



**HAL**  
open science

# Contributions to stochastic analysis for non-diffusive structures

Christophe Vuong

► **To cite this version:**

Christophe Vuong. Contributions to stochastic analysis for non-diffusive structures. Probability [math.PR]. Institut Polytechnique de Paris, 2023. English. NNT : 2023IPPAT054 . tel-04513594

**HAL Id: tel-04513594**

**<https://theses.hal.science/tel-04513594>**

Submitted on 20 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INSTITUT  
POLYTECHNIQUE  
DE PARIS

NNT : 2023IPPAT054

Thèse de doctorat



# Contributions to stochastic analysis for non-diffusive structures

Thèse de doctorat de l'Institut Polytechnique de Paris  
préparée à Télécom Paris

École doctorale n°574 École doctorale de mathématiques Hadamard (EDMH)  
Spécialité de doctorat : Mathématiques appliquées

Thèse présentée et soutenue à Palaiseau, le 18 décembre 2023, par

**CHRISTOPHE VUONG**

Composition du Jury :

Laure Coutin Professeure, Université de Toulouse	Présidente/Examinatrice
Nicolas Privault Professor, Nanyang Technological University	Rapporteur
Raphaël Lachièze-Rey Maître de conférences HDR, Université Paris Cité	Rapporteur
Hélène Halconrui Maîtresse de conférences, Télécom SudParis	Examinatrice
Laurent Decreusefond Professeur, Télécom Paris	Directeur de thèse
Marc Glisse Chargé de Recherche, INRIA Saclay - Île de France	Co-directeur de thèse



# Contents

<b>Remerciements</b>	<b>5</b>
<b>Notations</b>	<b>9</b>
<b>Introduction</b>	<b>10</b>
<b>I Stein-Malliavin-Dirichlet method and applications to statistics in hypergraph theory</b>	<b>21</b>
<b>1 Stein-Dirichlet-Malliavin method</b>	<b>23</b>
1.1 Background . . . . .	23
1.2 Probability distances . . . . .	24
1.3 Stein’s method principle . . . . .	25
1.4 Quantitative fourth moment theorems . . . . .	29
1.4.1 Origins . . . . .	29
1.4.2 The Markov triple approach . . . . .	31
1.4.3 Adaptations to non-diffusive structures . . . . .	32
1.5 Contributions . . . . .	33
<b>2 Malliavin calculus for conditionally independent random variables</b>	<b>34</b>
2.1 Motivation . . . . .	34
2.2 Discrete Malliavin-Dirichlet structure . . . . .	35
2.2.1 Malliavin operators . . . . .	37
2.2.2 Chaos decomposition . . . . .	39
2.2.3 Dirichlet structure . . . . .	42
2.3 Functional identities . . . . .	46
2.4 Applications to normal approximation . . . . .	50
2.4.1 Bounds in probability distance . . . . .	50
2.4.2 Rates in Lyapunov’s conditional central limit . . . . .	51
2.4.3 Abstract bounds for U-statistics . . . . .	53
2.5 Partial fourth moment theorems . . . . .	57
<b>3 Motif estimation</b>	<b>69</b>
3.1 General context . . . . .	69
3.2 Definitions . . . . .	70
3.3 Normal approximation . . . . .	73

3.4	A modified Hoeffding decomposition . . . . .	79
3.5	Future research directions . . . . .	84
<b>II</b>	<b>Invertibility of functionals of Poisson measures</b>	<b>87</b>
<b>4</b>	<b>Girsanov transforms of Poisson measures</b>	<b>91</b>
4.1	Background . . . . .	91
4.2	Preliminaries . . . . .	93
4.3	Random change of marks . . . . .	95
4.4	Change of measures and Girsanov theorem . . . . .	97
<b>5</b>	<b>Invertibility framework</b>	<b>103</b>
5.1	Overview . . . . .	103
5.2	Stochastic invertibility on the Poisson space with marks . . . . .	104
5.3	Entropy applications . . . . .	108
5.4	Solutions of SDEs driven by Poisson measures . . . . .	115
5.5	Further works . . . . .	116
	<b>Bibliography</b>	<b>119</b>

# Remerciements

Je tiens d'abord à remercier le jury pour avoir pris le temps d'examiner ma thèse, tant à travers les rapports rédigés sur le manuscrit qu'à travers les interrogations lors de la soutenance.

J'aimerais remercier Laurent et Marc, mes directeurs de thèse pour l'accompagnement durant ces trois années et notamment les nombreuses idées et pistes de recherches que vous m'avez proposées. Vous m'avez fait découvrir le monde de la recherche. À Laurent, j'adresse ma reconnaissance pour ta précieuse guidance et ta compréhension à mon égard. Ton expertise, ta patience et tes encouragements ont été essentiels pour mener à bien ma thèse. Je n'oublierai pas ta passion et ta bonne humeur durant ces trois ans. Tu as su me faire progresser et me motiver, même dans les moments difficiles.

Ensuite, je souhaiterais remercier mon laboratoire, et en particulier les membres de l'équipe Philippe, Anaïs et Jean-Sébastien pour l'accueil et les échanges enrichissants.

Un grand merci à mon collègue et ami Bruno, qui a partagé avec moi deux années de thèse. Tu as été un compagnon de route toujours présent, toujours prêt à aider. Tu as rendu cette aventure plus agréable et plus humaine. Je te souhaite le meilleur pour la suite.

Je n'oublie pas les professeurs qui m'ont formé durant mes études, au lycée, en classes préparatoires ou à Télécom Paris et au MVA. Vous avez su éveiller ma curiosité et mon intérêt pour les sciences jusqu'à me donner envie de poursuivre en thèse.

Enfin, je tiens à remercier du fond du cœur mes proches et en particulier ma sœur Caroline et mes parents Valérie et Jacques, pour votre soutien au quotidien et votre dévouement. Sans vous, rien n'aurait été possible. Merci pour tout.



# Notations

## Abbreviations

c.d.f.	Cumulative distribution function
i.i.d	Independent identically distributed
SBM	Stochastic block model
MPP	Marked point process
SDE	Stochastic differential equations

## Mathematical relations

$a \vee b$	$\max(a, b)$
$a \wedge b$	$\min(a, b)$
$\lesssim$	Lesser or equal up to a constant
$\gtrsim$	Greater or equal up to a constant
$\propto_Z$	Equal up to a constant depending on $Z$

## Usual sets

$\mathbb{N}$	Set of integers including 0
$\mathbb{R}$	Set of real numbers
$\mathbb{R}^+$	Set of non-negative real numbers
$\mathbb{R}^d$	Set of real valued vectors of dimension $d$
$\mathbb{Q}$	Set of rational numbers
$\mathcal{P}(A)$	Power set of the set $A$
$(A, p)$	Set of $p$ -tuples of $A$

## Functional analysis

Id	Identity on a given space
$\mathcal{F}_1 \otimes \mathcal{F}_2$	Product $\sigma$ -algebra of $\sigma$ -algebras $\mathcal{F}_1$ and $\mathcal{F}_2$
$L^p(E \rightarrow F, \mu)$	The set of $p$ -integrable functions $f : E \rightarrow F$ with respect to a measure $\mu$



$L^\infty(E \rightarrow F, \mu)$	The set of bounded functions $f : E \rightarrow F$ with respect to a measure $\mu$
$\langle \cdot, \cdot \rangle_E$	Inner product on Banach space $E$
$\  \cdot \ _E$	Norm on the Banach space $E$
$\  \cdot \ _\infty$	Infinity norm
$W_B$	Space identified to the Wiener space
$\langle \cdot, \cdot \rangle_{W_B^*, W_B}$	Dual bracket between element of $W_B^*$ and $W_B$
$L^{-1}$	Pseudoinverse of an operator $L$
$d_W(F, \mathcal{N}(0, 1))$	Wasserstein distance between the probability law of $F$ and the standardized normal distribution
$\ker$	Kernel
$\bigoplus_{k=0}^q \cdot$	Hilbert space direct sum
$\nu$	A Borel measure
$\mathbf{1}_B$	Indicator function of the set $B$
$\epsilon_x$	Dirac mass at $x$
<u>Probability</u>	
$\mathbb{E}[F]$	Expectation of random variable $F$
$\mathbb{E}[F   Z]$	Conditional expectation of $F$ given the $\sigma$ -algebra generated by the random variable $Z$
$\mathbb{P}$	Probability
$\mathbf{E}_\mu[\cdot]$	Expectation under the probability measure $\mu$
<i>a.s.</i>	Almost surely
$\mathcal{N}(0, 1)$	Standard Gaussian distribution
$\mathcal{N}(\mu, \Sigma)$	Gaussian distribution with mean vector $\mu$ and covariance matrix $\Sigma$
$\perp$	Independent
$\perp_Z$	Conditionally independent given a latent random variable $Z$
$X \stackrel{d}{=} Y$	The random variable $X$ is equal in distribution to a random variable $Y$
$P \stackrel{d}{=} \mathcal{B}(p_n)$	$P$ is the probability measure associated to the Bernoulli distribution
$X_n \xrightarrow{d} Y$	The sequence of random variables $(X_n)_{n \in \mathbb{N}}$ converge in distribution to a random variable $Y$
$\mathcal{S}$	The set of cylindrical random variables given a Malliavin framework
$D$	Gradient operator in a Malliavin framework
$\delta$	Divergence operator in a Malliavin framework (adjoint of $D$ )
$U_{\#}\mu$	The pushforward of the measure $\mu$ by a map $U$

$\left. \frac{d\pi'}{d\pi} \right|_{\mathcal{F}_t}$  Conditional expectation of the Radon-Nikodym derivative given  $\mathcal{F}_t$

Hypergraphs

$[n]$	Set $\{1, \dots, n\}$
$V(G)$	Set of vertices $V$ of the hypergraph $G = (V, E)$
$E(G)$	Set of hyperedges $E$ of $G = (V, E)$
$v_H$	Number of vertices of a hypergraph $H$
$e_H$	Number of hyperedges of a hypergraph $H$
$v(H)$	Number of vertices in the hypergraph $H$
$\binom{V}{k}$	Collection of $k$ -tuples of the set of vertices $V$
$ E $	Cardinality of the set $E$ otherwise absolute value
$I \simeq G$	The hypergraph $I$ is isomorphic to $G$

# Introduction

## Diffusion

Les *diffusions* sont des fonctions aléatoires, qui sont très utilisées en physique, chimie, biologie, statistique et en finance. Leur nature même en fait un outil de modélisation formidable : elle permet de capter des dynamiques instantanées entachées d'incertitude. Au-delà de leur intérêt descriptif, elles se prêtent aux utilisations quantitatives. D'un point de vue probabiliste, ce sont des processus solutions d'équations différentielles stochastiques d'un certain type. Hormis les processus gaussiens, il n'est pas possible de les décrire en spécifiant les marginales finidimensionnelles. Les diffusions connues peuvent être décrites selon leur dynamique, par exemple l'équation de Langevin:

$$dY(t) = \sigma dB(t) - bY(t) dt$$

où  $\sigma$  et  $b$  sont des paramètres réels de l'équation tandis que  $B$  est un mouvement Brownien. Pour simplifier nos propos dans cette introduction, on considère que le processus est à valeurs dans  $\mathbb{R}$ . Il est connu que le mouvement Brownien tout comme les diffusions sont des processus de Markov (Stroock et Varadhan (2006)). Un processus de Markov  $Y$  est caractérisé par son opérateur de semi-groupe  $P$ . En le laissant agir sur une classe suffisamment riche de fonctions  $f : \mathbb{R} \rightarrow \mathbb{R}$  boréliennes bornées :

$$(P_t f)(x) := \mathbb{E} [f(Y(t)) | Y(0) = x].$$

En général, on considère un espace de Banach séparable, muni de la norme  $\|\cdot\|_\infty$ , de fonctions continues tendant vers 0 à l'infini sur lesquelles le semi-groupe agit, dénoté  $\mathcal{C}_0 = \mathcal{C}_0(\mathbb{R}, \mathbb{R})$ . Nous rappelons le théorème suivant.

**Théorème**  $(P_t, t \geq 0)$  est un semi-groupe conjointement continu de contractions sur  $\mathcal{C}_0$ , i.e.,

1.  $\|P_t f\|_\infty \leq \|f\|_\infty, \forall t, f$ ;
2.  $P_{t+s} = P_t \circ P_s$ ;
3.  $P_0 = \text{Id}$ ;
4.  $(t, f) \mapsto P_t f$  est continu de  $\mathbb{R}_+ \times \mathcal{C}_0 \mapsto \mathcal{C}_0$ .

Ainsi, on définit le domaine du générateur comme le sous-espace vectoriel

$$\text{Dom } L := \{f \in \mathcal{C}_0 : \lim_{t \searrow 0} \frac{1}{t} (P_t f - f) \text{ existe}\},$$

et pour  $f \in \text{Dom } L$ , on définit  $Lf$  par la valeur de cette limite. L'opérateur linéaire  $L$  est appelé *générateur infinitésimal*. Le théorème suivant résulte du théorème de Hille-Yosida (Théorème 2.6 dans Ethier et Kurtz (1986)).

- Théorème**
1.  $\text{Dom } L$  est dense dans  $(\mathcal{C}_0, \|\cdot\|_\infty)$ .
  2.  $P_t$  laisse stable  $\text{Dom } L$  :  $P_t(\text{Dom } L) \subset \text{Dom } L$ .
  3.  $\forall f \in \text{Dom } L, \frac{d}{dt}P_t f = P_t Lf = LP_t f$ .

Puisque  $P_t$  est solution de l'équation différentielle  $\frac{d}{dt}P_t f = LP_t f$ , on écrit formellement:

$$P_t = \exp(tL). \quad (1)$$

Par l'inégalité de Jensen, pour toute fonction convexe  $\phi : \mathbb{R} \rightarrow \mathbb{R}$ , et tout  $t \geq 0$ , et toute fonction  $f \in \mathcal{C}_0(\mathbb{R}, \mathbb{R})$ ,

$$P_t(\phi(f)) \geq \phi(P_t f). \quad (2)$$

Par différentiation, en passant à la limite pour  $t \rightarrow 0$ , on a que:

$$L\phi(f) \geq \phi'(f)Lf. \quad (3)$$

Pour de nombreuses applications sur les processus de Markov, il est commode que (3) soit une égalité pour des fonctions  $\phi$  suffisamment régulières, c'est-à-dire que:

$$L\phi(f) = \phi'(f)Lf + \text{erreur}. \quad (4)$$

Pour écrire cette erreur dans le cadre d'une diffusion, on se réfère à la formule de transport suivante :

$$\begin{cases} \text{Var } f(C) = \text{Var}(C)f'^2(C) \\ \text{biais } f(C) = (\text{biais } C)f'(C) + \frac{1}{2} \text{Var}(C)f''(C) \end{cases}$$

où  $C$  est une quantité aléatoire. L'intuition de Bouleau consiste à identifier le biais et l'erreur quadratique dans cette formule avec les opérateurs  $L$  et  $\Gamma$  défini ci-dessous, qui forment une structure de Dirichlet. L'opérateur carré du champ est défini par

$$\Gamma(f, g) := \frac{1}{2} (L(fg) - gLf - fLg)$$

pour  $f, g \in \mathcal{A}$  une algèbre dans  $\text{Dom } L$  telle que le produit de fonctions  $fg$  est dans le domaine de  $L$ . Un générateur est dit *diffusif* si:

$$L\phi(f) = \phi'(f)Lf + \phi''(f)\Gamma(f, f). \quad (5)$$

La terminologie d'"opérateur carré du champ" a été introduite par Roth (1976) dans un contexte un peu différent au cours de ses années doctorales, and a été popularisée grâce à Meyer (1976) pour la caractérisation de semi-groupes de Feller en analyse stochastique. À l'origine, le terme provient de l'électrostatique.

## Liens entre la théorie des formes de Dirichlet et le calcul de Malliavin

Le monographe [Bakry et al. \(2013\)](#) a dédié son chapitre 2 à l'étude des semi-groupes associés parmi lesquels le semi-groupe d'Ornstein-Uhlenbeck, le semi-groupe de Laguerre et le semi-groupe de Jacobi. D'importantes conséquences en découlent comme des identités fonctionnelles et l'existence d'inégalités de Poincaré (voir aussi [Ledoux, 2000](#)). En particulier pour ces dernières, les résultats ont obtenus en considérant la décomposition spectrale de  $L$  dans l'espace de probabilité d'intérêt, c'est-à-dire que l'on se place dans un sous-espace de fonctions  $\mathfrak{C}$  tel qu'il existe  $\lambda \in \mathbb{R}^+$  pour lequel

$$f \in \mathfrak{C} \implies Lf = -\lambda f.$$

Ceux-ci s'inscrivent plus généralement dans la théorie des formes de Dirichlet apparue dans la fin des années 50 au travers des travaux de Anne Beurling et Jacques Deny ([Beurling et Deny \(1958\)](#); [Beurling et Deny \(1959\)](#)). C'était le point de départ d'une théorie florissante avec une interprétation en termes de processus de Markov. Dans les années 70, les idées de Paul Malliavin pour l'extension du raisonnement d'intégration par parties aux dimensions infinies et l'étude de la structure de l'espace de Wiener grâce au processus d'Ornstein-Uhlenbeck en dimension infinie, ont entraîné un nouveau vivier d'applications des formes de Dirichlet ([Bouleau et Hirsch \(1991\)](#)). Le calcul de Malliavin est un outil puissant qui permet d'étudier les propriétés des processus stochastiques, tels que les processus de Wiener (voir e.g. [Nualart, 2006](#)), les processus de Markov et les processus de Lévy (voir Section 9 de [Bichteler et al., 1987](#)). Il est basé sur la notion de dérivée stochastique, qui est une généralisation de la dérivée ordinaire aux processus stochastiques. Une structure de Malliavin est propre à l'espace de probabilité dans lequel est défini lesdits processus stochastiques. De manière générale, la construction d'une telle structure consiste en la définition d'un gradient  $D$ , d'une divergence  $\delta$  et de l'opérateur laplacien  $\hat{L} = -\delta D$  sur leurs domaines respectifs. Prenons l'espace de Wiener supporté par l'espace  $W_B = C^0([0, 1]; \mathbb{R})$  des fonctions continues sur  $[0, 1]$  et muni de la mesure de Wiener  $\mu_B$ , et  $L^2([0, 1] \rightarrow \mathbb{R}; \ell)$  où  $\ell$  est la mesure de Lebesgue sur  $[0, 1]$ , l'espace identifié au dual topologique de  $W_B$ , noté  $W_B^*$ . Soit  $B = (B_t)_{t \in [0, 1]}$  le mouvement Brownien. Il est connu que l'espace  $L^2(W_B \rightarrow \mathbb{R}; \ell)$  peut être décomposé en une somme d'Hilbert de sous-espaces fermés  $\mathfrak{C}_n$  générés par les variables aléatoires du type

$$I_n(f) = H_n \left( \int_0^1 f(t) dB(t) \right) \quad (6)$$

où  $H_n$  est le polynôme d'Hermite d'ordre  $n$  défini par:

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} (e^{-x^2/2})$$

pour  $n \geq 0$ . Définissons pour  $t \in (0, 1]$

$$\mathcal{T}_n = \{(t_1, \dots, t_n) \in [0, 1]^n : 0 \leq t_1 \leq \dots \leq t_n \leq 1\}.$$

Pour  $f \in L^2(\mathcal{T}_n \rightarrow \mathbb{R}; \ell)$ , on a:

$$I_n(f) = \int_0^1 dB(t_n) \int_0^{t_n} dB(t_{n-1}) \int_0^{t_{n-1}} f(t_1, \dots, t_n) dB(t_1), \quad (7)$$

où les intégrales sont des intégrales d'Itô. Elles vérifient la relation de récurrence :

$$I_n(f)(t) = \int_0^1 I_{n-1}(f(\cdot, t_p)) \, dB(t_p). \quad (8)$$

Étant donné  $n \geq 0$ , le sous-espace  $\mathfrak{C}_n$  est appelé  $n$ -ème chaos de Wiener.  $\mathfrak{C}_0$  est le sous-espace des fonctionnelles constantes presque sûrement. Les chaos sont commodes pour définir les opérateurs de Malliavin. Soit  $(h_k)_{k \in \mathbb{N}}$  une base hilbertienne de  $L^2([0, 1] \rightarrow \mathbb{R}; \ell)$ , l'opérateur de dérivée ou gradient est défini comme suit.

**Définition** — **Opérateur de dérivée.** Soit  $\mathcal{S}$  l'ensemble des variables aléatoires cylindriques du type

$$F = f \left( \int_0^1 h_0(t) \, dB(t), \dots, \int_0^1 h_n \, dB(t) \right),$$

où  $f$  appartient à  $C_p^\infty(\mathbb{R}^n)$ ,  $h_1, \dots, h_n \in L^2([0, 1] \rightarrow \mathbb{R}; \ell)$ , et  $n \geq 1$ . Sa dérivée est la variable aléatoire à valeurs dans  $L^2([0, 1] \rightarrow \mathbb{R}; \ell)$  donnée par :

$$DF = \sum_{i=1}^n \frac{\partial_i f}{\partial x_i} \left( \int_0^1 h_0(t) \, dB(t), \dots, \int_0^1 h_n(t) \, dB(t) \right) h_i. \quad (9)$$

Nous notons  $\delta$  son adjoint appelé communément opérateur de divergence.

Pour relier ce gradient à l'intégrale stochastique d'Itô, on peut utiliser une injection de  $\ell^2(\mathbb{N})$  dans  $L^2([0, 1] \rightarrow \mathbb{R}; \ell)$  que nous noterons  $i : \ell^2(\mathbb{N}) \rightarrow L^2([0, 1] \rightarrow \mathbb{R}; \ell)$ , définie par  $i(e_k) = h_k$   $k \in \mathbb{N}$ , où  $(e_k)_{k \in \mathbb{N}}$  est la base canonique de  $\ell_2(\mathbb{N})$ . La dérivée peut alors être définie sur les fonctionnelles régulières par  $\hat{D} = i \circ D$ . Alors l'adjoint de  $\hat{D}$  noté  $\delta$  est une extension de l'intégrale stochastique d'Itô, cf. Gaveau et Trauber (1982). Les opérateurs  $D$  et  $\delta$  sont liés par la relation de dualité qui correspond à l'intégration par parties établie par Stroock (1981) [155; 156; 157].

### Théorème

$$\mathbb{E}[\langle DF, \eta \rangle_{W_B^*, W_B}] = \mathbb{E}[F \delta U]$$

pour  $\eta \in W_B$ .

L'opérateur de Malliavin vérifie la règle de dérivation en chaîne suivante. Soit  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  une fonction continûment dérivable à dérivée bornée et  $F$  une fonctionnelle dans  $\text{Dom } D$ , alors:

$$D(\varphi(F)) = \varphi'(F)DF. \quad (10)$$

L'opérateur  $\hat{L}$  est défini sur chacun des espaces  $\mathfrak{C}_n = \{I_n(f); f \in L^2(\mathcal{T}_n \rightarrow \mathbb{R}; \ell)\}$  par:  $\hat{L}(I_n(f)) = -p I_p(f)$  pour  $p > 1$ . Cet opérateur coïncide avec le générateur infinitésimal associé au semi-groupe d'Ornstein-Uhlenbeck. C'est une des nombreuses connexions entre la théorie de Dirichlet et le calcul de Malliavin qui sont étudiées dans le monographe Bouleau et Hirsch (1991). Par abus de langage, dans la littérature sur les applications de calcul de Malliavin, l'opérateur d'Ornstein-Uhlenbeck désigne tout opérateur  $L = -\delta D$  où les opérateurs de gradient et divergence sont associés à une structure de Malliavin quelconque. La dénomination d'opérateur de nombre pour  $L$ , provenant de la théorie quantique des champs euclidienne est aussi couramment employée. Dans la plupart des papiers, celui-ci est défini par son action sur les chaos (Yan, 1987) et son domaine est défini comme :

$$\text{Dom } \hat{L} = \{F \in L^2(W_B \rightarrow \mathbb{R}; \ell) : \sum_{q=0}^{\infty} q^2 \|J_q(F)\|_{L^2(W_B \rightarrow \mathbb{R}; \ell)}^2 < \infty\}.$$

La propriété (10) est alors équivalente à la diffusivité du générateur  $\hat{L}$ . L'un des intérêts de considérer une formes de Dirichlet ou un calcul de Malliavin est de pouvoir utiliser plusieurs propriétés qui s'expriment en fonction de l'éventail d'opérateurs. Une propriété du semi-groupe d'Ornstein-Uhlenbeck défini formellement par (1) est la *formule de Mehler*.

**Théorème** Pour tout  $F \in L^2(W_B \rightarrow \mathbb{R}; \ell)$ ,

$$P_t F(\omega) = \int_{W_B} F(e^{-t}\omega + \sqrt{1 - e^{-2y}}y) d\mu_B(y). \quad (11)$$

La formule de Mehler présente le semi-groupe d'Ornstein-Uhlenbeck comme une sorte d'opérateur de convolution. Une conséquence importante est que le semi-groupe est *ergodique* et admet la mesure de Wiener comme mesure stationnaire. On en déduit aussi cette formule de commutativité entre le gradient et le semi-groupe dans le cas où le membre de gauche est défini :

$$P_t(DF) = e^{-t}D(P_t F). \quad (12)$$

Ce schéma de construction de calcul de Malliavin est analogue pour les structures de Dirichlet associées au semi-groupe de Laguerre et de Jacobi. La notion de calcul de Malliavin a été étendue sur d'autres espaces et pour diverses applications en probabilités. Nous notons l'existence d'une structure de Malliavin pour des fonctionnelles de suites de variables aléatoires indépendantes suivant la loi uniforme sur  $[-1, 1]$  (Privault, 1997) qui fait le lien avec les formes de Dirichlet. Contrairement aux structures de Malliavin qui suit, elle a l'avantage d'être diffuse, tout en constituant un calcul de Malliavin dit *discret*, c'est-à-dire qui s'applique à des fonctionnelles de processus indexé sur un espace dénombrable.

## Structures sans propriété de diffusion

Pourtant, les structures sans propriété de diffusion donnant lieu à des décompositions en chaos sont légion (Biane (1990); Nualart et Schoutens (2000); Privault (2009)).

Les fonctionnelles de processus de Poisson rentrent dans ce cadre étant donné que un processus de Poisson  $N$  (respectivement une mesure de Poisson) peut être vu comme un élément de l'espace des configurations sur  $\mathbb{R}^+$  (respectivement  $\mathbb{R}^d$  avec  $d \geq 1$ ). La structure de Malliavin qui lui est associée définit le gradient  $D$  comme un opérateur de différence:

$$D_x F(N) = F_x(N) - F(N)$$

pour  $x \in \mathbb{R}^+$  (respectivement  $\mathbb{R}^d$ ). Sur l'espace de Poisson  $F_x$  correspond à une modification de la trajectoire  $F(N)$  induite par un ajout de  $x$  dans la configuration  $N$  (Dermoune et al., 1988; Ito, 1988; Picard, 1996) :

$$F_x(N) = F(N \oplus x) = \begin{cases} F(N) & \text{si } x \in N \\ F(N \cup \{x\}) & \text{si } x \notin N. \end{cases}$$

À l'origine, Malliavin a introduit la notion de calcul qui porte son nom concernant l'analyse de processus continus. Bismut a été l'un des premiers à transposer les résultats connus pour les processus de diffusion aux processus de Markov avec sauts (Bismut, 1983). Il utilise la caractérisation des processus de Markov et leur générateur infinitésimal. Par ailleurs, les résultats sur les équations différentielles dirigées par une mesure de Poisson ne sont pas des extensions

directes des résultats pour celles dirigées par un mouvement Brownien (voir par exemple les résultats de la seconde partie du manuscrit). Néanmoins les résultats d'analyse stochastique évoqués plus haut sont valables sur l'espace de Poisson. Il en est de même pour l'espace de Rademacher. Les résultats se basent sur la représentation chaotique pour des suites de variables indépendantes identiquement distribuées (Privault, 2008).

Soit  $(p_k)_{k \in \mathbb{N}}$  une suite de réels appartenant à  $[0, 1]$ . Sur l'espace de Rademacher  $\Omega$ , pour des fonctionnelles de variables aléatoires indépendantes de Rademacher  $(X_k)_{k \in \mathbb{N}}$  tel que  $X_k - 1 \sim \mathcal{B}(p_k)$  pour  $k \in \mathbb{N}$ , le gradient de la structure de Malliavin est aussi défini comme un opérateur de différence (Privault et Schoutens (2002)). Soit

$$\omega = (\omega_0, \omega_1, \dots) \in \{-1, 1\}^{\mathbb{N}},$$

$$\omega_+^k = (\omega_1, \omega_1, \dots, \omega_{k-1}, +1, \omega_{k+1}, \dots)$$

et

$$\omega_-^k = (\omega_1, \omega_1, \dots, \omega_{k-1}, -1, \omega_{k+1}, \dots)$$

alors le gradient s'exprime comme suit:

$$D_k F(\omega) = \sqrt{p_k(1-p_k)}(F(\omega_+^k) - F(\omega_-^k)).$$

Cet opérateur apparaît plus tôt dans le contexte de probabilités quantiques sous la dénomination  $a_k^-$  (Meyer, 1993, §2-2). L'opérateur d'Ornstein-Uhlenbeck correspondant est de la forme:

$$LF = \sum_{k=0}^{\infty} D_k F.$$

Comme sur l'espace de Wiener et de Poisson, une fonctionnelle de carré intégrable admet une décomposition en chaos. Soit  $l^2(\mathbb{N})^{\otimes n}$  le sous-espace de  $l^2(N)^{\otimes n} = l^2(\mathbb{N}^n)$  des fonctions  $f_n$  qui sont symétriques en  $n$  variables, c'est-à-dire qu'à chaque permutation  $\sigma$  de  $\{1, \dots, n\}$ ,  $f_n(k_{\sigma(1)}, \dots, k_{\sigma(n)}) = f_n(k_1, \dots, k_n)$ , pour  $k_1, \dots, k_n \in \mathbb{N}$ , on a :

$$F = \mathbb{E}[F] + \sum_{n=1}^{\infty} J_n(f_n)$$

où  $J_n(f_n)$  sont de la forme:

$$\sum_{(i_1, \dots, i_n) \in \mathfrak{D}_n} f_n(i_1, \dots, i_n) Y_{i_1} \dots Y_{i_n}$$

avec  $\mathfrak{D}_n = \{(k_1, \dots, k_n) \in \mathbb{N}^n : k_i \neq k_j, 1 \leq i < j \leq n\}$ ,  $f_n \in \ell^2(\mathbb{N})^{\otimes n}$  et

$$Y_i = \frac{q_n - p_n + X_i}{2\sqrt{p_i q_i}}.$$

Comme sur l'espace de Wiener, l'intégrale multiple stochastique  $J_n$  est une intégrale itérée:

$$J_n(f_n) = n \sum_{k=1}^{\infty} Y_k f_n(*, k) \mathbf{1}_{[0, k-1]^{n-1}}(*).$$

La décomposition en chaos de Walsh et les prémices du calcul de Malliavin pour des fonctionnelles de variables de Rademacher sont introduits dans Holden et al. (1992, 1993).



## Approximations probabilistes

Cette boîte à outils a servi originellement à obtenir un résultat d'hypoellipticité sur un opérateur différentiel écrit sous la forme de Hörmander. comprenant celle relative aux approximations probabilistes par la méthode de Stein. La méthode Stein est un ensemble de techniques qui permettent de quantifier l'erreur en approximation d'une loi en termes de distances de probabilités, par exemple la distance de Kolmogorov ou la distance de Kantorovich-Rubinstein (appelée aussi distance Wasserstein-1).

Sous l'impulsion de Nourdin et Peccati, le calcul de Malliavin a connu un renouveau avec l'application de ces outils et de l'intégration par parties à l'approximation probabiliste par la méthode Stein (Decreusefond, 2022, et références ci-après), parmi lesquels la formule de multiplication qui s'exprime dans le cas gaussien à l'aide d'opérateurs de contractions (Nourdin et Peccati (2009)):

$$I_p(f)I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}(f \tilde{\otimes}_r g), \quad (13)$$

où  $f \tilde{\otimes}_r g$  est le résultat de l'opérateur de contraction appliqué aux noyaux  $f$  et  $g$  d'ordre  $r$ .

Un résultat majeur issu de ce mariage de branches des probabilités est un théorème limite quantitatif, dit de *quatrième moment*, en termes de distance de Kolmogorov :

$$d_{Kol}(F, \mathcal{N}(0, 1)) \leq C \sqrt{|\mathbb{E}[F^4] - 3|}.$$

Ce résultat peut s'obtenir de façon plus directe en utilisant les opérateurs donné par le triplet de Markov associé à la structure de Malliavin sur l'espace de Wiener. En fait, des bornes supérieures du même type peuvent être obtenues pour d'autres lois cibles en utilisant la propriété de diffusion du triplet de Markov associé à cette loi, vue comme une loi invariante d'un certain processus de Markov (Azmoodeh et al., 2014).

Sans cette propriété de diffusion, le théorème de quatrième moment ne s'obtient pas directement par le schéma de preuve de Azmoodeh et al. (2014). Plus que la nature discrète, c'est l'absence de propriété de diffusion sur l'espace qui est un frein pour appliquer la méthode de Stein à partir des opérateurs de Malliavin. Récemment, il a été question d'étendre des résultats à d'autres chaos, sans la propriété de diffusion du générateur, notamment l'espace de Poisson (Peccati et al. (2010); Lachièze-Rey et Peccati (2013I); Lachièze-Rey et Peccati (2013II); Lachièze-Rey et Reitzner (2016); Döbler et Peccati (2018)). Plusieurs stratégies existent mais sont équivalentes. Elles consistent à établir une approximation de la propriété de diffusion (Reinert et al., 2010). Privault a prouvé à partir du calcul de variations stochastique pour la mesure de densité uniforme, mentionné plus haut, des bornes supérieures de Berry-Esseen pour des fonctionnelles de variables aléatoires indépendantes exprimées comme des fonctionnelles de variables aléatoires uniformes indépendantes (Privault et Serafin (2022)).

Comme sur l'espace de Poisson, nous avons choisi de recourir à un gradient "discret" qui semble plus naturel par rapport à la nature de l'espace même si c'est au prix d'un reste à contrôler dans la règle de dérivation.

## Calcul de Malliavin pour les variables aléatoires conditionnellement indépendantes

La première partie de ce travail porte sur l'implémentation de la méthode Malliavin-Stein pour des variables aléatoires conditionnellement indépendantes avec en vue une application à la quantification de théorèmes limites concernant les hypergraphes aléatoires.

Nous nous intéressons aux fonctionnelles  $F$  de séquence de variables aléatoires  $X = (X_a)_{a \in A}$  indépendantes conditionnellement à une variables aléatoire  $Z$ . Le gradient discret dans notre structure de Malliavin est défini pour  $a \in A$  par:

$$D_a F := F - \mathbb{E} [F | (X_b)_{b \in A \setminus \{a\}}, Z].$$

Il constitue avec l'opérateur de divergence  $\delta$  et l'opérateur  $L = \delta D$  la boîte à outils fournie par la structure de Malliavin. Comme dans les travaux de [Decreusefond et Halconruy \(2019\)](#), celle-ci découle de la *dynamique de Glauber* qui peut être vu comme un échantillonnage de Gibbs étalé dans le temps. Nous pouvons effectivement décrire la loi jointe d'une séquence de variables aléatoires indépendantes, comme la loi liée à la mesure invariante d'un tel processus de Markov. La modification de la dynamique comme présentée dans la première partie du manuscrit consiste à ajouter une version des résultats obtenus pour les séquences de variables aléatoires conditionnellement indépendantes. Le générateur infinitésimal associé correspond bien à l'opérateur de Malliavin  $L$  au sens de Stroock ([Stroock, 1981](#)). L'opérateur de *semi-groupe* permet obtenir une identité de covariance conditionnellement étant donné  $Z$  en utilisant le fait que  $L^{-1}F = \int_0^{+\infty} P_t F dt$ . Les conséquences de celle-ci sont des inégalités de concentration de fonctionnelles de variables aléatoires conditionnellement indépendantes.

Du fait de cette caractérisation supplémentaire en tant que mesure invariante, on obtient un outil supplémentaire qui complète la boîte à outils, venant de la structure de Dirichlet sous-jacente, à savoir l'opérateur carré du champ. La *formule d'intégration par parties* qui en découle est primordiale, à l'instar d'une règle de dérivation avec reste, pour dérouler la méthode de Stein :

$$\mathbb{E}[\Gamma(F, G)] = -\mathbb{E}[FLG]. \quad (14)$$

Nous obtenons par ailleurs une décomposition en chaos qui rappelle celle pour des fonctionnelles des variables aléatoires indépendantes ([Duerinckx, 2021](#)) qui elle-même rappelle celle sur l'espace Walsh. Même si les intégrales stochastiques multiples ne s'expriment pas sous forme d'intégrales itérées, elles ont un sens en étant des statistiques dégénérées d'ordre  $p$  pour  $p \geq 1$ . Nous obtenons des théorèmes limites quantitatifs pour des U-statistiques de variables conditionnellement indépendantes. Une première application est le théorème quantitatif du théorème centrale limite de Lyapunov conditionnel. Des théorèmes de quatrième moment qui prennent appui sur des articles fondateurs de calcul de Malliavin dans une *structure de Dirichlet* ([Azmoodeh et al., 2014](#)) peuvent être déduits avec reste du type:

$$d_W(F, \mathcal{N}(0, 1)) \leq C \sqrt{|\mathbb{E}[F^4] - 3\mathbb{E}[F^2]|} + \text{remainder},$$

permettant de fournir des vitesses de convergence d'approximation normale des statistiques d'hypergraphes aléatoires. Les problèmes statistiques pour les graphes ont leur équivalent concernant les hypergraphes aléatoires. Il est à noter que ces problèmes reviennent à étudier des U-statistiques. Dans le cas où elles sont fonctions de variables aléatoires indépendantes, les théorèmes limites sont bien connus ([Gine, 1997](#)).

Nous l'appliquons pour la preuve de la normalité asymptotique du comptage de motifs dans un hypergraphe aléatoire  $\mathbb{T}^{(3)}(n, q_n, p_n)$  où les triangles sont tirés indépendamment avec une certaine probabilité  $p_n$  conditionnellement à la présence des arêtes dans un graphe Erdős-Rényi  $\mathbb{G}(n, q_n)$  avec  $q_n < 1$ . L'hypergraphe aléatoire donné est différent de l'hypergraphe aléatoire Erdős-Rényi  $\mathbb{H}_n = \mathbb{G}^{(3)}(n, p'_n) = \mathbb{T}(n, 1, p'_n)$  ([Lovász, 2012](#), section 23.3) avec  $p'_n \in (0, 1)$ .

Le théorème limite est un résultat nouveau dans le sens où la structure conditionnellement indépendantes était un frein pour la déduction de vitesse de convergence. Concernant l'application

au comptage de motifs pour un hypergraphe, la structure consistant en une couche d'arêtes indépendantes, et une couche de triangles conditionnellement indépendantes permet de déduire un théorème central limite de la variables aléatoire centrée réduite par rapport à son espérance conditionnelle, et par rapport à son espérance. Si on tire les triangles avec probabilité  $p_n = 1$  et  $q_n = \frac{1}{2}$ , alors le modèle  $\mathbb{T}^{(3)}(n, q_n, p_n)$  et  $\mathbb{G}^{(3)}(n, 1/8)$  ont la même densité d'arêtes  $1/8$ , mais  $\mathbb{G}^{(3)}(n, 1/8)$  est quasi-aléatoire tandis que  $\mathbb{T}_n^{(3)}$  ne l'est pas. Il a une petite intersection avec chaque 3-uniforme hypergraphe quasi-aléatoire (Lovász, 2012, Proposition 23.10). Les deux modèles se ressemblent néanmoins par leur homogénéité, car finalement issu du même modèle de graphe Erdős-Rényi étendu différemment. Rien qu'avec cette famille de graphe aléatoire, on peut créer plusieurs sous-familles d'hypergraphes aléatoires, ce qui donne une latitude de modélisation et de résultats de convergence en loi pour des statistiques sur ces modèles.

Pour avoir une étude asymptotique complète selon les différents régimes possibles, il aurait fallu un résultat de convergence en loi vers une loi de Poisson pour un régime parcimonieux comme pour les graphes aléatoires, mais la stratégie d'approximation poissonnienne en utilisant le gradient s'est révélée infructueuse sur l'exemple basique de la convergence de sommes de variables aléatoires suivant une loi de Bernoulli de paramètre  $\lambda/N$  pour  $N \rightarrow \infty$ . Nous nous référons aux résultats de Barbour (1982) sur l'approximation poissonnienne appliquée notamment à des statistiques dans des graphes aléatoires. Ces résultats sont étendus par Coulson et al. (2016) pour des statistiques dans des modèles aléatoires générés par des graphons qui sont des U-statistiques de variables aléatoires conditionnellement indépendantes.

Quant à l'extension des résultats de normalité asymptotique au cas d'un vecteur de fonctionnelles à valeurs dans  $\mathbb{R}^d$  pour  $d > 1$ , elle serait basée sur des techniques éprouvées (Nourdin et al., 2010). De nombreux papiers explorent ce type d'extension ainsi que le cas où la variable aléatoire  $F$  est un processus indexé sur  $\mathbb{R}^+$  (Döbler et al., 2022). Étant donné la richesse de la littérature relatif à ce sujet, nous avons étudié d'autres pistes de contributions.

## Inversibilité dans le contexte de fonctionnelles de mesure de Poisson

Autre que le calcul Malliavin, des résultats connexes sur l'espace de Wiener peuvent être étendus sur l'espace de Poisson (Bismut, 1983; Bichteler et al., 1987). Bismut a proposé une approche du calcul de Malliavin qui exploite le théorème de Girsanov et des flots, qui sont des semi-groupes déterministes sur l'espace de Wiener pour obtenir une formule d'intégration par parties "dirigée". Son approche est une alternative de celle de Malliavin-Stroock au calcul de variations, utilisant la propriété de quasi-invariance de la mesure de probabilité comparable à la formule de Girsanov. Une présentation simple des deux points de vue est exposée par Williams (1981). L'extension développée par Bismut concerne plus généralement l'espace de fonctions continues à droite et limites à gauche pour les processus à sauts à valeurs dans  $\mathbb{R}^n$ . La notion de transformée de Girsanov est la base de son approche afin de définir d'un calcul de variations qui permet notamment de déduire la représentation martingale, avec des applications à la perturbation d'équations différentielles stochastiques. A un changement de loi absolument continu, on associe une transformation  $\Gamma$  de l'élément canonique  $\omega$  qui compense celle-ci, c'est-à-dire tel que sous la nouvelle mesure,  $\Gamma(\omega)$  ait la même loi que  $\omega$  sous la mesure de référence. A partir de cette transformation donnée par la théorème de Girsanov, on étudie les règles de compositions de telles transformations des trajectoires  $\omega$  (voir l'approche de Bismut décrite dans Bichteler et al. (1987)). De plus, comme sur l'espace de Wiener, la dérive de Girsanov est liée à une entropie par une relation commode.

Sur l'espace de Wiener, les premiers travaux remontent aux années 80 (Föllmer (1986); Föllmer et Wakolbinger (1986)) puis les travaux d'Üstünel et Zakai sur le critère d'inversibilité des dérivées de Girsanov basé sur l'entropie. Üstünel fait la remarque que Theorem 3.5.3 de Üstünel et Zakai (2000) que la propriété de quasi-invariance est valide pour des perturbations non nécessairement adaptées perturbations de l'identité. Néanmoins, nous travaillons tout au long avec des transformations adaptées, en opposition à des transformations anticipatives. La seconde partie du mémoire s'intéresse à la notion d'inversibilité stochastique sur l'espace de Poisson qui reprend les principes de l'inversibilité stochastique sur l'espace de Wiener (Lassalle, 2012). C'est aussi la suite des travaux de Coutin et Decreusefond (2023). En particulier, on donnera un résultat de critère entropique pour l'inversibilité qui est lié à la représentation variationnelle de l'entropie sur l'espace de Poisson (Zhang, 2009). Un morphisme d'espaces de probabilité vers l'espace de Poisson, qui est de plus adapté, peut être associé canoniquement à certaines lois d'équations différentielles stochastiques, si inversible. Le cadre d'inversibilité stochastique donne lieu à un nouveau point de vue sur les solutions d'équations différentielles.

## Manuscrit

Cette thèse se compose de deux parties distinctes indépendantes. La première contribution est une collaboration avec Laurent Decreusefond, et va être soumis à *Stochastic Processes with Applications*. La deuxième sur l'inversibilité de fonctionnelles de mesures de Poisson est en cours. Certaines notations de la partie I désignent des objets différents dans la partie II. La notation  $A$  désigne un ensemble discret d'indexation de variables aléatoires tandis que dans la seconde partie elle correspond à un borélien de l'espace de marques.  $Z$  désigne une variable aléatoire latente dans la première partie tandis que dans la deuxième partie, la notation  $Z_n$  désigne la marque au  $n$ -ème saut d'un processus ponctuel marqué.  $X$  est une suite de variables aléatoires indexée par  $A$  tandis que dans la seconde partie,  $X$  désigne le processus de comptage de sauts associé à un processus ponctuel marqué. Dans la première partie,  $\Gamma$  est l'opérateur carré du champ associé à une dynamique de Glauber tandis que dans la seconde partie, la notation désigne une transformation de trajectoires sur l'espace de configuration. Quant à la notation  $\rho$ , dans la première partie, elle désigne une quantité d'influence maximale sur la variance d'une variable aléatoire fonctionnelle de suites de variables aléatoires, alors que dans la seconde partie c'est une densité de Radon-Nikodym par rapport à la mesure de Lebesgue. Dans la première partie, la notation  $x \sim \mu$  signifie que  $x$  est tiré selon la mesure de probabilité  $\mu$  tandis que dans la seconde partie, elle désigne l'équivalence entre deux mesures de probabilité.

Je donne des dernières indications pour la lecture du manuscrit. J'ai pris la liberté de mettre en valeur les résultats nouveaux avec des boîtes colorées, en espérant une lecture plus aisée.



## Part I

# Stein-Malliavin-Dirichlet method and applications to statistics in hypergraph theory



# Chapter 1

## Stein-Dirichlet-Malliavin method

This first chapter is dedicated to a brief introduction of the developments around Malliavin-Stein's method which leverages a Dirichlet structure and its application to non-diffusive structures, including our contribution to the construction of a Malliavin structure for conditionally independent random variables.

### 1.1 Background

The fundamental example of the type of result we address in the first part is the Berry-Esseen bound for the central limit theorem. Let  $X_1, \dots, X_n$  be i.i.d. random variables with  $\mathbb{E}|X_1|^3 < \infty$ ,  $\mathbb{E}[X_1] = 0$  and  $\text{Var}(X_1) = 1$ , if  $\Phi$  denotes the cumulative distribution function of a standard normal distribution and  $W_n = \sum_{i=1}^n \frac{X_i}{\sqrt{n}}$ , then there exists  $C > 0$  such that:

$$|\mathbb{P}(W_n \leq x) - \Phi(x)| \leq \frac{C \mathbb{E}|X_1|^3}{\sqrt{n}}.$$

That quantifies the error in the classical central limit theorem. More generally, a central theme of probability theory is proving distributional limit theorems. For the purpose of approximation it is of interest to estimate the rate of convergence in such results. The common techniques employed are method of moments, Fourier analysis and martingale theory, but in some important cases they lead to suboptimal bound on the rate.

Stein's method is a technique that can quantify the error in the approximation of one distribution by another in a variety of metrics. The Stein's method for normal approximation was invented in the groundbreaking paper [Stein \(1972\)](#) published in 1972 and extended for Poisson convergence ([Chen, 1975](#)) a few years later. Stein's method has proved powerful in particular for deriving explicit sharp bounds on distributional distances even when the underlying random element consists of structures with dependence (see for example [Stein, 1986](#); [Goldstein and Reinert, 1997](#); [Arratia et al., 2019](#); [Arras et al., 2020](#)). Moreover, it has been developed by a growing community to tackle an enlarging collection of approximation problems including beta, binomial, gamma, multinomial, variance-gamma, Wishart, and many more.

The common denominator of most of them is the existence of a generator whose invariant measure is the target distribution. The generator approach was a novel approach to Stein's method introduced in [Barbour \(1990\)](#) (see also [Barbour \(1988\)](#); [Götze \(1991\)](#)). Since then, it has been a common scheme for approximation. That is one of the first building blocks of the Malliavin-Stein method in which the semigroup associated to the generator plays a key role.



The powerful interactions of Malliavin calculus of variations and Stein's method were highlighted by Nourdin and Peccati. The seminal paper [Nourdin and Peccati \(2009\)](#) used the combination of Malliavin calculus and Stein's approach to obtain a rather simple proof of the striking fourth moment theorem for normal approximation, established earlier in [Nualart and Peccati \(2005\)](#); [Peccati and Tudor \(2005\)](#). Hence, that line of search has been popular as the fourth moment phenomenon emerges in many contexts including free probability, compressed sensing, time series analysis, stochastic geometry and motif estimation in random graphs.

The latest developments in computational statistics are summarized in [Anastasiou et al. \(2023\)](#). We also point to a website that regroups the papers using Stein's method over the years:

<https://sites.google.com/site/steinsmethod/articles>.

