



**HAL**  
open science

# Malliavin calculus and Dirichlet structures for independent random variables

Hélène Halconruy

► **To cite this version:**

Hélène Halconruy. Malliavin calculus and Dirichlet structures for independent random variables. Probability [math.PR]. Institut Polytechnique de Paris, 2020. English. NNT : 2020IPPAT016 . tel-03099427

**HAL Id: tel-03099427**

**<https://theses.hal.science/tel-03099427v1>**

Submitted on 6 Jan 2021

**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 : 2020IPPAT016

Thèse de doctorat



# Calcul de Malliavin et structures de Dirichlet pour des variables aléatoires indépendantes

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 Paris, le 14 septembre 2020, par

**HÉLÈNE HALCONRUY**

Composition du Jury :

Laure Coutin Professeur (Université de Toulouse)	Président
Elisa Alòs Professeur (Universitat Pompeu Fabra)	Rapporteur
Raphaël Lachièze-Rey Maître de conférences (Université Paris Descartes)	Rapporteur
Ivan Nourdin Professeur (Université du Luxembourg)	Examineur
Francesco Russo Professeur (ENSTA Paris)	Examineur
Laurent Decreusefond Professeur (Télécom Paris)	Directeur de thèse
Nicolas Marie Maître de conférences (Université Paris Nanterre)	Invité



# Contents

<b>Remerciements</b>	<b>5</b>
<b>Introduction</b>	<b>9</b>
Introduction (in English) . . . . .	9
Introduction (en Français) . . . . .	23
<b>I Malliavin calculus and Dirichlet structures for independent random variables</b>	<b>35</b>
<b>1 From Stein's to Stein-Dirichlet-Malliavin method</b>	<b>37</b>
1.1 Stein's method principle . . . . .	37
1.1.1 Taxonomy on probability metrics . . . . .	37
1.1.2 Stein's method principle . . . . .	39
1.2 The Stein-Dirichlet structure . . . . .	40
1.2.1 Conversion of the initial problem . . . . .	40
1.2.2 The Dirichlet structure point of view . . . . .	41
1.3 Dirichlet-Malliavin structures . . . . .	46
1.3.1 Some words about Malliavin calculus . . . . .	46
1.3.2 Malliavin calculus and Dirichlet structures: an intrinsic connection . . . . .	51
1.3.3 Examples and applications of Dirichlet-Malliavin structures . . . . .	52
1.4 Stein-Malliavin criterion on a Gaussian space . . . . .	56
<b>2 Discrete Dirichlet-Malliavin calculus</b>	<b>59</b>
2.1 Malliavin calculus for independent random variables . . . . .	60
2.1.1 Divergence . . . . .	63
2.1.2 Ornstein-Uhlenbeck semi-group and generator . . . . .	64
2.2 Functional identities . . . . .	66
2.3 Dirichlet structures . . . . .	70
2.3.1 Brownian motion . . . . .	70
2.3.2 Poisson point process . . . . .	72
2.4 Proofs of chapter 2 . . . . .	75
2.4.1 Proofs of Section 2.1 . . . . .	75
2.4.2 Proofs of Section 2.2 . . . . .	78

<b>3 Applications</b>	<b>85</b>
3.1 Representations . . . . .	85
3.2 Discrete Stein-Malliavin criterion . . . . .	88
3.3 Finance and Malliavin calculus . . . . .	94
3.4 Proofs of chapter 3 . . . . .	96
3.4.1 Proofs of Section 3.1 . . . . .	96
3.4.2 Proofs of Section 3.2 . . . . .	100
3.4.3 Proofs of Section 3.3 . . . . .	105
<b>II Malliavin calculus and enlargement of filtrations in discrete time: additional utility of an insider in a ternary model</b>	<b>107</b>
<b>4 Insider model and geometric Malliavin calculus</b>	<b>111</b>
4.1 Insider's trading and initial enlargement . . . . .	111
4.2 Hedging and portfolio's optimization in incomplete markets . . . . .	115
4.3 Malliavin calculus for jump processes . . . . .	119
4.4 Conclusion . . . . .	121
<b>5 Geometric Malliavin calculus</b>	<b>123</b>
5.1 Compound geometric processes . . . . .	123
5.2 Stochastic integrals . . . . .	125
5.3 Modified chaos representation . . . . .	128
5.4 Functional identities . . . . .	132
5.5 Link with Rademacher and Poisson settings . . . . .	134
5.6 Proofs of chapter 5 . . . . .	136
5.6.1 Proofs of Section 5.2 . . . . .	136
5.6.2 Proofs of Section 5.3 . . . . .	137
5.6.3 Proofs of Section 5.4 . . . . .	140
<b>6 The insider problem in the ternary model</b>	<b>143</b>
6.1 Framework: the ternary model . . . . .	143
6.1.1 Martingales measures in the ternary market model . . . . .	145
6.2 Enlargement of filtration . . . . .	146
6.2.1 The information drift process . . . . .	146
6.2.2 The martingale preserving measure . . . . .	148
6.3 Hedging formula in a ternary model . . . . .	149
6.4 Additional utility of the insider in a ternary model . . . . .	151
6.4.1 Theoretical results . . . . .	151
6.4.2 Comparison with the binomial and the Black-Scholes models . . . . .	154
6.4.3 Computations in the case $G = \mathbf{1}_{\{S_T \in [0, S_0]\}}$ . . . . .	154
6.5 Proofs of chapter 6 . . . . .	158
6.5.1 Proofs of Section 6.2 . . . . .	158
6.5.2 Proofs of Section 6.3 . . . . .	160
6.5.3 Proofs of Section 6.4 . . . . .	161

# Remerciements

Si c'est seule en scène qu'il faut me défendre ce travail, nombreux sont ceux qui m'ont aidée au montage, ou qui des coulisses continuent à me donner du courage. Aussi c'est de bonne grâce que je me plie à l'exercice des remerciements.

Les premiers que j'adresse vont à mon directeur de thèse, et metteur en scène d'un jour, Laurent Decreusefond. Un immense merci Laurent de n'avoir eu égard à l'atypicité de mon profil pour me suggérer la définition d'un gradient *discrètement* efficace, point de départ de l'expérience mathématique enthousiasmante que tu m'as donné la chance de vivre. Depuis les après-midis de début de thèse à noircir tableaux et feuilles de papier, jusqu'à cette soutenance finale, merci d'avoir distribué pendant ces années traits d'humour et pléthore d'encouragements, et de m'avoir fait visiter un monde pas toujours *chaotique* mais souvent *intégrable par parties*, pour qu'à la question de Leonard Gross :

*How far can you go with integration by parts?*

Je réponde finalement - et presque en chanson - Far far away.

À ses côtés le jour de la soutenance, cinq membres du jury que je tiens à remercier très chaleureusement pour le privilège qu'ils m'offrent d'en faire partie. I would like to thank Elisa Alòs for kindly agreeing to review this thesis. I am both pleased and honoured that you accepted to read my modest contributions. Un grand merci à Raphaël Lachièze-Rey d'avoir également accepté de rapporter cette thèse et de m'avoir suggéré au détour d'une conversation d'après séminaire à Descartes, de mettre en regard des points de vue "Hoeffding" et "Malliavin". Thank you both for writing a report on the current difficult working conditions of confinement; all the more generous!

Merci à François Roueff pour son suivi en tant que directeur adjoint de l'EDMH, dont j'ai apprécié conseils avisés et bienveillants et les agréables discussions que nous avons notamment eues à ces moments-là. Je tiens aussi à remercier Laure Coutin pour son chaleureux accueil à l'université Paul Sabatier lors d'un séjour très agréablement logistique à Toulouse. Je suis également honorée et très reconnaissante à Francesco Russo d'avoir gracieusement accepté de prendre part à mon jury. Enfin un très grand merci à Ivan Nourdin dont les travaux ont été particulièrement inspirants tout au long de cette thèse et dont je me réjouis d'honorer la généreuse invitation à exposer à l'Université du Luxembourg.

La grande modestie de mes travaux ne reflète pas la fascination que les mathématiques exercent en moi. J'y associe l'esthétique, pour ne pas dire l'artistique. Abstraites et distillées comme une statue de Giacometti, elles exhortent à une création dynamique et un remodelage infini. Symboliques aussi, si bien que la recherche en mathématique est une *Invitation au voyage* que

m'évoquent les deux derniers vers du poème éponyme de Charles Baudelaire *Les Fleurs du Mal* (1857) :

*Là, tout n'est qu'ordre et beauté,  
Calme, luxe et volupté.*

Je suis d'ailleurs infiniment reconnaissante à mon ami et collègue Nicolas Marie d'avoir été le véritable instigateur de ce projet (de voyage). Depuis l'idée de thèse que nos discussions a fait germer et tu t'es employé à semer en organisant la rencontre avec Laurent, jusqu'à ces invitations à participer à quelques de tes activités (RMR en 2018, cours dans la majeure finance etc.) et cette collaboration que tu me proposes aujourd'hui, un immense merci Nicolas !

Ma seule sortie mathématique hors du périphérique m'a conduite à Toulouse. En dépit d'une *information* pas encore tout à fait *mutuelle* et de quelques problèmes *logistiques* persistants et prometteurs de biens des tableaux encore à noircir, mille mercis à Anthony Réveillac et Lorick Huang pour leur accueil incomparable et un séjour exquis à bien des égards !

Mener cette thèse en parallèle de mon activité salariée de responsable pédagogique et enseignante à l'ESME Sudria fut un défi ; je remercie chaleureusement Marc Sellam, PDG de IONIS Education Group et Véronique Bonnet, directrice générale déléguée de l'ESME Sudria, d'avoir accompagné ce projet par leur confiance et leur soutien financier. J'en profite pour adresser un clin d'oeil à mes collègues de la *Team Montparnasse*, à nos discussions enflammées en "salle de réu", nos apéros du QG d'origine de la BAC à Falguière et qui se déclinent temporairement en virtuel ! Un grand merci également à son *coach* Hervé Coum, directeur des études du cycle prépa, qui par un cocktail détonant de confiance, de bienveillance, d'écoute attentive et de questions subtilement posées donne à chacun les clefs pour qu'il dessine ses objectifs et y apporte ses propres solutions ; ceci en toute *agilité* et sans *mais*.

Anaïs Nin écrit dans son *Journal* (1947-1955) :

*Les moments de courage et de divination sont tous nés de la passion.  
Les déserts qui suivent, je ne m'y attarde pas.*

J'ai eu la chance de croiser le chemin d'enseignants qui au-delà de l'expertise disciplinaire et des qualités de pédagogues que je leur ai toujours prêtées, m'ont aidée à ne pas trop m'attarder dans le désert. Je pense en particulier à Claire Laurent, avec qui les moments partagés, depuis les discussions passionnantes d'après-cours de géopolitique jusqu'à une inoubliable visite au musée Granet, en passant par les soirées mémorables en terre *lettreuse* ont scellé une amitié fidèle et complice. À Guillermo Moreno-Socías et Otared Kavian pour leur soutien constant et aussi précieux. À Rachel Ollivier, à nos discussions inspirantes d'il y a quelques années et au souvenir indélébile qu'il en reste aujourd'hui. À Martin Andler et Mourad Besbes, qui ont accompagné mes années d'étudiante, et ont aujourd'hui délicatement substitué à leur casquette de mentor celle d'ami fidèle cher à mes yeux.

Je n'oublie pas ceux et celles qui ont accompagné ces années (et celles d'avant!) par leur humour, sourires, diatribes et affection, et avec qui j'ai eu la joie de partager : soirées au Parc des Princes ou à l'Opéra, balades en forêt de Chantilly ou dans la cache à Luce, verres de Graves ou de Coup Franc, expositions, dîners confidentiels ou après-midis à refaire le monde,

soirées chez Paul, Sofia à la Belle Époque ou au Circus, signatures, théâtres ou (inclusif voire à l'intersection de) cinémas. . . Sur une frise presque-chronologique prennent place : Yvonne et Jacques Desvergez, Théo, Mathilde et Wallerand, la famille Laurent, Damien, Laure et Xavier, Florence, Lee, Camille, *Camiche* et Ming, Ingrid, Jean, Carole, la famille Larebière-Ducasse, Marie-Pascale, Julie et Fred, Adrien, la famille Bourgeois, Aurore, Jennifer, François, Bénédicte.

Les derniers remerciements sont destinés à ma famille pour son irremplaçable présence. Merci à mes oncle et tantes Marie-Pascale, Nicolas et Isabelle pour leurs encouragements et soutien toujours affectueux. Pensée toute particulière pour mon oncle Bernard, le premier à m'avoir parlé de Paul Malliavin lors d'une de nos discussions hebdomadaires ! Un grand merci enfin. . . À Antonin pour la (presque familiale) *entité* encadrant notre amitié, dont on trouve au dénominateur : Rachmaninoff et Dream Koala, matrices de Pauli et paire duale (et une topologie discutable), joutes verbales au bureau, au Zango ou sur Teams/Whatsapp/Outlook, à Tignes et au Revest, conférences du samedi et cinés du dimanche. . . et dont il m'est donné de mesurer quotidiennement (mais pas quantiquement) l'inestimable richesse ! À ma *Doc-Soeur* Stéphanie, dont j'admire les compétences professionnelles, suis une inconditionnelle des inégalables réparties, et apprécie le tempérament *complémentaire*, *partition* de nos jeux d'hier, de nos discussions d'aujourd'hui et de notre complicité de toujours. À mes parents pour leur soutien indéfectible, leur affection, tous les délicieux moments en famille dont ils ont été les instigateurs, et - si je ne pourrai jamais leur rendre - pour ce qu'ils m'ont donné : *presque sûrement* tout.





# Introduction (in English)

## Malliavin's tree

Malliavin calculus is at the heart of our work. Or rather, it is the *trunk* of the tree by which we choose to draw the outlines of this thesis. The anatomy of Malliavin's tree is described in the following sections; the *ground* on which it flourishes is made of analysis and probability theories. Malliavin calculus takes its *roots* in processes characterized by (semi/normal) martingales, sample-paths and/or independence of increments properties. Its formalism sits on what embodies its *basement/cornerstone*: the integration by parts formula. The meeting with other *branches* of analysis or probability has produced much *fruit*: the applications of Malliavin calculus.

## Roots and basement: Malliavin's framework and integration by parts

Historically, its initial development was not really to provide an infinite-dimensional differential calculus on the Wiener space, but rather to elaborate a probabilistic toolbox from an application perspective. The first ones were to prove the ellipticity of Hörmander operators and to rule on the existence and regularity of the density function of random vectors. The criterions stated by Paul Malliavin (see [85]) in both cases were based on what would become the *cornerstone* of his eponymous calculus: the integration by parts formula. Many works leaning on the seminal paper of Paul Malliavin were led around the hypoellipticity of second order degenerate elliptic differential operators; see for instance that of Jean-Michel Bismut [19], Hiroshi Kunita [77], Daniel W. Stroock [133] or Shinzo Watanabe [138]. In the same time, other probabilists focused on another perspective that it suggested: the possibility to formalize a *differentiation* on the Wiener space  $\mathscr{W}$  and to connect it with the preexisting integration notion.

Two main approaches lead to Malliavin calculus on the classical Wiener space supported by the Banach space  $C^0(\mathbf{T}; \mathbf{R})$  (where  $\mathbf{T} = [0, T]$  or  $\mathbf{T} = \mathbf{R}_+$ ) and equipped with the Wiener measure  $\mathbf{P}$ .

A possible *variational approach* consists in coupling the Wiener space with a *separable Hilbert space*  $\mathfrak{H} = L^2(\mathbf{T}, \mathcal{B}, \lambda)$ , where  $\lambda$  is the Lebesgue measure on  $\mathbf{T}$ . As explained in the reference books of David Nualart [98] or Svante Janson [71], there exists thus a *Gaussian random measure*  $\{\mathbf{W}(A), A \in \mathcal{B}, \lambda(A) < \infty\}$  independent on any family of disjoint subsets of  $\mathbf{T}$ , such that  $\mathbf{W}(A)$  has variance  $\lambda(A)$ , and that the paths of the Brownian motion can be reconstructed via the Centsov representation:

$$B_t = \mathbf{W}([0, t]) = \mathbf{W}(\mathbf{1}_{[0, t]}) ; t \in \mathbf{T}.$$

In fact,  $W(h)$  coincides with the Wiener integral of the function  $h \in \mathfrak{H}$  with respect to  $B$ . The Wiener space is thus entirely characterized by the so-called associated *isonormal Gaussian process*  $\{W(h), h \in \mathfrak{H}\}$ . Even if the Wiener integral naturally connects the Wiener space and a notion of integration, the challenge remains to construct a differentiation operator applicable to a wide class of Wiener functionals. A definition as a Fréchet derivative, if it seems natural at first, is not relevant: usual Wiener functionals such as Wiener integrals or solutions of stochastic differential equations with smooth coefficients are not even continuous with respect to the norm of the Wiener space. The suitable definition of a directional derivative on  $L^p(\mathbf{P})$  derives from the Cameron-Martin theorem that indicates - implicitly - that it makes sense only if the increases generating the derivative are performed in the directions of the eponymous vector space. The derivative operator  $D$  is thus defined on the space  $\mathcal{S}$  of *cylindrical random variables* composed of smooth random variables  $F$  of the form

$$F = f(W(h_1), \dots, W(h_n)); h_1, \dots, h_n \in \mathfrak{H}, \quad (0.0.1)$$

where  $f$  belongs to the Schwartz space  $\mathcal{S}(\mathbf{R}^n)$ . It is then extended to its domain  $\mathbf{D}^{1,2}$ , by density of the class  $\mathcal{S}$  in the space of square integrable Wiener functionals. The adjoint of the operator  $D$ , called *divergence* and denoted by  $\delta$ , coincides with the Wiener (resp. Skorohod) integral when applied to adapted (resp. non-adapted) processes. The operators  $D$  and  $\delta$  are connected by the relationship:

$$\mathbf{E}[\langle DF, U \rangle_{\mathfrak{H}}] = \mathbf{E}[F \delta U]; F \in \mathbf{D}^{1,2}, U \in \text{Dom } \delta.$$

As detailed for instance in the books of Paul-André Meyer [89] or Nobuaki Obata [107], an alternative approach, lying on the *Wiener-Itô chaotic representation property*, can also lead to Malliavin calculus. It is based on the development of an infinite dimensional analogue of the Schwartz theory where the Gaussian measure  $\mu$  on the dual  $E^*$  of a nuclear space  $E = \mathcal{S}(\mathbf{R})$  would supersede the Lebesgue measure on  $\mathbf{R}^n$ . The space  $\mathfrak{E}$  (resp.  $\mathfrak{E}^*$ ) of *test* (resp. *generalized*) *white noise functionals* is constructed by a continuous and dense embedding in  $L^2(E^*, \mu)$  (resp. by duality):

$$\mathfrak{E} \subset L^2(E^*, \mu) \subset \mathfrak{E}^*.$$

The space  $L^2(E^*, \mu)$  is in fact canonically isomorphic to the *Fock space* over  $L^2(\mathbf{R})$  through the Wiener-Itô-Sega isomorphism. Thus, any square integrable Wiener functionals admits a unique expansion in terms of multiple Wiener integrals. The so-called *chaotic decomposition* characterizes this alternative approach. In the general context of a Fock space the operators  $D$  and  $\delta$  coincide respectively with the *annihilation operator* and the *creation operator* met in quantum probability and satisfy a generalized canonical commutation relation. The equivalence of the approaches lies on the intrinsic relation between the standard Brownian motion and Hermite polynomials. Paul-André Meyer stated in [88] the equivalence of the norms associated to the different Sobolev spaces thus defined.

Even if the Malliavin calculus was associated with the Wiener space in a first place, it has since been extended to other classes of processes; to name just a few, Gaussian processes in general (see Nualart [98], Nourdin and Peccati [94]), Poisson processes (see Bichteler et al. [18] for a variational approach, Nualart and Vives [103] or Privault [116] for a chaotic approach), Lévy processes (see Nualart and Schoutens [102]) and Rademacher processes (see Privault [117]). The terminology of classical differential calculus in Banach spaces can be declined in the

paradigm of Malliavin calculus: the paths of the canonical process replace vectors, the functionals operating on the space of the paths take the place of functions, while the two key operators are called by the same names of *gradient* and *divergence*. The fundamental relationship between these latter appears as the **basement** of this calculus and conceptualizes the notion of *integration by parts in Malliavin's sense*:

$$\mathbf{E}[\langle DF, U \rangle_{\mathfrak{H}}] = \mathbf{E}[F \delta U], \quad (0.0.2)$$

where  $(\Omega, \mathcal{A}, \mathbf{P})$  is the canonical probability space and  $\mathfrak{H}$  is a generic Hilbert space seen as the *perturbations space* when  $DF$  is interpreted as a directional derivative (for instance  $\mathfrak{H}$  is the Cameron-Martin's space for Brownian motion).

During the preparation of this thesis work, we found two frameworks equipped with a Malliavin structure and of particular interest: the Poisson space and the Rademacher space. Their Malliavin formalism will be described and compared to our contributions later.

## Branches and fruits: applications of Mallavin calculus

The use of Malliavin's powerful toolbox in other **branches** of analysis or probability has been particularly **fruitful** for years.

### Partial/Stochastic differential equations and anticipative calculus

Having in mind the results of the seminal work of Paul Malliavin many authors re-used the non-degeneracy condition to study the regularity of solutions of partial differential equations and stochastic differential equations. This led to many applications reviewed for instance by Martha Sanz-Solé in [128] (resp. by Hu, Huang, Lê, Nualart and Tindel [65]) for partial differential equations driven by coloured noises (resp. by a white noise in time/rough noise in space with Hurst index  $H \in (1/4, 1/2)$ ). We can cite the works of Vlad Bally and Denis Talay [13], Yaozhong Hu, David Nualart and Jian Song [66], or Shigeo Kusuoka and Daniel W. Stroock [78] for stochastic differential equations driven by a Brownian motion and that of David Nualart and Bruno Sausserau [101] for SDEs driven by a fractional Brownian motion with Hurst index  $H \in (1/2, 1)$ .

Malliavin calculus results were also exploited to tackle with non-causal stochastic differential equations formulated in terms of anticipating stochastic integrals. Indeed the divergence operator allows to define an extension of Itô's stochastic integral to anticipating integrands in Brownian motion case (see Skorohod [131]). An anticipating formula is presented by David Nualart and Pardoux in [99] or in its earlier (and very unknown) version by Masayuki Hitsuda in [64]. Elisa Alòs and David Nualart improved it in [4]. An *anticipative calculus* was thus first designed through this approach for Brownian motion (see Berger [17] or the chapter 3 of Nualart [98]) and has since been extended to Poisson processes (see Decreusefond and Savy [40], Nualart and Vives [103]) and Lévy processes (see [3],[105]). Elisa Alòs, Olivier Mazet and David Nualart used also Malliavin calculus (in particular the divergence) to give a sense to stochastic integrals with respect to the fractional Brownian motion with Hurst index  $H \in (0, 1/2)$  and then to design a stochastic calculus in that framework.

## Statistics

A wide range of Malliavin tools were also used for statistical purposes. We give a quick overview through some examples. For instance, Ciprian A. Tudor and Frederi G. Viens [135] used the *chaotic decomposition* to provide an estimator of the self-similarity parameter  $H$  of the Rosenblatt process. Fabienne Comte and Nicolas Marie [29] showed that Skorohod's integral (defined from the *divergence*) was the only suitable extension of Itô's integral to study nonparametric estimation in SDEs driven by a fractional Brownian motion of Hurst index  $H \in (1/2, 1)$ . José M. Corcuera and Arturo Kohatsu-Higa [31] wielded Malliavin *integration by parts* to study asymptotic inference of stochastic (jump) processes. Nicolas Privault and Anthony Réveillac also used it with another aim: to provide estimators of Stein-type estimators of the drift of some Gaussian processes [118].

## Finance

A wide range of the *fruits* born of the intersection of Malliavin calculus and *finance* can be found in the monograph of Paul Malliavin and Anton Thalmaier (see [86]). Among them, three have sparked special interest throughout our work: the calculation of Greeks, portfolios hedging and insider's trading.

In order to frame our point in the vast universe of financial mathematics, we restrict first our frame to a simple financial market model composed of two assets: one risk-free asset  $A = (A_t)_{t \in \mathbf{T}}$  and one risky asset  $S = (S_t)_{t \in \mathbf{T}}$ . We assume that  $A$  is defined by an interest rate and that  $S$  satisfies a stochastic differential equation (in the Black-Scholes model for instance) or an equation in differences (in discrete models) on the trading interval  $\mathbf{T}$ . The *derivatives*, namely the financial products which value derives from the performance of  $S$ , we are focused on are called *vanilla options*. An option is a contract which gives the buyer (the owner or holder of the option) the right, but not the obligation, to buy or sell the risky asset at a specified strike price  $\mathcal{K}$  and at a specified date  $T$ . The value of an option at expiry, i.e. what the holder will receive, is called *payoff*; this is given for a vanilla option by a random variable  $\Phi(S_T)$ . On this market act buyers and sellers who are considered as *regular agents* (resp. *insiders*) if decisions as to the composition of their portfolio (with both types of assets at their disposal) are made based on public information (resp. with extra information).

The *Greeks*, so-called in reminiscence of the Greek letters by which they are denoted, designate the sensitivities of the price of an option with respect to a change in the underlying parameters on which its value depends. They are of capital use for the trader giving him the power to control his risk exposure. For instance, Delta and Rho are the first-order Greeks measuring respectively the sensitivity to the underlying asset price and to the interest rate whereas Gamma designates the second-order Greek quantifying the rate of change in the Delta with respect to changes in the underlying price. For a long time, these computations were the only prerogative of analysts who were treating them with finite difference methods in a Monte-Carlo frame. Without precisely introducing the formalism (that will be done in chapter 4) consider an option of payoff  $\mathcal{K} = \Phi(S_T)$  where  $(S_t)_{t \in \mathbf{T}}$  is the price of a risky asset defined by the Black-Scholes model:

$$S_t = 1 + \int_0^t r S_u du + \int_0^t \sigma S_u dB_u, \quad t \in \mathbf{T}, \quad (0.0.3)$$

where the parameters  $r$  and  $\sigma$  stand respectively for the risk-free interest rate and the volatility of returns of  $S$ . The value of the option is given by  $V_t$  at time  $t \in \mathbf{T}$ . Delta can be expressed

(in the Black-Scholes model) as

$$\Delta = \frac{\partial V_0}{\partial S_0} = e^{-rT} \mathbf{E}_{\mathbf{Q}} \left[ \Phi'(S_T) \frac{\partial S_T}{\partial S_0} \right], \quad (0.0.4)$$

where  $\mathbf{Q}$  denotes the risk-neutral probability measure. Generally, the function  $\Phi$  is not derivable; for instance the payoff of a european call is  $\Phi(S_T) = (S_T - \mathcal{K})_+$ . Therefore, it would be wise to write  $\Delta = e^{-rT} \mathbf{E}_{\mathbf{Q}} [\Phi(S_T) \pi]$ , where the *weight*  $\pi$  does not depend of the payoff. The basic idea introduced first by Éric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, Pierre-Louis Lions and Nizar Touzi in [54] and [55] making use of the integration by parts and the chain rule formulas leads to

$$\Delta = e^{-rT} \mathbf{E}_{\mathbf{Q}} [\Phi(S_T) \delta(u)],$$

so that the weight can be expressed as the Skorohod integral of a generator  $u$ , here deterministic and constant equal to  $(\sigma T)^{-1}$ . Éric Benamou gave in [16] necessary and sufficient conditions on  $u$  so that it generates a weight for the Greeks simulation. The pioneer work [54] was restated and transposed to financial jump models driven by Poisson processes by Nicolas Privault et al. (see [50],[121]) and by Lévy processes by (in particular) Marie-Pierre Bavouzet-Morel and Marouen Messaoud in [15].

Strongly connected to the semi-martingales theory, the valuation of options and in particular hedging problems have also benefited from the contributions of Malliavin calculus. Hedging is crucial since it gives an understanding of how sellers or buyers can manage dynamically and compose their portfolio to replicate the payoff of the option, i.e. to attain its value at maturity. Mathematically, this boils down to determine a  $\mathbf{R}^2$ -valued process  $\psi$  satisfying the self-financing condition (see chapter 4 for its expression in the Black-Scholes and the Cox-Ross-Rubinstein models) such that

$$V_T(\psi) = \Phi(S_T).$$

In discrete time models, the additional condition

$$V_t(\psi) \geq 0, ; \forall t \in \mathbf{N}_T, \mathbf{P} - \text{a.s.},$$

is required. In complete markets, where all claims are reachable, and in which the underlying model is equipped with a Malliavin structure, the process  $\psi$  can be elegantly described in terms of Malliavin derivative; this was suggested by the contemporary works (both published in 1991) of Ioannis Karatzas and Daniel Ocone [73] on the one hand and with Jinlu Li [108] on the other hand. The so-called Karatzas-Ocone hedging formula is actually derived from the Clark-Ocone one. Analogue formulas were provided for Poisson processes by Günter Last and Matthew Penrose [81] and for Lévy processes by Giulia Di Nunno [104].

Insider trading is a related topic to portfolio optimization. The frame is that of a financial market where trade two agents with different levels of information: an ordinary agent whose decisions are made in light of public flow and an insider enjoying an additional information from the start of the trading period. Two questions arise: how to quantify insider's additional expected utility? Does the additional information produce an arbitrage? This latter can be traduced by an enrichment of the filtration on which the insider can base his portfolio decisions and leads then to the theory of enlargement of filtrations. Following the pioneer work of Igor Pikovsky and Ioannis Karatzas [115], Jürgen Amendinger et al. [7],[6],[5], Axel Grorud and

Monique Pontier [61] precised criterions for optimization and computed the additional utility of the insider in their respective works. Peter Imkeller connected it to Malliavin calculus in [68] by expressing the *information drift* as the logarithmic Malliavin trace of a conditional density characterizing insider's advantage. Jorge A. León, Reyla Navarro and David Nualart use in [83] some techniques of Malliavin calculus to analyze the properties the *forward integral* (introduced by Russo and Vallois in [127]) and to maximize the expected logarithmic utility of the insider.

In some extensions of the Black-Scholes model allowing the volatility to be a stochastic process (see for instance Hull and White[67], or Heston [63]), in particular in some stochastic volatility diffusion models, where the volatility also follows a diffusion process, it can be observed that the *implied volatility* (forward-looking measure to estimate the future fluctuations of the underlying asset) behaves *roughly* at any reasonable timescale. Malliavin's toolbox whose effectiveness in anticipative calculus was mentioned, is also efficient to analyze future volatilities, which are non-adapted processes. In [2], Elisa Alòs, Jorge León and Josep Vives, on a seminal work on the nowadays named *rough volatilities*, give an expression for the short-time behavior of the implied volatility in a jump-diffusion models in terms of the Malliavin derivative of it.

### Stein's method

In a seminal paper ([132]), Charles Stein described a process to quantify the errors in the Normal approximation by sums of random variables having a stationary dependence structure. By the same, he laid the foundations of a very powerful method whose perspectives of application very quickly exceeded the framework of its birth; *Stein's method* notably stood out as one (not to say the one) efficient way to compute distance between two probability measures of the form

$$\text{dist}_{\mathcal{H}}(\mathbf{P}^*, \mathbf{Q}) = \sup_{h \in \mathcal{H}} \left| \int_{\mathfrak{F}} h d\mathbf{P}^* - \int_{\mathfrak{F}} h d\mathbf{Q} \right|,$$

where  $\mathbf{P}^*$  stands for the target measure (to approximate) and  $\mathbf{Q}$  the initial measure, both defined on the same metric space  $(\mathfrak{F}, c)$  and  $\mathcal{H}$  is a set of test functions. Even if Stein's method keeps a special bond with the Normal approximation for which it was born (refer to the reference book of Louis H.Y. Chen, Larry Goldstein and Qi-Man Shao [27]), its principles could be transposed to other target distributions. It first gave birth to Stein-Chen method for Poisson convergence (see Chen [26]). Many approximation criterions of more general random variables by varied distributions followed: Gamma (Arras and Swan [8], Döbler and Peccati [44], Peköz, Röllin and Ross [114]), Exponential (Chatterjee, Fulman and Röllin [25], Peköz and Röllin [112]), Geometric (Peköz, Röllin and Ross [113])... Besides, in the Section *Taxonomy* of the webpage [Stein's method](#) he has dedicated to it, Yvik Swan reviews all distributions which approximation can be quantified through this method. Many surveys were published to teach the proselyte and initiate the layman with the effectiveness of the concept. If it remains impossible to give an exhaustive list, we recommend the reference and pedagogical works of Benjamin Arras and Yvik Swan, [8], Andrew D. Barbour [14] or Nathan Ross [125].

Stein's method seems to be split in two stages; the first one consists in converting the problem of bounding the error in the approximation of the measure  $\mathbf{P}^*$  by  $\mathbf{Q}$  into a problem of bounding an expression of the form

$$\mathbf{E}_{\mathbf{Q}} [L\varphi(X)] = \mathbf{E} [L_1\varphi(X)] + \mathbf{E} [L_2\varphi(X)], \quad (0.0.5)$$

where  $L$  (resp. the class of  $\varphi$ ) is determined to characterize the target measure  $\mathbf{P}^*$  (resp.  $\mathbf{P}^*$  and  $\mathcal{H}$ ), and  $X$  is a random variable of law  $\mathbf{Q}$ . The second component of Stein's method gathers techniques to bound (0.0.13); by taking into account how  $X$  is defined this consists into transforming  $L_1\varphi(X)$  into  $-L_2\varphi(X) + \text{remainder}$ . This remainder is what gives the bound of the distance and in a problem of convergence, provides its rate. To make the transformation of  $L_1\varphi(X)$ , several approaches appeared along the years. One of the most popular approach (see for instance Barbour and Chen [14]) is to use exchangeable pairs. Alternative to it, are the size-biased (see Chen, Goldstein and Shao [27]) or zero biased (see Goldstein and Reinert [58]) couplings, which again conveniently transform  $L_1\varphi$ . In a path-breaking work, Ivan Nourdin and Giovanni Peccati (see [93],[94]) showed that the transformation step can be advantageously made simple using integration by parts in the sense of Malliavin calculus, and by the same gave an intersection to the two theories. This approach is efficient provided there exists a Malliavin gradient on the space on which  $X$  is defined. It was in particular applied to functionals of Rademacher (see Nourdin, Peccati and Reinert [96], Zheng [143], Poisson (see for instance Decreusefond, Schulte and Thäle [41], Lachièze-Rey and Peccati [79], Peccati, Solé, Taqqu and Utzet [111]) or Gaussian random variables (see Nualart and Peccati [100]) or processes (see Coutin and Decreusefond [32, 33]). A review of the works resulting from the Malliavin-Stein approach can be retrieved in the webpage designed by Ivan Nourdin.

### Dirichlet forms

Parallel to its connection with Stein's method, Malliavin calculus meets profitably Dirichlet forms in the frame of error calculus (see Bouleau and Hirsch [23]). The related topics have generated considerable interest (among the analyst community before probabilists take it) since the beginning of XX-th century. The use of Dirichlet forms is particularly efficient to perform those computations when errors are supposed to be infinitesimal and probabilistic. If  $Y$  stands for a scalar erroneous quantity, the error propagates by the function  $f$  according to the formulas:

$$\begin{cases} \text{bias of error on } f(Y) &= (\text{bias of error on } Y) f'(Y) + \frac{1}{2}(\text{var of error on } Y) f''(Y) & (E_1) \\ \text{var of error on } f(Y) &= (\text{var of error on } Y) f'^2(Y) & (E_2) \end{cases}$$

where, if the variance is of the same order of magnitude as the bias or if the bias is negligible with respect to the variance, the calculus has to be

- for the bias of error: a second order differential equation with variance and bias ( $E_1$ ),
- for the variance of error: a first order differential equation for the calculus of the variance which does not involve the bias ( $E_2$ ).

By considering up to now a quantity  $F = F(X_A)$  ( $A \subset \mathbf{N}^*$ ) function of the erroneous quantities  $\{X_k, k \in A\}$  (supposed to be known, small and independent) the potential quadratic error to expect on  $F$  denoted  $\sigma_F^2$  can be written as:

$$\sigma_F^2 = L(F^2) - 2FL(F), \quad (0.0.6)$$

where  $L$  is the differential operator

$$L = \frac{1}{2} \sum_{k \in A} \sigma_k^2 \frac{\partial^2}{\partial X_k^2}$$



and  $\{\sigma_k^2, k \in A\}$  is the family of quadratic errors on  $\{X_k, k \in A\}$ . Actually, the quadratic error can be expressed as

$$\sigma_F^2 = \Gamma(F) = \Gamma(F, F),$$

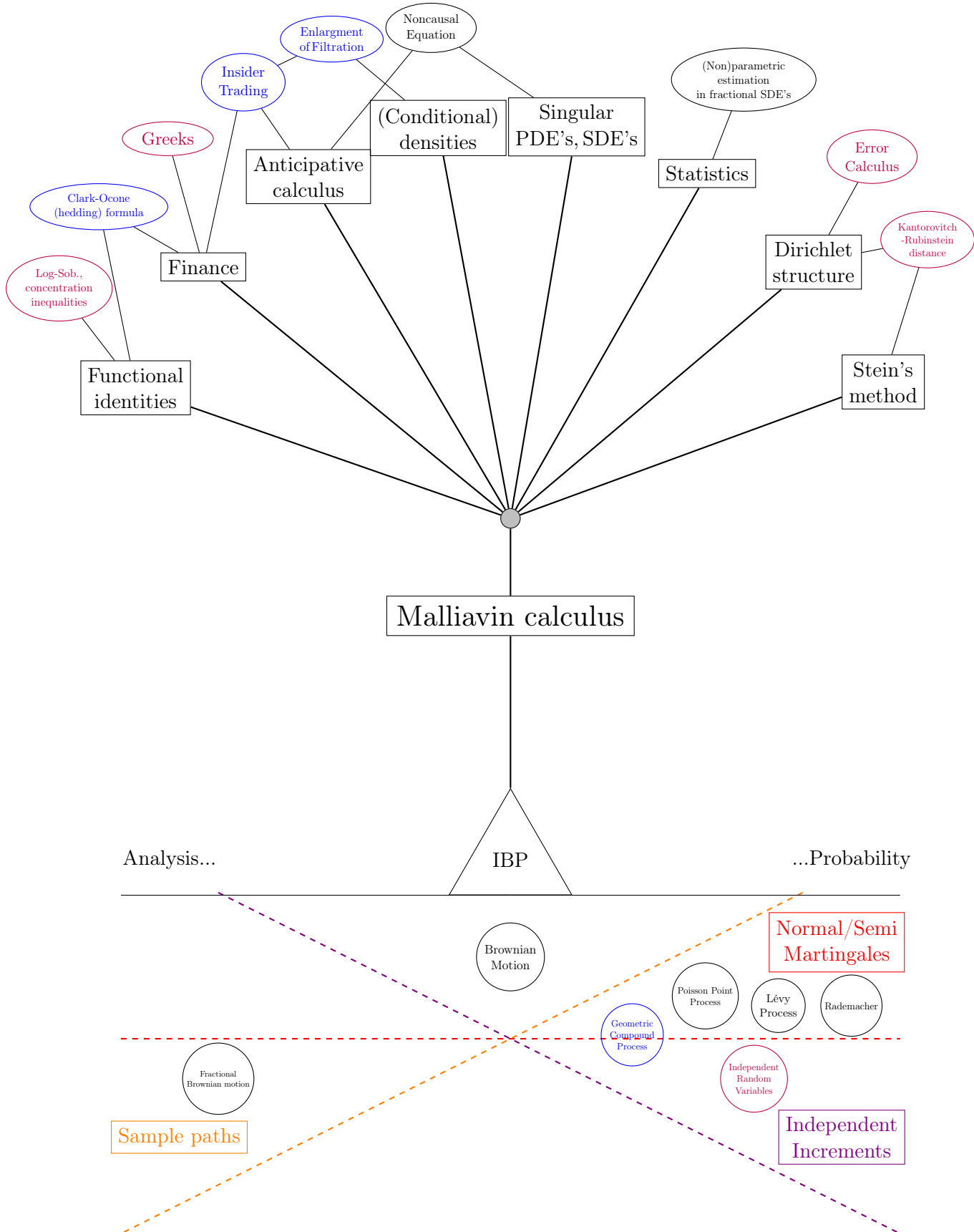
where  $\Gamma$  is the *carré du champ operator* defined by

$$\Gamma(F, G) = \frac{1}{2} \left( L(FG) - GLF - FLG \right). \quad (0.0.7)$$

The use of these tools makes most sense when tackling with the estimation of the convergence rate of a sequence  $(X_n)_{n \in \mathbf{N}^*}$  towards a random variable  $X$ , all assumed to be defined on a probability space  $(\Omega, \mathcal{A}, \mathbf{P})$ . Indeed, if the form  $\Gamma$  is closed, the error calculus theory deploys in this context by lying on the following principle: if the sequence of pairs  $(X_n, \text{error on } X_n)$  converges suitably, it converges necessarily to a pair  $(X, \text{error on } X)$ . When the underlying space is provided with a Malliavin structure the operator  $L$  appearing in (0.0.15) coincides with the Laplacian/number operator  $L = -\delta D$ . This key fact will be explained in Section 1.3.2 and illustrated in the case of the Wiener space in Example 1.3.6. This suggests that Dirichlet forms implicitly operate in the Stein-Malliavin combination to lead to a Stein-Dirichlet-Malliavin method (see Decreasefond [36]). Besides, Ehsan Azmoodeh, Simon Campese and Guillaume Poly exploited the connection between the three theories to revisit the fourth moment theorem in [11].

