined $\eta^{(q)}$ the factorial power of $\eta$ and that $\mathbb{E}\eta^{(q)} = \nu^q$. We write $\mathscr{H} = \mathscr{L}^2(\nu)$. Recall that we use $\mathscr{H}^{\odot q}$ to denote the ($\nu^q$-almost everywhere) symmetric functions in $\mathscr{L}^2(\nu^q)$. We also have defined the multiple stochastic integrals

(3.2.1.2) $$I_q \colon \mathscr{H}^{\odot q} \to \mathscr{L}^2(\mathfrak{W}), \quad q \in \mathbb{N},$$

such that

(3.2.1.3) $$\mathbb{E} I_q(h) I_p(\tilde{h}) = 1_{q=p} q! \nu^q(h\tilde{h}), \quad q, p \in \mathbb{N},\ h \in \mathscr{H}^{\odot q},\ \tilde{h} \in \mathscr{H}^{\odot p}.$$

The spaces $\mathscr{C}_q = \operatorname{im} I_q$ yield a chaotic decomposition of $\mathscr{L}^2(\mathfrak{W})$, and every $F \in \mathscr{L}^2(\mathfrak{W})$ can be written as

(3.2.1.4) $$F = \sum_{q \in \mathbb{N}} J_q F = \sum_{q \in \mathbb{N}} I_q(h_q), \quad h_q \in \mathscr{H}^{\odot q},\ \forall q \in \mathbb{N},$$

where $J_q$ is the projection onto $\mathscr{C}_q$ for all $q \in \mathbb{N}$. This decomposition gives rise to the two unbounded Malliavin operators:

(3.2.1.5) $$\begin{cases} \mathscr{D}\mathrm{om}\, L = \left\{ F \in \mathscr{L}^2(\mathfrak{W}), \text{ such that } \sum_{q \in \mathbb{N}} q^2 \mathbb{E}\left( (J_q F)^2 \right) < \infty \right\}, \\ L \colon \mathscr{D}\mathrm{om}\, L \to \mathscr{L}^2(\mathfrak{W}), \\ LF = -\sum_{q \in \mathbb{N}} q J_q F, \quad F \in \mathscr{D}\mathrm{om}\, L; \end{cases}$$

(3.2.1.6) $$\begin{cases} \mathscr{D}\mathrm{om}\, D = \left\{ F \in \mathscr{L}^2(\mathfrak{W}), \text{ such that } \sum_{q \in \mathbb{N}} q |h_q|^2_{\mathscr{H}^{\odot q}} < \infty \right\}, \\ D \colon \mathscr{D}\mathrm{om}\, D \to \mathscr{L}^2(\mathfrak{W}) \otimes \mathscr{H}, \\ D_z F = \sum_{q \in \mathbb{N}} q I_{q-1}(h_q(z, \cdot)), \quad F \in \mathscr{D}\mathrm{om}\, D,\ z \in Z. \end{cases}$$

By Theorems 2.7.2.5 and 2.7.3.3, we know that, for $F \in \mathscr{D}\mathrm{om}\, D$, $DF = D^+F$, where

(3.2.1.7) $$D_z^+ F = f(\eta + \delta_z) - f(\eta), \quad F = f(\eta) \in \mathscr{L}^0(\mathfrak{W}),\ z \in Z.$$

We also have that

(3.2.1.8) $$D_z^- F = (f(\eta) - f(\eta - \delta_z)) 1_{\eta(z) > 0}, \quad F = f(\eta) \in \mathscr{L}^0(\mathfrak{W}),\ z \in Z.$$

By the Mecke formula Theorem 2.7.3.2, the Campbell measure of the Poisson point process $\eta$ is $\eta$ itself, and by Theorem 2.4.2.4, the carré du champ is given by $\mathscr{D}\mathrm{om}\,\Gamma = \mathscr{D}\mathrm{om}\,D$ and

$$2\Gamma(F,G) = \int D_z^+ F D_z^+ G \nu(\mathrm{d}z) + \int D_z^- F D_z^- G \eta(\mathrm{d}z), \quad F,G \in \mathscr{D}\mathrm{om}\,D. \tag{3.2.1.9}$$

The carré du champ $\Gamma$ satisfies the pseudo-chain rule as stated in Lemma 2.4.3.1. Recall also that $\mathscr{A} = \mathscr{D}\mathrm{om}\,D \cap \mathscr{L}^\infty(\mathfrak{W})$ is an algebra stable by composition with Lipschitz and smooth functions.

We now give the following product formulae that we will us when carrying out explicit computations. Let $f \in \mathscr{L}^2_\sigma(\nu^p)$ and $g \in \mathscr{L}^2_\sigma(\nu^q)$ such that $I_p(f)I_q(g) \in \mathscr{L}^2(\mathfrak{W})$. By Lemma 2.4.2.1, there exists $h_r \in \mathscr{L}^2_\sigma(\nu^r)$ such that

$$I_p(f)I_q(g) = \sum_{r=0}^{p+q} I_r(h_r). \tag{3.2.1.10}$$

For $f \in \mathscr{L}^2_\sigma(\nu^p)$ and $g \in \mathscr{L}^2_\sigma(\nu^q)$, we define the *star contraction* of order $(l,r)$, $r \in \{0,\dots,p\wedge q\}$ and $l \in \{0,\dots,r\}$ by

$$f \star_r^l g(x_1,\dots,x_{p+q-r-l}) = \int f(y_{[l]}, x_{[p-l]}) g(y_{[l]}, x_{[r-l]}, x_{[p-l+1,p+q-r-l]}) \nu^l(\mathrm{d}y_{[l]}). \tag{3.2.1.11}$$

Then by [80, Proposition 5] for $f \in \mathscr{L}^2_\sigma(\nu^p)$ and $g \in \mathscr{L}^2_\sigma(\nu^q)$ such that $f \star_r^l g \in \mathscr{L}^2(\nu^{p+q-r-l})$,

$$I_p(f)I_q(g) = \sum_{r=0}^{p\wedge q}\sum_{l=0}^{r} r! \binom{p}{r}\binom{q}{r}\binom{r}{l} I_{p+q-r-l}(f \star_r^l g). \tag{3.2.1.12}$$

3.2.2. **Extended Malliavin operators.** As we will consider stable convergence and mixtures that involve random variables independent of $\mathfrak{W}$, we will extend the definition of $D^+$ and $D^-$ in that direction. Whenever $F \in \mathscr{L}^0(\mathfrak{W})$ and $a$ is independent of $\mathfrak{W}$, we write $D^\pm(aF) = aD^\pm F$. We extend $D$, $L$ and $\Gamma$ accordingly. Recall that whenever $F = (F_1,\dots,F_{d_1}) \in \mathscr{D}\mathrm{om}\,D$ and $G = (G_1,\dots,G_{d_2}) \in \mathscr{D}\mathrm{om}\,D$, one will write $\Gamma(F,G)$ to indicate the $d_1 \times d_2$ symmetric random matrix whose coefficient of index $(i,j)$ is given by

$$\frac{1}{2}(\Gamma(F_i,G_j) + \Gamma(F_j,G_i)). \tag{3.2.2.1}$$

For $F = (F_1,\dots,F_d) \in \mathscr{D}\mathrm{om}\,D$, we defined the matrix-valued Stein kernel of $F$ as the following $d \times d$ random symmetric matrix

$$S(F) = -\Gamma(L^{-1}F, F), \tag{3.2.2.2}$$

and recall that, in view of the hypothesis on $F$, $S(F) \in \mathscr{L}^1(\mathfrak{W})$ and

$$\mathbb{E}FF^T = \mathbb{E}S(F). \tag{3.2.2.3}$$

3.2.3. **Stable convergence.** We conclude this section by defining the notion of stable convergence that we will study in this chapter. (See [70, Chap VIII Section 5c].) A sequence of random variables $(F_n) \subset \mathscr{L}^2(\Omega, \mathfrak{A}, \mathbb{P})$ is said to *converge stably* towards $F_\infty \in \mathscr{L}^2(\Omega, \mathfrak{A}, \mathbb{P})$ whenever for all bounded random variables $Z$, measurable with respect to $\mathfrak{W}$,

$$(F_n, Z) \xrightarrow[n\to\infty]{law} (F_\infty, Z). \tag{3.2.3.1}$$

This convergence is denoted by

$$F_n \xrightarrow[n\to\infty]{stably} F_\infty. \tag{3.2.3.2}$$

Of course, stable convergence implies convergence in law but the reverse implication does not hold. In practice, we simply need to check the previous convergence for a smaller class of bounded random variables such that the $\mathbb{P}$-completion of the generated $\sigma$-algebra coincides with $\mathfrak{W}$. We let $\mathscr{G}$ be the linear span of the random vectors of the form $(\mathrm{e}^{-\eta(v_1)}, \dots, \mathrm{e}^{-\eta(v_d)})$ for some $v_i \colon Z \to [0, \infty]$ with $\nu(v > 0) < \infty$ for all $i \in [d]$. From the proof of Lemma 2.7.1.2, we know that $\sigma(\mathscr{G}) = \mathfrak{W}$ and we have that

**Proposition 3.2.3.1.** *Let $(F_n)_{n\in\mathbb{N}} \subset \mathscr{L}^2(\mathfrak{W})$. Then, the following are equivalent:*

$$F_n \xrightarrow[n\to\infty]{stably} F_\infty; \tag{3.2.3.3}$$

$$(F_n, G) \xrightarrow[n\to\infty]{law} (F_\infty, G), \quad \forall G \in \mathscr{G}. \tag{3.2.3.4}$$

## 3.3. Main abstract results

**Outline.** The main results of the section are Theorems 3.3.1.4 and 3.3.1.5 and are presented in Section 3.3.1. Theorem 3.3.1.5 gives sufficient conditions for the stable convergence of a sequence of Poisson functionals to a random variable of the type $\Sigma N$ for a matrix-valued $\Sigma \in \mathscr{L}^2(\mathbb{P})$ and $N \sim \mathbf{N}(0, id_{\mathbb{R}^d})$ independent of $\mathfrak{W}$. Informally, we refer to a random variable of the form $\Sigma N$ as a *Gaussian mixture*. Theorem 3.3.1.4 is a quantitative version of the previous theorem and provides bounds on the distance $d_2$ between the distribution of a Poisson functional and that of a random variable of the type $\Sigma N$ for a matrix-valued $\Sigma \in \mathscr{L}^2(\mathbb{P})$ and $N \sim \mathbf{N}(0, id_{\mathbb{R}^d})$ independent of $\mathfrak{W}$. The proof of Theorem 3.3.1.4 is rather technical and is based on Lemmas 3.3.1.2 and 3.3.1.3, that we prove in Section 3.3.3. We compare our theorems to existing theorems on the Gaussian space in Section 3.3.2.

3.3.1. **Main results.** For $F \in \mathscr{L}^0(\mathfrak{W})$, taking values in $\mathbb{R}$, we defined

$$\Delta(F) = \int (D_z^+ F)^4 \nu(\mathrm{d}z). \tag{3.3.1.1}$$

This quantity is well-defined, though possibly infinite. The polynomial remainder $\Delta(F)$ can be regarded as an equivalent, at our level, of the quantities appearing in the Lyapunov condition associated with the classical central theorem (see, e.g. [16, Theorem 27.3]). This remainder already appeared in the discrete setting (see Theorem 2.4.4.3). If moreover $\mathbb{E}F = 0$, we also define

(3.3.1.2) $$\gamma^2(F) = \int (D_z^+ L^{-1} F)^2 \nu(\mathrm{d}z).$$

Observe that, since $L^{-1}F \in \mathscr{D}\mathrm{om}\, D$, by Theorem 2.7.3.2, $\mathbb{E}\gamma^2(F) = \mathbb{E}\Gamma(L^{-1}F) < \infty$. The quantity $\gamma^2(F)$ plays a role in controlling remainders when dealing with infinite chaotic decompositions.

The next statement is a stable convergence result for Poisson functionals and motivates other results of this chapter.

**Proposition 3.3.1.1.** *Let $(F_n = (F_{n,1}, \dots, F_{n,d})) \subset \mathscr{D}\mathrm{om}\, D$, $\Sigma = (\Sigma_{ij})_{i,j=1}^d \in \mathscr{L}^2(\Omega)$, $N$ be a standard $d$-dimensional Gaussian independent of $\eta$. Let $C = \Sigma\Sigma^T$. Assume that*

(3.3.1.3) $$-\Gamma(L^{-1}F_{n,i}, F_{n,j}) \xrightarrow[n\to\infty]{\mathscr{L}^1(\mathfrak{W})} C_{ij}, \quad i, j \in [d].$$

*and that either one of the following conditions (3.3.1.4) or (3.3.1.5) is satisfied:*

(3.3.1.4a) $$\gamma^2(F_{n,i}) \xrightarrow[n\to\infty]{} 0, \quad i \in [d],$$

(3.3.1.4b) $$\sup_n \Delta(F_{n,i}) < \infty, \quad i \in [d];$$

*or*

(3.3.1.5a) $$\Delta(F_{n,i}) \xrightarrow[n\to\infty]{} 0, \quad i \in [d],$$

(3.3.1.5b) $$\sup_n \gamma^2(F_n) < \infty, \quad i \in [d].$$

*Then, $(F_n)$ converges stably to $\Sigma N$.*

*Proof.* Following [109, 63], we will use the characteristic function method. Let $G \in \mathscr{G}$. From Proposition 3.2.3.1 it is sufficient to show that,

(3.3.1.6) $$(F_n, G) \xrightarrow[n\to\infty]{law} (F_\infty, G),$$

where $F_\infty$ is a random vector satisfying

(3.3.1.7) $$\mathbb{E}[\mathrm{e}^{i\langle\lambda, F_\infty\rangle_{\ell^2}} | \eta] = \exp\left(-\frac{1}{2}|\Sigma\lambda|_{\ell^2}\right), \quad \lambda \in \mathbb{R}^d.$$

Let us define for all $n \in \mathbb{N}$,

(3.3.1.8) $$\phi_n(\lambda) = \mathbb{E}(\mathrm{e}^{i\langle\lambda, F_n\rangle_{\ell^2}} G), \quad \lambda \in \mathbb{R}^d.$$

Since by assumption,

(3.3.1.9) $$\mathbb{E}F_n^T F_n = \mathbb{E}S(F_n) \xrightarrow[n\to\infty]{} \mathbb{E}\Sigma\Sigma^T,$$

the sequence $(F_n, G)_{n\in\mathbb{N}}$ is bounded in $\mathscr{L}^2(\mathfrak{W})$, and, by the Markov inequality, such a sequence is tight. It is a well-known fact [71, Lemma 5.2 & Theorem 5.3], that up to passing to a sub-sequence, we obtain that there exists a random vector $F_\infty$ such that

(3.3.1.10) $$(F_n, G) \xrightarrow[n\to\infty]{law} (F_\infty, G),$$

and such that

$$\phi_n(\lambda) \xrightarrow[n\to\infty]{} \phi_\infty(\lambda) := \mathbb{E}(\mathrm{e}^{i\langle\lambda,F_\infty\rangle_{\ell^2}}\, G), \quad \lambda\in\mathbb{R}. \tag{3.3.1.11}$$

Let $j\in[d]$. Recall [71, Lemma 4.11], that is a sequence of random variables converges in law and is uniformly integrable, then the sequence of mean also converge. Since $(F_n)_{n\in\mathbb{N}}$ is bounded in $\mathscr{L}^2(\mathfrak{W})$, it is also uniformly integrable, and, by [71, Lemma 4.11], we find that

$$\partial_j\phi_n(\lambda) = i\mathbb{E}(F_{n,j}\,\mathrm{e}^{i\langle\lambda,F_n\rangle_{\ell^2}}\, G) \xrightarrow[n\to\infty]{} i\mathbb{E}(F_{\infty,j}\,\mathrm{e}^{i\langle\lambda,F_\infty\rangle_{\ell^2}}\, G) = \partial_j\phi_\infty(\lambda). \tag{3.3.1.12}$$

On the other hand by integration by parts (3.2.2.3), we obtain that

$$\partial_j\phi_n(\lambda) = i\mathbb{E}\Gamma(L^{-1}F_{n,j}, \mathrm{e}^{i\langle\lambda,F_n\rangle_{\ell^2}}\, G). \tag{3.3.1.13}$$

Combining (2.4.2.5) and (2.4.2.6) and Theorem 2.7.3.2 then using the pseudo chain rule Lemma 2.4.3.1 yields

$$\begin{aligned}\partial_j\phi_n(\lambda) &= i\mathbb{E}G\Gamma(L^{-1}F_{n,j}, \mathrm{e}^{i\langle\lambda,F_n\rangle_{\ell^2}}) + i\mathbb{E}\,\mathrm{e}^{i\langle\lambda,F_n\rangle_{\ell^2}}\,\Gamma(L^{-1}F_{n,j}, G)\\ &= -\sum_{k=1}^{d}\lambda_k\mathbb{E}G\,\mathrm{e}^{i\langle\lambda,F_n\rangle_{\ell^2}}\,\Gamma(L^{-1}F_{n,j}, F_{n,k})\\ &\quad + R_{\tau_\lambda}(F_n, L^{-1}F_{n,j}) + i\mathbb{E}\,\mathrm{e}^{i\langle\lambda,F_n\rangle_{\ell^2}}\,\Gamma(L^{-1}F_{n,j}, G),\end{aligned} \tag{3.3.1.14}$$

where $\tau_\lambda(x) = \mathrm{e}^{i\langle\lambda,x\rangle_{\ell^2}}$ for $x\in\mathbb{R}^d$. We fix $j$ and $k\in[d]$, and we write $S_n = -\Gamma(L^{-1}F_{n,j}, F_{n,k})$ $(n\in\mathbb{N})$. Let $\lambda\in\mathbb{R}^d$. By (3.3.1.3), we have that

$$\mathbb{E}G\,\mathrm{e}^{i\langle\lambda,F_n\rangle_{\ell^2}}\, S_n \xrightarrow[n\to\infty]{} \mathbb{E}G\,\mathrm{e}^{i\langle\lambda,F_\infty\rangle_{\ell^2}}\, C_{ij}. \tag{3.3.1.15}$$

By the properties of the set $\mathscr{G}$, the Cauchy-Schwarz inequality or by Hölder's inequality and Theorem 2.7.3.2, we have that

$$\mathbb{E}|\Gamma(L^{-1}F_{n,j}, G)| \le c\gamma^2(F_{n,j})^{\frac{1}{2}}; \tag{3.3.1.16}$$

$$\mathbb{E}|\Gamma(L^{-1}F_{n,j}, G)| \le c\Delta(F_{n,j})^{\frac{1}{4}}; \tag{3.3.1.17}$$

$$|R_{\tau_\lambda}(F_n, L^{-1}F_{n,j}) \le \lambda^2\left(\sum_{i=1}^{d}\gamma_2(F_j)^{\frac{1}{2}}\right)^2\sum_{k,l=1}^{d}\left(\Delta(F_k)^{\frac{1}{2}}\Delta(F_l)^{\frac{1}{2}}\right)^{\frac{1}{2}}. \tag{3.3.1.18}$$

Therefore, combining (3.3.1.15) with either (3.3.1.4) or (3.3.1.5), we find that

$$\partial_j\phi_\infty(\lambda) = -\sum_{k=1}^{d}\lambda_k\mathbb{E}(\mathrm{e}^{i\langle\lambda,F_\infty\rangle_{\ell^2}}\, C_{kj}G), \quad \lambda\in\mathbb{R}^d,\ j\in[d]. \tag{3.3.1.19}$$

In view of the density of $\mathscr{G}$ in $\mathscr{L}^\infty(\mathfrak{W})$, this yields the differential equation

$$\frac{\partial}{\partial\lambda_j}\mathbb{E}[\mathrm{e}^{i\langle\lambda,F_\infty\rangle_{\ell^2}}\,|\eta] = -\sum_{k=1}^{d}\lambda_k C_{kj}\mathbb{E}[\mathrm{e}^{i\langle\lambda,F_\infty\rangle_{\ell^2}}\,|\eta], \quad \lambda\in\mathbb{R}^d,\ j\in[d]. \tag{3.3.1.20}$$

The unique solution to this differential system is (3.3.1.7). □

The condition (3.3.1.3) involves $\mathscr{L}^1(\mathfrak{W})$ convergence. Such convergence is, in general, established by proving a convergence in $\mathscr{L}^2(\mathfrak{W})$. However, we do not expect, in general, convergence in $\mathscr{L}^2(\mathfrak{W})$. We subsequently develop more refined version of Proposition 3.3.1.1 that allows to deal with weaker form of convergence (such as the convergence in law). We start by proving a quantitative counterpart to Proposition 3.3.1.1. Note also that, in this case, we work with the symmetric random matrix $S(F)$ rather than its non-symmetric version (see the discussion in Section 3.3.2). Stating our results requires some further notations. We identify the real matrices of size $d \times d$ with vectors of $\mathbb{R}^{d^2}$. For a square-integrable random variable $F$, with values in $\mathbb{R}^d$, we introduce quantities related to $\Delta(F)$ and $\gamma^2(F)$, that we used in Proposition 3.3.1.1. We let

$$r^n(F) = \mathbb{E} \int |D_z^+ F|_{\ell^1}^n \nu(\mathrm{d}z), \tag{3.3.1.21}$$

and

$$\gamma^2(F) = r^2(L^{-1}F). \tag{3.3.1.22}$$

For every multivariate square-integrable $F$ with $\mathbb{E}F = 0$, we have that $L^{-1}F \in \mathscr{D}\mathrm{om}\, L \subset \mathscr{D}\mathrm{om}\, D$. Therefore, by the Cauchy-Schwarz inequality $D^+L^{-1}FD^+L^{-1}F \in \mathscr{L}^1(\mathbb{P} \otimes \nu)$ and $\gamma^2(F) < \infty$. The quantity $r^n(F)$ is always well-defined, although possibly infinite. It is finite if $D^+F \in \mathscr{L}^n(\mathfrak{W})$. Given $F = (F_1, \dots, F_d) \in \mathscr{D}\mathrm{om}\, D$, $\Sigma = (\Sigma_{ij}) \in \mathscr{D}\mathrm{om}\, D$, and $G = (G_1, \dots, G_d) \in \mathscr{D}\mathrm{om}\, D$, we introduce the following quantities that appear in Theorems 3.3.1.4 and 3.3.1.5.

$$\epsilon_4(F, \Sigma) = r^4(F) \left( r^{4/3}(\Sigma\Sigma^T) + r^{8/3}(F)^{1/2} + r^{8/3}(\Sigma)^{1/2} \right); \tag{3.3.1.23}$$

$$\epsilon_\infty(F, \Sigma) = \gamma^2(F) \left( r^2(\Sigma\Sigma^T) + r^4(\Sigma)^{1/2} + r^4(F)^{1/2} \right); \tag{3.3.1.24}$$

$$c_0 = (2^{\frac{1}{3}} + 2^{-\frac{2}{3}}) \left( 2\sqrt{d} + \mathbb{E}|F|_{\ell^2} + \sqrt{\mathbb{E}|\Sigma|_{\ell^2}^2} \right)^{\frac{2}{3}}; \tag{3.3.1.25}$$

$$\tau(\epsilon) = \max \left( c_0 \left( \frac{\sqrt{2\pi}}{4} \sqrt{d\epsilon} \right)^{\frac{1}{3}}, 12\sqrt{d\epsilon} \right), \ \forall \epsilon \geq 0. \tag{3.3.1.26}$$

Most of our analysis relies on the following central lemma from which we deduce several important bounds. Let us point out that we did not try to optimize the bounds, favoring instead clarity and conciseness.

**Lemma 3.3.1.2.** *Let $F = (F_1, \dots, F_d) \in \mathscr{D}\mathrm{om}\, D$, $\Sigma = (\Sigma_{ij}) \in \mathscr{D}\mathrm{om}\, D$, let $N$ be a standard $d$-dimensional Gaussian vector independent of $\eta$, and let $G \in \mathscr{D}\mathrm{om}\, D$. Then, for all $\phi \in \mathscr{C}^3(\mathbb{R}^d)$, the following bound holds*

(3.3.1.27)
$$\begin{aligned}
&|\mathbb{E}\phi(G+F) - \mathbb{E}\phi(G+\Sigma N)| \leq |\nabla^2\phi|_{\ell^\infty,\infty} (2\mathbb{E}|S(F) - \Sigma\Sigma^T|_{\ell^1} \\
&+ \mathbb{E} \int |D_z^+ L^{-1}F|_{\ell_1} |D_z^+ G|_{\ell_1} \nu(\mathrm{d}z)) \\
&+ 5|\nabla^3\phi|_{\ell^\infty,\infty} \mathbb{E} \int |D_z^+ L^{-1}F|_{\ell^1} \left( |D_z^+(\Sigma\Sigma^T)|_{\ell^1} + |D_z^+\Sigma|_{\ell^1}^2 + |D_z^+F|_{\ell^1}^2 + |D_z^+G|_{\ell^1}^2 \right) \nu(\mathrm{d}z)
\end{aligned}$$

*Moreover, the two following bounds are also valid*

$$
\begin{aligned}
&|\mathbb{E}\phi(F+G)-\mathbb{E}\phi(\Sigma N+G)| \le |\nabla^2\phi|_{\ell^\infty,\infty}\mathbb{E}|S(F)-\Sigma\Sigma^T| \\
&+|\nabla^2\phi|_{\ell^\infty,\infty} r^4(F)^{1/4} r^{4/3}(G) \\
&+|\nabla^3\phi|_{\ell^\infty,\infty} r^4(F)^{1/4}\left(r^{4/3}(\Sigma\Sigma^T)+r^{8/3}(F)^{1/2}+r^{8/3}(\Sigma)^{1/2}+r^{8/3}(G)^{1/2}\right);
\end{aligned}
\tag{3.3.1.28}
$$

$$
\begin{aligned}
&|\mathbb{E}\phi(F+G)-\mathbb{E}\phi(\Sigma N+G)| \le |\nabla^2\phi|_{\ell^\infty,\infty}\mathbb{E}|S(F)-\Sigma\Sigma^T| \\
&+\gamma^2(F)^{1/2}\left(|\nabla^2\phi|_{\ell^\infty,\infty} r^2(G)+|\nabla^3\phi|_{\ell^\infty,\infty}\left(r^2(\Sigma\Sigma^T)+r^4(F)^{1/2}+r^4(\Sigma)^{1/2}+r^4(G)^{1/2}\right)\right).
\end{aligned}
\tag{3.3.1.29}
$$

The next lemma appears implicitly in the proof of [110, Theorem 3.4] in the case $d=1$. We state it here for $d\in\mathbb{N}$ and give an outline of the proof (for completeness) in Section 3.3.3.

**Lemma 3.3.1.3.** *Let $F=(F_1,\dots,F_d)\in\mathscr{D}\mathrm{om}\,D$, $\Sigma=(\Sigma_{ij})\in\mathscr{D}\mathrm{om}\,D$ and $N$ be a standard $d$-dimensional Gaussian independent of $\eta$. Assume that for $a,b>0$, and all $\phi\in\mathscr{C}^3(\mathbb{R}^d)$ with bounded derivatives,*

$$|\mathbb{E}\phi(F)-\mathbb{E}\phi(\Sigma N)| \le a|\nabla^2\phi|_{\ell^\infty,\infty}+b|\nabla^3\phi|_{\ell^\infty,\infty}. \tag{3.3.1.30}$$

*Then,*

$$d_2(F,\Sigma N)\le \tau(a+b), \tag{3.3.1.31}$$

*where $\tau$ is defined in (3.3.1.26).*

The following statement is one of the main results of the chapter. Recall the definition of $\epsilon_4$, $\epsilon_\infty$ and $\tau$ in (3.3.1.23), (3.3.1.24) and (3.3.1.26).

**Theorem 3.3.1.4.** *Let $F=(F_1,\dots,F_d)\in\mathscr{D}\mathrm{om}\,D$, $\Sigma=(\Sigma_{ij})\in\mathscr{D}\mathrm{om}\,D$, $N$ be a standard $d$-dimensional Gaussian independent of $\eta$. Let $\epsilon\in\{\epsilon_4(F,\Sigma),\epsilon_\infty(F,\Sigma)\}$. Then,*

$$d_2(F,\Sigma N)\le\tau\left(\mathbb{E}|S(F)-\Sigma\Sigma^T|+\epsilon\right). \tag{3.3.1.32}$$

*Proof.* For simplicity, we do not specify the dependence in $F$ and $\Sigma$ in $\epsilon_4$ or $\epsilon_\infty$. The bounds (3.3.1.32) are direct consequences of Lemma 3.3.1.3, andLemma 3.3.1.2 in the case $G=0$ and. Indeed, when $\epsilon=\epsilon_\infty$, combining (3.3.1.28) and (3.3.1.31) produces (3.3.1.32). When $\epsilon=\epsilon_4$, combining Lemma 3.3.1.2 and (3.3.1.31) produces (3.3.1.32). This concludes the proof. □

As announced, we now establish the qualitative result similar to Proposition 3.3.1.1 with a weakened version of (3.3.1.3).

**Theorem 3.3.1.5.** *Let $(F_n)\subset\mathscr{D}\mathrm{om}\,D$ (possibly vector-valued), $\Sigma=(\Sigma_{ij})\in\mathscr{L}^0(\Omega)$, $N$ be a standard $d$-dimensional Gaussian independent of $\eta$. Let $\tau$ denotes either the topology of the convergence in law or the topology of the convergence stable. Assume that*

$$S(F_n)\xrightarrow[n\to\infty]{\tau}\Sigma\Sigma^T, \tag{3.3.1.33}$$

*and that there exists $n_0 \in \mathbb{N}$ such that $S(F_n) \geq 0$ for all $n \geq n_0$. Assume, moreover, that either one of the following conditions (3.3.1.34) or (3.3.1.35) is satisfied:*

(3.3.1.34a) $$\gamma^2(F_n) \xrightarrow[n\to\infty]{} 0,$$

(3.3.1.34b) $$\sup_{n\geq n_0} r^4(F_n) + r^2(S(F_n)) + r^4(S^{\frac{1}{2}}(F_n)) < \infty;$$

*or*

(3.3.1.35a)
$$r^4(F_n) \xrightarrow[n\to\infty]{} 0,$$
(3.3.1.35b)
$$\sup_{n\geq n_0} r^2(F_n) + r^4(F_n) + r^2(S^{\frac{1}{2}}(F_n)) + r^4(S^{\frac{1}{2}}(F_n)) + r^1(S(F_n)) + r^2(S(F_n)) < \infty.$$

*Then, $(F_n)$ converges to $\Sigma N$ for the topology $\tau$.*

*Proof.* Note that, when dealing with (3.3.1.35), we need to check that $\sup r^{8/3}(F_n) + r^{8/3}(\Sigma) + r^{4/3}(\Sigma\Sigma^T) < \infty$, under (3.3.1.35b). This is a consequence of the following classical interpolation result: for $p, q \geq 0$ and $\theta \in (0,1)$ such that $\frac{\theta}{p} + \frac{1-\theta}{q} = 1$, then,

(3.3.1.36) $$|\phi|_1 \leq |\phi|_p^{\theta} |\phi|_q^{1-\theta},$$

where $\phi$ is any measurable function and $|\cdot|_l$ means the $\mathscr{L}^p$-norm with respect to any measure. Applying this inequality with $\theta = 1/2$, $p = 3/2$ and $q = 3/4$, we obtain that

(3.3.1.37) $$r^{4/3}(\Sigma\Sigma^T) \leq r^2\left(\Sigma\Sigma^T\right)^{1/4} r^1\left(\Sigma\Sigma^T\right)^{1/2}; \quad \text{and} \quad r^{8/3}(\Sigma) \leq r^4(\Sigma)^{1/8} r^2(\Sigma)^{1/4}.$$

Let us prove the stable convergence. Let $\phi$ be smooth with bounded derivatives and $G \in \mathscr{G}$. We apply the triangle inequality to obtain
(3.3.1.38)
$$|\mathbb{E}\phi(G+F) - \mathbb{E}\phi(G+\Sigma N)| \leq |\mathbb{E}\phi(G+F) - \mathbb{E}\phi(G+S^{\frac{1}{2}}(F)N)| + |\mathbb{E}\phi(G+S^{\frac{1}{2}}(F_n)) - \mathbb{E}\phi(G+\Sigma N)|.$$

By either (3.3.1.35) or (3.3.1.34) and Lemma 3.3.1.2 we find that the first term on the right-hand side vanishes as $n \to \infty$. By assumption of stable convergence, the first term on the right-hand side also vanishes. Let $a, b \in \mathbb{R}^d$. Without loss of generality, we assume that $a_i \neq 0$, for all $i \in [d]$. We take $\phi(x) = \mathrm{e}^{ia\cdot x}$ and $G = (\frac{b_1}{a_1}\,\mathrm{e}^{-\eta(v_1)}, \dots, \frac{b_d}{a_d}\,\mathrm{e}^{-\eta(v_d)})$ where $v_i \geq 0$ and $\nu(v_i > 0) < \infty$ for all $i \in [d]$. Hence, we have proved that,

(3.3.1.39) $$\mathrm{e}^{i(a\cdot F_n + b\cdot \tilde{G})} \to \mathrm{e}^{i(a\cdot \Sigma N + b\cdot \tilde{G})}, \quad \text{for all } \tilde{G} \in \mathscr{G}.$$

This gives the stable convergence by Proposition 3.2.3.1. The proof for the topology of the convergence in law is the same as for the stable convergence but with $G = 0$. This concludes the proof. □

3.3.2. **Comparison with the results on Gaussian spaces.** In Theorems 3.3.1.4 and 3.3.1.5, when $S(F)$ is close to $\Sigma\Sigma^T$ and $\epsilon_4$ or $\epsilon_\infty$ is small then the distribution of the functional $F = (F_1, \dots, F_d)$ will be close to that of the Gaussian mixture $\Sigma N$. The comparison of $S(F)$ and $\Sigma\Sigma^T$ is similar to the Gaussian cases: the quantity $-\langle DF, DL^{-1}F\rangle$ where $D$ is the Malliavin derivative and $L^{-1}$ the pseudo-inverse of the Ornstein-Uhlenbeck generator on the Gaussian space controls the asymptotic variance of the functional $F$. In this respect, let us refer to [111, Theorem 5.3.1] for deterministic variance and to [109, Theorem 3.1], to [63, Theorem 3.2] and to [110, Theorem 5.1] for random asymptotic variances. Let us point out some differences of our results with the Gaussian results of [109, 63, 110].