In this manuscript, we deal with a branch of Stein's method combined with Malliavin calculus enriched by the structure given by Dirichlet forms, honing on the application to fourth moment limit theorems.

## 1.2 Probability distances

For the taxonomy of probability metrics and their history, the essential reference is the monograph ([Rachev, 1991](#)) that presents a unified and comprehensive approach to the theory of the more than seventy known probability metrics and their applications. Many of them are variants, particular cases or extensions of the Wasserstein and Lévy-Prokhorov metrics. We ponder on the first type of probability distances. For two probability measures  $P$  and  $Q$  on a measurable space  $(\mathcal{X}, \mathcal{A}, \mathbb{P})$ , the probability metrics we consider have the form:

$$d(P, Q) = \sup_{h \in \mathcal{H}} \left| \int h \, dP - \int h \, dQ \right|, \quad (1.1)$$

where  $\mathcal{H}$  is a space of test functions  $f : \mathcal{X} \rightarrow \mathbb{R}$ . In the remainder of this chapter, we assume  $\mathcal{X} = \mathbb{R}^d$  and  $\mathcal{A} = \mathcal{B}(\mathbb{R}^d)$ .

**Example 1.2.1.** For the Kolmogorov distance, the space is:

$$\mathcal{H} = \{\mathbb{1}_{(-\infty, x)} : x \in \mathbb{R}\}.$$

The Kolmogorov metric, denoted  $d_K$ , is the maximum distance between cumulative distribution functions, so a sequence of distributions converging to a fixed distribution in this metric implies weak convergence, although the converse is not true since weak convergence only implies pointwise convergence of cumulative distribution function (c.d.f. for short) at continuity points of the target c.d.f..

**Example 1.2.2.** For the total variation metric denoted  $d_{TV}$ , the space of test functions is:  $\mathcal{H} = \{\mathbb{1}_A : A \in \mathcal{A}\}$ . We use it for approximation by discrete distributions, for example the Poisson distribution.

**Example 1.2.3.** The Wasserstein distance is defined as the minimum cost of transporting  $P$  to  $Q$ , where the cost is measured by the distance that each mass unit must be moved. They are common metrics, occurring in many contexts as optimal transport, partial differential equations and even surprising ones as quantification of some variants of the uncertainty principle in quantum physics. Mathematically, the Wasserstein distance is defined as follows:

$$d_{W_p}(P, Q) = \inf_{\gamma \in \Gamma(P, Q)} \int_{M \times M} c(x, y) \gamma(dx, dy)$$

where  $\Gamma(P, Q)$  is the set of all couplings between  $P$  and  $Q$  and  $c$  is a cost function. In the literature, the  $p$ -Wasserstein distance refers to the cases where the cost functions are power of Euclidean distances, i.e.  $c(x, y) = \|x - y\|^p$ . By the Kantorovich-Rubinstein duality formula, the first-order Wasserstein distance can be written in the form (1.1) where the space of test functions is a subspace of Lipschitz-continuous functions:  $\mathcal{H} = \{f \in \text{Lip}(\mathcal{X}, \mathbb{R}) : \text{Lip}(f) = 1\}$ . In the sequel, we refer to it as the 1-Wasserstein distance with the notation  $d_W$ , also called in the literature Kantorovich-Rubinstein distance (or norm).

**Proposition 1.2.4** *Let  $\mu$  be absolutely continuous with respect to the Lebesgue measure  $\lambda$ . If the Radon-Nikodym derivative  $\frac{d\mu}{d\lambda}$  is bounded by  $C$ , then for any random variable  $W$ ,*

$$d_K(\mathcal{L}_W, \mu) \leq \sqrt{2C d_W(\mathcal{L}_W, \mu)}.$$

In the remainder,  $d_W$  is the main distance for our distributional approximations. The following theorem illustrates the interest in considering that distance between each element of a family of probability measures  $(\mu_n)_{n \in \mathbb{N}}$  and a target probability measure  $\mu$ .

**Theorem 1.2.5** *Let  $n \in \mathbb{N}$ , the following two properties are equivalent:*

1.  $d_W(\mu_n, \mu) \rightarrow 0$ .
2.  $\int_{\mathcal{X}} f d\mu_n \rightarrow \int_{\mathcal{X}} f d\mu$  for all bounded and continuous functions from  $\mathcal{X}$  to  $\mathbb{R}$ .

In our case,  $\mu_n$  is the law of functionals  $F_n$  of sequences of random variables, and  $\mu$  is the Gaussian distribution.

### 1.3 Stein's method principle

The Stein's method relies on the characterization of the target distribution in the scheme of approximation. It consists in the construction of so-called Stein identities or equivalently in our case Stein operators which act on a space of functions  $\mathcal{H}_*$  not to confuse with the starting space of test functions.

**Definition 1.3.1.** *The Stein operator  $L$  is defined such that for a given random variable  $Y$ ,*

$$\mathbb{E}[Lf(Y)] = 0 \quad \forall f \in \mathcal{H}_* \iff Y \text{ has distribution } Q. \quad (1.2)$$

This equation is called Stein equation, although it may take a different form in other papers. The first component in Stein's method is to transform the expression of the quantity

$|\int h dP - \int h dQ|$  of which we take the supremum over  $\mathcal{H}$  in (1.1) into a Stein identity of the form  $\mathbb{E}[L\varphi_h(Y)]$  for  $\varphi_h \in \mathcal{H}_*$ .

For the following, we fix  $\mathcal{H} = \{f \in \text{Lip}(\mathcal{X}, \mathbb{R}) : \text{Lip}(f) = 1\}$ . We will explain one way to obtain a Stein operator via the generator approach, using the formalism of Decreusefond (2015) which uses the underlying Dirichlet structure.

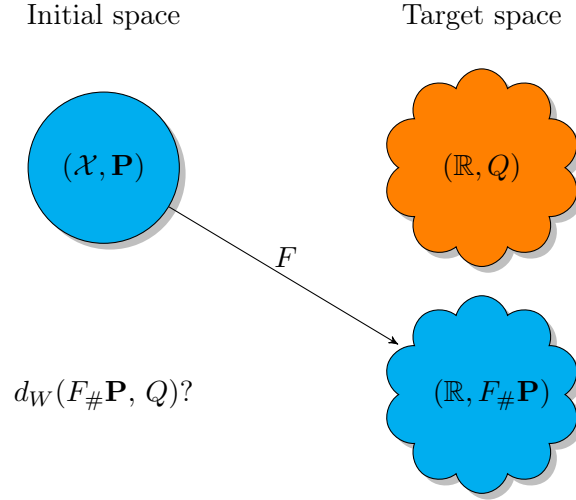


Figure 1.1 – Comparison between  $F_{\#}\mathbf{P}$  and  $Q$ .

The second component consists of bounding the expectation obtained in the first step. The earlier collection of techniques is related to the so-called auxiliary randomization which requires various strategies depending on the problem at hand. In a way, the resort to Dirichlet structures gives a systematic procedure for the first step, but also for the second step. Let  $E$  a metric space equipped with its  $\sigma$ -field  $\mathcal{A}$ , we recall some definitions.

**Definition 1.3.2** (Semigroup). Let  $(P_t)_{t \in \mathbb{R}^+}$  the family of operators defined on some set of real-valued measurable functions on  $(E, \mathcal{A})$ , and satisfying the following conditions:

1. For any  $t \in \mathbb{R}^+$ ,  $P_t$  is a linear operator sending bounded measurable functions on  $(E, \mathcal{A})$  to bounded measurable real functions.
2.  $P_0 = \text{Id}$  where  $\text{Id}$  is the identity operator (initial condition).
3. For every  $(s, t) \in (\mathbb{R}^+)^2$ ,  $P_{t+s} = P_t \circ P_s$ .
4. For any  $t \in \mathbb{R}^+$ ,  $P_t$  conserves the mass and preserves positivity (Markov property), i.e.  $P_t \mathbf{1} = \mathbf{1}$ . For any positive function  $f$ ,  $P_t f \geq 0$ .

**Definition 1.3.3.** A *Dirichlet structure* on  $(E, \mathcal{A}, \mu)$  is a quadruple  $(X^\circ, L^\circ, (P_t^\circ, t \geq 0), \mathcal{E}^\circ)$ , where  $X^\circ$  is a strong Feller process with values in  $E$  whose generator is  $L^\circ$ , and its semigroup is  $(P_t^\circ, t \geq 0)$  for  $f : E \rightarrow \mathbb{R}$  sufficiently regular such that:

$$\begin{aligned} P_t^\circ f(x) &= \mathbb{E}[f(X^\circ(t)) | X^\circ(0) = x] \\ \frac{d}{dt} P_t^\circ f(x) &= L^\circ P_t^\circ f(x) \\ &= P_t^\circ L^\circ f(x). \end{aligned}$$

Actually, we only need to define a Markov triple since the bilinear form  $\mathcal{E}^\circ$  is expressed as the expectation of the *carré du champ operator* defined for  $g$  sufficiently regular by:

$$\Gamma^\circ(f, g) = \frac{1}{2} (L^\circ(fg) - fL^\circ g - gL^\circ f)$$

Then, for  $X$  a random variable on  $(E, \mathcal{A}, \mu)$ ,  $f, g : E \rightarrow \mathbb{R}$  sufficiently regular, we have the *integration by parts formula*:

$$\mathbb{E}[L^\circ(fg)(X)] = \mathbb{E}[f(X)L^\circ g(X)] = -\mathbb{E}[\Gamma^\circ(f(X), g(X))]. \quad (1.3)$$

**Remark 1.3.4.** It must be noted that the knowledge of one of  $L^\circ$ ,  $P^\circ$  or  $X^\circ$  is equivalent to the knowledge of the other two. ■

**Definition 1.3.5** (Invariant measure). Let a family  $(P_t)_{t \in \mathbb{R}^+}$  of operators defined on  $(E, \mathcal{A})$  and satisfying all the properties above. A positive  $\sigma$ -finite measure  $\nu_0$  on  $(E, \mathcal{A})$  is said to be *invariant* for  $(P_t)_{t \in \mathbb{R}^+}$ , if for every bounded function  $f : E \rightarrow \mathbb{R}$ , and  $t \in \mathbb{R}^+$ ,

$$\int_E P_t f \, d\nu_0 = \int_E f \, d\nu_0. \quad (1.4)$$

In the following, we only consider semigroups that have invariant measure. We use that property in order to show the Stein's identity for normal approximation.

**Definition 1.3.6.** Let  $X^\dagger$  the Markov process defined as:

$$X^\dagger(t, x) = \begin{cases} X^\dagger(0, x) = x \\ dX^\dagger(t, x) = -X^\dagger(t, x) + \sqrt{2} \, dB(t), \end{cases} \quad (1.5)$$

where  $B$  is a standard Brownian motion. The solution to this stochastic differential equation is the Ornstein-Uhlenbeck process.

By the Itô formula, we have for  $t \in \mathbb{R}^+$  and  $x \in \mathcal{X}$ :

$$X^\dagger(t, x) = xe^{-t} + \sqrt{2} \int_0^t e^{-(t-s)} \, dB(s).$$

$X^\dagger(\cdot, x)$  is then a Gaussian process with parameters:

$$m(t) = xe^{-t} \quad K(s, t) = e^{-|t-s|} - e^{-(t+s)}. \quad (1.6)$$

Hence,  $X^\dagger(t, x) \sim \mathcal{N}(xe^{-t}, 1 - e^{-2t})$ . We have that  $X^\dagger(t, x) \implies \mathcal{N}(0, 1)$  when  $t \rightarrow \infty$ . The invariant distribution is  $Q = \mathcal{N}(0, 1)$ .

Let  $f \in \mathcal{S}$  the set of functions belonging to  $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$  such that  $f$  and all its derivatives have at most polynomial growth, and  $x \in \mathbb{R}$ . By the Mehler representation formula, the associated semigroup is defined as:

$$P_t^\dagger f(x) = \int_{\mathbb{R}} f\left(e^{-t}x + \sqrt{1 - e^{-2t}}y\right) dQ(y). \quad (1.7)$$

By the dominated convergence theorem, for  $f \in L^1(\mathbb{R} \rightarrow \mathbb{R}; Q)$ ,

$$P_t^\dagger f(x) \xrightarrow{t \rightarrow \infty} \int f(y) dQ(y). \quad (1.8)$$

Hence, we get the first component:

$$\begin{aligned} \int h(x) dP(x) - \int h(x) dQ(x) &= \int h(x) dP(x) - \int \int h(y) d\nu_0(y) dP(x) \\ &= \int P_0^\dagger h(x) dP(x) - \int \lim_{t \rightarrow \infty} P_t^\circ h(x) dP(x) \\ &= \int \int_0^{+\infty} \bar{L}^\dagger P_s^\dagger h(x) ds dP(x) \\ &= \mathbb{E}_{Y \sim P} \left[ L^\dagger \left( \int_0^{+\infty} P_s^\dagger h(Y) ds \right) \right], \end{aligned}$$

where  $\bar{L}^\dagger$  is the infinitesimal generator associated to  $P^\dagger$ .

For any  $f \in \mathcal{S}$ ,

$$\begin{aligned} \frac{d}{dt} P_t^\dagger f(x) &= -xe^{-t} \int f'(e^{-t}x + \sqrt{1 - e^{-2t}}y) dQ(y) \\ &\quad + \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} \int f'(e^{-t}x + \sqrt{1 - e^{-2t}}y)y dQ(y) \\ &= -xe^{-t} \int f'(e^{-t}x + \sqrt{1 - e^{-2t}}y) dQ(y) + e^{-2t} \int f''(e^{-t}x + \sqrt{1 - e^{-2t}}y) dQ(y), \end{aligned}$$

using the Itô formula. In particular,  $\bar{L}^\dagger f(x) = \frac{d}{dt} P_t^\dagger f(x)|_{t=0} = -xf'(x) + f''(x)$ . Thus,

$$d_W(P, Q) = \sup_{h^\dagger: h \in \mathcal{H}} \left| \mathbb{E}_{Y \sim P} \left[ \bar{L}^\dagger \bar{h}^\dagger(Y) \right] \right|, \quad (1.9)$$

with

$$\bar{h}^\dagger : x \mapsto \int_0^{+\infty} P_s^\dagger h(x) ds.$$

**Remark 1.3.7.** As long as one can find a semigroup with invariant measure, the procedure for the first component remains the same.  $\blacksquare$

We can characterize the space  $\mathcal{H}_*$  (Nourdin and Peccati, 2012, Chapter 3). As (1.9) is not the usual expression, we rewrite it as:

$$d(P, \mathcal{N}(0, 1)) \leq \sup_{\varphi_h \in \mathcal{H}_*} |\mathbb{E}_{Y \sim P}[L^\dagger \varphi_h(Y)]|, \quad (1.10)$$

with  $\varphi_h = h^\dagger$  being the derivative of  $\bar{h}^\dagger$  and  $L^\dagger h(x) = h'(x) - xh(x)$ . Hence, we have

- for the total variation,  $\mathcal{H}_* = \{h^\dagger : \|h^\dagger\|_\infty \leq \sqrt{\pi/2}, \|(h^\dagger)'\|_\infty \leq 2\}$ ;
- for the Kolmogorov distance,  $\mathcal{H}_* = \{h^\dagger : \|h^\dagger\|_\infty \leq \sqrt{2\pi}/4, \|(h^\dagger)'\|_\infty \leq 1\}$ ;
- for the 1-Wasserstein distance,  $\mathcal{H}_* = \{h^\dagger : \|(h^\dagger)'\|_\infty \leq \sqrt{2/\pi}\}$ .

The first step of Stein's method relies on the characterization of the target distribution. The second step boils down to bound (1.10) with target the normal distribution. There are various approaches to achieve that, using the structures in the construction of random variables, among them the method of exchangeable pairs (Stein, 1986), the dependency graph method (Barbour et al., 1989; Arratia et al., 1989, 1990), size-bias coupling (Goldstein and Rinott, 1996) and zero-bias coupling (Goldstein and Reinert, 1997). In the next section, we give a glimpse of the Malliavin-Stein's method, which is a ramification of Stein's method with several breakthroughs concerning normal approximation (see e.g. Nourdin and Peccati, 2012; Bourguin and Peccati, 2016) through a proof of a fourth moment theorem.

## 1.4 Quantitative fourth moment theorems

### 1.4.1 Origins

Malliavin calculus is also known as the stochastic calculus of variations. At the very core of it, it considers a gradient on a measured space. The link between these, the differential geometry and the measure is made through the so-called integration by parts formula. When the measured space is the Wiener space, i.e. the set of continuous functions with the Brownian measure, the gradient generalizes the usual gradient on  $\mathbb{R}^N$  and the integration by parts yields an extension of the Itô integral. The intersection with Stein's method originates from the pivotal paper by Nourdin and Peccati, who were able to associate a quantitative bound to the remarkable fourth moment theorem on the Wiener space established by Nualart and Peccati (2005).

Fourth moment theorems are simplifications of results using the Method of Moments. The Method of Moments in probability theory is one of the oldest versatile tool used by probabilists to prove limit theorems in non-standard problems. The principle stems from the property that the moments of a random variable are determined by the distribution. Although the converse is not true, we have that the distribution of some random variable  $X$  is determined by its moments if  $X$  has finite moments and every random variable with the same moments as  $X$  has the same distribution. The standard version of the method of moments can be stated as follows (Chung, 1974, Theorem 4.5.5).

**Theorem 1.4.1** *Let  $Y$  be a random variable with a distribution that is determined by its moments. If  $X_1, X_2, \dots$  are random variables with finite moments such that  $\mathbb{E}[X_k^n] \rightarrow \mathbb{E}[Y^k]$  as  $n \rightarrow \infty$  for every integer  $k \geq 1$ , then  $X_n \xrightarrow{d} Y$ .*

It has been used for results of asymptotic normality of statistics on random graphs (see Janson et al. (2000, Chapter 6)), and Poisson convergence as well, i.e. when the distribution of the

random variable  $Z$  is the standard Gaussian distribution  $\mathcal{N}(0, 1)$  and the Poisson distribution  $\text{Pois}(\lambda)$  with  $\lambda > 0$ . Overall, it is well adapted to combinatorial problems. The drawback is that it usually requires tedious calculations for the estimations. In some cases, the proofs can be reduced to the control of some moments and not all of them. This is the case of the fourth limit theorem.

**Definition 1.4.2** (Cumulants). Let  $F$  be a real-valued random variable such that  $\mathbb{E}[|F|^m] < \infty$  for some integer  $m \geq 1$ , and write  $\varphi_F(t) = \mathbb{E}[e^{itF}]$ ,  $t \in \mathbb{R}$ , for the characteristic function of  $F$ . Then, for  $r = 1, \dots, m$ , the  $r$ -th cumulant of  $F$  denoted by  $\kappa_r(F)$ , is given by

$$\kappa_r(F) = (-i)^r \frac{d^r}{dt^r} \log \varphi_F(t). \quad (1.11)$$

**Remark 1.4.3.** When  $\mathbb{E}[F] = 0$ , then the first four cumulants of  $F$  are the following:  $\kappa_1(F) = \mathbb{E}[F] = 0$ ,  $\kappa_2(F) = \mathbb{E}[F^2] = \text{Var}(F)$ ,  $\kappa_3(F) = \mathbb{E}[F^3]$ , and:

$$\kappa_4(F) = \mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2.$$

■

The fourth moment theorem states a rate of convergence in terms of cumulants for a sequence of multiple stochastic Wiener-Itô integrals of fixed order  $I_q(f)$  for  $q \geq 2$ .

**Theorem 1.4.4** Let  $(F_n)_{n \geq 1} = (I_q(f_n))_{n \geq 1}$  be a sequence of random elements in a fixed Wiener chaos of order  $q \geq 2$  such that  $\mathbb{E}[F_n^2] = q \|f_n\|^2 = 1$ . Then, as  $n$  tends to infinity, the following assertions are equivalent.

1.  $F_n \rightarrow \mathcal{N}(0, 1)$  in distribution.
2.  $\kappa_4(F_n) \rightarrow 0 = \kappa_4(N)$  where  $N \sim \mathcal{N}(0, 1)$ .

A few years later, Nourdin and Peccati made a huge breakthrough by providing an error bound in terms of the fourth moment (Nourdin and Peccati, 2009). The fourth moment phenomenon has since then emerged as a unifying principle governing the central limit theorems for various non-linear functionals of random fields. It has paved the way for a fruitful line of search, growing as far as obtaining the *universality phenomenon* according to which the asymptotic behavior of large random systems does not depend on the distribution of its components (see the comprehensive monograph Nourdin and Peccati, 2012). More precisely, one can use random variables in Wiener chaos instead of independent random variables in order to prove asymptotic normality of functionals of those provided some mild assumptions on the functionals (Nourdin et al., 2010, 2016; Simone, 2017). The intuition behind also led to a connection of the Malliavin-Stein's method and criterions of asymptotic normality in the literature such as De Jong's one (De Jong, 1989) which states a partial fourth moment theorem concerning multilinear forms, a particular class of functionals of independent random variables. The additional remainder in the convergence rate is expressed in terms of the maximal influence of the variables, roughly speaking the maximum over  $i \in \mathbb{N}$  contribution of a random variable  $X_i$  to the overall configuration of the multilinear forms. Due to the discrete nature, the techniques used for the Wiener chaos do not apply in a straightforward way, although they can be adapted. It is the object of the remainder and of my contribution detailed in the next chapter.

### 1.4.2 The Markov triple approach

The approach of Nourdin and Peccati entails the use of the so-called product formula for Wiener integrals. The use of this formula makes the proofs rather involved since it relies on subtle combinatorial arguments. Starting with the work of Ledoux (Ledoux, 2012), a new approach to the fourth moment theorem was developed by Azmoodeh, Campese and Poly (Azmoodeh et al., 2014) leveraging Dirichlet forms and adapted to normal approximation (Chen and Poly, 2015). An advantage of this new approach is that it provides a simpler proof of the theorem by avoiding completely the use of the product formula. The main assumption is the property of *diffusion* of the generator  $L$  on the initial space.

**Definition 1.4.5.** A Markov operator  $L$  is said to be a *diffusive*, if for every function  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ , and for every smooth enough function  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ , and for every function  $f \in \mathcal{H}$ ,

$$L\varphi(f) = \varphi'(f)Lf + \varphi''(f)\Gamma(f, f). \quad (1.12)$$

Equivalently, the associated carré du champ operator  $\Gamma$  is a derivation in the sense that:  $\Gamma(\varphi(X), X) = \varphi'(X)\Gamma(X, X)$  for any  $\varphi \in \mathcal{C}^\infty(\mathbb{R})$ .

That is a consequence of the chain rule, which leads to:

$$\Gamma(\psi(f), \theta(g)) = \psi'(f)\theta'(g)\Gamma(f, g). \quad (1.13)$$

**Definition 1.4.6.** A fourth moment structure is a triple  $(E, \mu, L)$  such that:

1.  $(E, \mu)$  is a probability space;
2.  $L$  is a symmetric unbounded operator defined on some dense subset of  $L^2(E, \mu)$  that we denote  $\text{Dom } L$ , the domain of  $L$ .
3. The Markov operator  $L$  is diffusive.

That holds in the following diffusion structures: Ornstein-Uhlenbeck, Laguerre and Jacobi. For those structures,  $L$  is diagonalizable with respect to an orthonormal basis. We assume that there exist  $(\lambda_k)_{k \geq 0}$  such that  $0 = \lambda_0 < \lambda_1, \lambda_2, \dots$  and:

$$L^2(\mathbb{R} \rightarrow \mathbb{R}; \mu) = \bigoplus_{k=0}^{+\infty} \ker(L + \lambda_k \text{Id}). \quad (1.14)$$

To illustrate the process of Malliavin-Stein's method from end to end, we suppose that the target distribution is the normal distribution.

**Theorem 1.4.7** *Let  $V \in L^2(E \rightarrow \mathbb{R}, \mu)$ , such that  $\mathbb{E}[V] = 0$  and  $\mathbb{E}[V^2] = 1$ . Then,*

$$d_W(\mathcal{L}(V), \mathcal{N}(0, 1)) \leq \sqrt{\frac{2}{\pi}} |\mathbb{E}[\Gamma(L^{-1}V, V) + 1]|. \quad (1.15)$$

*Proof.* We have to estimate:



$$\begin{aligned}
\sup_{f^\dagger: f \in \text{Lip}_1} \mathbb{E}[L^\dagger f^\dagger] &= \mathbb{E}[V f^\dagger(V) - (f^\dagger)'(V)] \\
&= \mathbb{E}[LL^{-1}V(f^\dagger)'(V)] - \mathbb{E}[(f^\dagger)''(V)] \text{ since } LL^{-1} = \text{Id} \\
&= -\mathbb{E}[\Gamma(L^{-1}V, f^\dagger(V))] - \mathbb{E}[(f^\dagger)'(V)] \text{ by (1.3)} \\
&= \mathbb{E}[(f^\dagger)'(V)\Gamma(L^{-1}V, V) - 1] \text{ by (1.13)}.
\end{aligned}$$

The result follows from (1.10).  $\square$

In the proof of this theorem, the integration by parts entailed by the Dirichlet structure was used. The *chaos decomposition* is one of the feature of Malliavin calculus that allows to compute  $L^{-1}V$ .

**Corollary 1.4.8** *If  $V \in \ker(L + \lambda_p \text{Id})$ ,*

$$d_W(\mathcal{L}(V), \mathcal{N}(0, 1)) \leq \frac{1}{\lambda_p} \sqrt{\frac{2}{\pi}} |\mathbb{E}[\Gamma(V, V) - \lambda_p]|. \quad (1.16)$$

**Definition 1.4.9** (Azmoodeh et al. (2014) p.6). An eigenfunction  $X$  of the generator  $-L$  with eigenvalue  $\lambda_p$  is called a *chaos eigenfunction of order  $p$* , if and only if:

$$X^2 \in \bigoplus_{k=0}^{2p} \ker(L + \lambda_k \text{Id}). \quad (1.17)$$

This property allows to estimate the variance of  $\Gamma(V, V)$ .

**Theorem 1.4.10** *For  $V \in \ker(L + \lambda_p \text{Id})$ ,*

$$\mathbb{E}[(\Gamma(V, V) - \lambda_p)^2] \leq \frac{\lambda_p^2}{3} (\mathbb{E}[V^4] - 6\mathbb{E}[V^2] + 3).$$

That theorem leads to a quantitative version of [Theorem 1.4.4](#).

### 1.4.3 Adaptations to non-diffusive structures

The control by the fourth cumulant in the Wiener chaos and the Poisson space of the quality of normal approximation is now well-understood. As stressed in the previous section in the context of diffusive structures, it is obtained in a straightforward way provided some mild assumptions. In the Poisson space, the generator of the corresponding Markov process is not diffusive. However, there is a rule of approximation of the derivation of the associated carré du champ operator where the residual terms, expressed in terms of add-one cost operators (Döbler and Peccati, 2018) can be bounded by a term proportional to the fourth cumulant. This paper refined a result found earlier (Lachièze-Rey and Peccati, 2013) which derived fourth moment bounds for positive kernels in the chaos expansion of a given random variable. The authors Lachièze-Rey and Peccati implemented the usual integration by parts formula with the Malliavin operators  $D$ ,  $\delta$  and  $L$  on the Poisson space which led to bounds in terms of contraction operators which turn out to be related to fourth cumulants. See also Eichelsbacher and Thäle

(2014); Schulte (2016) for normal approximation in Kolmogorov distance of Poisson functionals. The next natural random elements to deal with are sequences of independent random variables. Before being dealt in full generality, the case of interest was the Rademacher random variables (Privault, 2008, and references therein). Employing an approximate chain rule, Reinert, Nourdin, and Peccati derived a Stein bound for functionals in a fixed Walsh chaos in terms of norms of contraction operators. Those operators are used in many papers about discrete Malliavin calculus combined with Stein's method. Their probabilistic approximations are also based on multiplication formula for discrete multiple stochastic integrals (see also Proposition 5.1 and Section 9 of Privault and Torrisi, 2015).

That notion was tackled in full generality (Decreusefond and Halconruy, 2019). but without fourth moment-type of theorem. One of the difficulties is to use the appropriate approximate chain rule. For an exhaustive overview of the Malliavin-Stein method, we refer the reader to the website:

<https://sites.google.com/site/malliavinstein/home>.

## 1.5 Contributions

There are limit theorems for subhypergraph counts in random hypergraphs (De Jong, 1996). However, most of them consider simply random hypergraphs as extended random graphs, hence considering hyperedges as independent random variables. In Austin (2008), the taxonomy of exchangeable random hypergraphs concerns conditionally independent Bernoulli random variables. Stein's method is amenable to application of approximation of distribution of statistics on random graphs viewed as U-statistics. That is why we have focused on the branch of Stein's method leveraging a Malliavin-Dirichlet structure for conditionally independent random variables. It leads to Berry-Esseen bounds, including for functions of independent random variables. The development of this topic has been recent although there are earlier references that discuss Berry-Esseen bounds (Helmers, 1982; van Zwet, 1984; Friedrich, 1989), relying as usual on martingale representation.

Our contributions in this field are:

- A new Malliavin framework for conditionally independent random variables;
- A new proof of concentration inequalities for functionals of conditionally independent random variables;
- A pseudo chain rule for approximation of the derivation of the carré du champ operator;
- New bounds on the 1-Wasserstein bounds for U-statistics of conditionally independent random variables;
- A partial fourth moment theorem for functionals of conditionally independent random variables and a version for independent random variables;
- An application to asymptotic normality of subhypergraph counts in random exchangeable hypergraphs.

In the last chapter of the first part, we consider one of the applications in motif estimation.

## Chapter 2

# Malliavin calculus for conditionally independent random variables

On any denumerable product of probability spaces, we extend the discrete Malliavin structure for conditionally independent random variables. As a consequence, we obtain the chaos decomposition for functionals of conditionally independent random variables. We also show how to derive some concentration results in that framework. The Malliavin-Stein method yields Berry-Esseen bounds for U-Statistics of such random variables. It leads to quantitative statements of conditional limit theorems: Lyapunov's central limit theorem, De Jong's limit theorem for multilinear forms. The latter is related to the fourth moment phenomenon.

The chapter is organized as follows. The [Section 2.2](#) lays the foundations of the Malliavin framework. We define the Malliavin operators and especially the gradient that is related to the so-called *Glauber dynamics*. Our focus respectively lies on the independent setting and the conditional independent setting (i.e.  $X$  is a sequence of independent random variables, respectively conditionally independent random variables). We follow the original approach of completing it with a Dirichlet structure that naturally arises without further assumption that gives us another formula of integration by parts with the carré du champ operator. We derive some functional identities in [Section 2.3](#), specifically conditional versions of Poincaré inequality and McDiarmid's inequality. The [Section 2.4](#) presents results of normal approximation. We will see that using the carré du champ operators instead of the norms of Malliavin gradient will allow us to bypass at once all combinatorial difficulties, leading to a partial fourth moment theorem for U-statistics under mild assumptions (see [Section 2.5](#)) in the same vein of [Azmoodeh et al. \(2014\)](#) that shows the fourth moment theorem on the Wiener space. Finally, we show applications of such theorems to multilinear forms (also known as homogenous forms).

### 2.1 Motivation

It is only very recently that, concomitantly, the situation where the measured space is a product space, i.e. if we deal with independent random variables, has been addressed (see [Dung, 2018](#); [Decreusefond and Halconruy, 2019](#); [Duerinckx, 2021](#)) with a Malliavin framework. By order of complexity, the next situation which can be analyzed is that of conditionally independent random variables. This is a very common structure as de Finetti's theorem says that an infinite sequence of random variables is exchangeable if and only if these random variables are conditionally independent. This is the key theorem to develop a theory on random hypergraphs as in [Austin \(2008\)](#) that are the objects of the application.

The first definitions of gradient (denoted by  $D$ ) and divergence we introduce below for conditionally independent random variables, bear strong formal similarities with those of [Decreusefond and Halconruy \(2019\)](#). The difference lies into the computations which rely heavily on conditional distributions given the latent variable, which is here called  $Z$ . We can then follow the classical development of the Malliavin calculus apparatus: gradient, divergence, chaos, number operator and Ornstein-Uhlenbeck semigroup (denoted by  $P_t$ ). We can even describe the dynamics of the Markov process whose infinitesimal generator is the number operator. At a formal level, the computations are almost identical to those of [Decreusefond and Halconruy \(2019\)](#) with expectations replaced by expectations given  $Z$ . That notion of gradient is itself the extension of gradient in the particular discrete setting when  $X$  is a sequence of Rademacher random variables ([Privault, 2008](#), Proposition 10.1). Related works are [Reinert et al. \(2010\)](#); [Privault and Torrisi \(2015\)](#).

Nevertheless, for more advanced applications, namely functional identities like the covariance representation formula, we need to introduce a difference operator (see [Definition 2.2.24](#)) which appears more often than the gradient itself. It is in some sense a finer tool than the original gradient which is useful to define the Dirichlet structure (the Glauber process, the infinitesimal generator denoted by  $\mathbf{L}$ , etc.) but no more. This is due to the fact that  $D_a D_a = D_a$ , which entails that  $\mathbf{L}$  commutes with  $D$ , and thus we have  $D P_t = P_t D$  in place of the usual formula  $D P_t = e^{-t} P_t D$  which is the core formula to derive all functional inequalities in the Gaussian and Poisson cases. The difference operator  $\Delta$  allows to recover the crucial  $e^{-t}$  factor (see [Proposition 2.3.1](#)). As discussed earlier, the prevailing application of Malliavin calculus is nowadays, the evaluations of convergence rates via the Stein's method. Motivated by the applications to random graphs statistics, we focus here on normal approximations of  $U$ -statistics as in [Barbour et al. \(1989\)](#); [Krokowski et al. \(2017\)](#); [Privault and Serafin \(2020\)](#); [Röllin \(2022\)](#). In passing, we extend the notion of  $U$ -statistics by allowing the coefficients to depend on the latent variable instead of being only deterministic. Following the strategy of [Azmoodeh et al. \(2014\)](#), we establish a fourth moment theorem with remainder for such functionals as well as results on the control of the remainder.

## 2.2 Discrete Malliavin-Dirichlet structure

Set  $A$  be an at most denumerable set equipped with the counting measure, and define:

$$\ell^2(A) := \left\{ u : A \rightarrow \mathbb{R}, \sum_{a \in A} |u_a|^2 < \infty \right\} \text{ and } \langle u, v \rangle_{\ell^2(A)} := \sum_{a \in A} u_a v_a.$$

Let  $(\Omega, \mathcal{T}, \mathbb{P})$  be a probability space,  $E_0$  be a Polish space and  $((E_a, \Upsilon_a), a \in A)$  be a family of Polish spaces such that

$$\begin{aligned} E_A &= \prod_{a \in A} E_a \\ \Omega &= E_0 \times E_A. \end{aligned} \tag{2.1}$$

The product probability space  $E_A$  is endowed with its Borel  $\sigma$ -algebra denoted  $\Upsilon \subset \mathcal{T}$ . Let  $Z$  an  $E_0$ -valued random variable. By Theorem 10.2.2 ([Dudley, 2002](#)), all the subsequent conditional distributions in the chapter admit regular versions. For any subset  $B$  of  $A$ , we denote the set  $E_B := \prod_{b \in B} E_b$  and for  $x \in E_A$ ,  $x_B := (x_a, a \in B) \in E_B$  so that for  $a \in B$ ,  $x_a \in E_a$ . We denote  $x^B = (x_a, a \in A \setminus B)$ . Let  $X := (X_a)_{a \in A}$  be a sequence defined on  $(\Omega, \mathcal{T}, \mathbb{P})$  of conditionally

independent random variables given  $Z$  such that for all  $a \in A$ ,  $X_a$  is an  $E_a$ -valued random variable, i.e.:

$$X_a \perp\!\!\!\perp_Z (X_b, b \in A \setminus \{a\}),$$

or, equivalently:

$$\mathbb{P}(X_a \in \cdot \mid \sigma((X_b, b \neq a), Z)) = \mathbb{P}(X_a \in \cdot \mid \sigma(Z)).$$

We denote by  $\mathbf{P}$  the law of  $X$  and  $\mathbf{P}^Z$  the law  $\mathcal{L}(X|Z)$ . See chapter 6 of [Kallenberg \(2002\)](#) for a thorough review of conditional independence, and [Rao \(2009\)](#) for some limit theorems for conditionally independent random variables. We use the notation  $\mathbb{E}$  for the expectation of a random variable. By the disintegration theorem, for  $a \in A$ , the conditional probability distribution of  $X_a$  given  $\sigma(X^{\{a\}}) \vee \sigma(Z)$  admits a regular version  $\mathbf{P}_a$ . For  $p \geq 1$ , let us denote  $L^p(E_A \rightarrow \mathbb{R}, \mathbf{P})$  the set of  $p$ -th-integrable functions on  $E_A$  with respect to the measure  $\mathbf{P}$ . It is equipped with the norm  $\|\cdot\|_{L^p(E_A \rightarrow \mathbb{R}, \mathbf{P})}$  defined for  $f$  a measurable function on  $E_A$  by  $\|f\|_{L^p(E_A \rightarrow \mathbb{R}, \mathbf{P})} := \int |f(x)|^p \mathbf{P}(dx)$ . For the sake of notations,  $L^p(E_A)$  stands for the space of  $p$ -integrable functionals

$$L^p(E_A) := \left\{ \omega \mapsto F(X(\omega)) : \omega \in \Omega, F \in L^p(E_A \rightarrow \mathbb{R}, \mathbf{P}) \right\}.$$

In this respect,  $L^\infty(E_A)$  is the space of bounded functionals. We shall write  $F$  in place of  $F(X)$  for the sake of conciseness. We closely follow the usual construction of Malliavin calculus on that space.

**Definition 2.2.1.** A functional  $F$  is said to be cylindrical if there exists a finite subset  $I \subset A$  and a functional  $F_I$  in  $L^2(E_I)$  such that  $\mathbb{E}[|F_I|^2] < +\infty$  and  $F = F_I \circ r_I$ , where  $r_I$  is the restriction operator:

$$\begin{aligned} r_I &: E_A \longrightarrow E_I \\ (x_a, a \in A) &\longmapsto (x_a, a \in I). \end{aligned}$$

It is clear that the set of those functionals  $\mathcal{S}$  is dense in  $L^2(E_A)$ . We set  $L^2(A \times E_A)$  the Hilbert space of processes which are square-integrable with respect to the measure  $\sum_{a \in A} \Delta^{\{a\}} \otimes \mathbf{P}$ , i.e.

$$L^2(A \times E_A) = \left\{ U : \sum_{a \in A} \mathbb{E}[U_a(X)^2] < +\infty \right\},$$

equipped with the norm and inner product:

$$\|U\|_{L^2(A \times E_A)} := \sum_{a \in A} \mathbb{E}[U_a^2] \quad \text{and} \quad \langle U, V \rangle_{L^2(A \times E_A)} := \sum_{a \in A} \mathbb{E}[U_a V_a].$$

**Definition 2.2.2.** The set of simple processes, denoted  $\mathcal{S}_0(l^2(A))$  is the set of random variables defined on  $A \times E_a$  of the form

$$U = \sum_{a \in A} U_a \mathbf{1}_a,$$

for  $U_a \in \mathcal{S}$ .

### 2.2.1 Malliavin operators

**Definition 2.2.3** (Discrete gradient). For  $F \in \mathcal{S}$ ,  $DF$  is the simple process of  $L^2(A \times E_A)$  defined for all  $a \in A$  by:

$$D_a F := F - \mathbb{E} \left[ F | X^{\{a\}}, Z \right].$$

In particular,  $\mathcal{S} \subset \text{Dom } D$ . Define the  $\sigma$ -field  $\sigma(X^{\{a\}}) \vee \sigma(Z)$  by  $\mathcal{G}^a$ , so that

$$D_a F = F - \mathbb{E} [F | \mathcal{G}^a]. \quad (2.2)$$

Recall that for  $K \subset A$ ,  $X_K = (X_a, a \in K)$  and  $X^K = (X_a, a \in A \setminus K)$ . We shall write  $\mathcal{G}^K = \sigma(X^K) \vee \sigma(Z)$  and  $\mathcal{G}_K = \sigma(X_K) \vee \sigma(Z)$  for  $K$  a subset of  $A$ .

**Lemma 2.2.4** Let  $(a, b) \in A^2$ ,  $a \neq b$ , for  $F \in \text{Dom } D$ ,

1.  $D_a D_a F = D_a F$ ;
2.  $D_a D_b F = D_b D_a F$ ;
3.  $D_a \mathbb{E} [F | \mathcal{G}^b] = D_b \mathbb{E} [F | \mathcal{G}^a]$ .

*Proof of Lemma 2.2.4.* For  $(a, b) \in A^2$ , with  $b \neq a$ ,

$$\begin{aligned} D_a D_b F &= D_b F - \mathbb{E} [D_b F | \mathcal{G}^a] \\ &= F - \mathbb{E} [F | \mathcal{G}^b] - \mathbb{E} [F | \mathcal{G}^a] + \mathbb{E} \left[ \mathbb{E} [F | \mathcal{G}^b] | \mathcal{G}^a \right] \\ D_b D_a F &= D_a F - \mathbb{E} [D_a F | \mathcal{G}^b] + \mathbb{E} \left[ \mathbb{E} [F | \mathcal{G}^a] | \mathcal{G}^b \right] \\ &= F - \mathbb{E} [F | \mathcal{G}^a] - \mathbb{E} [F | \mathcal{G}^b] + \mathbb{E} \left[ \mathbb{E} [F | \mathcal{G}^a] | \mathcal{G}^b \right]. \end{aligned}$$

We note that:

$$\begin{aligned} &\mathbb{E} \left[ \mathbb{E} [F(X) | \mathcal{G}^a] | \mathcal{G}^b \right] \\ &= \int \int F(X_{A \setminus \{a, b\}}, x_a, x_b) \mathbf{P}_a((X_{A \setminus \{a, b\}}, Z), x_b, dx_a) \mathbf{P}_b((X_{A \setminus \{a, b\}}, Z), dx_b) \\ &= \int \int F(X_{A \setminus \{a, b\}}, x_a, x_b) \mathbb{P}^{X_b|Z}(Z, dx_b) \mathbb{P}^{X_a|Z}(Z, dx_a) \\ &= \mathbb{E} \left[ \mathbb{E} [F(X) | \mathcal{G}^b] | \mathcal{G}^a \right]. \end{aligned}$$

Hence, the equality follows. □

The key to the definition of the Malliavin framework is the so-called integration by parts.

**Theorem 2.2.5 — Integration by parts I.** Let  $F \in \mathcal{S}$ , for every simple process  $U$ ,

$$\langle DF, U \rangle_{L^2(E_A \times A)} = \mathbb{E} \left[ F \sum_{a \in A} D_a U_a \right]. \quad (2.3)$$

*Proof of Theorem 2.2.5.* We get:

$$\begin{aligned}
\langle DF, U \rangle_{L^2(E_A \times A)} &= \mathbb{E} \left[ \sum_{a \in A} D_a F U_a \right] \\
&= \mathbb{E} \left[ \sum_{a \in A} (F - \mathbb{E}[F | \mathcal{G}^a]) U_a \right] \\
&= \sum_{a \in A} \mathbb{E} [F(U_a - \mathbb{E}[U_a | \mathcal{G}^a])] \\
&= \sum_{a \in A} \mathbb{E} [F D_a U_a],
\end{aligned}$$

by self-adjointness of the conditional expectation.  $\square$

**Corollary 2.2.6 — Closability of the discrete gradient.** *The operator  $D$  is closable from  $L^2(E_A)$  into  $L^2(A \times E_A)$ .*

*Proof of Corollary 2.2.6.* The proof is analogous to the proof of closability of the gradient in (Decreusefond and Halconruy, 2019, Corollary 2.5)  $\square$

The domain of  $D$  in  $L^2(E_A)$  is the closure of cylindrical functionals with respect to the norm:

$$\|F\|_{1,2} := \sqrt{\|F\|_{L^2(E_A)}^2 + \|DF\|_{A \times L^2(E_A)}^2}.$$

The ensuing lemma gives a way to define square-integrable functionals in  $\text{Dom } D$  that are not in  $\mathcal{S}$ .

**Lemma 2.2.7** *If there exists a sequence  $(F_n)_{n \in \mathbb{N}}$  of elements of  $\text{Dom } D$  such that*

1. *the sequence converges to  $F$  in  $L^2(E_A)$ ;*
2.  $\sup_{n \in \mathbb{N}} \|DF_n\|_{L^2(E_A \times A)} < +\infty$ ;

*then  $F$  belongs to  $\text{Dom } D$  and  $DF = \lim_{n \rightarrow +\infty} DF_n$ .*

*Proof of Lemma 2.2.7.* Let  $(F_n)_{n \in \mathbb{N}}$  a sequence in  $L^2(E_A)$  with  $\mathbf{P}$ -a.s. limit  $F$ , then for  $a \in A$ ,

$$\begin{aligned}
\mathbb{E}[|D_a F - D_a F_n|^2] &\leq \mathbb{E}[|F - F_n|^2] + \mathbb{E}[|\mathbb{E}[F_n | \mathcal{G}^a] - \mathbb{E}[F | \mathcal{G}^a]|^2] \\
&\leq \mathbb{E}[|F - F_n|^2] + \mathbb{E}[\mathbb{E}[|F - F_n|^2 | \mathcal{G}^a]] \text{ by Jensen's inequality} \\
&= 2\mathbb{E}[|F - F_n|^2] \xrightarrow{n \rightarrow +\infty} 0.
\end{aligned}$$

Let  $(A_m)_{m \in \mathbb{N}}$  a family of subsets of  $A$  such that  $\bigcup_{m \geq 0} A_m = A$  and  $|A_m| = m$ , then for all  $m \in \mathbb{N}$ ,  $(\sum_{a \in A_m} D_a F_n)_{n \in \mathbb{N}}$  converges in  $L^2(E_A)$  to  $\sum_{a \in A_m} D_a F$ . We denote by  $D^m$  the operator on  $L^2(E_A \times A)$  such that for  $a \in A_m$ ,  $D_a^m = D_a$  and otherwise  $D_a^m$  is the null operator. For  $m \in \mathbb{N}$ ,  $(D^m F_n)_{n \in \mathbb{N}}$  converges to  $D^m F$  in  $L^2(E_A \times A)$ . Because of (2), by the uniform boundedness principle,  $DF$  is in  $L^2(E_A \times A)$ , and the result follows.  $\square$

**Definition 2.2.8** (Divergence operator). The domain of the divergence operator  $\text{Dom } \delta$  in  $L^2(E_A)$  is the set of processes  $U$  in  $L^2(E_A \times A)$  such that there exists  $\delta U$  satisfying the duality relation

$$\langle DF, U \rangle_{L^2(E_A \times A)} = \mathbb{E}[F\delta U], \text{ for all } F \in \text{Dom } D. \quad (2.4)$$

Moreover, for any process  $U$  belonging to  $\text{Dom } \delta$ ,  $\delta U$  is the unique element of  $L^2(E_A)$  characterized by that identity. The integration by parts formula entails that for every process  $U \in \text{Dom } \delta$ ,

$$\delta = \sum_{a \in A} D_a U_a. \quad (2.5)$$

**Definition 2.2.9** (Ornstein-Uhlenbeck operator). The Ornstein-Uhlenbeck operator, denoted by  $\mathsf{L}$  is defined on its domain

$$\text{Dom } \mathsf{L} = \left\{ F \in L^2(E_A) : \mathbb{E} \left[ \left| \sum_{a \in A} D_a F \right|^2 \right] < +\infty \right\} \supseteq \mathcal{S}$$

by

$$\mathsf{L}F := -\delta DF = -\sum_{a \in A} D_a F. \quad (2.6)$$

## 2.2.2 Chaos decomposition

The [Lemma 2.2.4](#) entails a chaos decomposition of  $L^2(E_A)$  similar to the one in [Duerinckx \(2021\)](#).

**Theorem 2.2.10 — Chaos decomposition.** For any  $F \in L^2(E_A)$ ,

$$F = \sum_{n=0}^{+\infty} \pi_n(F), \quad (2.7)$$

where  $(\pi_n)_{n \in \mathbb{N}}$  is a sequence of orthogonal projectors on  $L^2(E_A)$  with  $\pi_0(F) = \mathbb{E}[F|Z]$ .