We conclude this state of the art by drawing Malliavin's tree.

Malliavin's tree



## Contributions

By grabbing the tree metaphor again, we could sum up the contributions of this thesis to the addition of two roots: the construction of a Malliavin calculus for **independent random variables** on the one hand, and for **compound geometric processes** on the other one. The fruits resulting from these two formalisms are displayed in colour in the above figure: in **purple** for those relating to independent random variables framework, in **blue** for those relating to compound geometric processes framework.

### Malliavin calculus and Dirichlet structures for independent random variables

The motivation to develop a **Malliavin calculus for independent random variables** was twofold. As mentioned above, after some years of development, the Malliavin calculus has reached a certain maturity. The most complete and fruitful theories are for Wiener (see for instance Nualart [98]) and Poisson point processes (see for instance Albeverio, Kondratiev and Röckner [1], Privault [117]). The only Malliavin's foray in discrete settings had been so far for Rademacher processes (see Privault [117]); our initial aim was to generalize it. Malliavin calculus can be constructed in the three aforementioned frameworks (Gaussian, Poisson, Rademacher) via a chaotic approach. In those, the existence of normal martingales as solutions of structure equations (see Émery [51], Privault, Solé and Vives [120] for the continuous time, Privault [117] for Rademacher processes) entails directly a chaotic decomposition by induction of the martingale representation property (see Émery [51]). The existence of structure equations is not necessary to state a chaotic decomposition (see for instance Dasgupt and Kallianpur [35], Eddhabi and Vives [49] for chaotic decomposition and Clark formula for the fractional Brownian motion with index  $H \in (1/2, 1)$ ); but it considerably facilitates it. The independence and stationarity of increments seems to play a major role in the resolution of the so-called structure equations and thus in the effectiveness of the concepts. This motivated to investigate and explore what we can inherit in the simplest situation of all with independence (without requiring stationarity): that of a family of independent, non necessarily identically distributed, random variables. The closest situation to our investigations is that of the Rademacher space, namely  $\{-1, 1\}^{\mathbf{N}}$ , equipped with the product probability  $\otimes_{k \in \mathbf{N}} \mu_k$  where  $\mu_k$  is a Bernoulli probability on  $\{-1, 1\}$ . The chaotic approach that leads to a Malliavin calculus on it is intrinsically linked to the existence of a normal martingale or *novation* (in Emery's terminology [52]) coupled with the canonical process  $X$ . The existence of a normal martingale in a discrete setting also requires that the law of the random variables be only supported by two points (see Émery [51],[52], Privault and Schoutens [119]). This is also reflected in the very definition of the gradient (see Nourdin, Peccati and Reinert [96] or Privault [117]), usually defined as

$$\begin{aligned} \widehat{D}_k F(X_1, \dots, X_n) &= \mathbf{E}[X_k F(X_1, \dots, X_n) | X_\ell, \ell \neq k] \\ &= \mathbf{P}(X_k = 1) F(X_1, \dots, \underbrace{+1}_{k\text{-th}}, \dots, X_n) - \mathbf{P}(X_k = -1) F(X_1, \dots, \underbrace{-1}_{k\text{-th}}, \dots, X_n), \end{aligned} \quad (0.0.8)$$

and that implies to be meaningful, either that the random variables are real valued or that they only have two possible outcomes. Our initial aim was to provide any countable product of probability spaces with a Malliavin structure; it must be clear that in the present work all the random variables may leave on different spaces, which are then only supposed to be Polish

spaces. That means that in the definition of the gradient, we can not use any algebraic property of the underlying spaces. Although some of our applications do concern random variables with finite number of outcomes, it did not seem straightforward to devise what should be the weights, replacing  $\mathbf{P}(X_k = 1)$  and  $-\mathbf{P}(X_k = -1)$ . In that respect, we introduce a discrete gradient  $D$  as a difference operator which can be interpreted as the measure of the "influence" of the  $k$ -th component of the process  $X$  on  $F$ . The definition of the divergence operator  $\delta$ , as the adjoint of  $D$ , satisfies a discrete version of integration by parts formula (1.3.1). Last the introduction of the operator number  $L = -\delta D$ , also generator of a Markovian semi-group  $(P_t)_{t \in \mathbf{R}_+}$ , completes the definition of the modicum of Malliavin operators. We bring kind of legitimacy to our formalism from several perspectives. We connect our formalism to its elder, the Rademacher framework; not with the gradient but with the number operator. Indeed, many applications, notably those revolving around functional identities, rely not directly on the gradient  $D$  but rather on the operator number  $L$ . It turns out that for the Rademacher space, the operators  $\hat{L} = -\hat{\delta}\hat{D}$  defined according to (0.0.8) and  $L$  defined in our construction do coincide. Besides, we attach our construction to preexisting continuous time theories. Actually, we bring to light that the usual Poisson and Brownian Dirichlet structures associated to their respective gradient, could be retrieved as limits of the structures induced by our formalism. We obtained versions of almost all the classical functional inequalities in discrete settings, and revisit some of them to which we give a new point of view. To mention just a few, we show that the Efron-Stein inequality can be interpreted as a Poincaré inequality or that the Hoeffding decomposition of  $U$ -statistics can be viewed as an avatar of the Clark representation formula. We transpose the classical covariance identity, log-Sobolev inequality, and deviation inequality in our framework. Then our formalism finds its place within Malliavin's landscape; it fully generalizes what is known about Rademacher spaces, and connects with Brownian and Poisson frameworks with both limit procedure and similar functional identities.

The second main motivation of our first work came from a branch of the tree with which such formalism would be likely to combine advantageously: Stein's method. We provide here a "Stein-Malliavin criterion" analogue to the eponymous results in the Gaussian and Poisson cases (see Nourdin and Peccati [93], Peccati, Solé, Taqqu and Utzet [111]), or in the Rademacher space (see Nourdin, Peccati and Reinert [96]) giving a "Berry-Esseen" general bound for the Kolmogorov-Rubinstein distance between the law of a functional of independent random variables (without any additional hypothesis on the initial laws) and the Normal distribution, and this in terms of Malliavin operators. We establish a similar criterion for the Gamma approximation. Applied to the particular case of degenerate  $U$ -statistics of order two, it allows to provide a de Jong type estimate in the Gamma approximation of those  $U$ -statistics (see Döbler and Peccati [48]). To the best of our knowledge, there does not yet exist a Stein criterion for Gaussian or Gamma approximation which does not rely on exchangeable pairs or any other sort of coupling. Our results appear as a natural continuation of those stated by Ivan Nourdin, Giovanni Peccati and Gesine Reinert [96] for Rademacher functionals; we generalize thus the use of the Nourdin-Peccati approach (Nourdin and Peccati [94]) to general discrete settings.

For all these reasons, it seems that our Dirichlet-Malliavin structure gives a unified framework for many results scattered in the literature so that we hope to bring new insights on why these apparently disjoint results (Efron-Stein, exchangeable pairs, etc.) are in fact multiple sides of the same coin... with the "effigy" of the integration by parts formula.

## Malliavin calculus for compound geometric processes and insider trading in a ternary model

The initial aim of the second work was not to design a stochastic calculus for [geometric compound processes](#) but rather to use our discrete Malliavin calculus to address insider's trading issues. This is also the reason why the title of the thesis does not mention it. Indeed, we first tried to deploy our artillery to hit an identified target: the computation of the additional logarithmic expected utility of an insider in the trinomial model. This problem is of particular interest in several respects; first, and as mentioned above, financial mathematics is the flagship scope of Malliavin calculus. If option hedging has already been studied in the Cox-Ross-Rubinstein model (see chapter 1 in Privault [117]) using the Malliavin equipment in the Rademacher space, no similar project has been carried out so far in the framework of multinomial trees (with at least three branches at each stage). Moreover, while many works deal with enlargement of filtrations and insider issues in continuous time (see the works of Amendinger, Imkeller and Schweizer ([7]), very few were carried on discrete settings. Some results already exist for the enlargement of filtrations in discrete time. Catherine Blanchet-Scalliet, Monique Jeanblanc and Roméo Romo Romero showed in [20] that the results known in continuous time extend immediately in a discrete time setting. Most of them are easily obtained as a result of Doob's decomposition. In this respect, we wanted to discuss the possibility to interpret the increasing predictable process of Doob's decomposition as the *information drift* and to express it in terms of Malliavin derivative, as done by Peter Imkeller in continuous time [69].

The results of our investigations follow; while our formalism lends itself very well to the computation of Greeks in the trinomial model, it turns out to be impossible to state a Karatzas-Ocone hedging formula in this context. This is a direct consequence of the absence of a martingale representation theorem which requires that the law of the random variables be only supported by two points (see Émery [52]). Without completely losing sight of our original goal, we chose thus to substitute what we have called *a ternary model* to the trinomial one. Equivalent in law to the initial one, it is supported by a compound geometric process. This is based on a jump structure as the compound Poisson process, and thereby inherits number of properties from this latter. In particular, the definition of the measurable space  $(\Omega, \mathcal{A})$  on which it is defined is a crucial point to characterize the compound geometric process. Following the frame designed by Laurent Decreusefond and Nicolas Savy for filtered Poisson processes [40], we describe  $\Omega$  as the set composed of the couple  $(t, k)$  where  $t$  stands for the jump time and  $k$  the height of the corresponding jump. This definition plays a major role in the effectiveness of the concepts; the so-called *ternary model* lying on this new structure, yet equivalent in law to the trinomial model, best suits for hedging problems. As expected, it remains impossible to state a chaotic decomposition from a normal martingale (for the reasons evoked above); nevertheless this new paradigm enables to establish a *modified chaotic decomposition* in terms of multiple integrals with respect to a non-orthogonal family of random variables. The gradient is thus introduced as the annihilation operator acting on this pseudo-chaotic decomposition and such that the stochastic integral defined appears - up to a linear transformation - as the "inverse" operation of it and as its adjoint. These latter are linked via an integration by parts formula so that we equip compound geometric processes with a Malliavin calculus. Besides, the gradient coincides with a difference operator more prone to state functional identities from which derives the expected Clark-Ocone-Karatzas formula.

The second part of the work consists in making use of this new formalism to compute the additional utility of an insider in the new ternary model. The insider benefits from an extra information hidden in a  $\Gamma$ -valued random variable  $G$ . The insider's filtration  $\mathbb{G}$  does not coincide with the initial one, so that we need to use the techniques of enlargement of filtrations in a discrete setting. Using the toolbox of Catherine Blanchet-Scalliet, Monique Jeanblanc and Roméo Romo Romero [20], we define an analogue of the *information drift*. We express it by means of the Malliavin gradient (for compound geometric processes) applied to the *conditional density processes*  $p^c$  ( $c \in \Gamma$ ). Besides, we focus on a particular process  $1/p^G$  from which we define the *martingale preserving measure*  $(\mathbf{Q}_t)_{t \in \mathbf{T}}$ . Following the method of Freddy Delbaen et Walter Schachermayer [42], we compute and express the insider's optimal portfolio in terms of the  $\mathbb{G}$ -martingale  $(\mathbf{Q}_t)_{t \in \mathbf{T}}$ . We show that the insider's additional expected logarithmic utility can be expressed as the relative entropy of the initial measure  $\mathbf{P}$  with respect to  $\mathbf{Q}_t$ ; we retrieve exactly the result of Jürgen Amendinger, Peter Imkeller and Martin Schweizer [7] stated in the continuous case. We end the work by doing explicit computations in the specific case where the insider gets an extra information on the terminal value of the risky asset.

## Manuscript

The manuscript is organized as follows; it consists of two parts, each of which refers to one of the above-mentioned works. The first one, entitled *Malliavin and Dirichlet structures for independent random variables* is a collaboration with Laurent Decreusefond and was published in Stochastic Processes and their Applications in 2019 (see [38]). The second one will be submitted soon. Both parts are deployed according to the same scheme; they start with an introductory chapter where is giving the state of the art and presenting the formalism that need to be developed to respond to an identified problematic. The necessary tools are built in the chapters immediatly following them whereas the *fruits* are harvested in the last chapter of each part, as the response provided to the initial problem.

Before ending this introduction, let me mention that any new contribution (original definition or result) is flagged with a symbol ( $\star$ ). Despite my poor English (I apologize for that), I really enjoyed preparing this manuscript; I wish you a pleasant reading!



# Introduction (en français)

## L'arbre de Malliavin

Au coeur de notre travail, se situe *le calcul de Malliavin*. Ou plutôt, il constitue le *tronc* de l'arbre par lequel nous avons choisi de dessiner les contours de cette thèse. L'anatomie de cet arbre de Malliavin se décline au travers des sections suivantes ; la composition du *sol* sur lequel il s'érige a en héritage les théories analytique et probabiliste. Il prend ses *racines* dans les familles de processus caractérisés par des propriétés de (semi/normales) martingales, trajectorielles ou d'indépendance des incréments. Son formalisme repose sur un(e) *Pierre angulaire/socle* qu'incarne la formule d'intégration par parties. Enfin, sa rencontre avec d'autres *branches* de l'analyse ou des probabilités a donné de nombreux *fruits*, les applications du calcul de Malliavin.

## Racines et socle : le cadre du calcul de Malliavin et la formule d'intégration par parties

Historiquement, le développement du calcul de Malliavin était davantage motivé par la conception d'outils probabilistes à des fins applicatives que par l'élaboration - à proprement parler - d'un calcul différentiel en dimension infinie sur l'espace de Wiener. Dans cette perspective d'applications, les premiers résultats de Malliavin furent utilisés pour fournir une preuve probabiliste du théorème de Hörmander et statuer sur l'existence et la régularité de la densité de vecteurs aléatoires. Les critères établis par Paul Malliavin (voir [85]) dans le cadre de ces deux études reposaient sur ce qui deviendrait la pierre angulaire de sa théorie éponyme : la formule d'intégration par parties. Ce travail pionnier de Paul Malliavin inspira peu après de nombreux travaux autour de l'hypoellipticité d'opérateurs différentiels elliptiques dégénérés d'ordre 2 ; on peut notamment citer les travaux de Jean-Michel Bismut [19], Hiroshi Kunita [77], Daniel W. Stroock [133] or Shinzo Watanabe [138]. En parallèle, d'autres probabilistes s'intéressèrent alors à ce qu'impliquait notamment ce nouveau formalisme : la possibilité de construire une différentiation sur l'espace de Wiener  $\mathscr{W}$  et de la connecter avec la précédente notion d'intégration.

Deux approches principales menant au calcul de Malliavin classique sur l'espace de Wiener (supporté par l'espace de Banach  $\mathcal{C}^0(\mathbf{T}; \mathbf{R})$  avec  $\mathbf{T} = [0, T]$  ou  $\mathbf{T} = \mathbf{R}_+$ , et équipé de la mesure de Wiener  $\mathbf{P}$ ) peuvent être identifiées.

Une première approche, dite *variationnelle*, consiste à coupler l'espace de Wiener à un espace de Hilbert séparable du type  $\mathfrak{H} = L^2(\mathbf{T}, \mathcal{B}, \lambda)$ , où  $\lambda$  est la mesure de Lebesgue sur  $\mathbf{T}$ . Comme



expliqué dans les ouvrages de référence de David Nualart [98] ou Svante Janson [71], il existe alors *une mesure Gaussienne*  $\{\mathbf{W}(A), A \in \mathcal{B}, \lambda(A) < \infty\}$  indépendante sur toute famille de sous-ensembles disjoints de  $\mathbf{T}$ , telle que la variance de  $\mathbf{W}(A)$  est égale à  $\lambda(A)$  et que les trajectoires du mouvement brownien peuvent être reconstruites via la formule de représentation de Centsov :

$$B_t = \mathbf{W}([0, t]) = \mathbf{W}(\mathbf{1}_{[0, t]}) ; t \in \mathbf{T}.$$

En fait,  $W(h)$  coïncide avec l'intégrale de Wiener de la fonction  $h \in \mathfrak{H}$  relativement à  $B$ . L'espace de Wiener est donc entièrement caractérisé par le *processus isonormal Gaussien*  $\{W(h), h \in \mathfrak{H}\}$  qui peut lui être associé. Même si l'intégrale de Wiener lie naturellement l'espace de Wiener à une notion d'intégration dans le sens qui vient d'être défini, la possibilité de définir un opérateur de différentiation pour une large classe de fonctionnelles de Wiener est restée longtemps obscure. En effet, la définition en tant que dérivée de Fréchet, aussi naturelle qu'elle puisse paraître au premier abord, n'est pas pertinente : les fonctionnelles de Wiener usuelles comme les intégrales de Wiener ou les solutions d'équations différentielles stochastiques avec des coefficients réguliers ne sont même pas continues pour la norme de l'espace de Wiener. La définition appropriée d'une dérivée directionnelle dans  $L^p(\mathbf{P})$  provient du théorème de Cameron-Martin qui indique - implicitement - qu'elle n'a de sens que si les accroissements générant la dérivée sont réalisés dans les directions de l'espace vectoriel du même nom. L'opérateur  $D$  est alors défini sur l'espace  $\mathcal{S}$  variables aléatoires *cylindriques* de la forme

$$F = f(W(h_1), \dots, W(h_n)) ; h_1, \dots, h_n \in \mathfrak{H}, \quad (0.0.9)$$

où  $f$  appartient à l'espace de Schwartz  $\mathcal{S}(\mathbf{R}^n)$ . Par densité de l'espace  $\mathcal{S}$  dans l'espace des fonctionnelles de Wiener intégrables, cette définition peut être étendue au domaine de  $D$ , noté  $\mathbf{D}^{1,2}$ . L'adjoint de l'opérateur  $D$ , appelé *divergence* et noté  $\delta$ , coïncide avec l'intégrale de Wiener (resp. de Skorohod) pour des processus adaptés (resp. non-adaptés). Les opérateurs  $D$  et  $\delta$  sont liés par l'identité fondamentale :

$$\mathbf{E}[\langle DF, U \rangle_{\mathfrak{H}}] = \mathbf{E}[F \delta U] ; F \in \mathbf{D}^{1,2}, U \in \text{Dom } \delta.$$

Comme cela est détaillé notamment dans les ouvrages de Paul-André Meyer [89] ou Nobuaki Obata [107], il existe une approche alternative menant au calcul de Malliavin et reposant sur la *représentation chaotique de Wiener-Itô*. Elle est basée sur le développement de ce qui serait l'analogie (en dimension infinie) de la théorie de Schwartz où la mesure Gaussienne  $\mu$  sur le dual  $E^*$  d'un espace nucléaire  $E = \mathcal{S}(\mathbf{R})$  se substituerait à la mesure de Lebesgue sur  $\mathbf{R}^n$ . L'espace  $\mathfrak{E}$  (resp.  $\mathfrak{E}^*$ ) des *fonctionnelles de bruit blanc test* (resp. *généralisées*) est construit par un plongement dense et continu dans  $L^2(E^*, \mu)$  (resp. par dualité) :

$$\mathfrak{E} \subset L^2(E^*, \mu) \subset \mathfrak{E}^*.$$

L'espace  $L^2(E^*, \mu)$  est en fait canoniquement isomorphe à l'*espace de Fock* sur  $L^2(\mathbf{R})$  à travers l'isomorphisme de Wiener-Itô-Sega. Ainsi toute fonctionnelle de Wiener de carré intégrable admet une unique décomposition en termes d'intégrales de Wiener, de sorte à ce que cette *décomposition chaotique* caractérise cette approche alternative. Dans le contexte général des espaces de Fock,  $D$  et  $\delta$  coïncident respectivement avec l'*opérateur d'annihilation* et l'*opérateur de création* que l'on rencontre en théorie des probabilités quantique et satisfont une relation canonique de commutation généralisée. L'équivalence de ces deux approches repose sur la relation intrinsèque entre le mouvement Brownien standard et les polynômes de Hermite.

Paul-André Meyer met en lumière dans [88] l'équivalence des normes associées aux différents espaces de Sobolev ainsi définis.

Si le formalisme de Malliavin était circonscrit à l'espace Wiener à ses débuts il fut par la suite étendu à d'autres types de processus ; pour n'en citer que quelques uns, les processus Gaussiens en général (voir Nualart [98], Nourdin et Peccati [94]), les processus de Poisson (voir Bichteler et al. [18] pour un point de vue variationnel, Nualart and Vives [103] ou Privault [116] dans le cadre d'une approche chaotique), les processus de Lévy (voir Nualart and Schoutens [102]) et les processus de Rademacher (voir Privault [117]).

La multiplicité des approches et la variété des espaces canoniques sur lesquels il opère donc, semblent dissuader l'éventuel projet d'un travail unificateur. Si la formalisation d'une construction universelle paraît donc compromise, il demeure envisageable d'identifier un dénominateur commun à tous ces formalismes. Celui-ci serait plutôt à chercher du côté de la terminologie du calcul de Malliavin qui peut être introduite par analogie avec celle du calcul différentiel classique dans les espaces de Banach : aux vecteurs se substituent les trajectoires (indexées par le temps ou un espace de Hilbert) du processus canonique, aux fonctions les fonctionnelles opérant sur l'espace des trajectoires, tandis qu'on désigne volontiers par les termes *gradient* et *divergence*, les deux opérateurs phares de cette théorie : la dérivée au sens de Malliavin et son adjoint. Le semi-groupe d'Ornstein-Uhlenbeck et son générateur complètent la famille des opérateurs de Malliavin. L'existence d'une relation fondamentale liant gradient et divergence parachève la description du champ lexical du calcul Malliavin et apparaît comme sa pierre angulaire : la *formule d'intégration par parties* au sens de Malliavin,

$$\mathbf{E}[\langle DF, U \rangle_{\mathfrak{H}}] = \mathbf{E}[F \delta U], \quad (0.0.10)$$

où  $(\Omega, \mathcal{A}, \mathbf{P})$  est l'espace de probabilité canonique et  $\mathfrak{H}$  est un espace de Hilbert générique à voir comme *l'espace des perturbations* lorsque  $DF$  est interprété comme une dérivée directionnelle (par exemple  $\mathfrak{H}$  est l'espace de Cameron-Martin pour le mouvement Brownien).

Au cours de la préparation de cette thèse, deux espaces munis d'une structure de Malliavin suscitèrent spécialement notre intérêt : l'espace de Poisson et l'espace de Rademacher. Le formalisme de Malliavin développé dans ces deux cadres sera décrit puis comparé à nos contributions à la fin de cette introduction.

## Branches et fruits : les applications du calcul de Malliavin

La rencontre détonante du calcul de Malliavin avec d'autres branches des probabilités ou de l'analyse a fait entrevoir la puissance de cet outil et l'étendue des applications dont il est garant des bases théoriques. Nombreux sont les champs dans lesquels le calcul de Malliavin a trouvé sa place ; à titre d'exemple on peut citer l'étude des opérateurs de Hörmander (qui est l'application « historique »), plus généralement l'étude de la régularité des lois de probabilité de solutions d'EDS/EDPs, la finance, la méthode de Stein, le calcul d'erreur, le calcul anticipatif et plus récemment l'inférence statistique.

## Équations différentielles stochastiques, équations aux dérivées partielles et calcul anticipatif

Gardant à l'esprit les résultats du papier pionnier de Paul Malliavin, de nombreux auteurs réutilisèrent la condition de non-dégénérescence pour étudier la régularité des solutions d'équations

aux dérivées partielles et d'équations différentielles stochastiques. Cela donna lieu à de nombreuses applications passées en revue notamment par Martha Sanz-Solé dans [128] (resp. par Hu, Huang, Lê, Nualart and Tindel [65]) pour les EDP dirigées par un bruit coloré (resp. par un bruit blanc en temps /bruit rugueux en espace avec un indice de Hurst  $H \in (1/4, 1/2)$ ). On peut également citer les travaux de Vlad Bally et Denis Talay [13], Yaozhong Hu, David Nualart et Jian Song [66], ou Shigeo Kusuoka et Daniel W. Stroock [78] pour des équations différentielles stochastiques dirigées par le mouvement Brownien et ceux de David Nualart et Bruno Sausserau [101] pour les EDS dirigées par un mouvement Brownien fractionnaire avec un indice de Hurst  $H \in (1/2, 1)$ .

Les résultats du calcul de Malliavin furent également exploités pour étudier des équations différentielles stochastiques non-causales formulées en termes d'intégrales stochastiques anticipatives. En effet l'opérateur divergence permet de définir une extension de l'intégrale stochastique d'Itô à des intégrales anticipatives dans le cas du mouvement Brownien (voir Skorohod [131]). Une formule d'anticipation est donnée dans Nualart et Pardoux [99] ou dans une version antérieure (et méconnue) par Masayuki Hitsuda [64] qu'Elisa Alòs et David Nualart améliorèrent dans [4]. Un *calcul anticipatif* d'abord développé selon cette approche pour le mouvement Brownien (voir Berger [17] ou dans le chapitre 3 de Nualart [98]) a depuis été étendu aux processus de Poisson (voir Decreusefond and Savy [40], Nualart et Vives [103]) et aux processus de Lévy (voir [3],[105]). Elisa Alòs, Olivier Mazet et David Nualart utilisèrent également le calcul de Malliavin (en particulier la divergence) pour donner un sens à l'intégrale stochastique par rapport au mouvement Brownien fractionnaire avec un indice de Hurst  $H \in (0, 1/2)$  et construire un calcul stochastique dans ce contexte.

## Statistique

Un large éventail d'outils issus du calcul de Malliavin trouvèrent leur place dans des travaux en statistique ; nous en donnons un bref aperçu dans ce paragraphe. Par exemple, Ciprian A. Tudor et Frederic G. Viens [135] utilisèrent la *décomposition chaotique* pour proposer un adaptateur pour le paramètre d'auto-similarité  $H$  du processus de Rosenblatt. Fabienne Comte et Nicolas Marie [29] montrèrent que l'intégrale de Skorohod (définie à partir de la *divergence*) est l'unique extension de l'intégrale d'Itô valable pour l'estimation non-paramétrique dans des EDS dirigées par un mouvement Brownien fractionnaire d'indice de Hurst  $H \in (1/2, 1)$ . José M. Corcuera et Arturo Kohatsu-Higa [31] tirèrent profit de la *formule d'intégration par parties* pour étudier l'inférence asymptotique d'un processus (à sauts) stochastique, tandis que Nicolas Privault et Anthony Réveillac l'utilisèrent à une autre fin : celle de d'estimer le drift de processus Gaussiens à l'aide d'estimateurs de type Stein [118].

## Finance

Une grande variété des *fruits* nés de la rencontre du calcul de Malliavin et de la *finance* sont décrits la monographie de Paul Malliavin et Anton Thalmaier (voir [86]). Trois de ces applications ont particulièrement éveillé notre intérêt : le calcul de Grecques, la gestion de portefeuille, et le délit d'initié.

Afin de placer ces points d'intérêt dans le vaste univers des mathématiques financières, restreignons d'abord le cadre d'étude à celui d'un marché composé simplement de deux actifs : un actif sans risque modélisé par une suite déterministe  $A = (A_t)_{t \in \mathbf{T}}$  et un actif risqué dont la suite des cours est donnée par  $S = (S_t)_{t \in \mathbf{T}}$ . Supposons d'abord que  $A$  est défini par une

suite  $(r_t)_{t \in \mathbf{T}}$  de taux d'intérêt et que  $S$  satisfait une équation différentielle stochastique (dans le cadre du modèle de Black-Scholes par exemple) ou une équation aux différences (dans le cadre de modèles discrets) sur l'intervalle de trading  $\mathbf{T}$ . Nous nous intéressons à certains *produits dérivés* particuliers, dont la valeur dérive donc des performances de  $S$ , appelés *options (vanille)*. Une option est un contrat qui donne à l'acheteur (le propriétaire ou détenteur) le droit mais non l'obligation, d'acheter ou de vendre un actif risqué à un prix fixé d'avance  $\mathcal{K}$  et à une date spécifique  $T$ . La valeur d'une option à échéance, correspondant à ce qui est finalement perçu par l'acheteur, est appelée *payoff* ; pour une option, cette valeur est modélisée par une variable aléatoire  $\Phi(S_T)$ . Enfin, on considère que sur ce marché agissent des vendeurs et acheteurs considérés comme *des investisseurs ordinaires* (resp. des *initiés*) si les décisions relatives à la composition de leur portefeuille (avec les deux types d'actifs à leur disposition) sont prises sur la base d'informations publiques (resp. en bénéficiant d'informations supplémentaires).

Les *Grecques*, ainsi dénommées par allusion aux lettres grecques par lesquelles elles sont désignées, sont des instruments de calcul de sensibilité du prix d'un actif relativement à la modification de paramètres sous-jacents dont sa valeur dépend. Elles sont d'une utilité capitale pour l'agent car elle lui confèrent des éléments de contrôle de son exposition aux risques. Par exemple, Delta et Rho sont les Grecques mesurant respectivement la sensibilité au prix de l'actif sous-jacent et au taux d'intérêt alors que Gamma désigne un indicateur du second ordre quantifiant la variation du Delta relativement à une variation du prix sous-jacent. Pendant longtemps, le calcul de ces indicateurs a été la seule prérogative des analystes qui s'y employaient par des méthodes de différences finies dans le cadre de la méthode de Monte-Carlo. Sans introduire précisément le formalisme (ce qui sera fait au chapitre 4) considérons une option de payoff  $\mathcal{K} = \Phi(S_T)$  et dont la dynamique de la suite de prix de l'actif risqué sous-jacent  $(S_t)_{t \in \mathbf{T}}$  est donnée par le modèle de Black-Scholes :

$$S_t = 1 + \int_0^t r S_u du + \int_0^t \sigma S_u dB_u, \quad t \in \mathbf{T}, \quad (0.0.11)$$

où les paramètres  $r$  et  $\sigma$  représentent respectivement le taux d'intérêt sans risque et la volatilité des profits générés par l'actif risqué. La valeur de l'option à un instant donné  $t \in \mathbf{T}$  est notée  $V_t$ . Le Delta s'exprime (dans le modèle Black-Scholes) par

$$\Delta = \frac{\partial V_0}{\partial S_0} = e^{-rT} \mathbf{E}_{\mathbf{Q}} \left[ \Phi'(S_T) \frac{\partial S_T}{\partial S_0} \right], \quad (0.0.12)$$

où  $\mathbf{Q}$  correspond à la mesure de probabilité risque neutre. En général, la fonction  $\Phi$  n'est pas dérivable ; par exemple, le payoff d'un *call européen* (option d'achat à une date déterminée) est  $\Phi(S_T) = (S_T - \mathcal{K})_+$ . Il serait donc accommodant de pouvoir réécrire  $\Delta$  sous la forme  $\Delta = e^{-rT} \mathbf{E}_{\mathbf{Q}} [\Phi(S_T) \pi]$ , où le *poids*  $\pi$  ne dépendrait pas du payoff. L'idée de base fut d'abord introduite par Éric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, Pierre-Louis Lions et Nizar Touzi dans [54] et [55] qui, tirant avantage de la formule d'intégration par parties d'une part, de la *chain rule* ("règle de dérivation en chaîne") d'autre part, ont prouvé que

$$\Delta = e^{-rT} \mathbf{E}_{\mathbf{Q}} [\Phi(S_T) \delta(u)],$$

de telle sorte que le poids s'exprime comme l'intégrale de Skorohod de générateur  $u$ , ici déterministe et constant égal à  $(\sigma T)^{-1}$ . Éric Benamou ajouta dans [16] des conditions nécessaires et suffisantes sur  $u$  permettant d'automatiser la génération d'un poids dans la perspective

de simuler les Grecques. Le travail pionnier de Fournier et al. [54] a été transposé pour les modèles à sauts (en finance) dirigés par des processus de Poisson par Nicolas Privault et al. (voir [50],[121]) et par des processus de Lévy (notamment) par Marie-Pierre Bavouzet-Morel et Marouen Messaoud dans [15].

Fortement connectée à la théorie des semi-martingales, l'évaluation des options et en particulier les problèmes de couverture ont également tiré bénéfice des contributions du calcul de Malliavin. La réponse à la question de *hedging* ("couverture") est cruciale dans la mesure où elle donne des éléments de compréhension aux acheteurs ou vendeurs pour composer et ajuster de façon dynamique leur portefeuille dans le but de *répliquer* cette option, i.e. atteindre sa valeur à échéance. Mathématiquement, cela revient à déterminer un processus  $\psi$  à valeurs dans  $\mathbf{R}^2$  satisfaisant une condition d'auto-financement (voir le chapitre 4 du manuscrit pour son expression dans les modèles de Black-Scholes et de Cox-Ross-Rubinstein)

$$V_T(\psi) = \Phi(S_T).$$

À noter que dans les modèles à temps discret l'hypothèse supplémentaire,

$$V_t(\psi) \geq 0, ; \forall t \in \mathbf{N}_T, \mathbf{P} - \text{a.s.},$$

est requise. Dans les marchés complets où toutes les options sont répliquables, et dans lesquels l'espace probabilisé sous-jacent est muni d'une structure de Malliavin, le processus  $\psi$  s'exprime élégamment à l'aide de la dérivée de Malliavin ; cela fut suggéré dans les deux travaux indépendants et contemporains (tous deux publiés en 1991) de Ioannis Karatzas et Daniel Ocone [73] d'une part et avec Jinlu Li [108] d'autre part. La formule de couverture de Karatzas-Ocone hedging est en effet dérivée de celle de Clark-Ocone. Des formules analogues furent établies pour les processus de Poisson par Günter Last et Matthew Penrose [81] et pour les processus de Lévy par Giulia Di Nunno [104].

Le délit d'initié est un sujet intrinsèquement lié à l'optimisation de portefeuille. Le cadre est celui d'un marché financier où deux agents ayant des niveaux différents d'information coexistent : un agent ordinaire dont les décisions sont prises à la lumière des informations publiques et un initié bénéficiant d'un surplus d'information (confidentielle) dès le début de la période d'échanges. Deux questions se posent naturellement : comment quantifier l'utilité supplémentaire espérée par l'initié ? L'information additionnelle dont il bénéficie produit-elle un arbitrage ? Ce surplus d'information peut se traduire mathématiquement par un élargissement de la filtration sur lequel l'initié fonde ses décisions de gestion de portefeuille et conduit donc alors à la théorie de grossissement de filtrations. Suite aux travaux pionniers d'Igor Pikovsky et de Ioannis Karatzas [115] dans le domaine, Jürgen Amendinger et al. [7],[6],[5], Axel Gyorud et Monique Pontier [61] précisèrent les critères pour l'optimisation de portefeuille et calculèrent l'utilité additionnelle de l'initié dans leur travaux respectifs. Peter Imkeller connecta cette utilité au calcul de Malliavin dans [68] en exprimant le *drift d'information* comme la trace de Malliavin logarithmique de densités conditionnelles caractérisant le bénéfice de l'initié. Jorge A. León, Reyla Navarro et David Nualart utilisèrent dans [83] certaines techniques du calcul de Malliavin pour analyser les propriétés de *l'intégrale anticipative* (introduite par Russo et Vallois dans [127]) et pour maximiser l'utilité logarithmique espérée de l'initié.

Dans certaines extensions du modèle de Black-Scholes où la volatilité peut être un processus stochastique (voir par exemple Hull et White[67], ou Heston [63]), en particulier dans certains modèles de diffusion à volatilité stochastique, où la volatilité suit également un processus de diffusion, on peut observer que le *volatilité implicite* (mesure prospective pour estimer les

fluctuations futures de l'actif sous-jacent) se comporte *rugueusement* à toute échelle de temps raisonnable. La boîte à outils de Malliavin, dont l'efficacité en matière de calcul anticipatif a déjà été mentionnée, est alors également efficace pour analyser les volatilités futures, qui sont des processus non adaptés. Dans un travail précurseur sur les *volatilités rugueuses* [2], Elisa Alòs, Jorge León et Josep Vives, donnent une expression pour le comportement à court terme de la volatilité implicite dans un modèle de diffusion à sauts en fonction de sa dérivée de Malliavin.

### La méthode de Stein

Dans son article fondateur ([132]), Charles Stein décrit une procédure visant à quantifier les erreurs commises dans l'approximation Normale par des sommes de variables aléatoires ayant une structure de dépendance stationnaire. Par la même, il posa les bases d'une méthode très puissante dont les perspectives d'application dépassèrent rapidement le cadre de sa naissance ; la *méthode de Stein* est devenue l'une des (pour ne pas dire la) façons les plus efficaces de calculer la distance entre deux mesures de probabilité s'écrivant sous la forme :

$$\text{dist}_{\mathcal{H}}(\mathbf{P}^*, \mathbf{Q}) = \sup_{h \in \mathcal{H}} \left| \int_{\mathfrak{F}} h \, d\mathbf{P}^* - \int_{\mathfrak{F}} h \, d\mathbf{Q} \right|,$$

où  $\mathbf{P}^*$  représente la mesure cible (à approximer) et  $\mathbf{Q}$  la mesure initiale, toutes deux étant définies sur le même espace métrique  $(\mathfrak{F}, c)$  et  $\mathcal{H}$  est un ensemble de fonctions test. Même si la méthode de Stein conserve un lien particulier avec les questions d'approximation Normale qui ont motivé son développement (se référer au livre de référence de Louis H.Y. Chen, Larry Goldstein et Qi-Man Shao [27]), ses principes ont pu être transposés à d'autres distributions cibles. En particulier, la méthode de Stein-Chen a été conçue pour traiter les cas d'approximation de la loi Poisson (voir Chen [26]). De nombreux critères d'approximation plus généraux d'autres lois de variables aléatoires par des distributions variées ont suivi : Gamma (Arras et Swan [8], Döbler et Peccati [44], Peköz, Röllin et Ross [114]), Exponentiel (Chatterjee, Fulman et Röllin [25], Peköz et Röllin [112]), Géométrique (Peköz, Röllin et Ross [113]) ... En outre, dans la section *Taxonomie* de la page web [Malliavin-Stein approach](#) qu'il lui a consacré, Yvik Swan passe en revue toutes les distributions dont l'approximation peut être quantifiée par cette méthode. De nombreuses études ont été publiées pour enseigner au prosélyte et initier le profane à l'efficacité du concept. S'il reste impossible d'en donner une liste exhaustive, nous recommandons les ouvrages de référence et particulièrement pédagogiques de Benjamin Arras et Yvik Swan, [8], Andrew D. Barbour [14] ou Nathan Ross [125].

La méthode de Stein semble se déployer en deux étapes ; la première consiste à convertir le problème initial (difficile) de borne de l'erreur dans l'approximation de la mesure  $\mathbf{P}^*$  par  $\mathbf{Q}$  en celui du contrôle d'une expression de la forme :

$$\mathbf{E}_{\mathbf{Q}}[L\varphi(X)] = \mathbf{E}[L_1\varphi(X)] + \mathbf{E}[L_2\varphi(X)], \quad (0.0.13)$$

où  $L$  (resp. la classe d'appartenance de  $\varphi$ ) est déterminé par un procédé de caractérisation de  $\mathbf{P}^*$  (resp.  $\mathbf{P}^*$  et  $\mathcal{H}$ ), et  $X$  est une variable aléatoire de loi  $\mathbf{Q}$ . La seconde composante de la méthode de Stein regroupe des techniques permettant de borner (0.0.13) ; se basant sur la définition même de  $X$ , cela revient à transformer  $L_1\varphi(X)$  en  $-L_2\varphi(X) + \text{reste}$ . C'est ce reste qui donne une borne pour la distance entre lois et dans le cas d'un problème de convergence, en fournit la vitesse. Plusieurs approches ont émergé au fil des années pour indiquer comment

effectuer la transformation de  $L_1\varphi(X)$ . L'une des approches les plus populaires (voir par exemple Barbour et Chen [14]) consiste à utiliser des paires échangeables. Des solutions cousines, appelées *size-biased (coupling)* (voir Chen, Goldstein et Shao [27]) ou *zero-biased (coupling)* (voir Chen, Goldstein et Shao [27]) basées également sur des couplages tirent bénéfice des propriétés de la mesure initiale (voir Goldstein et Reinert [58]), pour, là encore, transformer commodément  $L_1\varphi$ . Dans un travail inédit, Ivan Nourdin et Giovanni Peccati (voir [93],[94]) ont montré que l'étape de transformation pouvait être avantageusement simplifiée en utilisant l'intégration par parties au sens de Malliavin ; ils ont donné par la même occasion une intersection aux deux théories. Cette approche est efficace sous couvert de l'existence d'un gradient de Malliavin sur l'espace sur lequel  $X$  est défini. Elle a notamment été appliquée aux fonctionnelles de Rademacher (voir Nourdin, Peccati and Reinert [96], Zheng [143], Poisson (voir par exemple Decreusefond, Schulte et Thäle [41], Lachièze-Rey et Peccati [79], Peccati, Solé, Taqqu et Utzet [111]) ou Gaussienne variables aléatoires (voir Nualart et Peccati [100]) ou processus (voir Coutin et Decreusefond [32, 33]). Une liste des travaux se rapportant à l'approche de Stein-Malliavin est disponible sur la page web entretenue par Ivan Nourdin.