First, on the Gaussian spaces, the authors of [109, 63, 110] work with iterated Skorokhod integrals of any order $q \in \mathbb{N}$. That is, given $u$ such that $F = \delta^q u$, they give analytical conditions on $u$ and $F$ for the stable convergence. In the particular case of $q = 1$, we can find our condition on the random covariance $S(F)$. Indeed, since $L = -\delta D$, we can always choose $u = -DL^{-1}F$ in their theorems (note that other choices are possible). Let us point out that, due to the lack of diffusiveness on the Poisson space, it does not seem possible to reach a result involving iterated Kabanov integrals, via our method of proof, that is, via integration by parts.

Second, we obtain two different bounds involving respectively $r^4(F)$ and $\gamma^2(F)$ to control the remainders while on the Gaussian space [109, 63, 110] obtain only one bound involving two terms as, e.g. the one appearing in [110, Corollary 3.2]. The reader can easily verify that applying the Cauchy-Schwarz inequality or the Hölder inequality on one of their bounds (for instance, in [110, Equation 3.1]) can lead to two bounds similar to our estimates. In our case, due to the lack of diffusiveness, we obtain extra terms with respect to [110, Corollary 3.2]. We gather all of those terms in the synthetic quantities $r^4(F)$ and $\gamma^2(F)$. Aggregating all the terms in one quantity helps us to to interpret the estimates. The bound involving $\epsilon_\infty$ and $\gamma^2(F)$ is useful whenever one deals with sequences of random variables $F_n = \sum_{q \in \mathbb{N}} I_q(f_{n,q})$ such that $\mathbb{E} I_q(f_{n,q})^2 = o(\mathbb{E}F_n^2)$. In that case, with $G_n = F_n/(\mathbb{E}F_n^2)$,there exists a sequence of integers $k_n$ converging to $\infty$ such that

$$\mathbb{E}\gamma^2(G_n) \leq \frac{1}{k_n}. \tag{3.3.2.1}$$

However, for variables of finite chaotic decomposition (or with a dominant term of finite order), $\gamma^2(F_n)$ and $S(F_n)$ are typically of the same order and we, therefore, need $\epsilon_4$ and $r^4(F)$. Note that bounds involving polynomial quantities such as $r^4(F) = \int |D_z^+ F|^4$ are quite common in the Poisson setting see, among others, [121, Equation 3.4], [123, Equation 23], or [40, Equation 1.8].

Thirdly, the authors of [109, 63, 110] obtained results involving the convergence in $\mathscr{L}^1(\mathfrak{W})$ of the Stein matrix $S(F)$. In our case, when the limiting covariance is non-negative, we can replace this condition by a weaker form of convergence such as the stable convergence. Note that, a priori, the quantitative results of [110] can be modified in order to obtain a result similar to Theorem 3.3.1.5. In fact, using the triangle inequality and selecting $u = DL^{-1}F$ in [110, Theorem 3.4], we obtain the following result, that is the equivalent of our Theorem 3.3.1.5 in the setting of isonormal Gaussian processes.

**Theorem 3.3.2.1.** *We work in the setting of an isonormal Gaussian processes over a separable Hilbert space $\mathscr{H} = \mathscr{L}^2(\nu)$ as described in Section 2.5. Let $(F_n) \in \mathscr{D}\mathrm{om}\, D$ assume that there exists $\Sigma \in \mathscr{L}^0(\Omega)$ such that*

$$S(F_n) \xrightarrow[n\to\infty]{law} \Sigma^2; \tag{3.3.2.2}$$

*and that either one of the following holds*

$$|DL^{-1}F_n|_{\mathscr{H}} \xrightarrow[n\to\infty]{\mathscr{L}^2(\mathfrak{W})} 0; \tag{3.3.2.3a}$$

$$\sup_{n\in\mathbb{N}} \mathbb{E}|DS(F_n)|^2_{\mathscr{H}} < \infty, \tag{3.3.2.3b}$$

*or*

$$\mathbb{E}\int (D_zF_n)^4 \nu(\mathrm{d}z) \xrightarrow[n\to\infty]{} 0; \tag{3.3.2.4a}$$

$$\sup_{n\in\mathbb{N}} \mathbb{E}\int (D_zS(F_n))^{\frac{4}{3}} \nu(\mathrm{d}z) < \infty. \tag{3.3.2.4b}$$

*Then,* $F_n \xrightarrow[n\to\infty]{law} \Sigma N$.

Fourthly, due to the lack of diffusiveness, our conditions in the quantitative theorem Theorem 3.3.1.4 involve $\Sigma$ and not just $\Sigma\Sigma^T$. This adds an extra difficulty because in practice, we only have access to $\Sigma\Sigma^T$ via the convergence of $S(F_n)$. We do not expect that this term can disappear in general.

Lastly, in the multidimensional case, our bound holds for every random matrix $\Sigma$ while in [110], the authors are limited to the case of a diagonal matrix. In [63], the authors were also able to deal with generic matrices but their method relies on the so-called method of the characteristic function that is not known to provide quantitative bounds. Also, in [63], in place of $S(F)$, the authors take the possibly non-symmetric random matrix $M$ whose coefficient $(i,j)$ is given by $\Gamma(L^{-1}F_i, F_j)$. In this case, for $d = 2$ with $F_1 = I_q(f)$ and $F_2 = I_p(f)$ with $p \neq q$ then $M_{12} = p/qM_{21}$. Thus, $M$ cannot be symmetric, hence an acceptable covariance matrix, unless it is diagonal. In many cases, the result of [63] has the same limitation as the one of [110] and that our result Theorem 3.3.1.4 does not have. Note that our result Theorem 3.3.1.5 has this limitation has we could not work with the symmetrized Stein matrix. Remark that for deterministic target covariance, this remark does not apply, as stochastic integrals living in different chaoses are uncorrelated.

3.3.3. **Proof of the lemmas.** We start by recalling the following useful bound for the operator $L^{-1}$. Namely, from [81, Lemma 3.4], we have that, for all $p \geq 1$,

$$r^p(L^{-1}F) \leq r^p(F), \tag{3.3.3.1}$$

By (2.4.2.5) and (2.4.2.6), we also obtain

$$D_z^+F^2 = 2FD_z^+F + (D_z^+F)^2 \quad \text{and} \quad D_z^-F^2 = 2FD_z^-F - (D_z^-F)^2. \tag{3.3.3.2}$$

We now give a proof of the key lemma.

*Proof of Lemma 3.3.1.2.* As usual, we work first with $F$, $G$ and $\Sigma \in \mathscr{A}$. For short, we write $\Phi_2 = |\nabla^2\phi|_{\ell^\infty,\infty}$ and $\Phi_3 = |\nabla^3\phi|_{\ell^\infty,\infty}$. Let $(u_t)_{t\in[0,1]}$ be a smooth $[0,1]$-valued path such that $u_0 = 0$ and $u_1 = 1$ and define

$$F_t = G + u_tF + u_{1-t}\Sigma N. \tag{3.3.3.3}$$

Let $g(t) = \mathbb{E}\phi(F_t)$. Then,

$$\mathbb{E}\phi(F+G) - \mathbb{E}\phi(\Sigma N + G) = \int_0^1 \dot{g}_t \mathrm{d}t. \tag{3.3.3.4}$$

An explicit computation yields

$$\dot{g}_t = \mathbb{E}[\nabla\phi(F_t)\cdot(\dot{u}_t F - \dot{u}_{1-t}\Sigma N)]. \tag{3.3.3.5}$$

Since $\mathscr{A}$ is a linear space, in view of the assumptions, $F_t \in \mathscr{A}$. Since $\nabla\phi$ is Lipschitz, $\nabla\phi(F_t) \in \mathscr{D}\mathrm{om}\, D$. Using the integration by part formula (3.2.2.3) and the pseudo-chain rule Lemma 2.4.3.1, we infer that

$$\mathbb{E}[\nabla\phi(F_t)\cdot F] = \mathbb{E}[\nabla^2\phi(F_t)\cdot\Gamma(L^{-1}F, F_t)] + \sum_i \mathbb{E}R_{\partial_i\phi}(F_t, L^{-1}F_i). \tag{3.3.3.6}$$

Owing to the bi-linearity of $\Gamma$,

$$\Gamma(L^{-1}F, F_t) = \Gamma(L^{-1}F, G) + u_t\Gamma(L^{-1}F, F) + u_{1-t}\Gamma(L^{-1}F, \Sigma N). \tag{3.3.3.7}$$

As $N$ is independent of $\eta$, we have that $D^+(\Sigma N) = (D^+\Sigma)N$ and $D^-(\Sigma N) = (D^-\Sigma)N$. Eventually,

$$\begin{aligned}\mathbb{E}[\nabla\phi(F_t)\cdot F] &= \mathbb{E}[\nabla^2\phi(F_t)\cdot\Gamma(L^{-1}F, G)] + u_t\mathbb{E}[\nabla^2\phi(F_t)\cdot\Gamma(L^{-1}F, F)] \\ &\quad + u_{1-t}\sum_{i,j,k}\mathbb{E}[\partial_{ij}\phi(F_t)N_k\Gamma(L^{-1}F_i, \Sigma_{jk})] + \sum_i \mathbb{E}R_{\partial_i\phi}(F_t, L^{-1}F_i).\end{aligned} \tag{3.3.3.8}$$

Recall that, by integration by parts, $\mathbb{E}N\psi(N) = \mathbb{E}\nabla\psi(N)$, for all smooth $\psi$. Let

$$\psi(x) = \partial_{ij}\phi(G + u_t F + u_{1-t}\Sigma x). \tag{3.3.3.9}$$

Then,

$$\partial_k\psi(x) = u_{1-t}\sum_l \Sigma_{lk}\partial_{ijl}(G + u_t F + u_{1-t}\Sigma x). \tag{3.3.3.10}$$

As a consequence, by the previous Gaussian integration by parts,

$$\mathbb{E}[\partial_{ij}\phi(F_t)N_k\Gamma(L^{-1}F_i, \Sigma_{jk})] = u_{1-t}\sum_l \mathbb{E}[\partial_{ijl}\phi(F_t)\Gamma(L^{-1}F_i, \Sigma_{jk})\Sigma_{lk}]. \tag{3.3.3.11}$$

Similarly, we obtain that

$$\mathbb{E}[\nabla\phi(F_t)\cdot\Sigma N] = u_{1-t}\mathbb{E}[\nabla^2\phi(F_t)\cdot\Sigma\Sigma^T]. \tag{3.3.3.12}$$

Combining (3.3.3.5), (3.3.3.8), (3.3.3.11) and (3.3.3.12), we find that

$$\begin{aligned}\dot{g}_t &= \mathbb{E}\left[\nabla^2\phi(F_t)\cdot\left(\dot{u}_t\Gamma(L^{-1}F, G) + u_t\dot{u}_t\Gamma(L^{-1}F, F) + u_{1-t}\dot{u}_{1-t}\Sigma\Sigma^T\right)\right] \\ &\quad + \dot{u}_t u_{1-t}^2\sum_{ijkl}\mathbb{E}[\partial_{ijk}\phi(F_t)\Gamma(L^{-1}F_i, S_{jl})S_{kl}] + \dot{u}_t\sum_i \mathbb{E}R_{\partial_i\phi}(F_t, L^{-1}F_i).\end{aligned} \tag{3.3.3.13}$$

We subsequently examine carefully the two last terms appearing on the right-hand side of (3.3.3.13). First, we focus on $\sum_i \mathbb{E}R_{\partial_i\phi}(F_t, L^{-1}F_i)$. From the definition of $R_\phi$ in Lemma 2.4.3.1, we derive that

$$\sum_i \mathbb{E}R_{\partial_i\phi}(F_t, L^{-1}F_i) = \frac{1}{2}\sum_{ijk}\mathbb{E}\int D_z^+L^{-1}F_iD_z^+F_{t,j}D_z^+F_{t,k}R_{ijk}^+(z)\nu(\mathrm{d}z) - \frac{1}{2}\sum_{ijk}\mathbb{E}\int D_z^-L^{-1}F_iD_z^-F_{t,j}D_z^-F_{t,k}R_{ijk}^-(z)\eta(\mathrm{d}z), \tag{3.3.3.14}$$

where

$$R_{ijk}^{\pm}(z) = \int_0^1\int_0^1 \alpha\partial_{ijk}\phi(F_t \pm \alpha\beta D_z^{\pm}F_t)\mathrm{d}\alpha\mathrm{d}\beta. \tag{3.3.3.15}$$

Note that $|R_{ijk}(z)| \leq |\partial_{ijk}\phi|_\infty$, and thus, by the Mecke formula,

$$\left|\sum_i \mathbb{E}R_{\partial_i\phi}(F_t, L^{-1}F_i)\right| \leq \sum_{ijk}|\partial_{ijk}\phi|_\infty\int |D_z^+L^{-1}F_i||D_z^+F_{t,j}D_z^+F_{t,k}|. \tag{3.3.3.16}$$

By Young's inequality and the fact that $(a+b)^2 \leq 2(a^2+b^2)$, we have that

$$|D_z^+F_{t,j}D_z^+F_{t,k}| \leq 2\left((D_z^+G_j)^2 + (D_z^+G_k)^2 + u_t^2(D_z^+F_j)^2 + u_t^2(D_z^+F_k)^2 + u_{1-t}^2(D_z^+(\Sigma N)_j)^2 + u_{1-t}^2(D_z^+(\Sigma N)_k)^2\right). \tag{3.3.3.17}$$

Therefore, we obtain that

$$\sum_i \mathbb{E}|R_{\partial_i\phi}(F_t, L^{-1}F_i)| \leq 4\Phi_3\mathbb{E}\int |D_z^+L^{-1}F|_{\ell^1}\left(|D_z^+G|_{\ell^2}^2 + u_t^2|D_z^+F|_{\ell^2}^2 + u_{1-t}^2|D_z^+\Sigma N|_{\ell^2}^2\right)\nu(\mathrm{d}z). \tag{3.3.3.18}$$

Since $N$ is independent of $\eta$ and all the other terms in the previous expression are measurable with respect to $\eta$, by expanding the squares, we find that

$$\sum_i \mathbb{E}|R_{\partial_i\phi}(F_t, L^{-1}F_i)| \leq 4\Phi_3\mathbb{E}\int |D_z^+L^{-1}F|_{\ell^1}(|D_z^+G|_{\ell^2}^2 + u_t^2|D_z^+F|_{\ell^2}^2 + u_{1-t}^2|D_z^+\Sigma|_{\ell^2}^2)\nu(\mathrm{d}z). \tag{3.3.3.19}$$

Now, we focus on $\sum_{ijkl}\mathbb{E}\partial_{ijk}\phi(F_t)\Gamma(L^{-1}F_i, \Sigma_{jl})\Sigma_{kl}$. By (2.4.2.5) and (2.4.2.6),

$$\Gamma(L^{-1}F_i, \Sigma_{jl})\Sigma_{kl} = \Gamma(L^{-1}F_i, \Sigma_{jl}\Sigma_{kl}) - \Gamma(L^{-1}F_i, \Sigma_{jl}, \Sigma_{kl}), \tag{3.3.3.20}$$

where

$$\Gamma(A,B,C) = \frac{1}{2}\int D_z^+AD_z^+BD_z^+C\nu(\mathrm{d}z) - \frac{1}{2}\int D_z^-AD_z^-BD_z^-C\eta(\mathrm{d}z).$$

Thus, by the Mecke formula, we deduce that

$$\left|\sum_{ijkl}\mathbb{E}\partial_{ijk}\phi(F_t)\Gamma(L^{-1}F_i, \Sigma_{jl})\Sigma_{kl}\right| \leq \Phi_3\mathbb{E}\int |D_z^+L^{-1}F|_{\ell^1}(|D_z^+(\Sigma\Sigma^T)|_{\ell^1} + |D_z^+\Sigma D_z^+\Sigma^T|_{\ell^1})\nu(\mathrm{d}z). \tag{3.3.3.21}$$

Notice that, for a matrix $M$,

$$|M|_{\ell^2}^2 = \mathrm{tr}(MM^T) \leq |MM^T|_{\ell^1} \leq |M|_{\ell^1}^2, \tag{3.3.3.22}$$

and that by the Mecke formula and the triangular inequality,

$$\mathbb{E}|\Gamma(L^{-1}F, G)|_{\ell^1} \leq \mathbb{E}\int |D_z^+ L^{-1}F|_{\ell^1} |D_z^+ G|_{\ell^1}. \tag{3.3.3.23}$$

Thus, in view of (3.3.3.19), (3.3.3.21), (3.3.3.22) and (3.3.3.23) choosing $u_t = t^{2/3}$ in (3.3.3.13) yields the bound

$$\begin{aligned} |\mathbb{E}\phi(G+F) - \mathbb{E}\phi(G+\Sigma N)| \leq{}& 2\Phi_2\mathbb{E}|S(F) - \Sigma\Sigma^T|_{\ell^1} \\ &+ \Phi_2\mathbb{E}\int |D_z^+ L^{-1}F|_{\ell_1} |D_z^+ G|_{\ell_1} \nu(\mathrm{d}z) \\ &+ \Phi_3\mathbb{E}\int |D_z^+ L^{-1}F|_{\ell^1} |D_z^+(\Sigma\Sigma^T)|_{\ell^1} \nu(\mathrm{d}z) \\ &+ 5\Phi_3\mathbb{E}\int |D_z^+ L^{-1}F|_{\ell^1} |D_z^+\Sigma|_{\ell^1}^2 \nu(\mathrm{d}z) \\ &+ 4\Phi_3\mathbb{E}\int |D_z^+ L^{-1}F|_{\ell^1} |D_z^+ G|_{\ell^2}^2 \nu(\mathrm{d}z) \\ &+ \frac{4}{3}\Phi_3\mathbb{E}\int |D_z^+ L^{-1}F|_{\ell^1} |D_z^+ F|_{\ell^2}^2 \nu(\mathrm{d}z). \end{aligned} \tag{3.3.3.24}$$

Indeed, the reader can immediately verify that with such a choice for $u$, we have that (3.3.3.25)

$$\int_0^1 \dot{u}_t = 1; \int_0^1 u_t\dot{u}_t \mathrm{d}t = \int_0^1 u_{1-t}\dot{u}_{1-t} = 2; \int_0^1 \dot{u}_t u_{1-t}^2 = \frac{4}{9}\frac{\pi}{\sqrt{3}} \leq 1; \int_0^1 u_t^2\dot{u}_t = \frac{1}{3}.$$

Thereby, we obtain (3.3.1.27). By applying the Cauchy-Schwarz inequality in (3.3.1.27), we obtain immediately Lemma 3.3.1.2. To obtain (3.3.1.28), we apply Hölder's inequality and (3.3.3.1). The proof is complete. □

Finally, we conclude the section with the proof of Lemma 3.3.1.3.

*Proof of Lemma 3.3.1.3.* Let $\phi \in \mathscr{C}^2(\mathbb{R})$ be 1-Lipschitz on $\mathbb{R}^d$ with $|\nabla^2\phi|_{op} \leq 1$. For $\epsilon \in [0,1]$, we write

$$\phi_\epsilon(x) = \mathbb{E}\phi(\sqrt{1-\epsilon}x + \sqrt{\epsilon}N), \quad x \in \mathbb{R}^d. \tag{3.3.3.26}$$

By Gaussian integration by parts, we have that

$$\partial_j\phi_\epsilon(x) = \sqrt{1-\epsilon}\mathbb{E}\partial_j\phi_\epsilon(x) = \sqrt{\frac{1-\epsilon}{\epsilon}}\mathbb{E}N_j\phi_\epsilon(x). \tag{3.3.3.27}$$

In a way that,

$$\partial_{ij}\phi_\epsilon(x) = \frac{\epsilon}{\sqrt{1-\epsilon}}\mathbb{E}N_j\partial_i\phi_\epsilon(x). \tag{3.3.3.28}$$

Hence, we conclude that

$$|\nabla^2\phi_\epsilon|_{\ell^2,\infty} \leq \frac{\epsilon}{\sqrt{1-\epsilon}}\sqrt{d} \leq \sqrt{d}\epsilon^{-1}. \tag{3.3.3.29}$$

In the same way (see [127] for details), we prove that

$$|\nabla^3\phi_\epsilon|_{\ell^2,\infty} \leq \frac{\sqrt{2\pi}}{4}\frac{\epsilon^{\frac{3}{2}}}{\sqrt{1-\epsilon}}\sqrt{d} \leq \frac{\sqrt{2\pi}}{4}\sqrt{d}\epsilon^{-1}. \tag{3.3.3.30}$$

Also we have that

$$|\phi(x) - \phi_\epsilon(x)| \leq \sqrt{\epsilon}(|N|_{\ell^2} + |x|_{\ell^2}). \tag{3.3.3.31}$$

Eventually, we obtained

$$|\nabla^2\phi_\epsilon|_{\ell^2,\infty} \leq \sqrt{d}\epsilon^{-1}; \tag{3.3.3.32a}$$

$$|\nabla^3\phi_\epsilon|_{\ell^2,\infty} \leq \frac{\sqrt{2\pi}}{4}\sqrt{d}\epsilon^{-1}; \tag{3.3.3.32b}$$

$$|\mathbb{E}\phi(G) - \mathbb{E}\phi_\epsilon(G)| \leq \epsilon^{\frac{1}{2}}\left(\sqrt{d} + \mathbb{E}|G|_{\ell_2}\right). \tag{3.3.3.32c}$$

By independence, $\mathbb{E}|\Sigma N|^2_{\ell^2} = \mathbb{E}|\Sigma|^2_{\ell^2}$. Thus, by Jensen's inequality, $\mathbb{E}|\Sigma N|_{\ell^2} \leq \sqrt{\mathbb{E}|\Sigma|^2_{\ell^2}}$. Hence, applying the triangle inequality and the bounds (3.3.1.30) and (3.3.3.32a) yields:

$$|\mathbb{E}\phi(F) - \mathbb{E}\phi(\Sigma N)| \leq \frac{\sqrt{2\pi}}{4}\sqrt{d}\epsilon^{-1}(a+b) + \sqrt{\epsilon}\left(2\sqrt{d} + \mathbb{E}|F|_{\ell^2} + \sqrt{\mathbb{E}|\Sigma|^2_{\ell^2}}\right). \tag{3.3.3.33}$$

The function $\mathbb{R} \ni \epsilon \mapsto \sqrt{\epsilon}u + \epsilon^{-1}v$ attains its minimum at $\epsilon_0 = \left(2\frac{v}{u}\right)^{2/3}$, where it takes the value $(2^{1/3} + 2^{-2/3})u^{2/3}v^{1/3}$. If $\epsilon_0 \leq 1$, we choose $\epsilon = \epsilon_0$ and we choose $\epsilon = 1$ otherwise. This yields the announced inequality.

□

## 3.4. Stable limit theorems for stochastic integrals

**Outline.** We explicit our main bound when the functional $F$ has finite chaotic decomposition. These computations are useful to deduce Proposition 3.4.1.1, that is a stable version of the fourth moment theorem of C. Döbler & G. Peccati (2018) [40] and C. Döbler, A. Vidotto & G. Zheng (2018) [41] and Proposition 3.4.2.2 that gives sufficient condition for a Poisson $U$-statistics of order 2 to converge to a Gaussian mixture. The crucial result is Theorem 2.4.2.4. It shows that our definition of the carré du champ in terms of $D^+$ and $D^-$ coincides with the usual representation in term of the generator of the Ornstein-Uhlenbeck semigroup à la Bakry-Emery [13, Section 1.4.2].

### 3.4.1. A stable fourth-moment theorem.

In a recent reference, C. Döbler, A. Vidotto & G. Zheng (2018) [41] proved a multidimensional fourth-moment theorem on the Poisson space, thus refining and generalizing the previous findings of C. Döbler & G. Peccati (2018) [40]. It is worth noting that taking $G = 0$ and $S$ deterministic in Lemma 3.3.1.2 yields the same bound as [40, Equation 4.2], that was also obtained in Theorem 2.4.4.2 with less restrictive assumptions on the functional $F$: the fact that we can achieve a statement with optimal assumptions is due to the fact that we use Theorem 2.4.2.4. In fact, as a crucial application of Theorem 3.3.1.4, we deduce a stable fourth-moment theorem on the Poisson space.

**Proposition 3.4.1.1** (Stable fourth-moment theorem). *For $n \geq 1$, let $(h_n^i) \subset \mathscr{L}^2(\nu^{q_i})$ and let $F_n = (I_{q_1}(h_n^1), \dots, I_{q_d}(h_n^d))$. Assume $F \in \mathscr{L}^4(\mathfrak{W})^d$. Let $\sigma$ be a deterministic $d \times d$ matrix. If:*

$$\mathbb{E} F_n F_n^T \xrightarrow[n\to\infty]{} \sigma\sigma^T; \tag{3.4.1.1}$$

$$\mathbb{E}(F_n^i)^4 \xrightarrow[n\to\infty]{} 3\big(\sigma\sigma_{ii}^T\big)^2,\ \forall i = 1, \dots, d. \tag{3.4.1.2}$$

*Then, $(F_n)$ converges stably to $\sigma N$.*

**Remark 20.** Proposition 3.4.1.1 is very close to [26, Theorem 2.22]. However, one condition of their theorem requires that the norms of each of the individual star-contractions vanish. This is strictly stronger than a fourth-moment converging to $3$ times the square of the second moment as, by the product formula, this condition translates in vanishing properly chosen linear combinations of the star-contractions (see for instance, [39]). It follows particularly that the statement of Proposition 3.4.1.1 is outside of the scope of [26].

*Proof.* We apply Proposition 3.3.1.1 with $\Sigma = \sigma$. First of all, $(I_{q_i}(f_n^i)) \subset \mathscr{D}\mathrm{om}\, L \subset \mathscr{D}\mathrm{om}\, D$. Let us check that (3.3.1.5) is satisfied. Moreover, by the isometry property of stochastic integrals,

$$\mathbb{E}\gamma^2(F_{n,i}) = \sum_i q_i q_i! \nu^{q_i}({h_n^i}^2) \leq \max q_i \mathbb{E}(F_{n,i})^2, \quad i \in [d]. \tag{3.4.1.3}$$

This quantity is bounded since, by assumption,

$$\mathbb{E}|F_n|^2 = \operatorname{tr} \mathbb{E} F_n F_n^T \xrightarrow[n\to\infty]{} \operatorname{tr} \sigma\sigma^T. \tag{3.4.1.4}$$

This shows (3.3.1.5b). From [40, Lemma 3.2] or Theorem 2.4.4.3, we know that when $F = I_q(f)$,

$$\mathbb{E}\Delta(F) \leq (4q-3)\Big(\mathbb{E}F^4 - 3(\mathbb{E}F^2)^2\Big). \tag{3.4.1.5}$$

Hence, by assumption

$$\Delta(F_{n,i}) \xrightarrow[n\to\infty]{} 0, \quad i \in [d]. \tag{3.4.1.6}$$

This shows (3.3.1.5a).

In order to conclude with Theorem 3.3.1.5, we are left to show (3.3.1.3), namely

$$\Gamma(L^{-1}F_{n,i}, F_{n,j}) \xrightarrow[n\to\infty]{\mathscr{L}^1(\mathfrak{W})} c_{ij}, \tag{3.4.1.7}$$

where $C = \sigma^T\sigma$. But we know from Theorem 2.4.4.3, that under our assumptions,

$$\mathbb{E}|\Gamma(L^{-1}F_{n,i}, F_{n,j}) - v_{i,j}^{(n)}| \xrightarrow[n\to\infty]{} 0, \tag{3.4.1.8}$$

where $v^{(n)} = \mathbb{E}F_n^T F_n$. This concludes the proof,as we assumed that $v^{(n)} \xrightarrow[n\to\infty]{} c$. $\square$

**Remark 21.** Let $\Sigma = (\Sigma_{ij}) \in \mathscr{D}\mathrm{om}\, D$ satisfying $r^3(\Sigma) + r^{8/3}(\Sigma) + r^4(\Sigma\Sigma^T) < \infty$. If we assume that

$$S(F_n) \xrightarrow[n\to\infty]{\mathscr{L}^1(\mathfrak{W})} \Sigma\Sigma^T; \tag{3.4.1.9}$$

$$\mathbb{E}(F_n^i)^4 \xrightarrow[n\to\infty]{} 3\big(\mathbb{E}(\Sigma\Sigma^T)_{ii}\big)^2,\ \forall i = 1, \ldots, d. \tag{3.4.1.10}$$

Then, from the previous computations, $\Sigma = \sigma$ is deterministic. This shows that the fourth-moment theorems cannot capture phenomena with asymptotic random variances.

3.4.2. **A threshold for non-normality of $U$-statistics of order 2: a tentative statement.** In view of the results of this chapter, and of the fact that our findings allow us to retrieve a stable version of the multidimensional fourth moment theorem, it is tempting to use our general stable convergence results in order to find conditions, for a sequence of $U$-statistics to converge towards a mixture of a Gaussian random variable -such a result might have applications, for instance in stochastic geometry. We will prove these results -that seem to us very natural- in Propositions 3.4.2.1 and 3.4.2.2. However, our subsequent discussion shows that these findings fail to capture some elementary stable convergence results on the Poisson space. Recall that a $U$-statistic of order 2 is simply a random variable of the form $\eta^{(2)}(f)$ for some $f \in \mathscr{L}^1_\sigma(\nu^2) \cap \mathscr{L}^2_\sigma(\nu^2)$ and that from the definition of stochastic integrals (2.7.2.13), we have (see also Section 2.7.3)

$$\eta^{(2)}(f) = I_2(f) + 2I_1(\tilde{f}) + \nu^2(f), \tag{3.4.2.1}$$

where

$$\tilde{f} = \int f(z,\cdot)\nu(\mathrm{d}z). \tag{3.4.2.2}$$

We, thus, immediately see that $\mathbb{E}\eta^{(2)}(f) = \nu^2(f)$ and $\mathbb{V}\mathrm{ar}\,(\eta^2(f)) = \nu^2(f^2) + \nu(\tilde{f}^2)$ and that there are two competing terms in $\eta^{(2)}(f) - \nu^2(f)$. The next two statements show that when the term $I_1(\tilde{f})$ dominates, the $U$-statistics typically exhibit a Gaussian behaviour, while when $I_2(f)$ dominates, the typical behaviour is close to those of a Gaussian mixture. For conciseness, we write $\mathscr{K}_q = \mathscr{L}^1_\sigma(\nu^q) \cap \mathscr{L}^2_\sigma(\nu^q) \cap \mathscr{L}^4_\sigma(\nu^q)$, for all $q \in \mathbb{N}$.

**Proposition 3.4.2.1.** *Let $(f_n)_{n\in\mathbb{N}} \subset \mathscr{K}_2$. Let $F_n = \eta^{(2)}(f_n) - \nu^2(f_n)$. Let $N$ be a standard univariate normal random variable independent of $\eta$. If $\nu^2(f_n^2) = o(\mathbb{E}F_n^2)$ and $\nu(\tilde{f}_n^4) = o\left(\nu(\tilde{f}_n^2)^2\right)$. Then,*

$$\frac{F_n}{\sqrt{\mathbb{E}F_n^2}} \xrightarrow[n\to\infty]{\text{stably}} N. \tag{3.4.2.3}$$

**Proposition 3.4.2.2.** *Let $(f_n)_{n\in\mathbb{N}} \subset \mathscr{K}_2$. Let $F_n = \eta^{(2)}(f_n) - \nu^2(f_n)$. Assume that*

$$\nu(\tilde{f}_n^2) = o(\mathbb{E}F_n^2); \tag{3.4.2.4}$$

$$\nu^2(f_n^4) = o(\nu^2(f_n^2)^2); \tag{3.4.2.5}$$

$$\nu((f_n \star_2^1 f_n)^2) = o(\nu^2(f_n)^2); \tag{3.4.2.6}$$

*and that there exists $g_\infty \in \mathscr{L}^2(\nu^2)$ such that*

$$\frac{1}{\nu^2(f_n^2)} f_n \star_1^1 f_n \xrightarrow[n\to\infty]{\mathscr{L}^2(\nu)} g_\infty. \tag{3.4.2.7}$$

*Let*

$$C = 2(I_2(g_\infty) + 1). \tag{3.4.2.8}$$

*Then, $C \geq 0$ (almost surely), and with $\Sigma = C^{\frac{1}{2}}$, we have that*

$$\frac{F_n}{\sqrt{\mathbb{E}F_n^2}} \xrightarrow[n\to\infty]{\text{stably}} \Sigma N. \tag{3.4.2.9}$$

**Remark 22.** Our proposition is in principle well-adapted to a setting of geometric random graphs, where $\nu$ is the Lebesgue measure and $f_n$ is some $\{0, 1\}$-valued kernels such that $\nu(f_n) \to \infty$. In this case, it is immediate that (3.4.2.5) always holds in that case.