*Proof.* One can notice that:

$$\mathbb{E}[D_a F(X)|\mathcal{G}^a] = D_a(\mathbb{E}[F|\mathcal{G}^a])F(X) = 0, \text{ for all } a \in A. \quad (2.8)$$

Let  $(A_m)_{m \in \mathbb{N}}$  a family of finite subsets of  $A$  such that  $|A_m| = m$  and  $\bigcup_{m \in \mathbb{N}} A_m = A$ . Let  $m \in \mathbb{N}$ ,  $\text{Id}_{L^2(E_{A_m})} = \prod_{a \in A_m} (D_a + \mathbb{E}[\cdot|\mathcal{G}^a])$ . Indeed, for all  $a \in A_m$ ,  $\text{Id}_{\text{Dom } D} = D_a + \mathbb{E}[\cdot|\mathcal{G}^a]$ . Hence, by distributivity and by using [Lemma 2.2.4](#), the identity also reads off:  $\text{Id}_{L^2(E_{A_m})} = \sum_{n=0}^m \pi_n^m$ , where

$$\pi_n^m := \sum_{J \subset A_m, |J|=n} \left( \prod_{b \in J} D_b \right) \left( \prod_{c \in A_m \setminus J} \mathbb{E}[\cdot|\mathcal{G}^c] \right) \quad \forall n \leq m. \quad (2.9)$$



Set  $n \leq m$ ,

$$\begin{aligned}
\pi_n^m \pi_n^m &= \sum_{\substack{I \subset A_m \\ |I|=n}} \sum_{\substack{J \subset A_m \\ |J|=n}} \left( \prod_{b \in I} D_b \right) \left( \prod_{c \in A_m \setminus I} \mathbb{E}[\cdot | \mathcal{G}^c] \right) \left( \prod_{d \in J} D_d \right) \left( \prod_{e \in A_m \setminus J} \mathbb{E}[\cdot | \mathcal{G}^e] \right) \\
&= \sum_{I \subset A_m, |I|=n} \sum_{J \subset A_m, |J|=n} \left( \prod_{b \in I} D_b \prod_{e \in A_m \setminus J} \mathbb{E}[\cdot | \mathcal{G}^e] \right) \left( \prod_{c \in A_m \setminus I} \mathbb{E}[\cdot | \mathcal{G}^c] \prod_{d \in J} D_d \right) \\
&= \sum_{\substack{I \subset A_m \\ |I|=n}} \left( \prod_{b \in I} D_b \prod_{e \in A \setminus I} \mathbb{E}[\cdot | \mathcal{G}^e] \right) \left( \prod_{c \in A_m \setminus I} \mathbb{E}[\cdot | \mathcal{G}^c] \prod_{d \in I} D_d \right) \text{ by Lemma 2.2.4} \\
&= \sum_{I \subset A_m, |I|=n} \left( \prod_{b \in I} \prod_{b \in I} D_b D_b \right) \left( \prod_{c \in A_m \setminus I} \mathbb{E}[\cdot | \mathcal{G}^c] \mathbb{E}[\cdot | \mathcal{G}^c] \right) = \pi_n^m.
\end{aligned} \tag{2.10}$$

By convention  $\pi_n^m(F) = 0$  for  $n > m$ . Analogously, for  $n' \neq n$ ,  $\pi_n^m \pi_{n'}^m = 0$ . The operator  $\pi_n^m$  is continuous on  $L^2(E_A)$ . Hence,  $(\pi_n^m)_{m \in \mathbb{N}}$  is a well-defined family of projectors on  $L^2(E_A)$ . Moreover, for all  $n \in \mathbb{N}$  and  $F \in L^2(E_A)$ , we have  $\sup_{m \in \mathbb{N}} \|\pi_n^m(F)\|_{L^2(E_A)} \leq \|F\|_{L^2(E_A)}$ . Then, by the uniform boundedness principle,

$$\sup_{\substack{m \in \mathbb{N} \\ \|F\|_{L^2(E_A)}}} \|\pi_n^m(F)\|_{L^2(E_A)} < +\infty.$$

The pointwise limits of  $(\pi_n^m(F))_{m \in \mathbb{N}}$  for  $F \in L^2(E_A)$  define a bounded linear operator  $\pi_n$  on  $L^2(E_A)$  for  $n \in \mathbb{N}$ . Thus:

$$L^2(E_A) = \bigoplus_{n=0}^{+\infty} \text{Im } \pi_n. \tag{2.11}$$

Given (2.9), for a functional  $F \in \text{Dom } \mathbf{L}$ , we have  $\pi_0(F) = \mathbb{E}[F | Z]$ .  $\square$

**Lemma 2.2.11 — Spectral decomposition.** *Let  $F \in L^2(E_A)$  of chaos decomposition*

$$F = \mathbb{E}[F | Z] + \sum_{n=1}^{+\infty} \pi_n(F).$$

1. *We say that  $F$  belongs to  $\text{Dom } \mathbf{L}$  whenever*

$$\sum_{n=1}^{+\infty} n^2 \|\pi_n(F)\|_{L^2(E_A)} < +\infty.$$

2. *The operator has a unit spectral gap, i.e. the spectrum of  $\mathbf{L}$  coincides with  $\mathbb{N}_0$ .*

$$L^2(E_A) = \bigoplus_{k=0}^{+\infty} \ker(\mathbf{L} + k\text{Id}). \tag{2.12}$$

3. *It is invertible from  $L_0^2(E_A) = \{F \in L^2(E_A), \mathbb{E}[F | Z] = 0\}$  into itself.*

*Proof of Lemma 2.2.11.* Let us show that  $\pi_n$  is in the domain of  $L$  for all  $n \in \mathbb{N}$ . By summability,

$$\begin{aligned}
\left| \sum_{a \in A} D_a \pi_n \right|^2 &= \left| \sum_{a \in A} D_a \sum_{I \subset A, |I|=n} \left( \prod_{b \in I} D_b \right) \left( \prod_{c \in A \setminus I} \mathbb{E}[\cdot | \mathcal{G}^c] \right) \right|^2 \\
&= \left| \sum_{a \in A} \mathbf{1}_I(a) \sum_{I \subset A, |I|=n} \left( \prod_{b \in I} D_b \right) \left( \prod_{c \in A \setminus I} \mathbb{E}[\cdot | \mathcal{G}^c] \right) \right|^2 \\
&= n^2 \left| \sum_{I \subset A, |I|=n} \left( \prod_{b \in I} D_b \right) \left( \prod_{c \in A \setminus I} \mathbb{E}[\cdot | \mathcal{G}^c] \right) \right|^2 \quad \text{since } |I| = n \\
&= n^2 |\pi_n|^2,
\end{aligned} \tag{2.13}$$

so for  $F \in L^2(E_A)$ ,  $\pi_n(F) \in \text{Dom } L$ . Hence, because of the orthogonality of  $(\text{Im } \pi_n)_{n \in \mathbb{N}}$ ,  $F \in \text{Dom } L \iff \sum_{n=1}^{+\infty} n^2 \|\pi_n(F)\|_{L^2(E_A)}^2 < +\infty$ . With the same calculations, we get  $L\pi_n = -n\pi_n$ . The spectrum of  $-L$  coincides with  $\mathbb{N}$ . Then, we deduce that:

$$L = \sum_{n=0}^{+\infty} -n\pi_n, \tag{2.14}$$

and  $\text{Im } \pi_n \subset \ker(L + n\text{Id})$ . Because of the orthogonality of the kernels, we get  $\text{Im } \pi_n = \ker(L + n\text{Id})$ . Now we prove the third item. The pseudoinverse  $L^{-1}$  is defined on its domain  $\{F \in L^2(E_A) : \mathbb{E}[F|Z] = 0\}$  and reads  $\sum_{n=1}^{+\infty} -\frac{\pi_n}{n}$ . Then for  $F \in \{G \in \text{Dom } L : \mathbb{E}[G|Z] = 0\}$ ,  $L^{-1}(LF) = F$ .  $\square$

**Corollary 2.2.12** For  $k > 0$  and  $J$  a subset of  $A$  of cardinal  $k$ , define by  $\mathfrak{C}_k$  the space of functionals  $\phi = \sum_{J \subset A, |J|=k} \psi_J$  such that:

- for every  $J \subset A$ ,  $\psi_J$  is  $\mathcal{F}_J$ -measurable;
- for every  $K \subset A$ ,  $\mathbb{E}[\psi_J | \mathcal{G}_K] = 0$  unless  $K \subset J$ ;

then  $\mathfrak{C}_k = \ker(L + k\text{Id})$ .

*Proof of Corollary 2.2.12.* From (2.9), for  $J = (a_1, \dots, a_n) \subset A$ , the component  $\psi_J$  is  $\mathcal{F}_J$ -measurable. Let us compute the expression of the iterated gradient for  $F$  a  $\mathcal{F}_J$ -measurable function:

$$\begin{aligned}
\prod_{a \in J} D_a F &= \sum_{k=0}^{|J|} (-1)^k \sum_{\substack{K \subset J \\ |K|=k}} \mathbb{E}[F | \mathcal{G}^K] \\
&= \sum_{L \subset J} (-1)^{|J|-|L|} \mathbb{E}[F | \mathcal{G}_L],
\end{aligned}$$

where  $\mathcal{G}^K = \sigma(X^K) \vee \sigma(Z)$  and  $\mathcal{G}_L = \sigma(X_L) \vee \sigma(Z)$ . In this view, we have the inclusion  $\ker(L + n\text{Id}) = \text{Im } \pi_n \subset \mathfrak{C}_n$  for  $n \in \mathbb{N}$ .

Conversely, let  $\phi$  for which the properties above hold.

$$\begin{aligned}
 \mathsf{L}\phi &= - \sum_{a \in A} D_a \sum_{J \subset A, |J|=n} \psi_J \\
 &= - \sum_{a \in A} \sum_{J \subset A, |J|=n} (\psi_J - \mathbb{E}[\psi_J | \mathcal{G}^a]) \\
 &= - \sum_{k \in A} \sum_{\substack{J \subset A, |J|=n \\ a \in J}} \psi_J \text{ because } \mathbb{E}[\psi_J | \mathcal{F}_{A \setminus \{a\}}] = 0 \text{ for } J \not\subset A \setminus \{a\} \\
 &= -n \sum_{J \in \mathcal{A}, |J|=n} \psi_J = -n\phi.
 \end{aligned}$$

Therefore,  $\mathfrak{C}_n = \ker(\mathsf{L} + n\text{Id})$  for  $n \geq 1$ . □

**Remark 2.2.13.** The first term in the decomposition is not  $\mathbb{E}[F]$ . It is known that there is no Hoeffding decomposition for functionals of exchangeable random variables unless we assume *weak independence* (Peccati, 2004). The decomposition at hand is not a *Hoeffding decomposition* since the kernels of the U-statistics also depend on  $Z$ . ■

### 2.2.3 Dirichlet structure

The map  $\mathsf{L}$  can be viewed as the generator of a Glauber dynamics where the index set is a finite set of random variables indexed by  $A_m$  for  $m > 1$ . For practical term, we introduce a new index  $\partial$  and  $X_\partial = Z$   $\mathbb{P}$ -a.s.

**Definition 2.2.14** (Modified Glauber process). Consider  $(N(t))_{t \geq 0}$  a Poisson process on the half-line  $[0, +\infty)$  of rate  $|A_m| + 1$ . Let  $(X^{\circ A_m}(t))_{t \geq 0} = (X_a^{\circ A_m}(t), t \geq 0, a \in A)$  the process valued in  $E_A$  starting with  $X^{\circ A_m}(0) = X$  which evolves according to the following rule. At jump time  $\tau$  of the process,

- Choose randomly an index  $a$  in  $A_m \sqcup \{\partial\}$  with equal probability.
- If  $a \neq \partial$ , replace  $X_a^{\circ A_m}(\tau)$  with a conditionally independent random variable  $X_a^\dagger$  distributed according to  $\mathbf{P}_a((X_{A \setminus \{a\}}^{\circ A_m}(\tau), Z), \cdot)$ , otherwise do nothing.

That Markov process has for infinitesimal generator  $\mathsf{L}^{A_m}$ :

$$\mathsf{L}^{A_m} F = - \sum_{a \in A_m} D_a F.$$

**Remark 2.2.15.** That is referred as Glauber dynamics because one can identify a coordinate, and change the random variable at this coordinate in function of the other ones. Here it does not simulate the Ising model, but a more general model. Our description is similar to the Glauber dynamics described in Adamczak et al. (2022, Section 4.3) and Duerinckx (2021) which deal with independent random variables. ■

Our aim is to show that the operator  $\mathsf{L}$  is an infinitesimal generator, letting  $m \rightarrow +\infty$ . We recall the Hille-Yosida theorem (Yosida, 1995).

**Proposition 2.2.16 — Hille-Yosida.** A linear operator  $L$  on  $L^2(E_A)$  is the generator of a strongly continuous contraction semigroup on  $L^2(E_A)$  if and only if

1.  $\text{Dom } L$  is dense in  $L^2(E_A)$ ;
2.  $L$  is dissipative, i.e. for any  $\lambda > 0$ ,  $F \in \text{Dom } L$ ,

$$\|\lambda F - LF\|_{L^2(E_A)} \geq \lambda \|F\|_{L^2(E_A)};$$

3.  $\text{Im}(\lambda \text{Id} - L)$  is dense in  $L^2(E_A)$ .

**Theorem 2.2.17**  $\mathbf{L}$  is an infinitesimal generator on  $E_A$  of a strongly continuous contraction semigroup on  $L^2(E_A)$ .

*Proof of Theorem 2.2.17.* We know that  $\mathcal{S}$  is dense in  $L^2(E_A)$ . As  $\text{Dom } \mathbf{L} \supset \mathcal{S}$ , it is also dense in  $L^2(E_A)$ . Let  $A_m$  an increasing sequence (with respect to  $\subset$ ) of subsets of  $A$  such that  $\bigcup_{n \geq 1} A_m = A \cup \partial$  and  $|A_m| = m$ . Then  $(\mathcal{F}_{A_m})_{m \in \mathbb{N}}$  is a filtration. For  $F \in L^2(E_A)$ , let  $F_m = \mathbb{E}[F | \mathcal{F}_{A_m}]$ . Since  $(F_m)_{m \in \mathbb{N}}$  is a square-integrable  $\mathcal{F}_A$ -martingale,  $(F_m)_{m \in \mathbb{N}}$  converges both almost surely and in  $L^2(E_A)$  to  $F$ . For any  $m \in \mathbb{N}$ ,  $F_m$  depends only on  $X_{A_m}$ . Because of the conditional independence of the random variables  $X_a$  given  $X_\partial$ , for all  $a \in A$ , we get that  $D_a F_m = \mathbb{E}[D_a F | \mathcal{F}_{A_m}]$ . Using that  $\mathbf{L}_{A_m}$  is dissipative for all  $m \in \mathbb{N}$ , we have:

$$\begin{aligned} \lambda^2 \|F_m\|_{L^2(E_A)}^2 &\leq \|\lambda F_m - \mathbf{L}^{A_m} F_m\|_{L^2(E_A)}^2 = \mathbb{E} \left[ \left( \lambda F_m + \sum_{a \in A_m} D_a F_m \right)^2 \right] \\ &= \mathbb{E} \left[ \left( \lambda F_m + \sum_{a \in A} D_a F_m \right)^2 \right] \text{ because } D_a F_m = 0, \text{ if } a \notin A_m. \\ &= \mathbb{E} \left[ \mathbb{E} \left[ \lambda F + \sum_{a \in A} D_a F \mid \mathcal{F}_{A_m} \right]^2 \right]. \end{aligned}$$

It means that the operator  $\mathbf{L}$  is dissipative. Thus, by the Hille-Yosida theorem,  $\mathbf{L}$  is the infinitesimal generator of a strongly continuous contraction semigroup on  $L^2(E_A)$  denoted  $P$ .  $\square$

**Lemma 2.2.18** Let  $F \in L^2(E_A)$ , then

$$\mathbb{E} [F(X^{\circ A_m}) | X, Z] = P_t^{A_m} F \xrightarrow{\mathbf{P}\text{-a.s.}} P_t F$$

and

$$X^{\circ A_m} \xrightarrow{d} X^\circ.$$

*Proof of Lemma 2.2.18.* The theorem 19.25 of (Kallenberg, 2002, Trotter, Sova, Kurtz, Mackevičius) gives the convergence in distribution of  $X^{\circ A_m}$  towards  $X^\circ$  the Markov process associated to  $\mathbf{L}$ , and the almost sure convergence of the semigroup.  $\square$

These are pieces of the Dirichlet structure with invariant measure  $\mathbf{P}$  that we complete with the carré du champ operator. Here, we note that  $\mathcal{S}$  is an algebra which is a core of  $\text{Dom } \mathbf{L}$ .

**Definition 2.2.19** (Carré du champ operator). Let  $F, G \in \mathcal{S}$ . The bilinear map

$$\Gamma(F, G) := \frac{1}{2} \{ \mathbf{L}(FG) - F\mathbf{L}G - G\mathbf{L}F \}$$

is well-defined, and called carré du champ operator of the Markov generator  $\mathbf{L}$ .

By an argument of density, there exists an algebra  $\mathcal{A} \supset \mathcal{S}$  maximal in the sense of inclusion such that the carré du champ operator acts on it.

**Definition 2.2.20** (Dirichlet structure). The associated Dirichlet structure defined on  $(E_A, \Upsilon, \mathbf{P})$  is given by the quadruple  $(X^\circ, \mathbf{L}, (P_t)_{t \geq 0}, \mathcal{E})$  where  $X^\circ$  is a Markov process with values in  $E_A$  whose infinitesimal generator is  $\mathbf{L}$  and its semigroup is  $P$ , i.e. for any  $F \in L^\infty(E_A)$ :

$$\frac{d}{dt} P_t F = (\mathbf{L}P_t)F.$$

Furthermore,  $\mathbf{P}^Z$  is the invariant (or stationary) distribution of  $X^\circ$  given  $Z$  and the Dirichlet form is defined by

$$\mathcal{E}(F, G) = \mathbb{E}[\Gamma(F, G)].$$

It comes with the classical properties entailed by the spectral decomposition of  $\mathbf{L}$ , including the Mehler's formula.

**Lemma 2.2.21 — Mehler's formula.** For any  $F \in L^2(E_A)$ ,

1.

$$\begin{aligned} P_t F &= \mathbb{E}[F|Z] + \sum_{n=1}^{\infty} e^{-nt} \pi_n(F) \\ &= \mathbb{E}[F(X^\circ(t))|X]. \end{aligned} \tag{2.15}$$

In particular  $P_t F \in \text{Dom } \mathbf{L} \cap \text{Dom } \mathbf{L}^{-1}$ .

2.

$$\lim_{t \rightarrow \infty} P_t F(X) = \mathbb{E}[F(X)|Z].$$

3. The pseudoinverse of  $\mathbf{L}$  can be written:

$$\mathbf{L}^{-1}F := - \int_0^{+\infty} P_t F \, dt.$$

*Proof of Lemma 2.2.21.* Since formally  $P_t = e^{-t\mathbf{L}}$ , we get the first line of (2.15) from the spectral decomposition of  $\mathbf{L}$ . The second line is deduced from the definition of the Glauber dynamics and by passing to the limit. Then,

$$\begin{aligned} \mathbb{E}[F|Z] - F &= \lim_{t \rightarrow +\infty} P_t F - P_0 F \\ &= \int_0^{+\infty} \frac{d}{dt} P_t F \, dt \\ &= \mathbf{L} \left( \int_0^{+\infty} P_t F \, dt \right). \end{aligned}$$

Taking  $\mathbb{E}[F|Z] = 0$ , we get the expression of the pseudoinverse.  $\square$

**Remark 2.2.22.** By the chaos expansion,  $P_t F$  can be defined as the limit in  $L^2(E_A)$  of elements  $(P_t F_n)_{n \in \mathbb{N}}$  for  $F_n$  in  $\mathcal{S}$ . Hence, it is sufficient to define the semigroup acting on a functional of some finite vector of random variables  $X_B$ , using the definition of the Glauber dynamics entailed by it.  $\blacksquare$

The infinitesimal generator satisfies another integration by parts formula due to the Dirichlet structure which is the key to investigating the so-called fourth moment phenomenon.

**Lemma 2.2.23 — Integration by parts II.** For  $(F, G) \in \mathcal{A}^2$ ,

$$\mathcal{E}(F, G) = -\mathbb{E}[FLG]. \quad (2.16)$$

We introduce to the difference operator which is associated to the Malliavin-Dirichlet structure at hand. That difference operator serves the same purpose as in [Lachièze-Rey and Peccati \(2017\)](#) and [Dung \(2021\)](#) for computations in the proofs of the limit theorems.

**Definition 2.2.24** (Difference operator). Let  $F : E_A \rightarrow \mathbb{R}$ , for  $a \in A$ , we introduce the operator

$$\begin{aligned} \Delta^{\{a\}} F : E_A \times E_a &\longrightarrow \mathbb{R} \\ (x, x'_a) &\longmapsto f(x) - f(x^{\{a\}}, x'_a). \end{aligned}$$

For the sake of conciseness, we shall write  $F^{\{a\}'} = F(X^{\{a\}}, X'_a)$ .

**Lemma 2.2.25** For  $F$  a functional in  $\text{Dom } D$ , the gradient also reads as:

$$D_a F = \mathbb{E} \left[ \Delta^{\{a\}} F(X, X'_a) | X, Z \right], \quad (2.17)$$

where  $X'_a$  has the law of  $X_a$  given  $Z$  and is conditionally independent of  $X^{\{a\}}$  given  $Z$ . Similarly,

$$\Gamma(F, G) = \frac{1}{2} \sum_{a \in A} \mathbb{E} \left[ \left( \Delta^{\{a\}} F(X, X'_a) \right) \left( \Delta^{\{a\}} G(X, X'_a) \right) | X, Z \right]. \quad (2.18)$$

*Proof of Lemma 2.2.25.* We have

$$\mathbb{E}[F | \mathcal{G}^a] = \int F(X^{\{a\}}, x_a) \mathbf{P}_a(dx_a).$$

Since  $\sigma(X_a)$  is independent of  $\sigma(X^{\{a\}})$  given  $\sigma(Z)$ , we obtain

$$\mathbb{E}[F | \mathcal{G}^a] = \int F(X^{\{a\}}, x_a) \mathbb{P}^{X_a|Z}(dx_a).$$

Eqn.(2.18) is proved similarly.  $\square$

## 2.3 Functional identities

This section is devoted to classical functional identities obtained in the Malliavin framework. We follow the approach of Houdré and Privault (2002) using a covariance identity based on difference operators to deduce concentration inequalities. We note that for functionals of independent random variables, Decreusefond and Halconruy (2019) obtained covariance identities based on the Clark-Ocone representation formula instead.

**Proposition 2.3.1** For  $F \in L^2(E_A)$  and  $a \in A$ , then:

$$D_a(P_t F) = e^{-t} \mathbb{E} \left[ \Delta^{\{a\}} F(X^\circ(t), X'_a) | X, Z \right] \quad (2.19)$$

where  $X'$  has the law of  $X$  given  $Z$ .

*Proof of Proposition 2.3.1.* We consider the Glauber dynamics with index set a finite subset  $A_m$  of  $A$ , as the construction of process  $(X^{\circ A_m}(t))_{t \in \mathbb{R}^+}$  is explicit in that case. Let  $a \in A_m$ , we denote by  $N_a$  the Poisson process of intensity 1 which represents the life duration of the  $a$ -th component in the dynamics of  $X^{\circ A_m}(t)$ , so:

$$X_a^{\circ A_m}(t) = \mathbb{1}_{\{\tau_a \geq t\}} X_a + \mathbb{1}_{\{\tau_a < t\}} X_a^\backslash,$$

where  $\tau_a = \inf\{t \geq 0, N_a(t) \neq N_a(0)\}$  is the life duration of the  $a$ -th component of the original sequence, exponentially distributed with parameter 1 (independent of everything else) and  $X_a^\backslash$  is conditionally independent of  $X$  given  $Z$ . Then:

$$\begin{aligned} D_a P_t^{A_m} F &= P_t^{A_m} F - \mathbb{E} \left[ P_t^{A_m} F | \mathcal{G}_a \right] \\ &= P_t^{A_m} F - \mathbb{E} \left[ \mathbb{E} \left[ F(X^{\circ A_m}(t)) | X, Z \right] \mathbb{1}_{\{t \leq \tau_a\}} | \mathcal{G}_a \right] - \mathbb{E} \left[ F(X^{\circ A_m}(t)) \mathbb{1}_{\{t > \tau_a\}} | X, Z \right] \\ &= \mathbb{E} \left[ F(X^{\circ A_m}(t)) \mathbb{1}_{\{t \leq \tau_a\}} | X, Z \right] - \mathbb{E} \left[ \mathbb{E} \left[ F(X^{\circ A_m}(t)) | X, Z \right] \mathbb{1}_{\{t \leq \tau_a\}} | \mathcal{G}_a \right] \\ &= e^{-t} \mathbb{E} \left[ \Delta^{\{a\}} F(X^{\circ A_m}(t), X'_a) | X, Z \right] \end{aligned}$$

because the law of  $X_a^\backslash$  given  $X$  is the same as the one of  $X'_a$  given  $X$ .

On one hand,

$$D_a P_t^{A_m} F \xrightarrow{\mathbb{P}\text{-a.s.}} D_a P_t F.$$

On the other hand, by the Skorohod's representation theorem, there exist copies of  $X^{\circ A_m}$  and  $X^\circ$  on a common probability space  $(\tilde{\Omega}, \tilde{\mathcal{T}}, \tilde{\mathbb{P}})$  such that the sequence  $(X^{\circ A_m})_{m \in \mathbb{N}}$  converges to  $X^\circ$   $\tilde{\mathbb{P}}$ -a.s. As the whole structure is invariant by copy, we can suppose the almost sure convergence on  $(\Omega, \mathcal{T}, \mathbb{P})$ , and the relation passes to the limit.  $\square$

**Remark 2.3.2.** In the case, we have only one random variable (or one particle), then the commutation relation simplifies to  $D_a(P_t F) = D_a$ , reminiscent of the one that holds for Malliavin operators in the Wiener and Poisson settings.  $\blacksquare$

**Corollary 2.3.3 — Conditional covariance identity.** For any  $F, G \in L^2(E_A)$ , then:

$$\text{Cov}(F, G | Z) = \int_0^\infty e^{-t} \sum_{a \in A} \mathbb{E} \left[ (D_a F)(\Delta^{\{a\}} G(X^\circ(t), X'_a)) | Z \right] dt. \quad (2.20)$$

*Proof of Corollary 2.3.3.* We use the following conditional covariance formula analogous to the covariance formula:

$$\text{Cov}(F, G|Z) = \mathbb{E}[FG|Z] = \mathbb{E}[FLL^{-1}G|Z]. \quad (2.21)$$

By the integration by parts I (2.3) which also holds with conditional expectation given  $Z$ , we get:

$$\begin{aligned} \mathbb{E}[FLL^{-1}G|Z] &= - \sum_{a \in A} \mathbb{E}[(D_a F)(D_a L^{-1}G)|Z] \\ &= - \sum_{a \in A} \mathbb{E}\left[(D_a F)(D_a \int_0^\infty P_t G dt)|Z\right] \\ &= - \sum_{a \in A} \mathbb{E}\left[(D_a F) \left(\int_0^\infty D_a P_t G dt\right)|Z\right] \\ &= - \int_0^\infty e^{-t} \sum_{a \in A} \mathbb{E}\left[(D_a F) \mathbb{E}[\Delta^{\{a\}} G(X^\circ(t), X'_a)|X, Z]|Z\right] dt, \end{aligned}$$

using (2.19). □

As an immediate consequence of the spectral gap, we find another proof of the Efron-Stein inequality which is of independent interest.

**Proposition 2.3.4** *If  $F \in \mathfrak{C}_p$  then*

$$\text{Var}[F] = \frac{1}{p} \mathcal{E}(F) = \frac{1}{p} \|DF\|_{L^2(E_A)}.$$

*Moreover, if there exist  $F_1, \dots, F_m \in L^2(E_A)$  such that  $F = \sum_{p=1}^m F_p$  with  $F_p \in \mathfrak{C}_p$  for  $p \in \llbracket 1, m \rrbracket$ , then:*

$$\text{Var}[F] \leq \|DF\|_{L^2(E_A)}. \quad (2.22)$$

*Proof of Proposition 2.3.4.* We use the previous covariance identity which yields:

$$\begin{aligned} \text{Var}[F] &= \text{Cov}(F, F) = \mathbb{E}[\Gamma(F, -L^{-1}F)] \\ &= \mathbb{E}\left[\Gamma\left(\sum_{p=1}^m F_p, \sum_{q=1}^m \frac{1}{q} F_q\right)\right] \\ &= \sum_{p=1}^m \sum_{q=1}^m \frac{1}{q} \mathbb{E}[\Gamma(F_p, F_q)] \\ &= \sum_{p=1}^m \frac{1}{p} \mathbb{E}[\Gamma(F_p, F_p)] \text{ because } \mathbb{E}[\Gamma(F_p, F_q)] = 0 \text{ for } q \neq p. \end{aligned}$$

It yields the inequality (2.22) noting that  $\Gamma(F_p, F_p) \geq 0$  for all  $p > 0$ . □

We now deduce the conditional first-order Poincaré inequality for functionals of conditionally independent random variables. The equivalent for functionals of independent random variables is rather known as the Efron-Stein inequality in the literature (Efron and Stein, 1981).



**Theorem 2.3.5 — Conditional Efron-Stein inequality.** For  $F \in L^2(E_A)$  such that  $\mathbb{E}[F|Z] = 0$ ,

$$\text{Var}[F|Z] \leq \mathbb{E}[\Gamma(F, F)|Z]. \quad (2.23)$$

*Proof of Theorem 2.3.5.* The conditional covariance formula yields

$$\begin{aligned} \text{Var}[F|Z] &= \int_0^\infty e^{-u} \sum_{a \in A} \mathbb{E} \left[ (D_a F)(\Delta^{\{a\}} F)(X_u^\circ, X_a') | Z \right] du \\ &\leq \int_0^\infty e^{-u} \sqrt{\sum_{a \in A} \mathbb{E}[(D_a F)^2 | Z]} \sqrt{\sum_{a \in A} \mathbb{E}[\mathbb{E}[(\Delta^{\{a\}} F)(X_u^\circ, X_a')^2 | X, Z]^2 | Z]} du. \end{aligned}$$

The invariance of  $\mathbf{P}^Z$  under the Glauber dynamics entails that

$$\sum_{a \in A} \mathbb{E} \left[ \mathbb{E} \left[ (\Delta^{\{a\}} F)(X_u^\circ, X_a')^2 | X, Z \right]^2 | Z \right] = \sum_{a \in A} \mathbb{E}[(D_a F)^2 | Z].$$

Hence,

$$\text{Var}[F|Z] \leq \mathbb{E}[\Gamma(F, F)|Z],$$

proving the theorem.  $\square$

**Remark 2.3.6** (Optimal constant in the Poincaré inequality). As mentioned in Bakry, Gentil, and Ledoux (2013), if a Poincaré inequality  $\mathcal{P}(C)$  with constant  $C$  holds for the Dirichlet structure at hand, the spectrum of the symmetric positive operator  $-\mathbf{L}$  is included in  $\{0\} \cup [\frac{1}{C}, \infty]$ . The existence of  $\mathcal{P}(C)$  does not tell whether the spectrum of  $\mathbf{L}$  is discrete or not. The consequence of  $\mathcal{P}(C)$  is the fact that  $\mathcal{E}(F) = 0 \implies F$  is constant.  $\blacksquare$

**Remark 2.3.7.** Even in the absence of chain rule, some papers derived second-order Poincaré inequalities (see e.g. Eichelsbacher et al. (2023)) in discrete setting. The principle of the latter article can be mobilized as to show similar results based on our version of Malliavin-Stein's method which is the topic of the next section. See Nourdin et al. (2009) for the original scheme of proof of second-order Poincaré inequalities on the Wiener space.  $\blacksquare$

We find a version of the McDiarmid's inequality for conditionally independent random variables.

**Theorem 2.3.8 — Conditional McDiarmid's inequality.** Let  $F$  be a square-integrable functional such that for all  $a \in A$ :

$$\sup_{\substack{x^{\{a\}} \in E_{A \setminus \{a\}} \\ x'_a \in E_a}} |F(x^{\{a\}}, x'_a) - F(x)| \leq d_a.$$

For any  $x > 0$ , we have the inequality:

$$\mathbb{P}(F(X) - \mathbb{E}[F(X)|Z] \geq x|Z) \leq \exp\left(-\frac{x^2}{2 \sum_{a \in A} d_a^2}\right). \quad (2.24)$$

Our strategy of proof is different from the original McDiarmid's original proof in McDiarmid (1989).

*Proof of Theorem 2.3.8.* We assume that  $F = F(X)$  is a bounded random variable verifying  $\mathbb{E}[F|Z] = 0$ . Using the inequality:

$$|e^{tx} - e^{ty}| \leq \frac{t}{2}|x - y|(e^{tx} + e^{ty}) \quad \forall x, y \in \mathbb{R}. \quad (2.25)$$

We have:

$$\begin{aligned} |\Delta^{\{a\}} e^{tF}(X, X'_a)| &= |e^{tF} - e^{tF^{\{a\}'}}| \\ &\leq \frac{t}{2} |\Delta^{\{a\}} F(X, X'_a)| (e^{tF} + e^{tF^{\{a\}'}}). \end{aligned}$$

Applying the covariance identity, it yields:

$$\begin{aligned} \mathbb{E}[F e^{tF} | Z] &= \int_0^\infty e^{-u} \sum_{a \in A} \mathbb{E}[D_a e^{tF} \Delta^{\{a\}} F(X'_u, X'_a) | Z] du \\ &\leq \int_0^\infty e^{-u} \sum_{a \in A} \mathbb{E} \left[ \mathbb{E} \left[ |\Delta^{\{a\}} e^{tF}(X, X'_a) | X, Z \right] \Delta^{\{a\}} F(X'_u, X'_a) | Z \right] du \\ &\leq \frac{t}{2} \int_0^\infty e^{-u} \sum_{a \in A} \mathbb{E} \left[ |\Delta^{\{a\}} F(X, X'_a)| e^{tF} |\Delta^{\{a\}} F(X'_u, X'_a)| | Z \right] du \\ &\quad + \frac{t}{2} \int_0^\infty e^{-u} \sum_{a \in A} \mathbb{E} \left[ |\Delta^{\{a\}} F(X, X'_a)| e^{tF^{\{a\}'}} |\Delta^{\{a\}} F(X'_u, X'_a)| | Z \right] du \end{aligned}$$

by using the Jensen's inequality for conditional expectation in the second inequality. Since  $|\Delta^{\{a\}} F(X, X'_a)|^2 \leq d_a$ ,  $|\Delta^{\{a\}} F(X'_u, X'_a)| \leq d_a$  for all  $u \in \mathbb{R}^+$  and  $\mathbb{E}[e^{tF^{\{a\}'}} | Z] = \mathbb{E}[e^{tF} | Z]$ , this shows that:

$$\mathbb{E}[F e^{tF} | Z] \leq \left( \sum_{a \in A} d_a^2 \right) t \mathbb{E}[e^{tF} | Z] = t K^2 \mathbb{E}[e^{tF} | Z],$$

where  $K^2 := \sum_{a \in A} d_a^2$ . Thus, in all generality for  $F$  bounded:

$$\begin{aligned} \log \mathbb{E}[e^{t(F - \mathbb{E}[F|Z])} | Z] &= \int_0^t \frac{\mathbb{E}[(F - \mathbb{E}[F|Z]) e^{s(F - \mathbb{E}[F|Z])} | Z]}{\mathbb{E}[e^{s(F - \mathbb{E}[F|Z])}]} ds \\ &\leq K^2 \int_0^t s ds = \frac{t^2}{2} K^2, \end{aligned}$$

hence:

$$\begin{aligned} e^{tx} \mathbb{P}(F - \mathbb{E}[F|Z] > x | Z) &\leq \mathbb{E}[e^{t(F - \mathbb{E}[F|Z])} | Z] \\ &= e^{t^2 K^2 / 2}, \quad t \geq 0, \end{aligned}$$

and:

$$\mathbb{P}(F - \mathbb{E}[F|Z] \geq x | Z) \leq e^{\frac{t^2}{2} K^2 - tx}, \quad t \geq 0.$$

The minimum of the right-hand side is obtained for  $t = x/K^2$ . If  $F$  is not bounded, the conclusion holds for  $F_n = \max(-n, \min(F, n))$ ,  $n \geq 0$ , and  $(F_n)_{n \in \mathbb{N}}$  converges  $\mathbb{P}$ -a.s. to  $F$ . Hence:

$$\mathbb{P}(F - \mathbb{E}[F|Z] \geq x|Z) \leq \exp\left(-\frac{x^2}{2K^2}\right) = \exp\left(-\frac{x^2}{2\sum_{a \in A} d_a^2}\right).$$

The proof is thus complete.  $\square$

**Remark 2.3.9.** The McDiarmid's inequality for conditionally independent random variables can be recovered from Azuma's inequality. Malliavin calculus offers an alternative for proof of concentration results to classical inequalities that use either or the martingale.  $\blacksquare$

## 2.4 Applications to normal approximation

We deal with the problem of approximations in law for functionals of the form:

$$F(X) := F(X_1, X_2, \dots). \quad (2.26)$$

When the functional only depends on a random vector  $(X_1, \dots, X_n)$  for a given  $n \in \mathbb{N}$ , the normal approximation of  $F(X_1, X_2, \dots, X_n)$  has been studied by Chatterjee (2008); Lachièze-Rey and Peccati (2017). More recently, those methods have been adapted to the infinite case when  $n \rightarrow \infty$  (Dung, 2018, 2019; Arras et al., 2019; Decreusefond and Halconruy, 2019; Privault and Serafin, 2020; Duerinckx, 2021; Privault and Serafin, 2022), explicitly stating quantitative limit theorems. They use Malliavin calculus for integration by parts in Stein's method. Earlier, Chen and Shao (2007) derived Berry-Esseen bound for nonlinear statistics such as U-statistics and L-statistics with an order of magnitude for the bounds which corresponds to the normalizing rate of central limit theorems. They used a concentration inequality approach by Stein's method as to control the non-linear part of the statistics. Still, the derived bounds were non-uniform except for particular cases, and the sequence requires to be independent identically distributed. The combination of Stein's method and Malliavin calculus leads to finer-grained upper bounds in probability distance for limit theorems of non-linear U-statistics, essentially relying on a *chaos expansion* of square-integrable functionals.

### 2.4.1 Bounds in probability distance

The goal is to bound for instance the 1-Wasserstein distance

$$d_W(\mathcal{L}(F(X)), \mathcal{L}(Y)) := \sup_{h \in \mathcal{H}} |\mathbb{E}[h(F(X))] - \mathbb{E}[h(Y)]|$$

for  $\mathcal{H}$  the set of 1-Lipschitz functions and  $Y$  the random variable following the target distribution. We recall the first component of Stein's method for normal approximation (see e.g. Chatterjee, 2008, Lemma 4.2).

**Lemma 2.4.1 — Normal approximation.** *Let  $L^\dagger h(x) := h'(x) - xh(x)$ . Then,*

$$d_W(\mathcal{L}(F(X)), \mathcal{N}(0, 1)) \leq \sup_{\varphi \in \mathcal{H}_*} \left| \mathbb{E}[L^\dagger \varphi(F(X))] \right|, \quad (2.27)$$

where  $\mathcal{H}_* := \{\varphi \in C^2(\mathbb{R}, \mathbb{R}) : \|\varphi'\|_\infty \leq \sqrt{\frac{2}{\pi}}, \|\varphi''\|_\infty \leq 2\}$ .

For notational convenience, we denote  $d_W(\mathcal{L}(F(X)), \mathcal{N}(0, 1))$  by  $d_W(F, \mathcal{N}(0, 1))$ . For sake of conciseness, we denote by  $\Delta^{\{a\}} F$  the quantity  $\Delta^{\{a\}} F(X, X'_a)$ .

### 2.4.2 Rates in Lyapunov's conditional central limit

**Lemma 2.4.2** For any  $F \in \mathcal{S}$  such that  $\mathbb{E}[F|Z] = 0$ . Then,

$$d_W(F, \mathcal{N}(0, 1)) \leq \sup_{\psi \in \mathcal{H}_*} \left| \mathbb{E} \left[ \sum_{a \in A} \psi(F(X^{\{a\}}, X'_a)) \Delta^{\{a\}'} F D_a(-\mathbb{L}^{-1}F) - \psi(F) \right] \right| + \sum_{a \in A} \mathbb{E}[(\Delta^{\{a\}'} F)^2 | D_a \mathbb{L}^{-1}F]. \quad (2.28)$$

*Proof of Proof of Lemma 2.4.2.* We compute:

$$\sup_{f^\dagger \in \mathcal{H}_*} |\mathbb{E}[F(f^\dagger)(F)] - (f^\dagger)'(F)|.$$

Since  $F$  is centered,

$$\begin{aligned} \mathbb{E}[F(f^\dagger)(F)] &= \mathbb{E}[\mathbb{L}(\mathbb{L}^{-1}F) f^\dagger(F)] \\ &= - \sum_{a \in A} \mathbb{E}[D_a \mathbb{L}^{-1}F D_a f^\dagger(F)] \text{ by integration by parts} \\ &= - \sum_{a \in A} \mathbb{E} \left[ D_a \mathbb{L}^{-1}F \mathbb{E} \left[ (f^\dagger)'(F) - f^\dagger(F^{\{a\}'}) | X, Z \right] \right] \\ &= - \sum_{a \in A} \mathbb{E}[D_a \mathbb{L}^{-1}F \Delta^{\{a\}'} f^\dagger(F)]. \end{aligned}$$

Then, we use the Taylor expansion taking the reference point to be  $F^{\{a\}'}$  instead of  $F$ , for all  $a \in A$  yielding:

$$\begin{aligned} \Delta^{\{a\}'} f^\dagger(F) &= f^\dagger(F) - f^\dagger(F^{\{a\}'}) \\ &= (f^\dagger)'(F^{\{a\}'}) \Delta^{\{a\}'} F + R_a, \end{aligned}$$

with  $|R_a| \leq \frac{\|(f^\dagger)''\|_\infty}{2} (\Delta^{\{a\}'} F)^2 = (\Delta^{\{a\}'} F)^2$ . Then,

$$\begin{aligned} |\mathbb{E}[F f^\dagger(F)] - (f^\dagger)'(F)| &\leq \left| \mathbb{E} \left[ \sum_{a \in A} \Delta^{\{a\}'} F (D_a(-\mathbb{L}^{-1}F)) \left( (f^\dagger)'(F^{\{a\}'}) - (f^\dagger)'(F) \right) \right] \right| \\ &\quad + \sum_{a \in A} \mathbb{E}[(\Delta^{\{a\}'} F)^2 | D_a \mathbb{L}^{-1}F]. \end{aligned}$$

Because  $(f^\dagger)''$  has Lipschitz-constant equal to 2, we get the result.  $\square$

We prove a quantitative Lyapunov's conditional central limit theorem for random variables with moments of order 3.

**Corollary 2.4.3** — **Lyapunov’s conditional central limit theorem.** *Let  $(X_n)_{n \in \mathbb{N}}$  be a sequence of thrice integrable, conditionally independent random variables given a latent random variable  $Z$ . Let us observe that*

$$\sigma_{j,Z}^2 = \text{Var}(X_j|Z), \quad s_{n,Z}^2 = \sum_{j=1}^n \sigma_{j,Z}^2 \quad \text{and} \quad \bar{X}_n = \frac{1}{s_{n,Z}} \sum_{j=1}^n (X_j - \mathbb{E}[X_j|Z]).$$

Then,

$$d_W(\bar{X}_n, \mathcal{N}(0, 1)) \leq 2(\sqrt{2} + 1) \mathbb{E} \left[ \frac{1}{s_{n,Z}^3} \sum_{i=1}^n |X_i - \mathbb{E}[X_i|Z]|^3 \right]. \quad (2.29)$$

The proof of the corollary follows the same steps as the one of [Decreusefond and Halconruy \(2019, Corollary 5.11\)](#), using [Lemma 2.4.2](#).

A version of conditional central limit theorem was first stated in [Rao \(2009\)](#) without proof and then was proved in [Grzenda and Zieba \(2008\)](#) for sequence of conditionally independent random variables (see [Yuan et al. \(2014\)](#); [Bulinski \(2017\)](#) for more "stringent" proofs). They also give a Lindebergh central limit theorem. We rule out a Lyapunov’s conditional central limit theorem, giving a quantitative limit theorem. We note that our approach differs from the one using Stein’s method for conditional central limit in [Dey and Terlov \(2023\)](#).

**Example 2.4.4** (Conditional Bernoulli random variables). Let  $(U_i)_{i \in \mathbb{N}}$  independent uniform random variables, and  $X_i = \mathbb{1}_{\{U_i \leq Z\}}$ , with  $Z$  an arbitrary random variable lying in  $[0, 1]$ , then  $(X_i)_{i \in \mathbb{N}}$  forms a sequence of conditionally independent random variables given  $Z$ . The law of  $\mathcal{L}(X_i|X^{\{i\}}, Z)$  is a Bernoulli law of parameter  $Z$ . We compute the right-hand side of the Lyapunov theorem in this case.

$$s_{n,Z}^2 = nZ(1 - Z)$$

$$\mathbb{E} [|X_i - \mathbb{E}[X_i|Z]|^3 | Z] = Z(1 - Z)(1 - 2Z).$$

Hence,

$$d_W(\bar{X}_n, \mathcal{N}(0, 1)) \leq 2(\sqrt{2} + 1) \mathbb{E} \left[ \frac{1 - 2Z + 2Z^2}{\sqrt{Z(1 - Z)}} \right] n^{-1/2}.$$

It is a classical quantitative theorem when one assumes mutual independence of the random variables. In some formulations, one supposes the random variables to have unit variance or normalization ([Reinert, 1998](#); [Chen et al., 2011](#)). There is quite a few papers investigating the lower constant before the absolute third moments in the upper bound for the Kolmogorov distance (more difficult to obtain usually because the solution of Stein equations has less bounded derivatives) and various other probability distances (see e.g. [Tyurin \(2009\)](#) and the references therein). One can consider the  $L^1$ -norm of the difference of cumulative function of  $\bar{X}_n$  and  $\mathcal{N}(0, 1)$  ([Goldstein, 2010](#)). Recently, with the same ambition to find general bounds for functional of independent random variables, [Privault and Serafin \(2020\)](#) derives a bound in Wasserstein-1 distance for sums of identically distributed independent random variables with a constant 2 which is lower than ours. We note that [Dung \(2018\)](#) uses the same gradient, but the constant here is better than 4 in his.

The strategy used in [Lemma 2.4.2](#) proves ineffective as to provide sharp bounds for a more

general class of functionals of random variables such as U-statistics.

Actually the integration by parts with carré du champ operator has been shown to be effective to bypass combinatorial difficulties with Malliavin derivatives when deriving fourth moment limit theorems.

### 2.4.3 Abstract bounds for U-statistics

The chaos decomposition has a natural interpretation as a decomposition in terms of degenerate U-statistics.

**Definition 2.4.5** (U-statistic (Hoeffding, 1948)). Let a family of measurable functions  $h_I : E_I \rightarrow \mathbb{R}$ . A U-statistic of degree (or order)  $p$  is defined for any  $n \geq p$  by:

$$U = \sum_{I \in (A, p)} h_I(X_I) = \sum_{I \in (A, p)} W_I.$$

**Definition 2.4.6** (Degenerate U-statistic). A degenerate U-statistic of order  $p > 1$  is a U-statistic of order  $p$  such that  $\mathbb{E} \left[ h_I(X_I^{\{a\}}, x_a) | Z \right] = 0$ , for all  $a \in A$  and  $x_a \in E_a$ .

The space of degenerate U-statistics is exactly  $\mathfrak{C}_p$ . Since we consider functionals given  $Z$  hereafter,  $h_I$  may be  $\sigma(Z)$ -measurable as well.

A convenient assumption in the proofs of quantitative limit theorems is the diffusiveness of the Markov generator at hand  $L$ , i.e. the associated carré du champ  $\Gamma_L$  satisfies for  $(F, G)$  in a dense algebra of  $\text{Dom } L$ :

$$\Gamma_L(\phi(F), G) = \phi'(F)\Gamma_L(F, G).$$

Due to the discreteness of the Malliavin structure, the operator  $L$  is not diffusive, but it is close to. We devise the pseudo chain rule below.