### Les formes de Dirichlet

En parallèle de sa connexion avec la méthode de Stein, le langage du calcul Malliavin se mêle avantageusement à celui des formes de Dirichlet le cadre du calcul d'erreur (voir Bouleau et Hirsch [23]). Les sujets connexes ont généré un intérêt croissant (au sein de la communauté des analystes avant que les probabilistes ne s'en emparent) depuis le début du XXe siècle. L'utilisation des formes Dirichlet s'avère particulièrement efficace pour effectuer des calculs d'erreurs lorsque celles-ci sont supposées être infinitésimales et probabilistes. Si  $Y$  représente une quantité scalaire erronée, la propagation de l'erreur par la fonction  $f$  est caractérisée par les formules :

$$\left\{ \begin{array}{l} \text{biais de l'erreur sur } f(Y) = (\text{biais de l'erreur sur } Y) f'(Y) + \frac{1}{2}(\text{var de l'erreur sur } Y) f''(Y) \\ \text{var de l'erreur sur } f(Y) = (\text{var de l'erreur sur } Y) f'^2(Y) \end{array} \right. \quad \begin{array}{l} (E_1) \\ (E_2) \end{array} ,$$

qui deviennent, si la variance est de même ordre de magnitude que le biais ou si le biais est négligeable par rapport à la variance :

- pour le biais de l'erreur : une équation différentielle du second ordre avec biais et variance (E<sub>1</sub>),
- pour la variance de l'erreur : une équation différentielle du premier ordre pour la variance n'incluant pour de terme de biais (E<sub>2</sub>).

Considérant  $F = F(X_A)$  ( $A \subset \mathbf{N}^*$ ) fonction de quantités erronées  $\{X_k, k \in A\}$  (supposées connues, petites et indépendantes) l'erreur quadratique potentielle à espérer sur  $F$  notée  $\sigma_F^2$  peut être écrite comme suit :

$$\sigma_F^2 = L(F^2) - 2F L(F), \quad (0.0.14)$$

où  $L$  est l'opérateur différentiel

$$L = \frac{1}{2} \sum_{k \in A} \sigma_k^2 \frac{\partial^2}{\partial X_k^2}$$

et  $\{\sigma_k^2, k \in A\}$  est la famille des erreurs quadratiques sur  $\{X_k, k \in A\}$ . En fait, l'erreur quadratique peut s'exprimer

$$\sigma_F^2 = \Gamma(F) = \Gamma(F, F),$$

où  $\Gamma$  est l'opérateur carré du champ défini par

$$\Gamma(F, G) = \frac{1}{2} \left( L(FG) - GLF - FLG \right). \quad (0.0.15)$$

L'utilisation de ces outils est plus pertinente lorsqu'il s'agit d'estimer la vitesse de convergence d'une suite de variables aléatoires  $(X_n)_{n \in \mathbf{N}^*}$  vers une variable aléatoire  $X$ , toutes supposées être définies sur un espace de probabilité  $(\Omega, \mathcal{A}, \mathbf{P})$ . En effet, si la forme  $\Gamma$  est fermée, la théorie du calcul d'erreur se déploie en s'appuyant sur le principe suivant : si la suite de couples  $(X_n, \text{erreur sur } X_n)$  converge convenablement, elle converge nécessairement vers un couple  $(X, \text{erreur sur } X)$ . Lorsque l'espace sous-jacent est muni d'une structure de Malliavin, l'opérateur  $L$  apparaissant dans (0.0.15) coïncide avec l'opérateur Laplacien/nombre  $L = -\delta D$ . Ce point clef sera expliqué dans la section 1.3.2 et illustré dans le cas de l'espace Wiener à travers l'exemple 1.3.6. Cette relation suggère en fait que les formes de Dirichlet s'insèrent implicitement dans la combinaison Stein-Malliavin pour aboutir à une méthode de Stein-Dirichlet-Malliavin (voir Decreusefond [36]). En outre, Ehsan Azmoodeh, Simon Campese et Guillaume Poly ont exploité le lien entre les trois théories pour revisiter le théorème du moment d'ordre 4 dans [11].

La version dessinée de cet état de l'ar-bre se trouve page 17 et conclut cette partie.

En reprenant cette métaphore, nous pouvons résumer les contributions de cette thèse à l'addition de deux racines à cet arbre : la construction d'un calcul de Malliavin pour les **suites de variables aléatoires indépendantes** d'une part, et pour les **processus géométriques composés** d'autre part. Au delà de l'intérêt propre que chaque construction a pu susciter, notre choix de développer un formalisme dans ces deux cas a été tout ou partie motivé par les perspectives d'applications qu'il laissait entrevoir. Aussi, les fruits résultant de ces deux formalismes sont représentés en couleur dans la figure page 17 : couleur **pourpre** pour ceux relatifs au cadre des variables aléatoires indépendantes, et en **bleu** pour ceux relatifs au cadre des processus géométriques composés.

## Calcul de Malliavin et structures de Dirichlet pour des variables aléatoires indépendantes

La motivation d'élaborer un calcul de Malliavin pour des **variables aléatoires indépendantes** était double ; d'une part, pour lui-même et pour généraliser ce qui avait été exploré jusqu'alors dans un cadre discret, et d'autre part pour les applications potentielles au contrôle des suites de variables aléatoires indépendantes auxquelles son élaboration pourrait conduire.

Jusqu'ici l'incursion du calcul de Malliavin dans un contexte discret était circonscrite à l'espace de Rademacher. L'approche chaotique par laquelle est développé un calcul de Malliavin dans cet espace s'appuie sur l'existence de martingales normales ou *novations* (pour reprendre la terminologie de Émery) associées au processus canonique. L'existence de novations requiert que les lois des variables aléatoires ne soient supportées que par deux points. L'expression du gradient (0.0.8) de Rademacher n'a d'ailleurs de sens que si les variables aléatoires sont réelles ou que leur espace d'état est réduit à deux points. Dans le cas général que nous avons choisi d'investiguer, celui de suites de variables aléatoires sur lesquelles une simple hypothèse d'indépendance est émise, il convient de préciser que les espaces d'états sous-jacents peuvent être différents (simplement supposés Polonais). Devant l'impossibilité de définir l'opérateur



gradient comme opérateur d'annihilation opérant sur un espace décomposable en chaos, nous définissons le gradient discret comme un *opérateur différence* dont la  $k$ -ème coordonnée traduit l'idée de mesure de l'influence de la  $k$ -ème composante du processus initial sur la fonction qui s'y applique. S'en suit la définition de son adjoint, appelé divergence (s'exprime comme somme des dérivées partielles) et qui satisfait une formule d'intégration par parties discrète. La définition de l'opérateur nombre (appelé Laplacien dans le cas Gaussien), qui est également générateur d'un semi-groupe de Markov, complète alors la définition des opérateurs de Malliavin.

Nous justifions notre construction à plusieurs égards et lui conférons ainsi une forme de légitimité. Tout d'abord, en la connectant à son aînée sur l'espace de Rademacher. Bien que les définitions de gradient soient distinctes (pour les raisons précédemment évoquées), il apparaît que les opérateurs nombre définis dans les deux contextes coïncident. L'*opérateur nombre* joue par ailleurs au rôle prépondérant dans l'établissement d'inégalités fonctionnelles. C'est donc par cette correspondance que nous pouvons en revisiter certaines.

D'autre part, notre formalisme se raccorde aux théories préexistantes en temps continu par les structures de Dirichlet. De telles structures sont de facto définies sur tout espace canonique muni d'un calcul de Malliavin formalisant ainsi le concept de *structure de Dirichlet-Malliavin*. En effet, les structures de Dirichlet Poissonienne et Brownienne usuelles associées à leur gradient respectif s'écrivent comme limites des structures induites par notre formalisme.

Enfin, nous dérivons de notre construction nombre d'identités fonctionnelles. Nous transposons ainsi l'identité de covariance, l'inégalité de log-Sobolev et l'inégalité de déviation classiques dans notre contexte. Nous revisitons d'autres auxquelles nous donnons une nouvelle interprétation. Par exemple, nous montrons que l'inégalité d'Efron-Stein peut être interprétée comme une inégalité de Poincaré ou que la décomposition de Hoeffding des  $U$ -statistiques peut être vue comme un avatar de la formule de représentation de Clark. Ainsi, notre formalisme trouve sa place dans le paysage du calcul de Malliavin ; il généralise ce qui avait été théorisé dans le cas Rademacher et se connecte avec les contextes Brownien et Poissonien par approximation des structures de Dirichlet induites d'une part, et par la ressemblances des identités fonctionnelles (comme la formule de Clark, inégalité de Poincaré...) obtenues dans les trois cas d'autre part.

La seconde motivation pour élaborer cette construction provenait de la méthode de Stein, avec qui, comme évoqué précédemment, branche des un tel formalisme serait susceptible de se combiner avantageusement. Nous donnons une borne générale de type "Berry-Esseen" pour la distance de Wasserstein entre la loi de fonctionnelles de variables aléatoires indépendantes et la loi Normale ; et par la même d'établir un critère de "Stein-Malliavin" analogue à celui rencontré dans les cas Gaussien et Poisson. Nous formulons un critère similaire pour l'approximation de la loi Gamma que nous appliquons au cas particulier de  $U$ -statistiques d'ordre 2 et en obtenons une estimation de type de Jong.

Pour toutes ces raisons, il semble que notre structure de Dirichlet-Malliavin donne un cadre unificateur à de nombreux résultats disséminés dans la littérature et suggère que des résultats sans lien apparent (Efron-Stein, paires échangeables etc.) sont en réalité les faces d'une même pièce... à "l'effigie" de la formule d'intégration par parties.

## Calcul de Malliavin pour des processus géométriques composés et délit d'initié dans un modèle ternaire

L'objectif initial du second travail n'était pas de concevoir un calcul stochastique pour les processus composés géométriques mais plutôt d'utiliser notre calcul discret de Malliavin pour traiter des problèmes de délit d'initié. C'est également la raison pour laquelle le titre de la thèse n'en fait pas mention. En effet, nous avons d'abord essayé de déployer notre artillerie pour atteindre une cible identifiée : le calcul de la valeur logarithmique additionnelle attendue l'utilité d'un initié dans le modèle trinomial. Ce problème présente un intérêt particulier à plusieurs égards ; tout d'abord, et comme mentionné plus haut, les mathématiques financières sont le champ d'application phare du calcul de Malliavin. Si la couverture d'options avait déjà été étudiée dans le modèle Cox-Ross-Rubinstein (voir le chapitre 1 dans Privault [117]) en utilisant la structure de Malliavin dont est équipé l'espace Rademacher, aucun projet similaire n'avait été réalisé jusqu'alors dans le cadre des arbres multinomiaux (avec au moins trois branches à chaque étape). En outre, si de nombreux travaux traitent de l'élargissement de filtrations et des questions de délit d'initié en temps continu (voir les travaux d'Amendinger, Imkeller et Schweizer ([7]), très peu ont été réalisés dans des modèles discrets. Certains résultats existent déjà pour l'élargissement de filtrations en temps discret. Catherine Blanchet-Scalliet, Monique Jeanblanc et Roméo Romo Romero ont montré dans [20] que les résultats connus en temps continu s'étendent immédiatement dans un contexte discret. La plupart d'entre eux sont facilement obtenus grâce à la décomposition de Doob. À cet égard, nous souhaitons discuter de la possibilité d'interpréter le processus prévisible croissant apparaissant dans la décomposition de Doob comme le *drift d'information* associé à l'initié et l'exprimer à l'aide de la dérivée de Malliavin, à l'instar du travail de Peter Imkeller en temps continu [69].

Si notre formalisme se prête bien au calcul de Grecques dans le modèle trinomial, il demeure en effet impossible d'en dériver une formule de Karatzas-Ocone pour la stratégie couverture et donc d'en tirer avantage pour l'évaluation d'options. Celle-ci est une conséquence directe du théorème de représentation de martingale qui ne peut être statué quand les lois des variables aléatoires (discrètes) sont portées par au moins trois points. Ce constat a motivé le développement d'un calcul de Malliavin pour les processus géométriques composés (dont les variables i.i.d. correspondantes sont à valeurs dans  $\{-1, 0, 1\}$ ), sur lesquels baser un modèle ternaire équivalent en loi au modèle trinomial.

De par la structure de saut qui les sous-tend, les processus géométriques héritent de nombreuses propriétés des processus de Poisson, source d'inspiration pour élaborer un calcul de Malliavin. S'il demeure, comme attendu, impossible d'établir une décomposition en chaos à partir de martingales normales, (le carré des variables aléatoires i.i.d. étant constant), il existe une *décomposition en pseudo-chaos* au moyen d'intégrales multiples relatives à une famille non-orthogonale (mais orthogonalisable) de variables aléatoires. Le gradient géométrique est alors défini comme l'*opérateur d'annihilation* agissant sur cette pseudo-décomposition chaotique et tel que l'intégration stochastique (définie par rapport à la famille orthogonalisable) apparaît - à une transformation linéaire près - comme l'opération "inverse". En outre, cet opérateur coïncide avec un opérateur différence dont l'expression est plus propice à statuer des identités fonctionnelles comme la formule de représentation martingale, la formule d'Itô, et la formule de Ocone-Karatzas espérée.

Nous appliquons alors ce formalisme au calcul de l'utilité additionnelle d'un initié dans un modèle ternaire construit sur la dynamique d'un processus géométrique composé. Reprenant

les outils développés pour l'élargissement de filtration en temps discret par Catherine Blanchet-Scalliett, Monique Jeanblanc et Roméo Romo Romero, nous définissons l'analogie du drift d'information dont nous donnons une expression en termes de gradient géométrique. Nous montrons alors que l'utilité logarithmique espérée additionnelle de l'initié au temps  $t$  peut être exprimée comme *l'entropie relative* de la mesure initiale par rapport à la mesure préservant les martingales sur  $[0, t]$ . Nous retrouvons ainsi exactement le résultat établi par Jürgen Amendinger, Peter Imkeller et Martin Schweizer dans le cas continu. Nous illustrons ce résultat et achevons ce travail en effectuant des calculs explicites dans le cas précis où l'initié obtient une information supplémentaire sur la valeur finale de l'actif risqué.

### Manuscrit

Le manuscrit est organisé comme suit ; il se compose de deux parties, chacune d'entre elles se référant à l'un des travaux mentionnés ci-dessus. La première, intitulée *Structures de Malliavin et Dirichlet pour les variables aléatoires indépendantes* renvoie à un travail mené en collaboration avec Laurent Decreusefond et publié dans Stochastic Processes and their Applications en 2019 (voir [38]). La deuxième sera soumise prochainement. Les deux parties se déploient selon le même schéma ; elles commencent par un chapitre où est présenté l'état de l'art et où l'on identifie le formalisme à développer pour répondre à une problématique donnée. Les outils nécessaires sont construits dans les chapitres qui suivent immédiatement alors que les **fruits** sont récoltés dans le dernier chapitre de chaque partie, comme la réponse apportée au problème initial.

Avant de clore cette introduction, je signale que toute nouvelle contribution (définition ou résultat original) sera signalée dans le manuscrit par le symbole (★). Malgré la piètre qualité de l'anglais dont souffre la rédaction (et pour laquelle je présente par avance mes excuses aux lecteurs), j'ai beaucoup apprécié la préparation de ce manuscrit de thèse et vous en souhaite une agréable lecture !

## Part I

# Malliavin calculus and Dirichlet structures for independent random variables



# Chapter 1

## From Stein's to Stein-Dirichlet-Malliavin method

Stein's method, initially developed to quantify the rate of convergence in the Central Limit Theorem [132] and then for Poisson convergence [26], has become a very popular not to say the most famous procedure to assess distances between two probability measures. Associated to a collection of probability metrics, it seems to deploy by two main stages: the first one consists in converting the initial problem into a more tractable expression before developing in the second one tools to derive from it the bound on the distance.

As announced by the title of the chapter, we are to start from Stein's method to get to the Stein-Dirichlet-Malliavin structure. The chapter is organized as follows; the first leg of our journey, by stating the principle of Stein's method, will be an opportunity to underscore its two identified main steps. We stop at Dirichlet theory in the second section and explain why the conversion of the initial problem means the existence of an underlying Dirichlet structure on both initial and target spaces. In the third section, we make a detour via Malliavin calculus. We show then its relevance to tract the second step in Stein's method (fourth section). The last section is devoted to detail one of the abundant results stated by following this path: a Stein-Malliavin criterion on the Gaussian space.

### 1.1 Stein's method principle

This first section is charged with bringing to light the efficiency of Stein's method in the estimation of distances between two probability measures.

#### 1.1.1 Taxonomy on probability metrics

In order to provide a mathematical sense to the notion "of proximity" between two probability measures, we give in this first part some elements of classification of probability metrics. The framework is that of two probability measures  $\mathbf{P}^*$  and  $\mathbf{Q}$ , respectively the *target* and the *initial* measures, defined on a same metric space  $(\mathfrak{F}, c)$ . The associated Borel  $\sigma$ -field is denoted  $\mathfrak{B}(\mathfrak{F})$ . It appears that we can basically pigeonhole the numerous existing metrics between probability measures in three classes.

The first distance we introduce, which is capital since its associated topology is precisely the one of the convergence in distribution is called *Prokhorov distance*:

$$\text{dist}_{\text{Pro}}(\mathbf{P}^*, \mathbf{Q}) = \inf\{\varepsilon > 0 : \mathbf{P}^*(A) \leq \mathbf{Q}(A^\varepsilon) + \varepsilon \text{ and } \mathbf{Q}(A) \leq \mathbf{P}^*(A^\varepsilon) + \varepsilon, \forall A \in \mathfrak{B}(\mathfrak{F})\},$$

where  $A^\varepsilon$  stands for the  $\varepsilon$ -neighborhood of  $A$  and is defined by  $A^\varepsilon = \{y \in \mathfrak{F} : \exists x \in A, d(x, y) \leq \varepsilon\}$ . The set  $\mathfrak{M}_1(\mathfrak{F})$  of probability measures on  $\mathfrak{F}$  is usually equipped with the weak convergence generated by the semi-norms

$$p_f(\mathbf{P}^*) = \left| \int_{\mathfrak{F}} f d\mathbf{P}^* \right|,$$

for any function  $f$  bounded and continuous from  $\mathfrak{F}$  to  $\mathbf{R}$ . The separability of  $\mathfrak{F}$  enables to find a countable family of bounded continuous functions  $\{f_n, n \in \mathbf{N}^*\}$  which generates the Borel  $\sigma$ -field on  $\mathfrak{F}$  and the topology of the weak convergence is metrizable by considering the distance:

$$\rho(\mathbf{P}^*, \mathbf{Q}) = \sum_{n=1}^{\infty} 2^{-n} \psi(p_{f_n}(\mathbf{P}^* - \mathbf{Q})) = \text{dist}_{\text{Pro}}(\mathbf{P}^*, \mathbf{Q}),$$

where the function  $\psi$  is defined on  $\mathbf{R}$  by  $\psi(x) = x/(1+x)$ . This definition is not prone to calculations in practice such that the possibility to propose alternatives has been investigated.

The second category of probability metrics are expressed as  $u$ -divergences which definition follows:

**Definition 1.1.1.** Let  $u$  be a convex function on  $\mathbf{R}$  such that  $u(1) = 0$ , and  $\mathbf{P}^*$  and  $\mathbf{Q}$  be two measures defined on the same Polish space  $\mathfrak{F}$ . The  $u$ -divergence of  $\mathbf{Q}$  with respect to  $\mathbf{P}^*$  is defined by:

$$\mathfrak{D}^u(\mathbf{P}^* || \mathbf{Q}) = \begin{cases} \int_{\mathfrak{F}} u\left(\frac{d\mathbf{P}^*}{d\mathbf{Q}}\right) d\mathbf{Q} & \text{if } \mathbf{P}^* \ll \mathbf{Q}, \\ \infty & \text{otherwise.} \end{cases}$$

Total variation between absolutely continuous measures is obtained when  $u : x \mapsto |x - 1|$ , whereas by letting  $u : x \mapsto x \log(x)$  we get the Kullback-Leibler distance which is of key importance in information theory. Besides, we will meet once more the latter distance when estimating the additional utility of an insider who holds extra... information (see part II).

The last class we deal with is related to the optimal transportation theory. The Monge-Kantorovitch problem between  $\mathbf{P}$  and  $\mathbf{Q}$  translates into the determination of

$$\inf_{\gamma \in \Gamma(\mathbf{P}^*, \mathbf{Q})} \int_{\mathfrak{F} \times \mathfrak{F}} \mathfrak{c}(x, y) d\gamma(x, y), \quad (1.1.1)$$

where  $\Gamma(\mathbf{P}^*, \mathbf{Q})$  denotes the space of probability measures on  $\mathfrak{F} \times \mathfrak{F}$  with first marginal  $\mathbf{P}^*$  and second marginal  $\mathbf{Q}$ , and  $\mathfrak{c}$  is a cost function, semi-continuous from  $\mathfrak{F} \times \mathfrak{F}$  to  $\mathbf{R}_+ \cup \{\infty\}$ . To connect it with the notion of probabilistic metric, we introduce, by letting  $\mathfrak{c} = c^p$ , where  $c$  is a distance on  $\mathfrak{F}$  and  $p$  a positive real number, the  $p$ -th *Wasserstein distance*  $\text{dist}_{W_p}$  between the measures  $\mathbf{P}^*$  and  $\mathbf{Q}$  by

$$\text{dist}_{W_p}(\mathbf{P}^*, \mathbf{Q}) = \left( \inf_{\gamma \in \Gamma(\mathbf{P}^*, \mathbf{Q})} \int_{\mathfrak{F} \times \mathfrak{F}} c^p(x, y) d\gamma(x, y) \right)^{1/p}.$$

By taking  $p = 1$ , the underlying metric  $c$  on space  $\mathfrak{F}$  can be viewed through the optimal transportation problem as a cost function, which argument of the minimum is realized by the best "coupling" between the measures  $\mathbf{P}^*$  and  $\mathbf{Q}$  or, from this point of view, their distance. The compactness of  $(\mathfrak{F}, c)$  ensures, via the duality theorem of Kantorovitch and Rubinstein (1958), or more precisely its extension to separable metric spaces by Dudley (see [45]) that

$$\text{dist}_{W_1}(\mathbf{P}^*, \mathbf{Q}) = \inf_{\gamma \in \Gamma(\mathbf{P}^*, \mathbf{Q})} \int_{\mathfrak{F} \times \mathfrak{F}} c(x, y) d\gamma(x, y) = \sup_{h \in \text{Lip}_1} \left\{ \left| \int_{\mathfrak{F}} h d\mathbf{P}^* - \int_{\mathfrak{F}} h d\mathbf{Q} \right| \right\},$$

where  $\text{Lip}_1$  is the set  $\text{Lip}_1 = \{h \in \mathbf{R}^{\mathfrak{F}} : |h(x) - h(y)| \leq c(x, y), \forall (x, y) \in \mathfrak{F}^2\}$ . That makes the 1-Wasserstein distance the dual representation of (1.1.1). Besides Theorem 11.3.1 of [45] states that the distances  $\text{dist}_{W_1}$  and  $\rho$  yield the same topology, and legitimizes the choice to retain this definition. In the case where  $\mathfrak{F} = \mathbf{R}^n$ , for technical reasons, it is often assumed that the test functions are more regular than simply Lipschitz continuous and we are led to compute

$$\mathfrak{d}_k(\mathbf{P}^*, \mathbf{Q}) = \sup_{h \in \text{Lip}_k} \left| \int_{\mathfrak{F}} h d\mathbf{P}^* - \int_{\mathfrak{F}} h d\mathbf{Q} \right|,$$

where  $\text{Lip}_k$  is a space included in  $\text{Lip}_1$  like the set of  $k$ -times differentiable functions with derivatives up to order  $k$  bounded by 1. Note that  $\mathfrak{d}_1$  coincides with the 1-Wasserstein distance. More broadly, we can define many other probabilistic metrics in the similar fashion as to say by letting

$$\text{dist}_{\mathcal{H}}(\mathbf{P}^*, \mathbf{Q}) = \sup_{h \in \mathcal{H}} \left| \int_{\mathfrak{F}} h d\mathbf{P}^* - \int_{\mathfrak{F}} h d\mathbf{Q} \right|, \quad (1.1.2)$$

where  $\mathcal{H}$  is a class of real-valued test functions which is separating, in the sense that  $\int_{\mathfrak{F}} h d\mathbf{P}^* = \int_{\mathfrak{F}} h d\mathbf{Q}$  for all  $h \in \mathcal{H}$  if and only if  $\mathbf{Q} = \mathbf{P}^*$ . To recite nobody else but them : if the set  $\mathcal{H} = \{\mathbf{1}_A, A \in \mathfrak{B}(\mathfrak{F})\}$ ,  $\text{dist}_{\mathcal{H}}$  is the *total-variation distance*, if  $\mathfrak{F} = \mathbf{R}$  and  $\mathcal{H} = \{\mathbf{1}_{(-\infty, z]}, z \in \mathbf{R}\}$ ,  $\text{dist}_{\mathcal{H}}$  coincides with the *Kolmogorov distance*.

### 1.1.2 Stein's method principle

Stein's method turns out to be particularly suitable to give upper bounds to distances between probability measures of the form (1.1.2). In particular, when  $\mathfrak{F} = \mathbf{R}$ , it is one efficient way to compute the distance between a real-valued measure and the Gaussian distribution.

In addition to the objects introduced above, we may consider two random variables  $W$  and  $Z$  respectively distributed by  $\mathbf{Q}$  and  $\mathbf{P}^*$ . When specified, the measure  $\mathbf{P}^*$  could designate the normal distribution on  $\mathbf{R}$ . Basically, we can identify two main steps through which Stein's method seems to deploy:

1. The conversion of the difficult initial problem (1.1.2) of the approximation of a target measure  $\mathbf{P}^*$  by another one  $\mathbf{Q}$  into a simpler one with a more tractable expression.
2. The development of tools to derive an upper-bound from the control of the new expression.



## 1.2 The Stein-Dirichlet structure

### 1.2.1 Conversion of the initial problem

In its very essence, Stein's method carries the idea that the laws of two random variables are close if they verify similar identities. The first step, which aims at the conversion of the initial problem is based on this idea. It comes in three stages : the characterization of the target distribution, the resolution of the so-called Stein's equation, and the reduction of the expression (1.1.2).

By identities we mean functional identities characterizing the target laws. For instance it is well-known that the random variable  $Z$  is a standard Gaussian on  $\mathbf{R}$  if and only if  $\mathbf{E}[Zh(Z)] = \mathbf{E}[h'(Z)]$ , for any function  $h$  of a class of real-valued test functions  $\mathcal{H}$ . By this, we introduce the operator  $L$ , the so-called *Stein's operator* which acts on  $\mathcal{H}$ . For the standard Normal distribution, Stein's lemma yields such an operator:

$$\mathcal{L}h(x) = Lh(x) = xh(x) - h'(x) ; \forall x \in \mathbf{R}.$$

In order to avoid any future confusion, when it designates the operator associated to the standard Gaussian law on  $\mathbf{R}$ , the Stein operator will be noted  $\mathcal{L}$ . In other cases, we keep the notation  $L$ .

The underlying idea is the following: if the laws of the variables  $Y$  and  $Z$  are "near" each other, the quantity

$$\mathbf{E}[h(Y)] - \mathbf{E}[h(Z)]$$

may be small for a wide range of functions  $h$ . Then, if  $Z \sim \mathcal{N}(0, 1)$ , we can expect the quantity

$$\mathbf{E}[Yh(Y) - h'(Y)] = \mathbf{E}[\mathcal{L}h(Y)]$$

to be small. We come then to solve the Stein equation. i.e. to rule on the existence and the regularity of the functions  $\varphi$  satisfying :

$$L\varphi(y) = h(y) - \mathbf{E}[h(Z)] ; \forall y \in \mathfrak{F}. \quad (1.2.1)$$

The class of test functions solutions of (1.2.1) satisfying

$$h \in \mathcal{H} \iff \varphi \in \mathcal{T},$$

is called the *Stein class*. For instance, the Stein class associated to the Normal distribution and for the Wasserstein distance, denoted by  $\mathcal{T}_W$ , is the class of twice differentiable functions, whose first derivative is bounded by 1 and whose second derivative is bounded by 2.

By integrating then equation (1.2.1) with respect to  $\mathbf{Q}$  and taking the supremum over the class  $\mathcal{T}$  we get:

$$\sup_{h \in \mathcal{H}} \left| \int_{\mathfrak{F}} h d\mathbf{P}^* - \int_{\mathfrak{F}} h d\mathbf{Q} \right| = \sup_{\varphi \in \mathcal{T}} \left| \mathbf{E}[L\varphi(Y)] \right| \text{ where } Y \sim \mathbf{Q}. \quad (1.2.2)$$

Thus, the initial problem of approximation of a target measure by an initial one has been converted into a simpler expression: that of the expectation of a certain functional on the initial space. This is in fact easier to handle ; various identities and couplings can be applied advantageously.

### 1.2.2 The Dirichlet structure point of view

Actually, we can address the problem from a distinct point of view. The computation of  $\text{dist}(\mathbf{P}^*, \mathbf{Q})$  can be interpreted as the measure of the *error* made by approximating  $\mathbf{P}^*$  by  $\mathbf{Q}$ . It turns out that the difficult steps of the functional characterization of  $\mathbf{P}^*$  and the resolution of the Stein equation can be got around by following a smart-paths approach. In fact, the target measure  $\mathbf{P}^*$  can be viewed as the stationary distribution of a Markovian semi-group  $(P_t)_{t \in \mathbf{R}_+}$  and of generator  $L$ .

#### Language elements of Dirichlet structures

We give some elements of terminology of Dirichlet structures by introducing some basic notions on Markov semigroups. A substantiated presentation on the subject can be found in the books of Bouleau and Hirsch [23] or Fukushima, Oshima and Takeda [56]. In the first part of this subsection, we successively provide the definitions of *Markov semi-group*, *Markov process* and *infinitesimal generator* before focusing on the fundamental relations that connect them to each other. Consider a measurable space  $(E, \mathcal{A})$ . Let  $(P_t)_{t \in \mathbf{R}_+}$  be 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  $t \in \mathbf{R}_+$ ,  $P_t$  is a linear operator; it sends 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  $s, t \in \mathbf{R}_+$   $P_{t+s} = P_t \circ P_s$  (semi-group property).
4. For any  $t \in \mathbf{R}_+$ ,  $P_t$  conserves the mass and preserves positivity (Markov property):  $P_t(\mathbb{1}) = \mathbb{1}$  and, for any positive function  $f$ ,  $P_t f$  is positive.

**Definition 1.2.1** (Invariant measure). Let a family  $(P_t)_{t \in \mathbf{R}_+}$  of operators defined on  $(E, \mathcal{A})$  and satisfying the properties (1)-(4). A positive  $\sigma$ -finite measure  $\mu$  on  $(E, \mathcal{A})$  is *invariant* for  $(P_t)_{t \in \mathbf{R}_+}$ , if for every bounded positive measurable function  $F : E \rightarrow \mathbf{R}$  and  $t \in \mathbf{R}_+$ ,

$$\int_E P_t F \, d\mu = \int_E F \, d\mu.$$

**Definition 1.2.2** (Markov semi-group). A family  $(P_t)_{t \in \mathbf{R}_+}$  of operators defined on the set of bounded measurable functions on  $(E, \mathcal{A})$  with invariant positive  $\sigma$ -finite measure  $\mu$  satisfying the properties (1)-(4) as well as the continuity property:

5. For any  $F \in L^2(\mu) = L^2(E, \mu)$ ,  $P_t F$  converges to  $F$  in  $L^2(E)$  as  $t$  tends to 0 (continuity property).

The semi-group is *symmetric* with respect to the invariant measure  $\mu$ , if for all  $F, G$  elements of  $E$  and  $t \in \mathbf{R}_+$ ,

$$\int_E F (P_t G) \, d\mu = \int_E (P_t F) G \, d\mu.$$

**Remark 1.2.3.** Thus defined, the Markov semi-group  $(P_t)_{t \in \mathbf{R}_+}$  is a semi-group of contractions on  $L^2(\Omega)$  (for more details on the subject refer to the Hille-Yosida theory) and satisfies the property of *contraction* and *strong continuity*:

$$\|P_t F\|_{L^2(\mu)} \leq \|F\|_{L^2(\mu)} \quad (\text{contraction}) \quad \text{and} \quad \lim_{t \rightarrow 0} P_t F = F \quad (\text{strong continuity}).$$

Up to now, consider a probability space  $(\Omega, \mathcal{F}, \mathbf{P}^*)$ . Let  $E = L^2(\Omega, \mathbf{P}^*)$  and  $\mu = \mathbf{P}^*$ .

**Definition 1.2.4** (Markov process). Consider a measurable process  $(X_t^x)_{t \in \mathbf{R}_+}$  on  $(\Omega, \mathcal{F}, \mathbf{P}^*)$  starting from  $x \in E$  at time  $t = 0$ . Denote by  $(\mathcal{F}_t^X)_{t \in \mathbf{R}_+}$  the filtration generated by the process  $(X_t^x)_{t \in \mathbf{R}_+}$ . The process  $(X_t^x)_{t \in \mathbf{R}_+}$  is a Markov process if it satisfies the Markov property: for any  $(s, t) \in (\mathbf{R}_+)^2$  such that  $t > s$ , the law of  $X_t^x$  given  $\mathcal{F}_s^X$  is the same as the law of  $X_t^x$  given  $X_s^x$  as well as the law of  $X_{t-s}^x$  given  $X_0^x$ .

In view of remark 1.2.3, the Markov semi-group  $(P_t)_{t \in \mathbf{R}_+}$  can be seen as a contraction semi-group; so that its infinitesimal generator can be introduced in the setting of Hille-Yosida theory.

**Definition 1.2.5** (Infinitesimal generator). The *generator* of a contraction semi-group  $(P_t)_{t \in \mathbf{R}_+}$  on  $L^2(\Omega)$  is defined by

$$\mathbf{D}_L = \left\{ F : \lim_{t \rightarrow 0} \frac{P_t F - F}{t} \text{ exists} \right\} \quad \text{and} \quad L F = \lim_{t \rightarrow 0} \frac{P_t F - F}{t}; \quad \forall F \in \mathbf{D}_L,$$

and such that  $\mathbf{D}_L$  is dense in  $L^2(\Omega)$ .

**Remark 1.2.6.** There exists a correspondence between the Markov semi-group  $(P_t)_{t \in \mathbf{R}_+}$  (definition 1.2.2), the Markov process  $(X_t^x)_{t \in \mathbf{R}_+}$  (definition 1.2.4) and the infinitesimal generator (definition 1.2.5) in the sense that the knowledge of one of  $(P_t)_{t \in \mathbf{R}_+}$ ,  $(X_t^x)_{t \in \mathbf{R}_+}$  or  $L$  implies the existence and gives the expression of the two others.

$$\begin{array}{ccc} (X_t^x)_{t \in \mathbf{R}_+} & \xrightarrow{P_t F = \mathbf{E}[F(X_t) | X_0 = \cdot]} & (P_t)_{t \in \mathbf{R}_+} & \xrightarrow{L F = \left. \frac{dP_t F}{dt} \right|_{t=0}} & L \\ & \xleftarrow{X_t^* \mathbf{P} = (P_t)^* \nu} & & \xleftarrow{P_t F = e^{tL}} & \end{array}$$

In fact, by duality the semi-group  $(P_t)_{t \in \mathbf{R}_+}$  acts on the set of measures  $\nu$  on  $\mathcal{A}$  via the formula

$$\int_E P_t F d\nu = \int_E F d((P_t)^* \nu),$$

so that if  $\nu$  denotes the law of  $X_0$ ,  $(P_t)^* \nu$  is the law of the variable  $X_t$ .

In order to introduce the carré du champ operator as well as the Dirichlet forms, assume the existence of a vector subspace  $\mathcal{K}$  of  $\mathbf{D}_L$  such that for any  $(F, G) \in \mathcal{K} \times \mathcal{K}$ , the product  $FG$  is an element of  $\mathbf{D}_L$ .

**Definition 1.2.7** (Carré du champ operator). The bilinear map defined on  $\mathbf{D}_\Gamma = \mathcal{K} \times \mathcal{K}$  by

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

is the *carré du champ operator* associated to the Markov generator  $L$ .

We can now transpose all this framework into the Dirichlet forms language. Consider a Markov semi-group  $(P_t)_{t \in \mathbf{R}_+}$ , symmetric with respect to a finite  $\sigma$ -measure  $\mu = \mathbf{P}^*$ , with infinitesimal generator  $L$  and carré du champ operator  $\Gamma$ .

**Definition 1.2.8.** Under previous assumptions, the energy function  $\mathcal{E}(F)$  is defined for any random variable  $F \in L^2(\mu)$  such that the limit

$$\lim_{t \rightarrow 0} \frac{1}{t} \int_{\mathbf{E}} F(F - P_t F) d\mu$$

exists; that shapes the domain  $\mathbf{D}_{\mathcal{E}}$  of  $\mathcal{E}$ . The Dirichlet form  $\mathcal{E}$  (similarly denoted for the purposes of notation) is the bilinear symmetric operator on  $\mathbf{D}_{\mathcal{E}}$  defined by

$$\mathcal{E}(F, F) = \mathcal{E}(F) \quad \text{and} \quad \mathcal{E}(F, G) = \int_{\mathbf{E}} \Gamma(F, G) d\mu; \quad \forall (F, G) \in \mathbf{D}_{\mathcal{E}} \times \mathbf{D}_{\mathcal{E}}. \quad (1.2.3)$$

**Remark 1.2.9.** The possibility to define  $\mathcal{E}$  in the fashion (1.2.3) is consequent of the integration by parts formula satisfied for any  $(F, G) \in \mathbf{D}_{\Gamma} \times \mathbf{D}_{\Gamma}$  by

$$\int_{\mathbf{E}} \Gamma(F, G) d\mu = \int_{\mathbf{E}} F(-LG) d\mu = \int_{\mathbf{E}} (-LF)G d\mu. \quad (1.2.4)$$

and the relations connecting the Markov semi-group  $(P_t)_{t \in \mathbf{R}_+}$  and its generator  $L$  (see remark 1.2.6). Nevertheless, the domain of the Dirichlet form  $\mathcal{E}$  is bigger than  $\mathbf{D}_{\Gamma}$ . The identity  $\mathcal{E}(F) = \mathcal{E}(F, F)$  ensures that  $\mathcal{E}$  can be defined on a set  $\mathbf{D}_{\mathcal{E}}$  such that  $\mathbf{D}_{\Gamma} \subset \mathbf{D}_{\mathcal{L}} \subset \mathbf{D}_{\mathcal{E}} \subset L^2(\Omega) = \mathbf{E}$ .

Within this definition, we can verify that  $\mathcal{E}$  thus defined is a non-negative definite symmetric form on  $L^2(\Omega)$  which domain  $\mathbf{D}_{\mathcal{E}}$  is dense in  $L^2(\Omega)$ . It is moreover *closed* in the sense that  $\mathbf{D}_{\mathcal{E}}$  equipped with the norm

$$\|F\|_{\mathbf{D}_{\mathcal{E}}} = \left( \|F\|_{L^2(\Omega)}^2 + \mathcal{E}(F) \right)^{\frac{1}{2}},$$

is a Hilbert space.

**Theorem 1.2.10.** *The generator of a closed form  $\mathcal{E}$  is defined by*

$$\mathbf{D}_{\mathcal{L}} = \{F \in \mathbf{D}_{\mathcal{E}} : \exists G \in L^2(\Omega), \forall H \in \mathbf{D}_{\mathcal{E}}, \mathcal{E}(F, H) = -\langle G, H \rangle_{L^2(\Omega)}\} \quad \text{and} \quad LF = G.$$

Moreover, if  $L$  is a negative self-adjoint operator and  $\mathbf{D}_{\mathcal{E}} = \mathbf{D}_{\sqrt{-L}}$ , then

$$\mathcal{E}(F) = \|\sqrt{-L}F\|_{L^2(\Omega)}^2; \quad \forall F \in \mathbf{D}_{\mathcal{E}}. \quad (1.2.5)$$

**Remark 1.2.11.** There is a one-to-one correspondence between the family of closed symmetric forms on  $L^2(\Omega)$  and the family of non-positive definite self-adjoint operators on  $L^2(\mathbf{E})$  and which is determined through Theorem 1.2.10.