*Proof of Proposition 3.4.2.1.* Without loss of generality, we can assume that $\mathbb{E}F_n^2 = 1$, for all $n \in \mathbb{N}$, that is we assume that $\nu(\tilde{f}_n^2) = 1$. We readily apply our stable fourth moment theorem on the Poisson space, that is our Proposition 3.4.1.1. By the assumption on the variance, we have that

$$I_2(f_n) \xrightarrow[n\to\infty]{\mathscr{L}^2(\mathfrak{W})} 0. \tag{3.4.2.10}$$

By the product formula (3.2.1.12) and the isometry property of stochastic integrals, we have that

$$\mathbb{E}I_1(\tilde{f}_n)^2 = \nu(\tilde{f}_n^2) \quad \text{and} \quad \mathbb{E}I_1(\tilde{f}_n)^4 = 3\nu^2(\tilde{f}_n)^2 + \nu(\tilde{f}_n^4). \tag{3.4.2.11}$$

We conclude by letting $n \to \infty$ in the previous expression and invoking our stable fourth moment theorem Proposition 3.4.1.1. □

*Proof of Proposition 3.4.2.2.* Again without loss of generality, we assume that $\mathbb{E}F_n^2 = 1$ for all $n \in \mathbb{N}$, that is we assume that $\nu^2(f_n^2) = 1$. By the condition on the variance (3.4.2.4), we have that

$$I_1(\tilde{f}_n) \xrightarrow[n\to\infty]{\mathscr{L}^2(\mathfrak{W})} 0. \tag{3.4.2.12}$$

In order to obtain the announced result it is, thus, sufficient to obtain the result for $F_n = I_2(f_n)$. From Theorem 2.4.2.4 and the product formula (3.2.1.12), we obtain

$$S(F_n) = \frac{1}{2} I_3(f_n \star_1^0 f_n) + I_2\left(f_n \star_1^1 f_n + \frac{1}{2} f_n^2\right) + \frac{3}{2} I_1(f_n \star_2^1 f_n) + \nu^2(f_n^2). \tag{3.4.2.13}$$

Let us show that under our assumptions some of the stochastic integrals appearing in $S(F_n)$ vanish as $n \to \infty$. First of all, observe that, by isometry, $\mathbb{E}I_2(f_n^2)^2 = 2\nu(f_n^4)$ so that this integral vanishes. We claim that for a kernel $f$

$$\nu^3\left((f \star_1^0 f)^2\right) = \int \left(\int f(z,x)^2 \nu(\mathrm{d}x)\right)^2 \nu(\mathrm{d}z) = \nu\left((f \star_2^1 f)^2\right). \tag{3.4.2.14}$$

Indeed,

$$f \star_1^0 f(x,y,z) = f(x,z)f(y,z), \tag{3.4.2.15}$$

and therefore

$$\nu^3\left((f \star_1^0 f)^2\right) = \int f(x,z)^2 f(y,z)^2 \nu(\mathrm{d}z)\nu(\mathrm{d}x)\nu(\mathrm{d}y). \tag{3.4.2.16}$$

This proves the first equality by Fubini's theorem for non-negative integrands. For the second inequality, observe that

$$f \star_2^1 f(x) = \int f(z,x)^2 \nu(\mathrm{d}z). \tag{3.4.2.17}$$

Taking squares and integrating yields the second inequality and proves (3.4.2.14). Therefore, by (3.4.2.6) and isometry, both $I_3(f_n \star_1^0 f_n)$ and $I_1(f_n \star_2^1 f_n)$ vanish. Using the vanishing of the aforementioned integrals, (3.4.2.7) and isometry, we proved that

$$S(F_n) \xrightarrow[n\to\infty]{\mathscr{L}^2(\mathfrak{W})} I_2(g_\infty) + 1 = C. \tag{3.4.2.18}$$

Since $S(F_n) = \frac{1}{2}\Gamma(F_n) \geq 0$, this shows that $C \geq 0$. In order to conclude we will invoke Proposition 3.3.1.1. Let us show that $\Delta(F_n)$ vanishes. Since $D = D^+$, we have that

$$\left(D_z^+ I_2(f_n)\right)^4 = (I_1(f_n(z,\cdot)))^2 (I_1(f_n(z,\cdot)))^2, \quad z \in Z, \tag{3.4.2.19}$$

and by the isometry property of Poisson integrals, one gets for $z \in Z$:

$$\begin{aligned} \mathbb{E}(D_z^+ F_n)^4 &= 4\nu^2\left((f_n(z,\cdot)\otimes f_n(z,\cdot))^2\right) + \nu\left(f_n(z,\cdot)^4\right) + (f_n \star_2^1 f)(z)^2 \\ &= 5(f_n \star_2^1 f_n)(z)^2 + \nu\left(f_n(z,\cdot)^4\right). \end{aligned} \tag{3.4.2.20}$$

Consequently,

$$\mathbb{E}\Delta(I_2(f_n)) = 5\nu^2\left((f_n \star_2^1 f_n)^2\right) + \nu^2\left(f_n^4\right), \tag{3.4.2.21}$$

that vanishes by (3.4.2.5) and (3.4.2.6). Similarly, we have that

$$\mathbb{E}\gamma^2(F_n) = \nu^2\left(f_n^2\right) + 4\nu\left(\tilde{f}_n^2\right). \tag{3.4.2.22}$$

The right-hand side of the previous expression is bounded by (3.3.1.3) and the fact that we assumed $\nu^2(f_n^2) = 1$. This concludes the proof. □

**Remark 23.** In view of the conclusion of the very natural question of what happens "at criticality", that is when the two terms $I_1(\tilde{f})$ and $I_2(f)$ are of the same order. Indeed our theorem could theoretically be applied to study the joint convergence of $(I_2(f_n), 2I_1(\tilde{f}_n))$. In this case, the reader can check that the limit of the random covariance would be

$$M_\infty = \begin{pmatrix} \int I_1(f_\infty(z,\cdot))^2\nu(\mathrm{d}z) & 3I_1\left(f_\infty \star_1^1 \tilde{f}_\infty\right) \\ 3I_1\left(f_\infty \star_1^1 \tilde{f}_\infty\right) & 4\nu\left(\tilde{f}_\infty^2\right) \end{pmatrix}. \tag{3.4.2.23}$$

And we would have to check that $M_\infty$ is always non-negative (or to be able to deal with negative covariance), a task that we cannot accomplish at the moment.

Remark, that in general, we do not expect the condition (3.4.2.7) to hold. Let us quickly explain why on the following example. We were not able to completely work out this example. The proof in the Poisson fails for a technical detail as we explain below. For technical reasons, we will also sometimes work in the Gaussian setting. In the following $W$ is a isonormal Gaussian process over the separable Hilbert space $\mathscr{H} = \mathscr{L}^2(\nu)$. In order to distinguish between the Poisson and the Gaussian case we will write $\Gamma^W$ and $I_q^W$ for the carré du champ and the stochastic integrals with respect to $W$.

3.4.3. **An example.** We let $(A_n)$ a collection of measurable sets such that $\nu(A_n) \xrightarrow[n\to\infty]{} \infty$. We let $h_n = 1_{A_n} \otimes 1_{A_n}$ and $f_n = \frac{h_n}{\sqrt{\nu^2(h_n^2)}}$. By the classical central limit theorem, we have that

(3.4.3.1) $$\eta^{(2)}(f_n) - \nu(f_n^2) \xrightarrow[n\to\infty]{law} N^2 - 1.$$

In that case, we have that

(3.4.3.2a) $$\nu^2(h_n^2) = \nu(A_n)^2;$$

(3.4.3.2b) $$\tilde{h}_n(x) = \nu(A_n)1_{A_n} = h_n \star_2^1 h_n;$$

(3.4.3.2c) $$\nu(\tilde{h}_n^2) = \nu(A_n)^3.$$

Thus, we see that (3.4.2.4), (3.4.2.5) and (3.4.2.6) are satisfied. However, we have that

(3.4.3.3) $$h_n \star_1^1 h_n = \nu(A_n) h_n.$$

Since $(f_n)$ converges to $0$ in $\mathscr{L}^4(\mathfrak{W})$, this shows that (3.4.2.7) cannot hold unless $g_\infty = 0$. However, with $F_n = I_2(f_n)$, we have that $\mathbb{E}F_n^2 = 1$. If now, we work with a subsequence converging in law to $F_\infty$, in view of (3.4.2.13), (3.4.3.2) and (3.4.3.3), we obtain that

(3.4.3.4) $$S(F_n) = I_2(f_n) + \nu^2(f_n^2) + \epsilon_n = F_n + 1 + \epsilon_n,$$

where $\epsilon_n \xrightarrow[n\to\infty]{\mathscr{L}^2(\mathfrak{W})} 0$. Hence, (3.4.2.18) is replaced by

(3.4.3.5) $$S(F_n) \xrightarrow[n\to\infty]{law} (F_\infty + 1).$$

Provided we can apply Theorem 3.3.1.5, we would have proved that $F_\infty$ satisfies

(3.4.3.6) $$law(F_\infty) = law(\sqrt{(F_\infty + 1)}N),$$

where $N \sim \mathbf{N}(0,1)$ independent of $\eta$. Remark that we do not know if (3.4.3.6) admits solution. In view of the computations for the proof of Proposition 3.4.2.2, to apply Theorem 3.3.1.5 we just have to check that we can control the quantities involving $S^{\frac{1}{2}}(F_n)$ in (3.3.1.35b). A task that we cannot accomplish. However, if we work rather with $F_n = I_2^W(f_n)$, we have that

(3.4.3.7) $$S(F_n) = F_n + 1 + \epsilon_n,$$

where $\epsilon_n \xrightarrow[n\to\infty]{\mathscr{L}^2(\mathfrak{W})} 0$. This shows that (3.3.2.4) is satisfied and, by Theorem 3.3.2.1, we obtain that

(3.4.3.8) $$I_2^W(f_n) \xrightarrow[n\to\infty]{} F_\infty,$$

where $F_\infty$ is solution of (3.4.3.6).

**Outline.** In Section 3.5.1, motivated by the regularity properties of the sample paths of the fractional Brownian motion (a continuous Gaussian process whose definitions and properties are collected below), we state Lemma 3.5.1.1 that, informally speaking, accounts for the weak regularity in law of a generic Gaussian process only in terms of its covariance function. From this lemma we deduce, with a straightforward proof, Theorem 3.5.1.2 that is a result on the asymptotic behaviour of quadratic functionals of Gaussian processes that were already obtained in [126] in the case of the Brownian motion, in [110] when the underlying Gaussian process if a regular fractional Brownian motion (see below for definitions) via a combination of Malliavin-Stein techniques and Itô's formula in law, and in [131] for a class of Gaussian processes satisfying some analytical condition (that includes the fractional Brownian motion of any Hurst parameters) via demanding computations. The Corollary 3.5.1.3 is obtained by specializing Theorem 3.5.1.2 to the fractional Brownian motion. In Section 3.5.2, we describe a natural Poisson-based counterpart of Gaussian processes with stationary increments, relying on the spectral decomposition of such processes via a white noise integration. All the needed material is recalled below. Then, we state and prove Theorem 3.5.2.3 giving conditions on the approximation such that the conclusion of Theorem 3.5.1.2 still holds when the Gaussian process (with stationary increments) is replaced by the Poisson approximation introduced in Section 3.5.2. This approximation procedure breaks the normality of some of the objects and the proof is not as straightforward as in the Gaussian case and thus requires Theorems 3.3.1.4 and 3.3.1.5 previously developed.

3.5.1. **Convergence of Gaussian functionals.** Understanding the path-wise behaviour of Gaussian processes has been one of the preeminent problems in the modern theory of probability, and is still very relevant nowadays. In this respect, the pioneering result in this field is the well-known Kolmogorov's continuity criterion [36, Theorem XXIII.19]. For Gaussian processes, this result links the properties of the covariance kernel with the path-wise Hölder-regularity of the process (or rather an equivalent version of it). See for instance [108, Lemma 1.1].

We start by presenting a family of Gaussian processes with stationary increments whose explicit covariance allows us to derive many properties: the fractional Brownian motion. Following [115, Section 5] or [108], the fractional Brownian motion with Hurst parameter $h \in (0,1)$ is the centered Gaussian process $B$ indexed by $\mathbb{R}_+$ with covariance

$$2\mathbb{E}B_tB_s = s^{2h} + t^{2h} - |t-s|^{2h},\ s,\ t \in \mathbb{R}_+. \tag{3.5.1.1}$$

The fractional Brownian motion of Hurst parameter $h = 1/2$ is the standard Brownian motion on $\mathbb{R}_+$. When $h \neq 1/2$, the fractional Brownian motion $B$ is not a semimartingale nor a Markov process nor has independent increments however it has stationary increments.

The aforementioned Kolmogorov criterion can be used to deduce that the fractional Brownian motion with Hurst parameter $h$ is almost surely $\alpha$-Hölder continuous for every $\alpha \in (0,h)$ [108, Proposition 1.6]. It is also straightforward to check that the sample paths of the fractional Brownian motion with Hurst parameter $h$ are almost surely not $\alpha$-Hölder continuous for all $\alpha \geq h$ and, in particular, they are, almost surely not $h$-Hölder. However, from the stationary increments and the self-similarity of the fractional Brownian motion [108, Proposition 2.2], we easily see that

$$\frac{B_s - B_t}{|s-t|^h} \sim \mathbf{N}(0,1), \tag{3.5.1.2}$$

where $B$ is a fractional Brownian motion with Hurst parameter $h$. By this argument and the change of variable $t \mapsto \mathrm{e}^{-u/n}$, we obtain that

(3.5.1.3)
$$\begin{aligned} n^{1+h}\int_0^1 t^{n-1}(B_1 - B_t)\mathrm{d}t &= \int_0^\infty \mathrm{e}^{-u}\left(n\left(1-e^{-u/n}\right)\right)^h (B_1 - B_{\mathrm{e}^{-u/n}})\left(1-\mathrm{e}^{-u/n}\right)^{-h} \\ &\xrightarrow[n\to\infty]{law} \Gamma(1+h)\mathbf{N}(0,1). \end{aligned}$$

The limit is justified since for positive $u$, as $n \to \infty$, $n(1-\mathrm{e}^{-u/n}) \to u^{-1}$ with a decreasing convergence and by the monotone convergence theorem. Hence, the previous convergence can be seen as a very weak Hölder property of the sample paths of the fractional Brownian motion. We now investigate a similar behaviour for generic Gaussian process.

Let $X = (X_t)_{t\in\mathbb{R}_+}$ be a Gaussian process with continuous covariance function $R(s,t) = \mathbb{E}X_sX_t$. We introduce the following quantities:

$$\begin{aligned} a_n^2 &= \mathbb{E}\left[\left(\int_0^1 t^{n-1}(X_1 - X_t)\mathrm{d}t\right)^2\right] \\ &= \int_0^1 s^{n-1}t^{n-1}(R(1,1) - R(t,1) - R(s,1) + R(t,s))\mathrm{d}t\mathrm{d}s; \end{aligned} \tag{3.5.1.4}$$

$$\begin{aligned} b_n &= \int t^{n-1}\mathbb{E}|X_1 - X_t|^2\mathrm{d}t \\ &= \int t^{n-1}(R(1,1) - 2R(t,1) + R(t,t)); \end{aligned} \tag{3.5.1.5}$$

$$\begin{aligned} f_n(s) &= \int t^{n-1}\mathbb{E}X_s(X_1 - X_t) \\ &= \int t^{n-1}(R(1,s) - R(t,s))\mathrm{d}t. \end{aligned} \tag{3.5.1.6}$$

First of all notice the following observation.

**Lemma 3.5.1.1.** *Assume that the sequence $(a_nf_n(s))$ converges to $0$ for all $s \in [0,1]$. Then,*

$$a_n^{-1}\int_0^1 t^{n-1}(X_1 - X_t)\mathrm{d}t \xrightarrow[n\to\infty]{stably} N, \tag{3.5.1.7}$$

*where $N$ is a standard Gaussian random variable independent of $(X_t)_{t\in[0,1]}$.*

*Proof.* Fix $n \in \mathbb{N}$. We let

$$G_n = a_n^{-1}\int_0^1 t^{n-1}(X_1 - X_t). \tag{3.5.1.8}$$

Since $X$ is a Gaussian process with covariance kernel given by $R$, by definition of $a_n$ in (3.5.1.4), the random variable $G_n$ is a standard Gaussian for every $n \in \mathbb{N}$. Hence, the only thing left to check is the asymptotic independence. Observe that, by definition of $f_n$ in (3.5.1.6), for all $s \in \mathbb{R}_+$,

$$\mathbb{E}G_n X_s = f_n(s). \tag{3.5.1.9}$$

By the fact that for Gaussian variables non-correlation is equivalent to independence, by the assumption on $(a_n f_n(s))$, we have that $a_n^{-1} G_n$ is asymptotically independent of $(X_t)_{t\in[0,1]}$. □

From this observation we can recover more refined results such as the ones of [126, 110, 131]. Note that for the sake of brevity, we only provide a qualitative result. We could obtain the quantitative result from [110, Theorem 3.4].

**Theorem 3.5.1.2.** *Let the previous notations prevail. Assume that the sequences $(a_n f_n(s))$ ($s \in [0,1]$) and $(a_n^{-1} b_n)$ converge to $0$, where $a$ and $b$ are defined in (3.5.1.4) and (3.5.1.5). Let*

$$F_n = \int_0^1 t^{n-1}(X_1^2 - X_t^2)\mathrm{d}t. \tag{3.5.1.10}$$

*Then,*

$$a_n^{-1} F_n \xrightarrow[n\to\infty]{\text{stably}} X_1 N, \tag{3.5.1.11}$$

*where $N$ is a standard Gaussian random variable independent of $X$ and the stable convergence is with respect $\sigma(X)$.*

*Proof.* We write

$$\begin{aligned} F_n &= 2X_1 \int t^{n-1}(X_1 - X_t)\mathrm{d}t - \int t^{n-1}(X_1 - X_t)^2 \mathrm{d}t \\ &= 2X_1 G_n - H_n, \end{aligned} \tag{3.5.1.12}$$

where

$$G_n = \int t^{n-1}(X_1 - X_t)\mathrm{d}t \tag{3.5.1.13}$$

$$H_n = \int t^{n-1}(X_1 - X_t)^2 \mathrm{d}t. \tag{3.5.1.14}$$

We compute $\mathbb{E}H_n = b_n$ so that $(a_n^{-1} H_n)$ converges to $0$ in $\mathscr{L}^1(\mathbb{P})$. We conclude by using Lemma 3.5.1.1. □

Finally, we obtain the announced result on fractional Brownian motion. Note that [110] could only deal with the case $h \geq \frac{1}{2}$.

**Corollary 3.5.1.3.** *Let $h \in (0,1)$ and $B$ be the fractional Brownian motion with Hurst parameter $h$. Then,*

$$\frac{n^{1+h}}{2} \int_0^1 t^{n-1}(B_1^2 - B_t^2)\mathrm{d}t \xrightarrow[n\to\infty]{\text{stably}} (h\Gamma(2h))^{1/2} B_1 N, \tag{3.5.1.15}$$

*where $N$ is a standard Gaussian random variable independent of $B$.*

*Proof.* Direct computations of $a_n$, $b_n$ and $f_n(s)$ and their asymptotic using the explicit covariance of the fractional Brownian motion. □

3.5.2. **Poisson-based approximation of continuous Gaussian processes.** We consider a centered Gaussian process $X = (X_t)_{t\in\mathbb{R}_+}$. We moreover assume that there exists a $\sigma$-finite measure $\nu$ on $\mathbb{R}_+$ and two families of function $(\phi_t^{re})_{t\in\mathbb{R}_+}$ and $(\phi_t^{im})_{t\in\mathbb{R}_+} \subset \mathscr{L}^2(\nu)$ such that

$$X_t = W_{re}(\phi_t^{re}) + W_{im}(\phi_t^{im}), \tag{3.5.2.1}$$

where $W_{re}$ and $W_{im}$ are two independent isonormal Gaussian processes over $\mathscr{L}^2(\nu)$ (that is, a $\nu$-white noise). In view of Examples 2.5.1.4 and 2.5.1.5 this is in particular the case if $X$ is a stationary process or a process with stationary increments such that $X_0 = 0$. In particular, we have that

$$R(t,s) = \nu(\phi_t^{re}\phi_s^{re} + \phi_t^{im}\phi_s^{im}). \tag{3.5.2.2}$$

We let $\eta_{re}^{\lambda}$ and $\eta_{im}^{\lambda}$ be two independent Poisson point processes with intensity measure $\lambda\nu$ ($\lambda > 0$). We define the *Poisson approximation of the Gaussian process* $X$

$$X_t^{\lambda} = \lambda^{-1/2} I_1^{\eta_{re}^{\lambda}}(\phi_t^{re}) + \lambda^{-1/2} I_1^{\eta_{im}^{\lambda}}(\phi_t^{im}), \tag{3.5.2.3}$$

where $I_1^{\eta}$ denotes the stochastic integral of order one with respect to a Poisson point process $\eta$ as defined in Section 2.7.3. As usual, the space $\mathscr{C}(0,1)$ of continuous functions on the unit interval is equipped with the norm of the uniform convergence. For continuous processes, we consider the corresponding convergence in law. We will show that $X^{\lambda}$ is an approximation of $X$. The main argument is that the moments of $X^{\lambda}$ are compatible as $\lambda \to \infty$ with Gaussian moments. It is a classical fact, see for instance [124, Chapter 7] (see also Lemma 2.7.2.3), that Poisson and Gaussian integrals satisfy some combinatorial moment formulae. The following lemma explicit this behaviour.

**Lemma 3.5.2.1.** *For $f$ and $g \in \mathscr{L}^2(\nu)$. We let*

$$F = W_{re}(f) + W_{im}(g). \tag{3.5.2.4}$$

*and for all $\lambda > 0$,we let*

$$F_{\lambda} = \lambda^{-1/2} I_1^{\eta_{re}^{\lambda}}(f) + \lambda^{-1/2} I_1^{\eta_{im}^{\lambda}}(g). \tag{3.5.2.5}$$

*Then, for all $p \in \mathbb{N}$,*

$$\mathbb{E}F^p = (p-1)(p-3)\dots 1(\nu(f^2+g^2))^{p/2} 1_{p\in 2\mathbb{N}}. \tag{3.5.2.6}$$

*Moreover, if $f$ and $g \in \cap_{p'\leq p}\mathscr{L}^{p'}(\nu)$, for all $\lambda > 0$,*

$$\mathbb{E}F_{\lambda}^p = \begin{cases} \mathbb{E}F^p + O(\lambda^{-1}), & \text{if } p \in 2n; \\ O(\lambda^{-1/2}), & \text{otherwise.} \end{cases} \tag{3.5.2.7}$$

*Proof.* To prove the identity for $F$, we simply remark that $F \sim \mathbf{N}(0, \nu(f^2+g^2))$ and we use a well-known formula for Gaussian moments that we can find in any probability textbook. For the moments of $F_{\lambda}$, we use [82, Theorem 12.7]. From this formula, we deduce that

$$\mathbb{E}\left(\lambda^{-1/2} I_1^{\eta_{re}^{\lambda}}(f)\right)^p = \lambda^{-p/2} \sum_{\sigma\in\Pi_{\geq 2}} \int f_{\sigma}^{\otimes p}(\lambda\nu)^{|\sigma|}, \tag{3.5.2.8}$$

where $\Pi_{\geq 2}$ is the set of all partitions of $[p]$ whose blocks contain at least $2$ elements, the symbol $|\sigma|$ designates the number of blocks in a partition, and $f_\sigma^{\otimes p}$ is a function whose explicit general definition is irrelevant here (see Section 2.7.2.1). If $p$ is even, when $\sigma$ is a partition that contains exactly $p/2$ blocks of size $2$, we have that

$$f_\sigma^{\otimes p} = \left(f^2\right)^{\otimes p/2}. \tag{3.5.2.9}$$

Thus we find that

$$\mathbb{E}\left(\lambda^{-1/2} I_1^{\eta_{re}^\lambda}(f)\right)^p = \begin{cases} \mathbb{E}W_{re}(f)^p + O(\lambda^{-1}) & \text{if } p \in 2n; \\ O(\lambda^{-1/2}) & \text{otherwise.} \end{cases} \tag{3.5.2.10}$$

We infer the announced formula for $F_\lambda$ by using the binomial formula and the independence of $\eta_{re}$ and $\eta_{im}$. $\square$

**Theorem 3.5.2.2.** *Let $X$ be a Gaussian process satisfying the representation (3.5.2.1) with a $\sigma$-finite measure $\nu$ and $X^\lambda$ be its Poisson approximation defined in (3.5.2.3). Then, the finite-dimensional distributions of $X^\lambda$ converge to the ones of $X$, in the sense that, for all $t_1, \dots, t_l \in \mathbb{R}_+$,*

$$(X_{t_1}^\lambda, \dots, X_{t_l}^\lambda) \xrightarrow[\lambda\to\infty]{\text{stably}} (X_{t_1}, \dots, X_{t_l}). \tag{3.5.2.11}$$

*Proof.* In order to prove the convergence of the finite-dimensional law, we will invoke the multidimensional fourth moment theorem on the Poisson space of [41] (see also our Proposition 3.4.1.1). Indeed, let $t_1, \dots, t_l \in (0,1)$. Then, since $X^\lambda$ admits $R$ for covariance function, we have that, for $i$ and $j \in [l]$:

$$\mathbb{E}X_{t_i}^\lambda X_{t_j}^\lambda = R(t_i, t_j) = \mathbb{E}X_{t_i} X_{t_J}. \tag{3.5.2.12}$$

By Lemma 3.5.2.1, we have that

$$\mathbb{E}\left(X_t^\lambda\right)^4 = \mathbb{E}X_t^4 + O(\lambda^{-1}). \tag{3.5.2.13}$$

Since, by assumption, the spectral measure $\nu$ is $\sigma$-finite, we can readily apply Proposition 3.4.1.1 to obtain the convergence of the finite-dimensional laws, as $\lambda \to \infty$. $\square$

The previous theorem somehow justifies the name of Poisson approximation that we gave to $X^\lambda$. In view of this result, its is natural to ask if the conclusion of Theorem 3.5.1.2 still holds when we replace $X$ by $X^\lambda$, where $\lambda \to \infty$ with $n$. The following theorem provides an affirmative answer to this question. We need the two additional notations:

(3.5.2.14)

$$A_n = \int_0^1 \left(\prod_{i=1}^4 t_i\right)^{n-1} \nu\left(\prod_{i=1}^4 (\phi_1 - \phi_{t_i})\right) \prod_{i=1}^4 \mathrm{d}t_i = \nu\left(\left(\int_0^1 t^{n-1}(\phi_1 - \phi_t)\mathrm{d}t\right)^4\right);$$

(3.5.2.15)

$$B_n = \int_0^1 s^{n-1} t^{n-1} \nu\left(\phi_1^2(\phi_1 - \phi_t)(\phi_1 - \phi_s)\right) \mathrm{d}s\mathrm{d}t.$$

**Theorem 3.5.2.3.** *Let $X$ be a continuous Gaussian process representable by (3.5.2.1) and let $X^\lambda$ be its Poisson approximation. Let $(\lambda_n)_{n\in\mathbb{N}} \subset \mathbb{R}_+$. We define*

$$F_n = \int t^{n-1}\left(\left(X_1^{\lambda_n}\right)^2 - \left(X_t^{\lambda_n}\right)^2\right)\mathrm{d}t \tag{3.5.2.16}$$

*Let*

$$\epsilon_n = a_n^{-1}f_n(1) + a_n^{-1}b_n + \lambda_n^{-1}a_n^{-2}B_n + \lambda_n^{-1}a_n^{-4}A_n + \lambda_n^{-1}. \tag{3.5.2.17}$$

*We take $N \sim \mathbb{N}(0,1)$ independent of $X$, then, there exists a positive constant $c$, such that*

$$d_2(a_n^{-1}F_n, NX_1) \leq c\epsilon_n. \tag{3.5.2.18}$$

*Moreover, if $\epsilon_n \to 0$, as $n \to 0$, then,*

$$a_n^{-1}F_n \xrightarrow[n\to\infty]{stably} NX_1. \tag{3.5.2.19}$$

**Remark 24.** In the Gaussian case, we need to assume that $(a_n^{-1}f_n(s))$ converges to $0$ in order to ensure the asymptotic independence of $\int t^{n-1}(X_1 - X_t)\mathrm{d}t$ and $X$. In the Poisson case, as the Poisson process $\eta$ and the Gaussian white noise $W$ are chosen to be independent, this condition can be relaxed and thus only concerns $(a_n f_n(1))$.

*Proof.* Without loss of generality, we assume that $R(1,1) = 1$. As in the proof of Theorem 3.5.1.2, we write

$$\begin{aligned} F_n &= 2X_1^{\lambda_n}\int t^{n-1}(X_1^{\lambda_n} - X_t^{\lambda_n})\mathrm{d}t - \int t^{n-1}(X_1^{\lambda_n} - X_t^{\lambda_n})^2\mathrm{d}t \\ &= 2X_1^{\lambda_n}G_n - H_n, \end{aligned} \tag{3.5.2.20}$$

where

$$G_n = \int t^{n-1}(X_1^{\lambda_n} - X_t^{\lambda_n})\mathrm{d}t \tag{3.5.2.21}$$

$$H_n = \int t^{n-1}I_1(X_1^{\lambda_n} - X_t^{\lambda_n})^2\mathrm{d}t. \tag{3.5.2.22}$$

Since the Poisson approximation construction preserves the covariance structure, we still have $\mathbb{E}H_n = b_n$, for all $n \in \mathbb{N}$, and the quantity $a_n^{-1}H_n$ converges to $0$ in $\mathscr{L}^1(\mathbb{P})$. We consider the random vector

$$T_n = \begin{pmatrix} X_1^{\lambda_n} \\ a_n^{-1}G_n \end{pmatrix}, \quad n \in \mathbb{N}. \tag{3.5.2.23}$$

It is immediate to compute that

$$\Gamma(X_t^{\lambda_n}, X_s^{\lambda_n}) = \nu(\phi_t\phi_s) + \frac{1}{2\lambda_n}\hat{\eta}^{\lambda_n}(\phi_t\phi_s) = R(t,s) + \frac{1}{2\lambda_n}\hat{\eta}^{\lambda_n}(\phi_t\phi_s), \quad n \in \mathbb{N}. \tag{3.5.2.24}$$

Consequently, we obtain for the Stein kernel

$$S(T_n) = \begin{pmatrix} 1 & a_n^{-1}f_n(1) \\ a_n^{-1}f_n(1) & 1 \end{pmatrix} + \frac{\epsilon_n}{2\lambda_n}, \quad n \in \mathbb{N}, \tag{3.5.2.25}$$

where

(3.5.2.26)
$$\epsilon_n = \begin{pmatrix} \hat{\eta}^{\lambda_n}(\phi_1^2) & a_n^{-1}\int_0^1 t^{n-1}\hat{\eta}^{\lambda_n}(\phi_1(\phi_1-\phi_t))\mathrm{d}t \\ a_n^{-1}\int_0^1 t^{n-1}\hat{\eta}^{\lambda_n}(\phi_1(\phi_1-\phi_t))\mathrm{d}t & a_n^{-2}\int_0^1 t^{n-1}s^{n-1}\hat{\eta}^{\lambda_n}((\phi_1-\phi_t)(\phi_1-\phi_s))\mathrm{d}t\mathrm{d}s \end{pmatrix}.$$

In the following equation the square of the matrix has to be understood entry-wise. Then, by Itô's isometry, we have that

(3.5.2.27)
$$\mathbb{E}\left[\left(\frac{\epsilon_n}{\lambda_n}\right)^2\right] = \lambda_n^{-1}\begin{pmatrix} 1 & a_n^{-2}B_n \\ a_n^{-2}B_n & a_n^{-4}A_n \end{pmatrix}, \quad n \in \mathbb{N}.$$

This shows that

(3.5.2.28)
$$S(T_n) \xrightarrow[n\to\infty]{\mathscr{L}^1(\mathfrak{W})} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In order to apply Theorems 3.3.1.4 and 3.3.1.5, since the target variance is deterministic we are left to bound $r^4(T)$. Direct computations yield

(3.5.2.29)
$$r^4(T_n) = \lambda_n^{-1}\nu(\phi_t^4) + \lambda_n^{-1}a_n^{-4}A_n.$$

That vanishes by assumptions. □

# Transport inequalities on the configuration space

## 4.1. Introduction

Until now, we recovered for some point processes, and in particular for mixed binomial processes, quantitative limit theorems and a modified logarithmic Sobolev inequality. In a diffusive setting, as we mentioned in Section 2.1 it is also customary to recover an inequality by Talagrand that compares the quadratic Wasserstein distance and the entropy (see Section 4.2.3 for details). The groundbreaking paper of F. Otto & C. Villani (2000) [118] proved by means of optimal transport that the classical logarithmic Sobolev inequality on $\mathbb{R}^d$ implies the Talagrand inequality. Shortly after, the seminal paper of S. G. Bobkov, I. Gentil & M. Ledoux (2001) [20] shed a new light on this implication by giving a new proof of this result based on some properties of the Hamilton-Jacobi semi-group. This approach is very flexible and can be extended to arbitrary metric spaces following the work of N. Gozlan, C. Roberto & P.-M. Samson (2014) [57].

On the space of configurations, it is not clear what distance to use. Given a function $\phi$, Y. Ma, S. Shen, X. Wang & L. Wu (2011) [91] proposed to define the distance $d_\phi$ by saying that a random variable $F$ (measurable with respect to an underlying Poisson point process) is Lipschitz if and only if

$$|D_z^+ F| \le \phi(z), \text{ for all } z \in Z, \tag{4.1.0.1}$$

where $D^+$ has been defined in an abstract way in Section 2.4.2 and concretely for Poisson point processes in Section 2.7. Building on the logarithmic Sobolev inequality for functionals of a Poisson point process of L. Wu (2000) [152] (see Theorem 2.4.4.1) and an argument of S. G. Bobkov & F. Götze (1999) [17], Y. Ma, S. Shen, X. Wang & L. Wu (2011) [91] proved a transport-entropy inequality involving the linear Monge-Kantorovich-Rubinstein distance with respect to $d_\phi$ provided that the underlying Poisson point process has finite intensity measure. Their motivations for studying these inequalities is that they provide concentration of measure while being tensorizable. It is, indeed, a very classical fact that goes back (at least) to K. Marton (1996) [94] that transport-entropy inequalities have consequences in terms of concentration of measure.

For Poisson point processes, M. Reitzner (2013) [132] showed, still in the finite intensity case, that Poisson functionals enjoy another form of concentration of measure, namely a Gaussian concentration of measure with respect to a convex distance introduced by M. Talagrand (1995) [145]. Finding a proof of this concentration of measure phenomenon via transport-entropy inequalities is the one of the goals of an ongoing work in collaboration with N. Gozlan & G. Peccati. The forthcoming section gathers the computations sufficiently matured to be presented here. So far, we obtain two original results:

(i) Theorem 4.3.2.1 that is a Talagrand inequality for mixed binomial processes under a Talagrand inequality for the sampling probability;

(ii) Corollary 4.3.3.2 that is a transport-entropy inequality à la K. MARTON (1996) [95] for binomial processes of fixed size without assumption on the sampling probability.