**Lemma 2.4.7 — First pseudo chain rule.** Let  $\psi \in C^1(\mathbb{R}, \mathbb{R})$ . Let  $G \in \mathcal{A}$  and  $F \in L^2(E_A)$  such that  $\psi(F) \in \mathcal{A}$ , then:

$$\Gamma(\psi(F), G) = \frac{1}{2} \sum_{a \in A} \psi'(F) \mathbb{E} \left[ (\Delta^{\{a\}'} F)(\Delta^{\{a\}'} G) | X, Z \right] + R_\psi(F, G), \quad (2.30)$$

where:

$$|R_\psi(F, G)| \leq \frac{\|\psi''\|_\infty}{4} \sum_{a \in A} \mathbb{E} \left[ |\Delta^{\{a\}'} G| (\Delta^{\{a\}'} F)^2 | X, Z \right].$$

*Proof of Lemma 2.4.7.* We write the Taylor expansion of  $\psi$ , and:

$$\begin{aligned} \mathbb{E} \left[ (\psi(F^{\{a\}'}) - \psi(F))(G^{\{a\}'} - G) | X, Z \right] &= \mathbb{E} \left[ \psi'(F)(\Delta^{\{a\}'} F)(\Delta^{\{a\}'} G) | X, Z \right] \\ &\quad + \mathbb{E} \left[ (G^{\{a\}'} - G)r_\psi(F, F^{\{a\}'} - F) | X, Z \right]. \end{aligned}$$

Then,

$$2\Gamma(\psi(F), G) = \psi'(F) \sum_{a \in A} \mathbb{E} \left[ (\Delta^{\{a\}'} F)(\Delta^{\{a\}'} G) \mid X, Z \right] \\ + \sum_{a \in A} \mathbb{E} \left[ (G^{\{a\}'} - G) r_\psi(F, F^{\{a\}'} - F) \mid X, Z \right]$$

where:

$$r_\psi(x, y) = \psi(x + y) - \psi(x) - \psi'(x)y = \int_0^y (y - s)\psi''(x + s) ds.$$

We note that  $r_\psi$  satisfies:

$$|r_\psi(x, y)| \leq \frac{\|\psi''\|_\infty}{2} y^2,$$

and we obtain the bound on the remainder.  $\square$

**Theorem 2.4.8 — Bounds in 1-Wasserstein distance.** *Assume that  $F \in L^3(E_A)$ , such that  $\mathbb{E}[F \mid Z] = 0$  and  $\mathbb{E}[F^2] = 1$ , then we get the bound:*

$$d_W(F, \mathcal{N}(0, 1)) \leq \sqrt{\frac{2}{\pi}} \mathbb{E} |\Gamma(F, -L^{-1}F) - 1| \\ + \frac{1}{2} \sum_{a \in A} \mathbb{E} [|\Delta^{\{a\}'} L^{-1}F| (\Delta^{\{a\}'} F)^2]. \quad (2.31)$$

Moreover, if  $F \in L^4(E_A)$ , then one has the further bound:

$$d_W(F, \mathcal{N}(0, 1)) \leq \sqrt{\frac{2}{\pi}} \sqrt{\text{Var}(\Gamma(F, L^{-1}F))} \\ + \frac{\sqrt{2}}{2} \sqrt{-\mathbb{E}[FLF]} \sqrt{\sum_{a \in A} \mathbb{E}[|\Delta^{\{a\}'} F|^4]}. \quad (2.32)$$

*Proof of Theorem 2.4.8.* We have:

$$\begin{aligned} \mathbb{E}[L^\dagger f^\dagger(F)] &= \mathbb{E}[F(f^\dagger)'(F) - (f^\dagger)''(F)] \\ &= \mathbb{E}[LL^{-1}F(f^\dagger)'(F)] - \mathbb{E}[(f^\dagger)''(F)] \\ &= \mathbb{E}[L^{-1}FL((f^\dagger)'(F))] - \mathbb{E}[(f^\dagger)''(F)] \\ &= \mathbb{E}[\Gamma(L^{-1}((f^\dagger)'(F)), -L^{-1}F)] - \mathbb{E}[(f^\dagger)''(F)] \end{aligned} \quad (2.33)$$

by integration by parts. We use [Lemma 2.4.7](#) and obtain that:

$$\mathbb{E}[\Gamma(L((f^\dagger)'(F)), -L^{-1}F)] \leq \mathbb{E}[(f^\dagger)''(F)\Gamma(F, -L^{-1}F)] + \mathbb{E}[R_{(f^\dagger)(3)}(F, -L^{-1}F)].$$

Thus,

$$\mathbb{E}[L^\dagger f^\dagger(F)] \leq \sqrt{\frac{2}{\pi}} \mathbb{E} |\Gamma(F, -L^{-1}F) - 1| + \frac{1}{2} \sum_{a \in A} \mathbb{E} [|\Delta^{\{a\}'} L^{-1}F| (\Delta^{\{a\}'} F)^2].$$

By Jensen's inequality for the first term and Cauchy-Schwarz inequality (for expectation of sum of random variables) for the second one, then by integration by parts, it yields:

$$\mathbb{E}[L^\dagger f^\dagger(F)] \leq \sqrt{\frac{2}{\pi}} \sqrt{\text{Var}(\Gamma(F, \mathbf{L}^{-1}F))} + \frac{1}{2} \sqrt{\sum_{a \in A} \mathbb{E}[|\Delta^{\{a\}'} \mathbf{L}^{-1}F|^2]} \sqrt{\sum_{a \in A} \mathbb{E}[(\Delta^{\{a\}'} F)^4]},$$

and the proof is complete.  $\square$

**Corollary 2.4.9** *If  $F = \sum_{p=1}^m F_p$  is four times integrable functional where  $F_p \in \ker(\mathbf{L} + p\text{Id})$ , then:*

$$\begin{aligned} d_W(F, \mathcal{N}(0, 1)) &\leq \sqrt{\frac{2}{\pi}} \sum_{p,q=1}^m \frac{1}{q} \sqrt{\text{Var}[\Gamma(F_p, F_q)]} \\ &\quad + \sqrt{2} \sum_{p=1}^m \frac{1}{p} \sqrt{\mathbb{E}[F_p^2]} \left\{ \sum_{p=1}^m p^{1/4} \left( \sum_{a \in A} \mathbb{E} \left| \Delta^{\{a\}'} F \right|^4 \right)^{1/4} \right\}^2. \end{aligned} \quad (2.34)$$

*Proof of Corollary 2.4.9.* We use the decomposition of  $\mathbf{L}^{-1}$  as to develop the first and second terms in (2.32). The final result is obtained after using Cauchy-Schwarz inequality.  $\square$

That is the starting point towards a partial fourth moment limit theorem.

**Remark 2.4.10.** With that approach, we cannot state a Berry-Esseen bound because of the lack of smoothness of the solutions of the Stein equation for the Kolmogorov distance.  $\blacksquare$

We turn to bounds in Kolmogorov distance which are based on the same computations. We recall the properties of the Stein equation for Kolmogorov bounds (Chen et al., 2011, Lemma 2.3).

**Lemma 2.4.11 — Inequalities for Kolmogorov test functions.** *Let  $z \in \mathbb{R}$ , the test functions  $f_z$  in the Stein equation for Kolmogorov distance is such that:*

$$\|f_z\|_\infty \leq \frac{\sqrt{2\pi}}{4}, \|f'_z\|_\infty \leq 1.$$

Moreover, for all  $u, v, w \in \mathbb{R}$ ,

$$|(w+u)f_z(w+u) - (w+v)f_z(w+u)| \leq \left( |w| + \frac{\sqrt{2\pi}}{4} \right) (|u| + |v|), \quad (2.35)$$

and the following local estimate holds for every  $x, h \in \mathbb{R}$ :

$$\begin{aligned} |f_z(x+h) - f_z(x) - hf'_z(x)| &\leq \frac{h^2}{2} \left( |x| + \frac{\sqrt{2\pi}}{4} \right) + h(\mathbf{1}_{\{[x, x+h]\}}(z) - \mathbf{1}_{\{x+h, x\}}(z)) \\ &= \frac{h^2}{2} \left( |x| + \frac{\sqrt{2\pi}}{4} \right) + |h|(\mathbf{1}_{\{[x, x+h]\}}(z) + \mathbf{1}_{\{x+h, x\}}(z)). \end{aligned} \quad (2.36)$$



**Proposition 2.4.12 — Kolmogorov bounds.** *Let  $F \in L^4(E_A)$ , then one has the bound:*

$$\begin{aligned}
d_{Kol}(F, \mathcal{N}(0, 1)) &\leq \mathbb{E}|\Gamma(F, -\mathbb{L}^{-1}F) - 1| + \frac{\sqrt{2\pi}}{16} \sum_{a \in A} \mathbb{E} \left[ |\Delta^{\{a\}'} \mathbb{L}^{-1}F| (F - F^{\{a\}'})^2 \right] \\
&\quad + \frac{1}{4} \sum_{a \in A} \mathbb{E} \left[ |\Delta^{\{a\}'} \mathbb{L}^{-1}F| |F| (\Delta^{\{a\}'} F)^2 \right] \\
&\quad + \frac{1}{2} \sup_{z \in \mathbb{R}} \sum_{a \in A} \mathbb{E} \left[ |\Delta^{\{a\}'} \mathbb{L}^{-1}F| (\Delta^{\{a\}'} F) \Delta^{\{a\}'} \mathbb{1}_{\{F > z\}} \right]. \tag{2.37}
\end{aligned}$$

*Proof of Proposition 2.4.12.* Using (2.36) in the case  $x = F$ , and  $h = -\Delta^{\{a\}'} F$ , one sees that, for every  $a \in A$ ,

$$\begin{aligned}
|f_z(F^{\{a\}'}) - f_z(F) - (F^{\{a\}'} - F)f'_z(F)| &\leq \frac{(F - F^{\{a\}'})^2}{2} \left( |F| + \frac{\sqrt{2\pi}}{4} \right) \\
&\quad + |F - F^{\{a\}'}| (\mathbb{1}_{\{F, F^{\{a\}'}\}}(z) - \mathbb{1}_{\{F^{\{a\}'}, F\}}(z)) \\
&\leq \frac{(F - F^{\{a\}'})^2}{2} \left( |F| + \frac{\sqrt{2\pi}}{4} \right) \\
&\quad + (F^{\{a\}'} - F) (\mathbb{1}_{\{F^{\{a\}'} > z\}} - \mathbb{1}_{\{F > z\}}).
\end{aligned}$$

Hence, plugging that into the Stein equation:

$$\begin{aligned}
\mathbb{E}[L^\dagger f^\dagger(F)] &= \mathbb{E}[F(f^\dagger)'(F) - (f^\dagger)''(F)] \\
&= \mathbb{E}[F f_z(F) - f'_z(F)]. \\
&= \mathbb{E}[F \mathbb{L} \mathbb{L}^{-1}(f_z(F))] - \mathbb{E}[f'_z(F)] \\
&= \mathbb{E}[\Gamma(-\mathbb{L}^{-1}F, f_z(F))] - \mathbb{E}[f'_z(F)] \text{ using the carré du champ integration by parts.}
\end{aligned}$$

The integration by parts allows us to express  $\mathbb{E}[L^\dagger f^\dagger(F)]$  without worrying about integration by  $\mathbb{E}'$ , instead of writing  $\mathbb{L} = -\delta D$ . Once more, it proves useful as to simplifying the bounds.

$$\begin{aligned}
2\mathbb{E}[F f_z(F)] &= \mathbb{E} \left[ \sum_{a \in A} (\mathbb{L}^{-1}F^{\{a\}'} - \mathbb{L}^{-1}F) [f_z(F) - f_z(F^{\{a\}'})] \right] \text{ because of (2.18)} \\
&\leq \sum_{a \in A} \mathbb{E} \left[ |\mathbb{L}^{-1}F - \mathbb{L}^{-1}F^{\{a\}'}| (F - F^{\{a\}'}) f'_z(F) \right] \\
&\quad + \mathbb{E} \left[ |\mathbb{L}^{-1}F - \mathbb{L}^{-1}F^{\{a\}'}| \frac{(F - F^{\{a\}'})^2}{2} \left( |F| + \frac{\sqrt{2\pi}}{4} \right) \right] \\
&\quad + \mathbb{E} \left[ |\mathbb{L}^{-1}F - \mathbb{L}^{-1}F^{\{a\}'}| \Delta^{\{a\}'} F (\mathbb{1}_{\{F > z\}} - \mathbb{1}_{\{F^{\{a\}'} > z\}}) \right] \\
&\leq 2\mathbb{E} [\Gamma(-\mathbb{L}^{-1}F, F) f'_z(F)] + \mathbb{E} \left[ |\mathbb{L}^{-1}F - \mathbb{L}^{-1}F^{\{a\}'}| \left( \frac{|F|}{2} + \frac{\sqrt{2\pi}}{8} \right) (F - F^{\{a\}'})^2 \right] \\
&\quad + \sup_{z \in \mathbb{R}} \sum_{a \in A} \mathbb{E} \left[ |\Delta^{\{a\}'} \mathbb{L}^{-1}F| \Delta^{\{a\}'} F \Delta^{\{a\}'} \mathbb{1}_{\{F > z\}} \right].
\end{aligned}$$

□

**Remark 2.4.13.** Given elements of a Malliavin calculus, [Döbler and Peccati](#); [Privault and Serafin](#) resort to the Skorohod isometry formula which is related to the Poincaré inequality as to bound the extra remainder term which is a supremum. In our case, the other terms can be bounded with similar strategies as for the Wasserstein metric using the exchangeability of  $F$  and  $F^{\{a\}'}$ , but the strategy for the extra remainder term was not fruitful. ■

Because of that, we focus exclusively on the Wasserstein-1 distance in the remainder of this chapter.

## 2.5 Partial fourth moment theorems

We adapt the main proofs of [Azmoodeh et al. \(2014\)](#), requiring a second pseudo chain rule that expresses the carré du champ operator as an approximation of a derivation operator in its two arguments.

**Lemma 2.5.1 — Second pseudo chain rule.** *Let  $\varphi, \psi$  be twice differentiable functions such that their second derivative is bounded Lipschitz-continuous. Assume that  $F$  a four times integrable functional such that  $\varphi(F) \in \mathcal{A}$ ,  $F \in \mathcal{A}$  and  $\mathbb{E}[F|Z] = 0$ , then one has:*

$$\begin{aligned} \Gamma(\varphi(F), \psi(F)) &= (\varphi'\psi')(F)\Gamma(F, F) \\ &\quad - \frac{1}{4}(\varphi''\psi' + \varphi'\psi'')(F) \sum_{a \in A} \mathbb{E} \left[ (\Delta^{\{a\}'} F)^3 | X, Z \right] + \sum_{a \in A} R_a, \end{aligned} \quad (2.38)$$

with:

$$R_a = \frac{1}{2} \left( \mathbb{E} \left[ R_{a, \varphi\psi}^{(4)}(F) | X, Z \right] - \varphi(F) \mathbb{E} \left[ R_{a, \psi}^{(4)}(F) | X, Z \right] - \psi(F) \mathbb{E} \left[ R_{a, \varphi}^{(4)}(F) | X, Z \right] \right),$$

and:

$$R_{a, \psi}^{(4)} \leq \frac{\|\psi^{(4)}\|_\infty}{24} \mathbb{E} \left[ (\Delta^{\{a\}'} F)^4 | X, Z \right] \text{ for any } \psi \text{ fourth times differentiable.}$$

*Proof of Lemma 2.5.1.* We have:

$$\begin{aligned} 2\Gamma(\varphi(F), \psi(F)) &= 2\varphi'(F)\psi'(F)\Gamma(F, F) - \frac{3}{6}(\varphi''\psi' + \varphi'\psi'')(F) \sum_{a \in A} \mathbb{E} \left[ (\Delta^{\{a\}'} F)^3 | X, Z \right] \\ &\quad + \sum_{a \in A} \mathbb{E} \left[ R_{a, \varphi\psi}^{(4)}(F) - \varphi(F)R_{a, \psi}^{(4)}(F) - \psi(F)R_{a, \varphi}^{(4)}(F) | X, Z \right], \end{aligned} \quad (2.39)$$

with:

$$R_{a, \phi}^{(4)} = \frac{1}{6} \mathbb{E} \left[ \int_F^{F^{\{a\}'}} \phi^{(4)}(x)(x - F)^4 dx | X, Z \right],$$

for  $\phi$  a four times differentiable function. □

We focus on functionals in the  $p$ -th chaos for  $p > 0$ , as to obtain such kind of bound:

$$\text{Var}[\Gamma(F, F)] \leq C(\mathbb{E}[F^4] - 3\mathbb{E}[F^2]^2) + \text{remainder.}$$

**Lemma 2.5.2** Let  $G \in \bigoplus_{k=0}^q \mathfrak{C}_k$ . Then for any  $\eta \geq q$ ,

$$\mathbb{E}[G(\mathbf{L} + \eta \text{Id})^2 G] \leq \eta \mathbb{E}[G(\mathbf{L} + \eta \text{Id})G] \leq c \mathbb{E}[G(\mathbf{L} + \eta \text{Id})^2 G], \quad (2.40)$$

where

$$c = \frac{1}{\eta - q} \wedge 1.$$

*Proof of Lemma 2.5.2.* Since  $G \in \bigoplus_{k=0}^q \mathfrak{C}_k$ , we write

$$G = \sum_{k=0}^q \pi_k(G) \text{ and } \mathbf{L}G = - \sum_{k=0}^q k \pi_k(G). \quad (2.41)$$

It follows that

$$\begin{aligned} \mathbb{E}[G(\mathbf{L} + \eta \text{Id})^2 G] &= \mathbb{E}[G\mathbf{L}(\mathbf{L} + \eta \text{Id})G] + \eta \mathbb{E}[G(\mathbf{L} + \eta \text{Id})G] \\ &= \mathbb{E}\left[G \sum_{k=0}^q k(k - \eta) \pi_k(G)\right] + \eta \mathbb{E}[G(\mathbf{L} + \eta \text{Id})G]. \end{aligned}$$

By orthogonality of the chaos,

$$\mathbb{E}\left[G \sum_{k=0}^q k(k - \eta) \pi_k(G)\right] = -\mathbb{E}\left[\sum_{k=0}^q k(\eta - k) \pi_k(G)^2\right] \leq 0,$$

and the inequality holds in view on the assumption on  $\eta$ . In the same vein,

$$\begin{aligned} \mathbb{E}[G(\mathbf{L} + \eta \text{Id})G] &= \sum_{k=0}^q (\eta - k) \mathbb{E}[\pi_k(G)^2] \\ &\leq c \sum_{k=0}^q (\eta - k)^2 \mathbb{E}[\pi_k(G)^2] \\ &= c \mathbb{E}[G(\mathbf{L} + \eta \text{Id})^2 G]. \end{aligned}$$

Thus, it yields the result.  $\square$

**Lemma 2.5.3** For  $F \in \mathfrak{C}_p \cap L^4(E_A)$  and  $Q$  a polynomial of degree two and  $a > 0$ ,

$$\mathbb{E}[Q(F)(\mathbf{L} + a p \text{Id})Q(F)] = p \mathbb{E}\left[aQ^2(F) - \frac{Q'(F)F}{3Q''(F)}\right] - \mathbb{E}[R_Q(F)], \quad (2.42)$$

where  $R_Q$  is a remainder term that depends on  $Q$ . For  $Q = H_2 = X^2 - 1$  the second Hermite polynomial, the remainder reads off:

$$\mathbb{E}[R_Q] = \mathbb{E}[R_{H_2}] = \frac{1}{6} \mathbb{E}\left[\sum_{\alpha \in A} |\Delta^{\{a\}'} F|^4\right]. \quad (2.43)$$

*Proof of Lemma 2.5.3.* We first integrate by parts, then use the pseudo chain rule of Lemma 2.5.1:

$$\begin{aligned}
 \mathbb{E}[Q(F)\mathbf{L}Q(F)] &= -\mathbb{E}[\Gamma(Q(F), Q(F))] \\
 &= -\mathbb{E}[Q'(F)^2\Gamma(F, F)] \\
 &\quad + \frac{1}{6}(Q^2)^{(3)}(F) \sum_{a \in A} \mathbb{E} \left[ (\Delta^{\{a\}'} F)^3 | X, Z \right] \\
 &\quad - \frac{1}{2} \sum_{a \in A} \mathbb{E} \left[ \mathbb{E} \left[ R_{a, Q^2}^{(4)}(F) | X, Z \right] - 2Q(F) \mathbb{E} \left[ R_{a, Q}^{(4)}(F) | X, Z \right] \right].
 \end{aligned} \tag{2.44}$$

Since  $Q^{(3)} = 0$ , we have:

$$\begin{aligned}
 \mathbb{E}[Q(F)\mathbf{L}Q(F)] &= -\mathbb{E} \left[ [Q'(F)^2\Gamma(F, F)] \right. \\
 &\quad \left. + \frac{1}{6} \mathbb{E} \left[ (Q^2)^{(3)}(F) \sum_{a \in A} \mathbb{E} \left[ (\Delta^{\{a\}'} F)^3 | X, Z \right] \right] \right. \\
 &\quad \left. - \frac{1}{2} \sum_{a \in A} \mathbb{E} \left[ \mathbb{E} \left[ R_{a, Q^2}^{(4)}(F) | X, Z \right] \right] \right].
 \end{aligned} \tag{2.45}$$

Moreover,

$$\left( \frac{Q'(F)^3}{3Q''(F)} \right)' = \frac{3Q'(F)Q''(F)^2}{3Q''(F)^2} = Q'(F)^2. \tag{2.46}$$

Subsequently, we use the pseudo chain rule of Lemma 2.5.1 taking  $\psi = \text{Id}$  and  $\varphi = \frac{Q'(\cdot)^3}{3Q''(\cdot)}$ :

$$\begin{aligned}
 \mathbb{E}[Q'(F)^2\Gamma(F, F)] &= \mathbb{E} \left[ \Gamma \left( \frac{Q'(F)^3}{3Q''(F)}, F \right) \right] \\
 &\quad + \frac{1}{4} \mathbb{E} \left[ (\varphi''\psi' + \varphi'\psi'')(F) \sum_{a \in A} \mathbb{E} \left[ (\Delta^{\{a\}'} F)^3 | X, Z \right] \right] \\
 &\quad - \sum_{a \in A} \mathbb{E} \left[ \mathbb{E} \left[ R_{a, \varphi\psi}^{(4)}(F) | X, Z \right] - \varphi(F) \mathbb{E} \left[ R_{a, \psi}^{(4)}(F) | X, Z \right] \right] \\
 &\quad - \mathbb{E} \left[ F \mathbb{E} \left[ R_{a, \varphi}^{(4)}(F) | X, Z \right] \right] \\
 &= \mathbb{E} \left[ \Gamma \left( \frac{Q'(F)^3}{3Q''(F)}, F \right) \right] + \frac{1}{4} \mathbb{E} \left[ (Q'(\cdot)^2)'(F) \sum_{a \in A} (\Delta^{\{a\}'} F)^3 \right] \\
 &\quad - \sum_{a \in A} \frac{1}{2} \mathbb{E} \left[ R_{a, \varphi\psi}^{(4)}(F) - F R_{a, \varphi}^{(4)}(F) \right].
 \end{aligned} \tag{2.47}$$

Finally,

$$\begin{aligned}
\mathbb{E}[Q(F)\mathbf{L}Q(F)] &= -\mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)}, F\right)\right] \\
&\quad + \mathbb{E}\left[\left(\frac{1}{4}(Q'(\cdot)^2)'(F) - \frac{1}{12}(Q^2)^{(3)}(F)\right) \sum_{a \in A} (\Delta^{\{a\}'} F)^3\right] \\
&\quad + \frac{1}{2} \sum_{a \in A} \mathbb{E}\left[R_{a, \varphi\psi}^{(4)}(F) - R_{a, Q^2}^{(4)}(F) - FR_{a, \varphi}^{(4)}(F)\right] \\
&= -\mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)}, F\right)\right] \\
&\quad + \mathbb{E}\left[\left(\frac{1}{4}(Q'(\cdot)^2)'(F) - \frac{1}{12}(Q^2)^{(3)}(F)\right) \sum_{a \in A} (\Delta^{\{a\}'} F)^3\right] \\
&\quad + \frac{1}{2} \sum_{a \in A} \mathbb{E}\left[R_{a, \varphi\psi}^{(4)}(F) - R_{a, Q^2}^{(4)}(F)\right].
\end{aligned} \tag{2.48}$$

Because  $F \in \mathfrak{C}_p$ , we have:  $-\mathbb{E}\left[\Gamma\left(\frac{Q'(F)^3}{3Q''(F)}, F\right)\right] = \mathbb{E}\left[\frac{Q'(F)^3}{3Q''(F)}\mathbf{L}F\right] = -p\mathbb{E}\left[\frac{Q'(F)^3}{3Q''(F)}F\right]$ . For  $Q = H_2 = X^2 - 1$  the second Hermite polynomial,

$$\frac{Q'(F)^3}{3Q''(F)} = \frac{4}{3}X^3,$$

so  $\left(\frac{Q'(\cdot)^3}{3Q''(\cdot)}\right)^{(4)} = 32$  and  $(Q^2)^{(4)} = 24$ . Thus,

$$\sum_{a \in A} \mathbb{E}\left[R_{a, \varphi\psi}^{(4)}(F) - R_{a, Q^2}^{(4)}(F)\right] = \frac{(32 - 24)}{24} \sum_{a \in A} \mathbb{E}\left[|\Delta^{\{a\}'} F|^4\right]. \tag{2.49}$$

Since  $(Q'(\cdot)^2)'(F) = 8F$ , and  $(Q^2)^{(3)}(F) = 24F$ , the result follows.  $\square$

The assumption under which a fourth moment theorem holds, is that  $F \in \mathfrak{C}_p$  is a chaos eigenfunction with respect to the Markov generator  $\mathbf{L}$  i.e.:

$$F^2 \in \bigoplus_{k=0}^{2p} \mathfrak{C}_k. \tag{EGF}$$

It is analog to the one in [Ledoux \(2012\)](#); [Azmoodeh et al. \(2014\)](#). We show that it holds for an important class of U-statistics, homogeneous sums. We shall use the notation  $(A, p)$  that stands for the set of  $p$ -tuples of distinct elements of  $A$ .

**Example 2.5.4** (Homogeneous sum of conditionally independent random variables). For  $p > 0$ , if there exists  $(a_I)_{I \subset A} \in \mathbb{R}^{\mathcal{P}(A)}$  such that

$$W = \sum_{k=1}^p \sum_{I \in (A,k)} a_I \prod_{i \in I} X_i, \quad (2.50)$$

then

1.  $W$  is square-integrable homogeneous sum of order  $p$  if  $X_i$  are  $2p$ -integrable. In that case,  $W \in \mathcal{S}$ .
2. Its conditional expectation given  $Z$

$$\mathbb{E}[W|Z] = \sum_{k=1}^p \sum_{I \in (A,k)} a_I \prod_{i \in I} \mathbb{E}[X_i|Z]$$

is a homogeneous sum of random variables  $\hat{X}_i = \mathbb{E}[X_i|Z]$  for  $i \in I$  with  $I \in (A, k)$  for  $k \leq p$ .

Remark that  $(a_I)_{I \subset A}$  may be a sequence of random variables, in which case there exists a family of functions  $(g_I)_{I \subset A}$  such that  $a_I = g_I(Z)$ .

**Lemma 2.5.5** *Let  $W$  a homogeneous sums of conditionally independent random variables given  $Z$ . Then (EGF) holds.*

*Proof of Lemma 2.5.5.* Let  $W_I$  represent the component of  $F$  in (2.50) proportional to  $\prod_{\alpha \in I} X_\alpha$ . We want to prove that there exist  $G_1, \dots, G_{2p}$  with  $G_i \in \mathfrak{C}_i \cup \{0\}$  such that  $W_I W_J = \sum_{i=1}^{2p} G_i$ . Note that if  $I \cap J = \emptyset$ , and  $a \in I$ , then  $a$  is not in  $J$  and vice versa. Therefore,  $W_I W_J \in \mathfrak{C}_{|I|+|J|}$ . In general,

$$\begin{aligned} W_I W_J &\propto \prod_{\alpha \in I} Y_\alpha \prod_{\beta \in J} Y_\beta \\ &= \prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in I \cap J} Y_\delta^2 \\ &= \prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in I \cap J} (Y_\delta^2 - \mathbb{E}[Y_\delta^2|Z] + \mathbb{E}[Y_\delta^2|Z]) \\ &= \sum_{K \subset I \cap J} \prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in K} (Y_\delta^2 - \mathbb{E}[Y_\delta^2|Z]) \prod_{\delta \in (I \cap J) \setminus K} \mathbb{E}[Y_\delta^2|Z]. \end{aligned}$$

For  $a \in A$ :

$$\begin{aligned} &\mathbb{E} \left[ \prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in K} (Y_\delta^2 - \mathbb{E}[Y_\delta^2|Z]) \prod_{\delta \in (I \cap J) \setminus K} \mathbb{E}[Y_\delta^2|Z] \middle| \mathcal{G}_a^Z \right] \\ &= \begin{cases} 0 & \text{if } a \in K \cup ((I \setminus J) \cup (J \setminus I)) \\ \prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in K} (Y_\delta^2 - \mathbb{E}[Y_\delta^2|Z]) \prod_{\delta \in (I \cap J) \setminus K} \mathbb{E}[Y_\delta^2|Z] & \text{otherwise.} \end{cases} \end{aligned}$$

Hence, we get

$$\prod_{\gamma \in (I \setminus J) \cup (J \setminus I)} Y_\gamma \prod_{\delta \in K} (Y_\delta^2 - \mathbb{E}[Y_\delta^2 | Z]) \prod_{\delta \in (I \cap J) \setminus K} \mathbb{E}[Y_\delta^2 | Z] \in \mathfrak{C}_{|K \cup ((I \setminus J) \cup (J \setminus I))|}$$

with  $|K \cup ((I \setminus J) \cap (J \setminus I))| \leq |I \cup J| \leq 2p$ . Thus, (EGF) holds.  $\square$

**Proposition 2.5.6** For  $F \in \mathfrak{C}_p \cap L^2(E_A)$  such that  $\mathbb{E}[F^2] = 1$  and (EGF) holds, one has:

$$\mathbb{E}[(\Gamma(F, F) - p)^2] \leq \frac{p^2}{3} |\mathbb{E}[F^4] - 3| + \frac{p}{12} \mathbb{E} \left[ \sum_{a \in A} |\Delta^{\{a\}'} F|^4 \right]. \quad (2.51)$$

*Proof of Proposition 2.5.6.* By the very definition of  $\Gamma$ , one has:

$$\begin{aligned} \Gamma(F, F) - p &= \frac{1}{2} \mathbb{L}(F^2) - F \mathbb{L}F - p = \frac{1}{2} \mathbb{L}(F^2) + pF^2 - p \text{ for } F \in \mathfrak{C}_p \\ &= \frac{1}{2} (\mathbb{L} + 2p\text{Id})(F^2 - 1). \end{aligned}$$

It follows that:

$$\mathbb{E}[(\Gamma(F, F) - p)^2] = \frac{1}{4} \mathbb{E}[(\mathbb{L} + 2p\text{Id})(F^2 - 1))^2].$$

Since  $\mathbb{L}$  is a self-adjoint operator, this yields:

$$\mathbb{E}[(\Gamma(F, F) - p)^2] = \frac{1}{4} \mathbb{E}[H_2(F)(\mathbb{L} + 2p\text{Id})^2 H_2(F)].$$

As (EGF) holds, we are in position to apply Lemma 2.5.2 with  $q = 2p$  and  $\eta = 2p$ :

$$\mathbb{E}[(\Gamma(F, F) - p)^2] \leq \frac{p}{2} \mathbb{E}[H_2(F)(\mathbb{L} + 2p\text{Id})H_2(F)]. \quad (2.52)$$

According to Lemma 2.5.3, with  $a = 2$ ,

$$\begin{aligned} \frac{p}{2} \mathbb{E}[H_2(F)(\mathbb{L} + 2p\text{Id})H_2(F)] &= \frac{p^2}{2} \mathbb{E} \left[ 2(F^2 - 1)^2 - \frac{4}{3} F^4 \right] + \frac{p}{2} \mathbb{E}[R_{H_2}(F)] \\ &= \frac{p^2}{6} \mathbb{E} [6(F^2 - 1)^2 - 4F^4] + \frac{p}{2} \mathbb{E}[R_{H_2}] \\ &= \frac{p^2}{3} \mathbb{E}[F^4 - 6F^2 + 3] + \frac{p}{2} \mathbb{E}[R_{H_2}]. \end{aligned}$$

Thus, it yields

$$\mathbb{E}[(\Gamma(F, F) - p)^2] \leq \frac{p^2}{3} |\mathbb{E}[F^4 - 6F^2 + 3]| + \frac{p}{2} |\mathbb{E}[R_{H_2}]|, \quad (2.53)$$

and the proof is complete, using again Lemma 2.5.3.  $\square$

The remainder is also a fourth moment term.

Many papers are devoted to find the optimal conditions for the asymptotic normality of U-statistics of independent random variables. The criterion established in De Jong (1990) is related to the fourth moment phenomenon. The extra assumption is a negligibility condition also known

as the Lindeberg-Feller condition. Fix  $A_m$  a finite subset of cardinal  $m$  such that  $F = F(X_{A_m})$  and  $\mathbb{E}[F^2] = 1$ , that means:

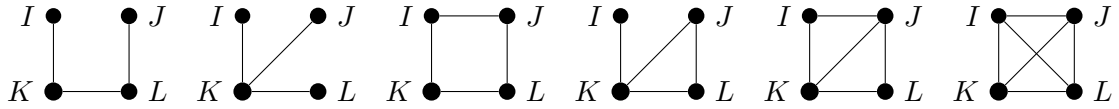
$$\rho_{A_m}^2 = \max_{i \in A_m} \sum_{I \ni i, I \subseteq A_m, |I|=p} \mathbb{E}[W_I^2] \xrightarrow{m \rightarrow +\infty} 0. \quad (2.54)$$

In some papers (Döbler and Peccati, 2017), the term  $\rho_{A_m}$  is called maximal influence of the random variables on the total variance of the degenerate U-statistics  $F$ . Subsequently, we shall denote it by  $\rho$ . The condition (2.54) is not necessary for asymptotic normality to hold, but there exist counterexamples for which the sequence of fourth cumulants of functionals of independent Rademacher random variables converges to 0 while (2.54) does not hold (see Döbler and Peccati, 2019).

We show that the quantity is related to the remainder above.

**Definition 2.5.7** (Connectedness of subsets). The  $r$ -tuple  $(I_1, \dots, I_r)$  subsets of  $A$  is connected if the intersection graph of  $\{I_1, \dots, I_r\}$  is connected, i.e. the graph  $G$  with vertex set  $\{I_1, \dots, I_r\}$  and edge set  $E(G) = \{\{I_i, I_j\} \mid i \neq j, I_i \cap I_j \neq \emptyset\}$  is connected.

In the case where  $r = 4$ , there are exactly six simple connected graphs with only four vertices (up to isomorphisms). There are listed in the figure below. The edges mean that the intersection between two vertices is non-empty.



In that order, they are respectively: linear (4-path), 4-star, square, kite (tadpole), diag (diamond), and 4-complete graph  $K_4$ .

**Lemma 2.5.8** If  $F \in \mathfrak{C}_p \cap L^4(E_A)$ , then:

$$\sum_{a \in A} \mathbb{E}[|\Delta^{\{a\}'} F|^4] \leq 16p \sum_{(I,J,K,L) \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|. \quad (2.55)$$

Moreover, assuming the hypercontractivity condition, i.e.

$$\sup_{J \in (A,p)} \frac{\mathbb{E}[W_J^4]}{\mathbb{E}[W_J^2]^2} < +\infty, \quad (\text{HC})$$

there exists a constant  $c_p$  that depends only on  $p$  such that:

$$\sum_{a \in A} \mathbb{E}[|\Delta^{\{a\}'} F|^4] \leq c_p \rho^2. \quad (2.56)$$



*Proof of Lemma 2.5.8.* Because  $(a + b)^4 \leq 8(a^4 + b^4)$ , one has:

$$\begin{aligned}
\sum_{a \in A} \mathbb{E} \left[ |\Delta^{\{a\}'} F|^4 \right] &\leq 8 \sum_{a \in A} \mathbb{E} \left[ \left( \sum_{I \ni a, |I| \leq p} W_I^{\{a\}'} \right)^4 + \left( \sum_{I \ni a, |I| \leq p} W_I \right)^4 \right] \\
&= 16 \sum_{a \in A} \mathbb{E} \left[ \left( \sum_{I \ni a, |I| \leq p} W_I \right)^4 \right] \\
&\leq 16 \sum_{I \cap J \cap K \cap L \neq \emptyset} |I \cap J \cap K \cap L| \mathbb{E}[W_I W_J W_K W_L] \\
&\leq 16p \sum_{I \cap J \cap K \cap L \neq \emptyset} |\mathbb{E}[W_I W_J W_K W_L]| \\
&\leq 16p \sum_{I, J, K, L \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|.
\end{aligned}$$

Then, we bound it by the maximal influence, using the generalized Hölder inequality:

$$\begin{aligned}
|\mathbb{E}[W_I W_J W_K W_L]| &\leq (\mathbb{E}[W_I^4] \mathbb{E}[W_J^4] \mathbb{E}[W_K^4] \mathbb{E}[W_L^4])^{1/4} \\
&\leq \max_{J \in A, |J|=p} \frac{\mathbb{E}[W_J^4]}{\mathbb{E}[W_J^2]^2} (\mathbb{E}[W_I^2]^2 \mathbb{E}[W_J^2]^2 \mathbb{E}[W_K^2]^2 \mathbb{E}[W_L^2]^2)^{1/4}
\end{aligned}$$

with  $\sigma_I^2 = \mathbb{E}[W_I^2]$ . Then the proposition 2.9 of Döbler and Peccati (2017) can be extended for functionals of conditionally independent random variables and implies that:

$$\sum_{I \cap J \cap K \cap L \neq \emptyset} \sigma_I \sigma_J \sigma_K \sigma_L \leq C_p \rho^2,$$

where the finite constant  $C_p$  only depends on  $p$ . It yields the existence of  $c_p > 0$  such that the inequality (2.56) holds true.  $\square$

We are now in position to state a partial fourth moment limit theorem.

**Theorem 2.5.9 — Quantitative De Jong’s limit theorem I.** *Let  $F \in L^4(E_A)$  a degenerate  $U$ -statistics of order  $p$  of conditionally independent random variables such that  $\mathbb{E}[F|Z] = 0$  and  $\mathbb{E}[F^2] = 1$ . If we suppose the hypercontractivity condition (HC) and the assumption (EGF), then one has the bound:*

$$d_W(F, \mathcal{N}(0, 1)) \leq \sqrt{\frac{2}{3\pi}} \sqrt{|\mathbb{E}[F^4] - 3|} + \tilde{C}_p \rho, \tag{2.57}$$

with  $\tilde{C}_p$  a positive constant that only depends on  $p$ .

*Proof.* By Corollary 2.4.9,

$$d_W(F, \mathcal{N}(0, 1)) \leq \sqrt{\frac{2}{\pi}} \frac{1}{p} \sqrt{\text{Var}[\Gamma(F, F)]} + \sqrt{2} \sqrt{\mathbb{E}[F^2]} \left( \sum_{a \in A} \mathbb{E} \left[ |\Delta^{\{a\}'} F|^4 \right] \right)^{1/2}.$$

The combination of (2.51) and Lemma 2.5.8 yields the final upper bound.  $\square$

The upper bound of the remainder expressed in terms of maximal influence is not used in the subsequent applications, so we drop the (HC) condition.

Similar results can be found in Döbler and Peccati (2017); Shao and Zhang (2019); Fang and Koike (2022) for independent random variables.

A related result to the fourth moment phenomenon appears in De Jong (1996) in the particular case that is the topic of the next chapter. We prove the associated quantitative statement for functionals of conditionally independent random variables. We prepare the proof with the proposition below.

**Proposition 2.5.10** *If  $F = \sum_{p=1}^m F_p$  where  $F_p = \sum_{|I|=p} W_I \in \mathfrak{C}_p$ , assuming there exists  $C \in \mathbb{R}^+$  such that for all  $I, J \subset A$ , and  $a \in A$ , that*

$$\frac{\mathbb{E}[W_I W_J | \mathcal{G}^a]}{W_{I \setminus \{a\}} W_{J \setminus \{a\}}} < C \quad \mathbb{P}\text{-a.s.}, \quad (\text{H1})$$

then for  $p \neq q$ :

$$\sqrt{\text{Var}[\Gamma(F_p, F_q)]} \lesssim \sqrt{\sum_{(I, J, K, L) \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|}, \quad (2.58)$$

for  $I, J, K, L$  sets of size less than  $\max(p, q)$ .

*Proof of Proposition 2.5.10.* The carré du champ reads for  $p \neq q$ :

$$\begin{aligned} \Gamma(F_p, F_q) &= \Gamma\left(\sum_{|I|=p} W_I, \sum_{|J|=q} W_J\right) \\ &= \sum_{|I|, |J|=p, q} \Gamma(W_I, W_J). \end{aligned}$$

Hence,

$$\begin{aligned} 2\Gamma(F_p, F_q) &= \sum_{|I|, |J|=p, q} (\mathbb{L}(W_I W_J) + (p+q)W_I W_J) \\ &= \sum_{|I|, |J|=p, q} \left( (p+q)W_I W_J - \sum_{a \in A} D_a(W_I W_J) \right) \\ &= \sum_{|I|, |J|=p, q} \left( (p+q)W_I W_J - \sum_{a \in I \cup J} D_a(W_I W_J) \right) \\ &= (p+q) \sum_{\substack{|I|, |J|=p, q \\ I \cap J = \emptyset}} W_I W_J + \sum_{\substack{|I|, |J|=p, q \\ I \cap J \neq \emptyset}} (|I| + |J| - |I \cup J|) W_I W_J \\ &\quad + \sum_{a \in I \cup J} \mathbb{E}[W_I W_J | \mathcal{G}_a]. \end{aligned}$$

Because of the spectral decomposition,  $\mathbb{E}[W_I | \mathcal{G}_a] = 0$  for  $a \in I$ . Let  $J$  such that  $a \notin J$ , then  $\mathbb{E}[W_I W_J | \mathcal{G}_a] = W_J \mathbb{E}[W_I | \mathcal{G}_a] = 0$ .

$$2\Gamma(F_p, F_q) = (p+q) \sum_{\substack{|I|, |J|=p, q \\ I \cap J = \emptyset}} W_I W_J + \sum_{\substack{|I|, |J|=p, q \\ I \cap J \neq \emptyset}} \sum_{a \in I \cap J} (W_I W_J + \mathbb{E}[W_I W_J | \mathcal{G}_a]).$$

Then for  $p \neq q$ , using the convexity of  $x \mapsto x^2$ ,

$$\begin{aligned} \text{Var}(\Gamma(F_p, F_q)) &\leq \frac{1}{2} \text{Var} \left[ (p+q) \sum_{\substack{|I|, |J|=p, q \\ I \cap J = \emptyset}} W_I W_J \right] \\ &\quad + \frac{1}{2} \text{Var} \left[ \sum_{\substack{|I|, |J|=p, q \\ I \cap J \neq \emptyset}} \sum_{a \in I \cap J} (W_I W_J + \mathbb{E}[W_I W_J | \mathcal{G}_a]) \right] \\ \text{Var}(\Gamma(F_p, F_q)) &\leq \frac{1}{2} \mathbb{E} \left[ \left( (p+q) \sum_{\substack{|I|, |J|=p, q \\ I \cap J = \emptyset}} W_I W_J \right)^2 \right] \\ &\quad + \frac{1}{2} \text{Var} \left[ \sum_{\substack{|I|, |J|=p, q \\ I \cap J \neq \emptyset}} \sum_{a \in I \cap J} (W_I W_J + \mathbb{E}[W_I W_J | \mathcal{G}_a]) \right] \end{aligned}$$

$$\begin{aligned} 2 \text{Var}(\Gamma(F_p, F_q)) &\leq \sum_{\substack{|I|, |J|=p, q \\ I \cap J = \emptyset}} \sum_{\substack{|K|, |L|=p, q \\ K \cap L = \emptyset}} \mathbb{E}[W_I W_J W_K W_L] \\ &\quad + \mathbb{E} \left[ \sum_{\substack{|I|, |J|=p, q \\ I \cap J \neq \emptyset}} \sum_{\substack{|K|, |L|=p, q \\ K \cap L \neq \emptyset}} \sum_{a \in I \cap J} \sum_{b \in K \cap L} W_I W_J W_K W_L \right] \\ &\quad + \mathbb{E} \left[ \sum_{\substack{|I|, |J|=p, q \\ I \cap J \neq \emptyset}} \sum_{\substack{|K|, |L|=p, q \\ K \cap L \neq \emptyset}} \sum_{a \in I \cap J} \sum_{b \in K \cap L} W_I W_J \mathbb{E}[W_K W_L | \mathcal{G}_b] \right] \\ &\quad + \mathbb{E} \left[ \sum_{\substack{|I|, |J|=p, q \\ I \cap J \neq \emptyset}} \sum_{\substack{|K|, |L|=p, q \\ K \cap L \neq \emptyset}} \sum_{a \in I \cap J} \sum_{b \in K \cap L} \mathbb{E}[W_I W_J | \mathcal{G}_a] W_K W_L \right] \\ &\quad + \mathbb{E} \left[ \sum_{\substack{|I|, |J|=p, q \\ I \cap J \neq \emptyset}} \sum_{\substack{|K|, |L|=p, q \\ K \cap L \neq \emptyset}} \sum_{a \in I \cap J} \sum_{b \in K \cap L} \mathbb{E}[W_I W_J | \mathcal{G}_a] \mathbb{E}[W_K W_L | \mathcal{G}_b] \right]. \end{aligned}$$

We shall write

$$|C_{I, J, a}| = \left| \frac{\mathbb{E}[W_I W_J | \mathcal{G}_a]}{W_{I \setminus \{a\}} W_{J \setminus \{a\}}} \right| \text{ for all } I, J, a$$

with the convention  $W_\emptyset = 1$ .

We deal with each term one by one:

- If  $I \cap J = \emptyset$ ,  $K \cap L = \emptyset$ , and if there is more than 2 other pairs with null intersection, the contribution of the term is 0, hence the first term is non-zero if  $(I, J, K, L)$  is connected, then:

$$\sum_{\substack{|I|,|J|=p,q \\ I \cap J = \emptyset}} \sum_{\substack{|K|,|L|=p,q \\ K \cap L = \emptyset}} \mathbb{E}[W_I W_J W_K W_L] \leq \sum_{I,J,K,L \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|.$$

- The second term consists of the sums of product of factors indexed by connected sets since there are at least two pairs that have non-null intersection. Since  $p \neq q$ ,  $\mathbb{E}[W_I W_J | Z] = 0$  for  $|I| = p$  and  $|J| = q$ , so if the terms are non-zero,  $W_I W_J$  and  $W_K W_L$  are not conditionally independent.
- For the third term, using self-adjointness, the terms are non-zero if  $b \in I \cap J$ , hence it is equivalent to:

$$|C_{I,J,a} \mathbb{E}[W_{I \setminus \{b\}} W_{J \setminus \{a\}} W_K W_L]| = |C_{I,J,a}| |\mathbb{E}[W_{I \setminus \{b\}} W_{J \setminus \{a\}} W_K W_L]|.$$

If  $b$  is the unique element that lies in the intersection, the contribution is 0, otherwise  $I, J, K, L$  are connected or the contribution is

$$\mathbb{E}[W_I W_J | Z] \mathbb{E}[W_K W_L | Z] = 0$$

because  $|I| \neq |J|$ .

- For the last term, it is the same argument.

Then, there exists a constant  $C$  independent of others such that

$$\text{Var}(\Gamma(F_p, F_q)) \leq (1 + m^2 + 2Cm^2 + C^2m^2) \sum_{I,J,K,L \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|.$$

**Remark 2.5.11.** It is clear that the constant  $C_m$  depends on the distributions of the random variables, hence it is not a universal constant as found in [Döbler and Peccati \(2017\)](#); [Döbler \(2023\)](#). It is a generalization of the main result for normal approximation in [Bhattacharya et al. \(2022\)](#) to homogeneous sums. In the case where the random variables are identically distributed, the conditions for the result are convenient. However, in the case of homogeneous sums, the [\(H1\)](#) condition additionally requires that the random variables are square-integrable. Moreover, the constant depends on the order of the U-statistic which can detrimental to the convergence rate for U-statistics with non-finite Hoeffding decomposition. ■

□

In [Privault and Serafin \(2022\)](#), Privault and Serafin established a partial fourth moment theorem for  $F$  a functional of independent random variables sum of element in the first and second chaos of their own Malliavin structure. To that end, we implement another strategy which is to reexpress the remainder in the partial fourth moment theorem as a fourth order term.

**Theorem 2.5.12 — Quantitative De Jong’s theorem II.** *If  $F = \sum_{p=1}^m F_p$  where  $F_p \in \mathfrak{C}_p$  and the following assumptions hold:*

- $F_p$  are chaos eigenfunctions (EGF);
- the condition (H1);
- 

$$\kappa = \sup_{I, J \subset A} \frac{\mathbb{E}[W_I^2] \mathbb{E}[W_J^2]}{\mathbb{E}[W_I^2 W_J^2]} < \infty \quad (\text{H2})$$

is independent of  $A$ .

Then:

$$d_W(F, \mathcal{N}(0, 1)) \leq C_m \sqrt{\sum_{(I, J, K, L) \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|}, \quad (2.59)$$

where the constant  $C_m$  grows quadratically with  $m$ , independent of all others.

*Proof of Theorem 2.5.12.* Let us prove the upper bound of  $\text{Var}[\Gamma(F_p, F_p)]$  by bounding the fourth cumulant:

$$\begin{aligned} \mathbb{E}[F_p^4] &= 3 \sum_{\substack{I, J, K, L \in (A, p) \\ (I \cup J) \cap (K \cup L) = \emptyset}} \mathbb{E}[W_I W_J] \mathbb{E}[W_K W_L] + \sum_{\substack{I, J, K, L \in (A, p) \\ I, J, K, L \text{ connected}}} \mathbb{E}[W_I W_J W_K W_L] \\ &= 3 \sum_{I, J \in (A, p)} \mathbb{E}[W_I^2] \mathbb{E}[W_J^2] - 3 \sum_{I \cap J \neq \emptyset} \mathbb{E}[W_I^2] \mathbb{E}[W_J^2] \\ &\quad + \sum_{\substack{I, J, K, L \in (A, p) \\ I, J, K, L \text{ connected}}} \mathbb{E}[W_I W_J W_K W_L] \\ &= 3\mathbb{E}[F_p^2]^2 + \sum_{\substack{I, J, K, L \in (A, p) \\ I, J, K, L \text{ connected}}} \mathbb{E}[W_I W_J W_K W_L] - 3 \sum_{\substack{I \cap J \neq \emptyset \\ I \neq J}} \mathbb{E}[W_I^2] \mathbb{E}[W_J^2]. \end{aligned}$$

Then, one has:

$$|\mathbb{E}[F_p^4] - 3\mathbb{E}[F_p^2]^2| \leq (1 + 3\kappa) \sum_{I, J, K, L \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]|. \quad (2.60)$$

□

The assumptions may seem cumbersome, but as shown in Lemma 2.5.5 concerning (EGF), they are valid for homogeneous sums.