**Definition 1.2.12** (Dirichlet form and Dirichlet structure). A *Dirichlet form* on  $(\Omega, \mathcal{F}, \mathbf{P})$  is a symmetric closed form  $\mathcal{E}$ , *Markovian* in the following sense :

$$F \in \mathbf{D}_{\mathcal{E}} \Rightarrow F \wedge 1 \in \mathbf{D}_{\mathcal{E}} \quad \text{and} \quad \mathcal{E}(F \wedge 1, F \wedge 1) \leq \mathcal{E}(F, F).$$

The quintuplet  $(\Omega, \mathcal{F}, \mathbf{P}, \mathbf{D}_{\mathcal{E}}, \mathcal{E})$  is a *Dirichlet structure*.

**Example 1.2.13.** The quintuplet  $(\mathbf{R}, \mathcal{B}(\mathbf{R}), \mathbf{P}, \mathbf{D}_{\mathcal{E}}, \mathcal{E})$  where  $\mathbf{P}$  denotes the standard Gaussian measure on  $\mathbf{R}$ ,

$$\mathbf{D}_{\mathcal{E}} = \{u \in L^2(\Omega) : u'_w \in L^2(\Omega)\} \quad \text{and} \quad \mathcal{E}(F) = \mathbf{E} [ |F'|^2 ],$$

( $u'_w$  stands for the derivative taken in the sense of distributions) defines the 1-dimensional Ornstein-Uhlenbeck structure.

**Definition 1.2.14** (Carré du champ operator). Let  $(P_t)_{t \in \mathbf{R}_+}$  a symmetric semi-group of generator  $L$  and  $\mathcal{E}$  a Dirichlet form defined as in (1.2.5). The *carré du champ operator*  $\Gamma$  associated to  $\mathcal{E}$  is defined on  $\mathbf{D}_{\mathcal{E}} \times \mathbf{D}_{\mathcal{E}}$  by

$$\Gamma(F, G) = \frac{1}{2} (L(FG) - G LF - F LG).$$

The quintuplet  $(\Omega, \mathcal{F}, \mathbf{P}, \mathbf{D}_{\mathcal{E}}, \Gamma)$  is an *error structure*.

**Example 1.2.15.** Within the notations of example (1.2.13),  $\Gamma(F, G) = F'G'$ .

### Dirichlet structures: a solution to the initial problem

We can now address the initial problem in terms of Dirichlet structures. As evokated above, there exists a strong ergodic Markov process  $(X_t^x)_{t \in \mathbf{R}_+}$  of invariant measure  $\mathbf{P}^*$  and generator  $L$ . These ones are connected by the relations

$$L\varphi(x) = \left. \frac{dP_t\varphi(x)}{dt} \right|_{t=0} \quad \text{and} \quad P_t\varphi(x) = \mathbf{E} [\varphi(X_t^x) | X_0^x = x]. \quad (1.2.6)$$

**Example 1.2.16** (Gaussian measure on  $\mathbf{R}$ ). Let  $\mu$  denote the standard Gaussian on  $E = \mathbf{R}$ , then  $X = (X_t)_{t \in \mathbf{R}_+}$  is the Ornstein-Uhlenbeck process defined by

$$dX_t = -X_t + \sqrt{2} dB_t \quad \text{and} \quad X_0 = x,$$

where  $(B_t)_{t \in \mathbf{R}_+}$  is a standard one-dimensional Brownian motion. It can be shown, via the Mehler representation formula, that the semi-group  $(P_t)_{t \in \mathbf{R}_+}$  is defined by

$$P_t\varphi(x) = \int_{\mathbf{R}} \varphi \left( e^{-t}x + \sqrt{1 - e^{-2t}}y \right) d\mu(y). \quad (1.2.7)$$

Its generator, the operator  $L$ , satisfies the reminding relation: for  $\varphi \in \mathcal{C}_b^2(\mathbf{R}; \mathbf{R})$ ,

$$L\varphi(x) = x\varphi'(x) - \varphi''(x); \quad \forall x \in \mathbf{R}.$$

In the general case, using (1.2.6) and noting that  $P_\infty\varphi = \mathbf{E}_{\mathbf{P}^*}[\varphi]$ ,  $P_0\varphi = \varphi$ , and taking the expectation with respect to  $\mathbf{Q}$  of the two members of the equality

$$P_\infty\varphi(\omega) - P_0\varphi(\omega) = \int_0^\infty \frac{dP_t}{dt}\varphi(\omega) dt,$$

we can state the so-called *Stein representation formula* or *smart-path formula*

$$\mathbf{E}_{\mathbf{P}^*}[\varphi] - \mathbf{E}_{\mathbf{Q}}[\varphi] = \int_{\mathfrak{F}} \int_0^\infty L(P_t\varphi) dt d\mathbf{Q} \quad (1.2.8)$$

It turns out that the reduction of the problem in the fashion of (1.2.2) which implies to characterize the target measure through a functional identity and to solve the Stein equation can be done here without all this work. In fact we are to bound

$$\mathbf{E}_{\mathbf{Q}} \left[ \int_0^\infty L(P_t \varphi) dt \right]. \quad (1.2.9)$$

In other words, this "smart-path"-like method gives another interpretation to the computation of distances between measures via Stein's method: that of the estimation of the distance between the underlying ergodic Markov process to its stationary measure. This will be illustrated in Subsection 1.3.2 through the example of the Normal approximation by a Poisson distribution (in dimension 1).

**Remark 1.2.17.** By reminiscence of Dirichlet forms vocabulary, we can interperate  $P_\infty \varphi - P_0 \varphi$  as the error made in the approximation of the measure  $\mathbf{P}^*$  by  $\mathbf{Q}$ . In this same setting (see Bouleau and Hirsch [23]) the carré du champ operator  $\Gamma$  (resp. the symmetric operator  $L$ ) represents the variance (resp. the bias) of the error. Thus, some convergence results proved by Stein's method can be interpreted in the language of error calculus. As the matter of fact, one of the existing proofs of the famous *Fourth Moment Theorem* stated by David Nualart and Giovanni Peccati (see [100]) lies on the existence of a deterministic constant  $c$  such that

$$\text{var}(\Gamma(F)) \leq c(\mathbf{E}[F^4] - 3\mathbf{E}[F^2]^2),$$

where  $\Gamma$  is the carré du champ operator associated with the generator of the Ornstein-Uhlenbeck semi-group and  $F$  a random variable in a Wiener chaos. The use of the Stein-Malliavin method (see Nourdin and Peccati [93]) yields another proof of the theorem and provides a estimation in the total variation distance w.r.t. the Normal distribution  $\mathbf{P}^*$

$$\text{dist}_{\text{TV}}(F, \mathbf{P}^*) \leq \frac{2}{\sqrt{3}} \sqrt{\mathbf{E}[F^4] - 3}.$$

The reader can refer to the work of Louis H.Y. Chen and Guillaume Poly (see [28]) for a statement of the links between Stein's method, Dirichlet forms and Malliavin's calculus through the Fourth Moment theorem.

## 1.3 Dirichlet-Malliavin structures

### 1.3.1 Some words about Malliavin calculus

As recalled in the introduction, the Malliavin calculus is a finite-dimensional stochastic variational calculus initially developed on the Wiener space.

One way to introduce the tools of Malliavin calculus is to transpose the terminology of classical differential calculus in Banach spaces to Malliavin's one. We choose to give the analogues for three processes families for which the theories are the most complete: the standard Brownian motion  $(B_t)_{t \in [0,1]}$ , the standard Poisson process  $(N_t)_{t \in \mathbf{R}_+}$  and a Rademacher process  $(X_n)_{n \in \mathbf{N}^*}$ . In any case, we assume the existence of a probability space  $(\Omega, \mathcal{A}, \mathbf{P})$  on which each process is defined. We denote the Hilbert space  $\mathfrak{H} = L^2(\mathbf{T})$  where  $\mathbf{T}$  is the parameter space:  $[0, 1]$  for the Brownian motion  $B$ ,  $\mathbf{R}_+$  for the Poisson point process  $N$  and  $\mathbf{N}^*$  for the Rademacher process  $X$ .

Differential Calculus on $(\mathbf{R}^n, \langle \cdot, \cdot \rangle)$	Malliavin Calculus on Wiener/Poisson/Rademacher spaces
Vectors $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$	Paths of the process ( $\omega \in \Omega$ ) $t \mapsto B_t(\omega), t \mapsto N_t(\omega), n \mapsto X_n(\omega)$
Functions defined on $\mathbf{R}^n$	Random variables = Functionals of the paths
Gradient $\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$	Malliavin derivative $\nabla_h F = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X(h_1), \dots, X(h_n)) h_i$ $D_t F(N) = F(N + \delta_t) - F(N)$ $D_k F(X_1, \dots, X_n) = \sqrt{p_k q_k} \times$ $F(X_1, \dots, +1, \dots, X_n) - F(X_1, \dots, -1, \dots, X_n)$
Integration by parts $\int_{\mathcal{D}} (\Delta f) g dx = - \int_{\mathcal{D}} \langle \nabla f, \nabla g \rangle + \int_{\partial \mathcal{D}} \frac{\partial f}{\partial n} g d\sigma$	Malliavin IBP $\mathbf{E}[F \delta U] = \mathbf{E}[\langle DF, U \rangle_{\mathfrak{H}}]$ where $\delta$ is the adjoint of $D$

We briefly present the main facts about Malliavin calculus on the Wiener space before giving some complementary elements in the more general case of Gaussian spaces. A more detailed discussion can be found in Nualart [98], or Üstünel and Zakai [137]. We hold the construction of Malliavin calculus for the jump processes in the second part of the manuscript as a motivation for our second work.

### The historical Malliavin calculus: the Wiener space case

As mentioned in the introduction, two approaches were developed to lead to a differential calculus on the Wiener space: a variational approach and a chaotic approach. In the first one, the derivative is defined for a class of smooth variables and then extended to the space of square integrable Wiener functionals by density. This can be interpreted as a directional derivative but only in certain directions: that of Cameron-Martin's space. In an alternative construction, the Malliavin derivative is obtained as an annihilation operator acting on square integrable random variables, for which a preliminary chaotic decomposition has been stated. The operators defined through these two different approaches coincide in the Gaussian space; this equivalence is due in particular to the intrinsic relation between the standard Brownian motion and the Hermite polynomials.

Consider the classical Wiener space supported by the Banach space  $\mathscr{W}_B = \mathcal{C}^0([0, 1]; \mathbf{R})$  and equipped with the Wiener measure  $\mathbf{P}$ . Let  $B = (B_t)_{t \in [0, 1]}$  the standard Brownian motion defined on the probability space  $(\mathscr{W}_B, \mathcal{A}_B, \mathbf{P})$ . Consider the family

$$\mathcal{E} = \{t \wedge \cdot, t \in [0, 1]\},$$

such that the closure with respect to the norm  $\|\cdot\|_{\mathfrak{H}^1}$  induced by the scalar product

$$\langle f, g \rangle_{\mathfrak{H}^1} = \langle \dot{f}, \dot{g} \rangle_{L^2(\mathbf{T})},$$

is a Hilbert space, called the *Cameron-Martin space* defined by

$$\mathfrak{H}^1 = \left\{ f : [0, 1] \rightarrow \mathbf{R}, \exists \dot{f} \in L^2(\mathbf{T}) \text{ with } f(t) = \int_0^t \dot{f}(s) ds \right\} \text{ and } \|f\|_{\mathfrak{H}^1} = \|\dot{f}\|_{L^2(\mathbf{T})}. \quad (1.3.1)$$

Define the mapping

$$\mathbf{W} : \begin{array}{l} \mathcal{E} \longrightarrow \mathscr{W}_B \\ h \longmapsto \int_0^1 \dot{h}(s) dB_s \end{array}$$

Within  $\mathbf{W}$ , the mapping  $\mathbf{1}_{[0, t]} \mapsto B_t$  naturally extend to an isometry between  $\mathfrak{H}^1$  and the space  $L^2(\mathscr{W}_B, \mathbf{P})$ . In particular we can retrieve the canonical process as the Wiener integral

$$B(t) = \mathbf{W}(t \wedge \cdot); t \in \mathbf{T}. \quad (1.3.2)$$

The process  $(W(h))_{h \in \mathfrak{H}^1}$  defined by

$$W(h) = \mathbf{W}(h); h \in \mathfrak{H}^1,$$

is an *isonormal Gaussian process* associated with the Hilbert space  $\mathfrak{H}^1$ .

The aim is to define the derivative  $\nabla F$  of a square integrable random variable  $F : \mathscr{W}_B \rightarrow \mathbf{R}$ . A natural idea is to explore the possibility to define it as a Fréchet derivative

$$\lim_{\varepsilon \rightarrow 0} \frac{F(\omega + \varepsilon \omega') - F(\omega)}{\varepsilon}.$$

Consider the stochastic differential equation defined for any  $t \in [0, T]$  by

$$X_t = x + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dB_s, \quad (1.3.3)$$



where  $b, \sigma \in \mathcal{C}^0(\mathbf{R})$ , and  $x \in \mathbf{R}$ . The application  $I : \mathbf{R} \times \mathcal{C}^0(\mathbf{R}^2, \mathbf{R}) \times \Omega \rightarrow \mathcal{C}^0([0, T], \mathbf{R})$  which associates  $(x, (b, \sigma), \omega)$  with - if it exists - the unique solution  $X$  of (1.3.3) is called the *Itô map*. This is not continuous on  $\mathscr{W}_B$  and a fortiori not Fréchet differentiable. The Cameron-Martin theorem which states that for two random variables  $F, G : \Omega \rightarrow \mathbf{R}$

$$F = G \text{ a.s.} \implies F(B + h) = G(B + h) ; \forall h \in \mathfrak{H}^1,$$

is prone to explore the possibility to narrow the set of admissible differentiation directions to a smaller one, the Cameron-Martin's space  $\mathfrak{H}^1$ . Let  $\mathcal{S}$  denote the class of smooth random variables  $F$  of the form

$$F = f(W(h_1), \dots, W(h_n))$$

where  $f$  belongs to the Schwartz space  $\mathcal{S}(\mathbf{T}^n)$  and  $h_1, \dots, h_n$  are elements of  $\mathfrak{H}^1$ . The derivative of  $F$  is the element of  $\mathfrak{H}^1$  defined by

$$\nabla F = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(W(h_1), \dots, W(h_n)) h_k, \quad (1.3.4)$$

which can be interpreted as a directional derivative:

$$\langle \nabla F, h \rangle_{\mathfrak{H}^1} = \lim_{\varepsilon \rightarrow 0} \frac{f(W(h_1) + \varepsilon \langle h_1, h \rangle_{\mathfrak{H}^1}, \dots, W(h_n) + \varepsilon \langle h_n, h \rangle_{\mathfrak{H}^1}) - f(W(h_1), \dots, W(h_n))}{\varepsilon}$$

The map  $\nabla$  is closable from  $L^2(\mathscr{W}_B)$  to  $L^2(\mathscr{W}_B; \mathfrak{H}^1) = L^2(\mathscr{W}_B \times \mathbf{T})$ . Thus, it is meaningful to define  $\mathbf{D}_B^{1,2}$  as the closure of cylindrical functions for the norm

$$\|F\|_{1,2} = \left( \|F\|_{L^2(\mathscr{W}_B)}^2 + \|\nabla F\|_{L^2(\mathscr{W}_B; \mathfrak{H}^1)}^2 \right)^{\frac{1}{2}}.$$

The derivative operator satisfies the useful *chain rule*.

**Theorem 1.3.1** (Chain rule). *Let  $\varphi \in (\mathbf{R}^n)$ , and a random vector  $F = (F_1, \dots, F_n)$  which components belong to  $\mathbf{D}^{1,q}$  for some  $q \in \mathbf{N}^*$ . Then  $\varphi(F)$  belongs to  $\mathbf{D}^{1,q}$  and*

$$\nabla(\varphi F) = \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i} \mathbf{D}F_i. \quad (1.3.5)$$

In the same way, it is possible to define for  $p \geq 1$ ,  $F \in \mathcal{S}$ , the  $p$ -th Malliavin derivative of  $F$  as the element of  $L^2(\mathscr{W}_B; (\mathfrak{H}^1)^{\odot p})$  ( $(\mathfrak{H}^1)^{\odot p}$  is the  $p$ -th symmetric tensor product of  $\mathfrak{H}^1$ ) defined by

$$\nabla^p F = \sum_{(i_1, \dots, i_p) \in [n]^p} \frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}}(W(h_1), \dots, W(h_n)) h_{i_1} \otimes \dots \otimes h_{i_p},$$

where  $[n] = \{1, \dots, n\}$ . For any  $q \in [1, \infty)$  and  $p \in \mathbf{N}^*$ , the set  $\mathbf{D}_B^{p,q}$  denotes the closure of  $\mathcal{S}$  with respect to the norm

$$\|F\|_{p,q} = \left( \mathbf{E}[|F|^q] + \mathbf{E}[\|\nabla F\|_{L^q(\mathfrak{H}^1)}^q] + \dots + \mathbf{E}[\|\nabla^p F\|_{L^q(\mathfrak{H}^{\otimes p})}^q] \right)^{\frac{1}{q}}$$

**Theorem 1.3.2** (Integration by parts formula). *For any cylindrical random variable  $F$  and  $h$  element of  $\mathfrak{H}^1$ ,*

$$\mathbf{E}[\langle \nabla F, h \rangle_{\mathfrak{H}^1}] = \mathbf{E}[F W(h)] \quad (1.3.6)$$

The second operator in stake called *divergence operator* is defined as the adjoint of the derivative operator.

**Definition 1.3.3.** The adjoint of the derivative operator  $\nabla$ , denoted  $\delta$  is the unbounded operator on  $L^2(\mathscr{W}_B; \mathfrak{H}_1)$  which domain is the set of  $\mathfrak{H}_1^1$ -valued square integrable random variables  $u$  such that

$$|\mathbf{E} [\langle \nabla F, u \rangle_{\mathfrak{H}_1}]| \leq c_u \|F\|_{L^2(\mathscr{W}_B)},$$

for all  $F$  element of  $\mathbf{D}_B^{1,2}$ . For any  $u \in \text{Dom } \delta$ ,  $\delta(u)$  is the element of  $L^2(\mathscr{W}_B)$  characterized by

$$\mathbf{E} [F \delta(u)] = \mathbf{E} [\langle \nabla F, u \rangle_{\mathfrak{H}_1}] ; \forall F \in \mathbf{D}_B^{1,2}. \quad (1.3.7)$$

**Remark 1.3.4.** The divergence operator coincides with the Wiener integral on  $\mathfrak{H}_1^1$  as highlighted in (1.3.6). Moreover, the special feature of Malliavin integration by parts is to include the adjoint of the derivative operator.

The Malliavin calculus on the Wiener space can be elaborated in a different but equivalent way; as a reminder, the family of *Hermite polynomials*  $\{H_n, n \in \mathbf{N}\}$  is such that the associated polynomial functions satisfy

$$H_0 = \mathbb{1} \quad \text{and} \quad H'_n = nH_n, \quad n \in \mathbf{N}^*.$$

The so-called *Wiener chaos* are designed in the following fashion;  $\mathcal{H}_0$  is the set of constants and for any  $n \geq 1$ , the  $n$ -th Wiener chaos, denoted  $\mathcal{H}_n$  linear subspace of  $L^2(\mathscr{W}_B, \mathcal{A}_B, \mathbf{P})$  generated by the random variables  $\{H_n(W(h)), h \in \mathfrak{H}_1^1\}$  such that the remarkable *chaotic decomposition* can be stated:

$$L^2(\mathscr{W}_B, \mathcal{A}_B, \mathbf{P}) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n$$

Note in particular that  $\mathcal{H}_1 = \{W(h), h \in \mathfrak{H}_1^1\} = W$ . Consequence of this decomposition, any random variable  $F \in L^2(\mathscr{W}_B)$ , admits an unique expansion of the form

$$F = \mathbf{E} [F] + \sum_{n \in \mathbf{N}} J_n(f_n), \quad (1.3.8)$$

where  $f_n \in L^2(\mathbf{T}^n, \mathcal{B}^n, \mu^{\otimes n})$  (equal to  $(n!)^{-1} \mathbf{E} [\nabla^{(n)} F]$  if  $F \in \mathbf{D}^{n,2}$ ) and the multiple stochastic integral of a function  $f \in L^2(\mathbf{T}^n, \mathcal{B}^n, \mu^{\otimes n})$  of the form

$$f = \sum_{(i_1, \dots, i_n) \in [m]^n} a_{i_1 \dots i_n} \mathbf{1}_{A_{i_1} \times \dots \times A_{i_n}}(t_1, \dots, t_n),$$

with  $A_1, A_2, \dots, A_n$  pairwise-disjoint sets belonging to  $\mathcal{B}_0$ , and the coefficients  $a_{i_1 \dots i_n}$  are zero if any two of the indices  $i_1, \dots, i_n$  are equal, is defined by

$$J_n(f) = \sum_{(i_1, \dots, i_n) \in [m]^n} a_{i_1 \dots i_n} W(A_{i_1}) \cdots W(A_{i_n})$$

The derivative operator  $\nabla$  coincides on  $\mathcal{S}$  with the *annihilation operator* (in reference to its usual name in the Fock space theory met in quantum probability)  $\tilde{\nabla}$  such that

$$\tilde{\nabla}_t J_n(f_n) = n J_{n-1}(f_n(\star, t)),$$

for any  $f_n$  element of  $L^2([0, 1])^{\otimes n}$  and  $t \in [0, 1]$ . Its adjoint  $\tilde{\delta}$ , called *creation operator* is such that

$$H_n = \tilde{\delta}^n(\mathbb{1}).$$

**Remark 1.3.5.** The very close connection between the standard Brownian motion and the Hermite polynomials is at the heart of the equivalence of variational and chaotic approaches. Indeed, the Wiener chaos are intrinsically produced by Hermite polynomials; on the other hand, the first Hermite polynomial subtly appears in the Cameron-Martin theorem

$$\mathbf{E}[F(B + th)] = \mathbf{E}\left[F e^{tW(h) - \frac{1}{2}t^2\|h\|_{\mathfrak{H}_1}^2}\right] = \mathbf{E}\left[F e^{tH_1(W(h)) - \frac{1}{2}t^2\|h\|_{\mathfrak{H}_1}^2}\right],$$

and gives rise to the integration by parts formula.

### The Malliavin calculus on infinite-dimensional Gaussian fields

A Malliavin calculus can be constructed in the same fashion on infinite-dimensional Gaussian fields i.e. for families  $Y = \{Y_t, t \in \mathbf{T}\}$  ( $\mathbf{T}$  is a parameter set) of real random variables defined on a probability space  $(\Omega, \mathcal{A}, \mathbf{P})$  such that any linear combination of the variables is a centered Gaussian variable. To encrypt the properties of a given Gaussian field  $Y$ , we consider its associated Gaussian subspace  $\mathfrak{H} \subset L^2(\Omega, \mathcal{A}, \mathbf{P})$  generated by the equivalence classes of  $Y$ . Then, Kolmogorov's theorem ensures the existence of an *isonormal Gaussian process* over  $\mathfrak{H}$ , denoted  $X$ ; that is a centered Gaussian family  $X = \{X(h), h \in \mathfrak{H}\}$  defined on  $(\Omega, \mathcal{A}, \mathbf{P})$  and which covariance structure is encoded by the inner product of  $\mathfrak{H}$  via the identity  $\mathbf{E}[X(h)X(g)] = \langle g, h \rangle_{\mathfrak{H}}$ , ( $g, h \in \mathfrak{H}$ ). We set  $\mathcal{F} = \sigma(X)$  and assume that  $\mathcal{A} = \mathcal{F}$ .

In this framework, the space  $\mathcal{S}$  of cylindrical random variables is composed of the variables  $F$  of the form

$$F = f(X(h_1), \dots, X(h_n)),$$

where  $f$  belongs to the Schwartz space  $\mathcal{S}(\mathbf{R}^n)$  and  $h_1, \dots, h_n$  are elements of  $\mathfrak{H}$ . The space  $\mathcal{S}$  is dense in  $L^q(\Omega)$  for any  $q \in \mathbf{N}^*$ .

For any  $p \in \mathbf{N}^*$  and  $F$  element of  $\mathcal{S}$ , the  $p$ -th Malliavin derivative of  $F$  (with respect to  $X$ ) is the element of  $L^2(\Omega, \mathfrak{H}^{\odot p})$  ( $\mathfrak{H}^{\odot p}$  is the  $p$ -th symmetric tensor product of  $\mathfrak{H}$ ) defined by

$$D^p F = \sum_{(i_1, \dots, i_p) \in [n]^p} \frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}}(X(h_1), \dots, X(h_n)) h_{i_1} \otimes \dots \otimes h_{i_p}.$$

For any  $q \in [1, \infty)$  and  $p \in \mathbf{N}^*$ , the set  $\mathbf{D}^{p,q}$  denotes the closure of  $\mathcal{S}$  with respect to the norm

$$\|F\|_{p,q} = \left( \mathbf{E}[|F|^q] + \mathbf{E}[\|DF\|_{\mathfrak{H}}^q] + \dots + \mathbf{E}[\|D^p F\|_{\mathfrak{H}^{\otimes p}}^q] \right)^{\frac{1}{q}}$$

The analogues of the *Wiener chaos* in this setting are designed in the following fashion;  $\mathcal{H}_0$  is the set of constants and for any  $n \in \mathbf{N}^*$ , the  $n$ -th Wiener chaos, denoted  $\mathcal{H}_n$  linear subspace of  $L^2(\Omega, \mathcal{A}, \mathbf{P})$  generated by the random variables  $\{H_n(X(h)), h \in \mathfrak{H}\}$ . Thus any random variable  $F \in L^2(\Omega)$  can be expanded in a unique way

$$F = \mathbf{E}[F] + \sum_{n \geq 1} \text{Proj}_{\mathcal{H}_n}(F),$$

where  $\text{Proj}_{\mathcal{H}_n}(F)$  designates the projection of  $F$  on the  $n$ -th chaos. If additionally  $F \in \mathbf{D}^{p,2}$  for some  $p \in \mathbf{N}^*$ , this can be equivalently written as (via Stroock's formula)

$$F = \mathbf{E}[F] + \sum_{n \geq 1} \frac{1}{n!} \delta^n \left( \mathbf{E}[\nabla^{(n)} F] \right).$$

### 1.3.2 Malliavin calculus and Dirichlet structures: an intrinsic connection

Due to the existence of an underlying Markovian frame, on which they both rest, Malliavin calculus and Dirichlet structures appear to be the two sides of the same coin. By construction, a Dirichlet structure is naturally induced when we are provided by a Malliavin frame. It becomes possible to express the integration by parts formula by means of the carré du champ operator associated in the coupled Dirichlet structure.

Let a probability space  $(\Omega, \mathcal{A}, \mathbf{P})$  supposed to be endowed by a Malliavin structure on a Hilbert space  $(\mathfrak{H}, \langle \cdot, \cdot \rangle)$  that is, a gradient operator  $\mathbf{D}$ , a divergence operator  $\delta$ , an Ornstein-Uhlenbeck operator satisfying  $\mathbf{L} = -\delta\mathbf{D}$  and a integration by parts formula including  $\delta$  (1.3.1). Let

$$\mathbf{D}^{1,2} = \{F \in L^2(\Omega) : \|F\|_{1,2} = \|F\|_{L^2(\Omega)} + \|\mathbf{D}F\|_{L^2(\Omega;\mathfrak{H})}\}$$

be the closure of cylindrical functions for the norm  $\|\cdot\|_{1,2}$ . Then the quintuplet  $(\Omega, \mathcal{A}, \mathbf{P}, \mathbf{D}_{\mathcal{E}}, \mathcal{E})$  where

$$\mathbf{D}_{\mathcal{E}} = \mathbf{D}^{1,2} \quad \text{and} \quad \mathcal{E}(F, G) = \mathbf{E}[\langle \mathbf{D}F, \mathbf{D}G \rangle_{\mathfrak{H}}] ; F, G \in \mathbf{D}_{\mathcal{E}} \quad (1.3.9)$$

is a Dirichlet structure. The associated carré du champ operator is defined on  $\mathbf{D}_{\mathcal{E}}$  by

$$\Gamma(F, G) = \langle \mathbf{D}F, \mathbf{D}G \rangle_{\mathfrak{H}}$$

The integration by parts formula is retrieved through

$$\mathbf{E}[\Gamma(F, G)] = \mathbf{E}[\langle \mathbf{D}F, \mathbf{D}G \rangle_{\mathfrak{H}}] = \mathbf{E}[F \delta(\mathbf{D}G)] = -\mathbf{E}[F \mathbf{L}G] ; \forall F, G \in \mathbf{D}^{1,2}.$$

The "reverse" construction as to say the elaboration of a Malliavin calculus from a given Dirichlet structure  $(\Omega, \mathcal{A}, \mathbf{P}, \mathbf{D}_{\mathcal{E}}, \mathcal{E})$  is due to Gabriel Mokobodzki (see [90]).

**Example 1.3.6.** (Dirichlet structure on the Wiener space) We now consider  $\mathbf{P}$  as the Wiener measure on  $\mathscr{W}_{\mathbf{B}} = \mathcal{C}_0([0, 1]; \mathbf{R})$ . The aim is to construct a Dirichlet structure of the form  $(\mathscr{W}_{\mathbf{B}}, \mathcal{A}_{\mathbf{B}}, \mathbf{P}, \mathbf{D}_{\mathcal{E}}, \mathcal{E})$ . As a remainder (see Subsection 1.3.1), for any  $F \in \mathcal{S}$  of the form

$$F = f(W(h_1), \dots, W(h_n)), \quad (1.3.10)$$

where  $f$  belongs to the Schwartz space  $\mathcal{S}(\mathbf{R}^n)$  and  $h_1, \dots, h_n$  belong to  $\mathfrak{H}^1$ ,

$$\nabla F = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(W(h_1), \dots, W(h_n)) h_k,$$

and  $\mathbf{D}_{\mathbf{B}}^{1,2}$  as the closure of cylindrical functions for the norm

$$\|F\|_{1,2} = \|F\|_{L^2(\mathscr{W}_{\mathbf{B}})} + \|\nabla F\|_{L^2(\mathscr{W}_{\mathbf{B}};\mathfrak{H}_1)}.$$

The Wiener space is endowed with the Dirichlet structure

$$S = (\mathscr{W}_{\mathbf{B}}, \mathcal{A}_{\mathbf{B}}, \mathbf{P}, \mathbf{D}_{\mathcal{E}}, \mathcal{E}).$$

It is constructed as the infinite product  $\prod_{n=1}^{\infty} S^n$ ; for any  $n \in \mathbf{N}^*$  where  $S^n$  is the product structure  $\prod_{i=1}^n S_i = (\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n), \otimes_{i=1}^n \mathbf{P}_i, \mathbf{D}_{\mathcal{E}}^n, \mathcal{E}^n)$  and  $S_i$  is the Dirichlet form  $(\mathbf{R}, \mathcal{B}(\mathbf{R}), \mathbf{P}_i, \mathbf{D}_{\mathcal{E}_i}, \mathcal{E}_i)$  with  $\mathbf{P}_i$  the standard Gaussian measure on  $\mathbf{R}$ ,

$$\mathbf{D}_{\mathcal{E}_i} = \left\{ F \in \mathcal{S}, f \in \mathcal{S}(\mathbf{R}^n) : \frac{\partial f}{\partial x_i} \in L^2(\mathbf{R}) \right\} \quad \text{and} \quad \mathcal{E}_i(F) = \mathbf{E} \left[ \left| \frac{\partial f}{\partial x_i} \right|^2 \right].$$

Then we can define for  $F$  of the form (1.3.10)

$$\Gamma^n(F, F) = \sum_{i=1}^n \Gamma_k(F, F) = \sum_{j=1}^n \sum_{k=1}^n \frac{\partial f}{\partial x_j} \frac{\partial f}{\partial x_k} \langle v_j, v_k \rangle_{L^2([0,1]; \mathbf{R})}$$

and for any  $(F, G) \in \mathbf{D}_B^{1,2}$  by a limit procedure,

$$\Gamma(F, G) = \nabla F \cdot \nabla G \quad \text{and} \quad \mathcal{E}(F, G) = \mathbf{E} [\nabla F \cdot \nabla G]$$

In order to manifest the intimate link between Malliavin calculus and Dirichlet forms, we introduce the definition of *Dirichlet-Malliavin structure*.

**Definition 1.3.7** ( $\star$  Dirichlet-Malliavin structure). Consider a probability space  $(\Omega, \mathcal{A}, \mathbf{P})$  endowed by a Malliavin framework, i.e. the modicum of operators derivative  $D$ , divergence  $\delta$ , Laplacian  $L$ , Markov semi-group  $(P_t)_{t \in \mathbf{R}_+}$  defined and connected to each other by (1.3.1), (1.3.2). Define a Dirichlet form satisfying on  $(\Omega, \mathcal{A}, \mathbf{P})$  (1.3.9). The quintuplet  $(D, \delta, L, (P_t)_{t \in \mathbf{R}_+}, \mathcal{E})$  thus defined is a *Dirichlet-Malliavin structure*.

**Example 1.3.8** (Dirichlet-Malliavin structure on the Wiener space). On the one hand, consider the Malliavin frame on the Wiener space described by the operators Malliavin derivative  $\nabla$  (2.3.1), divergence  $\delta$  (see Definition 1.3.3), Laplacian  $L$  and the relationships that bind them to each other, as to say the integration by parts formula (1.3.6) and the identity  $L = -\delta D$ . On the other hand, consider the functional operator  $L$  as the infinitesimal generator of a Markovian semi-group  $(P_t)_{t \in \mathbf{R}_+}$  (the Ornstein-Uhlenbeck one) and define the Dirichlet form (from a potentialist point of view)  $(D_{\mathcal{E}}, \mathcal{E})$  by

$$D_{\mathcal{E}} = \mathbf{D}_B^{1,2} \quad \text{and} \quad \mathcal{E}(F, G) = \mathbf{E} [\Gamma(F, G)] = \mathbf{E} [-LF G] ; F, G \in \mathbf{D}_B^{1,2},$$

where  $\Gamma$  is the associated carré du champ operator. Thus, the association of a Malliavin apparatus and a Dirichlet structure provides what we call a *Dirichlet-Malliavin structure* on the Wiener space described by the quintuplet  $(D, \delta, L, (P_t)_{t \in \mathbf{R}_+}, \mathcal{E})$  and which the operator  $L$  and the integration by parts are the keystones via the identity

$$\mathcal{E}(F, G) = \mathbf{E} [-LF G] = \mathbf{E} [\nabla F \nabla G] ; F, G \in \mathbf{D}_B^{1,2}.$$

Last, note in this frame, that the semi-group  $(P_t)_{t \in \mathbf{R}_+}$  satisfies the usual commutation property: for any  $F \in \mathbf{D}_B^{1,2}$ ,

$$\nabla P_t F = e^{-t} P_t \nabla F. \quad (1.3.11)$$

### 1.3.3 Examples and applications of Dirichlet-Malliavin structures

This close connection of construction is found at the level of the applications which it generates. In other words, most probability theorems have their analogues in terms of Dirichlet structures. In particular we can revisit some results initially established using Malliavin calculus tools, from a Dirichlet structures point of view. Pick two examples as an illustration. The first one (example 1.3.3), is due to Paul Malliavin and remains the famous result of his seminal paper. It provides a criteria for the absolute continuity of a random vector defined on a Gaussian space. The second one illustrates the combination of Stein's method and Dirichlet-Malliavin calculus in the (one-dimensional) example of the Normal approximation by a Poisson distribution.

**Absolute continuity of a random vector**

Consider a random vector  $F = (F^1, \dots, F^N)$  measurable with respect to an underlying isonormal Gaussian process  $\{W(h), h \in \mathfrak{H}\}$ .

**Theorem 1.3.9** (Malliavin). *Assume that  $F$  belongs to  $\mathbf{D}_{\text{loc}}^{1,1}$  componentwise and satisfies the following conditions:*

1.  $F^i \in \mathbf{D}_{\text{loc}}^{2,p}$  for all  $i = 1, \dots, N$ , for some  $p > 0$ ,
2. The Malliavin covariance matrix  $\gamma_F = (\langle F^i, F^j \rangle_{\mathfrak{H}})_{1 \leq i, j \leq N}$  is invertible a.s.

Then, the law of  $F$  is absolutely continuous with respect to the Lebesgue measure.

The analogue of theorem 1.3.9 in terms of Dirichlet structures lies on the notion of *admissible elements*; by definition an element  $\xi \in \Omega$  is admissible if  $\xi \neq 0$  and, for any  $t \in \mathbf{R}$ ,  $(\tau_{t\xi})^* \mathbf{P}$ , the pullback measure of  $\mathbf{P}$  by the translation application  $\tau_\xi : \omega \mapsto \omega + \xi$  is equivalent to  $\mathbf{P}$ . The element  $\omega$  is additionally said to be strictly admissible, if  $\xi$  is admissible and, for almost every  $\omega \in \Omega$ , the function  $t \mapsto k_\xi(\omega + t\xi)$ , where  $k_\xi$  denotes a strictly positive Borel representative of the density of the measure  $(\tau_{t\xi})^* \mathbf{P}$  with respect to  $\mathbf{P}$ , is locally integrable on  $\mathbf{R}$  for the Lebesgue measure. It is thus possible to define (see Bouleau and Hirsch section 4.2 chapter II for more details) the Fréchet derivative  $\nabla_\xi F$  for certain functions  $F$  defined on  $\Omega$  and to set, given a sequence  $(\xi_n)_{n \geq 0}$  such that

$$\forall \psi \in \Omega', \quad \sum_{n \geq 0} |\psi(\xi_n)|^2 < \infty,$$

the Dirichlet form  $(\mathbf{D}_\mathcal{E}, \mathcal{E})$  and its associated carré du champ operator  $\Gamma$  by

$$\mathcal{E}(F) = \frac{1}{2} \sum_{n \geq 0} \|\nabla_{\xi_n} F\|_{L^2(\Omega)}^2 \quad \text{and} \quad \Gamma(F, G) = \sum_{n \geq 0} (\nabla_{\xi_n} F)(\nabla_{\xi_n} G); \quad \forall (F, G) \in \mathbf{D}_\mathcal{E}^2 \quad (1.3.12)$$

Within those assumptions we get an absolute continuity criterion expressible in terms of Malliavin derivative (see [23]):

**Theorem 1.3.10** (Bouleau, Hirsch). *Let  $(\mathbf{D}_\mathcal{E}, \mathcal{E})$  a Dirichlet form which associated carré du champ operator  $\Gamma$  is of the form (1.3.12). Then, for all  $F \in \mathbf{D}_\mathcal{E}^N$ , the measure*

$$F^* [(\det \Gamma(F, F^t)) \mathbf{P}]$$

*is absolutely continuous with respect to the Lebesgue measure on  $\mathbf{R}^N$ .*

**Stein's method and Malliavin integration by parts: a perfect match**

As seen in the subsequent sections, the conversion of the initial problem leads to deal with

$$\sup_{\varphi \in \mathcal{T}} |\mathbf{E}[L\varphi(Y)]| \quad \text{where } Y \sim \mathbf{Q},$$

or  $\mathbf{E}\left[\int_0^\infty L(P_t \varphi(Y)) dt\right]$  by addressing the problem in terms of Dirichlet structures. In both cases, we can write

$$\sup_{\varphi \in \mathcal{T}} |\mathbf{E}[L\varphi(Y)]| = \sup_{\varphi \in \mathcal{T}} |\mathbf{E}[L_1 \varphi(Y)] + \mathbf{E}[L_2 \varphi(Y)]|,$$

where  $L_1$  and  $L_2$  are two operators acting on  $\mathfrak{F}$  and such that  $L = L_1 + L_2$ . For instance, if the target distribution is the Gaussian law on  $\mathbf{R}$ ,

$$L_1\varphi(x) = \mathcal{L}_1\varphi(x) = x\varphi'(x) \quad \text{and} \quad L_2\varphi(x) = \mathcal{L}_2\varphi(x) = -\varphi''(x).$$

In the next step, we have to take into account how  $Y$  is defined and transform  $L_1\varphi(Y)$  such that it can be written as  $-L_2\varphi(Y) + \text{remainder}$ . This remainder is what gives the rate of convergence. To make the transformation of  $L_1\varphi(Y)$ , several approaches appeared along the years. One of the most popular approach (see for instance [14]) is to use exchangeable pairs: construct a copy  $Y'$  of  $Y$  with good properties which gives another expression of  $L_1\varphi(Y)$ , suitable to a comparison with  $L_2\varphi(Y)$ . To be more specific, for the proof of the CLT, it is necessary to create an exchangeable pair  $(S, S')$  with  $S = \sum_{i=1}^n Y_i$ . This is usually done by first, choosing uniformly an index  $I \in \{1, \dots, n\}$  and then, replacing  $Y_I$  with  $Y'$  an independent copy of  $Y_I$ , so that the couple  $(S, S' = S - Y_I + Y')$  is an exchangeable pair. This means that

$$\mathbf{E}[F(S') | I = a; Y_b, b \neq a] = \mathbf{E}[F(S) | Y_b, b \neq a]. \quad (1.3.13)$$

Actually, it is the right-hand-side of (1.3.13) which gave us some clue on how to proceed when dealing with functionals more general than the sum of random variables. An alternative to exchangeable pairs, is the size-biased [27] or zero biased [58] couplings, which again conveniently transform  $L_1\varphi(Y) = \mathcal{L}_1\varphi(Y)$ . For Gaussian approximation, it amounts to find a distribution  $Y^*$  such that

$$\mathbf{E}[\mathcal{L}_1\varphi(Y)] = \mathbf{E}[\varphi''(Y^*)].$$

Note that for  $S$  as above, one can choose  $S^* = S'$ . If the distribution of  $W^*$  is absolutely continuous with respect to that of  $\mathscr{W}$ , with Radon derivative  $\Lambda$ , we obtain

$$\mathbf{E}[\mathcal{L}_1\varphi(Y)] = \mathbf{E}[\varphi''(Y) \Lambda(Y)],$$

which means that we are reduced to estimate how far  $\Lambda$  is from the constant random variable equal to 1. This kind of identity, where the second order derivative is multiplied by a weight factor, is reminiscent to what can be obtained via integration by parts. Actually, Nourdin and Peccati (see [94]) showed that the transformation step can be advantageously made simple using integration by parts in the sense of Malliavin calculus. This works well only if there exists a Malliavin gradient on the space on which  $W$  is defined (see for instance [41]). That is to say, that up to now, this approach is restricted to functionals of Rademacher [96], Poisson [41, 111] or Gaussian random variables [100] or processes [32, 33]. Then, strangely enough, the first example of applications of the Stein's method which was the CLT, cannot be handled through this approach. On the one hand, exchangeable pairs or size-biased coupling have the main drawback to have to be adapted to each particular version of  $Y$ . On the other hand, Malliavin integration by parts are in some sense more automatic but we need to be provided with a Malliavin structure.