From Corollary 4.3.3.2, we recover Corollary 4.3.3.3 that is the concentration of measure result of M. REITZNER (2013) [132]. Let us mention that the result for Poisson point processes is obtained by thinning of binomial processes as in the original paper [132] but we are confident that we will obtain a transport-entropy inequality for Poisson point processes that will imply directly the result of [132]. This not yet complete result is however presented in Theorem 4.3.4.1. Let us also mention that the idea of proof is very similar to [132]. Indeed our method of proof is to push a Talagrand or a Marton inequality for the sampling measure via contraction, while the idea of [132] is to push the concentration inequality of [145] for the convex distance.

## 4.2. PRELIMINARIES

4.2.1. **Reminders on point processes.** Recall that point processes have been introduced in Section 2.7 and that in Section 2.7.3, we studied two canonical examples of them: *Poisson point processes* and *binomial processes*. For the reader's convenience, we recall here some definitions and introduce some further notations. Let $(Z, \mathfrak{Z})$ be a measurable space. A *point process* or *random point measure* is a $\mathscr{M}_{\bar{\mathbb{N}}}(Z)$-valued random variable. We say that a point process $\eta$ is *finite* if $\mathbb{P}(\eta(Z) < \infty) = 1$ and we say that $\eta$ is *proper*, whenever there exists $Z$-valued random variables $X_1, X_2, \dots$ and a $\bar{\mathbb{N}}$-valued random variable $N$ such that

$$\eta = \sum_{i=1}^{N} \delta_{X_i}. \tag{4.2.1.1}$$

By extension, we say that a probability measure $\Pi \in \mathscr{P}(\mathscr{M}_{\bar{\mathbb{N}}}(Z))$ is *proper* if there exists a proper point process $\eta$ such that $\eta \sim \Pi$. We also write $\mathscr{M}_n(Z)$ for the space of point measures with total mass $n$.

Among the class of proper point processes, we shall be interested in the particular class of *binomial processes*, that is the class of proper point processes where, in (4.2.1.1), $N$ is taken independently of the $X_i$'s and the $X_i$'s are chosen independent and identically distributed according to a given probability measure $\nu \in \mathscr{P}(Z)$. In that case, the resulting point process is referred to as a *mixed binomial process with sampling distribution $\nu$ and size $N$* or, for short, simply a *mixed binomial process*, and its law is denoted by $B_{\nu,N}$. Note, that, of course, $B_{\nu,N}$ depends on $N$ only through its law. If $N \sim \kappa$, where $\kappa \in \mathcal{P}(\mathbb{N})$, we also write $B_{\nu,\kappa} = B_{\nu,N}$. The particular choice of $N = n$, for some $n \in \mathbb{N}$, yields the classical case of *binomial process of size $n$ and sampling distribution $\mu$* or simply *binomial process*. Provided that $\mathbb{E}N < \infty$, then $\eta \sim B_{\mu,N}$ has a finite *intensity measure* given by

$$\mathbb{E}\eta(A) = (\mathbb{E}N)\mu(A), \ A \in \mathscr{Z}. \tag{4.2.1.2}$$

Another important class of proper point processes, is the class of *Poisson point processes*. Given a measure $\nu$, we say that $\eta$ is a *Poisson point process with intensity measure* $\nu$ if: for all pairwise disjoint measurable sets $(A_1, \ldots, A_l)$ the random vector $(\eta(A_1), \ldots, \eta(A_l))$ is a vector of independent Poisson random variables with mean $(\nu(A_1), \ldots, \nu(A_l))$. If such a process exists is law will be denoted $\Pi_\nu$. Existence of Poisson point process with arbitrary reference measure is a non-trivial fact. However, one can easily check that if $\nu(Z) < \infty$ then,

$$\Pi_\nu = B_{\bar{\nu},N}, \tag{4.2.1.3}$$

where $\bar{\nu} = \frac{\nu}{\nu(Z)}$ and $N$ is a Poisson random variable with mean $\nu(Z)$. When $\nu$ is $\sigma$-finite (or more generally $s$-finite) a Poisson point process with reference measure $\nu$ can be obtained by a gluing procedure [82, Theorem 3.6] and its law is proper.

4.2.2. **Reminders on transportation distances.** We now recall some definitions and results about transport-entropy inequalities. Most of the content of this section is taken from the reference [58]. Let $E$ be a polish space. Recall that $E$ is endowed with its Borel $\sigma$-algebra $\mathfrak{E}$ and that $\mathscr{P}(E)$ is endowed with the $\sigma$-algebra generated by the evaluation maps

$$\mathscr{P}(E) \ni \mu \mapsto \mu(A), \quad A \in \mathfrak{E}. \tag{4.2.2.1}$$

Given a bi-measurable *cost function* $c\colon E \times \mathscr{P}(E) \to [0, \infty]$, the *(generalized) transportation cost* associated to $c$ from $\nu_1 \in \mathscr{P}(E)$ to $\nu_2 \in \mathscr{P}(E)$, noted $\mathcal{T}_c(\nu_1|\nu_2)$, is

$$\begin{aligned} \mathcal{T}_c(\nu_1|\nu_2) &= \inf_p \int c(x, p_x)\, \mathrm{d}\nu_2(x) \\ &= \inf_{X_1, X_2} \mathbb{E} c(X_2, \mathrm{Law}(X_1|X_2)), \end{aligned} \tag{4.2.2.2}$$

where the first infimum runs over the set of kernels $p\colon E \ni x \mapsto p_x \in \mathscr{P}(E)$ such that $\int p_x(A)\nu_2(\mathrm{d}x) = \nu_1(A)$, for all Borel set $A$, and the second infimum runs over all random variables $X_1 \sim \nu_1,\ X_2 \sim \nu_2$. We will implicitly assume that the bi-measurability is satisfies in the rest of the document. Note that, $\mathcal{T}_c$ is, in general, not symmetric with respect to $\nu_1$ and $\nu_2$. Let us give three canonical examples of transportation costs that we use pervasively throughout the paper. Consider a pseudo-distance $\rho\colon E \times E \to [0, \infty]$. The quadratic *Wasserstein cost*, $\mathcal{W}_{2,\rho}^2$, associated to $\rho$ is the transportation cost associated to the cost

$$c(x, p) = \int \rho(x, y)^2 p(\mathrm{d}y). \tag{4.2.2.3}$$

Namely,

$$\mathcal{W}_{2,\rho}^2(\nu_1, \nu_2) = \inf\{\mathbb{E}\rho(X_1, X_2)^2 | X_1 \sim \nu_1, X_2 \sim \nu_2\}. \tag{4.2.2.4}$$

The *Marton cost*, $\mathcal{M}_\rho^2$, associated to $\rho$ is the transportation cost associated to the cost

$$c(x, p) = \left( \int \rho(x, y) p(\mathrm{d}y) \right)^2. \tag{4.2.2.5}$$

Namely,

$$\mathcal{M}_\rho^2(\nu_1,\nu_2) = \inf\left\{\mathbb{E}\left((\mathbb{E}\left[\rho(X_1,X_2)|X_2\right])^2\right)|X_1\sim\nu_1, X_2\sim\nu_2\right\}. \tag{4.2.2.6}$$

Finally the *Monge-Kantorovich-Rubinstein cost*, $\mathcal{W}_{1,\rho}$, associated $\rho$ is the transportation cost associated to the cost

$$c(x,p) = \int \rho(x,y)p(\mathrm{d}y). \tag{4.2.2.7}$$

Namely,

$$\mathcal{W}_{1,\rho}(\nu_1,\nu_2) = \inf\{\mathbb{E}\rho(X_1,X_2)|X_1\sim\nu_1, X_2\sim\nu_2\}. \tag{4.2.2.8}$$

Note that $\mathcal{W}_{1,\rho}$ and $\mathcal{W}_{2,\rho}$ are both symmetric since $\rho$ is symmetric, while $\mathcal{M}_\rho$ is, in general, not symmetric. By Jensen's inequality, we easily check that, for all $\nu_1$ and $\nu_2\in\mathscr{P}(E)$,

$$\mathcal{W}_{1,\rho}^2(\nu_1,\nu_2) \le \mathcal{M}_\rho^2(\nu_1,\nu_2) \le \mathcal{W}_{2,\rho}^2(\nu_1,\nu_2). \tag{4.2.2.9}$$

Observe that when $\rho(x,y) = 1_{x\neq y}$, then
(4.2.2.10)

$$\mathcal{W}_{1,\rho}(\nu_1,\nu_2) = \mathcal{W}_{2,\rho}^2(\nu_1,\nu_2) = \inf\mathbb{P}(X_1\neq X_2) = \int\left|1-\frac{\mathrm{d}\nu_2}{\mathrm{d}\nu_1}(x)\right|_+ \nu_1(\mathrm{d}x) = TV(\nu_1,\nu_2),$$

where the infimum is running over all $X_1\sim\nu_1$ and $X_2\sim\nu_2$ and (see [95])

$$\mathcal{M}_\rho^2(\nu_1,\nu_2) = \int\left|1-\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_2}(x)\right|_+^2 \nu_2(\mathrm{d}x). \tag{4.2.2.11}$$

4.2.2.1. *Partial transport distances.* Observe that the definition in (4.2.2.2) of the transportation cost associated to $c$ can be extended to any two measures with same total mass (possibly infinite) and for two such measures $\nu_1$ and $\nu_2$, we write $\mathcal{T}_c(\nu_1|\nu_2)$ for this transport cost. We now extend this transportation cost to point measures with different total masses. Given a cost $c\colon E\times\mathscr{P}(E)$ and a function $\phi\colon E\to[0,\infty]$, the *partial transportation cost* between them $\nu_1$ and $\nu_2\in\mathscr{M}(E)$ associated to $c$ and $\phi$, denoted by $\mathcal{T}_{c,\phi}(\nu_1|\nu_2)$, is defined as follows:

$$\mathcal{T}_{c,\phi}(\nu_1|\nu_2) = \begin{cases}\inf\{\mathcal{T}_c(\nu_1|\bar{\nu}_2)+\tilde{\nu}_2(\phi)\}, & \text{if } \nu_1(E)\le\nu_2(E);\\ \inf\{\mathcal{T}_c(\bar{\nu}_1|\nu_2)+\tilde{\nu}_1(\phi)\}, & \text{if } \nu_2(E)\le\nu_1(E),\end{cases} \tag{4.2.2.12}$$

where the first infimum is running over all $\bar{\nu}_2$ and $\tilde{\nu}_2\in\mathcal{M}(E)$ such that $\nu_2=\bar{\nu}_2+\tilde{\nu}_2$ and $\bar{\nu}_2(E)=\nu_1(E)$ and a similar constraint for the second infimum. This partial transport cost consists in transporting the same amount of mass from one measure to the other while penalizing the remaining mass with $\phi$ and doing so in the most efficient way. When $c$ is given by one of the previously introduced cost (4.2.2.3), (4.2.2.5) and (4.2.2.7) the corresponding partial transport costs are denoted (respectively) $\mathcal{W}_{2,\rho,\phi}$, $\mathcal{M}_{\rho,\phi}$ and $\mathcal{W}_{1,\rho,\phi}$. There are two extremal choices for $\phi$:

(i) When $\phi=0$, the points of the measure $\nu_1$ with the smallest support are optimally paired with points in sub-configuration of $\nu_2$ of same size of $\nu_1$; the remaining points in $\nu_2$ play no role.

(ii) When $\phi = \infty$, if $\nu_1$ and $\nu_2$ are of same total mass, their points are paired optimally; otherwise the distance between is zero. In this case, the space $\mathscr{M}_{\mathbb{N}}(Z)$ has infinitely many connected components (with respect to to the topology induced by this partial transportation distance). These components are given by the family $\{\mathscr{M}_n(Z);\ n \in \mathbb{N} \cup \{\infty\}\}$.

Let us now give another representation of this partial transport cost when $\nu_1$ and $\nu_2 \in \mathcal{M}_{\mathbb{N}}(E)$. In the sequel, $\Sigma_n$ will denote the set of all permutations of $\{1, \ldots, n\}$. For $\xi = \sum_{i=1}^m \delta_{a_i} \in \mathscr{M}_{\mathbb{N}}(E)$, we define, for $n \le m$ and $\sigma \in \Sigma_m$, the two point measures:

$$\sigma^n \xi = \sum_{i=1}^n \delta_{a_{\sigma(i)}} \quad \text{and} \quad \sigma_n \xi = \sum_{i=n+1}^m \delta_{a_{\sigma(i)}}$$

(an empty sum is treated as 0). Then, given a cost $c$ and a function $\phi\colon E \to [0, \infty]$, the *partial transportation cost* associated to $c$ and $\phi$, denoted by $\mathcal{T}_{c,\phi}$, is defined as follows:

$$\mathcal{T}_{c,\phi}(\xi_1|\xi_2) = \begin{cases} \inf_{\sigma \in \Sigma_m} \mathcal{T}_c(\xi_1|\sigma^n\xi_2) + (\sigma_n\xi_2)(\phi), & \text{if } n = \xi_1(E) \le \xi_2(E) = m, \\ \inf_{\sigma \in \Sigma_m} \mathcal{T}_c(\sigma^n\xi_1|\xi_2) + (\sigma_n\xi_1)(\phi), & \text{if } n = \xi_2(E) \le \xi_1(E) = m. \end{cases} \tag{4.2.2.13}$$

where $\xi_1$ and $\xi_2 \in \mathscr{M}_{\mathbb{N}}(E)$. In (4.2.2.12), we use the convention that $0 \cdot \infty = 0$.

4.2.3. **Reminders on transport-entropy inequalities.** Recall that, given $\gamma \in \mathscr{P}(E)$ and $\nu \in \mathscr{P}(E)$ absolutely continuous with respect to $\gamma$ and with Radon-Nikodym density $f$, the *relative entropy* of $\nu$ with respect to $\gamma$ is defined by

$$\mathcal{H}(\nu|\gamma) = \gamma(f \log f). \tag{4.2.3.1}$$

If $\nu$ does not have a density with respect to $\gamma$, we set $\mathcal{H}(\nu|\gamma) = \infty$. The measure $\gamma \in \mathscr{P}(E)$ satisfies the *transport-entropy inequality* with the cost function $c$ if for all $\nu_1$ and $\nu_2 \in \mathscr{P}(E)$, it holds

$$\mathcal{T}_c(\nu_1|\nu_2) \le \mathcal{H}(\nu_1|\gamma) + \mathcal{H}(\nu_2|\gamma). \tag{4.2.3.2}$$

Transport-entropy inequalities enjoy the following tensorization property [58, Theorem 4.11].

**Proposition 4.2.3.1.** *If $\gamma$ satisfies the transport-entropy inequality (4.2.3.2) on $E$ with cost $c$ such that for all $x \in E$, $c(x, \cdot)$ is convex. Then for all $n \ge 1$, $\gamma^n$ satisfies the transport-entropy inequality (4.2.3.2) with cost $c^n$ on $E^n$, where*

$$c^n(x, p) = \sum_{i=1}^n c(x_i, p_i), \tag{4.2.3.3}$$

*with $x = (x_1, \ldots, x_n) \in E^n$ and $p_i$ is the $i$-th marginal of $p \in \mathscr{P}(E^n)$.*

We now introduce two particular classes of transport-entropy inequalities that are associated with the two costs of (4.2.2.3) and (4.2.2.5). These two inequalities were introduced in the seminal papers by M. TALAGRAND (1996) [144] and K. MARTON (1996) [95], whence they get their names. Given a distance $\rho\colon E \times E \to \mathbb{R}_+$ we say that $\gamma$ satisfies a *Talagrand inequality* with respect to $\rho$ and with constant $C > 0$ if

$$\mathcal{W}_\rho^2(\nu_1, \nu_2) \le C\mathcal{H}(\nu_1|\gamma) + C\mathcal{H}(\nu_2|\gamma), \qquad \forall \nu_1, \nu_2 \in \mathcal{P}(E). \tag{4.2.3.4}$$

We say that $\gamma$ satisfies a *Marton inequality* with respect to $\rho$ and with constant $C > 0$ if

$$\mathcal{M}_\rho^2(\nu_1,\nu_2) \le C\mathcal{H}(\nu_1|\gamma) + C\mathcal{H}(\nu_2|\gamma), \qquad \forall \nu_1,\nu_2 \in \mathcal{P}(E). \tag{4.2.3.5}$$

We also say that $\gamma$ satisfies an *infimum convolution inequality* (introduced by B. MAUREY (1991) [96]) with respect to $c$ when

$$\mu(\exp(Q_c\phi))\mu(\exp(-\phi)) \le 1, \tag{4.2.3.6}$$

where

$$Q_c\phi(x) = \inf_{p\in\mathcal{P}(E)} \{p(\phi) + c(x,p)\}. \tag{4.2.3.7}$$

Morally, (4.2.3.2) and (4.2.3.6) are equivalent. The equivalence between (4.2.3.4) and their infimum convolution form appeared in the work of S. G. BOBKOV & F. GÖTZE (1999) [17]; N. GOZLAN (2007) [54] showed the complete equivalence. The equivalence of (4.2.3.5) and its infimum convolution form was derived by N. GOZLAN, C. ROBERTO, P.-M. SAMSON & P. TETALI (2017) [58], where they also gave sufficient conditions on the cost $c$ for the equivalence of (4.2.3.2) and (4.2.3.6).

4.2.4. **Transport, entropy and concentration of measure.** Transport-entropy inequalities have consequences in terms of concentration of measure (and were initially introduced in order to study the concentration of measure phenomenon). To state them, let us recall some definitions from [58]. Given a cost $c$ and a Borel set $A$ we write

$$c_A(x) = \inf_{p(A)=1} c(x,p), \tag{4.2.4.1}$$

for the *Talagrand convex distance* associated to $A$ and $c$, and

$$A_t = \{x \in E, c_A(x) \le t\}, \tag{4.2.4.2}$$

for the *enlargement* of $A$ with respect to this convex distance. Note that, despite the name, it is not always a distance. Observe, however, that when $c$ is the Monge-Kantorovich-Rubinstein cost (4.2.2.7) associated to a distance $\rho$, then $c_A(x)$ is the distance from $x$ to $A$ and $A_t$ is the $t$-enlargement of $A$ with respect to $\rho$. With these notations, we state the following concentration result.

**Theorem 4.2.4.1** ([58, Theorem 5.1]). *Let $E$ be a Polish space and let $c$ be a cost such that $\mathscr{P}(E) \ni p \mapsto c(x,p)$ is convex. Assume $\gamma \in \mathscr{P}(E)$ satisfies the transport-entropy inequality* (4.2.3.2) *with cost $c$. Then,*

$$(1-\gamma(A_t))\gamma(A) \le e^{-t}, \qquad \forall t \ge 0. \tag{4.2.4.3}$$

**Remark 25.** The theorem in the original paper has the additional condition on $c$ that $c(x,\delta_x) = 0$ for all $x \in E$. However, one of the author of [58] pointed out to us that this condition is in fact irrelevant.

4.3.1. **Two ancillary results.** We will need the two following results.

**Proposition 4.3.1.1.** *Let $T : \mathcal{X} \to \mathcal{Y}$ be a measurable map between two polish spaces $\mathcal{X}$ and $\mathcal{Y}$. Suppose that $\gamma \in \mathscr{P}(\mathcal{X})$ satisfies the transport-entropy inequality (4.2.3.2) with some cost function $c : \mathcal{X} \times \mathscr{P}(\mathcal{X}) \to \mathbb{R}_+$. If the cost function $c$ is convex with respect to its second variable, then the probability measure $\bar{\gamma} = T_\#\gamma$ satisfies the transport-entropy inequality (4.2.3.2) with the cost function $\bar{c} : \mathcal{Y} \times \mathscr{P}(\mathcal{Y}) \to [0, \infty]$ defined for all $y \in \mathcal{Y}$ and $q \in \mathscr{P}(\mathcal{Y})$ by*

$$\bar{c}(y, q) = \inf\{c(x, p); T(x) = y,\ T_\# p = q\} \tag{4.3.1.1}$$

*(with the convention $\inf \emptyset = +\infty$).*

*Proof.* Let $\bar{\nu}_1, \bar{\nu}_2 \in \mathscr{P}(\mathcal{Y})$ be such that $H(\bar{\nu}_1|\bar{\gamma}) < \infty$ and $H(\bar{\nu}_2|\bar{\gamma}) < \infty$. Let $\bar{h}_1$ be the density of $\bar{\nu}_1$ with respect to $\bar{\gamma}$; then for all bounded continuous function $f$ on $\mathcal{Y}$

$$\int f\, d\bar{\nu}_1 = \int f \bar{h}_1\, d\bar{\gamma} = \int f(T)\bar{h}_1(T)\, d\gamma = \int f(T)\, d\nu_1, \tag{4.3.1.2}$$

denoting $d\nu_1 = \bar{h}_1(T)\, d\gamma$. Therefore, there exists at least one probability measure $\nu_1$ on $\mathcal{X}$ such that $\bar{\nu}_1 = T_\#\nu_1$. On the other hand,

$$H(\bar{\nu}_1|\bar{\gamma}) = \int \bar{h}_1 \log \bar{h}_1\, d\bar{\gamma} = \int \bar{h}_1(T) \log \bar{h}_1(T)\, d\gamma = H(\nu_1|\gamma). \tag{4.3.1.3}$$

Let us consider the function $\overline{\mathcal{T}}_c(\,\cdot\,|\,\cdot\,)$ defined on $\mathscr{P}(\mathcal{X})^2$ by

$$\overline{\mathcal{T}}_c(\bar{\nu}_1|\bar{\nu}_2) = \inf\left\{\mathcal{T}_c(\nu_1|\nu_2); \bar{\nu}_1 = T_\#\nu_1 \text{and} \bar{\nu}_2 = T_\#\nu_2\right\}. \tag{4.3.1.4}$$

According to what precedes, for all $\bar{\nu}_i$, $i = 1, 2$, such that $H(\bar{\nu}_i|\bar{\gamma}) < \infty$, there exist $\nu_i$, $i = 1, 2$, on $\mathscr{P}(\mathcal{X})$ such that $\bar{\nu}_i = T_\#\nu_i$ and so

$$\overline{\mathcal{T}}_c(\bar{\nu}_1|\bar{\nu}_2) \le \mathcal{T}_c(\nu_1|\nu_2) \le H(\nu_1|\gamma) + H(\nu_2|\gamma) = H(\bar{\nu}_1|\bar{\gamma}) + H(\bar{\nu}_2|\bar{\gamma}) \tag{4.3.1.5}$$

Now let us prove that

$$\overline{\mathcal{T}}_c(\bar{\nu}_1|\bar{\nu}_2) \ge \mathcal{T}_{\bar{c}}(\bar{\nu}_1|\bar{\nu}_2). \tag{4.3.1.6}$$

Let $\bar{\nu}_1, \bar{\nu}_2$ such that $H(\bar{\nu}_i|\bar{\gamma}) < +\infty$, $i = 1, 2$; there exist $\nu_1, \nu_2$ such that $\bar{\nu}_i = T_\#\nu_i$. Let $p$ be a kernel such that $\nu_1 = \nu_2 p$. Equivalently, there exists a pair of random variables $(X_1, X_2)$ with $\mathrm{Law}(X_2) = \nu_2$ and $\mathrm{Law}(X_1|X_2 = x) = p_x$. Consider $Y_1 = T(X_1)$ and $Y_2 = T(X_2)$; for all bounded continuous functions $f_1, f_2$ on $\mathcal{Y}$ it holds

$$\begin{aligned}
\mathbb{E}[f_1(Y_1)f_2(Y_2)] = \mathbb{E}[f_1(T(X_1))f_2(T(X_2))] &= \mathbb{E}\left[\int f_1(T(x_1))\, dp_{X_2}(x_1) f_2(T(X_2))\right] \\
&= \mathbb{E}\left[\int f_1(y_1)\, d(T_\# p_{X_2})(y_1) f_2(T(X_2))\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\int f_1(y_1)\, d(T_\# p_{X_2})(y_1) | T(X_2)\right] f_2(T(X_2))\right]
\end{aligned}$$

Consider a regular conditional probability $k$ for $\mathrm{Law}(X_2|Y_2)$; then

$$\begin{aligned}\mathbb{E}[f_1(Y_1)f_2(Y_2)] &= \mathbb{E}\left[\int\left(\int f_1(y_1)\,d(T_\# p_{x_2})(y_1)\right)k_{Y_2}(dx_2)f_2(Y_2)\right]\\ &= \iint f_1(y_1)f_2(y_2)\bar{p}_{y_2}(dy_1)\,\bar{\nu}_2(dy_2),\end{aligned}\tag{4.3.1.7}$$

with

$$\bar{p}_{y_2} = \int (T_\# p_{x_2})\,k_{y_2}(dx_2).\tag{4.3.1.8}$$

This proves that $\bar{\nu}_2\bar{p} = \nu_1$.

$$\begin{aligned}\int c(x_2, p_{x_2})\,\nu_2(dx_2) &\geq \int \bar{c}(T(x_2), T_\# p_{x_2})\,\nu_2(dx_2)\\ &= \iint \bar{c}(T(x_2), T_\# p_{x_2})\,k_{y_2}(dx_2)\bar{\nu}_2(dy_2)\\ &= \iint \bar{c}(y_2, T_\# p_{x_2})\,k_{y_2}(dx_2)\bar{\nu}_2(dy_2)\\ &\geq \int \bar{c}\left(y_2, \int T_\# p_{x_2}\,k_{y_2}\right)\bar{\nu}_2(dy_2)\\ &\geq \mathcal{T}_{\bar{c}}(\bar{\nu}_1|\bar{\nu}_2),\end{aligned}$$

where the first inequality comes from the definition of $\bar{c}$, the third is a consequence of the fact that $T(x_2) = y_2$ for $k_{y_2}$ almost all $x_2$, and the fourth follows from the convexity of $\bar{c}$ (which is itself a simple consequence of the convexity of $c$). Therefore, taking the infimum over $p$ yields to $\mathcal{T}_c(\nu_1|\nu_2) \geq \mathcal{T}_{\bar{c}}(\bar{\nu}_1|\bar{\nu}_2)$. Taking the infimum over all $\nu_1, \nu_2$ such that $T_\#\nu_i = \bar{\nu}_i$ finally gives (4.3.1.6) and completes the proof. □

The second result is a well-known result by K. Marton (1996) [95] about transport inequalities for product probability measures.

**Theorem 4.3.1.2.** *Let $\nu$ be a probability measure on $Z$; then for all positive integer $1 \leq n \leq \infty$ the probability measure $\nu^n$ on $Z^n$ satisfies the transport entropy inequality (4.2.3.2) with a cost $c$ given by*

$$c(x,p) = \frac{1}{4}\sum_{i=1}^{n} p(\{y \in Z^n,\ \textit{such that}\ \ y_i \neq x_i\}),\ x \in Z^n,\ p \in \mathscr{P}(Z^n).\tag{4.3.1.9}$$

**Remark 26.** In probabilistic notations, the content of Theorem 4.3.1.2 can be rewritten

$$\mathcal{M}_H^{(n)}(\nu_1|\nu_2) \leq 4H(\nu_1|\nu^n) + 4H(\nu_2|\nu^n), \qquad \forall \nu_1, \nu_2 \in \mathscr{P}(Z^n),\tag{4.3.1.10}$$

where $\mathcal{M}_H^{(n)}$ stands for the Marton distance with respect to the Hamming distance $\rho(x,y) = 1_{x\neq y}$ in dimension $n$, namely:

$$\mathcal{M}_H^{(n)}(\nu_1|\nu_2) = \inf \mathbb{E}\left[\sum_{i=1}^{n}\mathbb{P}(Y_i \neq X_i|X_i)^2\right],\tag{4.3.1.11}$$

where the infimum runs over the set of couples of random vectors $(X,Y)$ with $X \sim \nu_2$ and $Y \sim \nu_1$.

**Remark 27.** In view of the tensorization property Proposition 4.2.3.1, the content of Theorem 4.3.1.2 is just that when $\rho(x,y) = 1_{x\neq y}$ then every probability measure $\nu$ satisfies the Marton inequality $\mathcal{M}_\rho$.

4.3.2. **Talagrand inequality for mixed binomial processes.** The following main result states that Talagrand inequalities can be transferred from the sampling distribution $\nu$ to the law of the binomial process $B_{\nu,\kappa}$ for every distribution size $\kappa$.

**Theorem 4.3.2.1.** *Let $\kappa \in \mathscr{P}(\mathbb{N})$ and $\nu \in \mathscr{P}(Z)$. Assume $\nu$ satisfies Talagrand's inequality (4.2.3.4) on $Z$ for some pseudo-distance $\rho\colon Z \times Z \to [0,\infty]$ with constant $C > 0$. Then, the law of the mixed binomial process $B_{\nu,\kappa}$ satisfies Talagrand's inequality (4.2.3.4) on $E = \mathscr{M}_{\mathbb{N}}(Z)$ associated with the pseudo-distance $\mathcal{W}_{2,\rho,\infty}$, where this quantity is the partial transport cost defined in (4.2.2.12) from the Wasserstein transportation distance (4.2.2.4). Namely, for all $\Pi_1$ and $\Pi_2 \in \mathscr{P}(\mathscr{M}_{\mathbb{N}}(Z))$ such that $\mathcal{W}_{2,\rho,\infty}(\Pi_1,\Pi_2) < \infty$, we have that*

$$\mathcal{T}_{\mathcal{W}^2_{2,\rho,\infty}}(\Pi_1,\Pi_2) \leq C\mathcal{H}(\Pi_1|B_{\nu,\kappa}) + C\mathcal{H}(\Pi_2|B_{\nu,\kappa}). \tag{4.3.2.1}$$

**Remark 28.** When the cost $c$ can assume the value $\infty$, the left-hand side of (4.2.3.2) can assume the value $\infty$ for measures $\nu_1$ and $\nu_2$ that are both absolutely continuous with respect to reference $\gamma$ so that the right-hand side of (4.2.3.2) is finite. For instance, on $\mathscr{M}_{\mathbb{N}}(Z)$, if $n_1 \neq n_2$ we have that $\mathcal{T}_{\mathcal{W}^2_{2,\rho,\infty}}(B_{\nu,n_1}, B_{\nu,n_2}) = \infty$, while the two binomial laws are absolutely continuous with respect to the $\Pi_\nu$ with Radon-Nikodym given by $\mathrm{e}^{-1}\, n_i! 1_{\mathscr{M}_{n_i}(Z)}$ (this comes from the fact that a binomial process can be obtained by conditioning a Poisson point process to have $n$ points), and, thus, $\mathcal{H}(B_{\nu,n_1}|\Pi_\nu) < \infty$. In this case, it is reasonable to ask if a restricted transport-entropy inequality holds in the form of

$$\mathcal{T}_c(\nu_1,\nu_2) \leq c\mathcal{H}(\nu_1|\gamma) + c\mathcal{H}(\nu_2|\gamma), \quad \forall \nu_1,\nu_2,\ \mathcal{T}_c(\nu_1,\nu_2) < \infty. \tag{4.3.2.2}$$

We stress that we do not know if these inequalities tensorize or provide concentration of measure.