## Chapter 3

# Motif estimation

In this chapter, we consider the problem of subgraph counts in random hypergraphs. That problem found its roots in the landmark paper Erdős (1947) and has been tackled in numerous papers. The ones which use Stein’s method derived rates of convergence for the rescaled subgraph counts towards the normal distribution or the Poisson distribution as the number of vertices  $n$  grows to infinity. The overarching goal of our effort is to extend the existing results as to derive the bounds of normal approximation of subgraph counts in random exchangeable hypergraphs. Upon recalling the general context of motif estimation in Section 3.1 and definitions of random hypergraphs in Section 3.2, we establish a quantitative limit theorems concerning 2 classes of random hypergraphs, extending some earlier results for random graphs. In Section 3.5, we discuss future research directions using our version of the quantitative De Jong’s theorem.

### 3.1 General context

Hypergraphs are extensions of graphs in which an edge can join any number of vertices. Although many articles revolving around that notion are oriented in combinatorics, the hypergraphs give rise to structures that are used in big data such as simplicial complexes for topological data analysis. There are a lot of applications in network analysis which considers huge hypergraphs. Among them, the complexity of the 5G architecture requires modeling networks by 3-uniform hypergraphs, instead of graphs. For instance, the link between users, SAPs, and channels. Several works have discussed the hypergraph application in resource allocation for wireless communications, mainly to better account for the cumulative interference in a dense context (Zhang et al., 2017). Moreover, the hypergraph is the appropriate structure to tackle *folksonomy* (Ghoshal et al., 2009). Folksonomy is a classification system in which end users apply public tags to online items, typically to make those items easier for themselves or others to find later (Flickr, Instagram, etc.). There are fewer works of network analysis on hypergraphs, but the questions for graphs have equivalents for hypergraphs. In the era of big data, the observed data is a sample from a much larger parent network. The central statistical question in such studies is to estimate global features of the parent network with a control of the approximation made by sampling. Counting motifs (patterns of subgraphs) in a large network is a prominent statistical and computational problem as motifs are building blocks of complex networks. Motif counting can be used to identify anomalies in networks. It proves useful for tasks as community detection, link prediction, and network classification in a variety of fields including biology, computer science and economics. Although large sparse graphs are usually better models for networks, the theory of dense large graphs is more mature (Lovász, 2012). One major result about large graphs is the

*Szemerédi's regularity lemma* that tells that we can approximate very well a large graph by a structure of dense random graph. The latter is characterized by an average number of edges per vertex which is non-negligible. A simple example of such random graph is the *Erdős-Rényi model* (Erdős, 1947). The theory of exchangeable random graphs and random hypergraphs provides with a recipe of related random models.

## 3.2 Definitions

**Definition 3.2.1.** A hypergraph denoted by  $G = (V, E = (e_i)_{i \in \mathcal{P}(V)})$  on a finite set  $V = V(G)$  is a family of subsets of  $V$  called hyperedges. Vertices in a hypergraph are adjacent if there is a hyperedge which contains them. The vertices not in any edge are the isolated vertices of  $G$ . A hypergraph is connected if it contains no isolated vertices and if the intersection graph of  $E$  is connected.

We denote by  $[e]$  the set of vertices of the hyperedge  $e$ .

**Definition 3.2.2.** A  $k$ -uniform hypergraph  $G = (V, E)$  is a hypergraph where each hyperedge has cardinality  $k$ . In particular, such hypergraph has hyperedge set in  $\binom{V}{k}$ , the collection of  $k$ -tuples of the set of vertices  $V$ .

**Definition 3.2.3.** For  $k > 3$ , a subhypergraph (or simply subgraph) of a hypergraph  $G = (V, E)$  is a hypergraph  $H = (V', E')$  such that  $V' \subset V$  and  $E' \subset E \cap \binom{V}{k}$ .

We denote by  $v_H$  and  $e_H$  the number of vertices and number of hyperedges of a hypergraph  $H$  respectively.

**Definition 3.2.4.** The set  $\text{Aut}(G)$  is the *automorphism group* of  $G$  a  $r$ -uniform hypergraph that is, the permutations  $\sigma$  on the vertex set  $V(G)$  such that  $(x_1, x_2, \dots, x_r) \in E(G)$  if and only if  $(\sigma(x_1), \sigma(x_2), \dots, \sigma(x_r)) \in E(G)$ .

A 2-uniform hypergraph is a graph. A 3-uniform hypergraph is a hypergraph whose hyperedges are triangles only. We also denote by  $G^{(j)}$  the hypergraph induced by the hyperedges of cardinality  $j \leq k$  included in the hyperedges of the  $k$ -uniform hypergraph  $G$ .

Random graph models are various, but in the context of dense graphs, they all share a common property. The vertices of those random graphs are indistinguishable. That is to say, that they are informative of the structure of the random objects. It is related to the fact that the graph can be considered unlabeled in non-combinatorics applications. The random hypergraphs are natural extensions of random graphs. A vast majority of the literature deals with the Erdős-Rényi model and its generalization. It is an example of exchangeable random hypergraphs.

**Definition 3.2.5** (Exchangeable random graphs). Let  $\mathbf{G}$  a random graph with vertex set  $[n] = \{1, \dots, n\}$ , and its adjacency matrix  $X_{i,j} = \mathbb{1}_{\{ij \in E(\mathbf{G})\}}$ .  $\mathbf{G}$  is called exchangeable random graph if and only if the joint distribution  $(X_{i,j} : 1 \leq i, j, n)$  is the same as the joint distribution  $X_{\sigma(i), \sigma(j)}$  for  $\sigma$  a permutation on  $[n]$ .

**Definition 3.2.6.** A  $k$ -uniform exchangeable random hypergraph  $\mathbf{G}$  of vertex set  $V = [n]$  is defined by the set of  $\{0, 1\}$ -valued random variables  $(X_\alpha, \alpha \subset \binom{[n]}{k})$  such that:

- one associates each realization of the random variables a hypergraph  $([n], E)$  with  $\alpha \in E$  if and only if  $X_\alpha = 1$ ;
- $(X_\alpha)$  form an exchangeable array, i.e.  $X_{(\sigma(u))_{u \in \alpha}} \stackrel{d}{=} X_\alpha$ .

One can formulate a recipe for exchangeable random hypergraphs as done in (Austin, 2008, definition 2.8). Fix a sequence of ingredients which consist of a sequence of sample spaces and probability kernels that determine the presence of  $k$ -hyperedges in the hypergraph based on the indicators  $X_\beta$  for  $(k-1)$ -hyperedges  $\beta$ :

$$(\{*\}), (V, P_1), (\{0, 1\}, P_2), (\{0, 1\}, P_3), \dots, (\{0, 1\}, P_{k-1}), (\{0, 1\}, P_k)$$

where we write  $\{*\}$  for a one-point space,  $(P_k)_{k \in \mathbb{N}}$  is a family of probability kernels such that for all  $k \in \mathbb{N}$ ,  $P_k$  is a probability kernel from  $\prod_{j=0}^{k-1} \{0, 1\}^{\binom{V}{j}}$  to  $\{0, 1\}^{\binom{V}{k}}$ .

- Color each vertex  $s \in V$  by some  $x_s \in \{0, 1\}$  chosen independently according to  $P_1(*, \cdot)$ ;
- Color each edge  $a = \{s, t\} \in \binom{V}{2}$  by some  $x_a \in \{0, 1\}$  chosen independently according to  $P_2(*, x_s, x_t, \cdot)$ ;
- $\vdots$
- Color each  $(k-1)$ -hyperedge  $u \in \binom{V}{k-1}$  by some  $x_u \in \{0, 1\}$  chosen independently according to  $P_{k-1}(*, (x_s)_{s \in \binom{[u]}{1}}, *, \dots, *, (x_v)_{v \in \binom{[u]}{k-2}}, \cdot)$ ;
- Color each  $k$ -hyperedge  $e \in \binom{V}{k}$  by some color  $x_e \in \{0, 1\}$  chosen independently according to  $P_k(*, (x_s)_{s \in \binom{[e]}{2}}, *, \dots, *, (x_u)_{u \in \binom{[e]}{k-1}}, \cdot)$ .

**Example 3.2.7** (Erdős-Rényi random model). The randomness intervenes at the level of edges.  $P_1(*, \cdot)$  is the uniform distribution on  $V$ . We color each edge  $a = \{s, t\} \in \binom{V}{2}$  by some  $z_a \in \{0, 1\}$  chosen independently according to  $P_2(*, x_s, x_t, \cdot) \stackrel{d}{=} \mathcal{B}(p)$  the Bernoulli distribution with parameter  $p$  for some  $p \in [0, 1]$  which is called the *edge density*.

**Example 3.2.8** (Stochastic block model). A stochastic block model corresponds to a model where there are communities, and each edge has a probability of belonging to the model according to the community of the vertices that the edge links. Likewise, the randomness intervenes at the level of the edges. Set a partition  $V = C_1 \sqcup \dots \sqcup C_q$ . Let  $(p_{i,j})_{i,j \in \llbracket 1,q \rrbracket^2}$  a sequence of reals in  $[0, 1]$ . We can assign a community to each vertex  $s$ , and denote it by  $c(s)$ . Then:

- $P_1(*, \cdot)$  is the uniform distribution;
- $P_2(*, z_s, z_t, \cdot) \stackrel{d}{=} \mathcal{B}(p_{c(s), c(t)})$ .



In the following, we write  $\mathbf{G}^{(3)}$  a random hypergraph model  $\mathbf{G}$  as to distinguish the 3-uniform hypergraph random model from its equivalent graph random model. The natural extension of the Erdős-Rényi model denoted  $\mathbb{G}^{(3)}(n, p_n)$  consists of having

$$P_3(*, x_{st}, x_{tu}, x_{us}) \stackrel{d}{=} \mathcal{B}(p_n),$$

i.e. we draw every triangle of the hypergraph with probability  $p_n$ . We also consider another random model based on the recipe. Define  $(\mathbb{T}^{(3)}(n, q_n, p_n))_{n \in \mathbb{N}}$  the sequence of 3-uniform hypergraphs such that for  $(s, t, u) \in V^3$ :

•

$$P_2(*, x_s, x_t) \stackrel{d}{=} \mathcal{B}(q_n);$$

•

$$P_3(*, x_{st}, x_{tu}, x_{us}) \stackrel{d}{=} \mathcal{B}(p_n).$$

It differs from  $\mathbb{G}^{(3)}(n, p_n)$  in many ways as pointed out by (Lovász, 2012, Example 23.11), but we note that  $\mathbb{G}^{(3)}(n, p_n)$  and  $\mathbb{T}^{(3)}(n, 1, p_n)$  have the same law. The case  $q_n < 1$  has been understudied in the literature. The functional identities in Section 2.3 can be applied to random hypergraphs in the same way as for random graphs Janson et al. (2000, Corollary 2.27). The most intriguing discovery made by Erdős and Rényi in the course of investigating random graphs is the phenomenon of thresholds. For many graph properties, the limiting probability that a random graph possesses them jumps from 0 to 1 (or vice versa) abruptly, that is, with a rather small increase in the expected number of edges. To be more precise in the description of this phenomenon, let us introduce threshold functions (or just thresholds) for monotone graph properties.

**Definition 3.2.9** (Threshold function). Let a sequence of properties  $(\mathcal{Q}_n)_n$ . A threshold function is a sequence  $(s_n)_{n \in \mathbb{N}}$  such that:

1. for all sequence  $(p_n)_n$  such that  $p_n = o(s_n)$ ,

$$\mathbb{P}(\mathbb{G}(n, p_n) \in \mathcal{Q}_n) \rightarrow 0;$$

2. for all sequence  $(p_n)_n$  such that  $s_n = o(p_n)$

$$\mathbb{P}(\mathbb{G}(n, p_n) \in \mathcal{Q}_n) \rightarrow 1.$$

One of the oldest problem of motif estimation is *subgraph counting* in random graphs or random hypergraphs. In 1960, Erdős and Rényi pioneered in their foundational papers the containment problem. Let  $\mathbb{G}(n, M)$  the uniform random graph i.e. the random graph defined by taking as  $\Omega$  the vertex set  $[n]$  with exactly  $M$  edges, and  $\mathbb{P}$  the uniform probability on  $\Omega$ ,

$$\mathbb{P}(H) = \left( \binom{n}{2} \right)^{-1}, \quad H \in \Omega.$$

Does at least one copy of a given graph in  $G$  exist in  $\mathbb{G}(n, M)$ ? Small subgraph counts can be used as summary statistics for large random graphs. That can be formulated as the problem of finding the threshold for the property of containing  $G$ .

We recall the following theorem (Janson et al., 2000, Theorem 3.4).

**Theorem 3.2.10** For an arbitrary graph  $G$  with at least one edge,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\mathbb{G}(n, p_n) \supset G) = \begin{cases} 0 & \text{if } p_n \ll n^{-1/m(G)}, \\ 1 & \text{if } p_n \gg n^{-1/m(G)}, \end{cases}$$

where

$$m(G) = \max \left\{ \frac{e_H}{v_H} : H \subset G, v_H > 0 \right\}$$

is the maximum average degree of subgraphs in  $G$ .

Hence, a quantity of interest is:

$$\Psi_G = \min_{H \subset G, e_H > 0} n^{v_H} p_n^{e_H}. \quad (3.1)$$

In the sequel, we recall some results for  $M_G$  (Janson et al., 2000, Lemma 3.6), the number of subgraphs of  $\mathbb{G}(n, p_n)$  which are isomorphic to  $G$ . We denote by  $\bar{M}_G$  the normalized number.

**Lemma 3.2.11** Those assertions are equivalent, for any graph  $G$  with  $e_G > 0$

1.  $n^{v_H} p_n^{e_H} \xrightarrow{n \rightarrow \infty} \infty$  for every  $H \subset G$  with  $v_H > 0$ .
2.  $\mathbb{E}[M_H] \xrightarrow{n \rightarrow \infty} \infty$  for every  $H \subset G$  with  $v_H > 0$ .
3.  $\Psi_G \xrightarrow{n \rightarrow \infty} \infty$ .

Those results and their consequences prove that the rate of convergence in terms of  $\Psi_G$  is sharp. Friedgut and Kalai (1996), the appendix to Friedgut and Bourgain (1999) and Bourgain and Kalai (1997) provide much greater insight into the notion of sharp thresholds. Friedgut gives a survey of these aspects (Friedgut, 2005).

### 3.3 Normal approximation

There are two main asymptotic distributions for numerical characteristic of random graphs: Poisson distribution and normal distribution. The Poisson approximation of subgraph count has been extensively studied (see Janson et al. (2000, Example 6.26 p.161)) using Stein-Chen method and the well-known theorem for Poisson approximation that can be found for example in Janson et al. (2000, Theorem 6.22 p.159). For related works, see also the earlier reference Barbour et al. (1992) and the survey (Chatterjee et al., 2005).

Historically, both Poisson convergence and asymptotic normality for subgraph counting had been dealt with method of moments (Ruciński, 1988), which requires tedious computations, but is quite adapted to this application. In the 1990s, Barbour and Chen used Stein's method to establish the asymptotic normality of the number of triangles in a random graph  $\mathbb{G}(n, p)$ . They also showed that the asymptotic normality holds under a wide range of conditions on the edge probabilities. Barbour, Karoński, and Ruciński proved the following bound for the convergence rate in the Wasserstein distance, using Stein's method (see Theorem 2 in Barbour et al. (1989)):

$$d_W(\bar{M}_G, \mathcal{N}(0, 1)) \leq C_G \cdot \begin{cases} \frac{1}{n\sqrt{1-p_n}} & \text{if } p_n > \frac{1}{2}, \\ \frac{1}{\sqrt{\Psi_G}} & \text{if } p_n \leq \frac{1}{2}. \end{cases} \quad (3.2)$$

Subsequently, it has been a long standing problem, whether the same rate could be achieved in the Kolmogorov distance  $d_{Kol}$ . In an article submitted in 2017, Röllin used the Stein-Tikhomirov method to bound the convergence rate in the Kolmogorov distance for normal approximation of normalized triangle counts in the Erdős-Rényi random graph (Röllin, 2022). That was generalized by Eichelsbacher et al. (2023); Eichelsbacher and Rednoß (2023) as to obtain a Kolmogorov bound equivalent to (3.2). It is also an application of Krokowski et al. (2017). We note that Wasserstein and Kolmogorov bounds were obtained by an application of Malliavin-Stein's method in Privault and Serafin (2020, 2022). All those papers suppose that the random variables indicating the presence of edges are independent of each other. In order to deal with random hypergraphs depending on conditionally independent random variables (Definition 3.2.6), we use the Malliavin-Stein's method developed in the previous chapter. We implement here a different approach using our partial fourth moment theorem for normal approximation in the Wasserstein-1 distance, and extend the results for random hypergraphs.

The ancestor of fourth moment theorem was initiated by De Jong (1990), and was applied to problem of subgraph counting (De Jong, 1996). More recently, Bhattacharya et al. (2022) derived bounds in Wasserstein distance for the analogous problem of motif estimation. The fourth moment phenomenon that arise in there is the convergence of the fourth cumulant of the normalized graph statistics to the fourth cumulant of the standard reduced normal distribution, which is 3. However, in the case of the subgraph counting, the appropriate bound is an ersatz of the fourth cumulant. It was shown that the convergence of the fourth cumulant is not enough to prove asymptotic normality for functionals of Rademacher random variables (Döbler and Krokowski, 2019). This was known, at least for asymptotic normality of normalized subhypergraph counts in random hypergraphs, as De Jong proved.

We follow De Jong's approach to the problem, that leverages the Hoeffding decomposition. In the sequel, we treat the special cases of 3-uniform random hypergraphs, focusing on extensions of the Erdős-Rényi model introduced earlier. The symbol  $\simeq$  stands for hypergraph isomorphism.

The number of subhypergraphs of  $\mathbb{G}^{(3)}(n, p_n)$  isomorphic to  $G$  is

$$M_G = \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \prod_{\alpha \in H} \hat{X}_\alpha. \quad (3.3)$$

For  $\sigma \in \text{Aut}(G)$ ,  $(x, y, z) \in E(G)$  if and only if  $(\sigma(x), \sigma(y), \sigma(z)) \in E(G)$ . The random variable  $M_G$  has a finite Hoeffding decomposition (De Jong, 1996, p.11(115)). Since  $\hat{X}_\alpha = p_n + (\hat{X}_\alpha - p_n)$ ,  $M_G$  admits the decomposition:

$$M_G = \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{J \subseteq H} p_n^{|H|-|J|} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n), \quad (3.4)$$

where the summation extends over all subsets  $J$  of  $I$ , in virtue of the inclusion-exclusion principle.

By interchanging the sums, we find the chaotic decomposition of  $M_G - \mathbb{E}[M_G]$  that is:

$$\begin{aligned}
M_G - \mathbb{E}[M_G] &= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{\substack{J \subseteq H \\ J \neq \emptyset}} p_n^{|J|-|J|} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n), \\
&= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{j=1}^{e_G} p_n^{e_G-j} \sum_{\substack{J \subseteq H \\ |J|=j}} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n) \\
&= \sum_{j=1}^{e_G} p_n^{e_G-j} \sum_{|J|=j} \prod_{\alpha \in J} (\hat{X}_\alpha - p_n) \left( \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G, H \supseteq J}} 1 \right) \\
&= \sum_{j=1}^{e_G} \pi_j(M_G),
\end{aligned}$$

where

$$\pi_k(M_G) = p_n^{e_G-j} \sum_{|J|=j} \left( \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G, H \supseteq J}} 1 \right) \prod_{\alpha \in J} \hat{Y}_\alpha \quad (3.5)$$

with  $\hat{Y}_\alpha$  is the centered version of  $\hat{X}_\alpha$  for all  $\alpha$  hyperedges of  $K_n$ . We note that the decomposition above corresponds to the Hoeffding decomposition of the U-statistics with

$$W_J \propto \left( \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G, H \supseteq J}} 1 \right) \prod_{\alpha \in J} \hat{Y}_\alpha. \quad (3.6)$$

We proceed in the same manner in  $\mathbb{T}^{(3)}(n, q_n, p_n)$ . Define  $N_G$  the number of subhypergraphs isomorphic to  $G$ .

$$N_G = \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \prod_{\alpha \in H} X_\alpha, \quad (3.7)$$

Here,  $(X_\alpha)_{\alpha \in \binom{[n]}{3}}$  is a sequence of conditionally independent Bernoulli random variables given  $Z := \mathbb{G}(n, q_n)$ . The chaos decomposition yields:

$$\begin{aligned}
N_G &= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{J \subseteq H} \prod_{\beta \in H \setminus J} \mathbb{E}[X_\beta | \mathbb{G}(n, q_n)] \prod_{\alpha \in J} (X_\alpha - \mathbb{E}[X_\alpha | \mathbb{G}(n, q_n)]) \\
&= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{J \subseteq H} p_n^{|H|-|J|} \mathbf{1}_{\{(H \setminus J)^{(2)} \subset \mathbb{G}(n, q_n)\}} \prod_{\alpha \in J} (X_\alpha - \mathbb{E}[X_\alpha | \mathbb{G}(n, q_n)]).
\end{aligned} \quad (3.8)$$

Hence,  $N_G - \mathbb{E}[N_G | \mathbb{G}(n, q_n)]$  reads off:

$$\sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{\emptyset \neq J \subseteq I} p_n^{|H|-|J|} \mathbf{1}_{\{(H \setminus J)^{(2)} \subset \mathbb{G}(n, q_n)\}} \prod_{\alpha \in J} (X_\alpha - \mathbb{E}[X_\alpha | \mathbb{G}(n, q_n)]). \quad (3.9)$$

The corresponding degenerate U-statistics in the decomposition are given for  $J \subset \binom{[n]}{3}$  by

$$W_J = w_J \prod_{\alpha \in J} Y_\alpha, \quad (3.10)$$

where  $Y_\alpha$  is the centered version of  $X_\alpha$  given  $\mathbb{G}(n, q_n)$  and:

$$w_J = \left( \sum_{\substack{I \in \binom{[n]}{3} \\ H \simeq G, I \supseteq J}} p_n^{|H|-|J|} \mathbb{1}_{\{(H \setminus J)^{(2)} \subset \mathbb{G}(n, q_n)\}} \right).$$

Let us denote  $\bar{N}_G$  the rescaled statistic of the number of isomorphic copies of  $G$  with respect to their expectation, and  $\tilde{N}_G$  the rescaled statistic with respect to its conditional mean.

**Theorem 3.3.1 — Central Limit theorem (De Jong, 1996).** *Let  $X = (X_1, X_2, \dots)$  a sequence of independent random variables. Suppose  $F$  is  $\mathcal{F}_{[n]}$  measurable, with finite Hoeffding decomposition such that  $\text{Var}[F] = 1$  ( $F$  is a degenerate U-statistics of order  $d$ ):*

$$F = \sum_{I \subset [n], |I| \leq d} W_I, \quad (3.11)$$

satisfies the two conditions:

1.  $\max_{i \in [1, n]} \sum_{I \ni i} \mathbb{E}[W_I^2] \rightarrow 0$ , for  $n \rightarrow +\infty$ ;
2.  $\sum_{I, J, K, L \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]| \rightarrow 0$  for  $n \rightarrow +\infty$ .

Then  $F - \mathbb{E}[F] \xrightarrow{d} \mathcal{N}(0, 1)$ .

**Remark 3.3.2.** The condition 2 implies that the normed fourth moment of  $F$  tends to 3. The first condition is not very severe. It is a Lindeberg-Feller type condition as for  $d = 1$  it is the crucial assumption in the Lindeberg-Feller central limit theorem (see for example Theorem 5.12 in Kallenberg (2002)). In the homogeneous case where  $|I| = d$ , the fourth moment condition is sufficient, but otherwise it is not. ■

De Jong applied that theorem to subgraph counting in Erdős-Rényi random hypergraphs with edges defined as random subsets of a vertex set as to obtain the rate in (3.2). We prove a quantitative version of the Theorem 3.3.1 by using Theorem 2.5.12. The index set of the random variables is here, the discrete set of hyperedges.

**Theorem 3.3.3** *Let  $G$  a hypergraph without isolated vertices. Then,*

$$d_W(\tilde{N}_G, \mathcal{N}(0, 1)) \leq C_{eG} \sqrt{\sum_{(I, J, K, L) \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]| / \text{Var}[\tilde{N}_G]}. \quad (3.12)$$

*Proof of Theorem 3.3.3.* We check whether the conditions of Theorem 2.5.12 hold. For both

statistics, the (EGF) assumption holds. By conditional independence of  $(X_\alpha)_{\alpha \in \binom{[n]}{3}}$ , we have:

$$\begin{aligned} \frac{\mathbb{E}[W_I^2] \mathbb{E}[W_J^2]}{\mathbb{E}[W_I^2 W_J^2]} &\propto \frac{\mathbb{E} \left[ \prod_{\alpha \in I} \mathbb{E}[Y_\alpha^2 | Z] \prod_{\alpha \in J} \mathbb{E}[Y_\alpha^2 | Z] \right]}{\prod_{I \setminus J} \mathbb{E}[Y_\alpha^2] \prod_{J \setminus I} \mathbb{E}[Y_\alpha^2] \prod_{I \cap J} \mathbb{E}[Y_\alpha^2]} \\ &= (p_n(1-p_n))^{|I|+|J|-|I \cup J|} q_n^{|I^{(2)}|+|J^{(2)}|-|I^{(2)} \cup J^{(2)}|} \leq 1. \end{aligned}$$

Let us note that for all  $a$ ,  $W_{I \setminus \{a\}}$  is non-zero with the definition of  $N_G - \mathbb{E}[N_G | Z]$ . Let  $W_I = w_I \prod_{i \in I} X_i$ , then for  $a \in I \cap J$ :

$$\begin{aligned} \mathbb{E}[W_I W_J | \mathcal{G}_a] &= w_I w_J \prod_{i \in I \setminus \{a\}} Y_i \prod_{j \in J \setminus \{a\}} Y_j \mathbb{E}[Y_a^2 | Z] \\ &= \frac{w_I w_J}{w_{I \setminus \{a\}} w_{J \setminus \{a\}}} \mathbb{E}[Y_a^2 | Z] W_{I \setminus \{a\}} W_{J \setminus \{a\}} \\ &= C_{I,J,a} W_{I \setminus \{a\}} W_{J \setminus \{a\}}, \end{aligned}$$

with

$$C_{I,J,a} = \frac{w_I w_J}{w_{I \setminus \{a\}} w_{J \setminus \{a\}}} \mathbb{E}[Y_a^2 | Z] < +\infty \quad \mathbb{P}\text{-almost surely.}$$

□

By taking  $q_n = 1$ , we have that result.

**Corollary 3.3.4** *Let  $G$  a hypergraph without isolated vertices. Then,*

$$d_W(\bar{M}_G, \mathcal{N}(0, 1)) \leq C_{e_G} \sqrt{\sum_{(I,J,K,L) \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]| / \text{Var}[M_G]}. \quad (3.13)$$

We deduce those convergence rates for  $p_n < c < 1$  for some  $c$ .

**Theorem 3.3.5** *Let  $G$  a hypergraph without isolated vertices. Then, we have*

$$d_W(\tilde{N}_G, \mathcal{N}(0, 1)) \lesssim \left( \min_{\substack{H \subset G \\ e_H > 1}} \{n^{v_H} p_n^{e_H}\} \right)^{-1/2} \quad (3.14)$$

and

$$d_W(\tilde{N}_G, \mathcal{N}(0, 1)) \lesssim \left( \min_{\substack{H \subset G \\ e_H > 1}} \{n^{v_H} p_n^{e_H} q_n^{e_H^{(2)}}\} \right)^{-1/2}, \quad (3.15)$$

where  $e_H^{(2)}$  is the number of edges included in the hyperedges of  $H$ .

*Proof of Theorem 3.3.5.* We are left to upper bound the quantity:

$$\begin{aligned} &\sum_{(I,J,K,L) \text{ connected}} |\mathbb{E}[W_I W_J W_K W_L]| \\ &\propto_Z \sum_{(I,J,K,L) \text{ connected}} \left| \mathbb{E} \left[ \mathbb{E} \left[ \prod_{\alpha \in I} Y_\alpha \prod_{\alpha \in J} Y_\alpha \prod_{\alpha \in K} Y_\alpha \prod_{\alpha \in L} Y_\alpha \mid Z \right] \right] \right| \end{aligned}$$

where the notation  $\propto_Z$  accounts for an equality up to a factor depending only on  $Z$ . The terms are non-zero if and only if  $\alpha$  lies in at least two elements of the quadruple, i.e. if  $\alpha$  does not lie in  $I \setminus (J \cup K \cup L)$ , etc. Then, the number of non-zero terms is  $I \cup J \cup K \cup L$ . We recall that:

$$\begin{aligned}\mathbb{E}[Y_\alpha|Z] &= 0 \\ \mathbb{E}[Y_\alpha^2|Z] &= p_n(1-p_n)\mathbb{1}_{\{\alpha^{(1)} \in Z\}} \prod_{i=1}^3 \mathbb{1}_{\{\alpha^{(i)} \in Z\}} \\ \mathbb{E}[Y_\alpha^3|Z] &= p_n(1-p_n)(1-2p_n) \prod_{i=1}^3 \mathbb{1}_{\{\alpha^{(i)} \in Z\}} \lesssim p_n(1-p_n) \prod_{i=1}^3 \mathbb{1}_{\{\alpha^{(i)} \in Z\}} \\ \mathbb{E}[Y_\alpha^4|Z] &= p_n(1-p_n)(1-3p_n(1-p_n)) \prod_{i=1}^3 \mathbb{1}_{\{\alpha^{(i)} \in Z\}} \lesssim p_n(1-p_n) \prod_{i=1}^3 \mathbb{1}_{\{\alpha^{(i)} \in Z\}}.\end{aligned}$$

Now, we remark  $I, J, K, L$  are respectively isomorphic to  $A, B, C, D$  subhypergraphs of  $G$ . Hence, we can sum first over  $(A, B, C, D)$ , and then over all the quadruples  $(I, J, K, L)$  whose components are respectively isomorphic to the ones of the fixed quadruple  $(A, B, C, D)$ . We shall write:

$$\sum_{I, J, K, L} \cdot = \sum_{A, B, C, D \subset G} \sum_{\substack{I \simeq A, J \simeq B \\ K \simeq C, L \simeq D}} \cdot := \sum_{A, B, C, D} \sum_{I, J, K, L}^{*A, B, C, D} \cdot$$

$v(A)$  denotes the number of vertices in  $A$ . We have that  $|\{I, J, K, L \in \binom{[n]}{r} : I \simeq A, J \simeq B, K \simeq C, L \simeq D\}|$  is bounded by the number of collection of vertices of cardinal  $v(A \cup B \cup C \cup D)$ . By a counting argument, we see that is of order  $n^{v(A \cup B \cup C \cup D)}$ . Because  $(I, J, K, L)$  is connected and copies of subhypergraphs of  $G$ , we also have that  $|I \cup J \cup K \cup L| = |A \cup B \cup C \cup D|$  and  $|I^{(2)} \cup J^{(2)} \cup K^{(2)} \cup L^{(2)}| = |A^{(2)} \cup B^{(2)} \cup C^{(2)} \cup D^{(2)}|$ . Hence, for a fixed connected quadruple  $(I, J, K, L)$  associated to  $(A, B, C, D)$ ,

$$\begin{aligned}|\mathbb{E}[W_I W_J W_K W_L]| \\ \lesssim w_I w_J w_K w_L n^{v(A \cup B \cup C \cup D)} p_n^{|A \cup B \cup C \cup D|} q_n^{|A^{(2)} \cup B^{(2)} \cup C^{(2)} \cup D^{(2)}|}.\end{aligned}$$

Let us bound the variance of  $N_G - \mathbb{E}[N_G | \mathbb{G}(n, q_n)]$ :

$$\begin{aligned}\text{Var}^2[N_G - \mathbb{E}[N_G | \mathbb{G}(n, q_n)]] &= \left( \sum_{I \cap J \neq \emptyset} \mathbb{E}[W_I W_J] \right)^2 = \sum_{I \cap J \neq \emptyset} (\mathbb{E}[W_I^2] + \mathbb{E}[W_J^2])^2 \\ &= \frac{1}{2^2} \sum_{\substack{A, B \subset G \\ A \cap B \neq \emptyset}} \left( \sum_I^{*A} \mathbb{E}[W_I^2] + \sum_J^{*B} \mathbb{E}[W_J^2] \right)^2.\end{aligned}$$

For a fixed connected quadruple  $(A, B, C, D)$ , by applying repeatedly the inequality  $a^2 + b^2 \geq 2ab$ ,

we get:

$$\begin{aligned} \text{Var}^2[N_G - \mathbb{E}[N_G | \mathbb{G}(n, q_n)]] & \\ & \geq \frac{1}{16} \left( \sum_I^{*A} \mathbb{E}[W_I^2] + \sum_J^{*B} \mathbb{E}[W_J^2] + \sum_K^{*C} \mathbb{E}[W_K^2] + \sum_L^{*D} \mathbb{E}[W_L^2] \right)^2 \\ & \geq \frac{1}{16} \left( \sum_I^{*A} \mathbb{E}[W_I^2] \times \sum_J^{*B} \mathbb{E}[W_J^2] \times \sum_K^{*C} \mathbb{E}[W_K^2] \times \sum_L^{*D} \mathbb{E}[W_L^2] \right)^{1/2}. \end{aligned}$$

Then using that  $\mathbb{E}[W_I^2] = w_I^2 q_n^{|I^{(2)}|} (1-p_n)^{|I|} p_n^{|I|} = q_n^{|A^{(2)}|} (1-p_n)^{|A|} p_n^{|A|}$ , so

$$\sum_I^{*A} \mathbb{E}[W_I^2] = \sum_I^{*A} w_I^2 (1-p_n)^{|A|} p_n^{|A|} q_n^{|A^{(2)}|} = n^{v(A)} (1-p_n)^{|A|} p_n^{|A|} q_n^{|A^{(2)}|}.$$

In particular, one has for a fixed quadruple  $(I, J, K, L)$  and associated  $(A, B, C, D)$ :

$$\begin{aligned} \text{Var}^2[N_G - \mathbb{E}[N_G | \mathbb{G}(n, q_n)]] & \geq \\ & \frac{1}{16} w_I w_J w_K w_L \left( n^{v^*(A,B,C,D)} (p_n(1-p_n))^{e^*(A,B,C,D)} q_n^{e^{(2)*}(A,B,C,D)} \right)^{1/2} \end{aligned} \quad (3.16)$$

where  $v^*(A, B, C, D) = v(A) + v(B) + v(C) + v(D)$  and  $e^*(A, B, C, D) = |A| + |B| + |C| + |D|$  and  $e^{(2)*}(A, B, C, D) = |A^{(2)}| + |B^{(2)}| + |C^{(2)}| + |D^{(2)}|$ . It yields the result. Using the Lemma 9 of De Jong (1996), for any quadruple  $(I, J, K, L)$  of collections  $I, J, K, L$  isomorphic to (sub)collections of  $H$ , there exists two (not both empty) subcollections of  $G$ , say  $M$  and  $M'$ , which may contain a nonzero number of isolated vertices, say  $i_M$  and  $i_{M'}$ , such that

$$v(M) + v(M') + i_M + i_{M'} = v(I) + v(J) + v(K) + v(L) - 2v(I \cup J \cup K \cup L), \quad (3.17)$$

$$|M^{(2)}| + |M'^{(2)}| = |I^{(2)}| + |J^{(2)}| + |K^{(2)}| + |L^{(2)}| - 2|I^{(2)} \cup J^{(2)} \cup K^{(2)} \cup L^{(2)}| \quad (3.18)$$

and by extension:

$$|M| + |M'| = |I| + |J| + |K| + |L| - 2|I \cup J \cup K \cup L|. \quad (3.19)$$

As  $G$  does not have isolated vertices, so do  $M$  and  $M'$ . As  $M$  and  $M'$  are subcollections of  $G$ , their average degree does not exceed  $m(G)$ . Hence, by [Theorem 3.3.3](#),

$$d_W(\tilde{N}_G, \mathcal{N}(0, 1)) \lesssim (n^{v(M)+v(M')} p_n^{|M|+|M'|} q_n^{|M^{(2)}|+|M'^{(2)}|})^{1/2}.$$

Thus, (3.15) follows. The first result for  $M_G$  is obtained with  $q_n = 1$ . □

**Remark 3.3.6.** This bound is relevant only for the regime  $p_n \xrightarrow{n \rightarrow \infty} 0$ . ■

### 3.4 A modified Hoeffding decomposition

In that section, we readopt the notations of the previous chapter by denoting  $A$  the index set of the random variables. Let another set  $\hat{A}$  that index auxiliary random variables in addition to  $(\hat{X}_\beta)_{\beta \in \hat{A}}$ . We shall write the sequence of conditionally independent random variables given



$Z, \mathbf{X} = (X_\alpha, \dots, \hat{X}_\beta, \dots)_{\alpha \in A, \beta \in \hat{A}}$  where the subsequence  $(\hat{X}_\beta)_{\beta \in \hat{A}}$  is a sequence of independent random variables, and  $\sigma(Z) = \sigma(\hat{X}_a, a \in A)$ . This setting is new to the best of our knowledge, and is specifically tailored for the application in  $\mathbb{T}^{(3)}(n, q_n, p_n)$ . We assume that  $A$  is the set of 3-hyperedges,  $\hat{A}$  is the set of edges included in the hyperedges of  $A$  and

$$X_\alpha = g(U_\alpha) \prod_{b \subset \alpha} \hat{X}_b \quad (3.20)$$

where  $(U_\alpha)_{\alpha \in A}$  forms a sequence of conditionally independent random variables given  $\hat{X}$ , following the uniform distribution.

**Lemma 3.4.1** *The sequence  $\mathbf{X}$  is a sequence of conditionally independent random variables.*

*Proof.* Since, by assumption, for  $f$  bounded and  $(\alpha, \beta) \in \hat{A}^2$  such that  $\alpha \neq \beta$ :

$$\mathbb{E} \left[ f(\hat{X}_\beta) | \hat{X}_\alpha, Z \right] = \mathbb{E} \left[ f(\hat{X}_\beta) | Z \right],$$

the subsequence  $(\hat{X}_\beta)_{\beta \in \hat{A}}$  is a sequence of conditionally independent random variables given  $Z$ . Let  $\alpha \in A$  and  $\beta \in \hat{A}$ , by definition

$$\mathbb{E} \left[ f(X_\alpha) | \hat{X}_\beta, Z \right] = \mathbb{E} \left[ f(X_\alpha) | Z \right],$$

and:

$$\mathbb{E} \left[ f(\hat{X}_\beta) | X_\alpha, Z \right] = \mathbb{E} \left[ f(\hat{X}_\beta) | Z \right]$$

because  $\hat{X}_\beta$  is a function of  $Z$ . □

**Remark 3.4.2.** That type of sequence is a degenerate case of sequence of conditionally independent random variables since the  $\hat{X}_\alpha$  are constant given  $Z$ . ■

For our purpose, the following lemma shows the commutation relation.

**Lemma 3.4.3** *For  $F \in L^2(E_A)$  a homogeneous sum of conditionally independent random variables  $\mathbf{X}$  and  $\alpha \in A$  and  $\beta \in \hat{A}$  such that*

$$\mathbb{E} \left[ \mathbb{E} \left[ F | \hat{X}^{\{\beta\}}, X \right] | X^{\{\alpha\}}, \hat{X} \right] = \mathbb{E} \left[ \mathbb{E} \left[ F | X^{\{\alpha\}}, \hat{X} \right] | X, \hat{X}^{\{\beta\}} \right]. \quad (3.21)$$

*Proof of Lemma 3.4.3.* It suffices to consider functionals of the type:

$$X_\alpha X_{\alpha_1} \dots X_{\alpha_n}$$

for  $n \geq 1$ . If  $\beta$  is not included in the edges of  $\alpha$ , we have the property by independence of the associated random variables.

Let consider the case where  $\beta$  is one of the edge of  $\alpha$ .

$$\mathbb{E} \left[ F | X^{\{\alpha\}}, \hat{X} \right] = \mathbb{E} [g(U_\alpha)] \prod_{b \subset \alpha} \hat{X}_b \prod_{i=1}^n X_{\alpha_i}$$

and

$$\mathbb{E} \left[ F | \hat{X}^{\{\beta\}}, X \right] = g(U_\alpha) \mathbb{E} \left[ \hat{X}_\beta^{1 + \sum_{i=1}^n \mathbb{1}_{\{\beta \in \alpha_i\}}} \right] \prod_{i=1}^n \prod_{b \subset \alpha_i} \hat{X}_b.$$

Then,

$$\begin{aligned} \mathbb{E} \left[ \mathbb{E} \left[ F | \hat{X}^{\{\beta\}}, X \right] | X^{\{\alpha\}}, \hat{X} \right] &= \mathbb{E}[g(U_\alpha)] \mathbb{E} \left[ \hat{X}_\beta^{1 + \sum_{i=1}^n \mathbf{1}_{\{\beta \in \alpha_i\}}} \prod_{i=1}^n \prod_{b \subset \alpha_i} \hat{X}_b \right] \\ &= \mathbb{E} \left[ \mathbb{E} \left[ F | X^{\{\alpha\}}, \hat{X} \right] | X, \hat{X}^{\{\beta\}} \right]. \end{aligned}$$

□

Those commutation relations of [Lemma 3.4.3](#) entail a modified Hoeffding decomposition of functionals of Bernoulli random variables.

**Lemma 3.4.4** *Given  $X$ , the modified chaos decomposition is given by:*

$$F = \mathbb{E}[F] + \sum_{n=1}^{+\infty} \pi_n(F)$$

with

$$\pi_n(F) = \sum_{\substack{I \subset A \cup \hat{A} \\ |I|=n}} \left( \prod_{b \in I} D_b \right) \left( \prod_{c \in (A \cup \hat{A}) \setminus I} \mathbb{E}[\cdot | \mathcal{G}^c] \right) \quad (3.22)$$

with  $\mathcal{G}^c = \sigma(X^{\{c\}})$ .

*Proof of Lemma 3.4.4.* We redefine a gradient  $D$  and Ornstein-Uhlenbeck operator  $\mathbf{L}$  in the same fashion as [Definition 2.2.3](#) such that for  $a \in A \cup \hat{A}$ :

$$D_a F = F - \mathbb{E} \left[ F | X^{\{a\}} \right].$$

Then, we follow the same scheme of proof as [Lemma 2.2.11](#) with that modified gradient. Thus, we obtain that  $\ker \mathbf{L} = \{F \in \text{Dom } \mathbf{L} : \mathbb{E}[F] = 0\}$  and (3.22). □

The resulting Malliavin framework is analogous to the Malliavin-Dirichlet structure in [Decreusefond and Halconruy \(2019\)](#), whose underlying Markov process is the usual Glauber dynamics starting from  $X$ . It extends the scope to a particular type of sequences of conditionally independent random variables. That applies to  $N_G$ . We recall that in the application to motif estimation,  $Z$  is the underlying Erdős-Rényi random graph  $\mathbb{G}(n, q_n)$ . The decomposition is similar to (3.8) except that this time the decomposition involves the random variables  $(\hat{X}_b)_{b \in \hat{A}}$ . Using the inclusion-exclusion principle,

$$\begin{aligned} \mathbb{E} [N_G | \mathbb{G}(n, q_n)] &= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \prod_{\alpha \in H} \mathbb{E} [X_\alpha | \mathbb{G}(n, q_n)] = \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \prod_{\alpha \in H} \prod_{\beta \subset \alpha} \hat{X}_\beta \\ &= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} p_n^{|H|} \prod_{\beta \in H^{(2)}} ((\hat{X}_\beta - \mathbb{E}[\hat{X}_\beta]) + \mathbb{E}[\hat{X}_\beta]) \end{aligned}$$

Hence,

$$\mathbb{E}[N_G | \mathbb{G}(n, q_n)] - \mathbb{E}[N_G] = \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} p_n^{|H|} \sum_{\emptyset \neq J \subseteq H} q_n^{|H^{(2)}| - |J^{(2)}|} \prod_{\beta \in J^{(2)}} (\hat{X}_\alpha - q_n).$$

It entails that:

$$\begin{aligned} N_G - \mathbb{E}[N_G] &= (N_G - \mathbb{E}[N_G | \mathbb{G}(n, q_n)]) + (\mathbb{E}[N_G | \mathbb{G}(n, q_n)] - \mathbb{E}[N_G]) \\ &= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{\emptyset \neq J \subseteq I} p_n^{|H| - |J|} \mathbb{1}_{\{(H \setminus J)^{(2)} \subset \mathbb{G}(n, q_n)\}} \prod_{\alpha \in J} (X_\alpha - \mathbb{E}[X_\alpha | \mathbb{G}(n, q_n)]) \\ &+ \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} p_n^{|H|} \sum_{\emptyset \neq J \subseteq H} q_n^{|H^{(2)}| - |J^{(2)}|} \prod_{\beta \in J^{(2)}} (\hat{X}_\alpha - q_n) \\ &= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{\emptyset \neq J \subseteq I} p_n^{|H| - |J|} \prod_{\beta \in (H \setminus J)^{(2)}} \hat{X}_\beta \prod_{\alpha \in J} (X_\alpha - \mathbb{E}[X_\alpha | \mathbb{G}(n, q_n)]) \\ &+ \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} p_n^{|H|} \sum_{\emptyset \neq J \subseteq H} q_n^{|H^{(2)}| - |J^{(2)}|} \prod_{\beta \in J^{(2)}} (\hat{X}_\alpha - q_n) \\ &= \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{\emptyset \neq J \subseteq H} p_n^{|H| - |J|} \prod_{\beta \in (H \setminus J)^{(2)}} \hat{Y}_\beta \prod_{\alpha \in J} Y_\alpha \\ &+ \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{\emptyset \neq J \subseteq H} p_n^{|H| - |J|} q_n^{|(H \setminus J)^{(2)}|} \prod_{\alpha \in J} Y_\alpha \\ &+ \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \sum_{\emptyset \neq J \subseteq H} p_n^{|H|} q_n^{|H^{(2)}| - |J^{(2)}|} \prod_{\beta \in J^{(2)}} \hat{Y}_\beta \end{aligned}$$

where  $Y_\alpha = X_\alpha - \mathbb{E}[X_\alpha | \mathbb{G}(n, q_n)]$  and  $\hat{Y}_\beta = \hat{X}_\beta - \mathbb{E}[\hat{X}_\beta]$ . It can be rewritten as  $N_G - \mathbb{E}[N_G] = \sum_J W_J^{(1)} + W_J^{(2)} + W_J^{(3)}$  where

$$\begin{aligned} W_J^{[1]} &= p_n^{e_G - |J|} q_n^{e_G^{(2)} - |J^{(2)}|} \left( \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G, H \supseteq J}} 1 \right) \prod_{\alpha \in J} Y_\alpha; & W_J^{[2]} &= p_n^{e_G - |J|} q_n^{e_G^{(2)} - |J^{(2)}|} p_n^{|J|} \prod_{\beta \in J^{(2)}} \hat{Y}_\beta; \\ & & & \\ W_J^{[3]} &= p_n^{e_G - |J|} \left( \sum_{\substack{H \in \binom{[n]}{3} \\ H \simeq G}} \prod_{\beta \in (H \setminus J)^{(2)}} \hat{Y}_\beta \right) \prod_{\alpha \in J} Y_\alpha. \end{aligned} \tag{3.23}$$

We consider the Malliavin structure associated to  $\mathbf{Y} = (Y_\alpha, \dots, \hat{Y}_\beta, \dots)_{\alpha \in \binom{[n]}{3}, \beta \in \binom{[n]}{3}}$ . Then, for each  $J$ , there exists  $m \in \mathbb{N}$  such that  $W_J^{(i)} \in \mathfrak{C}_m$  for  $i \in \{1, 2, 3\}$ .