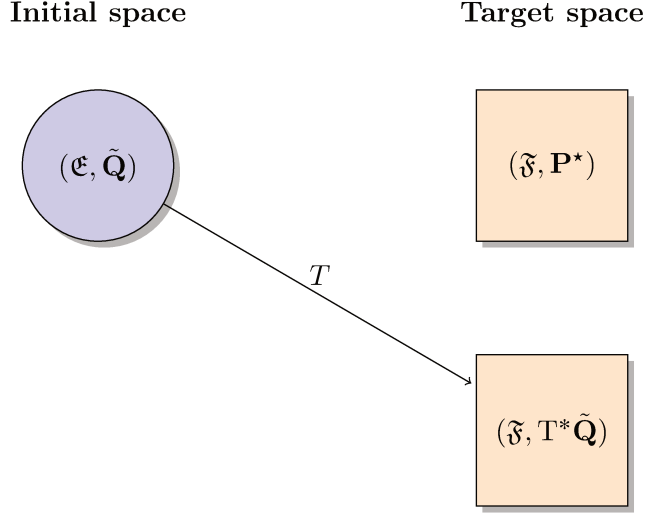
The setting in which we need to compute a KR distance is very often the situation in which we have another Polish space  $\mathfrak{E}$  with a probability measure  $\tilde{\mathbf{Q}}$  and a random variable  $T$  with value in  $\mathfrak{F}$ . The objective is then to compare some measure  $\mathbf{P}^*$  on  $\mathfrak{F}$  and  $T^*\tilde{\mathbf{Q}}$  the distribution of  $T$ , i.e. the push-forward of  $\tilde{\mathbf{Q}}$  by the application  $T$ . This means that we have to compute

$$\sup_{\varphi \in \mathcal{T}} \left| \int_{\mathfrak{E}} \varphi \circ T d\tilde{\mathbf{Q}} - \int_{\mathfrak{F}} \varphi d\mathbf{P}^* \right|. \quad (1.3.14)$$

As suggested by the figure below, the map  $T$  moves space  $\mathfrak{E}$  to  $\mathfrak{F}$  so that (1.3.14) can be rewritten in an equivalent way as

$$\sup_{\varphi \in \mathcal{T}} \left| \int_{\mathfrak{F}} \varphi dT^*\tilde{\mathbf{Q}} - \int_{\mathfrak{F}} \varphi d\mathbf{P}^* \right| = \sup_{\varphi \in \mathcal{T}} \left| \int_{\mathfrak{F}} \varphi d\mathbf{Q} - \int_{\mathfrak{F}} \varphi d\mathbf{P}^* \right|,$$

i.e. a problem of computation of distance between two measures defined on the same Polish space, prone to be solved by Stein's method.



**Example.** We choose to illustrate this with an example; that of estimation of the rate of convergence of

$$\hat{Z}_\lambda = \frac{Z_\lambda - \lambda}{\sqrt{\lambda}} \xrightarrow[\lambda \rightarrow \infty]{\text{law}} \mathcal{N}(0, 1), \tag{1.3.15}$$

where  $Z_\lambda$  is a Poisson random variable of parameter  $\lambda$ . We confront here with the situation where the target measure  $\mathbf{P}^*$  is defined on  $\mathbf{R}$  whereas the initial one  $\tilde{\mathbf{Q}}$  is integer valued. As highlighted by the figure, it is possible to get around the problem by defining a map  $T$  moving the initial space to the target one. It seems natural to set

$$\begin{aligned} T : \mathbf{N} &\longrightarrow \mathbf{R} \\ n &\longmapsto \frac{n - \lambda}{\sqrt{\lambda}} \end{aligned}$$

so that the image of  $Z_\lambda$  by the map  $T$  is  $\hat{Z}_\lambda$ , and we are at the point to evaluate the distance between the measures  $T^*\tilde{\mathbf{Q}}$  and  $\mathbf{P}^*$  both defined on  $\mathbf{R}$ . As described above, the problem is then reduced to the transformation of  $\mathbf{E} \left[ \mathcal{L}_1 \varphi(\hat{Z}_\lambda) \right]$  into  $\mathbf{E} \left[ \mathcal{L}_2 \varphi(\hat{Z}_\lambda) \right] + \text{remainder}$ . From the definition of the Stein operator associated to the standard Gaussian on  $\mathbf{R}$ , we are led to compute:

$$\mathbf{E} \left[ \hat{Z}_\lambda \varphi(\hat{Z}_\lambda) \right].$$

The trick consists in writing  $Z_\lambda = (\delta D)Z_\lambda$  in order to make use of Malliavin integration by parts formula. Remember that the gradient  $D$  (resp. its adjoint  $\delta$ ) for a Poisson measure on  $\mathbf{N}$  of parameter  $\lambda$  is given for  $F$  (resp.  $G$ ) function defined on  $\mathbf{N}$  by

$$DF(x) = F(x + 1) - F(x) \quad (\text{resp. } \delta G(x) = \frac{x}{\lambda} G(x - 1) - G(x))$$



$$\begin{aligned}
\mathbf{E} \left[ \widehat{Z}_\lambda \varphi(\widehat{Z}_\lambda) \right] &\stackrel{\text{I.P.P.}}{=} \mathbf{E} \left[ \delta(D\widehat{Z}_\lambda) \varphi(\widehat{Z}_\lambda) \right] \\
&= \mathbf{E} [DT(Z_\lambda) D(\varphi \circ T)(Z_\lambda)] \\
&= \mathbf{E} [(\varphi \circ T)(Z_\lambda + 1) - (\varphi \circ T)(Z_\lambda)] \\
&= \sqrt{\lambda} \mathbf{E} \left[ \varphi \left( \widehat{Z}_\lambda + \frac{1}{\sqrt{\lambda}} \right) - \varphi(\widehat{Z}_\lambda) \right] \\
&= \mathbf{E} \left[ \varphi'(\widehat{Z}_\lambda) + O\left(\frac{1}{\sqrt{\lambda}}\right) \right] = \mathbf{E} \left[ \mathcal{L}_2 \varphi(\widehat{Z}_\lambda) \right] + \text{remainder}.
\end{aligned}$$

We conclude that the rate of convergence in (1.3.15) is of order  $(\sqrt{\lambda})^{-1}$ .

It is possible to retrieve by using the smart-path formula (1.2.8). With the notations of Example 1.2.16, as explained in [36], we get

$$\begin{aligned}
\text{dist}_{\mathcal{T}}(\tilde{\mathbf{Q}}, \mathbf{P}^\star) &= \sup_{\varphi \in \mathcal{T}} \left| \mathbf{E} \left[ \int_0^\infty Z_\lambda (P_t \varphi)'(Z_\lambda) - (P_t \varphi)''(Z_\lambda) dt \right] \right| \\
&= \sqrt{\lambda} \sup_{\varphi \in \mathcal{T}} \left| \mathbf{E} \left[ \int_0^\infty (P_t \varphi)' \left( \widehat{Z}_\lambda + \frac{1}{\sqrt{\lambda}} \right) - (P_t \varphi)'(\widehat{Z}_\lambda) - (P_t \varphi)''(Z_\lambda) dt \right] \right| \\
&= \frac{1}{\sqrt{\lambda}} \sup_{\varphi \in \mathcal{T}} \left| \mathbf{E} \left[ \int_0^\infty \int_0^1 (1-r) (P_t \varphi)^{(3)} \left( \widehat{Z}_\lambda + \frac{1}{\sqrt{\lambda}} \right) dr dt \right] \right|,
\end{aligned}$$

where we use the integration by parts formula in the second line, and the Taylor expansion of  $P_t \varphi$  which is thrice differentiable for any  $t > 0$  by the regularizing properties of  $P_t$ . Then by choosing

$$\mathcal{T} = \{\varphi \in \mathcal{C}_b^2 : \|\varphi\|_{\mathcal{C}_b^2} \leq 1\},$$

and using the Mehler formula representation (1.2.7) of  $P_t \varphi$ , the commutation property, we get for any  $t \in \mathbf{R}_+$ ,

$$|(P_t \varphi)^{(3)}|_\infty \leq \frac{e^{-3t}}{\sqrt{1-e^{-2t}}} \int_{\mathbf{R}} |x| d\mathbf{P}^\star(x)$$

so that

$$\text{dist}_{\mathcal{T}}(\tilde{\mathbf{Q}}, \mathbf{P}^\star) \leq \frac{\sqrt{\pi}}{4\sqrt{2}\sqrt{\lambda}}.$$

Last, note that the distance can be expressed in the language of Dirichlet forms

$$\text{dist}_{\mathcal{T}}(\tilde{\mathbf{Q}}, \mathbf{P}^\star) = \int_0^\infty \mathbf{E} [LP_t \varphi] dt$$

## 1.4 As a conclusion: Stein-Malliavin criterion on a Gaussian space

As an illustration of the path taken from Stein's method to Malliavin calculus we choose to present in the last section of this chapter one significant and particularly inspiring infinite-dimensional result: a *Stein-Malliavin criterion on a Gaussian space*. It provides a "Berry-Esseen" general bound for the Wasserstein distance between the law of a functional of independent random variables (without any hypothesis on the laws of the sequence of initial variables) and the Normal distribution, and this in terms of Malliavin operators (see Nourdin and Peccati [93]). Throughout this section, we consider a Gaussian field embedded in an isonormal

process  $X = \{X(h), h \in \mathfrak{H}\}$  where  $\mathfrak{H} \subset L^2(\Omega, \mathcal{A}, \mathbf{P})$  is its associated Gaussian subspace. We assume that  $\mathcal{A} = \sigma(X)$  and denote  $\mathbf{Q}$  its distribution. The elaboration of Malliavin calculus within this framework is detailed in subsection 1.3.1 and we keep the notations introduced there.

**Theorem 1.4.1** (Nourdin, Peccati - 2009). *Let  $\mathbf{P}^*$  denote the standard Gaussian distribution on  $\mathbf{R}$ . For any  $F \in \mathbf{D}^{1,4}$  such that  $\mathbf{E}[F] = 0$ ,*

$$\mathfrak{d}_1(F^* \mathbf{Q}, \mathbf{P}^*) \leq \sqrt{\mathbf{E} \left[ |1 - \langle \nabla F, -\nabla L^{-1} F \rangle_{\mathfrak{H}}|^2 \right]}. \quad (1.4.1)$$

A look to the proof allows to extract the ingredients to gather to establish the result. As explained in the section devoted to Stein's method, the reduction of problem leads to tackle with

$$\mathbf{E}[\mathcal{L}_1 \varphi(F)] = \mathbf{E}[F \varphi(F)],$$

that we want to transform into  $-\mathbf{E}[\mathcal{L}_2 \varphi(F)] + \text{remainder}$ . Using the relations between the main Malliavin-type operators  $D$ ,  $\delta$  and  $L$ , we get:

$$\begin{aligned} \mathbf{E}[F \varphi(F)] &= \mathbf{E}[L(L^{-1} F) \varphi(F)] = \mathbf{E}[-\delta(DL^{-1} F) \varphi(F)] \\ &= \mathbf{E}[\langle D\varphi(F), -DL^{-1} F \rangle_{\mathfrak{H}}] \quad (\text{IPP formula}) \\ &= \mathbf{E}[\varphi'(F) \langle DF, -DL^{-1} F \rangle_{\mathfrak{H}}] \quad (\text{chain rule}) \\ &= \mathbf{E} \left[ \underbrace{\varphi'(F)}_{-\mathcal{L}_2 \varphi(F)} \underbrace{-\varphi'(F) (1 + \langle DF, -DL^{-1} F \rangle_{\mathfrak{H}})}_{\text{remainder}} \right] \end{aligned}$$

The last step which consists in using Cauchy-Schwarz inequality and requires that  $F \in \mathbf{D}^{1,4}$  to use the regularizing property of the semi-group  $(P_t)_{t \in \mathbf{R}_+}$ , i.e.

$$\begin{aligned} \mathbf{E}[\|DL^{-1} F\|_{\mathfrak{H}}^4] &= \mathbf{E} \left[ \left\| \int_0^\infty e^{-t} P_t DF dt \right\|_{\mathfrak{H}}^4 \right] \\ &\leq \mathbf{E}[\|DF\|_{\mathfrak{H}}^4]. \end{aligned}$$

As suggested in blue, the statement of this result requires to gather the following ingredients:

- A characterization of  $\mathbf{P}$  in terms of 1<sup>st</sup>-order differential operators.
- An underlying Dirichlet-Malliavin structure  $(D, \delta, L, (P_t)_{t \geq 0}, \mathcal{E})$ .
- An integration by parts formula including  $\delta$ .

One of the challenges of the next chapter is to develop the aforementioned tools to state an analogue of the Stein-Malliavin criterion in any countable product of probability spaces.



## Chapter 2

# Malliavin calculus and Dirichlet structures for independent random variables

In the first chapter we notably investigated the possibility to derive probabilistic approximations by regular functionals of processes for which there exists a Malliavin calculus. This means exploiting the underlying Dirichlet-Malliavin structure. The motivation of this second chapter is to shape the suitable tools to be able - in the third one - to provide an analogue of the "Stein-Malliavin criterion" for functionals of independent random variables. As suggested earlier, we need to define a modicum of Malliavin-type operators - such as derivative, divergence and Ornstein-Uhlenbeck operators - and the yielded Dirichlet structure in the general discrete context we choose to investigate: that of a family of independent, non necessarily identically distributed, random variables.

The design of our stochastic calculus of variations on any countable product of probability spaces enables to generalize to a certain extent what is known about Rademacher spaces (see Privault [117]), i.e.  $\{-1, 1\}^{\mathbf{N}}$  endowed by the probability product  $\otimes_{k \in \mathbf{N}} \mu_k$  where  $\mu_k$  is a Bernoulli distribution on  $\{-1, 1\}$ . Despite apparent similarities in their patterns (countable product of probability spaces in either case), the two approaches of constructing a Malliavin calculus nevertheless remain very different from each other. The Rademacher space, as product of two-components state spaces, shows a good range of algebraic specific properties which enable to define a normal martingale  $(Y_n)_{n \in \mathbf{N}}$  satisfying a structure equation of the form

$$Y_n^2 = 1 + \varphi_n Y_n. \quad (2.0.1)$$

where  $(\varphi_n)_{n \in \mathbf{N}}$  is  $\mathcal{F}$ -predictable process. This one satisfies a "discrete" predictable representation property such that any square integrable Rademacher functional can be expanded into chaos. The gradient operator expression comes thus from the definition of the annihilation operator as it is introduced. The gradient on the Rademacher space (see Nourdin, Peccati and Reinert [96], Privault [117]) is usually defined as

$$\begin{aligned} \widehat{D}_k F(X_1, \dots, X_n) &= \mathbf{E}[X_k F(X_1, \dots, X_n) | X_\ell, \ell \neq k] \\ &= \mathbf{P}(X_k = 1)F(X_1, \dots, +1, \dots, X_n) - \mathbf{P}(X_k = -1)F(X_1, \dots, -1, \dots, X_n), \end{aligned} \quad (2.0.2)$$

where the  $\pm 1$  are put in the  $k$ -th coordinate. It requires, for its very definition to be meaningful, either that the random variables are real valued or that they only have two possible outcomes.

Moreover, as highlighted by Privault and Schoutens there exists a discrete structure equation if and only if the law of  $X_k$  is supported by two points (see proposition 3 in Privault and Schoutens [119]). This leads, in a general context of unspecified Polish spaces, to the obvious impossibility of stating a chaotic decomposition and drives us to define an ad hoc gradient, which does not coincide with the usual one on the Rademacher space (2.0.2).

We proceed as follows. In Section 2.1, we define the gradient  $D$  and its adjoint  $\delta$ , which we call divergence as it appears as the sum of *the partial derivatives*, as in  $\mathbf{R}^n$ . It is defined thanks to what may arise as the Malliavin calculus cornerstone: the integration by parts formula. The following section is devoted to the exploitation of this new formalism to the statement of representation formulas and functional identities. We establish a Clark representation formula of square integrable random variables and an Helmholtz decomposition of vector fields. We establish a log-Sobolev inequality, strongly reminding that obtained for Poisson processes (see [140]), together with a concentration inequality. Then, we define the number operator  $L = -\delta D$ . It coincides with the generator of a Markov process whose stationary distribution is the tensor probability we started with. We show in Section 2.3 that we can retrieve the classical Dirichlet-Malliavin structures for Poisson processes and Brownian motion as limits of our structures. In that respect, our theory seems to find its place in the pre-existing Malliavin's landscape.

## 2.1 Malliavin calculus for independent random variables

Let  $A$  be an at most countable set equipped with the counting measure :

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

Let  $(E_a, a \in A)$  be a family of Polish spaces. For any  $a \in A$ , let  $\mathcal{E}_a$  and  $\mathbf{P}_a$  be respectively a  $\sigma$ -field and a probability measure defined on  $E_a$ . We consider the probability space  $E_A = \prod_{a \in A} E_a$  equipped with the product  $\sigma$ -field  $\mathcal{E}_A = \bigvee_{a \in A} \mathcal{E}_a$  and the tensor product measure  $\mathbf{P} = \bigotimes_{a \in A} \mathbf{P}_a$ .

The coordinate random variables are denoted by  $(X_a, a \in A)$ . For any  $B \subset A$ ,  $X_B$  denotes the random vector  $(X_a, a \in B)$ , defined on  $E_B = \prod_{a \in B} E_a$  equipped with the probability  $\mathbf{P}_B = \bigotimes_{a \in B} \mathbf{P}_a$ .

For any  $a \in A$ , let  $\mathcal{G}_a$  and  $\mathcal{F}_a$  be the  $\sigma$ -files defined by

$$\mathcal{G}_a = \sigma(X_b, b \neq a), \quad \mathcal{F}_a = \sigma(X_b, b \leq a),$$

and assume that  $\mathcal{E}_A = (\mathcal{F}_a)_{a \in A}$ .

A process  $U$  is a measurable random variable defined on  $(E_A \times A, \mathcal{P}(A) \otimes \mathcal{E}_A)$ .

We denote by  $L^2(E_A \times A)$  the Hilbert space of processes which are square integrable with respect to the measure  $\sum_{a \in A} \varepsilon_a \otimes \mathbf{P}_A$  (where  $\varepsilon_a$  is the Dirac measure at point  $a$ ) :

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

Our presentation follows closely the usual construction of Malliavin calculus from the class of *cylindrical* functionals, denoted in the following by  $\mathcal{S}$ .

**Definition 2.1.1.** A random variable  $F$  is said to be cylindrical if there exist a finite subset  $B \subset A$  and a function  $F_B : E_B \rightarrow L^2(A)$  such that  $F = F_B \circ r_B$ , where  $r_B$  is the restriction operator :

$$\begin{aligned} r_B : E_A &\longrightarrow E_B \\ (x_a, a \in A) &\longmapsto (x_a, a \in B). \end{aligned}$$

This means that  $F$  only depends on the finite set of random variables  $(X_a, a \in B)$ . It is clear that  $\mathcal{S}$  is dense in  $L^2(E_A)$ .

The very first tool to be considered is the discrete gradient, whose form is motivated by what follows ; throughout this work, it must be made clear that all the random variables may leave on different spaces, which are only supposed to be Polish spaces. That means that in the definition of the gradient, we cannot use any algebraic property of the underlying spaces. Though some of our applications does concern random variables with finite number of outcomes, it does not seem obvious to exhibit what should be the weights, replacing  $\mathbf{P}(X_k = 1)$  and  $-\mathbf{P}(X_k = -1)$  appearing in (2.0.2). We offer an alternative definition.

We first define the gradient of cylindrical functionals, for there is no question of integrability and then extend the domain of the gradient to a larger set of functionals by a limit procedure. In functional analysis terminology, we need to verify the closability of the gradient : If a sequence of functionals converges to 0 and the sequence of their gradients converges, then it should also converges to 0. This is the only way to guarantee in the limit procedure that the limit does not depend on the chosen sequence.

**Definition 2.1.2** ( $\star$  Discrete gradient). For  $F \in \mathcal{S}$ ,  $DF$  is the process of  $L^2(E_A \times A)$  defined by one of the following equivalent formulations : for all  $a \in A$ ,

$$\begin{aligned} D_a F(X_A) &= F(X_A) - \mathbf{E}[F(X_A) | \mathcal{G}_a] \\ &= F(X_A) - \int_{E_a} F(X_{A \setminus a}, X_a) d\mathbf{P}_a(X_a) \\ &= F(X_A) - \mathbf{E}'[F(X_{A \setminus a}, X'_a)], \end{aligned}$$

where  $X'_a$  is an independent copy of  $X_a$ .

**Remark 2.1.3.** During the preparation of this work, we found strong reminiscences of our gradient with the map  $\Delta$ , introduced by Stéphane Boucheron, Gábor Lugosi and Pascal Massart in [21], or WanSoo T. Rhee and Michel Talagrand in [124] for the proof of the Efron-Stein inequality, defined by

$$\Delta_k F(X_1, \dots, X_n) = \mathbf{E}[F | X_1, \dots, X_k] - \mathbf{E}[F | X_1, \dots, X_{k-1}].$$

Actually, our point of view diverges from that of these works as we do not focus on a particular inequality but rather on the intrinsic properties of our newly defined gradient. Its expression can be interpreted as the measure of the "influence" of the  $a$ -th component of the process  $X$  on  $F$ .

**Remark 2.1.4.** A straightforward calculation shows that for any  $F, G \in \mathcal{S}$ , any  $a \in A$ , we have

$$D_a(FG) = F D_a G + G D_a F - D_a F D_a G - \mathbf{E}[FG | \mathcal{G}_a] + \mathbf{E}[F | \mathcal{G}_a] \mathbf{E}[G | \mathcal{G}_a].$$

This formula has to be compared with the formula  $D(FG) = F(DG) + G(DF)$  for the Gaussian Malliavin gradient (see (2.3.1) below) and  $D(FG) = F(DG) + G(DF) + (DF)(DG)$  for the Poisson gradient (see (2.3.3) below).

For  $F \in \mathcal{S}$ , there exists a finite subset  $B \subset A$  such that  $F = F_B \circ r_B$ . Thus, for every  $a \notin B$ ,  $F$  is  $\mathcal{G}_a$ -measurable and then  $D_a F = 0$ . This implies that

$$\|DF\|_{L^2(A \times E_A)}^2 = \mathbf{E} \left[ \sum_{a \in A} |D_a F|^2 \right] = \mathbf{E} \left[ \sum_{a \in B} |D_a F|^2 \right] < \infty,$$

hence  $(D_a F, a \in A)$  defines an element of  $L^2(E_A \times A)$ .

**Definition 2.1.5.** The set of simple processes, denoted by  $\mathcal{S}_0(\ell^2(A))$  is the set of random variables defined on  $E_A \times A$  of the form

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

for  $B$  a finite subset of  $A$  and such that  $U_a$  belongs to  $\mathcal{S}$  for any  $a \in B$ .

The key formula for the sequel is the so-called integration by parts. It amounts to compute the adjoint of  $D$  in  $L^2(E_A \times A)$ .

**Theorem 2.1.6** (★ Integration by parts). *Let  $F \in \mathcal{S}$ . For every simple process  $U$ ,*

$$\langle DF, U \rangle_{L^2(E_A \times A)} = \mathbf{E} \left[ F \sum_{a \in A} D_a U_a \right]. \quad (2.1.1)$$

Thanks to the latter formula, we are now in position to prove the closability of  $D$ : for  $(F_n)_{n \in \mathbf{N}^*}$  a sequence of cylindrical functionals,

$$\left( F_n \xrightarrow[L^2(A)]{n \rightarrow \infty} 0 \text{ and } DF_n \xrightarrow[L^2(E_A \times A)]{n \rightarrow \infty} \eta \right) \implies \eta = 0.$$

**Corollary 2.1.7** (★). *The operator  $D$  is closable from  $L^2(A)$  into  $L^2(E_A \times A)$ .*

We denote the domain of  $D$  in  $L^2(A)$  by  $\mathbf{D}$ , the closure of the class of cylindrical functions with respect to the norm

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

We could as well define  $p$ -norms corresponding to  $L^p$  integrability. However, for the current applications, the case  $p = 2$  is sufficient and the apparent lack of hypercontractivity of the Ornstein-Uhlenbeck semi-group (see below Section 2.1.2) lessens the probable usage of other integrability order.

Since  $\mathbf{D}$  is defined as a closure, it is often useful to have a general criterion to ensure that a functional  $F$ , which is not cylindrical, belongs to  $\mathbf{D}$ . The following criterion exists as is in the settings of Wiener and Poisson spaces.

**Lemma 2.1.8** (★). *If there exists a sequence  $(F_n)_{n \in \mathbf{N}^*}$  of elements of  $\mathbf{D}$  such that*

1.  $F_n$  converges to  $F$  in  $L^2(E_A)$ ,
2.  $\sup_n \|DF_n\|_{\mathbf{D}}$  is finite,

*then  $F$  belongs to  $\mathbf{D}$  and  $DF = \lim_{n \rightarrow \infty} DF_n$  in  $\mathbf{D}$ .*

### 2.1.1 Divergence

We can now introduce the adjoint of  $\mathbf{D}$ , often called the divergence as for the Lebesgue measure on  $\mathbf{R}^n$ , the usual divergence is the adjoint of the usual gradient.

**Definition 2.1.9** (★ Divergence). Let

$$\text{Dom } \delta = \left\{ U \in L^2(\mathbf{E}_A \times A) : \exists c > 0, \forall F \in \mathbf{D}, |\langle \mathbf{D}F, U \rangle_{L^2(\mathbf{E}_A \times A)}| \leq c \|F\|_{L^2(A)} \right\}.$$

For any  $U$  belonging to  $\text{Dom } \delta$ ,  $\delta U$  is the element of  $L^2(A)$  characterized by the following identity

$$\langle \mathbf{D}F, U \rangle_{L^2(\mathbf{E}_A \times A)} = \mathbf{E}[F \delta U], \text{ for all } F \in \mathbf{D}.$$

The integration by parts formula (2.1.1) entails that for every  $U \in \text{Dom } \delta$ ,

$$\delta U = \sum_{a \in A} D_a U_a.$$

The expression of  $\delta$ , as the sum of *the partial derivatives*, as in  $\mathbf{R}^n$ , allows to legitimize the name of *divergence* we give it.

In the setting of Malliavin calculus for Brownian motion, the divergence of adapted processes coincides with the Itô integral and the square moment of  $\delta U$  is then given by the Itô isometry formula. We now see how this extends to our situation.

**Definition 2.1.10** (★). The Hilbert space  $\mathbf{D}(\ell^2(A))$  is the closure of  $\mathcal{S}_0(\ell^2(A))$  with respect to the norm

$$\|U\|_{\mathbf{D}(\ell^2(A))}^2 = \mathbf{E} \left[ \sum_{a \in A} |U_a|^2 \right] + \mathbf{E} \left[ \sum_{a \in A} \sum_{b \in A} |D_a U_b|^2 \right].$$

In particular, this means that the map  $\mathbf{D}U = (D_a U_b, a, b \in A)$  is Hilbert-Schmidt as a map from  $L^2(\mathbf{E}_A \times A)$  into itself. As a consequence, for two such maps  $\mathbf{D}U$  and  $\mathbf{D}V$ , the map  $\mathbf{D}U \circ \mathbf{D}V$  is trace-class (see Yosida [142]) with

$$\text{trace}(\mathbf{D}U \circ \mathbf{D}V) = \sum_{a, b \in A} (D_a U_b) (D_b V_a).$$

The next formula is the counterpart of the Itô isometry formula for the Brownian motion, sometimes called the Weitzenböck formula (see Privault [117, Eqn. (4.3.3)]) in the Poisson settings.

**Theorem 2.1.11** (★). *The space  $\mathbf{D}(\ell^2(A))$  is included in  $\text{Dom } \delta$ . For any  $U, V$  belonging to  $\mathbf{D}(\ell^2(A))$ ,*

$$\mathbf{E}[\delta U \delta V] = \mathbf{E}[\text{trace}(\mathbf{D}U \circ \mathbf{D}V)]. \quad (2.1.2)$$

**Remark 2.1.12.** It must be noted that compared to the analogue identity for the Brownian and the Poisson settings, the present formula is slightly different. For both processes, with corresponding notations, we have

$$\|\delta U\|_{L^2(A)}^2 = \|U\|_{L^2(\mathbf{E}_A \times A)}^2 + \text{trace}(\mathbf{D}U \circ \mathbf{D}U).$$

The absence of the term  $\|U\|_{L^2(A)}$  gives to our formula a much stronger resemblance to the analogue equation for the Lebesgue measure. As in this latter case, we do have here  $\delta \mathbf{1} = 0$



whereas for the Brownian motion, it yields the Itô integral of the constant function equal to one. If  $A = \mathbf{N}$ , let  $\mathcal{F}_n = \sigma\{X_k, k \leq n\}$  and assume that  $U$  is adapted, i.e. for all  $n \geq 1$ ,  $U_n \in \mathcal{F}_n$ . Then,  $D_n U_k = 0$  as soon as  $n > k$ , hence

$$\mathbf{E}[(\delta U)^2] = \mathbf{E} \left[ \sum_{n=1}^{\infty} \left( U_n - \mathbf{E}[U_n | \mathcal{F}_{n-1}] \right)^2 \right],$$

i.e.  $\mathbf{E}[(\delta U)^2]$  is the  $L^2(\mathbf{N} \times E_{\mathbf{N}})$ -norm of the innovation process associated to  $U$ , which appears in filtering theory.

### 2.1.2 Ornstein-Uhlenbeck semi-group and generator

Having defined a gradient and a divergence, one may consider the Laplacian-like operator defined by  $L = -\delta D$ , which is also called the number operator in the settings of Gaussian Malliavin calculus.

**Definition 2.1.13** (★). The number operator, denoted by  $L$ , is defined on its domain

$$\text{Dom } L = \left\{ F \in L^2(A) : \mathbf{E} \left[ \sum_{a \in A} |D_a F|^2 \right] < \infty \right\}$$

by

$$LF = -\delta DF = - \sum_{a \in A} D_a F. \quad (2.1.3)$$

The map  $L$  can be viewed as the generator of a symmetric Markov process  $X$ , which is ergodic, whose stationary probability is  $\mathbf{P}_A$ . Assume first that  $A$  is finite. Consider  $(Z_t)_{t \in \mathbf{R}_+}$  a Poisson process on the half-line of rate  $|A|$ , and the process  $X_t = (X_{1,t}, \dots, X_{N,t}, t \geq 0)$  which evolves according to the following rule: At a jump time of  $Z$ ,

- Choose randomly (with equiprobability) an index  $a \in A$ ,
- Replace  $X_a$  by an independent random variable  $X'_a$  distributed according to  $\mathbf{P}_a$ .

For every  $x \in E_A$ ,  $a \in A$ , set  $x^{-a} = (x_1, \dots, x_{a-1}, x_{a+1}, \dots, x_{|A|})$ . The generator of the Markov process  $X$  is clearly given by

$$|A| \sum_{a \in A} \frac{1}{|A|} \int_{E_a} \left( F(x^{-a}, x'_a) - F(x) \right) d\mathbf{P}_a(x'_a) = - \sum_{a \in A} D_a F(x). \quad (2.1.4)$$

The factor  $|A|$  is due to the intensity of the Poisson process  $Z$  which jumps at rate  $|A|$ , the factor  $|A|^{-1}$  is due to the uniform random choice of an index  $a \in A$ . Thus, for a finite set  $A$ ,  $L$  coincides with the generator of  $X$ .

**Remark 2.1.14.** This result is in fact transposable in terms of a regular Markov process of parameters  $(\nu, Q)$  where  $\nu$  is a probability measure on a state space  $E$  and  $Q$  a transition matrix. The Markov process is constructed in a trajectorial way (see for instance Decreusefond and Moyal [39]). Let  $E = E_A$ ,  $X_0$  be a random variable with law  $\nu$  and define the matrix  $Q = \left( q(x, (x^{-a}, z)), (x, (x^{-a}, z)) \in E_A \times E_A^{-a} \times E_a \right)$  as follows :

$$\begin{cases} q(x, x) &= |A| \\ q(x, (x^{-a}, z)) &= \frac{d\mathbf{P}_a(z)}{|A|}, \quad z \in E_a \end{cases}$$

that is well defined since  $\sum_{a \in A} \sum_{z \in E_a} q(x, (x^{-a}, z)) = 1$ .

As a reminder, the generator of a regular Markov process  $X$  of parameters  $(\nu, Q)$  is given for  $f \in L^\infty(E_A)$  by

$$A_Q F(x, y) = q(x, x) \sum_{y \in E_A^{\neq x}} (F(y) - F(x)) q(x, y) = q(x, x) \sum_{a \in A} \sum_{z \in E_a^{\neq x_a}} (F(x^{-a}, z) - F(x)) q(x, (x^{-a}, z))$$

so that we retrieve (2.1.4).

Denote by  $P = (P_t)_{t \in \mathbf{R}_+}$  the semi-group of  $X$ ; for any  $x \in E_A$ , for any bounded  $f : E_A \rightarrow \mathbf{R}$ ,

$$P_t F(x) = \mathbf{E} [F(X_t) | X_0 = x].$$

Then,  $(P_t, t \geq 0)$  is a strong Feller semi-group on  $L^\infty(E_A)$ . This result still holds when  $E_A$  is countable.

**Theorem 2.1.15** (★). *For any countable set  $A$ ,  $L$  defined as in (2.1.3) generates a strong Feller continuous semi-group  $(P_t)_{t \geq 0}$  on  $L^\infty(E_A)$ .*

*As a consequence, there exists a Markov process  $X$  whose generator is  $L$  as defined in (2.1.3). It admits as a core (a dense subset of its domain) the set of cylindrical functions.*

From the sample-path construction of  $X$ , the next result is straightforward for  $A$  finite and can be obtained by a limit procedure for  $A$  countable.

**Theorem 2.1.16** (★ Mehler formula). *For  $a \in A$ ,  $x_a \in E_a$  and  $t > 0$ , let  $X_{a,t}^{x_a}$  the random variable defined by*

$$X_{a,t}^{x_a} = \begin{cases} x_a & \text{with probability } e^{-t}, \\ X'_a & \text{with probability } 1 - e^{-t}, \end{cases}$$

*where  $X'_a$  is a  $\mathbf{P}_a$ -distributed random variable independent from everything else. In other words, if  $\mathbf{P}_{a,t}^{x_a}$  denotes the distribution of  $X_{a,t}^{x_a}$ ,  $\mathbf{P}_{a,t}^{x_a}$  is a convex combination of  $\delta_{x_a}$  and  $\mathbf{P}_a$  :*

$$\mathbf{P}_{a,t}^{x_a} = e^{-t} \delta_{x_a} + (1 - e^{-t}) \mathbf{P}_a.$$

*For any  $x \in E_A$ , any  $t > 0$ ,*

$$P_t F(x) = \int_{E_A} F(y) \otimes_{a \in A} d\mathbf{P}_{a,t}^{x_a}(y_a).$$

*It follows easily that  $(P_t)_{t \geq 0}$  is ergodic and stationary :*

$$\lim_{t \rightarrow \infty} P_t F(x) = \int_{E_A} F d\mathbf{P} \text{ and } X_0 \stackrel{\text{law}}{=} \mathbf{P} \implies X_t \stackrel{\text{law}}{=} \mathbf{P}.$$

We then retrieve the classical formula (in the sense that it holds as is for Brownian motion and Poisson process) of commutation between  $D$  and the Ornstein-Uhlenbeck semi-group.

**Theorem 2.1.17** (★). *Let  $F \in L^2(E_A)$ . For every  $a \in A$ ,  $x \in E_A$ ,*

$$D_a P_t F(x) = P_t D_a F(x). \quad (2.1.5)$$

## 2.2 Functional identities

This section is devoted to several functional identities which constitute the crux of the matter if we want to do some computations with our new tools. These can be associated to two areas of interest, which are also linked in our context: representation theorems and concentration inequalities.

The first part of this section is devoted to the statement of a representation identity, the analogue to the Clark-Ocone formula and its corollaries (for original functional identities see Bakry Gentil and Ledoux[12], Ledoux [82], Talagrand [134]): Poincaré's inequality and a covariance identity. The Clark formula is also the initial point to lead to a log-Sobolev inequality, strongly reminding that obtained for Poisson processes (see Wu [140]), together with a concentration inequality.

It is classical that the notion of adaptability is linked to the support of the gradient.

**Lemma 2.2.1 (★).** *Assume that  $A = \mathbf{N}$  and let  $\mathcal{F}_n = \sigma\{X_k, k \leq n\}$ . For any  $F \in \mathbf{D}$ ,  $F$  is  $\mathcal{F}_k$ -measurable if and only if  $D_n F = 0$  for any  $n > k$ . As a consequence,  $DF = 0$  if and only if  $F = \mathbf{E}[F]$ .*

It is also well known that, in the Brownian and Poisson settings,  $D$  and conditional expectation commute.

**Lemma 2.2.2 (★).** *For any  $F \in \mathbf{D}$ , for any  $k \geq 1$ , we have*

$$D_k \mathbf{E}[F | \mathcal{F}_k] = \mathbf{E}[D_k F | \mathcal{F}_k]. \quad (2.2.1)$$

The Brownian martingale representation theorem says that a martingale adapted to the filtration of a Brownian motion is in fact a stochastic integral. The Clark formula gives the expression of the integrand of this stochastic integral in terms of the Malliavin gradient of the terminal value of the martingale. We here have the analogue formula.

**Theorem 2.2.3 (★ Clark formula).** *For  $A = \mathbf{N}$  and  $F \in \mathbf{D}$ ,*

$$F = \mathbf{E}[F] + \sum_{k=1}^{\infty} D_k \mathbf{E}[F | \mathcal{F}_k].$$

Since Malliavin calculus is agnostic to any time reference, we do not even assume that we have an order on the product space. It is not a major feature since a countable set  $A$  is by definition one-to-one with the set of natural integers and thus inherits of at least one order structure. However, this added degree of freedom appears to be useful (see the Clark decomposition of the number of fixed points of a random permutations in Section 3.1) and bears strong resemblance with the different filtrations which can be put on an abstract Wiener space, via the notion of resolution of the identity (see Üstünel and Zakai[136]).