*Proof.* For short, we will write in this proof $\mathcal{T}_{\mathcal{W}^2} = \mathcal{T}_{\mathcal{W}^2_{2,\rho,\infty}}$. Let $\Pi_1$ and $\Pi_2 \in \mathscr{P}(\mathscr{M}_{\mathbb{N}}(Z))$ such that $\mathcal{T}_{\mathcal{W}^2}(\Pi_1,\Pi_2) < \infty$. Recall that $\mathscr{M}_b(Z)$ is endowed with the narrow topology and that it is a Polish for this topology. In view of [146, Cor 6.11, Rmk 6.12], the cost $\mathcal{W}^2_{2,\rho,\infty}$ is lower semi-continuous and by [146, Theorem 4.1] there exists be an optimal coupling in $\mathcal{T}_{\mathcal{W}^2}(\Pi_1,\Pi_2)$. We write $(\eta_1,\eta_2)$ for such an optimal coupling. From the finiteness assumption of the transport distance, we have that $\eta_1(Z) = \eta_2(Z) = N$ almost surely. We denote by $p$ their common law. For $k \in \mathbb{N}$, we write

$$\Pi^k = law((\eta_1,\eta_2)|N = k), \tag{4.3.2.3}$$

we consider $(\eta_1^k,\eta_2^k) \sim \Pi^k$, and we write $\Pi_1^k$ (resp. $\Pi_2^k$) for the law of $\eta_1^k$ (resp. $\eta_2^k$), that is $\Pi_1^k$ and $\Pi_2^k$ are the marginals of $\Pi^k$. Recall the following result taken from [146, Theorem 4.6]:

**Lemma 4.3.2.2.** *Let $X$ and $Y$ be two Polish spaces, let $c\colon X \times Y \to \mathbb{R}_+ \cup \{\infty\}$. Let $\mu \in \mathscr{P}(X)$ and $\nu \in \mathscr{P}(Y)$. Assume that $\mathcal{T}_c(\mu,\nu) < \infty$ and let $\pi$ be an optimal transport plan. Let $\tilde{\pi} \in \mathscr{M}_+(X \times Y)$ such that $\tilde{\pi} \leq \pi$ and $\tilde{\pi} \neq 0$, then the probability measure*

$$\pi' = \frac{\tilde{\pi}}{\tilde{\pi}(X \times Y)}, \tag{4.3.2.4}$$

*is an optimal transport plan between its marginals.*

In view of this lemma, we have that $(\eta_1^k, \eta_2^k)$ is the optimal coupling in $\mathcal{T}_{\mathcal{W}^2}(\Pi_1^k, \Pi_2^k)$. By construction, we have that

$$\begin{aligned}
\mathcal{T}_{\mathcal{W}^2}(\Pi_1, \Pi_2) &= \mathbb{E}\mathcal{W}^2_{\rho,2,\infty}(\eta_1, \eta_2) \\
&= \sum_{k\in\mathbb{N}} \mathbb{E}(\mathcal{W}^2_{\rho,2}(\eta_1, \eta_2)|N = k)\mathbb{P}(N = k) \\
&= \sum_{k\in\mathbb{N}} \mathbb{E}\mathcal{W}^2_{\rho,2}(\eta_1^k, \eta_2^k)\mathbb{P}(N = k) \\
&= \sum_{k\in\mathbb{N}} \mathcal{T}_{\mathcal{W}^2}(\Pi_1^k, \Pi_2^k)\mathbb{P}(N = k).
\end{aligned} \tag{4.3.2.5}$$

We consider the map $T\colon Z^k \to \mathscr{M}_k(Z)$, $T(z_1, \dots, z_k) = \sum_{i=1}^k \delta_{z_i}$. Observe that $T\sharp\nu^k = B_{\nu,k}$ and that

$$\mathcal{W}^2_{2,\rho}(\xi, \chi) = \inf\left\{\sum_{i=1}^k \rho(z_i, \tilde{z}_j)^2, \text{ such that } T(z_1, \dots, z_k) = \xi,\ T(\tilde{z}_1, \dots, \tilde{z}_k) = \chi\right\}. \tag{4.3.2.6}$$

Since, by tensorization, $\nu^k$ satisfies a Talagrand inequality (4.2.3.4) on $Z^k$ with respect to the distance

$$\rho_k(z, \tilde{z}) = \sqrt{\sum_{i=1}^k \rho(z_i, \tilde{z}_i)^2}, \quad z, \tilde{z} \in Z^k, \tag{4.3.2.7}$$

we obtain from Proposition 4.3.1.1 that $B_{\nu,k}$ satisfies a Talagrand inequality on $\mathscr{M}_k(Z)$ with respect to $\mathcal{W}_{2,\rho}$. Hence, we proved that

$$\mathcal{T}_{\mathcal{W}^2}(\Pi_1, \Pi_2) \le \sum_{k\in\mathbb{N}} C\left(\mathcal{H}(\Pi_1^k|B_{\nu,k}) + \mathcal{H}(\Pi_2^k|B_{\nu,k})\right)\mathbb{P}(N = k). \tag{4.3.2.8}$$

Now we claim that.

$$\mathcal{H}(\Pi_1|B_{\nu,\kappa}) = \mathcal{H}(p|\kappa) + \sum_{k\in\mathbb{N}} \mathcal{H}(\Pi_1^k|B_{\nu,k})\mathbb{P}(N = k). \tag{4.3.2.9}$$

Indeed, with $K \sim \kappa$,

$$\frac{\mathrm{d}\Pi_1^k}{\mathrm{d}B_{\nu,k}} = \frac{\mathrm{d}\Pi_1}{\mathrm{d}B_{\nu,\kappa}}\frac{\mathbb{P}(N = k)}{\mathbb{P}(K = k)}. \tag{4.3.2.10}$$

Hence,

$$\mathcal{H}(\Pi_1^k|B_{\nu,k})\mathbb{P}(N = k) = \int \frac{\mathrm{d}\Pi_1}{\mathrm{d}B_{\nu,\kappa}}(\xi)\log\frac{\mathrm{d}\Pi_1}{\mathrm{d}B_{\nu,\kappa}}(\xi)B_{\nu,k}(\mathrm{d}\xi)\mathbb{P}(K = k) + \mathbb{P}(K = k)\log\frac{\mathbb{P}(K = k)}{\mathbb{P}(N = k)}. \tag{4.3.2.11}$$

Summing the previous for $k \in \mathbb{N}$ yields (4.3.2.9). Using a similar equation for $\Pi_2$ and using the fact that the relative entropy is non-negative, we obtain that

$$\mathcal{T}_{\mathcal{W}^2}(\Pi_1, \Pi_2) \le C\mathcal{H}(\Pi_1|B_{\nu,\kappa}) + C\mathcal{H}(\Pi_2|B_{\nu,\kappa}). \tag{4.3.2.12}$$

This concludes the proof.

□

4.3.2.1. *Comparison with a results of Erbar & Huesmann.* On a Riemannian manifold $M$ with Ricci curvature bounded by $K \in \mathbb{R}$ with volume measure $\mathrm{vol}$ and Riemannian distance $d$, M. Erbar & M. Huesmann (2015) [44] have shown that

$$\mathcal{H}(\pi_t|\Pi_{\mathrm{vol}}) \le (1-t)\mathcal{H}(\pi_0|\Pi_{\mathrm{vol}}) + t\mathcal{H}(\pi_1|\Pi_{\mathrm{vol}}) - \frac{K}{2}t(1-t)\mathcal{W}^2_{2,d,\infty}(\pi_0,\pi_1), \tag{4.3.2.13}$$

where $\pi_t$ is any $\mathcal{W}_{2,d,\infty}$-geodesic from $\pi_0$ to $\pi_1$, where $\pi_0$ and $\pi_1 \in \mathscr{P}(\mathscr{M}_{\bar{\mathbb{N}}}(M))$. This means that the curvature properties of the metric measured space $(M, d, \mathrm{vol})$ are transferred to the metric measured space $(\mathscr{M}_{\bar{\mathbb{N}}}(M), \mathcal{W}_{2,d,\infty}, \Pi_{\mathrm{vol}})$. This result is very natural, if we think of $\Pi_{\mathrm{vol}}$ as the invariant measure of a system of non-interacting $N$ Brownian motions on $M$, where $N$ has a Poisson law with mean $\mathrm{vol}(M)$ (possibly infinite). Note that in the case where $K > 0$, the manifold is compact so that $\Pi_{\mathrm{vol}}$ has a representation as a mixed binomial process and taking $t = 1/2$ and using that the entropy is non-negative in (4.3.2.13) immediately yields our (4.3.2.1). Let us stress that our argument works under the sole assumption that the space $Z$ supports a Talagrand inequality and is quite straightforward. However, it is not clear how to adapt our proof for a Poisson point process with infinite intensity measure.

4.3.3. **Marton inequality for binomial processes.** Our Theorem 4.3.2.1 states that under Talagrand's inequality for $\nu$, the probability measure $B_{\nu,\kappa}$ satisfies a transport-entropy inequality for the cost function

$$c(\xi,\Pi) = \int \mathcal{W}^2_{2,\rho,\infty}(\xi,\chi)\Pi(\mathrm{d}\chi). \tag{4.3.3.1}$$

In view of (4.2.2.9), it is natural to conjecture that under Marton's inequality for $\nu$, $B_{\nu,\kappa}$ satisfies a transport inequality with one of the following costs:

$$c(\xi,\Pi) = \left(\int \mathcal{W}_{2,\rho,\infty}(\xi,\chi)\Pi(\mathrm{d}\chi)\right)^2 \quad \text{or} \quad c(\xi,\Pi) = \left(\int \mathcal{W}_{1,\rho,\infty}(\xi,\chi)\Pi(\mathrm{d}\chi)\right)^2. \tag{4.3.3.2}$$

So far we are only able to attain this result for binomial process of a fixed size and the weaker distance $\mathcal{W}_{1,\rho,0}$.

**Theorem 4.3.3.1.** *Let $n \in \mathbb{N}$ and $\nu \in \mathscr{P}(Z)$. Assume $\nu$ satisfies Marton's inequality (4.2.3.5) on $Z$ for a distance $\rho\colon Z \times Z \to [0,\infty]$. Then $B_{\nu,n}$ satisfies the transport-entropy inequality (4.2.3.2) on $\mathscr{M}_n(Z)$ with the cost function $c\colon \mathscr{M}_n(Z) \times \mathscr{P}(\mathscr{M}_n(Z)) \to \mathbb{R}_+$ defined, for all $\xi \in \mathscr{M}_n(Z)$ and $\Pi \in \mathscr{P}(\mathscr{M}_n(Z))$, by*

$$c(\xi,\Pi) = \int \frac{1}{\xi(x)^2}\left(\int \mathcal{W}_{1,\rho,0}(\chi, \xi(x)\delta_x)\Pi(\mathrm{d}\chi)\right)^2 \xi(\mathrm{d}x). \tag{4.3.3.3}$$

As corollaries, we obtain that binomial processes always satisfies a transport-entropy inequality and a concentration of measure inequality.

**Corollary 4.3.3.2.** *Let $\nu \in \mathscr{P}(Z)$ and $n \in \mathbb{N}$. Then, $B_{\nu,n}$ satisfies the transport-entropy inequality (4.2.3.2) with the cost function $c\colon \mathscr{M}_n(Z) \times \mathscr{P}(\mathscr{M}_n(Z)) \to \mathbb{R}_+$ defined, for all $\xi \in \mathscr{M}_n(Z)$ and $\Pi \in \mathscr{P}(\mathscr{M}_n(Z))$, by*

$$c(\xi,\Pi) = \frac{1}{4}\int\left(\int\left[1 - \frac{\chi(x)}{\xi(x)}\right]_+ \Pi(\mathrm{d}\chi)\right)^2 \xi(\mathrm{d}x). \tag{4.3.3.4}$$

*Proof.* By Theorem 4.3.1.2, the probability measure $\nu$ satisfies Marton's inequality (4.2.3.5) with the cost $\rho(x,y) = 1_{x\neq y}$. Let $\chi$ and $\xi \in \mathcal{M}_{\mathbb{N}}(Z)$. Then, we have that

$$\mathcal{W}_{1,\rho,0}(\chi, \xi(x)\delta_x) = \frac{1}{2}(\xi(x) - \chi(x))_+ 1_{\xi(x)\leq\chi(Z)}. \tag{4.3.3.5}$$

Now if $\chi(Z) = \xi(Z) = n$ the condition $\xi(x) \leq \chi(Z)$ is always true and therefore, by Theorem 4.3.3.1, we obtain that $B_{\nu,n}$ satisfies the transport-entropy inequality (4.2.3.2) with cost

$$c(\xi, \Pi) = \frac{1}{4}\int \left( \int \left[1 - \frac{\chi(x)}{\xi(x)}\right]_+ \Pi(\mathrm{d}\chi)\right)^2 \xi(\mathrm{d}x). \tag{4.3.3.6}$$

□

**Corollary 4.3.3.3.** *Let $\eta$ be a binomial process (of size $n$). Then for every measurable $A \subset \mathscr{M}_n(Z)$*

$$\mathbb{P}(\eta \in A)\mathbb{P}(\eta \notin A_t) \leq \mathrm{e}^{-t}, \tag{4.3.3.7}$$

*where $A_t$ is the enlargement of $A$ (in $\mathscr{M}_n(Z)$) given in (4.2.4.2) for the choice of $c$ given in (4.3.3.4), that is*

$$c(\xi, \Pi) = \frac{1}{4}\int \left( \int \left[1 - \frac{\chi(x)}{\xi(x)}\right]_+ \Pi(\mathrm{d}\chi)\right)^2 \xi(\mathrm{d}x), \quad \xi \in \mathscr{M}_n(Z),\, \Pi \in \mathscr{P}(\mathscr{M}_n(Z)). \tag{4.3.3.8}$$

*Let $\eta$ be a Poisson point process with finite intensity measure $\nu$. Then for every measurable $A \subset \mathscr{M}_{\mathbb{N}}(Z)$*

$$\mathbb{P}(\eta \in A)\mathbb{P}(\eta \notin A_t) \leq \mathrm{e}^{-t}, \tag{4.3.3.9}$$

*where $A_t$ is the enlargement of $A$ (in $\mathscr{M}_{\mathbb{N}}(Z)$) given in (4.2.4.2) for the choice of $c$ given in (4.3.3.4), that is*

$$c(\xi, \Pi) = \frac{1}{4}\int \left( \int \left[1 - \frac{\chi(x)}{\xi(x)}\right]_+ \Pi(\mathrm{d}\chi)\right)^2 \xi(\mathrm{d}x) \quad \xi \in \mathscr{M}_{\mathbb{N}}(Z),\, \Pi \in \mathscr{P}(\mathscr{M}_{\mathbb{N}}(Z)). \tag{4.3.3.10}$$

*Proof.* Combining Corollary 4.3.3.2 and Theorem 4.2.4.1, we obtain the announced result for binomial process. To go from binomial to Poisson, we recall the law of small numbers

$$B_{\nu,n}^{(\lambda/n)} \xrightarrow[n\to\infty]{TV} \Pi_{\lambda\nu}, \tag{4.3.3.11}$$

where for $t \in [0,1]$ and the law of a proper point process $B$, the symbol $B^{(t)}$ designates the $t$-thinning of $B$, that is the law of the point process obtained by discarding independently the point of $B$ with probability $1-t$. From the previous computations, we see that $B_{\nu,n}^{(\lambda/n)}$ also satisfies the concentration of measure and so does the Poisson point process by taking the limit. The reader can refer to [132] for details. □

*Proof of Theorem 4.3.3.1.* Recall that we now work on $\mathscr{M}_n(Z)$. By assumption and the tensorization property, the probability measure $\nu^n$ satisfies the transport inequality (4.2.3.2) on $Z^n$ with the cost function $c$

$$c(x,p) = \frac{1}{C}\sum_{i=1}^{n}\left(\int \rho(x_i,y)p_i(\mathrm{d}y)\right)^2, \qquad x \in Z^n,\ p \in \mathscr{P}(Z^n). \tag{4.3.3.12}$$

we consider the mapping $T\colon Z^n \to \mathscr{M}_n(Z)$ defined by

$$T(x) = \sum_{i=1}^{n}\delta_{x_i}, \quad x = (x_i) \in Z^n. \tag{4.3.3.13}$$

By construction of $T$, we have that $B_{\nu,n}$ is the push-forward of $\nu^n$ by $T$. Applying Proposition 4.3.1.1, we obtain that $B_{\nu,n}$ satisfies the transport-entropy inequality with the cost function $\bar{c}$ defined for all $(\xi,\Pi) \in \mathscr{M}_n(Z) \times \mathscr{P}(\mathscr{M}_n(Z))$ by

$$\bar{c}(\xi,\Pi) = \inf\left\{c(x,p),\ \text{such that } T(x)=\xi,\ T_{\#}p = \Pi\right\}. \tag{4.3.3.14}$$

In other words, if $\xi = \sum_{i=1}^{n}\delta_{a_i}$ and $(x_i)_{i=1}^{i=n} \in Z^n$, then

$$\bar{c}(\xi,\Pi) = \inf\left\{\frac{1}{C}\sum_{i=1}^{n}\left(\mathbb{E}\rho(Y_i,x_i)\right)^2\right\}, \tag{4.3.3.15}$$

where the infimum is running over all the random variables $Y = (Y_i)_{i=1}^{i=n} \in Z^n$ such that $\sum_{i=1}^{n}\delta_{Y_i} \sim \Pi$ and all sequences $(x_i)_{i=1}^{i=n} \in Z^n$ such that $\xi = \sum_{i=1}^{n}\delta_{x_i}$. Also, the constraint $\xi = \sum_{i=1}^{n}\delta_{x_i}$ determines the $x_i$'s (up to permutation). So that we obtain that

$$\bar{c}(\xi,\Pi) = \frac{1}{C}\inf\left\{\sum_{i=1}^{n}\left(\mathbb{E}\rho(Y_{\sigma(i)},a_i)\right)^2\right\}, \tag{4.3.3.16}$$

where the infimum runs over random variables $Y=(Y_i)$ such that $\sum_{i=1}^{n}\delta_{Y_i} \sim \Pi$ and $\sigma \in \Sigma_n$. For all $a \in Z$, such that $\xi(a) > 0$, we define $I(a) = \{i : a_i = a\}$. Then, by Jensen's inequality, we obtain

$$\begin{aligned}\sum_{i=1}^{n}\left(\mathbb{E}\rho(Y_{\sigma(i)},a_i)\right)^2 &= \sum_{a\in\xi}\xi(a)\frac{\sum_{i\in I(a)}\left(\mathbb{E}\rho(Y_{\sigma(i)},a_i)\right)^2}{\xi(a)}\\ &\geq \sum_{a\in\xi}\xi(a)\left(\frac{\sum_{i\in I(a)}\mathbb{E}\rho(Y_{\sigma(i)},a)}{\xi(a)}\right)^2.\end{aligned} \tag{4.3.3.17}$$

Let $\mu = \sum_{i=1}^{n}\delta_{Y_i}$. Eventually, by definition of (4.2.2.12) and since $\xi(a) \leq \xi(Z) = \mu(Z) = n$, we have that

$$\inf_{\sigma\in\Sigma_n}\sum_{i\in I(a)}\rho(Y_{\sigma(i)},a) = \mathcal{W}_{1,\rho,0}(\mu,\xi(a)\delta_a). \tag{4.3.3.18}$$

Combining (4.3.3.17) and (4.3.3.18), we obtain the announced cost. □

4.3.4. **Marton inequality for Poisson process?.** It is then very natural to conjecture the following theorem (that we cannot prove completely yet) but we will give some hints in this direction.

**Theorem 4.3.4.1.** *Let $\nu \in \mathscr{P}(Z)$. Then for all $\lambda > 0$, $\Pi_{\lambda\nu}$ satisfies the transport-entropy inequality (4.2.3.2) with the cost function defined in (4.3.3.4), that is $c\colon \mathscr{M}_{\mathbb{N}}(Z) \times \mathscr{P}(\mathscr{M}_{\mathbb{N}}(Z)) \to \mathbb{R}_+$ defined, for all $\xi \in \mathscr{M}_{\mathbb{N}}(Z)$ and $\Pi \in \mathscr{P}(\mathscr{M}_{\mathbb{N}}(Z))$, by*

$$c(\xi, \Pi) = \frac{1}{4} \int \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(\mathrm{d}\chi) \right)^2 \xi(\mathrm{d}x). \tag{4.3.4.1}$$

*Idea of proof of Theorem 4.3.4.1.* Applying Corollary 4.3.3.2 to $\hat{Z} = Z \cup \{\infty\}$, where $\infty$ is a cemetery point that is not in $Z$ and $\hat{\nu} = t\nu + (1-t)\delta_\infty$, we see that $B_{\nu,n}^{(t)}$ satisfies a transport-entropy inequality on $\mathscr{M}_{\leq n}(Z)$ with the cost given in (4.3.3.4), where the superscript $(t)$ indicates a thinning. We have that, for every $A \subset \mathscr{M}_{\mathbb{N}}(Z)$ measurable,

$$\Pi_n := B_{\nu,n}^{(\lambda/n)}(A) \xrightarrow[n\to\infty]{} \Pi_{\nu,N_\lambda}(A). \tag{4.3.4.2}$$

So that $\Pi_{\nu,n}^{\lambda/m} \to \Pi_{\lambda\nu}$ in this relatively strong sense. It is however not clear how to pass to the limit in the transport-entropy inequality

$$\mathcal{T}_c(P_1|P_2) \leq \mathcal{H}(P_1|\Pi_n) + \mathcal{H}(P_2|\Pi_n). \tag{4.3.4.3}$$

as $n \to \infty$. □

4.3.4.1. *Comparison with a result of Reitzner.* For further discussions, let us first recall some definitions and state a theorem of M. Reitzner (2013) [132] to be compared to Corollary 4.3.3.3. Given a point measure $\xi \in \mathscr{M}_{\mathbb{N}}(Z)$, we write $\mathscr{L}^2(\xi)$ for the Hilbert space of functions $\alpha\colon Z \to \mathbb{R}$ such that $\alpha(x) = 0$ for $x \notin \xi$ and we set

$$|\alpha|^2_{\mathscr{L}^2(\xi)} = \sum_{x\in\xi} \xi(x)\alpha(x)^2. \tag{4.3.4.4}$$

Note that because $\xi(Z) < \infty$, the previous quantity is always finite. Given another $\chi \in \mathscr{M}_{\mathbb{N}}(Z)$, we write $\xi \setminus \chi$ for the point measure given by $\sum_{x\in\xi} (\xi(x) - \chi(x))_+ \delta_x$. For $A \subset \mathscr{M}_{\mathbb{N}}(Z)$ and $\xi \in \mathscr{M}_{\mathbb{N}}(Z)$, M. Reitzner (2013) [132] defines

$$d_A(\xi) = \sup_{|\alpha|_{\mathscr{L}^2(\xi)} \leq 1} \inf_{\chi\in A} \int \alpha \mathrm{d}(\xi \setminus \chi), \tag{4.3.4.5}$$

where the supremum runs over non-negative $\alpha$ only and

$$A_t^d = \{\xi \in \mathscr{M}_{\mathbb{N}}(Z),\ d_A(\xi) \leq t\}, \qquad t \geq 0. \tag{4.3.4.6}$$

Then, we have:

**Theorem 4.3.4.2** ([132, Theorem 1.1])**.** *Let $\eta$ be a Poisson point process on $Z$ such that $\mathbb{E}\eta(Z) < \infty$. Then, for every measurable set $A \subset \mathscr{M}_{\mathbb{N}}(Z)$,*

$$\mathbb{P}(\eta \in A)\mathbb{P}(\eta \notin A_t^d) \leq e^{-t^2/4}. \tag{4.3.4.7}$$

Thanks to Corollary 4.3.3.3 and the following lemma, we obtain an immediate new proof of this theorem.

**Lemma 4.3.4.3.** *With the notation introduced before, let $A \subset \mathscr{M}_{\mathbb{N}}(Z)$ be measurable, then $A^d_{4\sqrt{t}} = A_t$, equivalently $c_A(\xi) = \frac{1}{4} d_A(\xi)^2$, for all $\xi \in \mathscr{M}_{\mathbb{N}}(Z)$.*

*Proof.* The proof is quite classical and goes back to Talagrand. We give a proof for completeness. First, we recall this well-known duality formula on Hilbert spaces: if $\mathfrak{H}$ is an Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $| \cdot |$, we have $|x| = \sup_{|y| \leq 1} \langle x, y \rangle$. Second, we recall this well-known fact about randomization of infimum

$$\inf_{x \in A} f(x) = \inf_{X \in A} \mathbb{E} f(X), \tag{4.3.4.8}$$

where the second infimum runs on random variables concentrated in $A$.
Third, for two linear spaces $V_1$ and $V_2$, a bilinear functional $\Lambda \colon V_1 \times V_2 \to \mathbb{R}$ and two convex sets $C_1$ and $C_2$ such that $C_1$ is compact we have by the min-max theorem of M. Sion (1958) [139, Corollary 3.3]

$$\sup_{C_1} \inf_{C_2} \Lambda = \inf_{C_2} \sup_{C_1} \Lambda. \tag{4.3.4.9}$$

From these three facts, we easily see that

$$\begin{aligned}
d_A(\xi) &= \sup_{|\alpha|_{\mathscr{L}^2(\xi)} \leq 1} \inf_{\Pi(A)=1} \int \int \alpha \mathrm{d}(\xi \setminus \nu) \, \Pi(d\nu) \\
&= \sup_{|\alpha|_{\mathscr{L}^2(\xi)} \leq 1} \inf_{\Pi(A)=1} \int \alpha(x) \left( \int \frac{(\xi \setminus \nu)(x)}{\xi(x)} \Pi(\mathrm{d}v) \right) \xi(\mathrm{d}x) \\
&= \inf_{\Pi(A)=1} \sup_{|\alpha|_{\mathscr{L}^2(\xi)} \leq 1} \int \alpha(x) \left( \int \frac{(\xi \setminus \nu)(x)}{\xi(x)} \Pi(\mathrm{d}v) \right) \xi(\mathrm{d}x) \\
&= \left( \int \left( \int \left( 1 - \frac{\nu(x)}{\xi(x)} \right)_+ \Pi(\mathrm{d}\nu) \right)^2 \xi(\mathrm{d}x) \right)^{1/2} \\
&= 2 c_A(\xi)^{1/2}.
\end{aligned}$$

Note that, since $\mathscr{L}^2(\xi)$ is a finite-dimensional space, the unit ball is compact. $\square$

# Multi set concentration of measure and multi-marginal transport

The following is a reproduction of an article to appear in Potential analysis.

## Introduction

Let $(M, g)$ be a smooth compact connected Riemannian manifold with its normalized volume measure $\mu$ and its geodesic distance $d$. The Laplace-Beltrami operator $\Delta$ is then a non-positive operator whose spectrum is discrete. Let us denote by $\lambda^{(k)}$, $k = 0, 1, 2 \ldots$, the eigenvalues of $-\Delta$ written in increasing order. With these notations $\lambda^{(0)} = 0$ (achieved for constant functions) and (by connectedness) $\lambda^{(1)} > 0$ is the so-called spectral gap of $M$.

The study of the spectral gap of Riemannian manifolds is, by now, a very classical topic which has found important connections with numerous geometrical and analytical questions and properties. The spectral gap constant $\lambda^{(1)}$ is for instance related to Poincaré type inequalities and governs the speed of convergence of the heat flow to equilibrium. It is also related to Ricci curvature via the classical Lichnerowicz theorem [88] and to Cheeger isoperimetric constant via Buser's theorem [29]. We refer to [13, 32] and the references therein for a complete picture.

Another important property of the spectral gap constant, first observed by Gromov and Milman [59], is that it controls exponential concentration of measure phenomenon for the reference measure $\mu$. The result states as follows. Define for all Borel sets $A \subset M$, its $r$-enlargement $A_r$ as the (open) set of all $x \in E$ such that there exists $y \in A$ with $d(x, y) < r$. Then, for any $A \subset M$ such that $\mu(A) \geq 1/2$ it holds

$$\mu(A_r) \geq 1 - be^{-a\sqrt{\lambda^{(1)}}r}, \qquad \forall r > 0, \tag{5.0.0.1}$$

where $a, b > 0$ are some universal constants (according to [85, Theorem 3.1], one can take $b = 1$ and $a = 1/3$). Note that this implication is very general and holds on any metric space supporting a Poincaré inequality (see [85, Corollary 3.2]). See also [19, 137, 2, 56] for alternative derivations, generalizations or refinements of this result.

This note is devoted to a multiple sets extension of the above result. Roughly speaking, we will see that if $A_1, \ldots, A_k$ are sets which are pairwise separated in the sense that $d(A_i, A_j) := \inf\{d(x, y) : x \in A_i, y \in A_j\} > 0$ for any $i \neq j$ and $A$ is their union then the probability of $A_r$ goes exponentially fast to $1$ at a rate given by $\sqrt{\lambda^{(k)}}$ as soon as $r$ is such that the sets $A_{i,r}$, $i = 1, \ldots, k$ remain separated. More precisely, it follows from Theorem 5.1.2.1 (whose setting is actually more general) that, if $A_1, \ldots, A_k$ are such that $\mu(A_i) \geq \frac{1}{k+1}$ and $d(A_{i,r}, A_{j,r}) > 0$ for all $i \neq j$, then, denoting $A = A_1 \cup \ldots \cup A_k$, it holds

$$\mu(A_r) \geq 1 - \frac{1}{k+1} \exp\left(-c \min(r^2 \lambda^{(k)}; r\sqrt{\lambda^{(k)}})\right), \tag{5.0.0.2}$$

for some universal constant $c$. This kind of probability estimates first appeared, in a slightly different but essentially equivalent formulation in the work of Chung, Grigor'yan and Yau [34] (see also the related paper [49] by Friedman and Tillich). Nevertheless, the method of proof we use to arrive at (5.0.0.2) (based on the Courant-Fischer min-max formula for the $\lambda^{(k)}$'s) is quite different from the one of [34] and seems more elementary and general. This is discussed in details in Section 5.1.5.

The paper is organized as follows. In Section 5.1, we prove (5.0.0.2) in an abstract metric space framework. This framework contains, in particular, the compact Riemannian case equipped with the Laplace operator presented above. The Section 5.1.5 contains a detailed discussion of our result with the one of Chung, Grigor'yan & Yau. In Section 5.2, we recall various bounds on eigenvalues on several non-negatively curved manifolds. Section 5.3 gives an extension of (5.0.0.2) to discrete Markov chains on graphs. In Section 5.4, we give a functional formulation of the results of Sections 5.1 and 5.3. As a corollary of this functional formulation, we obtain a deviation inequality as well as an estimate for difference of two Lipschitz extensions of a Lipschitz function given on $k$ subsets. Finally, Section 5.5 discusses open questions related to this type of concentration of measure phenomenon.

## 5.1. Multiple sets exponential concentration in abstract spaces

5.1.1. **Courant-Fischer formula and generalized eigenvalues in metric spaces.** Let us recall the classical Courant-Fischer min-max formula for the $k$-th eigenvalue ($k \in \mathbb{N}$) of $-\Delta$, noted $\lambda^{(k)}$, on a compact Riemannian manifold $(M, g)$ equipped with its (normalized) volume measure $\mu$:

$$\lambda^{(k)} = \inf_{\substack{V \subset \mathcal{C}^\infty(M) \\ \dim V = k+1}} \sup_{f \in V \setminus \{0\}} \frac{\int |\nabla f|^2 \, \mathrm{d}\mu}{\int f^2 \, \mathrm{d}\mu}, \tag{5.1.1.1}$$

where $\nabla f$ is the Riemannian gradient, defined through the Riemannian metric $g$ (see *e.g* [32]) and $|\nabla f|^2 = g(\nabla f, \nabla f)$. The formula (5.1.1.1) above does not make explicitly reference to the differential operator $\Delta$. It can be therefore easily generalized to a more abstract setting, as we shall see below.

In all what follows, $(E, d)$ is a complete, separable metric space and $\mu$ a reference Borel probability measure on $E$. Following [33], for any function $f \colon E \to \mathbb{R}$ and $x \in E$, we denote by $|\nabla f|(x)$ the *local Lipschitz constant* of $f$ at $x$, defined by

$$|\nabla f|(x) = \begin{cases} 0 \text{ if } x \text{ is isolated} \\ \limsup_{y \to x} \frac{|f(x)-f(y)|}{d(x,y)} \text{ otherwise.} \end{cases} \tag{5.1.1.2}$$

Note that when $E$ is a smooth Riemannian manifold, equipped with its geodesic distance $d$, then, the local Lipschitz constant of a differentiable function $f$ at $x$ coincides with the norm of $\nabla f(x)$ in the tangent space $T_x E$. With this notion in hand, a natural generalization of (5.1.1.1) is as follows (we follow [100, Definition 3.1]):

$$\lambda^{(k)}_{d,\mu} := \inf_{\substack{V \subset H^1(\mu) \\ \dim V = k+1}} \sup_{f \in V \setminus \{0\}} \frac{\int |\nabla f|^2 \, \mathrm{d}\mu}{\int f^2 \, \mathrm{d}\mu}, \qquad k \geq 0, \tag{5.1.1.3}$$

where $H^1(\mu)$ denotes the space of functions $f \in L^2(\mu)$ such that $\int |\nabla f|^2 \, \mathrm{d}\mu < +\infty$. In order to avoid heavy notations, we drop the subscript and we simply write $\lambda^{(k)}$ instead of $\lambda^{(k)}_{d,\mu}$ within this section.

5.1.2. **Statement of the main results.** To state our first main result, we need further notations: for any $k \geq 1$, we denote by $\Delta_k$ the set of vectors $(a_1, \dots, a_k) \in [0,1]^k$ satisfying the following linear constraints

$$\sum_{j=1}^{k} a_j \leq 1 \qquad \text{and} \qquad a_i + \sum_{j=1}^{k} a_j \geq 1,\ \forall i \in \{1, \dots, k\}. \tag{5.1.2.1}$$

Recall the classical notation $d(A, B) = \inf\{d(x, y) : x \in A, y \in B\}$ of the distance between two sets $A, B \subset E$.

The following theorem is the main result of the paper and is proved in Section 5.1.3.

**Theorem 5.1.2.1.** *There exists a universal constant $c > 0$ such that, for any $k \geq 1$ and for all sets $A_1, \dots, A_k \subset E$ such that $\min_{i \neq j} d(A_i, A_j) > 0$ and $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$, the set $A = A_1 \cup A_2 \cup \dots \cup A_k$ satisfies*

$$\mu(A_r) \geq 1 - (1 - \mu(A)) \exp\left(-c \min(r^2 \lambda^{(k)}; r\sqrt{\lambda^{(k)}})\right), \tag{5.1.2.2}$$

*for all $0 < r \leq \frac{1}{2} \min_{i \neq j} d(A_i, A_j)$, where $\lambda^{(k)} \geq 0$ is defined by (5.1.1.3).*

Note that, since $(1/(k+1), \dots, 1/(k+1)) \in \Delta_k$, Theorem 5.1.2.1 immediately implies Inequality (5.0.0.2).

Inverting our concentration estimate, we obtain the following statement that provides a bound on the $\lambda^{(k)}$'s.

**Proposition 5.1.2.2.** *Let $(E, d, \mu)$ be a metric measured space and $\lambda^{(k)}$ be defined as in (5.1.1.3). Let $A_1, \dots, A_k$ be measurable sets such that $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$, then, by letting $r = \frac{1}{2} \min_{i \neq j} d(A_i, A_j)$ and $A_0 = E \setminus (\cup A_i)_r$,*

$$\lambda^{(k)} \leq \frac{1}{r^2} \psi\left(\frac{1}{c} \min_i \ln \frac{\mu(A_i)}{\mu(A_0)}\right), \tag{5.1.2.3}$$

*where $\psi(x) = \max(x, x^2)$.*

*Proof.* Let $A = \cup_i A_i$. Inverting the formula in Theorem 5.1.2.1, we obtain

$$\lambda^{(k)} \leq \frac{1}{r^2} \psi\left(\frac{1}{c} \ln \frac{1 - \mu(A)}{1 - \mu(A_r)}\right), \tag{5.1.2.4}$$

where $\psi(x) = \max(x, x^2)$. By definition of $\Delta_k$,

$$1 - \mu(A) = 1 - \sum_i \mu(A_i) \leq \min_i \mu(A_i). \tag{5.1.2.5}$$

Therefore, letting $A_0 = E \setminus A_r$, we obtain the announced inequality by non-decreasing monotonicity of $\psi$ and $\ln$. $\square$

The collection of sets $\Delta_k$, $k \geq 1$ has the following useful stability property:

**Lemma 5.1.2.3.** *Let $I_1, I_2, \dots, I_n$ be a partition of $\{1, \dots, k\}$, $k \geq 1$. Let $a = (a_1, \dots, a_k) \in \mathbb{R}^k$ and define $b = (b_1, \dots, b_n) \in \mathbb{R}^n$ by setting $b_i = \sum_{j \in I_i} a_j$, $i \in \{1, \dots, n\}$. If $a \in \Delta_k$ then $b \in \Delta_n$.*

*Proof.* The proof is obvious and left to the reader. $\square$

Thanks to this lemma it is possible to iterate Theorem 5.1.2.1 and to obtain a general bound for $\mu(A_r)$ for all values of $r > 0$. This bound will depend on the way the sets $A_{1,r}, \ldots, A_{k,r}$ coalesce as $r$ increases. This is made precise in the following definition.