**Theorem 3.4.5** Let  $G$  a hypergraph without isolated vertices. Then, let  $p_n \xrightarrow{n \rightarrow +\infty} 0$  and  $q_n \xrightarrow{n \rightarrow +\infty} 0$ :

$$d_W(\bar{N}_G, \mathcal{N}(0, 1)) \lesssim \left( \min_{\substack{H \subset G \\ \epsilon_H > 1}} \{n^{v_H} p_n^{\epsilon_H} q_n^{\epsilon_H^{(2)}}\} \right)^{-1/2}. \quad (3.24)$$

*Proof of Theorem 3.4.5.* We follow the same lines as the proof of Theorem 3.3.5, with the difference that  $\pi_0(N_G) = \mathbb{E}[F]$ . The (EGF) assumption holds. We recall the bound in our context

$$d_W(\bar{N}_G, \mathcal{N}(0, 1)) \leq C_{e_G} \sqrt{\sum_{(I, J, K, L) \text{ connected}} \sum_{i_i, i_j, i_k, i_l=1}^3 |\mathbb{E}[W_I^{[i_i]} W_J^{[i_j]} W_K^{[i_k]} W_L^{[i_l]}]| / \text{Var}[N_G]}. \quad (3.25)$$

As each connected quadruple  $(I, J, K, L)$  is associated to  $(H_1, H_2, H_3, H_4)$  subhypergraphs of  $G$  such that  $I \simeq H_1$ ,  $J \simeq H_2$ ,  $K \simeq H_3$  and  $L \simeq H_4$ , from Theorem 3.3.5, we have:

$$\begin{aligned} & |\mathbb{E}[W_I^{[1]} W_J^{[i_j]} W_K^{[i_k]} W_L^{[i_l]}]| \\ & \leq w_I w_J w_K w_L n^{v(H_1 \cup H_2 \cup H_3 \cup H_4)} p_n^{|H_1 \cup H_2 \cup H_3 \cup H_4|} q_n^{|H_1^{(2)} \cup H_2^{(2)} \cup H_3^{(2)} \cup H_4^{(2)}|} \end{aligned} \quad (3.26)$$

for  $i_j, i_k, i_l \in \{1, 2\}$  as  $\prod_{\alpha \in J} \mathbb{E}[Y_\alpha | Z] \propto_Z \prod_{\beta \in J^{(2)}} \hat{Y}_\beta$  and  $\mathbb{E}[\hat{Y}_\beta^k] \propto p_n$  for  $k \geq 2$ . As  $\hat{X}_\beta \leq 1$  a.s., we also have (3.26) for all  $i_i, i_j, i_k, i_l \in \{1, 2, 3, 4\}$ . Likewise, the variance reads off in function of the quadruples:

$$\text{Var}[N_G] = \frac{1}{4} \sum_{\substack{H_1, H_2 \subset G \\ H_1 \cap H_2 \neq \emptyset}} \left( \sum_I^{*H_1} \mathbb{E}[W_I^2] + \sum_J^{*H_2} \mathbb{E}[W_J^2] \right)$$

where  $\sum_I^{*H_1} \dots$  stands for a sum over  $I$  such that  $I$  is isomorphic to  $H_1$ .

We follow the same lines of computations as those leading to (3.16). Then, for fixed quadruples  $(H_1, H_2, H_3, H_4)$ ,

$$\text{Var}[N_G] \geq \frac{1}{16} \left( \sum_I^{*H_1} \mathbb{E}[W_I^2] \times \sum_J^{*H_2} \mathbb{E}[W_J^2] \times \sum_K^{*H_3} \mathbb{E}[W_K^2] \times \sum_L^{*H_4} \mathbb{E}[W_L^2] \right)^{1/2}. \quad (3.27)$$

As  $\mathbb{E}[W_I^2] = p_n^{|H| - |H_1|} q_n^{|H^{(2)}| - |H_1^{(2)}|} q_n^{|H_1^{(2)}|} (1 - q_n)^{|H_1^{(2)}|} (1 - p_n)^{|H_1|} p_n^{|H_1|}$ , we have:

$$\begin{aligned} & \sum_{(I, J, K, L) \text{ connected}} \sum_{i_i, i_j, i_k, i_l=1}^3 |\mathbb{E}[W_I^{[i_i]} W_J^{[i_j]} W_K^{[i_k]} W_L^{[i_l]}]| / \text{Var}[N_G] \\ & \leq \frac{n^{v(H_1 \cup H_2 \cup H_3 \cup H_4)} p_n^{|H_1 \cup H_2 \cup H_3 \cup H_4|} q_n^{|H_1^{(2)} \cup H_2^{(2)} \cup H_3^{(2)} \cup H_4^{(2)}|}}{\left( n^{v(H_1) + v(H_2) + v(H_3) + v(H_4)} p_n^{|H_1| + |H_2| + |H_3| + |H_4|} q_n^{|H_1^{(2)}| + |H_2^{(2)}| + |H_3^{(2)}| + |H_4^{(2)}|} \right)^{1/2}}. \end{aligned}$$

At that point, we arrive at the same upper bound as in the proof of Theorem 3.3.5.  $\square$

Some recent works have tackled either graph statistics based on conditionally independent random variables (Kaur and Röllin, 2021; Dey and Terlov, 2023) or statistics on a family of random

simplicial complex [Temčinas et al. \(2023\)](#) which is a type of representation of random hypergraph. For example, in [Kaur and Röllin \(2021\)](#), the authors apply Stein's method for multivariate functionals of conditionally independent random variables for  $p_n = p$  fixed, because in that case, the summands are uncorrelated. When  $p_n \rightarrow 0$ , there are correlations between them. While in those works, the probability of keeping a hyperedge does not depend on the number of vertices, we let  $p_n$  tend to 0. As a consequence, we can state thresholds for subhypergraph containment that complement the ones in [Janson et al. \(2000, p.61\)](#).

### 3.5 Future research directions

As in [Zhang \(2022\)](#); [Kaur and Röllin \(2021\)](#) for random graphs, it should be possible to derive with our method the convergence rates for the equivalent problem of subgraph counts in an arbitrary exchangeable random hypergraph generated by a hypergraphon, the analog of graphon in graph limit theory. Let denote in this section,  $\alpha_G : V(F) \rightarrow \mathbb{R}^+$  a weight function. In that context, the statistic of interest is the homomorphism density. Let  $F$  a simple graph, we denote

$$t(F, G) = \frac{1}{\alpha_G^{|V(F)|}} \text{hom}(F, G) \quad (3.28)$$

the probability that a random map is a homomorphism, i.e. preserves adjacency, with  $G$  weighted graph ( $\alpha_G = 1$ ) and:

$$\text{hom}(F, G) = \sum_{\phi: V(F) \rightarrow V(G)} \prod_{i \in V(F)} \alpha_G(\phi(i)) \prod_{ij \in E(F)} \beta_G(\phi(i), \phi(j)), \quad (3.29)$$

where  $\beta_G$  is the adjacency matrix of a random graph  $G$  and  $\phi$  maps edges to edges. It does not necessarily map non-edges to non-edges. This quantity can be adapted to be a statistic on graphon. Let  $\mathcal{W}$  the space of graphons, i.e. the space of measurable functions on  $[0, 1]^2$  which are symmetric. That expression has an equivalent when  $G$  becomes a graph limit called graphon. Given a graphon  $W \in \mathcal{W}$ , and a finite subset  $S \subseteq [0, 1]$ , we define the weighted graph  $W[S]$  on that node set, all nodes with weight 1, in which  $\beta_{xy}(W[S]) = W(x, y)$  for  $(x, y) \in S^2$ . In practice, one generates  $n$  random variables  $X_1, \dots, X_n$  from the uniform distribution on  $[0, 1]$  and consider the random graph  $W[\{X_1, \dots, X_n\}]$ . Then, we can define two sampled objects.

**Definition 3.5.1.** A  $W$ -random weighted graph  $H(n, W)$  on node set  $\{1, \dots, n\}$  is generated as follows: Let  $(X_i)_{i \in [1, n]}$  random variables with  $X_i \sim \mathcal{U}([0, 1])$ , and consider a realization:  $W[\{X_1, \dots, X_n\}]$ .

In the case, where  $W \in [0, 1]$  (basically it is a continuous probability matrix), one can define a  $W$ -random graph.

**Definition 3.5.2.** A random simple graph  $\mathbb{G}(n, W)$  is generated as follows:

1. Draw  $X_1, \dots, X_n \sim \mathcal{U}[0, 1]$ ;
2. For every pair of vertices  $\{i, j\}$ , keep the edge with probability  $W(X_i, X_j)$ .

**Example 3.5.3.** A stochastic block model (SBM) is a block-wise constant graphon in that framework (Gao and Ma, 2021). One generates a SBM with vertex set  $[n]$  by:

1. sampling independent latent random variables  $(U_i)_{i \in [1, n]}$  drawn according to the uniform distribution  $\mathcal{U}([0, 1])$ ;
2. adding an edge  $\{i, j\}$  for  $(i, j) \in [1, n]$  with probability  $W(U_i, U_j) = p_{i, j}$ .

The  $(p_{i, j})_{(i, j) \in [1, n]}$  can be encoded in a  $n \times n$  matrix  $P$ . The object  $W$  is the graphon associated to the SBM, and can be random. It is a function in  $\mathcal{F}([0, 1]^2, [0, 1])$  that takes value  $p_{i, j}$  on  $[\frac{i-1}{n}, \frac{i}{n}] \times [\frac{j-1}{n}, \frac{j}{n}]$ . Beware, the matrix  $P$  needs to be symmetric as to encode an undirected SBM model.

Every graphon defines a simple graph parameter as follows. If  $F$  is a simple graph with  $V(F) = \{1, \dots, k\}$ , then let

$$t(F, W) = \int_{[0, 1]^k} \prod_{ij \in E(F)} W(x_i, x_j) \, dx_1, \dots, dx_k. \tag{3.30}$$

This is the natural extension of the notion of homomorphism density in a graphon  $W$ . In the literature, the convergence mode towards a graphon relative to it is called *left-convergence*.

**Theorem 3.5.4** *For every left-convergent sequence  $(G_n)$  of simple graphs there is a graphon  $W$  with values in  $[0, 1]$  such that:*

$$t(F, G_n) \xrightarrow{n \rightarrow +\infty} t(F, W), \tag{3.31}$$

*for every simple graph  $F$ . Moreover, for every graphon  $W$  with values in  $[0, 1]$  there is a left-convergent sequence of graphs satisfying this relation.*

We say, then that the left-limit of the sequence of graphs is  $W$ .

**Remark 3.5.5.** A sequence of  $W$ -random graphs is left-convergent with limit  $W$ . It was shown in Janson and Diaconis (2008) that left-convergence is equivalent to convergence in distribution of random subgraphs  $H_{j, k}$  drawn by selecting  $k$  vertices independently of  $G_j$  for  $j > 0$ . ■

The definition was originally given in Lovász and Szegedy (2006).

One can extend those notions for hypergraphs. The adjacency matrix is a  $k$ D matrix instead of 2D array. The principle obstacle is the limiting object, hypergraphon which does not enjoy the natural properties of a graphon. We describe the limit not in the  $k$ -dimensional but in the  $2^k$ -dimensional space. Actually, the limit object turns out to be a subset, rather than a function, which is a gain. It is of course very little relative to the increase in the number of coordinates. Can one overcome that lift in dimensionality as to extend the limit theorem for graphons?



## Part II

# Invertibility of functionals of Poisson measures





---

The starting point of our second work is not related to the first part except for the fact that the underlying Dirichlet structure in the Poisson space does not have the diffusion property. It presents a challenge when one aims at extending classical results that hold for diffusive structures, e.g. fourth moment theorems in the first part. It is also related to the Malliavin calculus for Poisson measures (see e.g. [Bouleau and Denis \(2015\)](#)). Another approach to the stochastic calculus of variations is the Bismut's approach ([Bismut, 1983](#)). That approach for Markov processes with jumps is analogous to the Malliavin calculus of variations on diffusions ([Bismut, 1981](#)). The latter leverages the fact that the Brownian motion possesses a strong quasi-invariance property under a Cameron-Martin-Marwama-Girsanov transform ([Stroock and Varadhan, 1979](#)). It is well known that the law of Brownian motion with an adapted, square integrable drift is absolutely continuous with respect to the law of the Brownian motion. In parallel to Bismut's work, Üstünel and Zakai devoted their whole energy to extend the family of admissible drifts, that is to say drifts such that the absolute continuity property still holds. The main question is to get rid of the adaptability. They showed that this can be replaced by, for instance, either monotony or some regularity on the Malliavin derivative of the drift. Most of their results are contained in [Üstünel and Zakai \(2000\)](#), most notably the entropic criterion that implies the strongness of a solution.

On the Poisson space, the analog of Girsanov transform is of different nature. The drift is replaced with is a change of locations of points in the configuration. The Bismut's approach has led to a collection of works on point processes which are described in [Bichteler et al. \(1987\)](#). Most results on the methods of transformations of point processes are covered in [Jacod \(1979\)](#). It is only recently that the point of view of Üstünel and Zakai has been reprised in the Poisson space ([Coutin and Decreusefond, 2023](#)). The studied transformation are time changes. Our work is the natural continuation when we consider a marked Poisson process as canonical element. Both approaches can be used as to define a Girsanov transform ([Decreusefond, 1998](#)). After an introduction of that type of transformation in the first chapter of that part, we study the framework of invertibility for such transforms.



## Chapter 4

# Girsanov transforms of Poisson measures

In [Section 4.1](#), we give the necessary background on the notion of Girsanov transform on the Poisson space as to approach the invertibility framework. In [Section 4.2](#), we recall some properties on marked point processes. Then, in [Section 4.3](#), we define the notion of random change of marks. Finally, in [Section 4.4](#), we provide with the Girsanov theorem for Poisson space.

### 4.1 Background

Denote  $W_B$  the Banach space of continuous functions on  $[0, 1]$ . We consider the stochastic differential equation:

$$X(0) = 0 \text{ and } dX(t) = \dot{u}(X(t)) dt + dB(t) \quad (4.1)$$

Consider the map

$$\begin{aligned} U : W_B &\longrightarrow W_B \\ \omega &\longmapsto \left( t \mapsto \omega(t) - \int_0^t \dot{u}(\omega(s)) ds \right). \end{aligned} \quad (4.2)$$

Let us assume that (4.1) has a weak solution whose law is noted  $\nu$ . Further, assume the weak solution is unique denoted  $V$ . It corresponds to an adapted morphism of probability between  $(W_B, \mathcal{F}, \nu)$  and  $(W_B, \mathcal{F}, \mu)$ . Then, it is equivalent to say that:

$$U \circ V = \text{Id}_{W_B}.$$

The celebrated Girsanov theorem (cf. [Girsanov \(1960\)](#); [Orey \(1974\)](#)) yields the change of variables formula, i.e. setting

$$\rho_U(w) = \exp \left( - \int_0^1 \dot{u}(s) dw(s) - \frac{1}{2} \int_0^1 |\dot{u}(s)|^2 ds \right)$$

and assuming  $\mathbb{E}[\rho_U] = 1$ , then, for smooth  $f$ , it holds true that:

$$\mathbb{E}[f \circ U \rho_U] = \mathbb{E}[f].$$

Hence, the image measure  $U_{\#}\mu$  is absolutely continuous with respect to the Wiener measure  $\mu$ . Let  $\Lambda_u$  be the corresponding Radon-Nikodym derivative. The Girsanov transform is related to

that transformation  $U$  which yields that Radon-Nikodym derivative  $\Lambda_u$ . That terminology is introduced in the monograph [Bichteler et al. \(1987\)](#) in the context of the Poisson space.

The Skorokhod theorem on invariance of measures ([Skorokhod, 1957](#)) gives the density with respect to Poisson measures of deterministic shifts of configuration. This theorem has an extension to random transformation of marks of a marked Poisson process  $N$ . The classical Girsanov theorem for Poisson process yields a Radon-Nikodym density for the change of measure, expressed in function of the Doléans-Dade exponential ([Kallenberg, 2002](#), Theorem 23.8). In that case, the key property is quasi-invariance with respect to anticipative transformation ([Privault, 1996](#); [Albeverio and Smorodina, 2006](#)) which is given by a classical application of the Girsanov theorem. The specificity of the Poisson setting is the identification of  $\mathfrak{M}$  the space of random measures and the configuration space. Namely, let  $\Gamma$  the transformation applied on a Poisson measure  $N$  under a probability measure  $\pi$  on  $\mathfrak{M}$ . As  $N$  can also be considered a marked point process,  $\Gamma$  transposes to a random transformation of an element  $\omega \in \mathfrak{N}$ , namely the marks  $Z_n$  of  $\omega = (T_n, Z_n)_{n \in \mathbb{N}}$  are changed to  $\gamma(T_n, Z_n)$  for  $n \in \mathbb{N}$ . Under some assumptions on  $\Gamma$ , there exists a probability measure  $\pi'$  under which both marked point processes  $\Gamma(N)$  and  $N$  share the same law. The Girsanov theorem ([Boel et al., 1975](#)) also gives the proof of the existence of weak solution of stochastic differential equation driven by Brownian motion or Poisson measures or both under weak assumptions. The stochastic differential equation (SDE for short) that we consider in this chapter, is of the form:

$$\begin{cases} Y(0) &= 0 \\ dY(t) &= \int \int b(Y(t_-), x) dN(t, x) \end{cases} \quad (4.3)$$

where  $N$  is a Poisson measure with control measure  $\nu(dx) \otimes dt$ . Denote by  $\mathfrak{M}$  the space of random counting processes on  $\mathbb{R}^+ \times \mathbb{R}$ . By identification of random measures and random counting processes,  $N$  is an element of  $\mathfrak{M}$ . Let  $\pi$  a probability on the space  $\mathfrak{M}$  such that  $N$  is the canonical counting process on it, namely  $\text{Id}_{\mathfrak{M}}$ . Let  $\mathbb{D}$  the space of càdlàg functionals. Consider the map

$$\begin{aligned} V : \mathfrak{M} &\longrightarrow \mathbb{D} \\ N &\longmapsto \left( t \mapsto \int_0^t \int x dN(s, x) \right). \end{aligned} \quad (4.4)$$

Let  $\Upsilon$  a perturbation of the sample path and by identification of  $\omega \in \mathfrak{N}$  such that:

$$\Upsilon(\omega)(t, z) = (T_n, b(\sum_{n \in \mathbb{N}: T_n < t} Z_n(\omega), z)),$$

The primary motivation of this work was to study the solution of the SDE (4.3) as defined by:

$$Y = V \circ \Upsilon(N).$$

It induces a map  $\mathbf{Y}$  on  $\mathbb{D}$  such that  $\mathbf{Y} \circ V(N) = Y$ . Let  $\mathbf{Z}$  a map on  $\mathbb{D}$  of the form  $V \circ \Gamma$  for  $\Gamma$  some random transformation, we have:

$$\mathbf{Y} \circ \mathbf{Z} = \text{Id}_{\mathbb{D}} \iff \Upsilon \circ \Gamma(N) = \text{Id}_{\mathfrak{M}}.$$

Solving (4.3) revolves to invert the map  $\Upsilon$ . That formulation of the problem puts the focus on the map  $\Upsilon$  on  $\mathfrak{M}$  instead of  $Y \in \mathbb{D}$ . To the best of our knowledge, this approach has been seldom addressed in the literature.

## 4.2 Preliminaries

Let  $\mathbb{M}$  a metric space and  $\mathcal{B}(\mathbb{M})$  its Borel  $\sigma$ -algebra. We recall the construction of marked point processes which is a sequence of points drawn at random in  $\mathbb{R}^+ \times \mathbb{M}$  in the configuration space  $\mathfrak{N}$  endowed with the vague convergence, its Borel  $\sigma$ -algebra  $\mathcal{B}(\mathfrak{N})$  and a probability measure  $\mu$  (see for instance [Kallenberg, 2017](#)).

**Definition 4.2.1** (Marked point process). A marked point process (MPP for short) is a sequence

$$\omega = (T_n(\omega), Z_n(\omega))_{n \in \mathbb{N}} \text{ of } \mathbb{R}^+ \times \mathbb{M} \text{ defined on } (\mathfrak{N}, \mathcal{B}(\mathfrak{N}), \mu)$$

such that for each  $n \geq 1$ ,  $T_n(\omega) < T_{n+1}(\omega)$ ,  $T_n(\omega)$  tends to infinity  $\mu$ -a.s. as  $n$  tends to infinity. The random variable  $T_n$  represents the  $n$ -th jump time and  $Z_n$  is the location associated to the  $n$ -th jump. It can be completely identified with a random counting measure  $\xi$ , viz.

$$\xi = \sum_{n \in \mathbb{N}: T_n(\omega) < \infty} \epsilon_{(T_n, Z_n)} \quad (4.5)$$

where  $\epsilon_{(T_n, Z_n)}(\omega) = \epsilon_{(T_n(\omega), Z_n(\omega))}$  is the Dirac mass on the product space  $((0, +\infty) \times \mathbb{M})$  at the point  $(T_n(\omega), Z_n(\omega))$ . We denote  $\mathfrak{M}$  the space of counting random measures that can be written in such form. We endow this Lusin space with the vague convergence as a family of positive integer-valued Radon measures  $(\xi(\omega, \cdot) : \omega \in \mathfrak{M})$  on  $(\mathbb{R}^+ \times \mathbb{M}, \mathcal{B}(\mathbb{R}^+ \times \mathbb{M}))$ . The random measure can be identified to a random counting process  $X$ . For  $A \in \mathcal{B}(\mathbb{M})$ , we shall write the process

$$X(\omega, t, A) = \xi(\omega, [0, t] \times A)$$

that counts the number of events on  $[0, t]$  matching a mark belonging to the set  $A$ . We define the ground process.

$$X_g(\omega, t) = X(\omega, t, \mathbb{M}).$$

Note that  $\omega$  can be retrieved as:

$$T_n(\omega) = \inf\{t \geq 0 : X_g(\omega, t) = n\};$$

$$(Z_n(\omega) \in A) = \bigcup_{K'=1}^{\infty} \bigcap_{K=K'}^{\infty} \bigcup_{k=1}^{\infty} \left( X_g(\omega, (k-1)/2^k = n-1, X_g(\omega, k/2^K, A) - X_g(\omega, (k-1)/2^K, A) = 1 \right).$$

That means that we can identify a sample-path  $X$  of  $\mathfrak{M}$  with an element  $\omega$  of  $\mathfrak{N}$ , and more generally elements of  $\mathcal{B}(\mathfrak{M})$  and  $\mathcal{B}(\mathfrak{N})$ . We redefine  $\mu$  as a probability on  $(\mathfrak{M}, \mathcal{B}(\mathfrak{M}))$ . Moreover,  $dX(s, z) = d\xi(s, z)$ .

Let  $(\mathcal{F}_t)_{t \in \mathbb{R}^+}$  be an increasing and right-continuous family of sub-algebras of  $\mathcal{B}(\mathfrak{M})$  such that each  $T_n$  is a stopping time, and each  $Z_n$ , is  $\mathcal{F}_{T_n}$ -measurable such that  $\mathcal{B}(\mathfrak{M})$  is included in  $\bigvee_{t \geq 0} \mathcal{F}_t = \mathcal{F}_{\infty}$ . We denote the following subsets of  $\mathcal{P}(\mathcal{F})$  by:

- $\mathcal{P}^+(\mathcal{F})$  the set of predictable non-negative real-valued processes  $(Y(t, z))_{(t, z) \in \mathbb{R}^+ \times \mathbb{M}}$  defined on  $(\mathfrak{M}, \mathcal{B}(\mathfrak{M}), \mu)$ ;
- $\mathcal{P}^{++}(\mathcal{F})$  the set of predictable positive real-valued processes;

$(Y(t, z))_{(t, z) \in \mathbb{R}^+ \times \mathbb{M}}$  defined on  $(\mathfrak{M}, \mathcal{B}(\mathfrak{M}), \mu)$ .

We recall the version of Theorem 2.1 of Jacod (1975) for random counting processes.

**Theorem 4.2.2** *Let  $\mu$  be a probability measure on  $(\mathfrak{M}, \mathcal{F}_\infty)$ . Let a random measure  $\xi$  on  $\mathbb{R}^+ \times \mathbb{M}$  such that  $t \mapsto \xi([0, t], \cdot)$  is adapted. Then there exists a unique predictable random measure denoted  $\xi^p$  such that*

$$\left( \xi([0, t], A) - \xi^p([0, t], A) \right)_{t \geq 0} \text{ is a } (\mathcal{F}, \mu)\text{-local martingale.}$$

*The random measure  $\xi^p$  is called the dual predictable projection of  $\xi$  along  $\mathcal{F}$ . Equivalently, it means that for each  $A \in \mathcal{B}(\mathbb{M})$ , the process  $t \mapsto \xi^p([0, t], A)$  is the so-called compensator of  $t \mapsto \xi([0, t], A)$ . We denote by  $\tilde{\xi}$  the random measure  $\xi - \xi^p$ .*

In this part, we denote expectation and probability respectively  $\mathbf{E}$  and  $\mathbf{P}$ .

**Definition 4.2.3.** Let  $\xi$  a random measure. We introduce  $\mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi)$  the space of all functions  $f$  in  $\mathcal{P}(\mathcal{F})$  such that for each  $t \in \mathbb{R}^+$ ,

$$\mathbf{E}_\mu \left[ \int_0^t \int_{\mathbb{M}} |f(\cdot, s, z)| \xi(ds, dz) \right] < +\infty.$$

The Bochner integral  $\int_0^t \int_{\mathbb{M}} f(\cdot, s, z) \xi(ds, dz)$  is a Lebesgue integral with respect to the measure  $\xi(\omega, \cdot)$  for every  $\omega \in \mathfrak{M}$  and is equal to the convergent sum

$$\int_0^t \int_{\mathbb{M}} f(\cdot, s, z) \xi(ds, dz) = \sum_{s \in (0, t] \cap \{s: \Delta X(s) \neq 0\}} f(\cdot, s, \Delta X(s)).$$

The stochastic integral of  $r$  with respect to  $\xi$  denoted by  $\delta_\xi r$ , is the process on  $\mathbb{R}^+$

$$\delta_\xi r(t) = \int_0^t \int_{\mathbb{M}} r(s, z) d\xi(s, z). \quad (4.6)$$

If  $f \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi^p) \cap \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi)$ , we define the integral with respect to the compensated random measure  $N$  by

$$\int_0^t \int_{\mathbb{M}} f(s, z) \tilde{\xi}(ds, dz) = \int_0^t \int_{\mathbb{M}} f(s, z) \xi(ds, dz) - \int_0^t \int_{\mathbb{M}} f(s, z) d\xi^p(s, z) \quad \mu\text{-a.s.} \quad (4.7)$$

**Remark 4.2.4.** In general, the stochastic integral cannot be divided as the difference of two integrals as in (4.7). ■

**Proposition 4.2.5** *If  $f \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi^p)$ , then we have  $f \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi)$  and for each  $t \geq 0$ ,*

$$\mathbf{E}_\mu \left[ \int_0^t \int_{\mathbb{M}} f(s, z) \xi(ds, dz) \right] = \mathbf{E}_\mu \left[ \int_0^t \int_{\mathbb{M}} f(s, z) \xi^p(ds, dz) \right].$$

We define the Doléans-Dade exponential associated to a random measure  $\xi$ .

**Definition 4.2.6.** For  $f \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \tilde{\xi})$ , additionally supposed to be non-negative, the Doléans-Dade exponential (Jacod, 1975), denoted by  $\mathcal{E}(\delta_{\tilde{\xi}}f)$ , is defined as the solution of

$$M(t) = \int_0^t \int M(s^-) f(s, z) d\tilde{\xi}(s, z),$$

explicitly given by

$$\tilde{\mathcal{E}}(\delta_{\tilde{\xi}}f)(t) = \exp\left(\int_0^t \int (1 - f(s, z)) d\tilde{\xi}(s, z)\right) \prod_{T_n \leq t} f((T_n, Z_n)). \quad (4.8)$$

In the case  $r > -1$  and  $r \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \xi^p)$ , it is explicitly given by:

$$\begin{aligned} \mathcal{E}(\delta_{\xi}f)(t) &= \exp\left(\int_0^t \int \log(f(s, z)) d\xi(s, z) - \int_0^t \int (f(s, z) - 1) d\xi^p(s, z)\right) \\ &= \exp\left(\delta_{\xi}(\log(f))(t) + \int_0^t \int (\log(f(s, z)) - f(s, z) + 1) d\xi^p(s, z)\right). \end{aligned} \quad (4.9)$$

An important assumption in the remainder is that  $\mathcal{F}$  is the minimal filtration to which  $X$  is adapted, viz.  $\mathcal{F}$  is the filtration on  $\mathfrak{M} \times \mathbb{R}^+$

$$\mathcal{F}^X = \sigma(X(t, A), t \in \mathbb{R}^+, A \in \mathcal{B}(\mathbb{M})). \quad (\text{A1})$$

The next result is the converse of [Theorem 4.2.2](#) (see Proposition 3.41 of [Jacod, 1979](#)).

**Theorem 4.2.7** *For any nonnegative random measure  $\xi^p$  on  $(0, \infty) \times \mathbb{M}$ , there exists a unique probability measure on  $(\mathfrak{M}, \mathcal{F}^X)$  such that  $X$  has  $\xi^p$  as dual predictable projection.*

### 4.3 Random change of marks

Following the Bismut's approach ([Bismut, 1983](#)), we devise a perturbation  $\gamma$  of the jump sizes such that the map  $\Gamma = \Gamma(N)$  under  $\pi$  has the same law as  $N$  under  $\pi_\phi$ , i.e.  $\Gamma_{\#}\pi = \pi_\phi$ . It consists of a random change of marks in analogy to the random change of time for point process ([Coutin and Decreusefond, 2023](#)).

**Definition 4.3.1** (Random transformation of mark). On  $(\mathfrak{M}, \mathcal{F}_\infty, \mu)$ , a random change of mark is a process  $(\gamma(s, z), (s, z) \in \mathbb{R}^+ \times \mathbb{M})$  such that:

- for any  $z \in \mathbb{M}$   $\nu$ -a.s.,  $s \mapsto \gamma(s, z) \in \mathcal{P}(\mathcal{F})$ ;
- $\forall (s, u) \in \mathbb{R}^+ \times \mathbb{M}$ ,  $\gamma(s, u) \in \mathbb{M}$ .

For any  $s \in \mathbb{R}^+$ , we say that  $\gamma$  is  $\mu$ -invertible with inverse denoted  $\gamma^* : \mathfrak{M} \times \mathbb{R}^+ \times \mathbb{M} \rightarrow \mathbb{M}$  if given  $z \in \mathbb{M}$ , for almost all  $\omega \in \mathfrak{M}$ ,  $s \in \mathbb{R}^+$ :

$$\gamma(\omega, s, \gamma^*(\omega, s, z)) = \gamma^*(\omega, s, \gamma(\omega, s, z)) = z \text{ for all } s \in \mathbb{R}^+ \quad dt \otimes \mu\text{-a.s.}$$

The transformation of mark is a very common operation on Poisson measures. It has been an important development in the theory of Malliavin calculus for perturbation analysis (see [Bichteler and Jacod, 1983](#)).



**Definition 4.3.2.** Let  $X$  a random counting process on  $(\mathbb{R}^+ \times \mathbb{M}, \mathcal{B}(\mathbb{R}^+ \times \mathbb{M}))$  and  $\gamma$  invertible transformation of mark. Then, the changed random process is defined by

$$\mathbf{\Gamma}(X)(t, A) = \int_0^t \int \mathbf{1}_A(\gamma(s, z)) X(ds, dz) \quad \text{for } (t, A) \in \mathbb{R}^+ \times \mathcal{B}(\mathbb{M}).$$

If  $X \in \mathfrak{M}$ ,  $\mathbf{\Gamma}(X) \in \mathfrak{M}$ .

**Lemma 4.3.3** *Let  $\gamma$  an invertible transformation, we have:*

$$\mathcal{F}_t^{\mathbf{\Gamma}(X)} \vee \sigma(\gamma^*(s, z), s \leq t, z \in \mathbb{M}) = \mathcal{F}_t^X \quad (4.10)$$

*Proof of Lemma 4.3.3.* We have by construction of  $\mathbf{\Gamma}(X)$ ,  $\mathcal{F}_t^{\mathbf{\Gamma}(X)} = \sigma(\mathbf{\Gamma}(X)(s, A), s \leq t, A \in \mathcal{B}(\mathbb{M})) \subset \mathcal{F}_t^X$  relative to the Poisson measure  $X$ . By predictability of  $\gamma^*$ ,  $\mathcal{F}_t^{\mathbf{\Gamma}(X)} \vee \sigma(\gamma^*(s, z), s \leq t, z \in \mathbb{M}) \subset \mathcal{F}_t^X$ . Conversely, for  $A \in \mathcal{F}_t^N$ , there exists a sequence of measurable functions  $(\chi_q)_{q \in \mathbb{N}}$  such that for any  $q$ ,  $\chi_q$  is measurable function from  $(\mathbb{R}^+ \times \mathbb{M})^q$  to  $\{0, 1\}$ :

$$\begin{aligned} \mathbf{1}_A &= \sum_{q=0}^{\infty} \chi_q(T_1(X), \dots, T_q(X), Z_1(X), \dots, Z_q(N)) \mathbf{1}_{\{T_q(X) < t \leq T_{q+1}(X)\}} \\ &= \sum_{q=0}^{\infty} \chi_q^{\mathbf{\Gamma}(X)} \mathbf{1}_{\{T_q(\mathbf{\Gamma}(X)) < t \leq T_{q+1}(\mathbf{\Gamma}(X))\}} \end{aligned}$$

where

$$\begin{aligned} \chi_q^{\mathbf{\Gamma}(X)} &= \chi_q(T_1(\mathbf{\Gamma}(X)), \dots, T_q(\mathbf{\Gamma}(X)), \gamma^*(X, T_1(\mathbf{\Gamma}(X)), Z_1(\mathbf{\Gamma}(X))), \\ &\quad \dots, \gamma^*(X, T_q(\mathbf{\Gamma}(X)), Z_q(\mathbf{\Gamma}(X)))). \end{aligned}$$

Because  $\gamma^*(N, \cdot)$  is measurable from  $\mathbb{R}^+ \times \mathbb{M}$  to  $\mathbb{R}^+$ , it is a limit of simple functions:

$$\begin{aligned} &\gamma^*(X, T_k(\mathbf{\Gamma}(X)), Z_k(\mathbf{\Gamma}(X))) \\ &= \lim_{n \rightarrow \infty} \sum_{i=0}^{2^n-1} \sum_{j=0}^{2^m-1} \gamma^*(X, \frac{i}{2^n}, \frac{j}{2^m}) \mathbf{1}_{[\frac{i-1}{2^n}, \frac{i}{2^n})} (T_k(\mathbf{\Gamma}(X))) \mathbf{1}_{[\frac{j-1}{2^m}, \frac{j}{2^m})} Z_k(\mathbf{\Gamma}(X)). \end{aligned}$$

Hence, the result holds.  $\square$

In the remainder, we consider as canonical marked point processes the stationary marked Poisson point process  $N$  with intensity measure  $ds \otimes \nu(dz)$  where  $\nu$  is a  $\sigma$ -finite measure on  $(\mathbb{M}, \mathcal{B}(\mathbb{M}))$  where  $ds$  is the Lebesgue measure on  $\mathbb{R}$  and the associated probability measure  $\pi$  on  $(\mathfrak{M}, \mathcal{F}_\infty)$  by [Theorem 4.2.7](#). We denote  $\Gamma$  the process  $\mathbf{\Gamma}(N)$ . Under  $\pi$ , we identify  $\omega$  and the canonical process  $N$ . We denote by  $\pi_\sigma$  the unique probability measure on  $(\mathfrak{M}, \mathcal{F}_\infty^N)$  such that  $N^p = \sigma(s, z)\nu(dz) ds$  for  $\sigma$  a random process on  $\mathfrak{M}$ . When  $\sigma = 1$ , we recover  $\pi = \pi_{\text{Id}}$ . We keep the notation  $\pi$  for sake of simplicity.

## 4.4 Change of measures and Girsanov theorem

Here we consider a probability  $\pi'$  with the basic hypothesis that  $\pi'$  is locally absolutely continuous with respect to  $\pi$  along  $\mathcal{F}$ . We write it  $\pi' \ll_{\text{loc}, \mathcal{F}} \pi$ . Our aim is to compute the dual predictable projection of  $N$  under  $\pi'$ . This is the object of the so-called Girsanov theorems. The next theorem is a combination of [Jacod \(1975, Theorem 4.5\)](#) and [Decreusefond \(1998, p.500\)](#). The proof of the criterion for the Radon-Nikodym derivative to be non-zero can be found in the comprehensive book [Jacod and Shiryaev \(2003, Theorem 5.19 and Corollary 5.22\)](#).

**Theorem 4.4.1** *Consider a filtration  $\mathcal{F} = (\mathcal{F}_t, t \geq 0)$  on  $\mathfrak{M}$ , and assume that under  $\pi$ , the process*

$$\tilde{N} : t \mapsto N(t, A) - \int_0^t \int_{\mathbb{M}} \mathbf{1}_A(z) \nu(dz) ds$$

*is a  $\mathcal{F}$ -local martingale. If a probability measure  $\pi'$  on  $(\mathfrak{M}, \mathcal{F}_\infty^N)$  is locally absolutely continuous with respect to  $\pi$  along  $\mathcal{F}$  then there exists a unique process  $\phi$  on  $\mathbb{R}^+ \times \mathbb{M}$  which is non-negative  $\pi$ -a.s. and  $\mathcal{F}$ -predictable such that*

$$\forall t \geq 0, \pi' \left( \int_0^t \int_{\mathbb{M}} \left(1 - \sqrt{\phi(s, z)}\right)^2 \nu(dz) ds < \infty \right) = 1 \quad (4.11)$$

*and for any  $A \in \mathcal{B}(\mathbb{M})$ ,*

$$t \mapsto N(t, A) - \int_0^t \int_A \phi(s, z) \nu(dz) ds \text{ is a } (\mathcal{F}, \pi_\phi)\text{-local martingale.} \quad (4.12)$$

*We refer to  $\phi$  as the Girsanov factor of  $\pi' := \pi_\phi$  with respect to  $\pi$ . Moreover, assume [\(A1\)](#), for any integer  $m \geq 1$ , let*

$$S_m = \inf \left\{ t \in \mathbb{R}^+, \int_0^t \int_{\mathbb{M}} \left(1 - \sqrt{\phi(s, z)}\right)^2 \nu(dz) ds \geq m \right\}.$$

*Then, with the previously introduced notations,*

$$\begin{aligned} \Lambda_\phi(t) &:= \frac{d\pi_\phi}{d\pi} \Big|_{\mathcal{F}_t} \\ &= \begin{cases} \mathcal{E}(\delta_{\tilde{N}}\phi)(t) & \text{if } t \leq S_m, \\ 0 & \text{if } t \geq \limsup_m S_m. \end{cases} \end{aligned}$$

*If  $\pi_\phi$  is absolutely continuous with respect to  $\pi$  on  $\mathcal{F}_\infty$  then*

$$\pi_\phi \left( \int_0^\infty \int_{\mathbb{M}} \left(1 - \sqrt{\phi(s, z)}\right)^2 \nu(dz) ds < \infty \right) = 1. \quad (4.13)$$

*For the converse part, let us suppose [\(A1\)](#). Consider  $\phi$  a non-negative  $\mathcal{F}$ -predictable process and  $\pi_\phi$  the probability measure on  $(\mathfrak{M}, \mathcal{F}_\infty^N)$ , which satisfies [\(4.11\)](#) and [\(4.12\)](#). Then,  $\pi_\phi$  is locally absolutely continuous with respect to  $\pi$  along  $\mathcal{F}$ . Finally, the probability measure  $\pi_\phi$  is absolutely continuous with respect to  $\pi$  on  $(\mathfrak{M}, \mathcal{F}_\infty)$  if and only if [\(4.13\)](#) is satisfied.*

**Definition 4.4.2.** Let  $\mathcal{F}$  a filtration relative to  $N$ , we introduce the following subset of  $\mathbb{F}^1(\mathcal{P}(\mathcal{F}), \nu \otimes ds)$  denoted  $\mathbb{F}_2^1(\mathcal{P}(\mathcal{F}), \nu \otimes ds; \pi)$  such that the following properties hold for an element  $\phi$ :

1.  $\pi_\phi \ll_{\text{loc}, \mathcal{F}} \pi$ ;
2.  $\pi_\phi \left( \int_0^\infty \int \left(1 - \sqrt{\phi(s, z)}\right)^2 \nu(dz) ds < \infty \right) = 1$ .

The change of measure is associated either with a change of time or a change of mark. The first one is covered in [Coutin and Decreusefond \(2023\)](#). The use of change of marks dates back to the approach pioneered by [Bismut \(1983\)](#). We take the latter approach. We fix  $\mathbb{M} = \mathbb{R}^+$  endowed with the Lebesgue measure, and  $\nu(dz) = \rho(z) dz$  with  $\rho : \mathbb{R} \rightarrow \mathbb{R}$  a non-negative integrable function. The central notion elaborated by Bismut is the Girsanov transform which is key to our investigations.

**Definition 4.4.3** (Girsanov transform). Let  $\phi \in \mathcal{P}^{++}(\mathcal{F})$ . Define  $\gamma_\phi^\rho$ , for  $\pi \otimes ds$  all  $\omega \in \mathfrak{M}$  and  $u \geq 0$  as a Lebesgue integral:

$$\gamma_\phi^\rho(\omega, s, z) = \int_0^z \phi(\omega, s, u) \rho(u) du.$$

Then, setting  $r(z) = \int_0^z \rho(u) du$ , the random transformation of mark

$$\gamma_\phi(\omega, s, \cdot) = r^{-1} \circ \gamma_\phi^\rho(\omega, s, \cdot).$$

is a  $C^1$ -diffeomorphism and called *Girsanov transform*. Moreover, it is  $\pi$ -invertible in the sense of [Definition 4.3.1](#) with inverse denoted  $\gamma_\phi^*$ . In the same vein, we define the process:

$$\phi^\dagger(\omega, s, z) = \frac{1}{\phi(\omega, s, \gamma_\phi^*(\omega, s, z))}.$$

Using the [Definition 4.3.2](#), given  $\phi$ , we have the triplet  $(\gamma_\phi, \mathbf{\Gamma}_\phi, \phi^\dagger)$  that characterizes the Girsanov transform.

**Remark 4.4.4.** We have for  $s \in \mathbb{R}^+$ :

$$(\gamma_\phi(s, \cdot))'(u) = \frac{\phi(s, u) \rho(u)}{\rho(\gamma_\phi(s, u))}$$

and

$$(\gamma_\phi^*(s, \cdot))'(u) = \frac{\rho(u)}{\rho(\gamma_\phi^*(s, u)) \phi(s, \gamma_\phi^*(s, u))} = \frac{\rho(u) \phi^\dagger(s, u)}{\rho(\gamma_\phi^*(s, u))}. \quad (4.14)$$

For instance, if  $\rho = 1$ , we have:

$$(\gamma_\phi(s, \cdot))'(z) = \phi(s, z) \text{ and } (\gamma_\phi^*(s, \cdot))'(z) = \phi^\dagger(s, z).$$

■

**Remark 4.4.5.** The generalization of the Girsanov transform to  $\mathbb{R}^d$  for  $d \geq 2$  is immediate as the remark 4.4.4 still holds. The difference with the case  $d = 1$  lies in the stochastic integration rules. Because  $\phi > 0$ ,  $\pi$ -a.s., it corresponds for any  $\omega \in \mathfrak{M}$   $\pi$ -a.s. and each  $s \in \mathbb{R}^+$  to:

$$\gamma'_\phi(\omega, s, x) = |\det(J_{\gamma_\phi(\omega, s, \cdot)}(x))|$$

for  $U$  an open set in  $\mathbb{R}^d$  and  $\gamma_\phi^*(s, \cdot) : U \rightarrow \mathbb{R}^d$  an injective differentiable function with continuous partial derivatives, the Jacobian of which is nonzero for  $x \in U$ . We notice that not all the components of  $\gamma_\phi^*$  need to be nonnegative for  $d \geq 2$ . For sake of readability, this generalization is omitted. ■

**Lemma 4.4.6** For  $f \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \nu \otimes dt)$ , then we have:

$$\begin{aligned} \Gamma_\phi \left( \int_0^\cdot \int f(N, s, z) \nu(dz) ds \right) \\ = \int_0^\cdot \int f(N, s, \gamma_\phi^*(N, s, z)) \phi^\dagger(N, s, z) \nu(dz) ds \end{aligned} \quad (4.15)$$

and:

$$\Gamma_\phi \left( \int_0^\cdot \int f(N, s, z) N(ds, dz) \right) = \int_0^\cdot \int f(N, s, \gamma_\phi^*(N, s, z)) d\Gamma_\phi(s, z). \quad (4.16)$$

*Proof of Lemma 4.4.6.* For  $t \in \mathbb{R}^+$  and  $A \in \mathcal{B}(\mathbb{R}^+)$ ,

$$\Gamma_\phi \left( \int_0^\cdot \int f(N, s, z) \nu(dz) ds \right) (N, t, A) = \int_0^\cdot \int \mathbf{1}_A(\gamma_\phi(N, s, z)) f(N, s, z) \nu(dz) ds.$$

In particular, for  $a, b \in \mathbb{R}^+$  with  $a < b$ ,  $C \in \mathcal{B}(\mathbb{M})$  and  $B \in \mathcal{F}_a$ , let

$$f(N, s, z) = B(N) \mathbf{1}_{(a, b]}(s) \mathbf{1}_C(z).$$

We have by the theorem of change of variables in terms of Lebesgue measure on  $\mathbb{R}$ , with  $u = \gamma_\phi(N, s, z)$  for each  $s \in [0, t]$ :

$$\int \mathbf{1}_A(\gamma_\phi(N, s, z)) \mathbf{1}_C(z) \nu(dz) = \int \mathbf{1}_A(u) \mathbf{1}_C(\gamma_\phi^*(N, s, u)) \phi^\dagger(N, s, u) \nu(du)$$

since  $(\gamma_\phi^*)'(N, s, \cdot)(u) \rho(\gamma_\phi^*(N, s, u)) = \frac{1}{\phi(N, s, \gamma_\phi^*(N, s, u))} \rho(u)$ . Then,

$$\begin{aligned} \int \int \mathbf{1}_A(\gamma_\phi(N, s, z)) f(N, s, \gamma_\phi^*(s, u)) \nu(dz) ds \\ = \int \int \mathbf{1}_A(u) f(N, s, \gamma_\phi^*(s, u)) \phi^\dagger(N, s, u) \nu(du) ds. \end{aligned}$$

By density of simple processes in  $\mathbb{F}^1(\mathcal{F}^N, \nu \otimes ds)$ , this yields (4.15). We apply the same change of variable for the proof of the second identity.

$$\begin{aligned} \Gamma_\phi \left( \int_0^\cdot \int f(N, s, z) N(ds, dz) \right) (t, A) &= \int_0^t \int \mathbf{1}_A(\gamma_\phi(N, s, z)) f(N, s, z) N(ds, dz) \\ &= \int_0^t \int \mathbf{1}_A(u) f(N, s, \gamma_\phi^*(N, s, u)) \Gamma(ds, du). \end{aligned}$$

Since it holds for each  $A \in \mathcal{B}(\mathbb{R}^+)$ , it yields (4.16).  $\square$

**Corollary 4.4.7** *Let  $\phi \in \mathcal{P}^{++}(\mathcal{F}^N)$ , then  $\phi^\dagger \in \mathcal{P}^{++}(\mathcal{F}^{\Gamma_\phi})$ .*

**Remark 4.4.8.** The factor  $\phi^\dagger$  is a Girsanov factor as well as  $\phi$ .  $\blacksquare$

**Theorem 4.4.9 — Girsanov theorem for Poisson measures.** *Let  $\pi' \ll_{loc, \mathcal{F}^N} \pi$  and  $\phi \in \mathcal{P}(\mathcal{F}^N)$  denote its Girsanov factor, i.e.  $\pi' = \pi_\phi$ . Assume that  $\phi$  belongs to  $\mathbb{F}_2^1(\mathcal{P}^{++}(\mathcal{F}^N), \nu \otimes dt, \pi_\phi)$ . Then, with our previous notations, the distribution of the process  $\Gamma_\phi$  under  $\pi_\phi$  is the distribution of  $N$  under  $\pi$ . This means that for any bounded measurable  $f : \mathfrak{M} \rightarrow \mathbb{R}$ , for any  $t \in \mathbb{R}^+$ ,*

$$\mathbf{E}_\pi [f(\Gamma_\phi^t) \Lambda_\phi(t)] = \mathbf{E}_\pi [f(N^t)], \quad (4.17)$$

where  $X^t$  is the process  $X$  stopped at time  $t$ .

*Proof of Theorem 4.4.9.* The Radon-Nikodym derivative is given by Theorem 4.4.1.

$$\begin{aligned} \Lambda_\phi(t) &= \frac{d\pi_\phi}{d\pi} \Big|_{\mathcal{F}_t^N} \\ &= \exp \left( \int_0^t \int \log(\phi(s, z)) dN(s, z) - \int_0^t \int (\phi(s, z) - 1) \nu(dz) ds \right). \end{aligned} \quad (4.18)$$

We recall that for each  $A \in \mathcal{B}(\mathbb{M})$ , the compensator of  $\Gamma_\phi(\cdot, A)$  under  $\pi$  is  $\int_0^\cdot \mathbf{1}_A(\gamma_\phi(s, z)) \nu(dz)$ . Thus, by Lemma 4.4.6 and Remark 4.4.4,

$$\begin{aligned} R(t) &= \int_0^t \int \mathbf{1}_A(\gamma_\phi(s, z)) \phi(s, z) (N(ds, dz) - \nu(dz) ds) \\ &= \int_0^t \int \mathbf{1}_A(\gamma_\phi(s, z)) \phi(s, z) N(ds, dz) - t \int \mathbf{1}_A(z) \nu(dz) \end{aligned}$$

is a  $(\mathcal{F}^N, \pi)$  local martingale. The standard Girsanov's theorem for local martingales Schuppen and Wong (1974, Theorem 3.2.) says that

$$R(t) - \int_0^t \frac{1}{\Lambda_\phi(s)} d[R, \Lambda_\phi](s)$$

is a  $(\mathcal{F}^N, \pi_\phi)$ -local martingale. Note that  $R$  and  $\Lambda_\phi$  have the same jump times as  $\Gamma_\phi$ , hence

$$\begin{aligned} \int_0^t \frac{1}{\Lambda_\phi(s)} d[R, \Lambda_\phi](s) &= \sum_{s \leq t, \Delta N(s) \neq 0} \frac{1}{\Lambda_\phi(s)} \Delta R(s) \Delta \Lambda_\phi(s) \\ &= \sum_{s \leq t, \Delta N(s) \neq 0} \left( 1 - \frac{\Lambda_\phi(s^-)}{\Lambda_\phi(s)} \right) \Delta R(s) \\ &= \sum_{T_n \leq t, Z_n \neq 0} \left( 1 - \frac{1}{\phi(T_n, Z_n)} \right) \mathbf{1}_A(\gamma_\phi(T_n, Z_n)) \phi(T_n, Z_n) \cdot \\ &= \int_0^t \int \mathbf{1}_A(\gamma_\phi(s, z)) \phi(s, z) dN(s, z) - \Gamma_\phi(t, A). \end{aligned}$$

Thus, we have

$$R(t) - \int_0^t \frac{1}{\Lambda_\phi(s)} d[R, \Lambda_\phi](s) = \Gamma_\phi(t, A) - t \int \mathbf{1}_A(z) \nu(dz).$$

This means that the random measure associated to  $\Gamma_\phi$  has  $(\mathcal{F}^N, \pi_\phi)$ -dual predictable projection  $\nu(dz) \otimes ds$ . According to [Theorem 4.2.2](#) by the characterization of homogeneous Poisson measure, the random measure is a  $\mathcal{F}^N$ -adapted homogeneous Poisson measure of control measure  $\nu(dz) ds$  under  $\pi_\phi$ .  $\square$

We use subsequently the version below of the quasi-invariance theorem.