**Corollary 2.2.4 (★).** *Within the assumptions of theorem 2.2.3 ; if  $A$  is finite and if there is no privileged order on  $A$ , we can write*

$$F = \mathbf{E}[F] + \sum_{B \subset A} \binom{|A|}{|B|}^{-1} \frac{1}{|B|} \sum_{b \in B} D_b \mathbf{E}[F | X_B].$$

**Remark 2.2.5.** The chaos decomposition is usually deduced from Clark formula by iteration. If we apply Clark formula to  $\mathbf{E}[F | \mathcal{F}_k]$ , we get

$$D_k \mathbf{E}[F | \mathcal{F}_k] = \sum_{j=1}^{\infty} D_k D_j \mathbf{E}[F | \mathcal{F}_{j \wedge k}] = \cdots = D_k \mathbf{E}[F | \mathcal{F}_k],$$

and we get back to our starting point, since  $j > k$  implies  $D_j \mathbf{E}[F | \mathcal{F}_k] = 0$  in view of Lemma 2.2.1. Furthermore, the same holds when  $k > j$  since it is easily seen that  $D_j D_k = D_k D_j$ . For  $j = k$ , simply remark that  $D_k D_k = D_k$ . Hence, it seems that we cannot go further this way to find a potential chaos decomposition.

As mentioned in the Introduction, it may be useful to reverse the time arrow. Choose an order on  $\mathbf{A}$  so that  $\mathbf{A}$  can be seen as  $\mathbf{N}$ . Then, let

$$\check{\mathcal{F}}_n = \sigma\{X_k, k > n\}.$$

and for any  $n \in \{0, \dots, N-1\}$ ,

$$\check{\mathcal{F}}_n^N = \check{\mathcal{F}}_n \cap \mathcal{F}_N \quad \text{and} \quad \check{\mathcal{F}}_k^N = \mathcal{F}_0 = \{\emptyset, E_A\} \text{ for } k \geq N.$$

Note that  $\check{\mathcal{F}}_0^N = \mathcal{F}_N$  and as in Lemma 2.2.1,  $F$  is  $\check{\mathcal{F}}_k$ -measurable if and only if  $D_n F = 0$  for any  $n \leq k$ .

**Theorem 2.2.6 (★).** *For every  $F$  in  $\mathbf{D}$ ,*

$$F = \mathbf{E}[F] + \sum_{k=1}^{\infty} D_k \mathbf{E}[F | \check{\mathcal{F}}_{k-1}].$$

In the present context, the next result is a Poincaré type inequality as it gives a bound for the variance of  $F$  in terms of the *oscillations* of  $F$ . In other context, it turns out to be called the Efron-Stein inequality (see Boucheron, Lugosi and Massart [21]). It can be noted that both the statement and the proof are similar in the Brownian and Poisson settings.

**Corollary 2.2.7 (★ Poincaré or Efron-Stein inequality).** *For any  $F \in \mathbf{D}$ ,*

$$\text{Var}(F) \leq \|DF\|_{L^2(E_A \times A)}^2.$$

Another corollary of the Clark formula is the following covariance identity.

**Theorem 2.2.8 (★ Covariance identity).** *For any  $F, G \in \mathbf{D}$ ,*

$$\text{cov}(F, G) = \mathbf{E} \left[ \sum_{k \in \mathbf{A}} D_k \mathbf{E}[F | \mathcal{F}_k] D_k G \right]. \quad (2.2.2)$$

As for the other versions of the Malliavin calculus (Brownian, Poisson and Rademacher), from (2.1.5), can be deduced another covariance identity.

**Theorem 2.2.9 (★).** *For any  $F, G \in \mathbf{D}$ ,*

$$\text{cov}(F, G) = \mathbf{E} \left[ \sum_{k \in \mathbf{A}} D_k F \int_0^{\infty} e^{-t} P_t \mathbf{E}[D_k G | \mathcal{F}_k] dt \right]. \quad (2.2.3)$$

Then, using the so-called Herbst principle, we can derive a concentration inequality, in its large deviation formulation, which, as usual, requires an  $L^\infty$  bound on the derivative of the functional to be valid.

**Theorem 2.2.10** (★ Concentration inequality). *Let  $F$  for which there exists an order on  $A$  with*

$$M = \sup_{X \in E_A} \sum_{k=1}^{\infty} |D_k F(X)| \mathbf{E} [|D_k F(X)| | \mathcal{F}_k] < \infty.$$

*Then, for any  $x \geq 0$ , we have*

$$\mathbf{P}(F - \mathbf{E}[F] \geq x) \leq \exp\left(-\frac{x^2}{2M}\right).$$

In the Gaussian case, the concentration inequality is deduced from the logarithmic Sobolev inequality. This does not seem to be feasible in the present context because  $D$  is not a derivation, i.e. does not satisfy  $D(FG) = F DG + G DF$ . There is so some interest to study an inequality such as

$$\text{Ent}_\mu[F] \leq C \mathbf{E}_\mu \left[ \frac{|DF|^2}{F} \right], \quad (2.2.4)$$

that can be stated if, for instance,  $\mu$  is the Bernoulli measure on  $\{0, 1\}$  or the Poisson measure on  $\mathbf{N}$  and  $D$  the associated usual discrete gradient in each case. We still have an LSI identity comparable to (2.2.4). For the proof of it, we follow closely the proofs of Nicolas Privault [116] and Liming Wu [140]. They are based on two ingredients: the Itô formula and the martingale representation theorem. We get an ersatz of the former but the latter seems inaccessible as we do not impose the random variables to live in the same probability space and to be real valued. Should it be the case, to the best of our knowledge, the martingale representation formula is known only for the Rademacher space (see Williams [139, Section 15.1]), which is exactly the framework of Nicolas Privault [116]. This lack of a predictable representation explains the conditioning in the denominator of (2.2.5).

**Theorem 2.2.11** (★ Logarithmic Sobolev inequality). *Let a positive random variable  $G \in L \log L(E_A)$ . Then,*

$$\mathbf{E}[G \log G] - \mathbf{E}[G] \log \mathbf{E}[G] \leq \sum_{k \in A} \mathbf{E} \left[ \frac{|D_k G|^2}{\mathbf{E}[G | \mathcal{G}_k]} \right]. \quad (2.2.5)$$

In the usual vector calculus on  $\mathbf{R}^3$ , the Helmholtz decomposition stands that a sufficiently smooth vector field can be resolved in the sum of a curl-free vector field and a divergence-free vector field. We have here the exact counterpart with our definition of gradient.

**Theorem 2.2.12** (★ Helmholtz decomposition). *Let  $U \in \mathbf{D}(\ell^2(A))$ . There exists a unique couple  $(\varphi, V)$  where  $\varphi \in L^2(E_A)$  and  $V \in L^2(E_A \times A)$  such that  $\mathbf{E}[\varphi] = 0$ ,  $\delta V = 0$  and*

$$U_a = D_a \varphi + V_a,$$

*for any  $a \in A$ .*

We choose to highlight the similarities and differences of the classical functional inequalities obtained in the Gaussian and Poisson settings and their analogues in our discrete framework through the following summary table. The identities relating to "Gaussian space", held on any Gaussian fields (we keep the notations used on subsection 1.3.1) except from the Clark formula stated here with respect to the standard Brownian motion. Those relating to "Poisson space" are true in any Poisson space constructed on an underlying measurable space  $(\mathcal{X}, \mu)$  where  $\mu$  is a  $\sigma$ -finite measure, apart from the Clark formula, given here for  $(N_t)_{t \in \mathbf{R}_+}$  a Poisson point process on  $\mathbf{R}_+$  with intensity 1. Most of the identities stated in Gaussian spaces can be retrieved in Bakry, Gentil and Ledoux [12], Ledoux [82], Nourdin and Peccati [94] and for Poisson spaces in Last and Penrose [80], Peccati and Reinert [110], Wu [140]. The logarithmic Sobolev inequality for differentiable Wiener functionals  $F$ ,

$$\text{Ent}[F^2] \leq 2\mathbf{E} [\|\nabla F\|_{\mathcal{H}}^2], \quad (2.2.6)$$

initially stated by Leonard Gross in [62] can be retrieved via the application of the Clark formula. It does not exist in its original form (2.2.6) due to the absence of chain rule formula for discrete gradients. This also explains the lack of hypercontractivity property in Poisson and discrete settings, actually equivalent to (2.2.6). Ivan Nourdin, Giovanni Peccati and Xiaochuan Yang nevertheless ruled on an restricted hypercontractivity in the Poisson space in a recent work [91].

In the table below we review the identities established above, and compare them to their analogues in the Gaussian/Brownian and Poisson cases.

Table 2.1: Functional identities - Similarities and differences

	Gaussian space	Poisson space	IR variables
Chaos decomposition	$F = \mathbf{E}[F] + \sum_{n \geq 1} \frac{1}{n!} \delta^n \left( \mathbf{E} \left[ \nabla^{(n)} F \right] \right)$	$F = \mathbf{E}[F] + \sum_{n \geq 1} \frac{1}{n!} \delta^n \left( \mathbf{E} [D^{(n)} F] \right)$	No decomposition
Clark formula	$F = \mathbf{E}[F] + \int_0^\infty \mathbf{E} [D_t F   \mathcal{F}_t] dB_t$	$F = \mathbf{E}[F] + \int_0^\infty \mathbf{E} [D_t F   \mathcal{F}_t] (dN_t - dt)$	$F = \mathbf{E}[F] + \sum_{k=1}^\infty D_k \mathbf{E} [F   \mathcal{F}_k]$ .
Isometry formula	$\ \delta U\ _{L^2(\Omega)} = \ U\ _{L^2(\Omega)}^2 + \mathbf{E} [\text{tr}(\nabla U \circ \nabla U)]$	$\ \delta U\ _{L^2(\Omega)}^2 = \ U\ _{L^2(\Omega)}^2 + \mathbf{E} [\text{tr}(DU \circ DU)]$	$\ \delta U\ _{L^2(\Omega)}^2 = \mathbf{E} [\text{tr}(DU \circ DU)]$
Modified log-Sobolev inequality	$\text{Ent}(F) \leq \mathbf{E} [\langle \nabla F, \nabla \log F \rangle_{\mathcal{H}}]$	$\text{Ent}(F) \leq \mathbf{E} \left[ \int_{\mathbf{E}} \min(F^{-1}  D_x F ^2, D_x F D_x \log F) d\mu(x) \right]$	$\text{Ent}(F) \leq \sum_{k=1}^\infty \mathbf{E} \left[ \frac{ D_k F ^2}{\mathbf{E} [F   \mathcal{G}_k]} \right]$
Hypercontractivity	$\ P_t F\ _{1+e^{2t}(p-1)} \leq \ F\ _p$ ( $F \in L^2(\Omega)$ )	restricted hypercontractivity $\ P_t F\ _{1+e^t(p-1)} \leq \ F\ _p$ ( $F \geq 0, DF \leq 0$ on $\mathcal{X}$ )	No regularizing property

## 2.3 Dirichlet structures

Despite apparent dissimilarities of structures, such as the lack of chaotic decomposition in our construction (see remark 2.2.5), we nevertheless succeeded in connecting our elaboration to the preexisting theories for Brownian motion and Poisson processes; this through Dirichlet structures. Inspired by the binding identities between Malliavin calculus and Dirichlet forms, we provide to any countable product of probability spaces a Dirichlet structure naturally induced by our construction. In the two following subsections, we state that the usual Poisson and Brownian Dirichlet structures (see Bouleau and Hirsch [23]), associated to their respective gradient, can be retrieved as limits of the structures induced by our formalism. We borrow for that the idea of convergence of Dirichlet structures to Bouleau [22].

On  $(E_A, \mathbf{P}_A)$ , we have already implicitly built a Dirichlet structure, i.e. a Markov process  $X$ , a semi-group  $P$  and a generator  $L$  (see subsection 2.1.2). It remains to define the Dirichlet form  $\mathcal{E}_A$  such that  $\mathcal{E}_A(F) = \mathbf{E}[F LF]$  for any sufficiently regular functional  $F$ .

**Definition 2.3.1** (★). For  $F \in \mathbf{D}$ , define

$$\mathcal{E}_A(F) = \mathbf{E} \left[ \sum_{a \in A} |D_a F|^2 \right] = \|DF\|_{L^2(E_A \times A)}^2.$$

The integration by parts formula means that this form is closed. Since we do not assume any property on  $E_a$  for any  $a \in A$  and since we do not seem to have a product rule formula for the gradient, we cannot assert more properties for  $\mathcal{E}_A$ . However, following [22], we now show that we can reconstruct the usual gradient structures on Poisson and Wiener spaces as well chosen limits of our construction. For these two situations, we have a Polish space  $\mathscr{W}$ , equipped with  $\mathcal{B}$  its Borelean  $\sigma$ -field and a probability measure  $\mathbf{P}$ . There also exists a Dirichlet form  $\mathcal{E}$  defined on a set of functionals  $\mathbf{D}$ . Let  $(E_N, \mathcal{A}_N)$  be a sequence of Polish spaces, all equipped with a probability measure  $\mathbf{P}_N$  and their own Dirichlet form  $\mathcal{E}_N$ , defined on  $\mathbf{D}_N$ . Consider maps  $U_N$  from  $E_N$  into  $\mathscr{W}$  such that  $(U_N)_* \mathbf{P}_N$ , the pullback measure of  $\mathbf{P}_N$  by  $U_N$ , converges in distribution to  $\mathbf{P}$ . We assume that for any  $F \in \mathbf{D}$ , the map  $F \circ U_N$  belongs to  $\mathbf{D}_N$ . The image Dirichlet structure is defined as follows. For any  $F \in \mathbf{D}$ ,

$$\mathcal{E}^{U_N}(F) = \mathcal{E}_N(F \circ U_N).$$

We adapt the following definition from Bouleau [22].

**Definition 2.3.2.** With the previous notations, we say that  $((U_N)_* \mathbf{P}_N, N \in \mathbf{N}^*)$  converges as a Dirichlet distribution whenever for any  $F \in \text{Lip} \cap \mathbf{D}$ ,

$$\lim_{N \rightarrow \infty} \mathcal{E}^{U_N}(F) = \mathcal{E}(F).$$

### 2.3.1 Brownian motion

We now consider  $\mathbf{P}$  as the Wiener measure on  $\mathscr{W}_B = \mathcal{C}_0([0, 1]; \mathbf{R})$ . Denote  $\mathbf{T} = [0, 1]$ . Let  $(h_k, k \in \mathbf{N}^*)$  be an orthonormal basis of the Cameron-Martin space  $\mathfrak{H}^1$  (defined by (1.3.1)). For any function  $F : \mathscr{W}_B \rightarrow \mathbf{R}$  cylindrical of the form

$$F = f(W(g_1), \dots, W(g_n)),$$

where  $W(g)$  coincides with the Wiener integral of  $g$  and  $f$  belongs to the Schwartz space  $\mathcal{S}(\mathbf{R}^n)$ . As a reminder, for  $t \in \mathbf{R}_+$ ,

$$\mathbf{W}(\mathbf{1}_{[0,t]}) = B(t),$$

where  $B$  stands for the standard Brownian motion on  $\mathbf{R}$  and for  $h \in \mathfrak{H}^1$ ,

$$\nabla_h F = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(\delta_B v_1, \dots, \delta_B v_n) h_k. \quad (2.3.1)$$

The map  $\nabla$  is closable from  $L^2(\mathscr{W}_B)$  to  $L^2(\mathscr{W}_B; \mathfrak{H}^1)$ . Thus, it is meaningful to define  $\mathbf{D}_B^{1,2}$  as the closure of cylindrical functions for the norm

$$\|F\|_{1,2} = \left( \|F\|_{L^2(\mathscr{W}_B)}^2 + \|\nabla F\|_{L^2(\mathscr{W}_B; \mathfrak{H}^1)}^2 \right)^{\frac{1}{2}}.$$

**Definition 2.3.3.** A function  $F : \mathscr{W} \rightarrow \mathbf{R}$  is said to be H-C<sup>1</sup> if

- for almost all  $\omega \in \mathscr{W}_B$ ,  $h \mapsto F(\omega + h)$  is a continuous function on  $\mathfrak{H}^1$ ,
- for almost all  $\omega \in \mathscr{W}_B$ ,  $h \mapsto F(\omega + h)$  is continuously Fréchet differentiable and this Fréchet derivative is continuous from  $\mathfrak{H}^1$  into  $\mathbf{R} \otimes \mathfrak{H}^1$ .

We still denote by  $\nabla F$  the element of  $\mathfrak{H}^1$  such that

$$\left. \frac{d}{d\tau} F(\omega + \tau h) \right|_{\tau=0} = \langle \nabla F(\omega), h \rangle_{\mathfrak{H}^1}.$$

For  $N \in \mathbf{N}^*$ , let

$$e_k^N(t) = \sqrt{N} \mathbf{1}_{[(k-1)/N, k/N)}(t) \text{ and } h_k^N(t) = \int_0^t e_k^N(s) ds.$$

The family  $(h_k^N, k = 1, \dots, N)$  is then orthonormal in  $\mathfrak{H}^1$ . For  $(M_k, k = 1, \dots, N)$  a sequence of independent identically distributed random variables, centered with unit variance, the random walk

$$\omega^N(t) = \sum_{k=1}^N M_k h_k^N(t), \text{ for all } t \in [0, 1],$$

is known to converge in distribution in  $\mathscr{W}_B$  to  $\mathbf{P}$ . Let  $E_N = \mathbf{R}^N$  equipped with the product measure  $\mathbf{P}_N = \otimes_{k=1}^N \nu$  where  $\nu$  is the standard Gaussian measure on  $\mathbf{R}$ . We define the map  $U_N$  as follows:

$$U_N : E_N \longrightarrow \mathscr{W}_B$$

$$m = (m_1, \dots, m_N) \longmapsto \sum_{k=1}^N m_k h_k^N.$$

It follows from our definition that:



**Lemma 2.3.4** (★). For any  $F \in L^2(\mathscr{W}_B; \mathbf{R})$ ,

$$\mathcal{E}^{U_N}(F) = \sum_{k=1}^N \mathbf{E} \left[ \left( F(\omega^N) - \mathbf{E}' \left[ F(\omega_{(k)}^N + M'_k h_k^N) \right] \right)^2 \right],$$

where  $\omega_{(k)}^N = \omega^N - M_k h_k^N$  and  $M'_k$  is an independent copy of  $M_k$ . The expectation is taken on the product space  $\mathbf{R}^{N+1}$  equipped with the measure  $\mathbf{P}_N \otimes \nu$ .

The definition of Lipschitz function we use here is the following:

**Definition 2.3.5.** A function  $F : \mathscr{W}_B \rightarrow \mathbf{R}$  is said to be Lipschitz if it is  $H-C^1$  and for almost all  $\omega \in \mathscr{W}_B$ ,

$$|\langle \nabla F(\omega), h \rangle| \leq \|\dot{h}\|_{L^1}.$$

In particular since  $e_k^N \geq 0$ , this implies that

$$|\langle \nabla F(\omega), h_k^N \rangle| \leq h_k^N(1) - h_k^N(0) = \frac{1}{\sqrt{N}}; \quad k = 1, \dots, N.$$

For  $F \in \mathbf{D}_B \cap H-C^1$ , we have

$$F(\omega + h) - F(\omega) = \langle \nabla F(\omega), h \rangle_{\mathfrak{H}^1} + \|\dot{h}\|_{L^1} \varepsilon(\omega, h), \quad (2.3.2)$$

where  $\varepsilon(\omega, h)$  is bounded and goes to 0 in  $L^2$ , uniformly with as  $\|\dot{h}\|_{L^1}$  tends to 0.

**Theorem 2.3.6** (★). For any  $F \in \mathbf{D}_B \cap H-C^1$ ,

$$\mathcal{E}^{U_N}(F) \xrightarrow{N \rightarrow \infty} \mathbf{E} [\|\nabla F\|_{\mathfrak{H}^1}^2] = \mathcal{E}(F).$$

**Remark 2.3.7.** The error calculus theory is provided by the following principle: the variance of the *error* on a random variable  $X$ , represented by  $\Gamma(X)$ , is thus attached to it. Then if the sequence of pairs  $(X_n, \Gamma(X_n))$  converges suitably, it converges necessarily to a pair  $(X, \Gamma(X))$ . We can interpret the result of Theorem 2.3.6 in this perspective: the approximation of the Brownian motion by a random walk is validated by the convergence of the associated error structures.

### 2.3.2 Poisson point process

Let  $\mathbf{Y}$  be a compact Polish space and  $\mathfrak{N}_{\mathbf{Y}}$  be the set of weighted configurations, i.e. the set of integer valued, locally finite measures on  $\mathbf{Y}$ . Such a measure is of the form

$$\omega = \sum_{n=1}^{\infty} p_n \varepsilon_{\zeta_n},$$

where  $(\zeta_n)_{n \in \mathbf{N}^*}$  is a set of distinct points in  $\mathbf{Y}$  with no accumulation point,  $(p_n)_{n \in \mathbf{N}^*}$  any sequence of positive integers. The topology on  $\mathfrak{N}_{\mathbf{Y}}$  is defined by the semi-norms

$$p_f(\omega) = \left| \sum_{n=1}^{\infty} p_n f(\zeta_n) \right|,$$

when  $f$  runs through the set of continuous functions on  $\mathbf{Y}$ . It is known (see for instance [72]) that  $\mathfrak{N}_{\mathbf{Y}}$  is then a Polish space for this topology. For some finite measure  $\mathbf{M}$  on  $\mathbf{Y}$ , we put on  $\mathfrak{N}_{\mathbf{Y}}$ , the probability measure  $\mathbf{P}$  such that the canonical process is a Poisson point process of control measure  $\mathbf{M}$ , which we consider without loss of generality, to have total mass  $\mathbf{M}(\mathbf{Y}) = 1$ .

On  $\mathfrak{N}_{\mathbf{Y}}$ , it is customary to consider the difference gradient (see [37, 103, 117]): For any  $x \in \mathbf{Y}$ , any  $\omega \in \mathfrak{N}_{\mathbf{Y}}$ ,

$$D_x F(\omega) = F(\omega + \varepsilon_x) - F(\omega). \quad (2.3.3)$$

Set

$$\mathbf{D}_P = \left\{ F : \mathfrak{N}_{\mathbf{Y}} \rightarrow \mathbf{R} \text{ such that } \mathbf{E} \left[ \int_{\mathbf{Y}} |D_x F|^2 d\mathbf{M}(x) \right] < \infty \right\},$$

and for any  $F \in \mathbf{D}_P$ ,

$$\mathcal{E}(F) = \mathbf{E} \left[ \int_{\mathbf{Y}} |D_x F|^2 d\mathbf{M}(x) \right]. \quad (2.3.4)$$

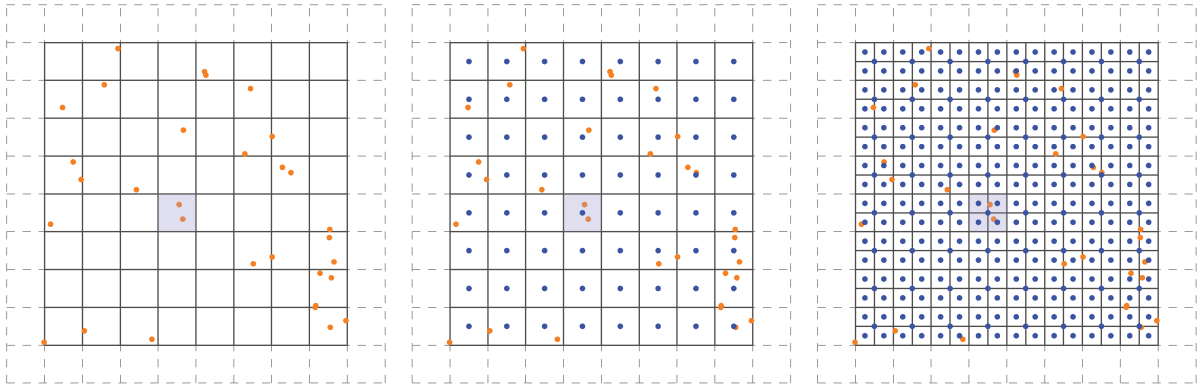


Figure 1

Figure 2

Figure 3

To see the Poisson point process as a Dirichlet limit, the idea is to partition the set  $\mathbf{Y}$  into  $N$  parts,  $C_1^N, \dots, C_N^N$  such that  $\mathbf{M}(C_k^N) = p_k^N$  (Figure 1) and then for each  $k \in \{1, \dots, N\}$ , take a point  $\zeta_k^N$  into  $C_k^N$  (Figure 2) so that the Poisson point process  $\omega$  on  $\mathbf{Y}$  with intensity measure  $\mathbf{M}$  is approximated by

$$\omega^N = \sum_{k=1}^N \omega(C_k^N) \varepsilon_{\zeta_k^N}.$$

We denote by  $\mathbf{P}_N$  the distribution of  $\omega^N$ . Repeat the procedure for  $N$  each time bigger (Figure 3). By computing its Laplace transform, it is clear that  $\mathbf{P}_N$  converges in distribution to  $\mathbf{P}$ . It remains to see this convergence holds in the Dirichlet sense for the sequence of Dirichlet structures induced by our approach for independent random variables.

Let  $(\zeta_k^N, k = 1, \dots, N)$  (respectively  $(p_k^N, k = 1, \dots, N)$ ) be a triangular array of points in  $\mathbf{Y}$  (respectively of non-negative numbers) such that the following two properties hold:

1) the  $p_k^N$ 's tends to 0 uniformly:

$$p^N = \sup_{k \leq N} p_k^N = O\left(\frac{1}{N}\right); \quad (2.3.5)$$

2) the  $\zeta_k^N$ 's are sufficiently well spread so that we have convergence of Riemann sums : For any continuous and  $\mathbf{M}$ -integrable function  $f : \mathbf{Y} \rightarrow \mathbf{R}$ , we have

$$\sum_{k=1}^N f(\zeta_k^N) p_k^N \xrightarrow{N \rightarrow \infty} \int f(x) d\mathbf{M}(x). \quad (2.3.6)$$

Take  $f = 1$  implies that  $\sum_k p_k^N$  tends to 1 as  $N$  goes to infinity.

For any  $N \in \mathbf{N}^*$  and any  $k \in \{1, \dots, N\}$ , let  $\mu_k^N$  be the Poisson distribution on  $\mathbf{N}$ , of parameter  $p_k^N$ . In this situation, let  $E_N = \mathbf{N}^N$  with  $\mu^N = \otimes_{k=1}^N \mu_k^N$ . That means we have independent random variables  $M_1^N, \dots, M_N^N$ , where  $M_k^N$  follows a Poisson distribution of parameter  $p_k^N$  for any  $k \in \{1, \dots, N\}$ . We turn these independent random variables into a point process by the map  $U_N$  defined as

$$\begin{aligned} U_N : \mathbf{N}^N &\longrightarrow \mathfrak{N}_{\mathbf{Y}} \\ (m_1, \dots, m_N) &\longmapsto \sum_{k=1}^N m_k \varepsilon_{\zeta_k^N}. \end{aligned}$$

**Lemma 2.3.8** (★). *For any  $F \in \mathbf{D}_P$ ,*

$$\mathcal{E}^{U_N}(F) = \sum_{m=1}^N \sum_{\ell=0}^{\infty} \mathbf{E} \left[ \left( \sum_{\tau=0}^{\infty} \left( F(\omega_{(m)}^N + \ell \varepsilon_{\zeta_m^N}) - F(\omega_{(m)}^N + \tau \varepsilon_{\zeta_m^N}) \right) \mu_m^N(\tau) \right)^2 \right] \mu_m^N(\ell), \quad (2.3.7)$$

where  $\omega_{(m)}^N = \sum_{k \neq m} M_k^N \varepsilon_{\zeta_k^N}$ .

*Proof.* According its very definition,

$$\mathcal{E}^{U_N}(F) = \sum_{m=1}^N \mathbf{E} \left[ \left( F(\omega_{(m)}^N + M_m^N \varepsilon_{\zeta_m^N}) - \sum_{\tau=0}^{\infty} F(\omega_{(m)}^N + \tau \varepsilon_{\zeta_m^N}) \mu_m^N(\tau) \right)^2 \right].$$

The result follows by conditioning with respect to  $M_m^N$ , whose law is  $\mu_m^N$ .  $\square$

Since the vague topology on  $\mathfrak{N}_{\mathbf{Y}}$  is metrizable, one could define Lipschitz functions with respect to this distance. However, this turns out to be not sufficient for the convergence to hold.

**Definition 2.3.9.** A function  $F : \mathfrak{N}_{\mathbf{Y}} \rightarrow \mathbf{R}$  is said to be TV – Lip if  $F$  is continuous for the vague topology and if for any  $\omega, \eta \in \mathfrak{N}_{\mathbf{Y}}$ ,

$$|F(\omega) - F(\eta)| \leq \text{dist}_{\text{TV}}(\omega, \eta),$$

where  $\text{dist}_{\text{TV}}$  represents the distance in total variation between two point measures, i.e. the number of distinct points counted with multiplicity.

**Theorem 2.3.10** (★). *For any  $F \in \text{TV – Lip} \cap \mathbf{D}_P$ , with the notations of Lemma [2.3.8] and (2.3.4),*

$$\mathcal{E}^N(F) \xrightarrow{N \rightarrow \infty} \mathcal{E}(F).$$

## 2.4 Proofs of chapter 2

### 2.4.1 Proofs of Section 2.1

*Proof of Theorem 2.1.6.* The process  $\text{trace}(\text{DU}) = (D_a U_a)_{a \in B}$  belongs to  $L^2(\mathbf{E}_A \times A)$ : using the Jensen inequality, we have:

$$\|\text{trace}(\text{DU})\|_{L^2(\mathbf{E}_A \times A)}^2 = \mathbf{E} \left[ \sum_{a \in B} |D_a U_a|^2 \right] \leq 2 \sum_{a \in B} \mathbf{E} [U_a^2] < \infty. \quad (2.4.1)$$

Moreover,

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

since the conditional expectation is a projection in  $L^2(\mathbf{E}_A)$ .  $\square$

*Proof of corollary 2.1.7.* Let  $(F_n)_{n \in \mathbf{N}^*}$  be a sequence of random variables defined on  $\mathcal{S}$  such that  $F_n$  converges to 0 in  $L^2(\mathbf{E}_A)$  and the sequence  $\text{DF}_n$  converges to  $\eta$  in  $L^2(\mathbf{E}_A \times A)$ . Let  $\text{U}$  be a simple process. From the integration by parts formula (2.1.1)

$$\mathbf{E} \left[ \sum_{a \in A} D_a F_n U_a \right] = \mathbf{E} \left[ F_n \sum_{a \in A} D_a U_a \right]$$

where  $\sum_{a \in A} D_a U_a \in L^2(\mathbf{E}_A)$  in view of (2.4.1). Then,

$$\langle \eta, \text{U} \rangle_{L^2(\mathbf{E}_A \times A)} = \lim_{n \rightarrow \infty} \mathbf{E} \left[ F_n \sum_{a \in A} D_a U_a \right] = 0,$$

for any simple process  $\text{U}$ . It follows that  $\eta = 0$  and then the operator  $\text{D}$  is closable from  $L^2(\mathbf{E}_A)$  to  $L^2(\mathbf{E}_A \times A)$ .  $\square$

*Proof of Lemma 2.1.8.* Since  $\sup_n \|\text{DF}_n\|_{\mathbf{D}}$  is finite, there exists a subsequence which we still denote by  $(\text{DF}_n, n \geq 1)$  weakly convergent in  $L^2(\mathbf{E}_A \times A)$  to some limit denoted by  $\eta$ . For  $k > 0$ , let  $n_k$  be such that  $\|F_m - F\|_{L^2} < 1/k$  for  $m \geq n_k$ . The Mazur's Theorem implies that there exists a convex combination of elements of  $(\text{DF}_m, m \geq n_k)$  such that

$$\left\| \sum_{i=1}^{M_k} \alpha_i^k \text{DF}_{m_i} - \eta \right\|_{L^2(\mathbf{E}_A \times A)} < 1/k.$$

Moreover, since the  $\alpha_i^k$  are positive and sums to 1,

$$\left\| \sum_{i=1}^{M_k} \alpha_i^k F_{m_i} - F \right\|_{L^2(\mathbf{E}_A)} \leq 1/k.$$

We have thus constructed a sequence

$$F^k = \sum_{i=1}^{M_k} \alpha_i^k F_{m_i}$$

such that  $F^k$  tends to  $F$  in  $L^2$  and  $DF^k$  converges in  $L^2(E_A \times A)$  to a limit. By the construction of  $\mathbf{D}$ , this means that  $F$  belongs to  $\mathbf{D}$  and that  $DF = \eta$ .  $\square$

*Proof of Lemma 2.1.11.* For  $U$  and  $V$  in  $\mathcal{S}_0(\ell^2(A))$ , from the integration by parts formula,

$$\begin{aligned} \mathbf{E} [\delta U \delta V] &= \langle D\delta(U), V \rangle_{L^2(E_A \times A)} \\ &= \mathbf{E} \left[ \sum_{a \in A} D_a(\delta U) V_a \right] \\ &= \mathbf{E} \left[ \sum_{(a,b) \in A^2} V_a D_a(D_b U_b) \right] \\ &= \mathbf{E} \left[ \sum_{(a,b) \in A^2} V_a D_b(D_a U_b) \right] \\ &= \mathbf{E} \left[ \sum_{(a,b) \in A^2} (D_b V_a) (D_a U_b) \right] = \mathbf{E} [\text{trace}(DU \circ DV)]. \end{aligned}$$

It follows that  $\mathbf{E} [\delta U^2] \leq \|U\|_{\mathbf{D}(\ell^2(A))}^2$ . Then, by density,  $\mathbf{D}(\ell^2(A)) \subset \text{Dom } \delta$  and Eqn. (2.1.2) holds for  $U$  and  $V$  in  $\text{Dom } \delta$ .  $\square$

*Proof of Theorem 2.1.15.* To prove the existence of  $(P_t)_{t \in \mathbf{R}_+}$  for a countable set, we apply the Hille-Yosida theorem:

**Theorem 2.4.1** (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  $L^2(E_A)$ .

We know that  $\mathcal{S} \subset \text{Dom } L$  and that  $\mathcal{S}$  is dense in  $L^2(E_A)$ , then so does  $\text{Dom } L$ .

Let  $(A_n)_{n \in \mathbf{N}^*}$  an increasing sequence of subsets of  $A$  such that  $\bigcup_{n \geq 1} A_n = A$ . For  $F \in L^2(E_A)$ , let  $F_n = \mathbf{E}[F | \mathcal{F}_{A_n}]$ . Since  $(F_n)_{n \in \mathbf{N}^*}$  is a square integrable  $\mathcal{F}$ -martingale,  $F_n$  converges to  $F$  both almost-surely and in  $L^2(E_A)$ . For any  $n \in \mathbf{N}^*$ ,  $F_n$  depends only on  $X_{A_n}$ . Abusing the notation, we still denote by  $F_n$  its restriction to  $E_{A_n}$  so that we can consider  $L_n F_n$  where  $L_n$  is

defined as above on  $E_{A_n}$ . Moreover, according to Lemma 2.2.2,  $D_a F_n = \mathbf{E}[D_a F | \mathcal{F}_{A_n}]$ , hence

$$\begin{aligned} \lambda^2 \|F_n\|_{L^2(E_A)}^2 &\leq \|\lambda F_n - L_n F_n\|_{L^2(E_{A_n})}^2 = \mathbf{E} \left[ \left( \lambda F_n + \sum_{a \in A} D_a F_n \right)^2 \right] \\ &= \mathbf{E} \left[ \mathbf{E} \left[ \lambda F + \sum_{a \in A} D_a F \mid \mathcal{F}_{A_n} \right]^2 \right] \xrightarrow{n \rightarrow \infty} \|\lambda F - L F\|_{L^2(E_A)}^2. \end{aligned}$$

Therefore, point (2) is satisfied.

Since  $A_n$  is finite, there exists  $G_n \in L^2(E_{A_n})$  such that

$$\begin{aligned} F_n &= (\lambda \text{Id} - L_n) G_n(X_{A_n}) = \lambda G_n(X_{A_n}) + \sum_{a \in A_n} D_a G_n(X_{A_n}) \\ &= \lambda \tilde{G}_n(X_A) + \sum_{a \in A_n} D_a \tilde{G}_n(X_A) = \lambda \tilde{G}_n(X_A) + \sum_{a \in A} D_a \tilde{G}_n(X_A), \end{aligned}$$

where  $\tilde{G}_n(X_A) = G_n(X_{A_n})$  depends only on the components whose index belongs to  $A_n$ . This means that  $F_n$  belongs to the range of  $\lambda \text{Id} - L$  and we already know it converges in  $L^2(E_A)$  to  $F$ .  $\square$

*Proof of Theorem 2.1.17.* For  $A$  finite,  $a \in A$ , let  $\mathcal{G}_{a,t} := \sigma(X_{b,t}^{x_b}, b \neq a)$  so that

$$\begin{aligned} D_a P_t F(x) &= P_t F(x) - \mathbf{E}[P_t F(x) | \mathcal{G}_{a,t}] \\ &= \mathbf{E}[F(X_t) | X_0 = x] - \mathbf{E}[\mathbf{E}[F(X_t) | X_0 = x] | \mathcal{G}_{a,t}] \\ &= \mathbf{E}[F(X_t) | X_0 = x] - \mathbf{E}[\mathbf{E}[F(X_t) | \mathcal{G}_{a,t}] | X_0 = x] \\ &= P_t D_a F(x). \end{aligned}$$

For  $A$  infinite, let  $(A_n)_{n \in \mathbf{N}^*}$  an increasing sequence of finite subsets of  $A$  such that  $\bigcup_{n \geq 1} A_n = A$ . For  $F \in L^2(E_A)$ , let  $F_n = \mathbf{E}[F | \mathcal{F}_{A_n}]$ . Since  $P$  is a contraction semi-group, for any  $t$ ,  $P_t F_n$  tends to  $P_t F$  in  $L^2(E_A)$  as  $n$  goes to infinity. From the Mehler formula, we know that  $P_t F_n = P_t^n F_n$  where  $P^n$  is the semi-group associated to  $P_n$ , hence

$$D_a P_t F_n = D_a P_t^n F_n = P_t^n D_a F_n. \quad (2.4.2)$$

Moreover,

$$\begin{aligned} \mathbf{E} \left[ \sum_{a \in A_n} |D_a P_t F_n|^2 \right] &= \sum_{a \in A_n} \mathbf{E}[|P_t D_a F_n|^2] \\ &\leq \sum_{a \in A_n} \mathbf{E}[|D_a F_n|^2] \\ &= \sum_{a \in A_n} \mathbf{E}[|\mathbf{E}[D_a F | \mathcal{F}_{A_n}]|^2] \\ &\leq \sum_{a \in A_n} \mathbf{E}[|D_a F|^2] \\ &\leq \|DF\|_{\mathbf{D}}^2. \end{aligned}$$

According to Lemma [2.1.8], this means that  $P_t F$  belongs to  $\mathbf{D}$ . Let  $n$  go to infinity in (2.4.2) yields (2.1.5).

**Remark 2.4.2.** An alternative proof consists in using the identity  $P_t F = e^{-tL} F$  and then the equivalent expression of  $P_t$  as

$$P_t F = \sum_{n \in \mathbf{N}} \frac{t^n}{n!} L^n F. \quad (2.4.3)$$

Moreover, as follows from  $D_a D_a = D_a$  and  $D_b D_a = D_a D_b$  for any  $a, b \in A$  that

$$L D_a F = D_a D_a F \sum_{b \in A, b \neq a} D_b D_a F = D_a \left( D_a F + \sum_{b \in A, b \neq a} D_b F \right) = D_a L F,$$

we get by induction that

$$L^n D_a = L^n D_a$$

on  $\mathcal{S}$ . We conclude by plotting this in (2.4.3) and using a density argument.  $\square$

## 2.4.2 Proofs of Section 2.2

*Proof of Lemma 2.2.1.* Let  $k \in A$ . Assume that  $F \in \mathcal{F}_k$ . Then, for every  $n > k$ ,  $F$  is  $\mathcal{G}_n$ -measurable and  $D_n F = 0$ .