**Definition 5.1.1** (Coalescence graph of a family of sets)**.** Let $A_1, \ldots, A_k$ be subsets of $E$. The *coalescence graph* of this family of sets is the family of graphs $G_r = (V, E_r)$, $r > 0$, where $V = \{1, 2, \ldots, k\}$ and the set of edges $E_r$ is defined as follows: $\{i, j\} \in E_r$ if $d(A_{i,r}, A_{j,r}) = 0$.

**Corollary 5.1.2.4.** *Let $A_1, \ldots, A_k$ be measurable subsets of $E$ such that $\min_{i \neq j} d(A_i, A_j) > 0$ and $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$. For any $r > 0$, let $N(r)$ be the number of connected components in the coalescence graph $G_r$ associated to $A_1, \ldots, A_k$. The function $(0, \infty) \to \{1, \ldots, k\} : r \mapsto N(r)$ is non-increasing and right-continuous. Define $r_i = \sup\{r > 0 : N(r) \geq k - i + 1\}$, $i = 1, \ldots, k$ and $r_0 = 0$ then it holds*

$$\mu(A_r) \geq 1 - (1 - \mu(A)) \exp\left( -c \sum_{i=1}^{k} \phi\left( [r \wedge r_i - r_{i-1}]_+ \sqrt{\lambda^{(k-i+1)}} \right) \right), \; \forall r > 0, \tag{5.1.2.6}$$

*where $\phi(x) = \min(x; x^2)$, $x \geq 0$ and $c$ is the universal constant appearing in Theorem 5.1.2.1.*

Observe that, contrary to usual concentration results, the bound given above depends on the geometry of the set $A$.

5.1.3. **Proofs.** First, we prove Corollary 5.1.2.4. The main argument is to repeatedly apply Theorem 5.1.2.1 until two sets or more coalesce.

*Proof of Corollary 5.1.2.4.* We proceed by induction over the number of components $k$. For $k = 1$, (5.1.2.6) follows immediately from Theorem 5.1.2.1. Let $k > 1$ and let us assume that (5.1.2.6) is true for any collection of subsets $B_1, \ldots, B_l$ satisfying the assumptions of Corollary 5.1.2.4 for all $l \in \{1, \ldots, k-1\}$. Let $A_1, A_2, \ldots, A_k$ be a collection of sets satisfying the assumptions of Corollary 5.1.2.4. According to Theorem 5.1.2.1, it holds

$$\mu(A_r) \geq 1 - (1 - \mu(A)) \exp\left( -c\phi(r\sqrt{\lambda^{(k)}}) \right), \tag{5.1.3.1}$$

for all $0 < r \leq \frac{1}{2} \min_{i \neq j} d(A_i, A_j)$.

Let $k_1 = N(\frac{1}{2} \min_{i \neq j} d(A_i, A_j))$ and let $i_1 = k - k_1$. Then, for all $i \in \{1, \ldots, i_1\}$, $r_i = \frac{1}{2} \min_{i \neq j} d(A_i, A_j)$. So that, for all $0 < r \leq r_{i_1}$, the preceding bound can be rewritten as follows (note that only the term of index $i = 1$ gives a non zero contribution)

$$\begin{aligned} \mu(A_r) &\geq 1 - (1 - \mu(A)) \exp\left( -c \sum_{i=1}^{i_1} \phi\left( [r \wedge r_i - r_{i-1}]_+ \sqrt{\lambda^{(k-i+1)}} \right) \right) \\ &= 1 - (1 - \mu(A)) \exp\left( -c \sum_{i=1}^{k} \phi\left( [r \wedge r_i - r_{i-1}]_+ \sqrt{\lambda^{(k-i+1)}} \right) \right) \end{aligned} \tag{5.1.3.2}$$

which shows that (5.1.2.6) is true for $0 < r \leq r_{i_1}$. Now let $I_1, \ldots, I_{k_1}$ be the connected components of $G_{r_1}$ and define, for all $i \in \{1, \ldots, k_1\}$, $B_i = \cup_{j \in I_i} A_{j,r_1}$. It follows easily from Lemma 5.1.2.3 that $(\mu(B_1), \ldots, \mu(B_{k_1})) \in \Delta_{k_1}$. Since $\min_{i\neq j} d(B_i, B_j) > 0$, the induction hypothesis implies that

(5.1.3.3)
$$\mu(B_s) \geq 1 - (1-\mu(B)) \exp\left(-c\sum_{i=1}^{k_1} \phi\left([s\wedge s_i - s_{i-1}]_+ \sqrt{\lambda^{(k_1-i+1)}}\right)\right), \qquad \forall s > 0,$$

where $B = B_1 \cup \cdots \cup B_{k_1} = A_{r_1}$ and $s_i = \sup\{s > 0 : N'(s) \geq k_1 - i + 1\}$, $i \in \{1, \ldots, k_1\}$ ($s_0 = 0$) with $N'(s)$ the number of connected components in the graph $G'_s$ associated to $B_1, \ldots, B_{k_1}$. It is easily seen that $r_{i_1+i} = r_{i_1} + s_i$, for all $i \in \{0, 1 \ldots, k_1\}$. Therefore, we have that, for $r > r_{i_1}$,

$$\begin{aligned}\mu(A_r) &\geq \mu(B_{r-r_{i_1}}) \\ &\geq 1 - (1-\mu(A_{r_{i_1}})) \exp\left(-c\sum_{i=i_1+1}^{k} \phi\left([r\wedge r_i - r_{i-1}]_+ \sqrt{\lambda^{(k-i+1)}}\right)\right) \\ &\geq 1 - (1-\mu(A)) \exp\left(-c\sum_{i=1}^{k} \phi\left([r\wedge r_i - r_{i-1}]_+ \sqrt{\lambda^{(k-i+1)}}\right)\right),\end{aligned}$$

where the last line is true by (5.1.3.2). □

To prove Theorem 5.1.2.1, we need some preparatory lemmas. Given a subset $A \subset E$, and $x \in E$, the minimal distance from $x$ to $A$ is denoted by

(5.1.3.4)
$$d(x, A) = \inf_{y\in A} d(x, y).$$

**Lemma 5.1.3.1.** *Let $A \subset E$ and $\epsilon > 0$, then $(E \setminus A_\epsilon)_\epsilon \subset E \setminus A$.*

*Proof.* Let $x \in (E \setminus A_\epsilon)_\epsilon$. Then, there exists $y \in E \setminus A_\epsilon$ (in particular $d(y, A) \geq \epsilon$) such that $d(x, y) < \epsilon$. Since the function $z \mapsto d(z, A)$ is 1-Lipschitz, one has

(5.1.3.5)
$$d(x, A) \geq d(y, A) - d(x, y) > 0$$

and so $x \in E \setminus A$. □

**Remark 29.** In fact, we proved that $(E \setminus A_\epsilon)_\epsilon \subset E \setminus \bar{A}$. The converse is, in general, not true.

**Lemma 5.1.3.2.** *Let $A_1, \ldots, A_k$ be a family of sets such that $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$ and $r := \frac{1}{2}\min_{i\neq j} d(A_i, A_j) > 0$. Let $0 < \epsilon \leq r$ and set $A = \cup_{1\leq i\leq k} A_i$ and $A_0 = E \setminus (A_\epsilon)$. Then,*

(5.1.3.6)
$$\max_{i=0,\ldots,k} \frac{\mu(A_{i,\epsilon})}{\mu(A_i)} \leq \frac{1-\mu(A)}{1-\mu(A_\epsilon)}.$$

*Proof.* First, this is true for $i = 0$. Indeed, by definition $A_0 = E \setminus (A_\epsilon)$ and, according to Lemma 5.1.3.1, $(A_0)_\epsilon \subset A^c$ (the equality is not always true), which proves (5.1.3.6) in this case. Now, let us show (5.1.3.6) for the other values of $i$. Since $\epsilon \leq r$, the $A_{j,\epsilon}$'s are disjoint sets. Thence, (5.1.3.6) is equivalent to

(5.1.3.7)
$$\left(1 - \sum_{j=1}^{k} \mu(A_{j,\epsilon})\right) \mu(A_{i,\epsilon}) \leq \left(1 - \sum_{j=1}^{k} \mu(A_j)\right) \mu(A_i).$$

This inequality is true as soon as

$$(1 - \mu(A_{i,\epsilon}) - m_i)\, \mu(A_{i,\epsilon}) \le (1 - \mu(A_i) - m_i)\, \mu(A_i), \tag{5.1.3.8}$$

denoting $m_i = \sum_{j \neq i}^k \mu(A_j)$. The function $f_i(u) = (1-u-m_i)u$, $u \in [0,1]$, is decreasing on the interval $[(1-m_i)/2, 1]$. We conclude from this that (5.1.3.6) is true for all $i \in \{1, \dots, k\}$, as soon as $\mu(A_i) \ge (1-m_i)/2$ for all $i \in \{1, \dots, k\}$ which amounts to $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$. □

For $p > 1$, we define the function $\chi_p \colon \mathbb{R}_+ \to [0,1]$ by

$$\chi_p(x) = (1 - x^p)^p, \quad \text{for} \quad x \in [0,1] \quad \text{and} \quad \chi_p(x) = 0 \quad \text{for} \quad x > 1. \tag{5.1.3.9}$$

It is easily seen that $\chi_p(0) = 1$, $\chi_p'(0) = \chi_p(1) = \chi_p'(1) = 0$, that $\chi_p$ takes values in $[0,1]$ and that $\chi_p$ is continuously differentiable on $\mathbb{R}_+$. We use the function $\chi_p$ to construct smooth approximations of indicator functions on $E$, as explained in the next statement.

**Lemma 5.1.3.3.** *Let $A \subset E$ and consider the function $f(x) = \chi_p(d(x,A)/\epsilon)$, $x \in E$, where $\epsilon > 0$ and $p > 1$. For all $x \in E$, it holds*

$$|\nabla f|(x) \le p^2 \epsilon^{-1} \mathbf{1}_{A_\epsilon \setminus A} \tag{5.1.3.10}$$

*Proof.* Thanks to the chain rule for the local Lipschitz constant (see e.g. [3, Proposition 2.1]),

$$\left| \nabla \chi_p \left( \frac{d(\cdot, A)}{\epsilon} \right) \right| (x) \le \epsilon^{-1} \chi_p' \left( \frac{d(\cdot, A)}{\epsilon} \right) |\nabla d(\cdot, A)|(x). \tag{5.1.3.11}$$

The function $d(\cdot, A)$ being Lipschitz, its local Lipschitz constant is $\le 1$ and, thereby,

$$|\nabla f|(x) \le \chi_p' \left( \frac{d(x, A)}{\epsilon} \right). \tag{5.1.3.12}$$

In particular, thanks to the aforementioned properties of $\chi$, $|\nabla f|$ vanishes on $A$ (and even on $\overline{A}$) and on $\{x \in E : d(x,A) \ge \epsilon\} = E \setminus A_\epsilon$. On the other hand, a simple calculation shows that $|\chi_p'| \le p^2$ which proves the claim. □

*Proof of Theorem 5.1.2.1.* Take Borel sets $A_1, \dots, A_k$ with $\frac{1}{2} \min_{i \neq j} d(A_i, A_j) \ge r > 0$ and $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$ and consider $A = A_1 \cup \dots \cup A_k$. Let us show that, for any $0 < \epsilon \le r$, it holds

$$\left(1 + \lambda^{(k)} \epsilon^2\right) (1 - \mu(A_\epsilon)) \le (1 - \mu(A)). \tag{5.1.3.13}$$

Let $A_0 = E \setminus (A_\epsilon)$ and set $f_i(x) = \chi_p(d(x, A_i)/\epsilon)$, $x \in E$, $i \in \{0, \dots, k\}$, where $p > 1$. According to Lemma 5.1.3.3 and the fact that $f_i = 1$ on $A_i$, we obtain

$$\int |\nabla f_i|^2 \,\mathrm{d}\mu = \frac{p^4}{\epsilon^2} \mu(A_{i,\epsilon} \setminus A_i) \quad \text{and} \quad \int f_i^2 \,\mathrm{d}\mu \ge \mu(A_i). \tag{5.1.3.14}$$

Since the $f_i$'s have disjoint supports they are orthogonal in $L^2(\mu)$ and, in particular, they span a $k+1$ dimensional subspace of $H^1(\mu)$. Thus, by definition of $\lambda^{(k)}$,

$$\lambda^{(k)} \leq \sup_{a\in\mathbb{R}^{k+1}} \frac{\int |\nabla \left(\sum_{i=0}^k a_i f_i\right)|^2 \, \mathrm{d}\mu}{\int \left(\sum_{i=0}^k a_i f_i\right)^2 \mathrm{d}\mu} \leq \sup_{a\in\mathbb{R}^{k+1}} \frac{\int \left(\sum_{i=0}^k |a_i||\nabla f_i|\right)^2 \mathrm{d}\mu}{\int \left(\sum_{i=0}^k a_i f_i\right)^2 \mathrm{d}\mu}, \tag{5.1.3.15}$$

where the second inequality comes from the following easy to check sub-linearity property of the local Lipschitz constant:

$$|\nabla (af+bg)| \leq |a||\nabla f| + |b||\nabla g|. \tag{5.1.3.16}$$

Since the $f_i's$ and the $|\nabla f_i|'s$ are two orthogonal families, we conclude using (5.1.3.14), that

$$\frac{\lambda^{(k)}\epsilon^2}{p^4} \leq \sup_{a\in\mathbb{R}^{k+1}} \frac{\sum_{i=0}^k a_i^2 \left(\mu(A_{i,\epsilon}) - \mu(A_i)\right)}{\sum_{i=0}^k a_i^2 \mu(A_i)}, \tag{5.1.3.17}$$

which amounts to

$$1 + \frac{\lambda^{(k)}\epsilon^2}{p^4} \leq \max_{i=0,\ldots,k} \frac{\mu(A_{i,\epsilon})}{\mu(A_i)}. \tag{5.1.3.18}$$

Applying Lemma 5.1.3.2 and sending $p$ to $1$ gives (5.1.3.13). Now, if $n \in \mathbb{N}$ and $0 < \epsilon$ are such that $n\epsilon \leq r$, then iterating (5.1.3.13) immediately gives

$$\left(1+\lambda^{(k)}\epsilon^2\right)^n (1-\mu(A_{n\epsilon})) \leq 1-\mu(A). \tag{5.1.3.19}$$

Optimizing this bound over $n$ for a fixed $\varepsilon$ gives

$$(1-\mu(A_r)) \leq (1-\mu(A)) \exp\left(-\sup\left\{\lfloor r/\epsilon \rfloor \log\left(1+\lambda^{(k)}\epsilon^2\right) : \epsilon \leq r\right\}\right). \tag{5.1.3.20}$$

Thus, letting

$$\Psi(x) = \sup\left\{\lfloor t \rfloor \log\left(1+\frac{x}{t^2}\right) : t \geq 1\right\}, \qquad x \geq 0, \tag{5.1.3.21}$$

it holds

$$(1-\mu(A_r)) \leq (1-\mu(A)) \exp\left(-\Psi\left(\lambda^{(k)} r^2\right)\right). \tag{5.1.3.22}$$

Using Lemma 5.1.3.4 below, we deduce that $\Psi\left(\lambda^{(k)} r^2\right) \geq c \min(r^2\lambda^{(k)}; r\sqrt{\lambda^{(k)}})$, with $c = \log(5)/4$, which completes the proof. □

**Lemma 5.1.3.4.** *The function $\Psi$ defined by* (5.1.3.21) *satisfies*

$$\Psi(x) \geq \frac{\log(5)}{4} \min(x; \sqrt{x}), \qquad \forall x \geq 0. \tag{5.1.3.23}$$

*Proof.* Taking $t = 1$, one concludes that $\Psi(x) \geq \log(1 + x)$, for all $x \geq 0$. The function $x \mapsto \log(1 + x)$ being concave, the function $x \mapsto \frac{\log(1+x)}{x}$ is non-increasing. Therefore, $\log(1+x) \geq \frac{\log(5)}{4} x$ for all $x \in [0, 4]$. Now, let us consider the case where $x \geq 4$. Observe that $\lfloor t \rfloor \geq t/2$ for all $t \geq 1$ and so, for $x \geq 4$,

$$\Psi(x) \geq \frac{1}{2} \sup_{t \geq 1} \left\{ t \log\left(1 + \frac{x}{t^2}\right) \right\} \geq \frac{\log(5)}{4} \sqrt{x}, \tag{5.1.3.24}$$

by choosing $t = \sqrt{x}/2 \geq 1$. Thereby,

$$\Psi(x) \geq \frac{\log(5)}{4} \left[ x \mathbf{1}_{0 \leq x \leq 4} + \sqrt{x} \mathbf{1}_{x > 4} \right] \geq \frac{\log(5)}{4} \min(x; \sqrt{x}), \tag{5.1.3.25}$$

which completes the proof. $\square$

**Remark 30.** The conclusion of Lemma Lemma 5.1.3.4 can be improved. Namely, it can be shown that
(5.1.3.26)

$$\Psi(x) = \max\left( \left(1 + \lfloor \frac{\sqrt{x}}{a} \rfloor \right) \log\left( 1 + \frac{x}{\left(1 + \lfloor \frac{\sqrt{x}}{a} \rfloor\right)^2} \right) ; \left( \lfloor \frac{\sqrt{x}}{a} \rfloor \right) \log\left( 1 + \frac{x}{\left( \lfloor \frac{\sqrt{x}}{a} \rfloor \right)^2} \right) \right),$$

(the second term in the maximum being treated as $0$ when $\sqrt{x} < a$) where $0 < a < 2$ is the unique point where the function $(0, \infty) \to \mathbb{R} : u \mapsto \log(1 + u^2)/u$ achieves its supremum. Therefore,

$$\Psi(x) \sim \frac{\log(1 + a^2)}{a} \sqrt{x} \tag{5.1.3.27}$$

when $x \to \infty$. The reader can easily check that $\frac{\log(1+a^2)}{a} \simeq 0.8$. In particular, it does not seem possible to reach the constant $c = 1$ in Theorem 5.1.2.1 using this method of proof.

5.1.4. **Two more multi-set concentration bounds.** The condition $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$ can be seen as the multi-set generalization of the condition, standard in concentration of measure, that the size of the enlarged set has to be bigger than $1/2$. Indeed, the reader can easily verify that $(\frac{1}{k+1}, \ldots, \frac{1}{k+1}) \in \Delta_k$. However, in practice, this condition can be difficult to check. We provide two more multi-set concentration inequalities that hold in full generality. The method of proof is the same as for Theorem 5.1.2.1 and is based on (5.1.3.18).

**Proposition 5.1.4.1.** *Let $(E, d, \mu)$ be a metric measured space and $\lambda^{(k)}$ be defined as in (5.1.1.3). Let $(A_1, \ldots, A_k)$ be $k$ Borel sets, $A = \cup_i A_i$ and $A_0 = E \backslash A_r$. Then, with $a_{(1)} = \min_{1 \leq i \leq k} \mu(A_i)$, the following two bounds hold:*

$$1 - \mu(A_r) \leq (1 - \mu(A)) \frac{1}{\prod_{i=1}^k \mu(A_i)} \exp\left( -c \min\left( r^2 \lambda^{(k)}, r\sqrt{\lambda^{(k)}} \right) \right);$$
$$1 - \mu(A_r) \leq (1 - \mu(A)) \frac{1}{\mu(A)^{\mu(A)/a_{(1)}}} \exp\left( -c \min\left( r^2 \lambda^{(k)}, r\sqrt{\lambda^{(k)}} \right) \right).$$

*Proof.* Fix $N \in \mathbb{N}$ and $\epsilon > 0$ such that $N\epsilon \leq r$. For $i = 1, \ldots, k$ and $n \leq N$, we define

$$\begin{aligned}
\alpha_i(n) &= \frac{\mu(A_{i,n\epsilon})}{\mu(A_{i,(n-1)\epsilon})};\\
M_n &= \max_{1\leq i\leq k} \alpha_i(n) \vee \frac{1-\mu(A_{(n-1)\epsilon})}{1-\mu(A_{n\epsilon})};\\
L_n &= \{i \in \{1,\ldots,k\} | M_n = \alpha_i(n)\};\\
N_i &= \sharp\{n \in \{1,\ldots,N\} | i = \inf L_n\};\\
N_0 &= N - \sum_{i=1}^{k} N_i.
\end{aligned}$$

Roughly speaking, the number $N_i$ ($0 \leq i \leq k$) counts the number of time where the set $A_i$ growths in iterating (5.1.3.18). Lemma 5.1.3.2 asserts that in the case where $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$, then $N_0 = N$. However, we still obtain from (5.1.3.18), for $1 \leq i \leq k$,

(5.1.4.1) $$\frac{1}{\mu(A_i)} \geq \prod_{n=1}^{N} \alpha_i(n) \geq \left(1 + \lambda^{(k)}\epsilon^2\right)^{N_i}.$$

The first inequality is true because $\mu(A_{i,N\epsilon}) \leq 1$ and a telescoping argument. The second inequality is true because, as $n$ ranges from $1$ to $N$, by definition of the number $N_i$ and (5.1.3.18), there are, at least $N_i$ terms appearing in the product that can be bounded by $(1 + \lambda^{(k)}\epsilon^2)$. The other terms are bounded above by $1$. The case of $i = 0$ is handled in a similar fashion and we obtain:

(5.1.4.2) $$\begin{aligned}
1 - \mu(A_{N\epsilon}) &\leq (1-\mu(A))\left(1+\lambda^{(k)}\epsilon^2\right)^{-N_0}\\
&= (1-\mu(A))\left(1+\lambda^{(k)}\epsilon^2\right)^{-N}\prod_{i=1}^{k}\left(1+\lambda^{(k)}\epsilon^2\right)^{N_i}.
\end{aligned}$$

The announced bounds will be obtain by bounding the product appearing in the right-hand side and an argument similar to the end of the proof of Theorem 5.1.2.1. From (5.1.4.1), we have that,

(5.1.4.3) $$\prod_{i=1}^{k}\left(1+\lambda^{(k)}\epsilon^2\right)^{N_i} \leq \frac{1}{\prod_{i=1}^{k}\mu(A_i)}.$$

Also, from (5.1.4.1),

(5.1.4.4) $$\mu(A_{i,N\epsilon}) \geq \left(1+\lambda^{(k)}\epsilon^2\right)^{N_i}\mu(A_i).$$

Because $N\epsilon \leq r$, the sets $A_{1,N\epsilon}, \ldots, A_{k,N\epsilon}$ are pairwise disjoint and, thereby,

(5.1.4.5) $$1 \geq \sum \mu(A_{i,N\epsilon}) \geq \sum_{i=1}^{k}\left(1+\lambda^{(k)}\epsilon^2\right)^{N_i}\mu(A_i).$$

Fix $\theta > 0$ to be chosen later. By convexity of $\exp$,

$$\begin{aligned}
1 + (1-\mu(A))\left(1+\lambda^{(k)}\epsilon^2\right)^{\theta} &\geq \exp\left(\left(\sum_{i=1}^{k}\mu(A_i)N_i + (1-\mu(A))\theta\right)\log\left(1+\lambda^{(k)}\epsilon^2\right)\right)\\
&\geq \exp\left(\left(a_{(1)}\sum_{i=1}^{k}N_i + (1-\mu(A))\theta\right)\log\left(1+\lambda^{(k)}\epsilon^2\right)\right).
\end{aligned}$$

Finally, with $p = 1 - \mu(A)$ and $t = \theta \log(1 + \lambda^{(k)}\epsilon^2)$, we obtain

$$\prod_{i=1}^{k} \left(1 + \lambda^{(k)}\epsilon^2\right)^{N_i} \leq \left(\mathrm{e}^{-pt} + p\,\mathrm{e}^{(1-p)t}\right)^{1/a_{(1)}}. \tag{5.1.4.6}$$

We easily check that, the quantity in the right-hand side is minimal for $t = \log \frac{1}{1-p}$ at which it takes the value $(1-p)^{p-1} = \mu(A)^{-\mu(A)/a_{(1)}}$. Thus,

$$\prod_{i=1}^{k} (1 + \lambda^{(k)}\epsilon^2)^{N_i} \leq \frac{1}{\mu(A)^{\mu(A)/a_{(1)}}}. \tag{5.1.4.7}$$

Combining (5.1.4.3) and (5.1.4.7) with (5.1.4.2) and the same argument as for (5.1.3.21), we obtain the two announced bounds. □

From Proposition 5.1.4.1, we can derive bounds on the $\lambda^{(k)}$'s. The proof is the same as the one of Proposition 5.1.2.2 and is omitted.

**Proposition 5.1.4.2.** *Let $(E, d, \mu)$ be a metric measured space and $\lambda^{(k)}$ be defined as in (5.1.1.3). Let $A_1, \ldots, A_k$ be measurable sets, then, with $r = \frac{1}{2}\min_{i \neq j} d(A_i, A_j)$ and $A_0 = E \setminus (\cup A_i)_r$,*

$$\lambda^{(k)} \leq \frac{1}{r^2}\psi\left(\frac{1}{c}\ln\frac{a_{(1)}}{\mu(A_0)} + \frac{1}{c}k\ln\frac{1}{a_{(1)}}\right);$$
$$\lambda^{(k)} \leq \frac{1}{r^2}\psi\left(\frac{1}{c}\ln\frac{a_{(1)}}{\mu(A_0)} + \frac{1}{c}\frac{\mu(A)}{a_{(1)}}\ln\frac{1}{\mu(A)}\right),$$

*where $\psi(x) = \max(x, x^2)$ and $a_{(1)} = \min_{1 \leq i \leq k} \mu(A_i)$.*

5.1.5. **Comparison with the result of Chung-Grigor'yan-Yau.** In [34], the authors obtained the following result:

**Theorem 5.1.5.1** (Chung-Grigoryan-Yau [34]). *Let $M$ be a compact connected smooth Riemannian manifold equipped with its geodesic distance $d$ and normalized Riemannian volume $\mu$. For any $k \geq 1$ and any family of sets $A_0, \ldots, A_k$, it holds*

$$\lambda^{(k)} \leq \frac{1}{\min_{i \neq j} d^2(A_i, A_j)} \max_{i \neq j} \log\left(\frac{4}{\mu(A_i)\mu(A_j)}\right)^2, \tag{5.1.5.1}$$

*where $1 = \lambda^{(0)} \leq \lambda^{(1)} \leq \cdots \lambda^{(k)} \leq \cdots$ denotes the discrete spectrum of $-\Delta$.*

Let us translate this result in terms of concentration of measure. Let $A_1, \ldots, A_k$ be sets such that $r = \frac{1}{2}\min_{1 \leq i < j \leq k} d(A_i, A_j) > 0$ and define $A = A_1 \cup \cdots \cup A_k$ and $A_0 = M \setminus A_s$, for some $0 < s \leq r$. Then, applying (5.1.5.1) to this family of $k+1$ sets gives the following inequality

$$\min\left(a_{(2)}; 1 - \mu(A_s)\right) \leq \frac{4}{a_{(1)}}\exp(-\sqrt{\lambda^{(k)}}s), \tag{5.1.5.2}$$

with $a_{(1)}$ and $a_{(2)}$ being respectively the smallest number and the second smallest number among $(\mu(A_1), \ldots, \mu(A_k))$ (counted with multiplicity). Note that the right hand side is less than or equal to $a_{(2)}$ if and only if $s \geq s_o := \frac{1}{\sqrt{\lambda_k}} \log\left(\frac{4}{a_{(1)}a_{(2)}}\right)$, so that (5.1.5.2) is equivalent to the following statement:

$$\mu(A_s) \geq 1 - \frac{4}{a_{(1)}} \exp(-\sqrt{\lambda^{(k)}}s), \qquad \forall s \in [\min(s_o, r); r]. \tag{5.1.5.3}$$

We note that (5.1.5.3) holds for any family of sets, whereas the inequality given in Theorem 5.1.2.1 is only true when $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$. Also due to the fact that the constant $c$ appearing in Theorem 5.1.2.1 is less than 1, (5.1.5.3) is asymptotically better than ours (see also Remark 30 above). On the other hand, one sees that (5.1.5.3) is only valid for $s$ large enough (and its domain of validity can thus be empty when $s_o > r$) whereas our inequality is true on the whole interval $(0, r)$. It does not seem also possible to iterate (5.1.5.3) as we did in Corollary 5.1.2.4. Finally, observe that the method of proof used in [34] is based on heat kernel bounds and is very different from ours.

Let us translate Theorem 5.1.5.1 in a form closer to our Proposition 5.1.2.2. Fix $k$ sets $A_1, \ldots, A_k$ such that $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$. Let $2r = \min d(A_i, A_j)$, where the infimum runs on $i, j = 1, \ldots, k$ with $i \neq j$. We have to choose a $(k+1)$-th set. In view of Theorem 5.1.5.1, the most optimal choice is to choose $A_0 = E \setminus (\cup A_i)_r$. Indeed, it is the biggest set (in the sense of inclusion) such that $\min d(A_i, A_j) = r$ where this time the infimum runs on $i, j = 0, \ldots, k$ and $i \neq j$. We let $a_{(0)} = \mu(A_0)$ and we remark that if $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$ then $a_{(0)} \leq a_{(1)}$. The bound (5.1.5.1) can be read: for all $r > 0$,

$$\lambda^{(k)} \leq \frac{1}{r^2}\left(\log \frac{4}{a_{(1)}a_{(0)}}\right)^2. \tag{5.1.5.4}$$

Therefore, to compare it to our bound, we need to solve

$$\phi^{-1}\left(\frac{1}{c}\log\frac{a_{(1)}}{a_{(0)}}\right)^2 \leq \left(\log\frac{4}{a_{(1)}a_{(0)}}\right)^2. \tag{5.1.5.5}$$

Because the right-hand side is always $\geq 1$, taking the square root and composing with the non-decreasing function $\phi$ yields

$$\frac{1}{c}\log\frac{a_{(1)}}{a_{(0)}} \leq \log\frac{4}{a_{(1)}a_{(0)}}. \tag{5.1.5.6}$$

That is

$$a_{(1)}^{1+c} \leq 4^c a_{(0)}^{1-c}. \tag{5.1.5.7}$$

In other words, on some range our bound is better and in some other range their bound is better. However, if the constant $c = 1$ could be attained in Theorem 5.1.2.1, this would show that our bound is always better. Note that comparing the bounds obtained in Proposition 5.1.4.2 and the one of [34] is not so clear as, without the assumption that $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$ it is not necessary that $a_{(0)} \leq a_{(1)}$ and in that case we would have to compare different sets.

## 5.2. Eigenvalue estimates for non-negatively curved spaces

We recall the values of the $\lambda^{(k)}$'s that appear in Theorem 5.1.2.1 in the case of two important models of positively curved spaces in geometry. Namely:

(i) The $n$-dimensional sphere of radius $\sqrt{\frac{n-1}{\rho}}$, $\mathbb{S}^{n,\rho}$ endowed with the natural geodesic distance $d_{n,\rho}$ arising from its canonical Riemannian metric and its normalized volume measure $\mu_{n,\rho}$ which has constant Ricci curvature equals to $\rho$ and dimension $n$.

(ii) The $n$-dimensional Euclidean space $\mathbb{R}^n$ endowed with the $n$-dimensional Gaussian measure of covariance $\rho^{-1}\mathrm{Id}$,

$$\gamma_{n,\rho}(\mathrm{d}x) = \frac{\rho^{n/2}\,\mathrm{e}^{-\rho|x|^2/2}}{(2\pi)^{n/2}}\mathrm{d}x. \tag{5.2.0.1}$$

This space has dimension $\infty$ and curvature bounded below by $\rho$ in the sense of [12].

These models arise as weighted Riemannian manifolds without boundary having a purely discrete spectrum. In that case, it was proved in [100, Proposition 3.2] that the $\lambda_k$'s of (5.1.1.3) are exactly the eigenvalues (counted with multiplicity) of a self-adjoint operator that we give explicitly in the following. Using a comparison between eigenvalues of [100], we obtain an estimates for eigenvalues in the case of log-concave probability measure over the Euclidean $\mathbb{R}^n$.

**Example 5.2.0.1** (Spheres)**.** On $\mathbb{S}^{n,\rho}$, the eigenvalues of minus the Laplace-Beltrami operator (see for instance [8, Chapter 3]) are of the form $\rho^{-2}(n-1)^2 l(l+n-1)$ for $l \in \mathbb{N}$ and the dimension of the corresponding eigenspace $H_{l,n}$ is

$$\dim H_{l,n} = \frac{2l+n-1}{l}\binom{l+n-2}{l-1}, \text{ if } l > 0 \text{ and } \dim H_{l,n} = 1, \text{ if } l = 0. \tag{5.2.0.2}$$

Consequently,

$$D_{l,n} := \dim \bigoplus_{l'=0}^{l} H_{l',n} = \binom{n+l}{l} + \binom{n+l-1}{l-1}, \tag{5.2.0.3}$$

and $\lambda^{(k)} = \rho^{-2}(n-1)^2 l(l+n-1)$ if and only if $D_{l-1,n} < k \leq D_{l,n}$ where $\lambda^{(k)}$ is the $k$-th eigenvalues of $-\Delta_{\mathbb{S}^{n,\rho}}$ and coincides with the variational definition given in (5.1.1.3).