**Theorem 4.4.10 — Quasi-invariance.** *Let  $\pi_\phi \ll_{loc, \mathcal{F}^\Gamma} \pi$ . Then the distribution of  $\Gamma_\phi$  under  $\pi_\phi$  is the distribution of  $N$  under  $\pi$ , i.e.:*

$$\mathbf{E}_\pi [f(\Gamma_\phi^t) \Lambda_{\phi^\dagger}^\dagger(t)] = \mathbf{E}_\pi [f(N^t)], \quad (4.19)$$

with

$$\Lambda_{\phi^\dagger}^\dagger(t) = \frac{d\pi_\phi}{d\pi} \Big|_{\mathcal{F}_t^\Gamma} \quad (4.20)$$

$$= \exp \left( - \int_0^t \int \log(\phi^\dagger(s, u)) d\Gamma_\phi(s, u) + \int_0^t \int (\phi^\dagger(s, u) - 1) \nu(du) ds \right). \quad (4.21)$$

*Proof of Theorem 4.4.10.* The Girsanov theorem yields for any bounded measurable  $f : \mathfrak{M} \rightarrow \mathbb{R}$ , for any  $t \in \mathbb{R}^+$ :

$$\mathbf{E}_{\pi_\phi} [f(\Gamma_\phi) \Lambda_\phi(t)] = \mathbf{E}_\pi [f].$$

We proceed with the change of variable  $z = \gamma_\phi(s, u)$  over  $\mathbb{R}^+$ . By [Lemma 4.4.6](#),

$$\begin{aligned} \log \Lambda_\phi(t) &= \int_0^t \int \log(\phi(s, \gamma_\phi^*(s, u))) d\Gamma_\phi(s, u) - \int_0^t \int (\phi(s, \gamma_\phi^*(s, u)) - 1) \phi^\dagger(s, u) \nu(du) ds \\ &= - \int_0^t \int \log(\phi^\dagger(s, u)) d\Gamma_\phi(s, u) + \int_0^t \int (\phi^\dagger(s, u) - 1) \nu(du) ds. \end{aligned}$$

By [Corollary 4.4.7](#),  $\Lambda_{\phi^\dagger}^\dagger(t)$  is  $\mathcal{F}_t^{\Gamma_\phi}$ -measurable. The proof is thus complete.  $\square$

We now prove that well-behaved mark changes induce locally absolutely continuous probability on  $\mathfrak{M}$ .

**Theorem 4.4.11** *Let  $\phi$  belong to  $\mathcal{P}^{++}(\mathcal{F}^N)$  such that  $\pi_\phi$  is equivalent to  $\pi$  on  $\mathcal{F}_\infty$ . Then  $\Gamma_{\phi \# \pi}$  is equivalent to  $\pi$  on  $\mathcal{F}_\infty$ .*

*Proof of Theorem 4.4.11.* According to [Lemma 4.4.12](#), the martingale

$$(\Lambda_\phi(t), t \geq 0)$$

is uniformly integrable, and we can let  $t$  go to infinity in [\(4.17\)](#) to obtain

$$\Gamma_{\phi \# \pi_\phi} = \pi.$$

Since  $\pi_\phi$  is equivalent to  $\pi$  on  $\mathcal{F}_\infty$ ,  $\mathbf{\Gamma}_{\phi\#}\pi_\phi$  is equivalent to  $\mathbf{\Gamma}_{\phi\#}\pi$ . As a consequence,

$$\pi = \mathbf{\Gamma}_{\phi\#}\pi_\phi \sim \mathbf{\Gamma}_{\phi\#}\pi.$$

The proof is thus complete.  $\square$

We conclude this section by introducing a sufficient condition such that the equivalence between probability measures holds.

**Lemma 4.4.12** *Assume that  $\mu'$  is absolutely continuous with respect to  $\mu$  on  $\mathcal{F}_\infty^N$  and set*

$$\Lambda(t) = \left. \frac{d\mu'}{d\mu} \right|_{\mathcal{F}_t^N}. \quad (4.22)$$

*Then  $\mu'$  and  $\mu$  are equivalent if and only the following two conditions are satisfied:*

- i) The local martingale  $(\Lambda(t), t \geq 0)$  is uniformly integrable, i.e. there exists  $\Lambda \in L^1(\mu)$  such that*

$$\Lambda(t) = \mathbf{E}_\mu [\Lambda | \mathcal{F}_t^N].$$

- ii) The random variable  $\Lambda$  is positive  $\mu$ -a.s.*

In view of this theorem, we introduce the following sets of processes.

**Definition 4.4.13.** The most restricted class we consider is  $\mathbb{F}_\infty(\mathcal{P}^{++}(\mathcal{F}), \pi_\sigma)$  of processes  $\varphi \in \mathcal{P}^{++}(\mathcal{F})$  for which there exist  $c \in (0, 1)$  such that

$$c \leq \varphi(s, z) \leq \frac{1}{c}, \quad \forall s \geq 0, z \in \mathbb{M} \quad \pi_\sigma - \text{a.s.}$$

**Remark 4.4.14.** Note that if  $\phi$  belongs to  $\mathbb{F}_\infty(\mathcal{P}^{++}(\mathcal{F}^N), \pi_\sigma)$  then there exists  $C < +\infty$  such that  $\pi - \text{a.s.}, 0 < \Lambda_\phi(N, t) \leq C$ . Hence,  $(\Lambda_\phi(N, t), t \geq 0)$  is uniformly integrable.  $\blacksquare$

# Chapter 5

## Invertibility framework

Based on the extension of the invertibility framework introduced by Üstünel (Üstünel, 2009) and recently in the Poisson space by Coutin and Decreusefond (2023), built in Section 5.2, we find an entropy criterion under which a random change of mark of marked Poisson point process is invertible using the Girsanov theorem. As a consequence, we devise a new proof of the variational representation of the entropy on the extended Poisson space in Section 5.3. Finally, we provide with a construction of solutions of stochastic differential equations driven by Poisson measures and open problems on that topic relative to the invertibility framework.

### 5.1 Overview

This chapter is devoted to the search of conditions of invertibility of some class of mappings on the Poisson space. Sufficient conditions of invertibility were identified on the Wiener space in Üstünel and Zakai (2007). Combining all their earlier results, Üstünel and Zakai extend the entropy criterion due to Föllmer (see entropy criterion due to Föllmer and Wakolbinger (1986, Lemma 3.6) and Föllmer (1986, Proposition 2.11)) which ensures the invertibility of a shift transformation on the Wiener space: If the kinetic energy of the drift  $u$  is equal to the entropy of the measure induced by the corresponding shift transformation, then the map  $U$  is invertible. Such a result can be interpreted as a construction of a strong solution of the SDE (4.1) for a general  $u$ .

Recently, using a similar idea, the paper Coutin and Decreusefond (2023) shows that solving SDEs with constant volatility on the Wiener space is the analog of constructing Hawkes processes, i.e. self excited point process, on the Poisson space. If  $N$  is a Poisson process and  $\mathbf{Y}$  an invertible transformation such that  $\mathbf{Y}(N)$  is the process changed of time according to an increasing process  $y$ , then we have the inverse is a Hawkes process (Hawkes, 1971). That new point of view of Hawkes processes, including non-linear processes depending on the process  $y$  at hand, allows to consider the Hawkes problem as an SDE driven by a Poisson process. The key is the quasi-invariance property from the Girsanov theorem. In the Wiener case, random non-adapted transformations of Brownian motion have been considered by several authors in the context of Malliavin calculus, cf. Üstünel and Zakai (2000) and references therein. On the Poisson space without marks, the consequences of the quasi-invariance property have been determined with respect to anticipative transformations (Privault, 1996) and in the general case of metric spaces (Albeverio and Smorodina, 2006). In this chapter, we consider adapted transformations of Poisson measures as presented in the previous one.



## 5.2 Stochastic invertibility on the Poisson space with marks

Those definitions are analog to the ones of Üstünel (2014) relative to the Wiener space. We add the absolute continuity condition as we consider Girsanov transforms as absolute continuous transformations.

**Definition 5.2.1.** For a probability  $\mu$  on  $(\mathfrak{M}, \mathcal{F}_\infty)$ , a map  $\mathfrak{V} : \mathfrak{M} \rightarrow \mathfrak{M}$  is  $\mu$ -left invertible if and only if  $\mathfrak{V}_\# \mu \ll \mu$  along  $\mathcal{F}_\infty$  and there exists  $\mathfrak{Z} : \mathfrak{M} \rightarrow \mathfrak{M}$  such that  $\mathfrak{Z} \circ \mathfrak{V} = \text{Id}_{\mathfrak{M}}$ ,  $\mu$ -a.s.

The map  $\mathfrak{V} : \mathfrak{M} \rightarrow \mathfrak{M}$  is  $\mu$ -right invertible if and only if there exists  $\mathfrak{Z} : \mathfrak{M} \rightarrow \mathfrak{M}$  such that  $\mathfrak{Z}_\# \mu \ll \mu$  along  $\mathcal{F}_\infty$  and  $\mathfrak{V} \circ \mathfrak{Z} = \text{Id}_{\mathfrak{M}}$ ,  $\mu$ -a.s.

The map  $\mathfrak{V}$  is  $\mu$ -invertible if it is both  $\mu$ -left and  $\mu$ -right invertible.

For notational convenience, as the invertibility of sample paths remains with respect to  $\pi$ , we drop the dependency in it.

**Lemma 5.2.2** *If there exists  $\mathfrak{Z}$  such that  $\mathfrak{Z} \circ \mathfrak{V} = \text{Id}_{\mathfrak{M}}$ ,  $\pi$ -a.s. then  $\mathfrak{V} \circ \mathfrak{Z} = \text{Id}_{\mathfrak{M}}$ ,  $\mathfrak{V}_\# \pi$ -a.s. If additionally,  $\mathfrak{V}_\# \pi$  is equivalent to  $\pi$  and  $\mathfrak{Z}_\# \pi \ll \pi$ , then  $\mathfrak{V}$  is invertible and  $\mathfrak{Z}_\# \pi$  is equivalent to  $\pi$ .*

*Proof of Lemma 5.2.2.* We have

$$\begin{aligned} \mathfrak{V}_\# \pi \left( \mathfrak{V} \circ \mathfrak{Z} = \text{Id}_{\mathfrak{M}} \right) &= \pi \left( \mathfrak{V} \circ \mathfrak{Z} \circ \mathfrak{V} = \mathfrak{V} \right) \\ &= \pi \left( \mathfrak{V} = \mathfrak{V} \right) \\ &= 1. \end{aligned}$$

The first assertion follows.

If the two measures  $\mathfrak{V}_\# \pi$  and  $\pi$  are equivalent, then  $\mathfrak{V} \circ \mathfrak{Z} = \text{Id}_{\mathfrak{M}}$   $\pi$ -almost-surely thus  $\mathfrak{V}$  is invertible. Let  $A$  such that  $\mathfrak{Z}_\# \pi(A) = 0$ . This means

$$\mathbf{E}_\pi [\mathbf{1}_A \circ \mathfrak{Z}] = 0.$$

Since  $\mathfrak{V}_\# \pi$  is equivalent to  $\pi$ , we get

$$0 = \mathbf{E}_\pi [\mathbf{1}_A \circ \mathfrak{Z} \circ \mathfrak{V}] = \mathbf{E}_\pi [\mathbf{1}_A]$$

hence  $\pi \ll \mathfrak{Z}_\# \pi$  and the equivalence follows. □

We begin by a technical lemma which states the composition rules by the change of marks  $\Gamma_\phi$ .

**Lemma 5.2.3** For  $\psi \in \mathcal{P}^{++}(\mathcal{F}^N, \pi)$ , we define the Girsanov transform  $\gamma_\psi$  according to Definition 4.4.3. Then, we have

$$(\mathbf{\Gamma}_\psi \circ \mathbf{\Gamma}_\phi(N))(t, A) = \int \int \mathbf{1}_{[0,t]}(s) \mathbf{1}_A(\gamma_\psi(\mathbf{\Gamma}_\phi, s, \gamma_\phi(N, s, z))) \, dN(s, z). \quad (5.1)$$

Moreover, for  $f \in \mathbb{F}^1(\mathcal{P}(\mathcal{F}), \nu \otimes dt)$ ,

$$\left( \int_0^\cdot \int f(N, s, z) \nu(dz) \, ds \right) \circ \mathbf{\Gamma}_\phi = \int_0^\cdot \int f(\mathbf{\Gamma}_\phi, s, z) \nu(dz) \, ds \quad (5.2)$$

and

$$\left( \int_0^\cdot \int f(N, s, z) \, dN(s, z) \right) \circ \mathbf{\Gamma}_\phi = \int_0^\cdot \int f(\mathbf{\Gamma}_\phi, s, z) \, d\mathbf{\Gamma}_\phi(s, z). \quad (5.3)$$

*Proof of Lemma 5.2.3.* For  $t \geq 0$  and  $A \in \mathcal{B}(\mathbb{M})$ , we have by definition:

$$(\mathbf{\Gamma}_\psi \circ \mathbf{\Gamma}_\phi)(X)(t, A) = \int \int \mathbf{1}_{[0,t]}(s) \mathbf{1}_A(\gamma_\psi(\mathbf{\Gamma}_\phi(X), s, z)) \kappa_\gamma(X)(ds, dz) \quad (5.4)$$

$$= \int \int \mathbf{1}_{[0,t]}(s) \mathbf{1}_A(\gamma_\psi(\mathbf{\Gamma}_\phi(X), s, \gamma_\phi(X, s, z))) \, dX(s, z). \quad (5.5)$$

Because of the density of simple processes in  $\mathbb{F}^1(\mathcal{F}^N, \nu \otimes ds)$ , we can apply the composition rule (5.4) to processes

$$(t, A) \mapsto \int_0^t \int \mathbf{1}_A(z) f(N, s, z) \, dN^p(s, z)$$

and

$$(t, A) \mapsto \int_0^t \int \mathbf{1}_A(z) f(N, s, z) \, dN(s, z)$$

evaluated at  $(t, \mathbb{M})$  as to obtain (5.2) and (5.3).  $\square$

**Corollary 5.2.4** For  $\psi \in \mathcal{P}^{++}(\mathcal{F}^N, \pi)$ ,  $\mathbf{\Gamma}_\psi$  is the  $\pi$ -left inverse of  $\mathbf{\Gamma}_\phi$  if and only if  $\mathbf{\Gamma}_{\phi\#}\pi \ll \pi$  and:

$$\gamma_\psi(\mathbf{\Gamma}_\psi, s, \gamma_\phi(N, s, z)) = z \, \forall (t, z) \in \mathbb{R}^+ \times \mathbb{M} \, \pi\text{-a.s.} \quad (5.6)$$

or equivalently either of those equations:

•

$$\gamma_\psi^*(\mathbf{\Gamma}_\phi, s, z) = \gamma_\phi(N, s, z) \, \forall (t, z) \in \mathbb{R}^+ \times \mathbb{M} \, \pi\text{-a.s.}; \quad (5.7)$$

•

$$\psi(\mathbf{\Gamma}_\phi, s, \gamma_\phi(N, s, z)) \times \phi(N, s, z) = 1 \, \forall (t, z) \in \mathbb{R}^+ \times \mathbb{M} \, \pi\text{-a.s.} \quad (5.8)$$

*Proof of Corollary 5.2.4.* We have:

$$\mathbf{\Gamma}_\psi \circ \mathbf{\Gamma}_\phi = \text{Id}_{\mathbb{M}} \, \pi\text{-a.s.} \iff \gamma_\psi(\mathbf{\Gamma}_\phi, s, \gamma_\phi(N, s, z)) = z \, \pi \otimes ds \otimes dz\text{-a.s.} \quad (5.9)$$

Since  $\gamma_\psi$  is  $\pi$ -invertible, it is equivalent to

$$\gamma_\psi^*(\Gamma_\phi, s, z) = \gamma_\phi(N, s, z) \pi \otimes ds \otimes dz\text{-a.s..}$$

Let us show that it is also equivalent to (5.8). By differentiation of (5.6), we obtain

$$(\gamma_\psi(\Gamma_\phi, s, \cdot))'(\gamma_\phi(N, s, z)) \times (\gamma_\phi(N, s, \cdot))'(z) = 1, \pi - \text{a.s.}$$

Using Remark 4.4.4, we obtain

$$\begin{aligned} (\gamma_\psi(\Gamma_\phi, s, \cdot))'(\gamma_\phi(N, s, z)) \times (\gamma_\phi^*(N, s, \cdot))'(z) \\ &= \frac{\rho(\gamma_\phi(N, s, z))\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z))}{\rho(\gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z)))} \times \frac{\rho(z)\phi(N, s, z)}{\rho(\gamma_\phi(N, s, z))} \\ &= \psi(\Gamma_\phi, s, \gamma_\phi(N, s, z))\phi(N, s, z) = 1. \end{aligned}$$

Conversely, assume there exists  $\psi \in \mathcal{P}^{++}(\mathcal{F}^N)$  such that

$$\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z))\phi(N, s, z) = 1 \quad \forall (t, z) \in \mathbb{R}^+ \times \mathbb{M} \text{ } \pi\text{-a.s..}$$

Given  $\gamma_\psi$  defined according to Definition 4.4.3, it yields:

$$\begin{aligned} (\gamma_\psi(\Gamma_\phi, s, \cdot))'(\gamma_\phi(N, s, z)) \times (\gamma_\phi(N, s, \cdot))'(z) &= \frac{\rho(\gamma_\phi(N, s, z))\rho(z)}{\rho(\gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z)))\rho(\gamma_\phi(N, s, z))} \\ &= \frac{\rho(z)}{\rho(\gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z)))}. \end{aligned}$$

Hence,

$$\begin{aligned} \gamma_\psi(\Gamma_\phi, s, \cdot)'(\gamma_\phi(N, s, \cdot)) \times \rho(\gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, z))) &= \rho(z) \\ (r \circ \gamma_\psi(\Gamma_\phi(N), s, \gamma_\phi^*(N, s, \cdot)))' &= r'. \end{aligned}$$

By integration,

$$r \circ \gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, \cdot)) = r.$$

Because  $r$  is a bijection from  $\mathbb{M}$  onto itself, it follows that

$$\gamma_\psi(\Gamma_\phi, s, \gamma_\phi(N, s, \cdot)) = \text{Id}_{\mathbb{M}}.$$

□

By definition,  $\mathcal{F}^N$  is smaller than  $\mathcal{F}^\Gamma$ . If we assume left invertibility of the map  $\Gamma$ , we have a stronger result.

**Theorem 5.2.5** *Let  $\phi \in \mathcal{P}^{++}(\mathcal{F}^N)$ . Then,  $\mathcal{F}^{\Gamma\phi} = \mathcal{F}^N$  if and only if  $\Gamma_\phi$  admits a left inverse.*

*Proof of Theorem 5.2.5.* If  $\Gamma_\phi$  is left invertible then there exists  $\gamma_\psi$  such that (5.7) holds:

$$\gamma_\psi(\Gamma_\phi(\omega), s, z) = \gamma_\phi^*(\omega, s, z) \pi \otimes ds \otimes \nu\text{-a.s..}$$

This means that:

$$\gamma_\phi^*(N, t, \cdot) \text{ is } \mathcal{F}_t^{\Gamma\phi}\text{-measurable.}$$

Hence by [Lemma 4.3.3](#),  $\mathcal{F}^{\Gamma_\phi} = \mathcal{F}^N$ .

Conversely, if, for all  $t \in \mathbb{R}^+$ ,  $\mathcal{F}_t^N = \mathcal{F}_t^{\Gamma_\phi}$ , then for all  $t \in \mathbb{R}^+$  there exists a sequence of predictable processes  $(\bar{\gamma}_\psi^t)_{t \in \mathbb{R}^+}$  such that:

$$\bar{\gamma}_\psi^t(\Gamma_\phi, z) = \gamma_\phi^*(N, t, z), \quad \pi \otimes dz - \text{a.s.}$$

Denoting  $\bar{\gamma}_\psi(\mathbf{\Gamma}_\phi(\omega), t, z) = \bar{\gamma}_\psi^t(\Gamma_\phi, z)$  for all  $t \in \mathbb{R}^+$ , there exists a full probability set  $B$  such that

$$\bar{\gamma}_\psi(\mathbf{\Gamma}_\phi(\omega), t, z) = \gamma_\phi^*(\omega, t, z), \quad \forall t \in \mathbb{Q}, \quad \forall z \in \mathbb{Q}, \quad \forall \omega \in B. \quad (5.10)$$

Let

$$\gamma_\psi(\Gamma_\phi, t, z) = \begin{cases} \gamma_\phi^*(N, t, z) & \text{if } t \in \mathbb{Q} \\ \lim_{r_n \rightarrow t, r_n \in \mathbb{Q} \cap [0, t]} \gamma_\phi^*(N, t, z) & \text{if } t \notin \mathbb{Q}. \end{cases} \quad \pi \otimes dz - \text{a.s.}$$

By the sample-path left-continuity of  $\gamma_\phi^*$  in the time variable, (5.10) holds for any  $t \in \mathbb{R}^+$  with probability 1. Hence, [Corollary 5.2.4](#) implies that  $\mathbf{\Gamma}_\phi$  admits a left inverse.  $\square$

The construction of an invertible mapping is straightforward in a specific case. We recall the notation  $X^t$  which stands for the process  $X$  stopped at time  $t$  for  $t \in \mathbb{R}^+$ .

**Lemma 5.2.6** *Let  $\phi \in \mathbb{F}_\infty(\mathcal{P}^{++}(\mathcal{F}^N), \pi)$  be defined by time cases, viz. consider a partition of  $\mathbb{R}^+$ ,  $0 = t_0 < t_1 < \dots < t_k < t_{k+1} = +\infty$  and assume that there exist positive functions  $g_j : \mathfrak{M} \times \mathbb{M} \rightarrow \mathbb{M}$  such that*

$$\phi(N, s, z) = \sum_{j=0}^k \mathbf{1}_{(t_j, t_{j+1}]}(s) g_j(N^{t_j}, z), \quad g_j : \mathfrak{M} \times \mathbb{M} \rightarrow \mathbb{M}.$$

*Then,  $\mathbf{\Gamma}_\phi$  is invertible.*

*Proof of Lemma 5.2.6.* Let  $\phi \in \mathcal{P}^{++}(\mathcal{F}^N)$  be defined by time cases. We recall that it is associated to the triplet  $(\gamma_\phi, \Gamma_\phi, \phi^\dagger)$  with  $\gamma_\phi$   $\pi$ -invertible and defined by time cases. Hence, there exist positive functions  $\tilde{g}_j : \mathfrak{M} \times \mathbb{M} \rightarrow \mathbb{M}$  such that:

$$\max(\|\tilde{g}_j\|_\infty, \|\frac{1}{\tilde{g}_j}\|_\infty) < +\infty$$

and

$$\phi^\dagger(N, s, z) = \sum_{j=0}^k \mathbf{1}_{(t_j, t_{j+1}]}(s) \tilde{g}_j(N^{t_j}, z).$$

Let us construct  $\hat{g}_j$  such that for  $z \in \mathbb{M}$ :

$$\hat{g}_0(N^{t_0}, u) = \frac{1}{g_0(\mathbf{\Gamma}_\phi^{t_0}, z)},$$

and for  $j = 1, \dots, k-1$

$$\hat{g}_j(N, u) = \frac{1}{\tilde{g}_j(\mathbf{\Gamma}_\phi^{t_{j-1}}, z)}.$$

Then,

$$\psi(N, s, z) := \sum_{j=0}^k \mathbf{1}_{(t_j, t_{j+1}]}(s) \hat{g}_j(N, z) = \frac{1}{\phi(\Gamma_\phi, s, \gamma_\phi^*(s, z))}$$

is defined by time cases in  $\mathcal{P}^{++}(\mathcal{F}^N)$ . It follows by [Corollary 5.2.4](#) that  $\Gamma_\psi \circ \Gamma_\phi = \text{Id}_{\mathfrak{M}}$   $\pi$ -a.s.. As in [Remark 4.4.14](#),  $\pi_\phi$  and  $\pi$  are equivalent which entails, by [Theorem 4.4.11](#), that  $\Gamma_{\phi\#}\pi$  is equivalent to  $\pi$  on  $\mathcal{F}_\infty$ , and its left invertibility follows. With the same argument as for  $\Gamma_\phi$ ,  $\Gamma_{\psi\#}$  is equivalent to  $\pi$ . By [Lemma 5.2.2](#),  $\Gamma_\phi \circ \Gamma_\psi = \text{Id}_{\mathfrak{M}}$   $\Gamma_{\psi\#}\pi$ -a.s., hence  $\pi$ -a.s.. Thus,  $\Gamma_\phi$  is also right invertible, which concludes the proof.  $\square$

As a consequence of the new framework, we devise a new proof of the variational representation of the entropy on the extended Poisson space. Finally, we establish a new criterion for solutions of stochastic differential equations driven by Poisson measures.

### 5.3 Entropy applications

The left invertibility has important consequences on the relative entropy of  $\Gamma_{\phi\#}\pi$  with respect to the probability measure of reference.

**Definition 5.3.1.** For  $\mu$  and  $\mu'$  two probability measures on  $(\mathfrak{M}, \mathcal{F}_\infty)$ , the relative entropy of  $\mu'$  with respect to  $\mu$  is given by

$$H(\mu' | \mu) = \begin{cases} \mathbf{E}_{\mu'} \left[ \log \left( \frac{d\mu'}{d\mu} \Big|_{\mathcal{F}_\infty} \right) \right] & \text{if } \mu' \ll \mu, \\ +\infty & \text{otherwise.} \end{cases}$$

Consider  $\mathfrak{m}$ , the smooth, convex, non-negative function defined on  $[-1, \infty)$  by

$$\mathfrak{m}(x) = \begin{cases} (x+1) \log(x+1) - x & \text{if } x > -1, \\ 1 & \text{if } x = -1, \end{cases}$$

and  $L_{\mathfrak{m}}$ , the corresponding Orlicz space ([Adams and Fournier, 2003](#)):

$$L_{\mathfrak{m}} = \left\{ f : \mathbb{R}^+ \rightarrow [-1, \infty), \int_0^\infty \mathfrak{m}(f(s)) \, ds < \infty \right\}.$$

Since  $\mathfrak{m}(2x) \leq 4\mathfrak{m}(x)$ ,  $L_{\mathfrak{m}}$  is a separable Banach space when equipped with the Luxemburg norm (see [Adams and Fournier, 2003](#)).

In the same vein of [Lassalle \(2012, Proposition 2.1\)](#), we obtain the next lemma.

**Lemma 5.3.2** *Let  $\psi \in \mathcal{P}^{++}(\mathcal{F}^N)$ , the following assertions are equivalent:*

- (i) *There exists  $\phi \in \mathbb{F}_2^1(\mathcal{P}^{++}(\mathcal{F}^N), \nu \otimes ds, \pi_\psi)$  such that  $\Gamma_\phi$  is left invertible.*
- (ii)  *$\Gamma_{\phi\#}\pi = \pi_\psi$  on  $\mathcal{F}_\infty^N$  and we have the following identity:*

$$\log(\Lambda_\psi \circ \Gamma_\phi) = -\log \Lambda_\phi^\dagger \pi - a.s.. \quad (5.11)$$

*Proof of Lemma 5.3.2.* The left invertibility of  $\Gamma_\phi$  entails that  $\Gamma_\phi \circ \Gamma_\psi = \text{Id}_{\mathfrak{M}}$   $\Gamma_{\phi\#\pi}$ -a.s. according to Lemma 5.2.2. Recall that the Girsanov theorem says that  $\Gamma_{\psi\#\pi_\psi} = \pi$  hence  $\Gamma_{\phi\#\pi} = \pi_\psi$ . According to Lemma 5.2.3, consider  $\Lambda_\psi$  as a process indexed on  $\mathbb{R}^+ \times \mathbb{M}$  evaluated at  $(t, \mathbb{M})$ . For  $t \geq 0$ , by linearity of  $\Gamma_\phi$  and Lemma 5.2.3,

$$\begin{aligned} & \log \Lambda_\psi \circ \Gamma_\phi(N)(t, \mathbb{M}) \\ &= \int_0^t \int \log(\psi(\Gamma_\phi, s, z)) \, d\Gamma_\phi(s, z) - \int_0^t \int (\psi(\Gamma, s, z) - 1) \nu(dz) \, ds \\ &= \int_0^t \int \log\left(\frac{1}{\phi(N, s, \gamma_\phi^*(N, s, z))}\right) \, d\Gamma_\phi(s, z) - \int_0^t \int \left(\frac{1}{\phi(N, s, \gamma_\phi^*(N, s, z))} - 1\right) \nu(dz) \, ds, \\ &= \int_0^t \int \log(\phi^\dagger(N, s, z)) \, d\Gamma_\phi(s, z) - \int_0^t \int (\phi^\dagger(N, s, z) - 1) \nu(dz) \, ds, \end{aligned}$$

using Corollary 5.2.4 and invertibility of  $\gamma_\phi^*$ . Hence, (5.11) follows.

For the converse, by (5.11), taking the conditional expectation with respect to  $\mathcal{F}_t^N$  in both sides of the equality:

$$\begin{aligned} & \int_0^t \int \left( \log(\psi(\Gamma, s, z)) + \log(\phi(N, s, \gamma_\phi^*(s, z))) \right) \, dN(s, z) \\ &= \int_0^t \int (\psi(\Gamma, s, z) + \phi(N, s, \gamma_\phi^*(s, z))) \nu(dz) \, ds. \end{aligned}$$

Equating the jumps yields:

$$\psi(\Gamma, s, z) = \phi(N, s, \gamma_\phi^*(s, z)) \quad \forall (s, z) \in [0, t] \times \mathbb{M} \quad \pi\text{-a.s.}$$

By (5.6), it is equivalent to  $\Gamma_\psi \circ \Gamma_\phi = \text{Id}_{\mathfrak{M}}$   $\pi$ -a.s..

It remains to show that  $\Gamma_{\phi\#\pi} \ll \pi$  along  $\mathcal{F}^N$ . Eqn (5.11) amounts to:

$$\Lambda_\phi^\dagger = \frac{1}{\Lambda_\psi} \circ \Gamma_\phi.$$

Since  $\Gamma_{\phi\#\pi} = \pi_\psi$ , we have:

$$\begin{aligned} \mathbf{E}_\pi \left[ \Lambda_\phi^\dagger \right] &= \mathbf{E}_{\pi_\psi} \left[ \frac{1}{\Lambda_\psi} \right] \\ &= \mathbf{E}_\pi \left[ \frac{1}{\Lambda_\psi} \Lambda_\psi \right] \text{ by the definition of the Radon-Nikodym density } \Lambda_\psi \\ &= 1. \end{aligned}$$

Hence, by Lemma 4.4.12,  $\Gamma_{\phi\#\pi} = \pi'_\phi$  is equivalent to  $\pi$ .  $\square$

**Lemma 5.3.3** *Let  $\phi \in \mathbb{F}_2^1(\mathcal{P}(\mathcal{F}^N), \nu \otimes ds; \pi)$  such that  $\mathbf{E}_\pi[\Lambda_\phi] = 1$  and*

$$\mathbf{E}_\pi \left[ \int_0^\infty \int \mathbf{m}(\phi^\dagger(s, z) - 1) \nu(dz) \, ds \right] < \infty.$$

*Then,*

$$\mathbf{E}_\pi \left[ -\log \Lambda_\phi^\dagger \right] \leq \mathbf{E}_\pi \left[ \int_0^\infty \int \mathbf{m}(\phi^\dagger(s, z) - 1) \nu(dz) \, ds \right]. \quad (5.12)$$

*Proof of Lemma 5.3.3.* From Proposition 4.2.5, as  $t \mapsto \Gamma(\cdot, A)$  has compensator

$$t \mapsto \int_0^t \int \phi^\dagger(s, z) \nu(dz) ds,$$

we have:

$$\begin{aligned} \mathbf{E}_\pi \left[ -\log \Lambda_{\phi^\dagger}^\dagger(t) \right] &= \mathbf{E}_\pi \left[ \int_0^t \int \log(\phi^\dagger(s, z)) d\Gamma_\phi(s, z) - \int_0^t \int (\phi^\dagger(s, z)^{-1} - 1) \nu(dz) ds \right] \\ &= \mathbf{E}_\pi \left[ \int_0^t \int \left( \log(\phi^\dagger(s, z)) \phi^\dagger(s, z) - \phi^\dagger(s, z) + 1 \right) \nu(dz) ds \right] \\ &\leq \mathbf{E}_\pi \left[ \int_0^\infty \int \mathbf{m}(\phi^\dagger(s, z) - 1) \nu(dz) ds \right]. \end{aligned}$$

It remains to prove that we can pass to the limit in the left-hand-side. Consider the non-negative, convex function  $\psi(x) = x - \log x$ . From Fatou's Lemma, we have

$$\begin{aligned} \mathbf{E}_\pi \left[ \psi(\Lambda_{\phi^\dagger}^\dagger) \right] &\leq \liminf_{t \rightarrow \infty} \mathbf{E}_\pi \left[ \psi(\Lambda_{\phi^\dagger}^*(t)) \right] \\ &\leq 1 + \mathbb{E} \left[ \int_0^\infty \int \mathbf{m}(\phi^\dagger(s, z) - 1) \nu(dz) ds \right]. \end{aligned}$$

This means that the non-negative submartingale  $(\psi(\Lambda_{\phi^\dagger}^\dagger(t)), t \geq 0)$  is uniformly integrable. Thus,

$$-\log \Lambda_{\phi^\dagger}^\dagger(t) \xrightarrow[t \rightarrow \infty]{L^1} -\log \Lambda_{\phi^\dagger}^\dagger.$$

This means that

$$1 + \mathbf{E}_\pi \left[ -\log \Lambda_{\phi^\dagger}^\dagger \right] \leq 1 + \mathbf{E}_\pi \left[ \int_0^\infty \int \mathbf{m}(\phi^\dagger(s, z) - 1) \nu(dz) ds \right].$$

The proof is thus complete. □

**Theorem 5.3.4** *Let  $\phi \in \mathbb{F}_2^1(\mathcal{P}^{++}(\mathcal{F}^N), \nu(dz) \otimes ds)$  such that  $\mathbf{E}_\pi[\Lambda_\phi] = 1$  and*

$$\mathbf{E}_\pi \left[ \int_0^\infty \int \mathbf{m}(\phi^\dagger(s, z) - 1) \nu(dz) ds \right] < \infty.$$

*If  $\mathbf{\Gamma}_{\phi\#\pi} \ll \pi$ , we always have*

$$H(\mathbf{\Gamma}_{\phi\#\pi} | \pi) \leq \mathbf{E}_\pi \left[ \int_0^\infty \int \mathbf{m}(\phi^\dagger(s, z) - 1) \nu(dz) ds \right]. \quad (5.13)$$

*Moreover, the map  $\mathbf{\Gamma}_\phi$  is left invertible if and only if*

$$H(\mathbf{\Gamma}_{\phi\#\pi} | \pi) = \mathbf{E}_\pi \left[ \int_0^\infty \int \mathbf{m}(\phi^\dagger(s, z) - 1) \nu(dz) ds \right]. \quad (5.14)$$

*Proof of Theorem 5.3.4.* Assume that  $\mathbf{\Gamma}_{\phi\#\pi}$  is absolutely continuous with respect to  $\pi$  on  $\mathcal{F}_\infty$ . According to Theorem 4.4.1, there exists  $\psi \in \mathbb{F}_{2,loc}^1(\mathcal{P}^{++}(\mathcal{F}^N), \nu \otimes ds, \mathbf{\Gamma}_{\phi\#\pi})$  such that

$$\left. \frac{d\mathbf{\Gamma}_{\phi\#\pi}}{d\pi} \right|_{\mathcal{F}_t^N} = \Lambda_\psi(N, t).$$

hence for  $t \in \mathbb{R}^+$

$$\begin{aligned}\mathbf{E}_\pi [f \circ \mathbf{\Gamma}_\phi^t(N)] &= \mathbf{E}_\pi [f(\mathbf{\Gamma}^t)] \\ &= \mathbf{E}_\pi [f\Lambda_\psi(N, t)].\end{aligned}$$

By applying the quasi-invariance Theorem, for  $f : \mathfrak{M} \rightarrow \mathbb{R}$  bounded and continuous, we have that  $f\Lambda_\psi(\cdot, t)$  is bounded and continuous, then

$$\mathbf{E}_\pi [f \circ \mathbf{\Gamma}_\phi^t(N)] = \mathbf{E}_\pi \left[ f \circ \mathbf{\Gamma}_\phi^t(N) \Lambda_\psi \circ \mathbf{\Gamma}_\phi^t(N) \Lambda_{\phi^\dagger}^\dagger(N, t) \right].$$

Since  $|f \circ \mathbf{\Gamma}_\phi^t(N)| \leq |f \circ \mathbf{\Gamma}_\phi|$  which is  $L^1(\mathfrak{M} \rightarrow \mathbb{R}, \pi)$  then, for  $t \rightarrow \infty$ , we obtain

$$\mathbf{E}_\pi \left[ f \circ \mathbf{\Gamma}_\phi^t \Lambda_\psi \circ \mathbf{\Gamma}_\phi \Lambda_{\phi^\dagger}^\dagger \right] \leq \mathbf{E}_\pi [f \circ \mathbf{\Gamma}_\phi].$$

Hence,  $\pi$ -a.s. we have

$$\Lambda_\psi \circ \mathbf{\Gamma}_\phi \times \Lambda_{\phi^\dagger}^\dagger \leq 1. \quad (5.15)$$

It follows that

$$\begin{aligned}0 \leq H(\mathbf{\Gamma}_{\phi\#\pi} | \pi) &= \mathbf{E}_{\mathbf{\Gamma}_{\phi\#\pi}} [\log \Lambda_\psi] \\ &= \mathbf{E}_\pi [\log \Lambda_\psi \circ \mathbf{\Gamma}_\phi] \\ &\leq -\mathbf{E}_\pi \left[ \log \Lambda_{\phi^\dagger}^\dagger \right].\end{aligned} \quad (5.16)$$

Then the first part holds.

Assume now that (5.14) holds. Then (5.16) is an equality as a consequence of (5.15), which in turn is an equality. According to Lemma 5.3.2,  $\mathbf{\Gamma}_\phi$  is left invertible.

Conversely, if  $\mathbf{\Gamma}_\phi$  is left invertible. According to the Definition 5.2.1,  $\mathbf{\Gamma}_{\phi\#\pi}$  is absolutely continuous with respect to  $\pi$ . Let us denote by  $\mathfrak{J}$  the map such that

$$\mathfrak{J} \circ \mathbf{\Gamma}_\phi = \text{Id}_{\mathfrak{M}}, \quad \pi - \text{a.s.}$$

The Lemma 5.3.2 entails that:

$$\log \Lambda_\psi \circ \mathbf{\Gamma}_\phi = -\log \Lambda_\phi. \quad (5.17)$$

Let

$$R = \frac{d\mathbf{\Gamma}_{\phi\#\pi}}{d\pi}.$$

For any  $f : \mathfrak{M} \rightarrow \mathbb{R}$  continuous and bounded, for any  $t > 0$ , we have

$$\begin{aligned}\mathbf{E}_\pi [fR] &= \mathbf{E}_\pi [f \circ \mathbf{\Gamma}_\phi] \\ &= \mathbf{E}_\pi \left[ (f\Lambda_\psi) \circ \mathbf{\Gamma}_\phi \Lambda_{\phi^\dagger}^\dagger \right]\end{aligned}$$

according to (5.17) and by the quasi-invariance Theorem.

$$\mathbf{E}_\pi [fR] = \mathbf{E}_\pi [f\Lambda_\psi]$$



It follows that  $R = \Lambda_\psi$ ,  $\pi$  - a.s.. Plug this identity into (5.17) to obtain

$$\begin{aligned} H(\Gamma_{\phi\#\pi} | \pi) &= \mathbf{E}_\pi [\log R \circ \Gamma_\phi] \\ &= \mathbf{E}_\pi [\log \Lambda_\psi \circ \Gamma_\phi] \\ &= \mathbf{E}_\pi \left[ -\log \Lambda_{\phi^\dagger}^\dagger \right] \\ &= \mathbf{E}_\pi \left[ \int_0^\infty \mathfrak{m}(\phi^\dagger(s, z) - 1) \nu(dz) ds \right], \end{aligned}$$

using Lemma 5.3.3. The entropy criterion is thus satisfied.  $\square$

We reuse the previous notations of Girsanov factor  $\phi$  and associated triplet  $(\gamma_\phi, \Gamma_\phi, \phi^\dagger)$ . Let

$$\begin{aligned} \mathcal{P}_m^{++}(\mathcal{F}^N) &= \left\{ \phi \in \mathcal{P}^{++}(\mathcal{F}^N) \text{ and } (\phi^\dagger - 1) \in L^1(\mathfrak{M} \rightarrow \mathbb{R}, \pi) \cap L_m \right\} \\ \mathcal{P}_{\infty, \text{pc}}^{++}(\mathcal{F}^N) &= \mathcal{P}_{\infty}^{++}(\mathcal{F}^N, \pi) \cap \{ \phi \text{ piecewise constant} \} \\ \mathfrak{M}_m(\mathcal{F}^N) &= \{ \mu, \exists \phi \in \mathcal{P}_m^{++}(\mathcal{F}^N) \text{ such that } \mu = \Gamma_{\phi\#\pi} \}. \end{aligned}$$

It is well known that there is a Legendre duality between relative entropy and logarithmic Laplace transform (Lehec, 2013).

**Proposition 5.3.5** *Let  $f : \mathfrak{M} \rightarrow \mathbb{R}$  such that*

$$\mathbf{E}_\pi \left[ |f|(1 + e^f) \right] < \infty. \quad (\text{B1})$$

*Then,*

$$\log \mathbf{E}_\pi \left[ e^f \right] = \sup_{\mu \in \mathfrak{M}_{\ll \pi}} (\mathbf{E}_\mu[f] - H(\mu | \pi)) \quad (5.18)$$

*where  $\mathfrak{M}_{\ll \pi}$  is the set of probability measures on  $\mathfrak{M}$  which are absolutely continuous with respect to  $\pi$  on  $\mathcal{F}_\infty$ . Furthermore, the supremum is attained at the measure  $\mu_f$  whose  $\pi$ -density is given by*

$$\frac{d\mu_f}{d\pi} = \frac{e^f}{\mathbf{E}_\pi[e^f]}. \quad (5.19)$$

The theorem of representation of the entropy reads as follows:

**Theorem 5.3.6 — Variational representation of the entropy.** *Let  $f : \mathfrak{M} \rightarrow \mathbb{R}$  satisfying (B1).*

*Then,*

$$\log \mathbf{E}_\pi \left[ e^f \right] = \sup_{\phi \in \mathcal{P}_m^{++}(\mathcal{F}^N)} \left( \mathbf{E}_\pi [f(\Gamma_\phi)] - \mathbf{E}_\pi \left[ \int_0^\infty \int \mathfrak{m}(\phi^\dagger(s, z) - 1) \nu(dz) ds \right] \right)$$

*where  $(\gamma_\phi, \Gamma_\phi, \phi^\dagger)$  is the triplet given by Definition 4.4.3.*

*Proof of Theorem 5.3.6.* In view of (5.18), we evidently have

$$\log \mathbf{E}_\pi \left[ e^f \right] \geq \sup_{\mu \in \mathfrak{M}_m(\mathcal{F}^N)} (\mathbf{E}_\mu[f] - H(\mu | \pi)),$$

and

$$\begin{aligned}
\sup_{\mu \in \mathfrak{M}_{\ll \pi}} \mathbf{E}_{\mu} [f] - H(\mu | \pi) &\geq \sup_{\phi \in \mathcal{P}_{\mathfrak{m}}^{++}(\mathcal{F}^N)} \mathbf{E}_{\Gamma_{\phi \# \pi}} [f] - H(\mu | \pi). \\
&\geq \sup_{\phi \in \mathcal{P}_{\infty, \text{pc}}^{++}(\mathcal{F}^N)} \mathbf{E}_{\Gamma_{\phi \# \pi}} [f] - H(\Gamma_{\phi \# \pi} | \pi) \\
&= \sup_{\phi \in \mathcal{P}_{\infty, \text{pc}}^{++}(\mathcal{F}^N)} \mathbf{E}_{\Gamma_{\phi \# \pi}} [f] - \mathbf{E}_{\pi} \left[ \int_0^{\infty} \int \mathfrak{m}(\phi^{\dagger}(s, z) - 1) \nu(dz) ds \right]
\end{aligned}$$

by using [Theorem 5.3.4](#) as such  $\Gamma_{\phi}$  is left invertible in virtue of [Lemma 5.2.6](#). It remains to prove that we can find  $(\phi_n, n \geq 1)$ , a sequence of elements of  $\mathcal{P}_{\infty, \text{pc}}^{++}(\mathcal{F}^N, \pi)$  such that

$$\begin{aligned}
\mathbf{E}_{\Gamma_{\phi_n \# \pi}} [f] &\xrightarrow{n \rightarrow \infty} \mathbf{E}_{\mu_f} [f] \\
H(\Gamma_{\phi_n \# \pi} | \pi) &\xrightarrow{n \rightarrow \infty} H(\mu | \pi)
\end{aligned}$$

to conclude. The proof of those limits is analogous to the ones in the proofs of [Lemma 2.2](#) and [Theorem 4.11](#) of [Zhang \(2009\)](#) (see [Theorem 5.3.7](#)).  $\square$

Let denote  $L^r(\pi) = L^r(\mathfrak{M} \rightarrow \mathbb{R}; \pi)$  for  $r \geq 1$ .

**Theorem 5.3.7** *Let  $\pi'$  be a probability measure on  $\mathfrak{M}$  absolutely continuous with respect to  $\pi$  and*

$$L = \frac{d\pi'}{d\pi}.$$

*Assume that*

$$\begin{aligned}
L \log L &\in L^1(\pi) \\
\log L &\in L^r(\pi)
\end{aligned}$$

*for some  $r > 1$ . Then, there exists  $(\phi_n, n \geq 1)$  a sequence of elements of  $\mathcal{P}_{\infty, \text{pc}}^{++}(\mathcal{F}^N, \pi)$  such that*

$$L_n \log L_n \xrightarrow[n \rightarrow \infty]{L^1(\pi)} L \log L \tag{5.20}$$

*and*

$$L_n \log L \xrightarrow[n \rightarrow \infty]{L^1(\pi)} L \log L \tag{5.21}$$

*where*

$$L_n = \frac{d\Gamma_n \# \pi}{d\pi}.$$

*Proof.* We first show that we can suppose  $L$  lower and upper bounded. Consider

$$\Phi_n = (L \wedge n) \vee \frac{1}{n}.$$

We have

$$|\Phi_n| \leq L + 1,$$

hence by dominated convergence,  $\Phi_n$  converges in  $L^1(\pi)$  to  $L$  and in particular,

$$\mathbf{E}_{\pi} [\Phi_n] \xrightarrow{n \rightarrow \infty} \mathbf{E}_{\pi} [L] = 1.$$

Let

$$L_n = \frac{\Phi_n}{\mathbf{E}_\pi[\Phi_n]}.$$

For any  $\alpha \in (0, 1)$ , for  $n$  sufficiently large,  $\mathbf{E}_\pi[\Phi_n] \geq \alpha$ . Moreover, for  $x \geq 0$ , we have

$$|x \log(x)| \leq \frac{1}{e} + \left| \frac{x}{\alpha} \log\left(\frac{x}{\alpha}\right) \right| \mathbf{1}_{x \geq \alpha}.$$

Hence,

$$|L_n \log L_n| \leq \frac{1}{e} + \left| \frac{L}{\alpha} \log\left(\frac{L}{\alpha}\right) \right|.$$

By dominated convergence again,  $L_n \log L_n$  converges to  $L \log L$  in  $L^1(\pi)$ . Similarly,

$$|L_n \log L| \leq \left| \frac{L}{\alpha} \log L \right|$$

and  $L_n \log L$  converges to  $L \log L$  in  $L^1(\pi)$ .