Let  $F \in \mathbf{D}$  such that  $D_n F = 0$  for every  $n > k$ . Then  $F$  is  $\mathcal{G}_n$ -measurable for any  $n > k$ . From the equality  $\mathcal{F}_k = \bigcap_{n > k} \mathcal{G}_n$ , it follows that  $F$  is  $\mathcal{F}_k$ -measurable.  $\square$

*Proof of Lemma 2.2.2.* For any  $k \in \mathbf{N}^*$ ,  $\mathcal{F}_k \cap \mathcal{G}_k = \mathcal{F}_{k-1}$ , hence

$$D_k \mathbf{E}[F | \mathcal{F}_k] = \mathbf{E}[F | \mathcal{F}_k] - \mathbf{E}[F | \mathcal{F}_{k-1}] = \mathbf{E}[D_k F | \mathcal{F}_k].$$

The proof is thus complete.  $\square$

*Proof of Theorem 2.2.3.* Let  $F$  a  $\mathcal{F}_n$ -measurable random variable. It is clear that

$$F - \mathbf{E}[F] = \sum_{k=1}^n (\mathbf{E}[F | \mathcal{F}_k] - \mathbf{E}[F | \mathcal{F}_{k-1}]) = \sum_{k=1}^n D_k \mathbf{E}[F | \mathcal{F}_k].$$

For  $F \in \mathbf{D}$ , apply this identity to  $F_n = \mathbf{E}[F | \mathcal{F}_n]$  to obtain

$$F_n - \mathbf{E}[F] = \sum_{k=1}^n D_k \mathbf{E}[F | \mathcal{F}_k].$$

Remark that for  $\ell > k$ , in view of Lemma 2.2.1,

$$\mathbf{E}[D_k \mathbf{E}[F | \mathcal{F}_k] D_\ell \mathbf{E}[F | \mathcal{F}_\ell]] = \mathbf{E}[D_\ell D_k \mathbf{E}[F | \mathcal{F}_k] \mathbf{E}[F | \mathcal{F}_\ell]] = 0, \quad (2.4.4)$$

since  $D_k \mathbf{E}[F | \mathcal{F}_k]$  is  $\mathcal{F}_k$ -measurable. Hence, we get

$$\mathbf{E}[|F - \mathbf{E}[F]|^2] \geq \mathbf{E}[|F_n - \mathbf{E}[F]|^2] = \sum_{k=1}^n \mathbf{E}[D_k \mathbf{E}[F | \mathcal{F}_k]^2].$$

Thus, the sequence  $(D_k \mathbf{E}[F | \mathcal{F}_k], k \in \mathbf{N}^*)$  belongs to  $\ell^2(\mathbf{N})$  and the result follows by a limit procedure.  $\square$

*Proof of Corollary 2.2.4.* We now analyze the non-ordered situation. If  $A$  is finite, each bijection between  $A$  and  $\{1, \dots, n\}$  defines an order on  $A$ . Hence, there are  $|A|!$  possible filtrations. Each term of the form

$$D_{i_k} \mathbf{E} [F | X_{i_1}, \dots, X_{i_k}]$$

appears  $(k-1)! (|A| - k)!$  times since the order of  $X_{i_1}, \dots, X_{i_{k-1}}$  is irrelevant to the conditioning. The result follows by summation then renormalization of the identities obtained for each filtration.  $\square$

*Proof of Theorem 2.2.6.* Remark that

$$D_k \mathbf{E} [F | \check{\mathcal{F}}_{k-1}^N] = \mathbf{E} [F | \check{\mathcal{F}}_{k-1}^N] - \mathbf{E} [F | \check{\mathcal{F}}_{k-1}^N \cap \mathcal{G}_k] = \mathbf{E} [F | \check{\mathcal{F}}_{k-1}^N] - \mathbf{E} [F | \check{\mathcal{F}}_k^N].$$

For  $F \in \mathcal{F}_N$ , since the successive terms collapse, we get

$$F - \mathbf{E} [F] = \mathbf{E} [F | \check{\mathcal{F}}_0^N] - \mathbf{E} [F | \check{\mathcal{F}}_N^N] = \sum_{k=1}^N D_k \mathbf{E} [F | \check{\mathcal{F}}_{k-1}^N] = \sum_{k=1}^{\infty} D_k \mathbf{E} [F | \check{\mathcal{F}}_{k-1}^N],$$

by the very definition of the gradient map. As in (2.4.4), we can show that for any  $N$ ,

$$\mathbf{E} [D_k \mathbf{E} [F | \check{\mathcal{F}}_{k-1}^N] D_\ell \mathbf{E} [F | \check{\mathcal{F}}_{\ell-1}^N]] = 0, \text{ for } k \neq \ell.$$

Consider  $F_N = \mathbf{E} [F | \mathcal{F}_N]$  and proceed as in the proof of Lemma 2.2.3 to conclude.  $\square$

*Proof of Corollary 2.2.7.* According to (2.4.4) and (2.2.1), we have

$$\begin{aligned} \text{Var}(F) &= \mathbf{E} \left[ \left| \sum_{k \in A} D_k \mathbf{E} [F | \mathcal{F}_k] \right|^2 \right] \\ &= \mathbf{E} \left[ \sum_{k \in A} |D_k \mathbf{E} [F | \mathcal{F}_k]|^2 \right] \\ &= \mathbf{E} \left[ \sum_{k \in A} |\mathbf{E} [D_k F | \mathcal{F}_k]|^2 \right] \\ &\leq \mathbf{E} \left[ \sum_{k \in A} \mathbf{E} [ |D_k F|^2 | \mathcal{F}_k ] \right] = \mathbf{E} \left[ \sum_{k \in A} |D_k F|^2 \right], \end{aligned}$$

where the inequality follows from then Jensen inequality.  $\square$

*Proof of Theorem 2.2.8.* Let  $F, G \in \mathbf{D}$ , the Clark formula entails

$$\begin{aligned} \text{cov}(F, G) &= \mathbf{E} [(F - \mathbf{E} [F])(G - \mathbf{E} [G])] \\ &= \mathbf{E} \left[ \sum_{k, \ell \in A} D_k \mathbf{E} [F | \mathcal{F}_k] D_\ell \mathbf{E} [G | \mathcal{F}_\ell] \right] \\ &= \mathbf{E} \left[ \sum_{k \in A} D_k \mathbf{E} [F | \mathcal{F}_k] D_k \mathbf{E} [G | \mathcal{F}_k] \right] \\ &= \mathbf{E} \left[ \sum_{k \in A} D_k F D_k \mathbf{E} [G | \mathcal{F}_k] \right] \end{aligned}$$



where we have used (2.4.4) in the third equality and the identity  $D_k D_k = D_k$  in the last one.  $\square$

*Proof of Theorem 2.2.9.* Let  $F, G \in L^2(E_A)$ .

$$\begin{aligned} \text{cov}(F, G) &= \mathbf{E} \left[ \sum_{k \in A} (D_k \mathbf{E}[F | \mathcal{F}_k]) (D_k \mathbf{E}[G | \mathcal{F}_k]) \right] \\ &= \mathbf{E} \left[ \sum_{k \in A} D_k \mathbf{E}[F | \mathcal{F}_k] \left( - \int_0^\infty \text{LP}_t \mathbf{E}[G | \mathcal{F}_k] dt \right) \right] \\ &= \int_0^\infty \mathbf{E} \left[ \sum_{k \in A} D_k \mathbf{E}[F | \mathcal{F}_k] \left( \sum_{\ell \in A} D_\ell P_t \mathbf{E}[G | \mathcal{F}_k] dt \right) \right] \\ &= \int_0^\infty e^{-t} \mathbf{E} \left[ \sum_{k \in A} (D_k F) (P_t D_k \mathbf{E}[G | \mathcal{F}_k]) \right] dt, \end{aligned}$$

when we have used the orthogonality of the sum, (2.1.5) and the  $\mathcal{F}_k$ -measurability of  $P_t D_k \mathbf{E}[G | \mathcal{F}_k]$  to get the last equality.  $\square$

*Proof of Theorem 2.2.10.* Assume with no loss of generality that  $F$  is centered. Apply (2.2.2) to  $\theta F$  and  $e^{\theta F}$  ( $\theta \in \mathbf{R}_+^*$ ),

$$\begin{aligned} \theta \left| \mathbf{E}[F e^{\theta F}] \right| &= \theta \left| \mathbf{E} \left[ \sum_{k \in A} D_k F D_k \mathbf{E}[e^{\theta F} | \mathcal{F}_k] \right] \right| \\ &\leq \theta \sum_{k \in A} \mathbf{E} \left[ |D_k F| \left| D_k \mathbf{E}[e^{\theta F} | \mathcal{F}_k] \right| \right]. \end{aligned}$$

Recall that

$$\begin{aligned} D_k \mathbf{E}[e^{\theta F} | \mathcal{F}_k] &= \mathbf{E}' \left[ \mathbf{E}[e^{\theta F} | \mathcal{F}_k] - \mathbf{E}[e^{\theta F(X_{-k}, X'_k)} | \mathcal{F}_k] \right] \\ &= \mathbf{E} \left[ \mathbf{E}'[e^{\theta F} - e^{\theta F(X_{-k}, X'_k)}] | \mathcal{F}_k \right] \\ &= \mathbf{E} \left[ e^{\theta F} \mathbf{E}'[1 - e^{-\theta \Delta_k F}] | \mathcal{F}_k \right] \end{aligned}$$

where  $\Delta_k F = F - F(X_{-k}, X'_k)$  so that  $D_k F = \mathbf{E}'[\Delta_k F]$ .

Since  $(x \mapsto 1 - e^{-x})$  is concave, we get

$$D_k \mathbf{E}[e^{\theta F} | \mathcal{F}_k] \leq \mathbf{E} \left[ e^{\theta F} (1 - e^{-\theta D_k F}) | \mathcal{F}_k \right] \leq \theta \mathbf{E} \left[ e^{\theta F} |D_k F| | \mathcal{F}_k \right].$$

Thus,

$$\left| \mathbf{E}[F e^{\theta F}] \right| \leq \theta \mathbf{E} \left[ e^{\theta F} \sum_{k=1}^\infty |D_k F| \mathbf{E}[|D_k F| | \mathcal{F}_k] \right] \leq M \theta \mathbf{E}[e^{\theta F}].$$

By Gronwall lemma, this implies that

$$\mathbf{E}[e^{\theta F}] \leq \exp \left( \frac{\theta^2}{2} M \right).$$

Hence,

$$\mathbf{P}(F - \mathbf{E}[F] \geq x) = \mathbf{P}(e^{\theta(F - \mathbf{E}[F])} \geq e^{\theta x}) \leq \exp\left(-\theta x + \frac{\theta^2}{2} M\right).$$

Optimize with respect to  $\theta$  gives  $\theta_{\text{opt}} = x/M$ , hence the result.  $\square$

*Proof of Theorem 2.2.11.* We follow closely the proof of [140] for Poisson process. Let  $G \in L^2(\mathbf{E}_A)$  be a positive random variable such that  $DG \in L^2(\mathbf{E}_A \times A)$ . For any non-zero integer  $n$ , define  $G_n = \min(\max(\frac{1}{n}, G), n)$ , for any  $k$ ,  $L_k = \mathbf{E}[G_n | \mathcal{F}_k]$  and  $L_0 = \mathbf{E}[G_n]$ . We have,

$$\begin{aligned} L_n \log L_n - L_0 \log L_0 &= \sum_{k=0}^{n-1} L_{k+1} \log L_{k+1} - L_k \log L_k \\ &= \sum_{k=0}^{n-1} \log L_k (L_{k+1} - L_k) + \sum_{k=0}^{n-1} L_{k+1} (\log L_{k+1} - \log L_k). \end{aligned}$$

Note that  $(\log L_n (L_{n+1} - L_n))_{n \in \mathbf{N}}$  and  $(L_{n+1} - L_n)_{n \in \mathbf{N}}$  are  $(\mathcal{F}_n)_{n \in \mathbf{N}}$ -martingales, hence

$$\begin{aligned} &\mathbf{E}[L_n \log L_n - L_0 \log L_0] \\ &= \mathbf{E}\left[\sum_{k=0}^{n-1} L_{k+1} \log L_{k+1} - L_{k+1} \log L_k - L_{k+1} + L_k\right] \\ &= \mathbf{E}\left[\sum_{k=0}^{n-1} L_{k+1} \log L_{k+1} - L_k \log L_k - (\log L_k + 1)(L_{k+1} - L_k)\right] \\ &= \mathbf{E}\left[\sum_{k=0}^{n-1} \ell(L_k, L_{k+1} - L_k)\right], \end{aligned}$$

where the function  $\ell$  is defined on  $\Theta = \{(x, y) \in \mathbf{R}^2 : x > 0, x + y > 0\}$  by

$$\ell(x, y) = (x + y) \log(x + y) - x \log x - (\log x + 1)y.$$

Since  $\ell$  is convex on  $\Theta$ , it comes from the Jensen inequality for conditional expectations that

$$\begin{aligned} \sum_{k=0}^{n-1} \mathbf{E}[\ell(L_k, L_{k+1} - L_k)] &= \sum_{k=0}^{n-1} \mathbf{E}[\ell(\mathbf{E}[G_n | \mathcal{F}_k], D_{k+1} \mathbf{E}[G_n | \mathcal{F}_{k+1}])] \\ &= \sum_{k=1}^n \mathbf{E}[\ell(\mathbf{E}[G_n | \mathcal{F}_{k-1}], \mathbf{E}[D_k G_n | \mathcal{F}_k])] \\ &\leq \sum_{k=1}^n \mathbf{E}[\mathbf{E}[\ell(\mathbf{E}[G_n | \mathcal{G}_k], D_k G_n) | \mathcal{F}_k]] \\ &= \sum_{k=1}^n \mathbf{E}[\ell(\mathbf{E}[G_n | \mathcal{G}_k], D_k G_n)] \\ &= \sum_{k=1}^{\infty} \mathbf{E}[\ell(\mathbf{E}[G_n | \mathcal{G}_k], D_k G_n)]. \end{aligned}$$

We know from [140] that for any non-zero integer  $k$ ,  $\ell(\mathbf{E}[G_n | \mathcal{G}_k], D_k G_n)$  converges increasingly to  $\ell(\mathbf{E}[G | \mathcal{G}_k], D_k G)$   $\mathbf{P}$ -a.s., hence by Fatou Lemma,

$$\mathbf{E}[G \log G] - \mathbf{E}[G] \log \mathbf{E}[G] \leq \sum_{k=1}^{\infty} \mathbf{E}[\ell(\mathbf{E}[G | \mathcal{G}_k], D_k G)].$$

Furthermore, for any  $(x, y) \in \Theta$ ,  $\ell(x, y) \leq |y|^2/x$ , then,

$$\mathbf{E}[G \log G] - \mathbf{E}[G] \log \mathbf{E}[G] \leq \sum_{k=1}^{\infty} \mathbf{E}\left[\frac{|D_k G|^2}{\mathbf{E}[G | \mathcal{G}_k]}\right].$$

The proof is thus complete.  $\square$

*Proof of Theorem 2.2.12.* We first prove the uniqueness. Let  $(\varphi, V)$  and  $(\varphi', V')$  two convenient couples. We have  $D_a(\varphi - \varphi') = V'_a - V_a$  for any  $a \in A$  and  $\sum_{a \in A} D_a(V'_a - V_a) = 0$ , hence

$$0 = \mathbf{E}\left[(\varphi - \varphi') \sum_{a \in A} D_a(V'_a - V_a)\right] = \mathbf{E}\left[\sum_{a \in A} D_a(\varphi - \varphi')(V'_a - V_a)\right] = \mathbf{E}\left[\sum_{a \in A} (V'_a - V_a)^2\right].$$

This implies that  $V = V'$  and  $D(\varphi - \varphi') = 0$ . The Clark formula (Theorem 2.2.3) entails that  $0 = \mathbf{E}[\varphi - \varphi'] = \varphi - \varphi'$ .

We now prove the existence. Since  $\mathbf{E}[D_a \varphi | \mathcal{G}_a] = 0$ , we can choose

$$V_a = \mathbf{E}[U_a | \mathcal{G}_a],$$

which implies  $D_a \varphi = D_a U_a$ , and guarantees  $\delta V = 0$ . Choose any ordering of the elements of  $A$  and remark that, in view of (2.4.4),

$$\begin{aligned} \mathbf{E}\left[\left(\sum_{k=1}^{\infty} \mathbf{E}[D_k U_k | \mathcal{F}_k]\right)^2\right] &= \mathbf{E}\left[\left(\sum_{k=1}^{\infty} D_k \mathbf{E}[U_k | \mathcal{F}_k]\right)^2\right] \\ &= \mathbf{E}\left[\sum_{k=1}^{\infty} \left(D_k \mathbf{E}[U_k | \mathcal{F}_k]\right)^2\right] \leq \sum_{k=1}^{\infty} \mathbf{E}[|D_k U_k|^2] \leq \|U\|_{\mathbf{D}(\ell^2(A))}^2, \end{aligned}$$

hence

$$\varphi = \sum_{k=1}^{\infty} \mathbf{E}[D_k U_k | \mathcal{F}_k],$$

defines a square integrable random variable of null expectation, which satisfies the required property.  $\square$

### *Proofs of Section 2.3*

*Proof of Theorem 2.3.6.* For  $F \in \mathbf{D}_B \cap H-C^1$ , in view of (2.3.2), we have

$$F(\omega^N) - F(\omega_{(k)}^N + M'_k h_k^N) = (M_k - M'_k) \langle \nabla F(\omega_{(k)}^N), h_k^N \rangle_{\mathfrak{H}^1} + \frac{M_k - M'_k}{\sqrt{N}} \varepsilon(\omega_{(k)}^N, h_k^N).$$

Hence,

$$\begin{aligned} & \sum_{k=1}^N \mathbf{E} \left[ \left( F(\omega^N) - \mathbf{E}' \left[ F(\omega_{(k)}^N) + M'_k h_k^N \right] \right)^2 \right] \\ &= \sum_{k=1}^N \mathbf{E} \left[ \left( M_k \langle \nabla F(\omega_{(k)}^N), h_k^N \rangle_{\mathfrak{H}^1} + \mathbf{E}' \left[ \frac{M_k - M'_k}{\sqrt{N}} \varepsilon(\omega_{(k)}^N, h_k^N) \right] \right)^2 \right] \\ &= \sum_{k=1}^N \mathbf{E} \left[ \langle \nabla F(\omega_{(k)}^N), h_k^N \rangle_{\mathfrak{H}^1}^2 \right] + \text{Rem}, \end{aligned}$$

and

$$\text{Rem} \leq \frac{c}{N} \sum_{k=1}^N \mathbf{E} \left[ \varepsilon(\omega_{(k)}^N, h_k^N)^2 \right] \xrightarrow{N \rightarrow \infty} 0,$$

by the Césaro theorem. It follows that  $\mathcal{E}^{\text{UN}}(F)$  has the same limit as

$$\sum_{k=1}^N \mathbf{E} \left[ \langle \nabla F(\omega_{(k)}^N), h_k^N \rangle_{\mathfrak{H}^1}^2 \right].$$

As  $N$  goes to infinity, we add more and more terms to the random walk, so that the influence of one particular term becomes negligible. The following result is well known (see Bouleau, Proposition 3 in [22]): for any  $k \in \{1, \dots, N\}$ , for any bounded  $\psi$  and  $\varphi$ ,

$$\mathbf{E} \left[ \psi(M_k) \varphi(\omega^N) \right] \xrightarrow{N \rightarrow \infty} \mathbf{E} \left[ \psi(M_k) \right] \mathbf{E} \left[ \varphi(\omega) \right].$$

Since  $\|\nabla F\|_{\mathfrak{H}^1}$  belongs to  $L^\infty$  and  $\|h_k^N\|_\infty$  tends to 0, this entails that for any  $k$ ,

$$\lim_{N \rightarrow \infty} \mathbf{E} \left[ \langle \nabla F(\omega_{(k)}^N), h_k^N \rangle_{\mathfrak{H}^1}^2 \right] = \lim_{N \rightarrow \infty} \mathbf{E} \left[ \langle \nabla F(\omega^N), h_k^N \rangle_{\mathfrak{H}^1}^2 \right] = \lim_{N \rightarrow \infty} \mathbf{E} \left[ \|\pi_{V_N} \nabla F(\omega^N)\|_{\mathfrak{H}^1}^2 \right],$$

where  $\pi_{V_N}$  is the orthogonal projection in  $\mathfrak{H}^1$  onto  $\text{span}\{h_k^N, k = 1, \dots, N\}$ . We conclude by dominated convergence.  $\square$

*Proof of Theorem 2.3.10.* Starting from (2.3.7), the terms with  $\tau = 0$  can be decomposed as

$$e^{-2p_m^N} \sum_{m=1}^N \mathbf{E} \left[ \left( F(\omega_{(m)}^N) + \varepsilon_{\zeta_m^N} - F(\omega_{(m)}^N) \right)^2 \right] \mu_m^N(1) + R_0^N.$$

Since  $F$  belongs to  $\text{TV} - \text{Lip}$ ,

$$R_0^N \leq \sum_{m=1}^N \sum_{\ell=2}^{\infty} \ell^2 \mu_m^N(\ell) \leq c_1 N (p^N)^2 \mathbf{E} \left[ (\text{Poisson}(p^N) + 2)^2 \right] \leq c_2 N (p^N)^2,$$

where the  $c_1$  and  $c_2$  are irrelevant constants. As  $Np^N$  is bounded,  $R_0^N$  goes to 0 as  $N$  grows to infinity. For the very same reasons, the sum of the terms of (2.3.7) with  $\tau \geq 1$  converge to 0, thus

$$\lim_{N \rightarrow \infty} \mathcal{E}^{\text{UN}}(F) = \lim_{N \rightarrow \infty} \sum_{m=1}^N e^{-2p_m^N} \mathbf{E} \left[ \left( F(\omega_{(m)}^N) + \varepsilon_{\zeta_m^N} - F(\omega_{(m)}^N) \right)^2 \right] p_m^N.$$

Consider now the space  $\mathfrak{N}_{\mathbf{Y}}^{\zeta} = \mathfrak{N}_{\mathbf{Y}} \times \{\zeta_k^N, k = 1, \dots, N\}$  with the product topology and probability measure  $\tilde{\mathbf{P}}_N = \mathbf{P}_N \otimes \sum_k p_k^N \varepsilon_{\zeta_k^N}$ . Let

$$\begin{aligned} \psi : \mathfrak{N}_{\mathbf{Y}} \times \{\zeta_k^N, k = 1, \dots, N\} &\longrightarrow \mathbf{E} \\ (\omega, \zeta) &\longmapsto \left( \mathbf{F}(\omega - (\omega(\zeta) - 1)\varepsilon_{\zeta}) - \mathbf{F}(\omega - \omega(\zeta)\varepsilon_{\zeta}) \right)^2. \end{aligned}$$

Then, we can write

$$\sum_{m=1}^N \mathbf{E} \left[ \left( \mathbf{F}(\omega_{(m)}^N + \varepsilon_{\zeta_m^N}) - \mathbf{F}(\omega_{(m)}^N) \right)^2 \right] p_m^N = \int_{\mathfrak{N}_{\mathbf{Y}}^{\zeta}} \psi(\omega, \zeta) d\tilde{\mathbf{P}}_N(\omega, \zeta).$$

Under  $\tilde{\mathbf{P}}_N$ , the random variables  $\omega$  and  $\zeta$  are independent. Equation (2.3.6) means that the marginal distribution of  $\zeta$  tends to  $\mathbf{M}$  (assumed to be a probability measure at the very beginning of this construction). Moreover, we already know that  $\mathbf{P}_N$  converges in distribution to  $\mathbf{P}$ . Hence,  $\tilde{\mathbf{P}}_N$  tends to  $\mathbf{P} \otimes \mathbf{M}$  as  $N$  goes to infinity. Since  $\mathbf{F}$  is in  $\text{TV} - \text{Lip}$ ,  $\psi$  is continuous and bounded, hence the result.  $\square$

## Chapter 3

# Applications : representation formulas, convergence theorems, finance

The last chapter is a patchwork which pieces are made of the applications inherited from our construction. The diversity of these comes from the variety of the fields explored. The first subsection deals with the declination of Clark formula for  $U$ -statistics, such that we show that the Hoeffding decomposition can be interpreted as an avatar of. The second one is motivated by the statement of a discrete analogue of the Stein-Malliavin criterion (1.4.1). In that perspective, we follow the footsteps of the path "from Stein's method to Stein-Dirichlet-Malliavin structures" described in the first chapter. In the last part we challenge our formalism with an application in finance : the computation of the greek Rho in the trinomial model.

### 3.1 Representations

We now show that our Clark formula yields an interesting decomposition of random variables. For  $U$ -statistics, it boils down to the Hoeffding decomposition.

**Definition 3.1.1.** For  $m, n \in \mathbf{N}^*$ , let  $h : \mathbf{R}^m \rightarrow \mathbf{R}$  be a symmetric function, and  $X_1, \dots, X_n$ ,  $n$  random variables supposed to be independent and identically distributed. The  $U$ -statistics of degree  $m$  and kernel  $h$  is defined, for any  $n \geq m$  by

$$U_n = U(X_1, \dots, X_n) = \binom{n}{m}^{-1} \sum_{A \in ([n], m)} h(X_A)$$

where  $([n], m)$  denotes the set of ordered subsets  $A \subset [n] = \{1, \dots, n\}$ , of cardinality  $m$ . More generally, for a set  $A \subset [n]$ ,  $(A, m)$  denotes the set of subsets of  $A$  with  $m$  elements.

If  $\mathbf{E}[|h(X_1, \dots, X_m)|]$  is finite, we define  $h_m = h$  and for  $1 \leq k \leq m - 1$ ,

$$h_k(X_1, \dots, X_k) = \mathbf{E}[h(X_1, \dots, X_m) | X_1, \dots, X_k].$$

Let  $\theta = \mathbf{E}[h(X_1, \dots, X_m)]$ , consider  $g_1(X_1) = h_1(X_1) - \theta$ , and

$$g_k(X_1, \dots, X_k) = h_k(X_1, \dots, X_k) - \theta - \sum_{j=1}^{k-1} \sum_{B \in ([k], j)} g_j(X_B),$$

for any  $1 \leq k \leq m$ . Since the variables  $X_1, \dots, X_n$  are independent and identically distributed, and the function  $h$  is symmetric, the equality

$$\mathbf{E}[h(X_{A \cup B}) | X_B] = \mathbf{E}[h(X_{C \cup B}) | X_B],$$

holds for any subsets  $A$  and  $C$  of  $[n] \setminus B$ , of cardinality  $n - k$ .

**Theorem 3.1.2** (★ Hoeffding decomposition of  $U$ -statistics, [76]). *For any integer  $n$ , we have*

$$U_n = \theta + \sum_{k=1}^m H_n^{(k)} \quad (3.1.1)$$

where  $H_n^{(k)}$  is the  $U$ -statistics based on kernel  $g_k$ , i.e. defined by

$$H_n^{(k)} = \binom{n}{k}^{-1} \sum_{B \subset [n], |B|=k} g_k(X_B).$$

As mentioned above, reversing the natural order of  $A$ , provided that it exists, can be very fruitful. We illustrate this idea by the decomposition of the number of fixed points of a random permutation under Ewens distribution. It could be applied to more complex functionals of permutations but to the price of increasingly complex computations. For every integer  $N$ , denote by  $\mathfrak{S}_N$  the space of permutations on  $\{1, \dots, N\}$ . We always identify  $\mathfrak{S}_N$  as the subgroup of  $\mathfrak{S}_{N+1}$  stabilizing the element  $N + 1$ . For every  $k \in \{1, \dots, N\}$ , define  $\mathcal{J}_k = \{1, \dots, k\}$  and

$$\mathcal{J} = \mathcal{J}_1 \times \mathcal{J}_2 \times \dots \times \mathcal{J}_N.$$

The coordinate map from  $\mathcal{J}$  to  $\mathcal{J}_k$  is denoted by  $I_k$ . Following the construction designed by Serguei Kerov [74], we have:

**Theorem 3.1.3** (★). *There exists a natural bijection  $\Gamma$  between  $\mathcal{J}$  and  $\mathfrak{S}_N$ .*

*Proof.* To a sequence  $(i_1, \dots, i_N)$  where  $i_k \in \mathcal{J}_k$ , we associate the permutation

$$\Gamma(i_1, \dots, i_N) = (N, i_N) \circ (N-1, i_{N-1}) \dots \circ (2, i_2).$$

where  $(i, j)$  denotes the transposition between the two elements  $i$  and  $j$ . To an element  $\sigma_N \in \mathfrak{S}_N$ , we associate  $i_N = \sigma_N(N)$ . Then,  $N$  is a fixed point of  $\sigma_{N-1} = (N, i_N) \circ \sigma_N$ , hence it can be identified as an element  $\sigma_{N-1}$  of  $\mathfrak{S}_{N-1}$ . Then,  $i_{N-1} = \sigma_{N-1}(N-1)$  and so on for decreasing indices. It is then clear that  $\Gamma$  is one-to-one and onto.  $\square$

In Kerov [74],  $\Gamma$  is described by the following rule: start with permutation  $\sigma_1 = (1)$ , if at the  $N$ -th step of the algorithm, we have  $i_N = N$  then the current permutation is extended by leaving  $N$  fixed, otherwise,  $N$  is inserted in  $\sigma_{N-1}$  just before  $i_N$  in the cycle of this element. This construction is reminiscent of the Chinese restaurant process (see Arratia, Barbour and Tavaré [9]) where  $i_N$  is placed immediately after  $N$ . An alternative construction of permutations is known as the Feller coupling (see Arratia, Barbour and Tavaré [9]). In our notations, it is given by

$$\sigma_1 = (1); \quad \sigma_N = \sigma_{N-1} \circ (\sigma_{N-1}^{-1}(i_N), N).$$

**Definition 3.1.4** (Ewens distribution). For some  $t \in \mathbf{R}^+$ , for any  $k \in \{1, \dots, N\}$ , consider the measure  $\mathbf{P}_k$  defined on  $\mathcal{J}_k$  by

$$\mathbf{P}_k(\{j\}) = \begin{cases} \frac{1}{t+k-1} & \text{if } j \neq k, \\ \frac{t}{t+k-1} & \text{for } j = k. \end{cases}$$

Under the distribution  $\mathbf{P} = \otimes_k \mathbf{P}_k$ , the random variables  $(I_k, k = 1, \dots, N)$  are independent with law given by  $\mathbf{P}(I_k = j) = \mathbf{P}_k(\{j\})$ , for any  $k$ . The Ewens distribution of parameter  $t$  on  $\mathfrak{S}_N$ , denoted by  $\mathbf{P}^t$ , is the push-forward of  $\mathbf{P}$  by the map  $\Gamma$ .

A moment of thought shows that a new cycle begins in the first construction for each index where  $i_k = k$ . Moreover, it can be shown that:

**Theorem 3.1.5** (see Kerov [74]). For any  $\sigma \in \mathfrak{S}_N$ ,

$$\mathbf{P}^t(\{\sigma\}) = \frac{t^{cyc(\sigma)}}{(t+1)(t+2) \times \dots \times (t+N-1)},$$

where  $cyc(\sigma)$  is the number of cycles of  $\sigma$ .

For any  $F$ , a measurable function on  $\mathfrak{S}_N$ , we have the following diagram

$$\begin{array}{ccc} (\mathcal{J}, \otimes_{k=1}^N \mathbf{P}_k) & & \\ \Gamma \downarrow & \searrow \tilde{F} = F \circ \Gamma & \\ (\mathfrak{S}_N, \mathbf{P}^t) & \xrightarrow{F} & \mathbf{R} \end{array}$$

We denote by  $i = (i_1, \dots, i_N)$  a generic element of  $\mathcal{J}$  and by  $\sigma = \Gamma(i)$ . Let  $C_1(\sigma)$  denote the number of fixed points of the permutation  $\sigma$  and  $\tilde{C}_1 = C_1 \circ \Gamma$ . For any  $k \in \mathcal{J}_N$ , the random variable  $U_k(\sigma)$  is the indicator of the event ( $k$  is a fixed point of  $\sigma$ ) and let  $\tilde{U}_k^N = U_k \circ \Gamma$ . The Clark formula with reverse filtration shows that we can write  $\tilde{U}_k^N$  as a sum of centered orthogonal random variables as in the Hoeffding decomposition of  $U$ -statistics (see Theorem 3.1.2).

**Theorem 3.1.6** ( $\star$ ). For any  $k \in \{1, \dots, N\}$ ,

$$\tilde{U}_k^N = \mathbf{1}_{(I_k=k)} \mathbf{1}_{(I_m \neq k, m \in \{k+1, \dots, N\})}. \quad (3.1.2)$$

and under  $\mathbf{P}^t$ ,  $\tilde{U}_k^N$  is Bernoulli distributed with parameter  $tp_k \alpha_k$ , where for any  $k \in \{1, \dots, N\}$ ,

$$p_k = \frac{1}{t+k-1} \text{ and } \alpha_k = \prod_{j=k+1}^N \frac{j-1}{t+j-1}.$$

Moreover,

$$\begin{aligned} \tilde{U}_k^N &= tp_k \alpha_k + \left( \mathbf{1}_{(I_k=k)} - tp_k \right) \prod_{m=k+1}^N \mathbf{1}_{(I_m \neq k)} \\ &\quad - tp_k \sum_{j=1}^{N-k-1} \frac{t+k-1}{t+k+j-2} \left( \mathbf{1}_{(I_{k+j}=k)} - p_{k+j} \right) \prod_{\ell=j+1}^{N-k} \mathbf{1}_{(I_{k+\ell} \neq k)}. \end{aligned}$$



Since

$$\tilde{\mathcal{C}}_1 = \sum_{k=1}^N \tilde{U}_k^N,$$

we retrieve the result of [10]:

$$\mathbf{E}[\tilde{\mathcal{C}}_1] = \frac{tN}{t+N-1},$$

and the following decomposition of  $\tilde{\mathcal{C}}_1$  can be easily deduced from the previous theorem.

**Theorem 3.1.7** (★). *We can write*

$$\begin{aligned} \tilde{\mathcal{C}}_1 &= t \left( 1 - \frac{t-1}{N+t-1} \right) + \sum_{\ell=1}^N D_\ell \tilde{U}_\ell^N + \sum_{\ell=2}^N \frac{t}{t+\ell-2} D_\ell \left( \sum_{k=1}^{\ell-1} \prod_{m=\ell}^N \mathbf{1}_{(I_m \neq k)} \right) \\ &= t \left( 1 - \frac{t-1}{N+t-1} \right) + \sum_{\ell=1}^N \left( \mathbf{1}_{(I_\ell = \ell)} - \frac{t}{t+\ell-1} \right) \prod_{m=\ell+1}^N \mathbf{1}_{(I_m \neq \ell)} \\ &\quad - \sum_{\ell=2}^{N-1} \frac{t}{t+\ell-2} \sum_{k=1}^{\ell-1} \left( \mathbf{1}_{(I_\ell = k)} - \frac{1}{t+\ell-1} \right) \prod_{m=\ell+1}^N \mathbf{1}_{(I_m \neq k)}. \end{aligned}$$

**Remark 3.1.8.** Note that such a decomposition with the natural order on  $\mathbf{N}$  would be infeasible since the basic blocks of the definition of  $\tilde{\mathcal{C}}_1$ , namely the  $\tilde{U}_k$ , are anticipative (following the vocabulary of Gaussian Malliavin calculus), i.e.  $\tilde{U}_k \in \sigma(I_{k+\ell}, \ell = 0, \dots, N-k)$ .

This decomposition can be used to compute the variance of  $\tilde{\mathcal{C}}_1$ . To the best of our knowledge, this is the first explicit, i.e. not asymptotic, expression of it.

**Theorem 3.1.9.** *For any  $t \in \mathbf{R}$ , we get*

$$\text{Var}[\tilde{\mathcal{C}}_1] = \frac{Nt}{t+N-1} \left( \frac{t}{t+N-1} + 1 - \frac{2t^2}{N} \sum_{k=1}^N \frac{1}{t+k-1} \right).$$

We retrieve

$$\text{Var}[\tilde{\mathcal{C}}_1] \xrightarrow{N \rightarrow \infty} t,$$

as can be expected from the Poisson limit.

### 3.2 Stein-Malliavin criterion for Normal and Gamma approximations

In this section we confront our model to the Stein-Dirichlet-Malliavin method in order to provide in particular a discrete analogue of the Stein-Malliavin criterions for Normal and Gamma approximations. Such approximation results derived from this method for functionals of Rademacher (see for instance Nourdin, Peccati and Reinert [96], Zheng [143]), Poisson (see for instance Decreusefond, Schulte and Thäle [41] Lachièze-Rey and Peccati [79], Peccati et al. [111]) or Gaussian random variables (see Nualart and Peccati [100]) or processes (see Coutin and Decreusefond [32, 33]). Our results take place among them; however, to the best of our knowledge, when  $\mathbf{Q}$  is the distribution of a family of independent random variables,

the distance  $\text{dist}(F^*\tilde{\mathbf{Q}}, \mathbf{P})$  (where  $\text{dist} = \mathfrak{d}_1, \mathfrak{d}_2$  or  $\text{dist}_{\text{Kol}}$  as the case may be) is evaluated through exchangeable pairs or couplings, which means to construct an ad-hoc structure for each situation at hand. We intend to give here an exact analogue of (1.4.1) in this situation using only our newly defined operator  $D$ , that makes our approach original.

**Remark 3.2.1.** In what follows, we deal with functions  $F$  defined on  $E_A$ , that means that  $F$  is a function of  $X_A$  and as such, we should use the notation  $F(X_A)$ . For the sake of notations, we identify  $F$  and  $F(X_A)$ . The distribution of  $F$  is denoted  $\mathbf{P}_F = F^*\mathbf{P}_A$ .

**Theorem 3.2.2 (★).** *Let  $\mathbf{P}$  denote the standard Gaussian distribution on  $\mathbf{R}$ . For any  $F : E_A \rightarrow \mathbf{R}$  such that  $\mathbf{E}[F] = 0$  and  $F \in \text{Dom } D$ . Then,*

$$\mathfrak{d}_1(\mathbf{P}, \mathbf{P}_F) \leq \mathbf{E} \left[ \left| 1 - \sum_{a \in A} D_a F (-D_a L^{-1}) F \right| \right] + \sum_{a \in A} \mathbf{E} \left[ \int_{E_A} \left( F - F(X_{A-a}; x) \right)^2 d\mathbf{P}_a(x) |D_a L^{-1} F| \right].$$

The proof of this version follows exactly the lines of the proof of Theorem 3.1 in Nourdin and Peccati [93], Peccati, Solé, Taqqu and Utzet [111] but we can do slightly better by changing a detail in the Taylor expansion.

**Theorem 3.2.3 (★).** *Let  $\mathbf{P}$  denote the standard Gaussian distribution on  $\mathbf{R}$ . For any  $F : E_A \rightarrow \mathbf{R}$  such that  $\mathbf{E}[F] = 0$  and  $F \in \text{Dom } D$ . Then,*

$$\mathfrak{d}_1(\mathbf{P}, \mathbf{P}_F) \leq \sup_{\psi \in \text{Lip}_2} \mathbf{E} \left[ \psi(F) - \sum_{a \in A} \psi(F(X'_{-a})) D_a F (-D_a L^{-1}) F \right] + \sum_{a \in A} \mathbf{E} \left[ \int_{E_A} \left( F - F(X_{A-a}; x) \right)^2 d\mathbf{P}_a(x) |D_a L^{-1} F| \right], \quad (3.2.1)$$

where  $X'_{-a} = X_{A-a} \cup \{X'_a\}$ .

**Remark 3.2.4.** This theorem provides a Berry-Essen type bound in the Wasserstein distance for the Normal approximation by any functional of random variables and generalises in the certain sense that got by Sourav Chatterjee ([24]). Raphaël Lachièze-Rey and Giovanni Peccati stated a remarkable bound in the Kolmogorov distance for the same approximation in terms of another difference operator  $\Delta$ ; this is defined for any ordered subset  $C$  of  $\llbracket 1, N \rrbracket$  by

$$\Delta_C F(X_A, X'_A) = F(X_A) - F^C(X_A, X'_A),$$

and constructed by iteration of the operator

$$\Delta_{\{a\}} F(X_A, X'_A) = F(X_A) - F(X'_1, \dots, X'_{a-1}, X_a, X'_{a+1}, \dots, X'_N).$$

**Remark 3.2.5.** We can point out however that, for lack of chaos decomposition, we could not rule the equivalent of the fourth order moment theorem (see [100], [95]), appearing in the normal approximations of number of nonlinear functionalities acting on random fields, such as Poisson ones (see [47]) or relative to fractional Brownian motion (see [97]) to mention but two examples.

This formulation may seem cumbersome, but it easily gives a close to the usual bound in the Lyapounov central limit theorem, with a non optimal constant (see [57]).

**Corollary 3.2.6** ( $\star$  Lyapounov). *Let  $(X_n)_{n \in \mathbf{N}^*}$  be a sequence of thrice integrable, independent random variables. Denote*

$$\sigma_n^2 = \text{Var}(X_n), s_n^2 = \sum_{j=1}^n \sigma_j^2 \text{ and } Y_n = \frac{1}{s_n} \sum_{j=1}^n (X_j - \mathbf{E}[X_j]).$$

Then,

$$\mathfrak{d}_1(\mathbf{P}, \mathbf{P}_{Y_n}) \leq \frac{2(\sqrt{2} + 1)}{s_n^3} \sum_{j=1}^n \mathbf{E}[|X_j - \mathbf{E}[X_j]|^3].$$

**Remark 3.2.7.** If we use Theorem 3.2.2, we get

$$\mathfrak{d}_1(\mathbf{P}, \mathbf{P}_{Y_n}) \leq \mathbf{E} \left[ \left| 1 - \sum_{j=1}^n \frac{X_j^2}{s_n^2} \right| \right] + \frac{2}{s_n^3} \sum_{j=1}^n \mathbf{E}[|X_j - \mathbf{E}[X_j]|^3],$$

and the quadratic term is easily bounded only if the  $X_i$ 's are such that  $\mathbf{E}[X_i^4]$  is finite, which in view of Corollary 3.2.6 is a too stringent condition.