**Example 5.2.0.2** (Gaussian spaces)**.** On the Euclidean space $\mathbb{R}^n$, equipped with the Gaussian measure $\gamma_{n,\rho}$, the corresponding weighted Laplacian is $\Delta_{\gamma_{n,\rho}} = \Delta_{\mathbb{R}^n} - \rho x \cdot \nabla$. The eigenvalues of $-\Delta_{\gamma_{n,\rho}}$ are exactly of the form $\rho^2 q$ and the dimension of the associated eigenspace $H_{q,n}$ is

$$\dim H_{q,n} = \binom{n+q-1}{q}. \tag{5.2.0.4}$$

Consequently,

$$D_{q,n} := \dim \bigoplus_{q'=0}^{q} H_{q',n} = \binom{n+q}{q}, \tag{5.2.0.5}$$

and $\lambda^{(k)} = \rho^{-2}q$ if and only if $D_{q-1,n} < k \leq D_{q,n}$ where $\lambda^{(k)}$ is the $k$-th eigenvalues of $-\Delta_{\gamma_{n,\rho}}$ and coincides with the variational definition given in (5.1.1.3).

**Example 5.2.0.3** (Log-concave Euclidean spaces)**.** We study the case where $E = \mathbb{R}^n$, $d$ is the Euclidean distance and $\mu$ is a strictly log-concave probability measure. By this we mean that $\mu(\mathrm{d}x) = \mathrm{e}^{-V(x)}\,\mathrm{d}x$, where $V\colon \mathbb{R}^n \to \mathbb{R}$ such that $V$ is $\mathcal{C}^2$ and satisfying $\nabla^2 V \geq K$ for some $K > 0$. It is a consequence of [12, Proposition 4] that such a condition on $V$ implies that the semigroup generated by the solution of the stochastic differential equation

$$\mathrm{d}X_t = \sqrt{2}\mathrm{d}B_t - \nabla V(X_t)\mathrm{d}t, \tag{5.2.0.6}$$

where $B$ is a Brownian motion on $\mathbb{R}^n$, satisfies the curvature-dimension $CD(\infty, K)$ of Bakry-Emery and, therefore, holds the log-Sobolev inequality, for all $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$,

$$\mathbf{Ent}_\mu f^2 \leq \frac{2}{K}\int |\nabla f(x)|^2 \mu(\mathrm{d}x). \tag{5.2.0.7}$$

Such an inequality implies the super-Poincaré of [147, Theorem 2.1] that in turns implies that the self-adjoint operator $L = -\Delta + \nabla V \cdot \nabla$ has a purely discrete spectrum. In that case, the $\lambda^{(k)}$ of (5.1.1.3) corresponds to these eigenvalues and [100] showed that

$$\lambda^{(k)} \geq \lambda^{(k)}_{\gamma_{n,\rho}}, \tag{5.2.0.8}$$

where $\lambda^{(k)}_{\gamma_{n,\rho}}$ is the eigenvalues of $-\Delta_{\gamma_{n,\rho}}$ of the previous example.

## 5.3. Extension to Markov chains

As in the classical case (see [85, Theorem 3.3]), our continuous result admits a generalization on finite graphs or more broadly in the setting of Markov chains on a finite state space. We consider a finite set $E$ and $X = (X_n)$ be a irreducible time-homogeneous Markov chain with state space $E$. We write $p(x,y) = \mathbb{P}(X_1 = y | X_0 = x)$ and we regard $p$ as a matrix. We assume that $X$ admits a reversible probability measure $\mu$ on $E$ such that $p(x,y)\mu(x) = p(y,x)\mu(y)$ and $\mu(y) = \sum_x p(x,y)\mu(x)$. This induces a graph structure on $E$ by the following procedure. Set the elements of $E$ as the vertex of the graph and for $x, y \in E$ connect them with an edge if $p(x,y) > 0$. As the chain is irreducible, this graph is connected. We equip $E$ with the induced graph distance $d$. We write $L = p - I$, where $I$ stands for the identity. The operator $-L$ is a symmetric positive operator on $\mathcal{L}^2(\mu)$. We let $\lambda^{(k)}$ be the eigenvalues of this operator. Then, our Theorem 5.1.2.1 extends as follows:

**Theorem 5.3.0.1.** *For any $k \geq 1$ and for all sets $A_1, \ldots, A_k \subset E$ such that $\min_{i\neq j} d(A_i, A_j) \geq 1$ and $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$ the set $B = A_1 \cup A_2 \cup \cdots \cup A_k$ satisfies*

$$\mu(B_n) \geq 1 - (1 - \mu(B))\big(1 + \lambda^{(k)}\big)^{-n}, \tag{5.3.0.1}$$

*for all $1 \leq n \leq \frac{1}{2}\min_{i\neq j} d(A_i, A_j)$ where $\lambda^{(k)}$ is the $k$-th eigenvalue of the operator $-L$ acting on $\mathcal{L}^2(\mu)$.*

*Proof.* We let $\Pi(x,y) = p(x,y)\mu(x)$ and

$$\mathscr{E}(f,g) = \frac{1}{2}\sum (f(y) - f(x))(g(y) - g(x))\Pi(x,y) = \langle f, -Lg\rangle_\mu. \tag{5.3.0.2}$$

For any set $A$, we define the discrete boundary of $A$ as $\partial A = A_1 \setminus A \cup (A^C)_1 \setminus A^C$. Let $(X_n)$ be the Markov chain with transition kernel $p$ and initial distribution $\mu$. By reversibility of $\mu$, $(X_0, X_1)$ is an exchangeable pair of law $\Pi$ whose the marginals are given by $\mu$. Then, for a set $U$, we have

(5.3.0.3)
$$\mathscr{E}(1_U) = \mathbb{E}1_U(X_0)(1_U(X_0) - 1_U(X_1)) = \mathbb{P}(X_0 \in U, X_1 \notin U) \le \mathbb{P}(X_1 \in \partial U) = \mu(\partial U).$$

Observe that if $d(U,V) \ge 1$, $U$ and $V$ are disjoint and $U \times V \notin \operatorname{supp} \Pi$ so that $\mathscr{E}(1_U, 1_V) = 0$. By Courant-Fischer's min-max theorem

$$\lambda^{(k)} = \min_{\dim V = k+1} \max_{f \in V} \frac{\mathscr{E}(f,f)}{\mu(f^2)}. \tag{5.3.0.4}$$

Choose sets $A_1, \dots, A_k$ with $d(A_i, A_j) \ge 2n$ ($i \ne j$) and $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$. Set $f_i = 1_{A_i}$. The $f_i$'s have disjoint support and so they are orthogonal in $L^2(\mu)$. By the previous variational representation of $\lambda^{(k)}$, we have

$$\lambda^{(k)} \le \sup_{a_i} \frac{\mathscr{E}\left(\sum_{i=0}^k a_i f_i\right)}{\int \left(\sum_{i=0}^k a_i f_i\right)^2 \mathrm{d}\mu} = \sup_{a_i} \frac{\sum a_i a_{i'} \mathscr{E}(f_i, f_{i'})}{\sum a_i a_{i'} \int f_i f_{i'} \mathrm{d}\mu} = \sup_{a_i} \frac{\sum_{i=0}^k a_i^2 \mathscr{E}(f_i)}{\sum_{i=0}^k a_i \int f_i^2 \mathrm{d}\mu}. \tag{5.3.0.5}$$

In other words,

$$\lambda^{(k)} \le \max_{i=0,\dots,k} \frac{\mu((A_i)_1) + \mu((A_i^C)_1) - 1}{\mu(A_i)} \le \frac{\mu((A_i)_1) - \mu(A_i)}{\mu(A_i)}, \tag{5.3.0.6}$$

where the last inequality comes from the fact that, by Lemma 5.1.3.1, $\mu(E \setminus (E \setminus A)_1) \ge \mu(A)$. Consider the set $B = \cup_{i=1}^k A_i$ and choose $A_0 = E \setminus B_1$. In that case, by Lemma 5.1.3.2 with $\epsilon = 1$, we have

$$\max_{i=0,\dots,k} \frac{\mu((A_i)_1)}{\mu(A_i)} \le \frac{1 - \mu(B)}{1 - \mu(B_1)}. \tag{5.3.0.7}$$

Thus, we proved that

$$(1 + \lambda^{(k)})(1 - \mu(B_1)) \le (1 - \mu(B)). \tag{5.3.0.8}$$

We derive the announced result by an immediate recursion. □

## 5.4. Functional forms of the multiple sets concentration property

We investigate the functional form of the multi-sets concentration of measure phenomenon results obtained in Sections 5.1 and 5.3.

**Proposition 5.4.0.1.** *Let $(E,d)$ be a metric space equipped with a Borel probability measure $\mu$. Let $\alpha_k : \mathbb{R}_+ \to \mathbb{R}_+$. The following properties are equivalent:*

1. *For all Borel sets $A_1, \dots, A_k \subset E$ such that $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$, the set $A = A_1 \cup \dots \cup A_k$ satisfies*

$$\mu(A_r) \ge 1 - (1 - \mu(A))\alpha_k(r), \qquad \forall 0 < r \le \frac{1}{2} \min_{i \ne j} d(A_i, A_j). \tag{5.4.0.1}$$

2. *For all 1-Lipschitz functions $f_1, \ldots, f_k : E \to \mathbb{R}$ such that the sublevel sets $A_i = \{f_i \leq 0\}$ are such that $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$, the function $f^* = \min(f_1, \ldots, f_k)$ satisfies*

$$\mu(f^* < r) \geq 1 - \mu(f^* \leq 0)\alpha_k(r), \qquad \forall 0 < r \leq \frac{1}{2}\min_{i \neq j} d(A_i, A_j). \tag{5.4.0.2}$$

Together with Theorem 5.1.2.1 or Theorem 5.3.0.1, one thus sees that the presence of multiple wells can improve the concentration properties of a Lipschitz function.

*Proof.* It is clear that (2) implies (1) when applied to $f_i(x) = d(x, A_i)$, in which case $A_i = \{f_i \leq 0\}$ and $f^*(x) = d(x, A)$. The converse is also very classical. First, observe that $\{f^* < r\} = \cup_{i=1}^k \{f_i < r\}$. Then, since $f_i$ is 1-Lipschitz, it holds $A_{i,r} \subset \{f_i < r\}$ with $A_i = \{f_i \leq 0\}$ and so letting $A = A_1 \cup \cdots \cup A_k$, it holds $A_r \subset \{f^* < r\}$. Therefore, applying (1) to this set $A$ gives (2). □

When (5.4.0.1) holds, we will say that the probability metric space $(E, d, \mu)$ satisfies the multi-set concentration of measure property of order $k$ with the concentration profile $\alpha_k$.

In the usual setting ($k = 1$), the concentration of measure phenomenon implies deviation inequalities for Lipschitz functions around their median. The next result generalizes this well known fact to $k > 1$.

**Proposition 5.4.0.2.** *Let $(E, d, \mu)$ be a probability metric space satisfying the multi-set concentration of measure property of order $k$ with the concentration profile $\alpha_k$ and $f : E \to \mathbb{R}$ be a 1-Lipschitz function. If $I_1, \ldots, I_k \subset \mathbb{R}$ are $k$ disjoint Borel sets such that $(\mu(f \in I_1), \ldots, \mu(f \in I_k)) \in \Delta_k$, then it holds*

$$\mu\left(f \in \cup_{i=1}^k I_{i,r}\right) \geq 1 - (1 - \mu(f \in \cup_{i=1}^k I_i))\alpha_k(r), \qquad \forall 0 < r \leq \frac{1}{2}\min_{i \neq j} d(I_i, I_j) \tag{5.4.0.3}$$

*Proof.* Let $\nu$ be the image of $\mu$ under the map $f$. Since $f$ is 1-Lipschitz, the metric space $(\mathbb{R}, |\cdot|, \nu)$ satisfies the multi-set concentration of measure property of order $k$ with the same concentration profile $\alpha_k$ as $\mu$. Details are left to the reader. □

Let us conclude this section by detailling some application of potential interest in approximation theory.

Suppose that $f : E \to \mathbb{R}$ is some 1-Lipschitz function and $A_1, \ldots A_k$ are (pairwise disjoint) subsets of $E$ such that $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$. Let us assume that the restrictions $f_{|A_i}$, $i \in \{1, \ldots, k\}$ are known and that one wishes to estimate or reconstruct $f$ outside $A = \cup_{i=1}^k A_i$. To that aim, one can consider an explicit 1-Lipschitz extension of $f_{|A}$, that is to say a 1-Lipschitz function $g : E \to \mathbb{R}$ (constructed based on our knowledge of $f$ on $A$ exclusively) such that $f = g$ on $A$. There are several canonical ways to perform the extension of a Lipschitz function defined on a sub domain (known as Kirszbraun-McShane-Whitney extension [72, 98, 148]). One can consider for instance the functions

$$g_+(x) = \inf_{y \in A}\{f(y) + d(x, y)\} \qquad \text{or} \qquad g_-(x) = \sup_{y \in A}\{f(y) - d(x, y)\}, \qquad x \in E. \tag{5.4.0.4}$$

It is a very classical fact that functions $g_-$ and $g_+$ are 1-Lipschitz extensions of $f_{|A}$ and moreover that any extension $g$ of $f_{|A}$ satisfies $g_- \leq g \leq g_+$ (see *e.g* [64]).

The following simple result shows that, for any 1-Lipschitz extension $g$ of $f_{|A}$, the probability of error $\mu(|f-g|>r)$ is controlled by the multi-set concentration profile $\alpha_k$. In particular, in the framework of our Theorem 5.1.2.1, this probability of error is controled by $\lambda^{(k)}$.

**Proposition 5.4.0.3.** *Let $(E,d,\mu)$ be a probability metric space satisfying the multi-set concentration of measure property of order $k$ with the concentration profile $\alpha_k$ and $f\colon E\to\mathbb{R}$ be a 1-Lipschitz function. Let $A_1,\dots A_k$ be subsets of $E$ such that $(\mu(A_1),\dots,\mu(A_k))\in\Delta_k$; then for any 1-Lipschitz extension $g$ of $f_{|A}$, it holds*

$$\mu(|f-g|\geq r)\leq(1-\mu(A))\alpha_k(r/2),\qquad\forall 0<r\leq\min_{i\neq j}d(A_i,A_j).\tag{5.4.0.5}$$

*Proof.* The function $h:E\to\mathbb{R}$ defined by $h(x)=|f-g|(x)$, $x\in E$, is 2-Lipschitz and vanishes on $A$. Therefore, for any $x\in E$ and $y\in A$, it holds $h(x)\leq h(y)+2d(x,y)=2d(x,y)$. Optimizing over $y\in A$ gives that $h(x)\leq 2d(x,A)$. Therefore $\{h\geq r\}\subset\{x:d(x,A)\geq r/2\}=\left(A_{r/2}\right)^c$ and so, if $0<r\leq\min_{i\neq j}d(A_i,A_j)$, it holds

$$\mu(|f-g|\geq r)\leq(1-\mu(A))\alpha_k(r/2).\tag{5.4.0.6}$$

□

**Remark 31.** Let us remark that Propositions 5.4.0.1, 5.4.0.2 and 5.4.0.3 can be immediately extended under the following more general (but notationally heavier) multi-set concentration of measure assumption: there exists functions $\alpha_k:\mathbb{R}_+\to\mathbb{R}_+$ and $\beta_k:\mathbb{R}_+^k\to[0,\infty]$ such that for all Borel sets $A_1,\dots,A_k\subset E$, the set $A=A_1\cup\dots\cup A_k$ satisfies

$$\mu(A_r)\geq 1-\beta_k(\mu(A_1),\dots,\mu(A_k))\alpha_k(r),\qquad\forall 0<r\leq\frac{1}{2}\min_{i\neq j}d(A_i,A_j).\tag{5.4.0.7}$$

This framework contains the preceding one, by choosing $\beta_k(a)=1-\sum_{i=1}^k a_i$ if $a=(a_1,\dots,a_k)\in\Delta_k$ and $+\infty$ otherwise. It also contains the concentration bounds obtained in Proposition 5.1.4.1, corresponding respectively to

$$\beta_k(a)=\frac{1-\sum_{i=1}^k a_i}{\prod_{i=1}^k a_i},\quad\text{and}\quad\beta_k(a)=\frac{1-\sum_{i=1}^k a_i}{\left(\sum_{i=1}^k a_i\right)^{\sum_{i=1}^k a_i/\min(a_1,\dots,a_k)}},\qquad a=(a_1,\dots,a_k).\tag{5.4.0.8}$$

## 5.5. Open questions

We list open questions related to the multi-set concentration of measure phenomenon.

5.5.1. **Gaussian multi-set concentration.** Using the terminology introduced in Section 5.4, Theorem 5.1.2.1 and the material exposed in Section 5.2 tell us that, if $\mu$ has a density of the form $e^{-V}$ with respect to Lebesgue measure on $\mathbb{R}^n$ with a smooth function $V$ such that $\operatorname{Hess}V\geq\rho>0$, then the probability metric space $(\mathbb{R}^n,|\cdot|,\mu)$ satisfies the multi-set concentration of measure property of order $k$ with the concentration profile

$$\alpha_k(r)=\exp\left(-c\min(r^2\lambda^{(k)}_{\gamma_n,\rho};r\sqrt{\lambda^{(k)}_{\gamma_n,\rho}})\right),\qquad r\geq 0,\tag{5.5.1.1}$$

where $\lambda^{(k)}_{\gamma_n,\rho}$ denotes the $k$th eigenvalue of the $n$-dimensional centered Gaussian measure with covariance matrix $\rho^{-1}\mathrm{Id}$. Since the measure $\mu$ satisfies the log-Sobolev inequality, it is well known that it satisfies a (classical) Gaussian concentration of measure inequality. Therefore, it is natural to conjecture that $\mu$ satisfies a multi-set concentration of measure property of order $k \geq 1$ with a profile of the form

$$\beta_k(r) = \exp\left(-C_{k,\rho,n} r^2\right), \qquad r \geq 0, \tag{5.5.1.2}$$

for some constant $C_{k,\rho,n}$ depending solely on its arguments. In addition, it would be interesting to see how usual functional inequalities (Log-Sobolev, transport-entropy, ...) can be modified to catch such a concentration of measure phenomenon. To that extent, let us mention some ongoing work with N. GOZLAN & P-M. SAMSOM investigating some of those questions. It is well-known that, if $(Z, d)$ is a geodesic metric space (say a Riemannian manifold to fix the idea) then the space $\mathscr{P}_2(Z)$ of probability measures on $Z$ with finite second-moment is also a geodesic metric space when equipped with the quadratic Wasserstein distance $\mathcal{W}_{d,2}$: for all $\nu_0$ and $\nu_1 \in \mathscr{P}_2(Z)$, there exists a family $(\nu_t)_{t\in[0,1]}$ such that

$$\mathcal{W}_{d,2}(\nu_0, \nu_1) = \mathcal{W}_{d,2}(\nu_0, \nu_t) + \mathcal{W}_{d,2}(\nu_t, \nu_1), \quad \text{for all } t \in (0,1). \tag{5.5.1.3}$$

The curve $(\nu_t)_{t\in[0,1]}$ is called the *Wasserstein geodesic* joining $\nu_0$ to $\nu_1$. Many fundamental papers about optimal transport (see in particular Y. BRENIER (1991) [28], R. J. MCCANN (1997) [97], F. OTTO (2001) [119]), upheld the idea that the geometry of $(Z, d)$ should be linked to the geometry of $(\mathscr{P}_2(Z), \mathcal{W}_{2,d})$. This point of view was formalized in the two seminal contributions of K.-T. STURM (2006) [141] & [142] and J. LOTT & C. VILLANI (2009) [90]. They noticed that the property of displacement convexity of the relative entropy $\mathcal{H}(\cdot|\gamma)$ along the Wasserstein geodesics was characterizing the geometry of the metric measured space $(Z, d, \gamma)$. In particular, they proved that, if for all $\nu_0$ and $\nu_1 \in \mathscr{P}_2(Z)$, there exists a Wasserstein geodesic $(\nu_t)$ joining $\nu_0$ to $\nu_1$ such that

$$\mathcal{H}(\nu_t|\gamma) \leq (1-t)\mathcal{H}(\nu_0|\gamma) + t\mathcal{H}(\nu_1|\gamma) - \frac{1}{2}t(1-t)\mathcal{W}^2_{2,d}(\nu_0, \nu_1), \tag{5.5.1.4}$$

then the functional inequalities on $(Z, d, \gamma)$ are (at least) as good as the ones on a $(\mathbb{R}, |\cdot|, \mathbf{N}(0,1))$. In particular we can recover for those spaces a logarithmic Sobolev inequality and a spectral gap inequality. Also observe that (5.5.1.4) implies the Talagrand inequality

$$\mathcal{W}^2_{2,d}(\nu_0, \nu_1) \leq (\mathcal{H}(\nu_0|\gamma) + \mathcal{H}(\nu_1|\gamma)), \tag{5.5.1.5}$$

and that Gaussian concentration of measure can be deduced from it (see M. TALAGRAND (1996) [144]). It is therefore very natural to ask if our concentration result Theorem 5.1.2.1 could be related to an higher dimensional version of the Talagrand inequality (5.5.1.5) or from an higher dimensional version of (5.5.1.4). We work over $\mathbb{R}^d$. Recall (see W. GANGBO & A. ŚWIECH (1998) [52] or M. AGUEH & G. CARLIER (2011) [1]) that given $\nu_1, \ldots, \nu_k \in \mathscr{P}_2(\mathbb{R}^d)$ and $t_1, \ldots, t_k \in [0,1]$ such that $\sum_{i=1}^k t_i = 1$ we define the *multidimensional transport cost*

$$T^t(\nu_1, \ldots, \nu_k) = \inf\left\{\int \sum_{1\leq i<j\leq k} t_i t_j |x_i - x_j|^2 \pi(\mathrm{d}x)\right\}, \tag{5.5.1.6}$$

where the infimum is over all the couplings $\pi$ of $\nu_1, \dots, \nu_k$. It can be shown [52] that this infimum is achieved for a unique coupling $\pi^*$. The *Wasserstein barycenters* of $\nu = (\nu_1, \dots, \nu_k)$ with coefficients $t = (t_1, \dots, t_k)$ is defined by

$$bar(\nu, t) = \left(\sum_{i=1}^{k} t_i x_i\right) \sharp \pi^*. \tag{5.5.1.7}$$

The curvature inequality (5.5.1.4) has the following multi-marginal equivalent

$$\mathcal{H}(bar(\nu,t)|\gamma) \le \sum_{i=1}^{k} t_i \mathcal{H}(\nu_i|\gamma) - \frac{1}{2} T_2^t(\nu_1, \dots, \nu_k), \tag{5.5.1.8}$$

Following [1, Proposition 7] and [5, Theorem 9.4.11], the normal measure $\mathbf{N}(0,1)$ satisfies such an inequality but other measures $\gamma$ might also satisfy (5.5.1.8). The curvature inequality implies the multi-marginal version of (5.5.1.5)

$$\frac{1}{2} T_2^t(\nu_1, \dots, \nu_k) \le \sum_{i=1}^{k} t_i \mathcal{H}(\nu_i|\gamma). \tag{5.5.1.9}$$

We are currently investigating the links between (5.5.1.8) or (5.5.1.9) and the multi-set concentration property.

5.5.2. **Equivalence between multi-set concentration and lower bounds on eigenvalues in non-negative curvature.** Let us quickly recall the main finding of E. Milman [103, 101], that is, under non-negative curvature assumptions, a concentration of measure estimate implies a bound on the spectral gap. Let $\mu$ be a probability measure with a density of the form $e^{-V}$ on a smooth connected Riemannian manifold $M$ with $V$ a smooth function such that

$$\mathrm{Ric} + \mathrm{Hess}\, V \ge 0. \tag{5.5.2.1}$$

Assume that $\mu$ satisfies a concentration inequality of the form: for all $A \subset M$ such that $\mu(A) \ge 1/2$

$$\mu(A_r) \ge 1 - \alpha(r), \qquad r \ge 0, \tag{5.5.2.2}$$

where $\alpha$ is a function such that $\alpha(r_o) < 1/2$ for at least one value $r_o > 0$. Then, letting $\lambda_1$ be the first non zero eigenvalue of the operator $-\Delta + \nabla V \cdot \nabla$, it holds $\lambda_1 \ge \frac{1}{4}\left(\frac{1-2\alpha(r_o)}{r_o}\right)^2$. It would be very interesting to extend Milman's result to a multi-set concentration setting. More precisely, if $\mu$ satisfies the curvature condition (5.5.2.1) and the multi-set concentration of measure property of order $k$ with a profile of the form $\alpha_k(r) = \exp(-\min(ar^2, \sqrt{a}r))$, $r \ge 0$, can we find a universal function $\varphi_k$ such that $\lambda_k \ge \varphi_k(a)$?

This question already received some attention in recent works by Funano and Shioya [50, 51]. In particular, let us mention the following improvement of the Chung-Grigor'yan-Yau inequality obtained in [50]. There exists a universal constant $c > 1$ such that if $\mu$ is a probability measure satisfying the non-negative curvature assumption (5.5.2.1), it holds: for any family of sets $A_0, A_1, \dots, A_l$ with $1 \le l \le k$

$$\lambda^{(k)} \le c^{k-l+1} \frac{1}{\min_{i \ne j} d^2(A_i, A_j)} \max_{i \ne j} \log\left(\frac{4}{\mu(A_i)\mu(A_j)}\right)^2. \tag{5.5.2.3}$$

Note that the difference with (5.1.5.1) is that $\lambda^{(k)}$ is estimated by a reduced number of sets. Using (5.5.2.3) (with $l = 1$) together with Milman's result recalled above, Funano showed that there exists some constant $C_k$ depending only on $k$ such that under the curvature condition (5.5.2.1), it holds $\lambda_k \leq C_k \lambda_0$ (recovering the main result of [51]). The constant $C_k$ is explicit (contrary to the constant of [51]) and grows exponentially when $k \to \infty$. This result has been then improved by Liu [89], where a constant $C_k = O(k^2)$ has been obtained. As observed by Funano [50], a positive answer to the open question stated above would yield that under (5.5.2.1) the ratios $\lambda_{k+1}/\lambda_k$ are bounded from above by a universal constant.

# Bibliography


[1] M. Agueh & G. Carlier. “Barycenters in the Wasserstein space”. In: *SIAM J. Math. Anal.* 43.2 (2011), pp. 904–924. ISSN: 0036-1410. URL: https://doi.org/10.1137/100805741.

[2] S. Aida & D. Stroock. “Moment estimates derived from Poincaré and logarithmic Sobolev inequalities”. In: *Math. Res. Lett.* 1.1 (1994), pp. 75–86. ISSN: 1073-2780.

[3] L. Ambrosio & R. Ghezzi. “Sobolev and bounded variation functions on metric measure spaces”. In: *Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. II*. Ed. by A. Laptev. EMS Ser. Lect. Math. Eur. Math. Soc., Zürich, 2016, pp. 211–273.

[4] L. Ambrosio, N. Gigli & G. Savaré. “Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds”. In: *Ann. Probab.* 43.1 (2015), pp. 339–404. ISSN: 0091-1798. URL: https://doi.org/10.1214/14-AOP907.

[5] L. Ambrosio, N. Gigli & G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Second. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008, pp. x+334. ISBN: 978-3-7643-8721-1.

[6] C. Ané et al. *Sur les inégalités de Sobolev logarithmiques*. Vol. 10. Panoramas et Synthèses [Panoramas and Syntheses]. With a preface by Dominique Bakry and Michel Ledoux. Société Mathématique de France, Paris, 2000, pp. xvi+217. ISBN: 2-85629-105-8.

[7] S. Artstein, K. M. Ball, F. Barthe & A. Naor. “On the rate of convergence in the entropic central limit theorem”. In: *Probab. Theory Related Fields* 129.3 (2004), pp. 381–390. ISSN: 0178-8051. DOI: 10.1007/s00440-003-0329-4. URL: https://doi.org/10.1007/s00440-003-0329-4.

[8] K. Atkinson & W. Han. *Spherical harmonics and approximations on the unit sphere: an introduction*. Vol. 2044. Lecture Notes in Mathematics. Springer, Heidelberg, 2012, pp. x+244. ISBN: 978-3-642-25982-1. DOI: 10.1007/978-3-642-25983-8. URL: https://doi.org/10.1007/978-3-642-25983-8.

[9] E. Azmoodeh, S. Campese & G. Poly. “Fourth Moment Theorems for Markov diffusion generators”. In: *J. Funct. Anal.* 266.4 (2014), pp. 2341–2359. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2013.10.014. URL: http://dx.doi.org/10.1016/j.jfa.2013.10.014.

[10] L. Bachelier. “Théorie de la spéculation”. In: *Ann. Sci. École Norm. Sup. (3)* 17 (1900), pp. 21–86. ISSN: 0012-9593. URL: http://www.numdam.org/item?id=ASENS_1900_3_17__21_0.

[11] L. Bachelier. “Théorie mathématique du jeu”. In: *Ann. Sci. École Norm. Sup. (3)* 18 (1901), pp. 143–209. ISSN: 0012-9593. URL: http://www.numdam.org/item?id=ASENS_1901_3_18__143_0.

[12] D. Bakry & M. Émery. “Diffusions hypercontractives”. In: *Séminaire de probabilités, XIX, 1983/84*. Ed. by J. Azéma & M. Yor. Vol. 1123. Lecture Notes in Math. Springer, Berlin, 1985, pp. 177–206. DOI: 10.1007/BFb0075847. URL: https://doi.org/10.1007/BFb0075847.

[13] D. Bakry, I. Gentil & M. Ledoux. *Analysis and geometry of Markov diffusion operators*. Vol. 348. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Cham, 2014, pp. xx+552. ISBN: 978-3-319-00226-2; 978-3-319-00227-9. URL: https://doi.org/10.1007/978-3-319-00227-9.

[14] A. Barron & O. Johnson. “Fisher information inequalities and the central limit theorem”. In: *Probab. Theory Related Fields* 129.3 (2004), pp. 391–409. ISSN: 0178-8051. DOI: 10.1007/s00440-004-0344-0. URL: https://doi.org/10.1007/s00440-004-0344-0.

[15] A. C. Berry. “The accuracy of the Gaussian approximation to the sum of independent variates”. In: *Trans. Amer. Math. Soc.* 49 (1941), pp. 122–136. ISSN: 0002-9947. DOI: 10.2307/1990053. URL: https://doi.org/10.2307/1990053.

[16] P. Billingsley. *Probability and measure*. Third. Wiley Series in Probability and Mathematical Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1995, pp. xiv+593. ISBN: 0-471-00710-2.

[17] S. G. Bobkov & F. Götze. “Exponential integrability and transportation cost related to logarithmic Sobolev inequalities”. In: *J. Funct. Anal.* 163.1 (1999), pp. 1–28. ISSN: 0022-1236. DOI: 10.1006/jfan.1998.3326. URL: https://doi.org/10.1006/jfan.1998.3326.

[18] S. G. Bobkov & M. Ledoux. “On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures”. In: *J. Funct. Anal.* 156.2 (1998), pp. 347–365. ISSN: 0022-1236. DOI: 10.1006/jfan.1997.3187. URL: https://doi.org/10.1006/jfan.1997.3187.

[19] S. Bobkov & M. Ledoux. “Poincaré’s inequalities and Talagrand’s concentration phenomenon for the exponential distribution”. In: *Probab. Theory Related Fields* 107.3 (1997), pp. 383–400. ISSN: 0178-8051. DOI: 10.1007/s004400050090. URL: https://doi.org/10.1007/s004400050090.

[20] S. G. Bobkov, I. Gentil & M. Ledoux. “Hypercontractivity of Hamilton-Jacobi equations”. In: *J. Math. Pures Appl. (9)* 80.7 (2001), pp. 669–696. ISSN: 0021-7824. DOI: 10.1016/S0021-7824(01)01208-9. URL: https://doi.org/10.1016/S0021-7824(01)01208-9.

[21] S. G. Bobkov & P. Tetali. “Modified logarithmic Sobolev inequalities in discrete settings”. In: *J. Theoret. Probab.* 19.2 (2006), pp. 289–336. ISSN: 0894-9840. DOI: 10.1007/s10959-006-0016-3. URL: https://doi.org/10.1007/s10959-006-0016-3.

[22] S. Bochner. “Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse”. In: *Math. Ann.* 108.1 (1933), pp. 378–410. ISSN: 0025-5831. DOI: 10.1007/BF01452844. URL: https://doi.org/10.1007/BF01452844.

[23] N. Bouleau & F. Hirsch. *Dirichlet forms and analysis on Wiener space*. Vol. 14. De Gruyter Studies in Mathematics. Walter de Gruyter & Co., Berlin, 1991, pp. x+325. ISBN: 3-11-012919-1. DOI: 10.1515/9783110858389. URL: http://dx.doi.org/10.1515/9783110858389.

[24] N. Bourbaki. *Intégration. Chapitre 9 Intégration sur les espaces topologiques séparés*. Eléments de mathématiques. Springer-Verlag Berlin Heidelberg, 2006, pp. vi+127. DOI: 10.1007/978-3-540-34391-2. URL: https://doi.org/10.1007/978-3-540-34391-2.

[25] N. Bourbaki. *Éléments de mathématique. Algèbre. Chapitre 1 à 3.* Springer-Verlag, Berlin, 2007, pp. viii+216. ISBN: 978-3-540-33849-9; 978-3-540-33850-5. DOI: 10.1007/978-3-540-33850-5.

[26] S. Bourguin & G. Peccati. “Portmanteau inequalities on the Poisson space: mixed regimes and multidimensional clustering”. In: *Electron. J. Probab.* 19 (2014), no. 66, 42. ISSN: 1083-6489. DOI: 10.1214/EJP.v19-2879. URL: https://doi.org/10.1214/EJP.v19-2879.

[27] S. Bourguin & G. Peccati. “The Malliavin-Stein method on the Poisson space”. In: *Stochastic analysis for Poisson point processes*. Ed. by G. Peccati & M. Reitzner. Vol. 7. Bocconi Springer Ser. Bocconi Univ. Press, [place of publication not identified], 2016, pp. 185–228.

[28] Y. Brenier. “Polar factorization and monotone rearrangement of vector-valued functions”. In: *Comm. Pure Appl. Math.* 44.4 (1991), pp. 375–417. ISSN: 0010-3640. DOI: 10.1002/cpa.3160440402. URL: https://doi.org/10.1002/cpa.3160440402.