Assume now that  $L$  is lower and upper bounded by respectively  $m$  and  $M$ . We know that there exists  $\phi \in \mathcal{P}^{++}(\mathcal{F}^N, \pi)$  such that

$$L = \Lambda_\phi.$$

Set

$$L_n = \Lambda_{1+(\phi^n-1)} = \mathbf{E}_\pi[L | \mathcal{F}_n^N].$$

Since  $L$  is bounded, it is clear that (5.20) and (5.21) hold. We can then assume that there exists  $T > 0$  such that  $\phi(s, z) = 1$  for  $s \geq T$  and  $z \in \mathbb{M}$ .

Moreover, from the Malliavin calculus for Poisson process, we know (see Decreusefond et al. (2010)) that

$$\phi(s, z) = \frac{\mathbf{E}_\pi[D_s L | \mathcal{F}_s^N]}{\mathbf{E}_\pi[L | \mathcal{F}_s^N]}$$

where  $D_{s,z}F(N) = F(N + \epsilon_{s,z}) - F(N)$ . We then have

$$0 \leq \phi(s, z) \leq \frac{M}{m}.$$

Consider  $\phi_n = \phi \vee n^{-1}$ , it is straightforward that (5.20) and (5.21) hold.

Finally, assume that  $\phi$  is lower and upper bounded on some interval  $[0, T]$  and equal to 1 above  $T$ . Set

$$\begin{aligned} \phi_n(s, z) &= 0 \text{ if } s \in [0, T/n) \\ \phi_n(s, z) &= n \int_{(i-1)/n}^{i/n} \phi(u, z) \, du \text{ if } s \in [iT/n, (i+1)T/n) \end{aligned}$$

for  $i \in \{1, \dots, n-1\}$ . We see that  $\phi_n$  belongs to  $\mathcal{P}_{\infty, \text{pc}}^{++}(\mathcal{F}^N, \pi)$ . We know (see Nualart (2006)) that  $\phi_n$  converges in  $L^2(\mathfrak{N} \times [0, T], \pi \otimes ds)$  to  $\dot{y}$ . Moreover, it is easy to see that

$$\sup_n \mathbf{E}_\pi \left[ \Lambda_{\phi_n}^p \right] < \infty \text{ for any } p \geq 1.$$

Thus, (5.20) and (5.21) hold.  $\square$

**Remark 5.3.8.** Zhang showed a similar result with a different point of view. In his work, the supremum is taken with respect of the Girsanov factor  $\phi$  instead of  $\phi^\dagger$ . Moreover,  $f$  is only bounded. We prove that the equality holds if and only if  $\mathbf{\Gamma}_\phi$  is left invertible whereas Zhang states a sufficient condition.  $\blacksquare$

## 5.4 Solutions of SDEs driven by Poisson measures

In this section, we specialize the result of right invertibility to characterizing solutions of SDEs under mild assumptions.

**Remark 5.4.1.** We can in fact start from a Girsanov factor to define  $\phi^\dagger$  as to define the associated Girsanov transform  $\gamma_\phi^*$ , and then  $\gamma_\phi$  as inverse. This point of view will be preferred in that section. ■

**Lemma 5.4.2** *Let  $\phi \in \mathcal{P}^{++}(\mathcal{F}^N)$  such that:*

$$\phi^\dagger(\omega, s, z) = \frac{\rho((b \circ \alpha)(\omega, s, z))}{\rho(z)} \times ((b \circ \alpha)(\omega, s, \cdot))'(z) \quad (5.22)$$

with  $b : \mathbb{D} \times \mathbb{M} \rightarrow \mathbb{R}$  a differentiable and increasing function in  $z$  and

$$\alpha : \mathfrak{M} \times \mathbb{R}^+ \times \mathbb{M} \rightarrow \mathbb{R}^+ \times \mathbb{M} \\ (\omega, s, z) \mapsto \left( \sum_{(T_n(\omega), Z_n(\omega)) : T_n(\omega) < t} Z_n(\omega), z \right). \quad (5.23)$$

Then, the process

$$Y(t) = \int_0^t \int x \, d\Gamma_\psi(s, x) \quad (5.24)$$

is a solution of the SDE

$$Z(t) = \int_0^t \int b(Z(t_-), x) \, dN(s, x) \quad (5.25)$$

if and only if  $\Gamma_\psi$  is the right inverse of  $\Gamma_\phi$ .

*Proof of Lemma 5.4.2.* Let assume  $\Gamma_\phi$  is right invertible with right inverse  $\Gamma_\psi$ . Then for  $t \in \mathbb{R}^+$ ,

$$\begin{aligned} \int_0^t \int x \, d\Gamma_\psi(s, x) &= \int_0^t \int \gamma_\psi(s, x) \, dN(s, x) \\ &= \int_0^t \int \gamma_\phi^*(\Gamma_\psi, s, x) \, dN(s, x). \end{aligned}$$

Since

$$(r \circ \gamma_\phi^*(\omega, \cdot))'(s, u) = \phi^\dagger(s, u)\rho(u),$$

we have:

$$\gamma_\phi^*(\omega, s, z) = b \circ \alpha(\omega, s, z).$$

Hence, rewriting Eqn (5.24), we get:

$$Y(t) = \int_0^t \int b(Y(t_-), x) \, dN(s, x).$$

Conversely, let assume that there exists a process  $Y$  for which (5.24) and (5.25) hold. Then there exists a predictable process  $U$  such that:

$$\int_0^t \int x \, d\Gamma_\psi(s, x) = \int_0^t \int U(\Gamma_\psi, s, x) \, dN(s, x).$$

As  $\Gamma_\psi$  is a marked point process, we have:  $\gamma_\psi^*(N, s, x) = U(\Gamma_\psi, s, x) \pi \otimes ds \otimes dx$ -a.s.. Hence, it admits a left inverse, i.e. there exists  $\phi \in \mathcal{P}^{++}(\mathcal{F}^N)$  such that  $\gamma_\phi = U$ .  $\square$

That characterization of SDEs as an integral of marked point process is based on assumptions that are different of the usual ones (see e.g. [Applebaum, 2009](#)).

## 5.5 Further works

Let  $(\check{\Omega}, \check{\mathcal{F}}, \check{\mathbf{P}}) = (W_B \times \mathfrak{N}, \mathcal{F}^B \otimes \mathcal{F}^N, \mathbf{P}^B \otimes \mathbf{P}^N)$ . We consider  $\mathbf{P}$  a probability measure on  $(\check{\Omega}, \check{\mathcal{F}}_\infty)$ . We denote by  $L^p(\check{\Omega} \rightarrow \mathbb{R}, \mathbf{P})$  the set of  $p$ -th integrable functionals on  $(\check{\Omega}, \check{\mathcal{F}}_\infty, \mathbf{P})$  for  $p > 0$ .

**Definition 5.5.1** (Lévy measure). Let  $\nu(dx)$  a Borel measure defined on  $\mathbb{R} \setminus \{0\}$ . We say that it is a Lévy measure if

$$\int_{\mathbb{R}} (1 \wedge x^2) \nu(dx) < +\infty. \quad (5.26)$$

It is worth noting that every Lévy measure on  $\mathbb{R}^d \setminus \{0\}$  is a  $\sigma$ -finite measure (see [Applebaum, 2009](#), p.29).

**Example 5.5.2** (Stable processes). The associated Lévy measure is

$$\nu(dx) = \frac{1}{|x|^{\alpha+1}} \mathbf{1}_{\{x \neq 0\}} dx$$

where  $0 < \alpha < 2$ .

Let  $X$  be a  $d$ -dimensional Lévy process under  $\mathbf{P}$ , càdlàg, with Lévy measure  $\nu$  on  $\mathbb{R}$  such that:  $X \in L^2(\check{\Omega} \rightarrow \mathbb{R}, \mathbf{P})$  then the *Lévy-Itô decomposition* ([Itô, 1942](#); [Lévy, 1954](#)) yields:

$$X = X^{(1)} + X^{(2)} + X^{(3)}, \quad (5.27)$$

where

- $X^{(1)}$  is a Brownian motion.
- $X^{(2)}$  is a compound Poisson process which is a stochastic integral of a Poisson measure  $N$  on  $[0, T] \times (\mathbb{R}^d \setminus \{0\})$  independent of  $B$  with intensity measure  $dt \otimes \nu(\mathbb{R} \setminus (-1, 1))$ .
- $X^{(3)}$  is the  $\mathbf{P}$ -square-integrable process of compensated small jumps of  $N$  with  $\mathbf{P}$  almost surely a countable number of jumps on every finite interval, possibly infinite.

$$X(t) = u_0 t + \sigma_0 B(t) + \int_0^t \int_{|x|>1} x N(ds, dx) + \lim_{\epsilon \downarrow 0} \int_{(0,t]} \int_{\epsilon < |x| \leq 1} x \tilde{N}(ds, dx) \quad (5.28)$$

such that

$$\int_{|x| \leq 1} |x|^2 \nu(dx) < \infty.$$

Let  $(\mathfrak{M}, \mathcal{F}^N, \mathbf{P}^N, (\bar{N}(t, A), t \geq 0, B \in \mathbb{M}))$  be the canonical space of the Poisson measure. Finally, consider  $(W_B \times \mathfrak{M}, \mathcal{F}^B \otimes \mathcal{F}^N, \mathbf{P}^B \otimes \mathbf{P}^N)$  and write:

$$B(\omega', \omega, t) = \bar{B}(\omega', t) \text{ and } N(\omega', \omega, t) = \bar{N}(\omega, t). \quad (5.29)$$

**Theorem 5.5.3** For any nonnegative  $\check{\mathcal{F}}$ -predictable random measure  $\mu$ , there exists a unique probability measure  $\pi$  on  $(\check{\Omega}, \check{\mathcal{F}})$  such that

- there exists two measures  $\mu^B$  and  $\mu^N$  on  $W_B$  and  $\mathfrak{M}$  so that  $\pi(d\omega', d\omega) = \mu^B(d\omega') \otimes \mu^N(d\omega)$ ;
- $N(\omega', \omega, \cdot)$  has  $\mu(\omega', \omega, \cdot)$  as dual predictable projection under  $\pi$ .

*Proof of Theorem 5.5.3.* We apply the previous theorem to  $\varphi(\omega', \cdot)$  for  $\omega' \in W_B$ . Let  $\mu^B$  a probability measure on  $W_B$ , Then, for  $\mu^B$  almost all  $\omega' \in W_B$ , for any  $A \in \mathcal{B}(\mathbb{M})$ , there exists a unique measure  $\mu^N$  on  $(\mathfrak{M}, \mathcal{F}_N^\infty)$  under which

$$N(\omega, t, A) - \varphi(\omega', \omega, (0, t], A) \text{ is a } \mathcal{F}^N\text{-local martingale,}$$

i.e.

$$\mathbf{E}_\mu [N(t, A) - N(s, A) | \mathcal{F}_s^N] = \mathbf{E}_{\mu^N} [\phi(\omega', (0, t], A) - \phi(\omega', (0, s], A) | \mathcal{F}_s^N].$$

Then on  $(W_B \times \mathfrak{M}, \mathcal{F}_\infty^B \otimes \mathcal{F}_\infty^N)$ , using that  $N(\omega, \cdot) = N(\omega', \omega, \cdot)$ :

$$\mathbf{E}_{\mu^B \otimes \mu^N} [N(t, A) - N(s, A) | \mathcal{F}_s^B \otimes \mathcal{F}_s^N] = \mathbf{E}_\mu [N(t, A) - N(s, A) | \mathcal{F}_s^N].$$

Hence, the map  $\omega' \mapsto \mathbf{E}_{\mu^N} [\phi(\omega', (0, t], A) - \phi(\omega', (0, s], A) | \mathcal{F}_s^N]$  is constant and:

$$\mathbf{E}_{\mu^B \otimes \mu^N} [N(t, A) - N(s, A) | \mathcal{F}_s^N] = \mathbf{E}_{\mu^B \otimes \mu^N} [\phi(\omega', (0, t], A) - \phi(\omega', (0, s], A) | \mathcal{F}_s^B \otimes \mathcal{F}_s^N].$$

Thus, under  $\mu^B \otimes \mu^N$ ,

$$t \mapsto N(t, A) - \varphi((0, t], A) \text{ is a } \mathcal{F}^B \otimes \mathcal{F}^N\text{-local martingale.}$$

Henceforth, the result holds. □

Using that theorem, it should be possible to merge the results of invertibility frameworks as to tackle solutions of stochastic differential equations defined as Lévy-type integrals.

The contribution of a Brownian component on the perturbation  $\gamma$  on the Poisson space is examined in [Privault \(2003\)](#), even without assumption of independence of the Brownian motion and the Poisson measure. We would like to finish with some open questions which are the following ones:

1. Can we describe, in the invertibility framework, solution of that following type of equations?

$$X(t) = \int_0^t u(X(s)) ds + B(t) + Z(t), \quad (5.30)$$

where  $B$  is a Brownian motion and  $Z$  is an integral with respect to a marked point process as previously introduced.

2. Is it possible to implement a limit procedure for [Lemma 5.4.2](#)?



# Bibliography

- [1] R. Adamczak, B. Polaczyk, and M. Strzelecki. Modified log-Sobolev inequalities, Beckner inequalities and moment estimates. *Journal of Functional Analysis*, 282(7):109349, 2022. (Cited on p. 42).
- [2] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003. (Cited on p. 108).
- [3] S. Albeverio and N. Smorodina. A distributional approach to multiple stochastic integrals and transformations of the Poisson measure. *Acta Applicandae Mathematica*, 94:1–19, 2006. (Cited on p. 92 and 103).
- [4] A. Anastasiou, A. Barp, F.-X. Briol, B. Ebner, R. E. Gaunt, F. Ghaderinezhad, J. Gorham, A. Gretton, C. Ley, and Q. Liu. Stein’s method meets computational statistics: A review of some recent developments. *Statistical Science*, 38(1):120–139, 2023. (Cited on p. 24).
- [5] D. Applebaum. *Lévy processes and stochastic calculus*, volume 116 of *Cambridge studies in advanced mathematics*. Cambridge University Press, second edition, 2009. (Cited on p. 116).
- [6] B. Arras, E. Azmoodeh, G. Poly, and Y. Swan. A bound on the Wasserstein-2 distance between linear combinations of independent random variables. *Stochastic processes and their Applications*, 129(7):2341–2375, 2019. (Cited on p. 50).
- [7] B. Arras, J.-C. Breton, A. Deshayes, O. Durieu, and R. Lachièze-Rey. Some recent advances for limit theorems. In *ESAIM: Proceedings and Surveys*, volume 68, pages 73–96. EDP Sciences, 2020. (Cited on p. 23).
- [8] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *The Annals of Probability*, 17(1):9–25, 1989. (Cited on p. 29).
- [9] R. Arratia, L. Goldstein, and L. Gordon. Poisson Approximation and the Chen-Stein Method. *Statistical Science*, 5(4):403 – 424, 1990. (Cited on p. 29).
- [10] R. Arratia, L. Goldstein, and F. Kochman. Size bias for one and all. *Probability Surveys*, 16:1 – 61, 2019. (Cited on p. 23).
- [11] T. Austin. On exchangeable random variables and the statistics of large graphs and hypergraphs. *Probability Surveys*, 5:80–145, 2008. (Cited on p. 33, 34, and 71).
- [12] E. Azmoodeh, S. Campese, and G. Poly. Fourth moment theorems for Markov diffusion generators. *Journal of Functional analysis*, 266(4):2341–2359, 2014. (Cited on p. 16, 17, 31, 32, 34, 35, 57, and 60).



- [13] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2013. (Cited on p. 12 and 48).
- [14] A. Barbour, L. Holst, and S. Janson. *Poisson Approximation*. Number 2 in Oxford Studies in Probability. Clarendon Press, 1992. (Cited on p. 73).
- [15] A. D. Barbour. Poisson convergence and random graph. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 92, pages 349–359. Cambridge University Press, 1982. (Cited on p. 18).
- [16] A. D. Barbour. Stein’s method and poisson process convergence. *Journal of Applied Probability*, 25(A):175–184, 1988. (Cited on p. 23).
- [17] A. D. Barbour. Stein’s method for diffusion approximations. *Probability theory and related fields*, 84(3):297–322, 1990. (Cited on p. 23).
- [18] A. D. Barbour, M. Karoński, and A. Ruciński. A central limit theorem for decomposable random variables with applications to random graphs. *Journal of Combinatorial Theory, Series B*, 47(2):125–145, 1989. (Cited on p. 29, 35, and 73).
- [19] A. Beurling and J. Deny. Espaces de Dirichlet: I. le cas élémentaire. *Acta Mathematica*, 99(1):203–224, 1958. (Cited on p. 12).
- [20] A. Beurling and J. Deny. Dirichlet spaces. *Proceedings of the National Academy of Sciences*, 45(2):208–215, 1959. (Cited on p. 12).
- [21] B. B. Bhattacharya, S. Das, and S. Mukherjee. Motif estimation via subgraph sampling: The fourth-moment phenomenon. *The Annals of Statistics*, 50(2):987–1011, 2022. (Cited on p. 67 and 74).
- [22] P. Biane. Chaotic representation for finite Markov chains. *Stochastics and Stochastic Reports*, 30(1):61–68, 1990. (Cited on p. 14).
- [23] K. Bichteler and J. Jacod. Calcul de Malliavin pour les diffusions avec sauts : existence d’une densité dans le cas unidimensionnel. *Séminaire de probabilités de Strasbourg*, 17: 132–157, 1983. (Cited on p. 95).
- [24] K. Bichteler, J. Gravereaux, and J. Jacod. *Malliavin Calculus for Processes with Jumps*, volume 2 of *Stochastic Monographs: Theory and Applications of Stochastic Processes*. Gordon and Breach Science Publishers, 1987. (Cited on p. 12, 18, 89, and 92).
- [25] J.-M. Bismut. Martingales, the Malliavin calculus and hypoellipticity under general Hörmander’s conditions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 56(4):469–505, 1981. (Cited on p. 89).
- [26] J.-M. Bismut. Calcul des variations stochastique et processus de sauts. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 63(2):147–235, 1983. (Cited on p. 14, 18, 89, 95, and 98).
- [27] R. Boel, P. Varaiya, and E. Wong. Martingales on Jump Processes. I: Representation Results; II: Applications. *SIAM Journal on Control*, 13(5):999–1021, 1975. (Cited on p. 92).

- [28] N. Bouleau and L. Denis. *Dirichlet Forms Methods for Poisson Point Measures and Lévy Processes*, volume 76 of *Probability Theory and Stochastic Modelling*. Springer, 2015. (Cited on p. 89).
- [29] N. Bouleau and F. Hirsch. *Dirichlet Forms and Analysis on Wiener Space*, volume 14 of *de Gruyter Studies in Mathematics*. De Gruyter, 1991. (Cited on p. 12 and 13).
- [30] J. Bourgain and G. Kalai. Influences of variables and threshold intervals under group symmetries. *Geometric and Functional Analysis*, 7(3):438–461, 1997. (Cited on p. 73).
- [31] S. Bourguin and G. Peccati. The Malliavin–Stein method on the Poisson space. In *Stochastic analysis for Poisson point processes: Malliavin calculus, Wiener-Itô chaos expansions and stochastic geometry*, volume 7 of *Bocconi & Springer Series*, pages 185–228. Springer, 2016. (Cited on p. 29).
- [32] A. V. Bulinski. Conditional central limit theorem. *Theory of Probability & Its Applications*, 61(4):613–631, 2017. (Cited on p. 52).
- [33] S. Chatterjee. A new method of normal approximation. *The Annals of Probability*, 36(4):1584–1610, 2008. (Cited on p. 50).
- [34] S. Chatterjee, P. Diaconis, and E. Meckes. Exchangeable pairs and Poisson approximation. *Probability Surveys*, 2:64 – 106, 2005. (Cited on p. 73).
- [35] L. H. Y. Chen. Poisson Approximation for Dependent Trials. *The Annals of Probability*, 3(3):534 – 545, 1975. (Cited on p. 23).
- [36] L. H. Y. Chen and G. Poly. Stein’s method, Malliavin calculus, Dirichlet forms and the fourth moment theorem. In *Festschrift Masatoshi Fukushima: In Honor of Masatoshi Fukushima’s Sanju*, volume 17 of *Interdisciplinary Mathematical Sciences*, chapter 6, pages 107–130. World Scientific Publishing Company, 2015. (Cited on p. 31).
- [37] L. H. Y. Chen and Q.-M. Shao. Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli*, 13(2):581–599, 2007. (Cited on p. 50).
- [38] L. H. Y. Chen, L. Goldstein, and Q.-M. Shao. *Normal approximation by Stein’s method*. Probability and Its Applications. Springer, 2011. (Cited on p. 52 and 55).
- [39] K. L. Chung. *A course in probability theory*. Academic Press, 1974. (Cited on p. 29).
- [40] M. Coulson, R. E. Gaunt, and G. Reinert. Poisson approximation of subgraph counts in stochastic block models and a graphon model. *ESAIM: Probability and Statistics*, 20:131–142, 2016. (Cited on p. 18).
- [41] L. Coutin and L. Decreusefond. Invertibility of functionals of the Poisson process and applications. In preparation, 2023. (Cited on p. 19, 89, 95, 98, and 103).
- [42] P. De Jong. *Central limit theorems for generalized multilinear forms*, volume 61 of *CWI Tracts*. Centrum voor Wiskunde & Informatica, Amsterdam, 1989. (Cited on p. 30).
- [43] P. De Jong. A central limit theorem for generalized multilinear forms. *Journal of Multivariate Analysis*, 34(2):275–289, 1990. (Cited on p. 62 and 74).

- [44] P. De Jong. A central limit theorem with applications to random hypergraphs. *Random Structures & Algorithms*, 8(2):105–120, 1996. (Cited on p. 33, 65, 74, 76, and 79).
- [45] L. Decreasefond. Perturbation Analysis and Malliavin Calculus. *Annals of Applied Probability*, 8(2):496–523, 1998. (Cited on p. 97).
- [46] L. Decreasefond. Perturbation analysis and Malliavin calculus. *The Annals of Applied Probability*, 8(2):496 – 523, 1998. (Cited on p. 89).
- [47] L. Decreasefond. The Stein-Dirichlet-Malliavin method. In *ESAIM: Proceedings and Surveys*, volume 51, pages 49–59. EDP Sciences, 2015. (Cited on p. 26).
- [48] L. Decreasefond. *Selected Topics in Malliavin Calculus: Chaos, Divergence and So Much More*, volume 10 of *Bocconi & Springer Series*. Springer, 2022. (Cited on p. 16).
- [49] L. Decreasefond and H. Halconruy. Malliavin and Dirichlet structures for independent random variables. *Stochastic Processes and their Applications*, 129(8):2611–2653, 2019. (Cited on p. 17, 33, 34, 35, 38, 46, 50, 52, and 81).
- [50] L. Decreasefond, A. Joulin, and N. Savy. Upper bounds on Rubinstein distances on configuration spaces and applications. *Communications on stochastic analysis*, 4(3):377–399, 2010. (Cited on p. 114).
- [51] A. Dermoune, P. Krée, and L.-M. Wu. Calcul stochastique non adapté par rapport à la mesure aléatoire de Poisson. *Séminaire de probabilités de Strasbourg*, 22:477–484, 1988. (Cited on p. 14).
- [52] P. S. Dey and G. Terlov. Stein’s method for conditional central limit theorem. *The Annals of Probability*, 51(2):723–773, 2023. (Cited on p. 52 and 83).
- [53] C. Döbler. Normal approximation via non-linear exchangeable pairs. *Latin American Journal of Probability and Mathematical Statistics*, 20:167, 2023. (Cited on p. 67).
- [54] C. Döbler and K. Krokowski. On the fourth moment condition for Rademacher chaos. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 55(1):61 – 97, 2019. (Cited on p. 74).
- [55] C. Döbler and G. Peccati. Quantitative de Jong theorems in any dimension. *Electronic Journal of Probability*, 22, 2017. (Cited on p. 63, 64, 65, and 67).
- [56] C. Döbler and G. Peccati. The fourth moment theorem on the Poisson space. *Annals of Probability*, 46(4):1878–1916, 2018. (Cited on p. 16 and 32).
- [57] C. Döbler and G. Peccati. Quantitative CLTs for symmetric  $U$ -statistics using contractions. *Electronic Journal of Probability*, 24, 2019. (Cited on p. 57 and 63).
- [58] C. Döbler, M. J. Kasprzak, and G. Peccati. Functional convergence of sequential  $U$ -processes with size-dependent kernels. *The Annals of Applied Probability*, 32(1):551–601, 2022. (Cited on p. 18).
- [59] R. M. Dudley. *Real Analysis and Probability*, volume 74 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 2002. (Cited on p. 35).

- [60] M. Duerinckx. On the size of chaos via Glauber calculus in the classical mean-field dynamics. *Communications in Mathematical Physics*, pages 1–41, 2021. (Cited on p. 17, 34, 39, 42, and 50).
- [61] N. T. Dung. Poisson and normal approximations for the measurable functions of independent random variables. *arXiv preprint arXiv:1807.10925*, 2018. (Cited on p. 34, 50, and 52).
- [62] N. T. Dung. Explicit rates of convergence in the multivariate CLT for nonlinear statistics. *Acta Mathematica Hungarica*, 158(1):173–201, 2019. (Cited on p. 50).
- [63] N. T. Dung. Rates of convergence in the central limit theorem for nonlinear statistics under relaxed moment conditions. *Acta Mathematica Vietnamica*, pages 1–26, 2021. (Cited on p. 45).
- [64] B. Efron and C. Stein. The Jackknife Estimate of Variance. *The Annals of Statistics*, 9(3):586 – 596, 1981. (Cited on p. 47).
- [65] P. Eichelsbacher and B. Rednoß. Kolmogorov bounds for decomposable random variables and subgraph counting by the Stein–Tikhomirov method. *Bernoulli*, 29(3):1821–1848, 2023. (Cited on p. 74).
- [66] P. Eichelsbacher and C. Thäle. New Berry-Esseen bounds for non-linear functionals of Poisson random measures. *Electronic Journal of Probability*, 19:1 – 25, 2014. (Cited on p. 32).
- [67] P. Eichelsbacher, B. Rednoß, C. Thäle, and G. Zheng. A simplified second-order Gaussian Poincaré inequality in discrete setting with applications. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 59(1):271 – 302, 2023. (Cited on p. 48 and 74).
- [68] P. Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53:292–294, 1947. (Cited on p. 69 and 70).
- [69] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc, 1986. (Cited on p. 11).
- [70] X. Fang and Y. Koike. New error bounds in multivariate normal approximations via exchangeable pairs with applications to Wishart matrices and fourth moment theorems. *The Annals of Applied Probability*, 32(1):602–631, 2022. (Cited on p. 65).
- [71] H. Föllmer. Time reversal on Wiener space. In S. A. Albeverio, P. Blanchard, and L. Streit, editors, *Stochastic Processes – Mathematics and Physics*, pages 119–129. Springer, 1986. (Cited on p. 19 and 103).
- [72] H. Föllmer and A. Wakolbinger. Time reversal of infinite-dimensional diffusions. *Stochastic Processes and their Applications*, 22(1):59–77, 1986. (Cited on p. 19 and 103).
- [73] E. Friedgut. Hunting for sharp thresholds. *Random Structures & Algorithms*, 26(1-2): 37–51, 2005. (Cited on p. 73).
- [74] E. Friedgut and J. Bourgain. Sharp thresholds of graph properties, and the  $k$ -sat problem. *Journal of the American mathematical Society*, 12(4):1017–1054, 1999. (Cited on p. 73).

- [75] E. Friedgut and G. Kalai. Every monotone graph property has a sharp threshold. *Proceedings of the American mathematical Society*, 124(10):2993–3002, 1996. (Cited on p. 73).
- [76] K. O. Friedrich. A Berry-Esseen Bound for Functions of Independent Random Variables. *The Annals of Statistics*, 17(1):170 – 183, 1989. (Cited on p. 33).
- [77] C. Gao and Z. Ma. Minimax rates in network analysis: Graphon estimation, community detection and hypothesis testing. *Statistical Science*, 36(1):16–33, 2021. (Cited on p. 85).
- [78] B. Gaveau and P. Trauber. L’intégrale stochastique comme opérateur de divergence dans l’espace fonctionnel. *Journal of Functional Analysis*, 46(2):230–238, 1982. (Cited on p. 13).
- [79] G. Ghoshal, V. Zlatić, G. Caldarelli, and M. E. Newman. Random hypergraphs and their applications. *Physical Review E*, 79(6):066118, 2009. (Cited on p. 69).
- [80] E. Giné. Decoupling and limit theorems for  $U$ -statistics and  $U$ -processes. In *Lectures on Probability Theory and Statistics*, volume 1665 of *Lecture Notes in Mathematics*, pages 1–35. Springer, 1997. (Cited on p. 17).
- [81] I. V. Girsanov. On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability & Its Applications*, 5(3):285–301, 1960. (Cited on p. 91).
- [82] L. Goldstein. Bounds on the constant in the mean central limit theorem. *Annals of Probability*, 38(4):1672–1689, 2010. (Cited on p. 52).
- [83] L. Goldstein and G. Reinert. Stein’s method and the zero bias transformation with application to simple random sampling. *The Annals of Applied Probability*, 7(4):935 – 952, 1997. (Cited on p. 23 and 29).
- [84] L. Goldstein and Y. Rinott. Multivariate normal approximations by Stein’s method and size bias couplings. *Journal of Applied Probability*, 33(1):1–17, 1996. (Cited on p. 29).
- [85] F. Götze. On the Rate of Convergence in the Multivariate CLT. *The Annals of Probability*, 19(2):724 – 739, 1991. (Cited on p. 23).
- [86] W. Grzenda and W. Zieba. Conditional central limit theorem. In *Int. Math. Forum*, volume 3, pages 1521–1528, 2008. (Cited on p. 52).
- [87] A. G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58:83–90, 1971. (Cited on p. 103).
- [88] R. Helmers. The Berry-Esseen bound for  $U$ -statistics. *Theory and Related Topics III*, 1: 497–512, 1982. (Cited on p. 33).
- [89] W. Hoeffding. A Class of Statistics with Asymptotically Normal Distribution. *The Annals of Mathematical Statistics*, 19(3):293 – 325, 1948. (Cited on p. 53).
- [90] H. Holden, T. Lindstrøm, B. Øksendal, and J. Ubøe. Discrete Wick calculus and stochastic functional equations. *Potential Analysis*, 1(3):291–306, 1992. (Cited on p. 15).

- [91] H. Holden, T. Lindstrøm, B. Øksendal, and J. Ubøe. Discrete Wick products. *Stochastic analysis and related topics (Oslo, 1992)*, pages 123–148, 1993. (Cited on p. 15).
- [92] C. Houdré and N. Privault. Concentration and deviation inequalities in infinite dimensions via covariance representations. *Bernoulli*, 8(6):697–720, 2002. (Cited on p. 46).
- [93] Y. Ito. Generalized Poisson Functionals. *Probability Theory and Related Fields*, 77(1): 1–28, 1988. (Cited on p. 14).
- [94] K. Itô. On stochastic processes. (I). (Infinitely divisible laws of probability). *Japanese Journal of Mathematics*, 18:261–301, 1942. (Cited on p. 116).
- [95] J. Jacod. Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 31(3):235–253, 1975. (Cited on p. 94, 95, and 97).
- [96] J. Jacod. *Calcul Stochastique et Problèmes de Martingales*, volume 714 of *Lecture Notes in Mathematics*. Springer, 1979. (Cited on p. 89 and 95).
- [97] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der mathematischen Wissenschaften*. Springer, second edition, 2003. (Cited on p. 97).
- [98] S. Janson and P. Diaconis. Graph limits and exchangeable random graphs. *Rendiconti di Matematica e delle sue Applicazioni. Serie VII*, pages 33–61, 2008. (Cited on p. 85).
- [99] S. Janson, T. Luczak, and A. Ruciński. *Random Graphs*. John Wiley & Sons, Inc., 2000. (Cited on p. 29, 72, 73, and 84).
- [100] O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer, second edition, 2002. (Cited on p. 36, 43, 76, and 92).
- [101] O. Kallenberg. *Random Measures, Theory and Applications*, volume 77 of *Probability Theory and Stochastic Modelling*, chapter 2. Springer, 2017. (Cited on p. 93).
- [102] G. Kaur and A. Röllin. Higher-order fluctuations in dense random graph models. *Electronic Journal of Probability*, 26:1–36, 2021. (Cited on p. 83 and 84).
- [103] K. Krokowski, A. Reichenbachs, and C. Thäle. Discrete Malliavin–Stein method: Berry–Esseen bounds for random graphs and percolation. *The Annals of Probability*, 45(2): 1071–1109, 2017. (Cited on p. 35 and 74).
- [104] R. Lachièze-Rey and G. Peccati. Fine Gaussian fluctuations on the Poisson space, I: contractions, cumulants and geometric random graphs. *Electronic Journal of Probability*, 18, 2013. (Cited on p. 16 and 32).
- [105] R. Lachièze-Rey and G. Peccati. Fine Gaussian fluctuations on the Poisson space II: rescaled kernels, marked processes and geometric  $U$ -statistics. *Stochastic Processes and their Applications*, 123(12):4186–4218, 2013. (Cited on p. 16).
- [106] R. Lachièze-Rey and G. Peccati. New Berry–Esseen bounds for functionals of binomial point processes. *The Annals of Applied Probability*, 27(4):1992–2031, 2017. (Cited on p. 45 and 50).



- [107] R. Lachièze-Rey and M. Reitzner.  $U$ -statistics in stochastic geometry. In *Stochastic Analysis for Poisson Point Processes*, volume 7 of *Bocconi & Springer Series*, pages 229–253. Springer, 2016. (Cited on p. 16).
- [108] R. Lassalle. *Stochastic Invertibility and Related Topics*. PhD thesis, Telecom ParisTech, 2012. (Cited on p. 19 and 108).
- [109] M. Ledoux. The geometry of Markov diffusion generators. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, Ser. 6, 9(2):305–366, 2000. (Cited on p. 12).
- [110] M. Ledoux. Chaos of a Markov operator and the fourth moment condition. *Annals of Probability*, 40(6):2439–2459, 2012. (Cited on p. 31 and 60).
- [111] J. Lehec. Representation formula for the entropy and functional inequalities. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 49(3):885–899, 2013. (Cited on p. 112).
- [112] P. Lévy. *Théorie de l’addition des variables aléatoires*. Gauthier-Villars, 2 edition, 1954. (Cited on p. 116).
- [113] L. Lovász. *Large networks and graph limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012. (Cited on p. 17, 18, 69, and 72).
- [114] L. Lovász and B. Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, 2006. (Cited on p. 85).
- [115] C. McDiarmid. On the method of bounded differences. *Surveys in combinatorics*, 141(1):148–188, 1989. (Cited on p. 48).
- [116] P.-A. Meyer. Démonstration probabiliste de certaines inégalités de Littlewood-Paley. Exposé II : l’opérateur carré du champ. *Séminaire de probabilités de Strasbourg*, 10:142–161, 1976. (Cited on p. 11).
- [117] P.-A. Meyer. *Quantum Probability for Probabilists*. Number 1538 in Lecture Notes in Mathematics. Springer, 1993. (Cited on p. 15).
- [118] I. Nourdin and G. Peccati. Stein’s method on Wiener chaos. *Probability Theory and Related Fields*, 145(1-2):75–118, 2009. (Cited on p. 16, 24, and 30).
- [119] I. Nourdin and G. Peccati. *Normal approximations with Malliavin calculus: from Stein’s method to universality*. Number 192 in Cambridge Tracts in Mathematics. Cambridge University Press, 2012. (Cited on p. 29 and 30).
- [120] I. Nourdin, G. Peccati, and G. Reinert. Second order Poincaré inequalities and CLTs on Wiener space. *Journal of Functional Analysis*, 257(2):593–609, 2009. (Cited on p. 48).
- [121] I. Nourdin, G. Peccati, and G. Reinert. Invariance principles for homogeneous sums: Universality of Gaussian Wiener chaos. *Annals of probability*, 38(5):1947–1985, 2010. (Cited on p. 30).
- [122] I. Nourdin, G. Peccati, and A. Réveillac. Multivariate normal approximation using Stein’s method and Malliavin calculus. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 46(1):45–58, 2010. (Cited on p. 18).

- [123] I. Nourdin, G. Peccati, G. Poly, and R. Simone. Classical and free fourth moment theorems: Universality and thresholds. *Journal of Theoretical Probability*, 29(2):653–680, 2016. (Cited on p. 30).
- [124] D. Nualart. *The Malliavin calculus and related topics*, volume 1995 of *Probability and Its Applications*. Springer, second edition, 2006. (Cited on p. 12 and 114).
- [125] D. Nualart and G. Peccati. Central limit theorems for sequences of multiple stochastic integrals. *Annals of Probability*, 33(1):177–193, 2005. (Cited on p. 24 and 29).
- [126] D. Nualart and W. Schoutens. Chaotic and predictable representations for Lévy processes. *Stochastic processes and their applications*, 90(1):109–122, 2000. (Cited on p. 14).
- [127] S. Orey. *Radon-Nikodym Derivatives of Probability Measures: Martingale Methods*. Department of the Foundations of Mathematical Sciences, Tokyo University of Education, 1974. (Cited on p. 91).
- [128] G. Peccati. Hoeffding-ANOVA decompositions for symmetric statistics of exchangeable observations. *The Annals of Probability*, 32(3):1796–1829, 2004. (Cited on p. 42).
- [129] G. Peccati and C. A. Tudor. *Gaussian Limits for Vector-valued Multiple Stochastic Integrals*, volume 1857 of *Lecture Notes in Mathematics*, pages 247–262. Springer, 2005. (Cited on p. 24).
- [130] G. Peccati, J. L. Solé, M. S. Taqqu, and F. Utzet. Stein’s method and Normal approximation of Poisson functionals. *The Annals of Probability*, 38(2):443 – 478, 2010. (Cited on p. 16).
- [131] J. Picard. Formules de dualité sur l’espace de Poisson. *Annales de l’Institut Henri Poincaré Probabilités et statistiques*, 32(4):509–548, 1996. (Cited on p. 14).
- [132] N. Privault. Girsanov theorem for anticipative shifts on Poisson space. *Probability Theory and Related Fields*, 104(1):61–76, 1996. (Cited on p. 92 and 103).
- [133] N. Privault. Calcul des variations stochastique pour la mesure de densité uniforme. *Potential Analysis*, 7(2):577–601, 1997. (Cited on p. 14).
- [134] N. Privault. Quasi-invariance for Lévy processes under anticipating shifts. In *Stochastic Analysis and Related Topics VIII*, pages 181–202. Birkhäuser Basel, 2003. (Cited on p. 117).
- [135] N. Privault. Stochastic analysis of Bernoulli processes. *Probability Surveys*, 5:435–483, 2008. (Cited on p. 15, 33, and 35).
- [136] N. Privault. *Stochastic Analysis in Discrete and Continuous settings with normal martingales*, volume 1982 of *Lectures Notes in Mathematics*. Springer, 2009. (Cited on p. 14).
- [137] N. Privault and W. Schoutens. Discrete chaotic calculus and covariance identities. *Stochastics and Stochastic Reports*, 72(3-4):289–316, 2002. (Cited on p. 15).
- [138] N. Privault and G. Serafin. Normal approximation for sums of discrete  $U$ -statistics—application to Kolmogorov bounds in random subgraph counting. *Bernoulli*, 26(1):587–615, 2020. (Cited on p. 35, 50, 52, and 74).



- [139] N. Privault and G. Serafin. Berry-Esseen bounds for functionals of independent random variables. *Electronic Journal of Probability*, 27:1–37, 2022. (Cited on p. 16, 50, 57, 67, and 74).
- [140] N. Privault and G. L. Torrisi. The Stein and Chen-Stein methods for functionals of non-symmetric Bernoulli processes. *ALEA Lat. Am. J. Probab. Math. Stat*, 12(1):309–356, 2015. (Cited on p. 33 and 35).
- [141] S. Rachev. *Probability Metrics and the Stability of Stochastic Models*, volume 269 of *Wiley Series in Probability and Statistics - Applied Probability and Statistics Section*. Wiley, 1991. (Cited on p. 24).
- [142] B. P. Rao. Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics*, 61(2):441–460, 2009. (Cited on p. 36 and 52).
- [143] G. Reinert. Couplings for normal approximations with Stein’s method. In *Microsurveys in discrete probability: DIMACS workshop*, volume 41 of *Series in Discrete Mathematics and Theoretical Computer Science*, pages 193–207. American Mathematical Society, DIMACS, 1998. (Cited on p. 52).
- [144] G. Reinert, I. Nourdin, and G. Peccati. Stein’s method and stochastic analysis of Rademacher functionals. *Electronic Journal of Probability*, 15:1703–1742, 2010. (Cited on p. 16, 33, and 35).
- [145] J.-P. Roth. Opérateurs dissipatifs et semi-groupes dans les espaces de fonctions continues. *Annales de l’Institut Fourier*, 26(4):1–97, 1976. (Cited on p. 11).
- [146] A. Ruciński. When are small subgraphs of a random graph normally distributed? *Probability Theory and Related Fields*, 78(1):1–10, 1988. (Cited on p. 73).
- [147] A. Röllin. Kolmogorov bounds for the normal approximation of the number of triangles in the Erdos-Rényi random graph. *Probability in the Engineering and Informational Sciences*, 36(3):747–773, 2022. (Cited on p. 35 and 74).
- [148] M. Schulte. Normal approximation of Poisson functionals in Kolmogorov distance. *Journal of theoretical probability*, 29(1):96–117, 2016. (Cited on p. 33).
- [149] J. Schuppen and E. Wong. Transformation of Local Martingales Under a Change of Law. *The Annals of Probability*, 2(5):879–888, 1974. (Cited on p. 100).
- [150] Q.-M. Shao and Z.-S. Zhang. Berry–Esseen bounds of normal and non-normal approximation for unbounded exchangeable pairs. *Annals of Probability*, 47(1):61–108, 2019. (Cited on p. 65).
- [151] R. Simone. *Universality and Fourth Moment Theorem for homogeneous sums. Orthogonal polynomials and apolarity*. PhD thesis, Università degli Studi della Basilicata, 2017. (Cited on p. 30).
- [152] A. V. Skorokhod. On the differentiability of measures which correspond to stochastic processes. i. processes with independent increments. *Theory of Probability & Its Applications*, 2(4):407–432, 1957. (Cited on p. 92).

- [153] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*, volume 6, pages 583–603. University of California Press, 1972. (Cited on p. 23).
- [154] C. Stein. *Approximate Computation of Expectations*, volume 7 of *IMS Lecture Notes - Monograph Series*. Institute of Mathematical Statistics, 1986. (Cited on p. 23 and 29).
- [155] D. W. Stroock. The Malliavin calculus, a functional analytic approach. *Journal of Functional Analysis*, 44(2):212–257, 1981. (Cited on p. 13 and 17).
- [156] D. W. Stroock. The Malliavin calculus and its applications. In D. Williams, editor, *Stochastic Integrals*, volume 851 of *Lecture Notes in Mathematics*, pages 394–432. Springer, 1981. (Cited on p. 13).
- [157] D. W. Stroock. The Malliavin calculus and its application to second order parabolic differential equations: Part I. *Mathematical systems theory*, 14(1):25–65, 1981. (Cited on p. 13).
- [158] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1979. (Cited on p. 10 and 89).
- [159] T. Temčinas, V. Nanda, and G. Reinert. Multivariate central limit theorems for random clique complexes. *Journal of Applied and Computational Topology*, 2023. (Cited on p. 84).
- [160] I. Tyurin. New estimates of the convergence rate in the Lyapunov theorem, 2009. (Cited on p. 52).
- [161] A. S. Üstünel. Entropy, invertibility and variational calculus of adapted shifts on Wiener space. *Journal of Functional Analysis*, 257(11):3655–3689, 2009. (Cited on p. 103).
- [162] A. S. Üstünel. Variational calculation of Laplace transforms via entropy on Wiener space and applications. *Journal of Functional Analysis*, 267(8):3058 – 3083, 2014. (Cited on p. 104).
- [163] A. S. Üstünel and M. Zakai. *Transformation of Measure on Wiener Space*. Springer Monographs in Mathematics. Springer, 2000. (Cited on p. 19, 89, and 103).
- [164] A. S. Üstünel and M. Zakai. Sufficient conditions for the invertibility of adapted perturbations of identity on the Wiener space. *Probability theory and related fields*, 139(1-2): 207–234, 2007. (Cited on p. 103).
- [165] W. R. van Zwet. A Berry-Esseen bound for symmetric statistics. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 66(3):425–440, 1984. (Cited on p. 33).
- [166] D. Williams. “to begin at the beginning: ...”. In D. Williams, editor, *Stochastic Integrals*, volume 851 of *Lecture Notes in Mathematics*, pages 1–55. Springer, 1981. (Cited on p. 18).
- [167] J.-A. Yan. Développement des distributions suivant les chaos de Wiener et applications à l’analyse stochastique. *Séminaire de probabilités de Strasbourg*, 21:27–33, 1987. (Cited on p. 13).

- 
- [168] K. Yosida. *Functional Analysis*. Classics in Mathematics. Springer, 1995. (Cited on p. 42).
- [169] D.-M. Yuan, L.-R. Wei, and L. Lei. Conditional central limit theorems for a sequence of conditional independent random variables. *Journal of the Korean Mathematical Society*, 51(1):1–15, 2014. (Cited on p. 52).
- [170] H. Zhang, L. Song, Y. Li, and G. Y. Li. Hypergraph theory: Applications in 5G heterogeneous ultra-dense networks. *IEEE Communications Magazine*, 55(12):70–76, 2017. (Cited on p. 69).
- [171] X. Zhang. Clark-Ocone formula and variational representation for Poisson functionals. *The Annals of Probability*, 37(2):506–529, 2009. (Cited on p. 19 and 113).
- [172] Z.-S. Zhang. Berry–Esseen bounds for generalized  $U$ -statistics. *Electronic Journal of Probability*, 27:1–36, 2022. (Cited on p. 84).

**Titre :** Contributions à l'analyse stochastique pour structures sans propriété de diffusion

**Mots clés :** Calcul de Malliavin, Méthode de Stein, Variables aléatoires conditionnellement indépendantes, Mesure de Poisson, Théorème de Girsanov, Critère entropique, Equations différentielles stochastiques

**Résumé :** Cette thèse a pour sujet l'étude de structures sans propriété de diffusion. Nous nous intéressons à deux classes de telles structures.

Le premier sujet traite du calcul de Malliavin pour les variables aléatoires conditionnellement indépendantes qui est un cas de calcul de Malliavin discret. Il généralise aussi celui théorisé sur des produits dénombrables d'espaces de probabilité, pour les variables aléatoires indépendantes. Dans notre cas, l'intérêt d'un tel calcul est de venir compléter des résultats d'analyse stochastique avec des preuves d'inégalités fonctionnelles (inégalité de Poincaré, inégalité de McDiarmid) et de théorèmes limites. Une des applications phares est la détermination de la vitesse de convergence de théorèmes centraux limites via la méthode de Stein. En combinant le calcul de Malliavin avec la structure de Dirichlet sous-jacente aux variables aléatoires, nous obtenons une formule d'intégration par parties cruciale pour déterminer des

bornes supérieures sur les vitesses de convergence. Nous montrons des théorèmes limites quantitatifs, dont un théorème de quatrième moment avec reste. En particulier, nous discutons d'une application à la normalité asymptotique du comptage de motifs dans des hypergraphes aléatoires échangeables.

Le deuxième sujet étudie les fonctionnelles d'une mesure de Poisson en utilisant la notion d'inversibilité de transformations de cette mesure sur l'espace échantillon des mesures aléatoires. Nous utilisons l'identification de ces mesures et des processus ponctuels marqués associés. Les transformations inversibles sont obtenues via le théorème de Girsanov, en respectant l'absolue continuité par rapport à la mesure de référence. Il en résulte un critère entropique pour l'inversibilité des transformations. Enfin, nous faisons le lien avec les équations différentielles stochastiques dirigées par des mesures de Poisson.

**Title :** Contributions to stochastic analysis for non-diffusive structures

**Keywords :** Malliavin calculus, Stein's method, Conditionally independent random variables, Poisson measure, Girsanov's theorem, Entropy criterion, Stochastic differential equations

**Abstract :** This thesis is concerned with the study of non-diffusive structures. We focus on two classes of such structures.

The first subject deals with Malliavin calculus for conditionally independent random variables, which is a special case of discrete Malliavin calculus. It also generalizes the calculus that has been developed for countable products of probability spaces, for independent random variables. In our case, the interest of such a calculus is to complement results in stochastic analysis with proofs of functional inequalities (Poincaré inequality, McDiarmid's inequality) and limit theorems. One of the main applications is the determination of the convergence rate of central limit theorems via the Stein method. By combining Malliavin calculus with the underlying Dirichlet structure of the random variables, we obtain an integration by parts for-

mula which is key to the derivations of so-called Stein bounds of the rates of convergence. We show quantitative limit theorems, including a fourth moment theorem with remainder. In particular, we discuss an application to the asymptotic normality of motif counting in exchangeable random hypergraphs.

The second subject studies functionals of a Poisson measure using the notion of invertibility of transformations of that measure on the sample space of random measures. We use the identification of these measures and the associated marked point processes. Invertible transformations are obtained via the Girsanov's theorem, respecting absolute continuity with respect to the reference measure. This results in an entropy criterion for the invertibility of transformations. Finally, we make the connection with stochastic differential equations driven by Poisson measures.