[29] P. Buser. “A note on the isoperimetric constant”. In: *Ann. Sci. École Norm. Sup. (4)* 15.2 (1982), pp. 213–230. ISSN: 0012-9593. URL: http://www.numdam.org/item?id=ASENS_1982_4_15_2_213_0.

[30] S. Campese, I. Nourdin, G. Peccati & G. Poly. “Multivariate Gaussian approximations on Markov chaoses”. In: *Electron. Commun. Probab.* 21 (2016), Paper No. 48, 9. ISSN: 1083-589X. DOI: 10.1214/16-ECP4615. URL: https://doi.org/10.1214/16-ECP4615.

[31] D. Chafaï. “Entropies, convexity, and functional inequalities: on Φ-entropies and Φ-Sobolev inequalities”. In: *J. Math. Kyoto Univ.* 44.2 (2004), pp. 325–363. ISSN: 0023-608X. DOI: 10.1215/kjm/1250283556. URL: https://doi.org/10.1215/kjm/1250283556.

[32] I. Chavel. *Eigenvalues in Riemannian geometry*. Vol. 115. Pure and Applied Mathematics. Including a chapter by Burton Randol, With an appendix by Jozef Dodziuk. Academic Press, Inc., Orlando, FL, 1984, pp. xiv+362. ISBN: 0-12-170640-0.

[33] J. Cheeger. “Differentiability of Lipschitz functions on metric measure spaces”. In: *Geom. Funct. Anal.* 9.3 (1999), pp. 428–517. ISSN: 1016-443X. DOI: 10.1007/s000390050094. URL: https://doi.org/10.1007/s000390050094.

[34] F. R. K. Chung, A. Grigor'yan & S.-T. Yau. "Upper bounds for eigenvalues of the discrete and continuous Laplace operators". In: *Adv. Math.* 117.2 (1996), pp. 165–178. ISSN: 0001-8708. DOI: 10.1006/aima.1996.0006. URL: http://dx.doi.org/10.1006/aima.1996.0006.

[35] C. Dellacherie & P.-A. Meyer. *Probabilités et potentiel. Chapitres I à IV*. Édition entièrement refondue, Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. XV, Actualités Scientifiques et Industrielles, No. 1372. Hermann, Paris, 1975, pp. x+291.

[36] C. Dellacherie & P.-A. Meyer. *Probabilités et potentiel. Chapitres XVII à XXIV, Processus de Markov (fin), Compléments de calcul stochastique*. Hermann, Paris, 1992, pp. xi+429.

[37] P. Dirac. "The quantum theory of the emission and absorption of radiation". In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 114.767 (1927), pp. 243–265. ISSN: 0950-1207. DOI: 10.1098/rspa.1927.0039. eprint: http://rspa.royalsocietypublishing.org/content/114/767/243.full.pdf. URL: http://rspa.royalsocietypublishing.org/content/114/767/243.

[38] C. Döbler & K. Krokowski. "On the fourth moment condition for Rademacher chaos". In: *ArXiv e-prints* (June 2017). arXiv: 1706.00751 [math.PR].

[39] C. Döbler & G. Peccati. "Fourth moment theorems on the Poisson space: analytic statements via product formulae". In: *ArXiv e-prints* (Aug. 2018). arXiv: 1808.01836 [math.PR].

[40] C. Döbler & G. Peccati. "The fourth moment theorem on the Poisson space". In: *Ann. Probab.* 46.4 (2018), pp. 1878–1916. ISSN: 0091-1798. DOI: 10.1214/17-AOP1215. URL: https://doi.org/10.1214/17-AOP1215.

[41] C. Döbler, A. Vidotto & G. Zheng. "Fourth moment theorems on the Poisson space in any dimension". In: *Electron. J. Probab.* 23 (2018), Paper No. 36, 27. ISSN: 1083-6489. DOI: 10.1214/18-EJP168. URL: https://doi.org/10.1214/18-EJP168.

[42] R. M. Dudley. "The sizes of compact subsets of Hilbert space and continuity of Gaussian processes". In: *J. Functional Analysis* 1 (1967), pp. 290–330.

[43] N. Dunford & J. T. Schwartz. *Linear operators. Part II*. Wiley Classics Library. Spectral theory. Selfadjoint operators in Hilbert space, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1963 original, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1988, i–x, 859–1923 and 1–7. ISBN: 0-471-60847-5.

[44] M. Erbar & M. Huesmann. "Curvature bounds for configuration spaces". In: *Calc. Var. Partial Differential Equations* 54.1 (2015), pp. 397–430. ISSN: 0944-2669. DOI: 10.1007/s00526-014-0790-1. URL: https://doi.org/10.1007/s00526-014-0790-1.

[45] C.-G. Esseen. "On the Liapounoff limit of error in the theory of probability". In: *Ark. Mat. Astr. Fys.* 28A.9 (1942), p. 19. ISSN: 0004-2080.

[46] F. Faà di Bruno. "Sullo sviluppo delle funzioni". In: *Annali di Scienze Matematiche e Fisiche* 6 (1855), pp. 479–480.

[47] T. S. Ferguson. *Mathematical statistics: A decision theoretic approach*. Probability and Mathematical Statistics, Vol. 1. Academic Press, New York-London, 1967, pp. xi+396.

[48] V. Fock. “Konfiguationsraum und zweite Quantelung”. In: *Z. Physik* 75 (9–10 1932), pp. 622–647. ISSN: 0001-5962. DOI: 10.1007/BF01344458. URL: https://doi.org/10.1007/BF01344458.

[49] J. Friedman & J.-P. Tillich. “Laplacian eigenvalues and distances between subsets of a manifold”. In: *J. Differential Geom.* 56.2 (2000), pp. 285–299. ISSN: 0022-040X. URL: http://projecteuclid.org/euclid.jdg/1090347645.

[50] K. Funano. “Estimates of eigenvalues of the Laplacian by a reduced number of subsets”. In: *Israel J. Math.* 217.1 (2017), pp. 413–433. ISSN: 0021-2172. DOI: 10.1007/s11856-017-1453-7. URL: http://dx.doi.org/10.1007/s11856-017-1453-7.

[51] K. Funano & T. Shioya. “Concentration, Ricci curvature, and eigenvalues of Laplacian”. In: *Geom. Funct. Anal.* 23.3 (2013), pp. 888–936. ISSN: 1016-443X. DOI: 10.1007/s00039-013-0215-x. URL: https://doi.org/10.1007/s00039-013-0215-x.

[52] W. Gangbo & A. Święch. “Optimal maps for the multidimensional Monge-Kantorovich problem”. In: *Comm. Pure Appl. Math.* 51.1 (1998), pp. 23–45. ISSN: 0010-3640. DOI: 10.1002/(SICI)1097-0312(199801)51:1<23::AID-CPA2>3.0.CO;2-H. URL: https://doi.org/10.1002/(SICI)1097-0312(199801)51:1%3C23::AID-CPA2%3E3.0.CO;2-H.

[53] N. Gozlan & R. Herry. “Multiple sets exponential concentration and higher order eigenvalues”. In: *ArXiv e-prints* (Apr. 2018). arXiv: 1804.06133 [math.PR].

[54] N. Gozlan. “Characterization of Talagrand’s like transportation-cost inequalities on the real line”. In: *J. Funct. Anal.* 250.2 (2007), pp. 400–425. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2007.05.025. URL: https://doi.org/10.1016/j.jfa.2007.05.025.

[55] N. Gozlan & C. Léonard. “Transport inequalities. A survey”. In: *Markov Process. Related Fields* 16.4 (2010), pp. 635–736. ISSN: 1024-2953.

[56] N. Gozlan, C. Roberto & P.-M. Samson. “From dimension free concentration to the Poincaré inequality”. In: *Calc. Var. Partial Differential Equations* 52.3-4 (2015), pp. 899–925. ISSN: 0944-2669. DOI: 10.1007/s00526-014-0737-6. URL: http://dx.doi.org/10.1007/s00526-014-0737-6.

[57] N. Gozlan, C. Roberto & P.-M. Samson. “Hamilton Jacobi equations on metric spaces and transport entropy inequalities”. In: *Rev. Mat. Iberoam.* 30.1 (2014), pp. 133–163. ISSN: 0213-2230. DOI: 10.4171/RMI/772. URL: https://doi.org/10.4171/RMI/772.

[58] N. Gozlan, C. Roberto, P.-M. Samson & P. Tetali. “Kantorovich duality for general transport costs and applications”. In: *J. Funct. Anal.* 273.11 (2017), pp. 3327–3405. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2017.08.015. URL: https://doi.org/10.1016/j.jfa.2017.08.015.

[59] M. Gromov & V. D. a. Milman. “A topological application of the isoperimetric inequality”. In: *Amer. J. Math.* 105.4 (1983), pp. 843–854. ISSN: 0002-9327. DOI: 10.2307/2374298. URL: http://dx.doi.org/10.2307/2374298.

[60] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*. English. Modern Birkhäuser Classics. Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates. Birkhäuser Boston, Inc., Boston, MA, 2007, pp. xx+585. ISBN: 978-0-8176-4582-3; 0-8176-4582-9.

[61] L. Gross. "Logarithmic Sobolev inequalities". In: *Amer. J. Math.* 97.4 (1975), pp. 1061–1083. ISSN: 0002-9327. URL: https://doi.org/10.2307/2373688.

[62] M. Hairer. *Advanced stochastic calculus*. Lecture notes. 2016.

[63] D. Harnett & D. Nualart. "Central limit theorem for a Stratonovich integral with Malliavin calculus". In: *Ann. Probab.* 41.4 (2013), pp. 2820–2879. ISSN: 0091-1798. DOI: 10.1214/12-AOP769. URL: https://doi.org/10.1214/12-AOP769.

[64] J. Heinonen. *Lectures on Lipschitz analysis*. Vol. 100. Report. University of Jyväskylä Department of Mathematics and Statistics. University of Jyväskylä, Jyväskylä, 2005, pp. ii+77. ISBN: 951-39-2318-5.

[65] W. Hoeffding. "The strong law of large numbers for $U$-statistics". In: *Mimeo. Series* 302 (1961), pp. 1–10. URL: http://www.lib.ncsu.edu/resolver/1840.4/2128.

[66] L. Hörmander. "Hypoelliptic second order differential equations". In: *Acta Math.* 119 (1967), pp. 147–171. ISSN: 0001-5962. DOI: 10.1007/BF02392081. URL: https://doi.org/10.1007/BF02392081.

[67] K. Itô. "Multiple Wiener integral". In: *J. Math. Soc. Japan* 3 (1951), pp. 157–169. ISSN: 0025-5645. URL: https://doi.org/10.2969/jmsj/00310157.

[68] K. Itô. "Spectral type of the shift transformation of differential processes with stationary increments". In: *Trans. Amer. Math. Soc.* 81 (1956), pp. 253–263. ISSN: 0002-9947. URL: https://doi.org/10.2307/1992916.

[69] K. Itô. "Stochastic integral". In: *Proc. Imp. Acad. Tokyo* 20 (1944), pp. 519–524. URL: http://projecteuclid.org/euclid.pja/1195572786.

[70] J. Jacod & A. N. Shiryaev. *Limit theorems for stochastic processes*. Second. Vol. 288. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 2003, pp. xx+661. ISBN: 3-540-43932-3. URL: https://doi.org/10.1007/978-3-662-05265-5.

[71] O. Kallenberg. *Foundations of modern probability*. Second. Probability and its Applications (New York). Springer-Verlag, New York, 2002, pp. xx+638. ISBN: 0-387-95313-2. DOI: 10.1007/978-1-4757-4015-8. URL: https://doi.org/10.1007/978-1-4757-4015-8.

[72] M. Kirszbraun. "Uber die zusammenziehende und Lipschitzsche Transformationen". In: *Fundamenta Mathematicae* 22 (1934), pp. 77–108. DOI: DOI:10.4064/fm-22-1-77-108.

[73] A. N. Kolmogorov. "Curves in a Hilbert space invariant with respect to a one-parameter group of motions". In: *C. R. (Doklady) Acad. Sci. URSS (N.S.)* 26 (1940), pp. 6–9.

[74] A. N. Kolmogorov. *Selected works of A. N. Kolmogorov. Vol. I*. Vol. 25. Mathematics and its Applications (Soviet Series). Mathematics and mechanics, With commentaries by V. I. Arnold, V. A. Skvortsov, P. L. Ulyanov et al, Translated from the Russian original by V. M. Volosov, Edited and with a preface, foreword and brief biography by V. M. Tikhomirov. Kluwer Academic Publishers Group, Dordrecht, 1991, pp. xx+551. ISBN: 90-277-2796-1.

[75] R. Lachièze-Rey & G. Peccati. "Fine Gaussian fluctuations on the Poisson space II: rescaled kernels, marked processes and geometric $U$-statistics". In: *Stochastic Process. Appl.* 123.12 (2013), pp. 4186–4218. ISSN: 0304-4149. DOI: 10.1016/j.spa.2013.06.004. URL: https://doi.org/10.1016/j.spa.2013.06.004.

[76] R. Lachièze-Rey & G. Peccati. "Fine Gaussian fluctuations on the Poisson space, I: contractions, cumulants and geometric random graphs". In: *Electron. J. Probab.* 18 (2013), no. 32, 32. ISSN: 1083-6489. DOI: 10.1214/EJP.v18-2104. URL: https://doi.org/10.1214/EJP.v18-2104.

[77] P.-S. de Laplace. *Théorie analytique des probabilités*. Mme Ve Courcier, 1812, [6]+464. URL: http://catalogue.bnf.fr/ark:/12148/cb35795753v.

[78] P.-S. de Laplace. *Théorie analytique des probabilités. Vol. I*. Introduction: Essai philosophique sur les probabilités. [Introduction: Philosophical essay on probabilities], Livre I: Du calcul des fonctions génératrices. [Book I: On the calculus of generating functions], Reprint of the 1819 fourth edition (Introduction) and the 1820 third edition (Book I). Éditions Jacques Gabay, Paris, 1995, pp. clxx+194. ISBN: 2-87647-161-2.

[79] P.-S. de Laplace. *Théorie analytique des probabilités. Vol. II*. Livre II: Théorie générale des probabilités. [Book II: General probability theory], Suppléments. [Supplements], Reprint of the 1820 third edition (Book II) and of the 1816, 1818, 1820 and 1825 originals (Supplements). Éditions Jacques Gabay, Paris, 1995, i–xviii and 195–691. ISBN: 2-87647-161-2.

[80] G. Last. "Stochastic analysis for Poisson processes". In: *Stochastic analysis for Poisson point processes*. Vol. 7. Bocconi Springer Ser. Bocconi Univ. Press, [place of publication not identified], 2016, pp. 1–36. URL: https://doi.org/10.1007/978-3-319-05233-5_1.

[81] G. Last, G. Peccati & M. Schulte. "Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization". In: *Probab. Theory Related Fields* 165.3-4 (2016), pp. 667–723. ISSN: 0178-8051. DOI: 10.1007/s00440-015-0643-7. URL: https://doi.org/10.1007/s00440-015-0643-7.

[82] G. Last & M. Penrose. *Lectures on the Poisson process*. Vol. 7. Institute of Mathematical Statistics Textbooks. Cambridge University Press, Cambridge, 2018, pp. xx+293. ISBN: 978-1-107-45843-7; 978-1-107-08801-6.

[83] G. Last & M. D. Penrose. "Poisson process Fock space representation, chaos expansion and covariance inequalities". In: *Probab. Theory Related Fields* 150.3-4 (2011), pp. 663–690. ISSN: 0178-8051. DOI: 10.1007/s00440-010-0288-5. URL: http://dx.doi.org/10.1007/s00440-010-0288-5.

[84] M. Ledoux. "Chaos of a Markov operator and the fourth moment condition". In: *Ann. Probab.* 40.6 (2012), pp. 2439–2459. ISSN: 0091-1798. DOI: 10.1214/11-AOP685. URL: https://doi.org/10.1214/11-AOP685.

[85] M. Ledoux. *The concentration of measure phenomenon*. Vol. 89. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2001, pp. x+181. ISBN: 0-8218-2864-9.

[86] M. Ledoux. "The geometry of Markov diffusion generators". In: *Ann. Fac. Sci. Toulouse Math. (6)* 9.2 (2000). Probability theory, pp. 305–366. ISSN: 0240-2963. URL: http://www.numdam.org/item?id=AFST_2000_6_9_2_305_0.

[87] M. Ledoux, I. Nourdin & G. Peccati. "Stein's method, logarithmic Sobolev and transport inequalities". In: *Geom. Funct. Anal.* 25.1 (2015), pp. 256–306. ISSN: 1016-443X. DOI: 10.1007/s00039-015-0312-0. URL: https://doi.org/10.1007/s00039-015-0312-0.

[88] A. Lichnerowicz. *Géométrie des groupes de transformations*. Travaux et Recherches Mathématiques, III. Dunod, Paris, 1958, pp. ix+193.

[89] S. Liu. "An optimal dimension-free upper bound for eigenvalue ratios". In: *ArXiv e-prints* (May 2014). arXiv: 1405.2213 [math.DG].

[90] J. Lott & C. Villani. "Ricci curvature for metric-measure spaces via optimal transport". In: *Ann. of Math. (2)* 169.3 (2009), pp. 903–991. ISSN: 0003-486X. URL: https://doi.org/10.4007/annals.2009.169.903.

[91] Y. Ma, S. Shen, X. Wang & L. Wu. "Transportation inequalities: from Poisson to Gibbs measures". In: *Bernoulli* 17.1 (2011), pp. 155–169. ISSN: 1350-7265. DOI: 10.3150/00-BEJ268. URL: https://doi.org/10.3150/00-BEJ268.

[92] P. Malliavin. "Stochastic calculus of variation and hypoelliptic operators". In: *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)*. Wiley, New York-Chichester-Brisbane, 1978, pp. 195–263.

[93] K. Marton. "A simple proof of the blowing-up lemma". In: *IEEE Trans. Inform. Theory* 32.3 (1986), pp. 445–446. ISSN: 0018-9448. DOI: 10.1109/TIT.1986.1057176. URL: https://doi.org/10.1109/TIT.1986.1057176.

[94] K. Marton. "Bounding $\overline{d}$-distance by informational divergence: a method to prove measure concentration". In: *Ann. Probab.* 24.2 (1996), pp. 857–866. ISSN: 0091-1798. DOI: 10.1214/aop/1039639365. URL: https://doi.org/10.1214/aop/1039639365.

[95] K. Marton. "A measure concentration inequality for contracting Markov chains". In: *Geom. Funct. Anal.* 6.3 (1996), pp. 556–571. ISSN: 1016-443X. DOI: 10.1007/BF02249263. URL: https://dx-doi-org.fennec.u-pem.fr/10.1007/BF02249263.

[96] B. Maurey. "Some deviation inequalities". In: *Geom. Funct. Anal.* 1.2 (1991), pp. 188–197. ISSN: 1016-443X. DOI: 10.1007/BF01896377. URL: https://doi.org/10.1007/BF01896377.

[97] R. J. McCann. "A convexity principle for interacting gases". In: *Adv. Math.* 128.1 (1997), pp. 153–179. ISSN: 0001-8708. DOI: 10.1006/aima.1997.1634. URL: https://doi.org/10.1006/aima.1997.1634.

[98] E. J. McShane. “Extension of range of functions”. In: *Bull. Amer. Math. Soc.* 40.12 (1934), pp. 837–842. ISSN: 0002-9904. DOI: 10.1090/S0002-9904-1934-05978-0. URL: https://doi.org/10.1090/S0002-9904-1934-05978-0.

[99] P.-A. Meyer et al. *Séminaire de probabilités*. 1967. URL: http://sites.mathdoc.fr/SemProba/.

[100] E. Milman. “Spectral Estimates, Contractions and Hypercontractivity”. In: *ArXiv e-prints* (Aug. 2015). arXiv: 1508.00606 [math.SP].

[101] E. Milman. “Isoperimetric and concentration inequalities: equivalence under curvature lower bound”. In: *Duke Math. J.* 154.2 (2010), pp. 207–239. ISSN: 0012-7094. DOI: 10.1215/00127094-2010-038. URL: http://dx.doi.org/10.1215/00127094-2010-038.

[102] E. Milman. “On the role of convexity in isoperimetry, spectral gap and concentration”. In: *Invent. Math.* 177.1 (2009), pp. 1–43. ISSN: 0020-9910. DOI: 10.1007/s00222-009-0175-9. URL: https://doi.org/10.1007/s00222-009-0175-9.

[103] E. Milman. “On the role of convexity in isoperimetry, spectral gap and concentration”. In: *Invent. Math.* 177.1 (2009), pp. 1–43. ISSN: 0020-9910. DOI: 10.1007/s00222-009-0175-9. URL: http://dx.doi.org/10.1007/s00222-009-0175-9.

[104] A. de Moivre. “Approximatio ad summam terminorum binomii (a+b)n in seriem expansi”. In: (1733), p. 7.

[105] A. de Moivre. *The doctrine of chances or, a method of calculating the probabilities of events in play*. New impression of the second edition, with additional material. Cass Library of Science Classics, No. 1. Frank Cass & Co., Ltd., London, 1967, pp. xxxii+257.

[106] E. Nelson. “The free Markoff field”. In: *J. Functional Analysis* 12 (1973), pp. 211–227.

[107] I. Nourdin. *Malliavin-Stein approach*. 2018. URL: https://sites.google.com/site/malliavinstein/home.

[108] I. Nourdin. *Selected aspects of fractional Brownian motion*. Vol. 4. Bocconi & Springer Series. Springer, Milan; Bocconi University Press, Milan, 2012, pp. x+122. ISBN: 978-88-470-2822-7; 978-88-470-2823-4. DOI: 10.1007/978-88-470-2823-4. URL: https://doi.org/10.1007/978-88-470-2823-4.

[109] I. Nourdin & D. Nualart. “Central limit theorems for multiple Skorokhod integrals”. In: *J. Theoret. Probab.* 23.1 (2010), pp. 39–64. ISSN: 0894-9840. DOI: 10.1007/s10959-009-0258-y. URL: http://dx.doi.org/10.1007/s10959-009-0258-y.

[110] I. Nourdin, D. Nualart & G. Peccati. “Quantitative stable limit theorems on the Wiener space”. In: *Ann. Probab.* 44.1 (2016), pp. 1–41. ISSN: 0091-1798. DOI: 10.1214/14-AOP965. URL: http://dx.doi.org/10.1214/14-AOP965.

[111] I. Nourdin & G. Peccati. *Normal approximations with Malliavin calculus*. Vol. 192. Cambridge Tracts in Mathematics. From Stein's method to universality. Cambridge University Press, Cambridge, 2012, pp. xiv+239. ISBN: 978-1-107-01777-1. DOI: 10.1017/CBO9781139084659. URL: http://dx.doi.org/10.1017/CBO9781139084659.

[112] I. Nourdin & G. Peccati. "Stein's method on Wiener chaos". In: *Probab. Theory Related Fields* 145.1-2 (2009), pp. 75–118. ISSN: 0178-8051. DOI: 10.1007/s00440-008-0162-x. URL: https://doi.org/10.1007/s00440-008-0162-x.

[113] I. Nourdin & J. Rosiński. "Asymptotic independence of multiple Wiener-Itô integrals and the resulting limit laws". In: *Ann. Probab.* 42.2 (2014), pp. 497–526. ISSN: 0091-1798. DOI: 10.1214/12-AOP826. URL: https://doi.org/10.1214/12-AOP826.

[114] D. Nualart & S. Ortiz-Latorre. "Central limit theorems for multiple stochastic integrals and Malliavin calculus". In: *Stochastic Process. Appl.* 118.4 (2008), pp. 614–628. ISSN: 0304-4149. DOI: 10.1016/j.spa.2007.05.004. URL: https://doi.org/10.1016/j.spa.2007.05.004.

[115] D. Nualart. *Malliavin calculus and its applications*. Vol. 110. CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2009, pp. viii+85. ISBN: 978-0-8218-4779-4. URL: https://doi.org/10.1090/cbms/110.

[116] D. Nualart & G. Peccati. "Central limit theorems for sequences of multiple stochastic integrals". In: *Ann. Probab.* 33.1 (2005), pp. 177–193. ISSN: 0091-1798. DOI: 10.1214/009117904000000621. URL: http://dx.doi.org/10.1214/009117904000000621.

[117] H. Ogura. "Orthogonal functionals of the Poisson process". In: *IEEE Trans. Information Theory* IT-18 (1972), pp. 473–481. ISSN: 0018-9448.

[118] F. Otto & C. Villani. "Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality". In: *J. Funct. Anal.* 173.2 (2000), pp. 361–400. ISSN: 0022-1236. DOI: 10.1006/jfan.1999.3557. URL: https://doi.org/10.1006/jfan.1999.3557.

[119] F. Otto. "The geometry of dissipative evolution equations: the porous medium equation". In: *Comm. Partial Differential Equations* 26.1-2 (2001), pp. 101–174. ISSN: 0360-5302. DOI: 10.1081/PDE-100002243. URL: https://doi.org/10.1081/PDE-100002243.

[120] G. Peccati. "The Chen-Stein method for Poisson functionals". In: *ArXiv e-prints* (Dec. 2011). arXiv: 1112.5051 [math.PR].

[121] G. Peccati, J. L. Solé, M. S. Taqqu & F. Utzet. "Stein's method and normal approximation of Poisson functionals". In: *Ann. Probab.* 38.2 (2010), pp. 443–478. ISSN: 0091-1798. URL: https://doi.org/10.1214/09-AOP477.

[122] G. Peccati & M. Reitzner, eds. *Stochastic analysis for Poisson point processes*. Vol. 7. Bocconi & Springer Series. Malliavin calculus, Wiener-Itô chaos expansions and stochastic geometry. Bocconi University Press, [place of publication not identified]; Springer, [Cham], 2016, pp. xv+346. ISBN: 978-3-319-05232-8; 978-3-319-05233-5. DOI: 10.1007/978-3-319-05233-5. URL: http://dx.doi.org/10.1007/978-3-319-05233-5.

[123] G. Peccati & M. S. Taqqu. "Central limit theorems for double Poisson integrals". In: *Bernoulli* 14.3 (2008), pp. 791–821. ISSN: 1350-7265. URL: https://doi.org/10.3150/08-BEJ123.

[124] G. Peccati & M. S. Taqqu. *Wiener chaos: moments, cumulants and diagrams*. Vol. 1. Bocconi & Springer Series. A survey with computer implementation, Supplementary material available online. Springer, Milan; Bocconi University Press, Milan, 2011, pp. xiv+274. ISBN: 978-88-470-1678-1. DOI: 10.1007/978-88-470-1679-8. URL: https://doi.org/10.1007/978-88-470-1679-8.

[125] G. Peccati & C. Thäle. "Gamma limits and $U$-statistics on the Poisson space". In: *ALEA Lat. Am. J. Probab. Math. Stat.* 10.1 (2013), pp. 525–560. ISSN: 1980-0436.

[126] G. Peccati & M. Yor. "Four limit theorems for quadratic functionals of Brownian motion and Brownian bridge". In: *Asymptotic methods in stochastics*. Vol. 44. Fields Inst. Commun. Amer. Math. Soc., Providence, RI, 2004, pp. 75–87.

[127] G. Peccati & C. Zheng. "Multi-dimensional Gaussian fluctuations on the Poisson space". In: *Electron. J. Probab.* 15 (2010), no. 48, 1487–1527. ISSN: 1083-6489. DOI: 10.1214/EJP.v15-813. URL: https://doi.org/10.1214/EJP.v15-813.

[128] G. Perelman. "The entropy formula for the Ricci flow and its geometric applications". In: *ArXiv Mathematics e-prints* (Nov. 2002). eprint: math/0211159.

[129] V. Pipiras & M. S. Taqqu. "Integration questions related to fractional Brownian motion". In: *Probab. Theory Related Fields* 118.2 (2000), pp. 251–291. ISSN: 0178-8051. DOI: 10.1007/s440-000-8016-7. URL: https://doi.org/10.1007/s440-000-8016-7.

[130] S.-D. Poisson. *Recherches sur la probabilité des jugements en matière criminelle et en matière civile ; précédées des Règles générales du calcul des probabilités*. Bachelier (Paris), 1837, pp. ix+415.

[131] L. Pratelli & P. Rigo. "Total variation bounds for Gaussian functionals". In: *Stochastic Process. Appl.* (), to appear. URL: http://www-dimat.unipv.it/~rigo/frac.pdf.

[132] M. Reitzner. "Poisson point processes: large deviation inequalities for the convex distance". In: *Electron. Commun. Probab.* 18 (2013), no. 96, 7. ISSN: 1083-589X. DOI: 10.1214/ECP.v18-2851. URL: https://dx-doi-org.fennec.u-pem.fr/10.1214/ECP.v18-2851.

[133] M. Reitzner & M. Schulte. "Central limit theorems for $U$-statistics of Poisson point processes". In: *Ann. Probab.* 41.6 (2013), pp. 3879–3909. ISSN: 0091-1798. DOI: 10.1214/12-AOP817. URL: https://doi.org/10.1214/12-AOP817.

[134] A. Rényi. "On mixing sequences of sets". In: *Acta Math. Acad. Sci. Hungar.* 9 (1958), pp. 215–228. ISSN: 0001-5954. DOI: 10.1007/BF02023873. URL: https://doi.org/10.1007/BF02023873.

[135] A. Rényi. "On stable sequences of events". In: *Sankhyā Ser. A* 25 (1963), p. 293 302. ISSN: 0581-572X.

[136] G.-C. Rota & T. C. Wallstrom. "Stochastic integrals: a combinatorial approach". In: *Ann. Probab.* 25.3 (1997), pp. 1257–1283. ISSN: 0091-1798. DOI: 10.1214/aop/1024404513. URL: https://doi.org/10.1214/aop/1024404513.

[137] M. Schmuckenschläger. "Martingales, Poincaré type inequalities, and deviation inequalities". In: *J. Funct. Anal.* 155.2 (1998), pp. 303–323. ISSN: 0022-1236. DOI: 10.1006/jfan.1997.3218. URL: http://dx.doi.org/10.1006/jfan.1997.3218.

[138] M. Schulte. "A central limit theorem for the Poisson-Voronoi approximation". In: *Adv. in Appl. Math.* 49.3-5 (2012), pp. 285–306. ISSN: 0196-8858. DOI: 10.1016/j.aam.2012.08.001. URL: https://doi.org/10.1016/j.aam.2012.08.001.

[139] M. Sion. "On general minimax theorems". In: *Pacific J. Math.* 8 (1958), pp. 171–176. ISSN: 0030-8730. URL: http://projecteuclid.org/euclid.pjm/1103040253.

[140] A. J. Stam. "Some inequalities satisfied by the quantities of information of Fisher and Shannon". In: *Information and Control* 2 (1959), pp. 101–112. ISSN: 0890-5401. DOI: 10.1016/S0019-9958(59)90348-1. URL: https://doi.org/10.1016/S0019-9958(59)90348-1.

[141] K.-T. Sturm. "On the geometry of metric measure spaces. I". In: *Acta Math.* 196.1 (2006), pp. 65–131. ISSN: 0001-5962. URL: https://doi.org/10.1007/s11511-006-0002-8.

[142] K.-T. Sturm. "On the geometry of metric measure spaces. II". In: *Acta Math.* 196.1 (2006), pp. 133–177. ISSN: 0001-5962. URL: https://doi.org/10.1007/s11511-006-0003-7.

[143] D. Surgailis. "On multiple Poisson stochastic integrals and associated Markov semigroups". In: *Probab. Math. Statist.* 3.2 (1984), pp. 217–239. ISSN: 0208-4147.

[144] M. Talagrand. "Transportation cost for Gaussian and other product measures". In: *Geom. Funct. Anal.* 6.3 (1996), pp. 587–600. ISSN: 1016-443X. DOI: 10.1007/BF02249265. URL: https://doi.org/10.1007/BF02249265.

[145] M. Talagrand. "Concentration of measure and isoperimetric inequalities in product spaces". In: *Inst. Hautes Études Sci. Publ. Math.* 81 (1995), pp. 73–205. ISSN: 0073-8301. URL: http://www.numdam.org/item?id=PMIHES_1995__81__73_0.

[146] C. Villani. *Optimal transport*. Vol. 338. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Old and new. Springer-Verlag, Berlin, 2009, pp. xxii+973. ISBN: 978-3-540-71049-3. DOI: 10.1007/978-3-540-71050-9. URL: https://doi.org/10.1007/978-3-540-71050-9.

[147] F.-Y. Wang. “Functional inequalities for empty essential spectrum”. In: *J. Funct. Anal.* 170.1 (2000), pp. 219–245. ISSN: 0022-1236. DOI: 10.1006/jfan.1999.3516. URL: https://doi.org/10.1006/jfan.1999.3516.

[148] H. Whitney. “Analytic extensions of differentiable functions defined in closed sets”. In: *Trans. Amer. Math. Soc.* 36.1 (1934), pp. 63–89. ISSN: 0002-9947. DOI: 10.2307/1989708. URL: https://doi.org/10.2307/1989708.

[149] N. Wiener. “Generalized harmonic analysis”. In: *Acta Math.* 55.1 (1930), pp. 117–258. ISSN: 0001-5962. DOI: 10.1007/BF02546511. URL: https://doi.org/10.1007/BF02546511.

[150] N. Wiener. “The Homogeneous Chaos”. In: *Amer. J. Math.* 60.4 (1938), pp. 897–936. ISSN: 0002-9327. DOI: 10.2307/2371268. URL: https://doi.org/10.2307/2371268.

[151] N. Wiener & A. Wintner. “The discrete chaos”. In: *Amer. J. Math.* 65 (1943), pp. 279–298. ISSN: 0002-9327. DOI: 10.2307/2371816. URL: https://doi.org/10.2307/2371816.

[152] L. Wu. “A new modified logarithmic Sobolev inequality for Poisson point processes and several applications”. In: *Probab. Theory Related Fields* 118.3 (2000), pp. 427–438. ISSN: 0178-8051. URL: https://doi.org/10.1007/PL00008749.

[153] G. Zheng. “A Peccati-Tudor type theorem for Rademacher chaoses”. In: *ArXiv e-prints* (Aug. 2017). arXiv: 1708.05283 [math.PR].