The functional which appears in the central limit theorem is the basic example of  $U$ -statistics or homogeneous sums. If we want to go further and address the problem of convergence of more general  $U$ -statistics (or homogeneous sums), we need to develop a similar apparatus for the Gamma distribution. Recall that the Gamma distribution of parameters  $r$  and  $\lambda$  has density

$$f_{r,\ell}(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} \mathbf{1}_{\mathbf{R}^+}(x).$$

Let  $Y_{r,\ell} \sim \Gamma(r, \ell)$ , it has mean  $r/\ell$  and variance  $r/\ell^2$ . Denote by  $\bar{Y}_{r,\ell} = Y_{r,\ell} - r/\ell$ . As described in Graczyk et al. [60],  $Z \sim \bar{Y}_{r,\ell} = Y_{r,\ell} - r/\ell$  if and only if  $\mathbf{E}[L_{r,\ell}f(Z)] = 0$  for any  $f$  once differentiable, where

$$L_{r,\ell}f(y) = \frac{1}{\lambda} \left( y + \frac{r}{\ell} \right) f'(y) - yf(y).$$

The Stein equation

$$L_{r,\ell}f(y) = g(y) - \mathbf{E}[g(\bar{Y}_{r,\ell})] \tag{3.2.2}$$

has a solution  $f_g$  which satisfies

$$\|f_g\|_\infty \leq \|g'\|_\infty, \|f'_g\|_\infty \leq 2\lambda \max\left(1, \frac{1}{r}\right) \|g'\|_\infty$$

$$\text{and } \|f''_g\|_\infty \leq 2\lambda \left( \max\left(\lambda, \frac{\lambda}{r}\right) \|g'\|_\infty + \|g''\|_\infty \right), \tag{3.2.3}$$

noting that  $f_g$  is solution of (3.2.2) if and only if  $h_g : x \mapsto \frac{1}{\lambda} f\left(x - \frac{r}{\lambda}\right)$  solves

$$xh'(x) + (r - \lambda x)h(x) = g(x) - \mathbf{E}[g(Y_{r,\lambda})],$$

studied in [8, 44].

**Theorem 3.2.8 (★).** *Let  $\mathcal{F}$  is the set of twice differentiable functions with first and second derivative bounded by 1. There exists  $c > 0$  such that for any  $F \in \text{Dom } D$  with  $\mathbf{E}[F] = 0$ ,*

$$\begin{aligned} \mathfrak{d}_1(\mathbf{P}_F, \mathbf{P}_{\bar{Y}_{r,\lambda}}) \leq & c \mathbf{E} \left[ \left| \frac{1}{\lambda} F + \frac{r}{\lambda^2} - \sum_{a \in A} D_a F (-D_a L^{-1}) F \right| \right] \\ & + c \sum_{a \in A} \mathbf{E} \left[ \int_{E_A} \left( F(X_A) - F(X_{A \setminus a}; x) \right)^2 d\mathbf{P}_a(x) |D_a L^{-1} F| \right]. \end{aligned} \quad (3.2.4)$$

This theorem reads exactly as Theorem 1.5 in [44] for Poisson functionals and is proved in a similar fashion.

**Remark 3.2.9.** The generalization of this result to multivariate Gamma distribution could be considered in a forthcoming paper. The difficulty lies in the regularity estimates of the solution of the Stein equation associated to multivariate Gamma distribution, which require lengthy calculations.

An homogeneous sum of order  $d$  is a functional of independent identically distributed random variables  $(X_1, \dots, X_{N_n})$ , of the form

$$F_n(X_1, \dots, X_{N_n}) = \sum_{1 \leq i_1, \dots, i_d \leq N_n} f_n(i_1, \dots, i_d) X_{i_1} \dots X_{i_d}$$

where  $(N_n, n \geq 1)$  is a sequence of integers which tends to infinity as  $n$  does and the functions  $f_n$  are symmetric on  $\{1, \dots, N_n\}^d$  and vanish on the diagonal. The asymptotics of these sums have been widely investigated and depend on the properties of the function  $f_n$ . For  $d = 2$ , see for instance Götze and Tikhomirov [59]. In Nourdin, Peccati and Reinert [96], the case of any value of  $d$  is investigated through the prism of universality: roughly speaking (see Theorem 4.1), if  $F_n(G_1, \dots, G_{N_n})$  converges in distribution when  $G_1, \dots, G_{N_n}$  are standard Gaussian random variables then  $F_n(X_1, \dots, X_{N_n})$  converges to the same limit whenever the  $X_i$ 's are centered with unit variance and finite third order moment and such that

$$\max_i \sum_{1 \leq i_2, \dots, i_d \leq N_n} f_n^2(i, i_2, \dots, i_d) \xrightarrow{n \rightarrow \infty} 0.$$

For Gaussian random variables, the functional  $F_n$  belongs to the  $d$ -th Wiener chaos. Combining the algebraic rules of multiplication of iterated Gaussian integrals and the Stein-Malliavin method, it is proved in Nourdin and Peccati [92] that  $F_n(G_1, \dots, G_{N_n})$  converges in distribution to a chi-square distribution of parameter  $\nu$  if and only if

$$\mathbf{E}[F_n^2] \xrightarrow{n \rightarrow \infty} 2\nu \text{ and } \mathbf{E}[F_n^4] - 12\mathbf{E}[F_n^3] - 12\nu^2 + 48\nu \xrightarrow{n \rightarrow \infty} 0.$$

We obtain here a related result for  $d = 2$  (for the sake of simplicity though the method is applicable for any value of  $d$ ) and a general distribution without resorting to universality.

Let  $A = \{1, \dots, n\}$ . For  $f, g : A^2 \rightarrow \mathbf{R}$ , symmetric functions vanishing on the diagonal, define the two contractions by

$$\begin{aligned} (f \star_1^1 g)(i, j) &= \sum_{k \in A} f(i, k) g(j, k), \\ (f \star_2^1 g)(i) &= \sum_{j \in A} f(i, j) g(i, j). \end{aligned} \quad (3.2.5)$$

**Theorem 3.2.10 (★).** *Let  $X_A = \{X_i, 1 \leq i \leq n\}$  be a collection of centered independent random variables with unit variance and finite moment of order 4. Define*

$$F(X_A) = \sum_{(i,j) \in A^\#} f(i,j) X_i X_j$$

where  $(i,j) \in A^\#$  means that we enumerate all the couples  $(i,j)$  in  $A^2$  with distinct components and  $f$  is a symmetric function which vanishes on the diagonal. Let  $\nu = \sum_{(i,j)} f^2(i,j)$ . Then, there exists  $c_\nu > 0$  such that

$$\begin{aligned} \mathfrak{d}_2^2(\mathbf{P}_F, \mathbf{P}_{\bar{Y}_{\nu/2,1/2}}) &\leq c_\nu \mathbf{E}[X_1^4]^2 \\ &\times \left[ \sum_{(i,a) \in A^2} f^4(i,a) + \|f \star_2^1 f\|_{L^2(A)}^2 + \|f - f \star_1^1 f\|_{L^2(A^2)}^2 \right]. \end{aligned} \quad (3.2.6)$$

We now introduce  $\text{Inf}_a(f)$ , called the influence of the variable  $a$ , by

$$\text{Inf}_a(f) = \sum_{i \in A} f^2(i,a).$$

Remark that

$$\begin{aligned} \sum_{i \in A} f^4(i,a) &\leq \sum_{a \in A} \sum_i f^2(i,a) \sum_j f^2(j,a) \\ &= \sum_{a \in A} \sum_i f^2(i,a) \text{Inf}_a(f) \\ &\leq \nu \max_{a \in A} \text{Inf}_a(f). \end{aligned}$$

The same kind of computations can be made for  $\|f \star_2^1 f\|_{L^2(A)}^2$ . As a consequence, we get the following corollary.

**Corollary 3.2.11 (★).** *With the same notations as above,*

$$\mathfrak{d}_2^2(\mathbf{P}_F, \mathbf{P}_{\bar{Y}_{\nu/2,1/2}}) \leq c_\nu \mathbf{E}[X_1^4]^2 \left[ \max_{a \in A} \text{Inf}_a(f) + \|f - f \star_1^1 f\|_{L^2(A^2)}^2 \right].$$

The supremum of the influence is the quantity which governs the distance between the distributions of  $F_n(G_1, \dots, G_{N_n})$  and  $F_n(X_1, \dots, X_{N_n})$  in Nourdin, Peccati and Reinert [96], thus it is not surprising that it still appears here.

**Remark 3.2.12.** It may be interesting to compare the result of Theorem 3.2.10 with Theorem 4.3 (by taking a dimension  $d = 2$ ) of Nguyen Tien Dung [46], which led to a publication slightly subsequent to ours. Similar assumptions are done on the "contractions" (without being explicitly mentioned in [46]) of the components of matrix  $(f(i,j))_{i,j}$  but more regularity on test functions that must be three differentiable is needed in [46] whereas we request twice differentiability in our statement. This multivariate CLT for quadratic forms is in fact studied as an application of a general work on multivariate CLT for nonlinear statistics, which Nguyen Tien Dung gives explicit bounds in Theorems 3.1 and 3.2. These are obtained using Stein's method and by means of a difference operator which definition coincides with our discrete

gradient (see Definition 2.1.2), reminiscent also of the works of G. Lugosi, P. Massart, C. Houdré, N. Privault, or S. Chatterjee. Without constructing a discrete Malliavin calculus as in our formalism, Nguyen Tien Dung stated a covariance formula and an approximate chain rule on which are based his results. The use of the Slepian's interpolation method cleverly allows to bypass the computational difficulties involved in upper dimension, and its combination with Stein's method is advantageous.

**Theorem 3.2.13.** *The moments of  $F$  (up to the fourth one) are linked by the relation:*

$$\begin{aligned} \mathbf{E}[F^4] - 12\mathbf{E}[F^3] - 12\nu^2 + 48\nu &= \sum_{(i,j) \in A^\neq} f^4(i,j) \mathbf{E}[X^4]^2 + 6 \sum_{(i,j,k) \in A^\neq} f^2(i,j) f^2(i,k) \mathbf{E}[X^4] \\ &+ 12 \mathbf{E}[X^3]^2 \left\{ \sum_{(i,j,k) \in A^\neq} f^2(i,j) f(i,k) f(k,j) - \sum_{(i,j) \in A^\neq} f^3(i,j) \right\} \\ &- 48 \left\{ \sum_{(i,j,k) \in A^\neq} f(i,j) f(i,k) f(k,j) - f^2(i,j) \right\} - 12 \sum_{(i,j) \in A^\neq} f^4(i,j). \end{aligned} \quad (3.2.7)$$

**Remark 3.2.14.** The Cauchy-Schwarz inequality entails that the properties

$$\mathbf{E}[F_n^4] - 12\mathbf{E}[F_n^3] - 12\nu^2 + 48\nu \xrightarrow{n \rightarrow \infty} 0$$

and

$$\mathfrak{d}_2(\mathbf{P}_F, \mathbf{P}_{\bar{Y}_{\nu/2, 1/2}}) \xrightarrow{n \rightarrow \infty} 0$$

share the same sufficient condition:

$$\sum_{(i,a) \in A^\neq} f^4(i,a) + \|f \star_2^1 f\|_{L^2(A)}^2 + \|f - f \star_1^1 f\|_{L^2(A^2)}^2 \xrightarrow{n \rightarrow \infty} 0.$$

However, we cannot go further and state a *fourth moment theorem* as we know, that for Benoulli random variables,  $F_n$  may converge to  $\bar{Y}_{\nu/2, 1/2}$  while the RHS of (3.2.6) does not converge to 0.

As another corollary of Theorem 3.2.10, we obtain the KR distance between a degenerate  $U$ -statistics of order 2 and a Gamma distribution. Compared to the more general result stated by Christian Döbler and Giovanni Peccati (Theorem 1.1 in [44]), the computations are here greatly simplified by the absence of exchangeable pairs.

**Theorem 3.2.15 (★).** *Let  $A = \{1, \dots, n\}$  and  $(X_i, i \in A)$  a family of independent and identically distributed real-valued random variables such that*

$$\mathbf{E}[X_1] = 0, \quad \mathbf{E}[X_1^2] = \sigma^2 \quad \text{and} \quad \mathbf{E}[X_1^4] < \infty.$$

*Consider the random variable*

$$F = \frac{2}{n-1} \sum_{(i,j) \in A^\neq} X_i X_j.$$

*Then, there exists  $c > 0$ , independent of  $n$ , such that*

$$\mathfrak{d}_2(\mathbf{P}_F, \mathbf{P}_{\bar{Y}_{1/2, 1/2\sigma^2}}) \leq c \frac{\sigma^2}{\sqrt{n}} \mathbf{E}[X_1^4]. \quad (3.2.8)$$

*Proof.* Take  $f_n(i, j) = 2/(n - 1)$  and apply Theorem 3.2.10.  $\square$

**Remark 3.2.16.** The proof of Theorem 3.2.10 is rich of insights. In Gaussian, Poisson or Rademacher contexts, the computation of  $L^{-1}F$  is easily done when there exists a chaos decomposition since  $L$  operates as a dilation on each chaos (see [93, 94, 111]). In Lemma 3.4 of [123], Matthias Schulte and Matthias Reitzner provide a formula for the operator  $L^{-1}$  of Poisson driven  $U$ -statistics, not resorting to the chaos decomposition. It is based on the fact that  $L$  applied to a  $U$ -statistics  $F$  of order  $k$  yields  $kF$  plus a  $U$ -statistics of order  $(k - 1)$ . Then, the construction of an inverse formula can be made by induction. In our framework, the action of  $L$  on a  $U$ -statistics yields  $kF$  plus a  $U$ -statistics of order  $k$  so that no induction seems possible. However, for an order  $k$   $U$ -statistics which is degenerate of order  $(k - 1)$ , we have  $LF = kF$ . For  $k = 2$ , this hypothesis of degeneracy is exactly the sufficient condition to have a convergence towards a Gamma distribution.

### 3.3 Finance and Malliavin calculus

The first one is the computation of one of the *Greeks*, so named with reference to the Greek letters by which they are denoted. In the framework of a financial market, these quantities simulate the sensitivity of the price of derivatives - such as options - to a change in underlying parameters of the model. Their computation is performed by the use of the Malliavin integration by parts formula; in the framework of Black-Scholes model, see for instance Nualart [98], Montero and Kohatsu-Higa [75]. The idea is to rewrite Rho (the one that interests us here)

$$\rho = \frac{\partial \Phi(S_T)}{\partial r} = \mathbf{E}_{\mathbf{Q}} \left[ \frac{\partial f(S_T)}{\partial r} \right] \quad (3.3.1)$$

where  $\mathbf{Q}$  denotes the risk-neutral probability measure and  $f : \mathbf{R}_+ \rightarrow \mathbf{R}$  is a function, as  $\mathbf{E}_{\mathbf{Q}} [f(S_T) \pi]$ , where the *weight*  $\pi$  does not depend of the payoff. The transformation step can be advantageously made simple using integration by parts in the sense of Malliavin calculus as shown first by Éric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, Pierre-Louis Lions and Nizar Touzi in [54] and [55].

The application of our formalism to the computation of Rho takes shape in the trinomial model described as follows ; given  $T \in \mathbf{N}^*$ , denote  $\mathbf{N}_T = \mathbf{N} \cap [0, T]$  and consider the state space  $\mathbf{E}^T = \prod_{t=1}^{\mathbf{N}_T} \mathbf{E}_t$ , where the  $\mathbf{E}_t$  are the 3-elements sets  $\mathbf{E}_t = \{-1, 0, 1\}$  ( $d = 1$ ), equipped with the product  $\bigotimes_{t \in \mathbf{N}_T} \mu_t$  where

$$\mu_t(\{1\}) = p, \quad \mu_t(\{-1\}) = q \quad \text{and} \quad \mu_t(\{0\}) = 1 - p - q.$$

The distribution of the canonical process  $(X_t)_{t \in \mathbf{N}_T}$  is given by

$$\mathbf{P}(X_t = 1) = p, \quad \mathbf{P}(X_t = -1) = q \quad \text{and} \quad \mathbf{P}(X_t = 0) = 1 - p - q.$$

The trinomial model, is, within this frame, a simplified discrete market model consisting of two assets:

- a riskless asset  $(A_t)_{t \in \mathbf{N}_T}$ ,
- a risky asset  $(S_t)_{t \in \mathbf{N}_T}$ ,

where  $(A_t)_{t \in \mathbf{N}_T}$  modelises the riskless asset with deterministic initial value  $A_0 = a_0$  and is defined by

$$A_n = a_0(1+r)^n, \quad (3.3.2)$$

whereas the stock price  $(S_t)_{t \in \mathbf{N}_T}$  with (deterministic) initial value  $S_0 = 1$  is given by

$$S_t = \begin{cases} (1+b)S_{t-1} & \text{if } X_t = 1 \\ S_{t-1} & \text{if } X_t = 0 \\ (1+a)S_{t-1} & \text{if } X_t = -1 \end{cases} \quad (\mathcal{P}^{\text{Tri}})$$

with  $a$  and  $b$  real such that  $-1 < a < r < b$ .

Define the sequence  $(\bar{S}_t)_{t \in \mathbf{N}_T}$  of discounted stock prices by

$$\bar{S}_t = \frac{1}{(1+r)^n} S_t, \quad (0 < r < b).$$

Assume in this section that  $\mathbf{P}$  verifies the risk-neutral condition i.e.

$$bp + aq = r, \quad (3.3.3)$$

which ensures that the sequence  $(\bar{S}_t)_{t \in \mathbf{N}_T}$  is a  $(\mathbf{P}, \mathbb{F})$ -martingale.

The aim of this subsection is to compute, in the frame of trinomial model, the greek called  $\rho$ , i.e. the sensitivity of the price with respect to the parameter  $r$  that is

$$\frac{d}{dr} \mathbf{E}[f(S_N(r))] = \lim_{h \rightarrow 0} \frac{1}{h} (\mathbf{E}[f(S_N(r+h))] - \mathbf{E}[f(S_N(r))]),$$

by considering the price as a function of  $r$ . The other ones, with respect to which it could have been interesting to study the variations, such as the initial value  $s_0$  (if  $S_0 = s_0$ ) or the volatility  $\sigma$  are fixed here.

To this end, consider  $\mathbf{P}$  the reference probability under which corresponds to the undisturbed evolution of the price and  $\mathbf{P}_h$  the probability corresponding to a  $h$ -perturbation. Denote by  $\mathbf{E}$  and  $\mathbf{E}_h$  the expectations respectively taken with respect to  $\mathbf{P}$  and  $\mathbf{P}_h$ .

This leads thus to compute, for a given functional  $F$ ,

$$\lim_{h \rightarrow 0} \frac{1}{h} (\mathbf{E}_h[F] - \mathbf{E}[F]). \quad (3.3.4)$$

Considering a small enough variation  $h$ , the existence of a random variable  $L_h$ , nonnegative  $\mathbf{P}$ -p.s. such that

$$d\mathbf{P}_h = L_h d\mathbf{P},$$

and so that (3.3.4) is equal to

$$\mathbf{E} \left[ F \frac{dL_h}{dh} \Big|_{h=0} \right].$$

We can thus establish:



**Theorem 3.3.1 (★).** *Let the process  $(S_t)_{0 \leq t \leq N_T}$  whose evolution is given by  $(\mathcal{P}^{\text{Tri}})$ . For any bounded functional  $f \in \mathcal{F}_N$ ,*

$$\frac{d}{dr} \mathbf{E}[f(S_N(r))] = \mathbf{E} \left[ f(S_N(r)) \frac{dL_h}{dh} \Big|_{h=0} \right],$$

where the expression of  $\frac{dL_h}{dh} \Big|_{h=0}$  is provided by

$$\frac{dL_h}{dh} \Big|_{h=0} = \sum_{t=1}^{N_T} \left\{ \left( \frac{k(r) + g(r)}{p} \mathbf{1}_{(U_t=1)} + \frac{k(r) - g(r)}{q} \mathbf{1}_{(U_t=-1)} - \frac{g(r)}{s} \mathbf{1}_{(U_k=0)} \right) \prod_{\ell=1}^{t-1} L_\ell^h(U_\ell) \right\}. \quad (3.3.5)$$

where  $k$  and  $g$  are the functions defined on  $\mathbf{R}_+$  by

$$k(r) = \frac{\sqrt{\Delta t} \lambda(\theta) (\lambda(\theta) - \theta \lambda'(\theta))}{2\lambda^3(\theta)\sigma} \quad \text{and} \quad g(r) = \frac{\Delta t \theta \lambda(\theta) - \lambda'(\theta) (1 + \Delta t \theta^2)}{\lambda^3(\theta)\sigma}.$$

## 3.4 Proofs of chapter 3

### 3.4.1 Proofs of Section 3.1

*Proof of Theorem 3.1.2.* Take care that in the argument of  $h$ , all the sets are considered as ordered : when we write  $B \cup C$ , we implicitly reorder its elements, for instance

$$h(X_{\{1,3\} \cup \{2\}}) = h(X_1, X_2, X_3).$$

Apply the Clark formula,

$$\begin{aligned} U_n - \theta &= \binom{n}{m}^{-1} \sum_{A \in ([n], m)} \sum_{B \subset A} \binom{m}{|B|}^{-1} \frac{1}{|B|} \sum_{b \in B} D_b \mathbf{E}[h(X_A) | X_B] \\ &= \binom{n}{m}^{-1} \sum_{B \subset [n]} \binom{m}{|B|}^{-1} \frac{1}{|B|} \sum_{b \in B} \sum_{\substack{A \supset B \\ A \in ([n], m)}} D_b \mathbf{E}[h(X_A) | X_B] \\ &= \binom{n}{m}^{-1} \sum_{B \subset [n]} \binom{m}{|B|}^{-1} \frac{1}{|B|} \sum_{b \in B} \sum_{C \in ([n] \setminus B, m - |B|)} D_b \mathbf{E}[h(X_{B \cup C}) | X_B]. \end{aligned}$$

It remains to prove that

$$\begin{aligned} \sum_{k=1}^m \binom{m}{k} H_n^{(k)} &= \binom{n}{m}^{-1} \sum_{B \subset [n], |B| \leq m} \binom{m}{|B|}^{-1} \frac{1}{|B|} \sum_{b \in B} \sum_{C \in ([n] \setminus B, m - |B|)} D_b \mathbf{E}[h(X_{B \cup C}) | X_B]. \quad (3.4.1) \end{aligned}$$

for any integer  $n$ . For  $n = 1$ , it is straightforward that

$$g_1(X_1) = h(X_1) - \theta = D_1 \mathbf{E}[h(X_1) | X_1].$$

Assume the existence of an integer  $n$  such that (3.4.1) holds for any set of cardinality  $n$ . In particular, for any  $l \in [n+1]$

$$\sum_{k=1}^m \binom{m}{k} H_{A_\ell}^{(k)} = \binom{n}{m}^{-1} \sum_{B \subset [A_\ell], |B| \leq m} \binom{m}{|B|}^{-1} \frac{1}{|B|} \sum_{b \in B} \sum_{C \in ([A_\ell] \setminus B, m-|B|)} D_b \mathbf{E} [h(X_{B \cup C}) | X_B],$$

where  $A_\ell = [n+1] \setminus \{\ell\}$ . Let  $m$  such that  $m \leq n$ . Then,

$$\begin{aligned} & \sum_{k=1}^m \binom{m}{k} H_{n+1}^{(k)} \\ &= \sum_{k=1}^m \binom{m}{k} \binom{n+1}{k}^{-1} \frac{1}{n+1-k} \sum_{l=1}^{n+1} \sum_{B \in ([A_\ell], k)} g_k(X_B) \\ &= \frac{1}{n+1} \sum_{l=1}^{n+1} \sum_{k=1}^m \binom{m}{k} \binom{n}{k}^{-1} \sum_{B \in ([A_\ell], k)} g_k(X_B) \\ &= \frac{1}{n+1} \sum_{l=1}^{n+1} \binom{n}{m}^{-1} \times \sum_{B \subset [A_\ell], |A_\ell| \leq m} \binom{m}{|B|}^{-1} \frac{1}{|B|} \sum_{b \in B} \sum_{C \in ([A_\ell] \setminus B, m-|B|)} D_b \mathbf{E} [h(X_{B \cup C}) | X_B] \\ &= \frac{n+1-m}{n+1} \binom{n}{m}^{-1} \times \sum_{B \subset [n+1], |B| \leq m} \binom{m}{|B|}^{-1} \frac{1}{|B|} \sum_{b \in B} \sum_{C \in ([n+1] \setminus B, m-|B|)} D_b \mathbf{E} [h(X_{B \cup C}) | X_B] \\ &= \binom{n+1}{m}^{-1} \times \sum_{B \subset [n+1], |B| \leq m} \binom{m}{|B|}^{-1} \frac{1}{|B|} \sum_{b \in B} \sum_{C \in ([n+1] \setminus B, m-|B|)} D_b \mathbf{E} [h(X_{B \cup C}) | X_B], \end{aligned}$$

where we have used in the first line that each subset  $B$  of  $[n+1]$  of cardinality  $k$  appears in  $n+1-k$  different subsets  $A_\ell$  (for  $l \in [n+1] \setminus B$ ), and in the same way, in the penultimate line, that each subset  $B \cup C$  of  $[n+1]$  of cardinality  $m$  appears in  $n+1-m$  different subsets  $A_\ell$  (for  $l \in [n+1] \setminus B \cup C$ ). Eventually, the case  $m = n+1$  follows from

$$\begin{aligned} & \sum_{k=1}^{n+1} \sum_{B \in ([n+1], k)} g_k(X_B) = h(X_{[n+1]}) - \theta \\ &= \sum_{B \subset [n+1]} \binom{n+1}{|B|}^{-1} \frac{1}{|B|} \sum_{b \in B} D_b \mathbf{E} [h(X_{[n+1]}) | X_B], \end{aligned}$$

by applying the Clark formula to  $h$ . □

*Proof of Theorem 3.1.6.* By the previous construction, for

$$i = (i_1, \dots, i_N) \in (I_k = k) \cap \bigcap_{m=k+1}^N (I_m \neq k),$$

the permutation  $\sigma = \Gamma(i)$  admits  $k$  as a fixed point. Hence,

$$\left\{ (I_k = k) \cap \bigcap_{m=k+1}^N (I_m \neq k) \right\} \subset (\tilde{U}_k^N = 1).$$

As both events have cardinality  $(N-1)!$ , they do coincide. The values of  $p_k$  and  $\alpha_k$  are easily computed since the random variables  $(I_m, k \leq m \leq N)$  are independent. According to Theorem 2.2.6,

$$\tilde{U}_k^N = \mathbf{E} \left[ \tilde{U}_k^N \right] + \sum_{\ell=1}^N D_\ell \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right] = \mathbf{E} \left[ \tilde{U}_k^N \right] + \sum_{\ell=1}^N \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right] - \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_\ell \right].$$

Since  $\tilde{U}_k^N \in \check{\mathcal{F}}_{k-1}$ ,  $D_\ell \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right] = 0$  for  $\ell < k$ . For  $\ell = k$ , we get

$$\begin{aligned} \mathbf{E} \left[ \mathbf{1}_{(I_k=k)} \prod_{m=k+1}^N \mathbf{1}_{(I_m \neq k)} | I_k, I_{k+1}, \dots \right] - \mathbf{E} \left[ \mathbf{1}_{(I_k=k)} \prod_{m=k+1}^N \mathbf{1}_{(I_m \neq k)} | I_{k+1}, I_{k+2}, \dots \right] \\ = \left( \mathbf{1}_{(I_k=k)} - \mathbf{P}_k(\{k\}) \right) \prod_{m=k+1}^N \mathbf{1}_{(I_m \neq k)}. \end{aligned}$$

For  $\ell = k+1$ ,

$$\begin{aligned} \mathbf{E} \left[ \mathbf{1}_{(I_k=k)} \prod_{m=k+1}^N \mathbf{1}_{(I_m \neq k)} | I_{k+1}, I_{k+2}, \dots \right] - \mathbf{E} \left[ \mathbf{1}_{(I_k=k)} \prod_{m=k+1}^N \mathbf{1}_{(I_m \neq k)} | I_{k+2}, I_{k+3}, \dots \right] \\ = tp_k \left( \mathbf{1}_{(I_{k+1} \neq k)} - \mathbf{P}_{k+1}(\{k\}^c) \right) \prod_{m=k+2}^N \mathbf{1}_{(I_m \neq k)} \\ = -tp_k \left( \mathbf{1}_{(I_{k+1}=k)} - \mathbf{P}_{k+1}(\{k\}) \right) \prod_{m=k+2}^N \mathbf{1}_{(I_m \neq k)}. \end{aligned}$$

The subsequent terms are handled similarly and the result follows.  $\square$

*Proof of Theorem 3.1.7.* By the very definition of  $\tilde{C}_1$ , we have

$$\tilde{C}_1 = \mathbf{E} \left[ \tilde{C}_1 \right] + \sum_{k=1}^N \sum_{\ell=k}^N D_\ell \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right]. \quad (3.4.2)$$

For  $k = \ell$ ,  $\mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right] = \tilde{U}_k^N$  and for  $\ell > k$ ,

$$\begin{aligned} \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right] &= \frac{t}{t+k-1} \left( 1 - \frac{1}{t+k} \right) \dots \left( 1 - \frac{1}{t+\ell-2} \right) \prod_{m=\ell}^N \mathbf{1}_{(I_m \neq k)} \\ &= \frac{t}{t+\ell-2} \prod_{m=\ell}^N \mathbf{1}_{(I_m \neq k)}. \end{aligned}$$

It is straightforward that  $\ell > k$ ,

$$\begin{aligned} D_\ell \left( \prod_{m=\ell}^N \mathbf{1}_{(I_m \neq k)} \right) &= \left( \mathbf{1}_{(I_\ell \neq k)} - \left( 1 - \frac{1}{t+\ell-1} \right) \right) \prod_{m=\ell+1}^N \mathbf{1}_{(I_m \neq k)} \\ &= - \left( \mathbf{1}_{(I_\ell=k)} - \frac{1}{t+\ell-1} \right) \prod_{m=\ell+1}^N \mathbf{1}_{(I_m \neq k)}. \end{aligned}$$

The result then follows by direct computations.  $\square$

*Proof of Theorem 3.1.9.* Recall that for  $j \neq l$ ,  $D_\ell \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right]$  and  $D_j \mathbf{E} \left[ \tilde{U}_m^N | \check{\mathcal{F}}_{j-1} \right]$  are orthogonal in  $L^2$ . In view of (3.4.2), according to the integration by parts formula, we have

$$\begin{aligned} \text{Var}[\tilde{C}_1] &= \sum_{k=1}^N \sum_{m=1}^N \sum_{\ell=k}^N \sum_{j=m}^N \mathbf{E} \left[ D_\ell \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right] D_j \mathbf{E} \left[ \tilde{U}_m^N | \check{\mathcal{F}}_{j-1} \right] \right] \\ &= \sum_{k=1}^N \sum_{m=1}^N \sum_{\ell=k \vee m}^N \mathbf{E} \left[ D_\ell \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right] D_\ell \mathbf{E} \left[ \tilde{U}_m^N | \check{\mathcal{F}}_{\ell-1} \right] \right] \\ &= 2 \sum_{k=1}^N \sum_{m=k+1}^N \sum_{\ell=m}^N \mathbf{E} \left[ \tilde{U}_k^N D_\ell \mathbf{E} \left[ \tilde{U}_m^N | \check{\mathcal{F}}_{\ell-1} \right] \right] + \mathbf{E} \left[ \sum_{k=1}^N \sum_{\ell=k}^N \tilde{U}_k^N D_\ell \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right] \right]. \end{aligned}$$

Then, for  $\ell \geq m > k$ ,

$$\begin{aligned} &\mathbf{E} \left[ \tilde{U}_k^N D_\ell \mathbf{E} \left[ \tilde{U}_m^N | \check{\mathcal{F}}_{\ell-1} \right] \right] \\ &= -\frac{t}{t+\ell-2} \mathbf{E} \left[ \mathbf{1}_{(I_k=k)} \prod_{p=k+1}^N \mathbf{1}_{(I_p \neq k)} \left( \mathbf{1}_{(I_\ell=m)} - \frac{1}{t+\ell-1} \right) \prod_{j=\ell+1}^N \mathbf{1}_{(I_j \neq m)} \right] \\ &= -\frac{t \mathbf{P}_k(\{k\})}{t+\ell-2} \left( \mathbf{P}_\ell(\{m\}) - \frac{1}{t+\ell-1} \right) \mathbf{E} \left[ \prod_{p=k+1}^{\ell-1} \mathbf{1}_{(I_p \neq k)} \right] \mathbf{E} \left[ \prod_{p=\ell+1}^N \mathbf{1}_{(I_p \notin \{k, m\})} \right] \\ &= 0, \end{aligned}$$

since, for any  $l \geq m > k$

$$\mathbf{E} \left[ \mathbf{1}_{(I_\ell=m)} \mathbf{1}_{(I_\ell \neq k)} \right] = \mathbf{E} \left[ \mathbf{1}_{(I_\ell=m)} \right] = \mathbf{P}_\ell(\{m\}) = \frac{1}{t+\ell-1}.$$

Furthermore, for  $\ell > k$ ,

$$\begin{aligned} &\mathbf{E} \left[ \tilde{U}_k^N D_\ell \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right] \right] \\ &= -\frac{t}{t+\ell-2} \mathbf{E} \left[ \mathbf{1}_{(I_k=k)} \prod_{p=k+1}^N \mathbf{1}_{(I_p \neq k)} \left( \mathbf{1}_{(I_\ell=k)} - \frac{1}{t+\ell-1} \right) \prod_{p=\ell+1}^N \mathbf{1}_{(I_p \neq k)} \right] \\ &= \frac{t}{(t+\ell-1)(t+\ell-2)} \mathbf{P}_k(\{k\}) \mathbf{E} \left[ \prod_{p=k+1}^N \mathbf{1}_{(I_p \neq k)} \right] \\ &= \frac{t^2}{(t+\ell-1)(t+\ell-2)(t+N-1)}, \end{aligned}$$

as  $\prod_{p=k+1}^N \mathbf{1}_{(I_p \neq k)} \mathbf{1}_{(I_\ell=k)} = 0$ , for  $\ell > k$ . Finally, for  $\ell = k$ , we get

$$\begin{aligned} \mathbf{E} \left[ \tilde{U}_k^N D_\ell \mathbf{E} \left[ \tilde{U}_k^N | \check{\mathcal{F}}_{\ell-1} \right] \right] &= \mathbf{E} \left[ \mathbf{1}_{(I_k=k)} \prod_{p=k+1}^N \mathbf{1}_{(I_p \neq k)} \left( \mathbf{1}_{(I_k=k)} - \frac{t}{t+k-1} \right) \prod_{p=k+1}^N \mathbf{1}_{(I_p \neq k)} \right] \\ &= \left( \frac{t}{t+k-1} - \frac{t^2}{(t+k-1)^2} \right) \frac{t+k-1}{t+N-1} \\ &= \frac{t(k-1)}{(t+k-1)(t+N-1)}. \end{aligned}$$

It follows that

$$\begin{aligned} \text{Var}[\tilde{C}_1] &= \frac{t^2}{t+N-1} \sum_{k=1}^N \sum_{\ell=k+1}^N \frac{1}{(t+\ell-1)(t+\ell-2)} + \frac{t}{t+N-1} \sum_{k=1}^N \frac{k-1}{t+k-1} \\ &= \frac{t}{t+N-1} \left( \frac{Nt}{t+N-1} + N - 2t \sum_{k=1}^N \frac{1}{t+k-1} \right). \end{aligned}$$

The proof is thus complete.  $\square$

### 3.4.2 Proofs of Section 3.2

*Proof of Theorem 3.2.3.* We have to compute

$$\sup_{\varphi \in \mathcal{T}} \mathbf{E} [\varphi'(F) - F\varphi(F)],$$

where  $\mathcal{T}$  is the set of twice differentiable functions with second order derivative bounded by 2. Since  $F$  is centered,

$$\mathbf{E} [F\varphi(F)] = \mathbf{E} [LL^{-1}F\varphi(F)] = \sum_{a \in A} \mathbf{E} [(-D_a L^{-1})F D_a \varphi(F)].$$

The trick is to use the Taylor expansion taking the reference point to be  $X'_{-a}$  instead of  $X_A$ . This yields

$$D_a \varphi(F) = \mathbf{E}' [\varphi(F(X_A)) - \varphi(F(X'_{-a}, X'_a))] = \varphi'(F(X'_{-a})) D_a F + R,$$

where

$$R = \frac{1}{2} \int_0^1 \mathbf{E}' \left[ \varphi''(\theta F(X'_{-a}) + (1-\theta)F(X_A)) (F(X_A) - F(X'_{-a}))^2 \right] d\theta.$$

Hence

$$\begin{aligned} \mathbf{E} [\varphi'(F) - F\varphi(F)] &= \mathbf{E} \left[ \varphi'(F) - \sum_{a \in A} \varphi'(F(X'_{-a})) D_a F (-D_a L^{-1}) F \right] \\ &\quad + \sum_{a \in A} \mathbf{E} [R (-D_a L^{-1}) F]. \end{aligned}$$

The rightmost term of the latter equation easily yields the rightmost of (3.2.1). Since  $\|\varphi''\|_\infty < 2$ , it is clear that  $\varphi'$  belongs to  $\text{Lip}_2$  hence the formulation of the distance with a supremum.  $\square$

*Proof of Corollary 3.2.6.* Without loss of generality, we can assume that  $X_i$  is centered for any  $i \geq 1$ . Remark that

$$D_j X_k = \begin{cases} 0 & \text{if } j \neq k, \\ X_k & \text{if } j = k. \end{cases}$$

Hence  $LY_n = Y_n$  and  $Y_n = L^{-1}Y_n$ . According to Theorem 3.2.3,

$$\begin{aligned} \mathfrak{d}_1(\mathbf{P}, \mathbf{P}_{Y_n}) &\leq \sup_{\psi \in \text{Lip}_2} \mathbf{E} \left[ \psi(\mathbf{F}) - \frac{1}{s_n^2} \sum_{i \in A} \psi \left( \mathbf{F} \left( Y_n - \frac{X_i - X'_i}{s_n} \right) \right) X_i^2 \right] \\ &\quad + \frac{1}{s_n^3} \sum_{j=1}^n \mathbf{E} \left[ \int_{\mathbf{E}_A} (X_i - x)^2 d\mathbf{P}_i(x) |X_i| \right]. \end{aligned}$$

By independence, since  $\psi$  is 2-Lipschitz continuous,

$$\begin{aligned} &\left| \mathbf{E} \left[ \psi(\mathbf{F}) - \frac{1}{s_n^2} \sum_{i \in A} \psi \left( \mathbf{F} \left( Y_n - \frac{X_i - X'_i}{s_n} \right) \right) X_i^2 \right] \right| \\ &= \left| \frac{1}{s_n^2} \sum_{i \in A} \sigma_i^2 \mathbf{E} \left[ \psi(\mathbf{F}) - \psi \left( \mathbf{F} \left( Y_n - \frac{X_i - X'_i}{s_n} \right) \right) \right] \right| \\ &\leq \frac{2}{s_n^3} \sum_{i \in A} \sigma_i^2 \mathbf{E} [|X_i - X'_i|] \leq \frac{2\sqrt{2}}{s_n^3} \sum_{i \in A} \sigma_i^3. \end{aligned}$$

Moreover,

$$\mathbf{E} \left[ \int_{\mathbf{E}_A} (X_i - x)^2 d\mathbf{P}_i(x) |X_i| \right] = \mathbf{E} [|X_i|^3] + \sigma^2 \mathbf{E} [|X_i|] \leq \mathbf{E} [|X_i|^3] + \sigma^3 \leq 2 \mathbf{E} [|X_i|^3],$$

according to the Hölder inequality. Hence the result.  $\square$

*Proof of Theorem 3.2.8.* According to the principle of Stein's method, we have to estimate

$$\mathbf{E} \left[ \frac{1}{\lambda} \left( \varphi(\mathbf{F}) + \frac{r}{\ell} \right) - \mathbf{F} \varphi'(\mathbf{F}) \right], \quad (3.4.3)$$

where  $\varphi$  and its derivatives satisfy (3.2.3). For any  $a \in A$ , thanks to the Taylor expansion,

$$-D_a \varphi(\mathbf{F}) = \mathbf{E}' [\varphi(\mathbf{F}(X^{-a}, X'_a)) - \varphi(\mathbf{F}(X))] = -\varphi'(\mathbf{F}) D_a \mathbf{F} + \mathbf{R}, \quad (3.4.4)$$

where

$$\mathbf{R} = \frac{1}{2} \int_0^1 (1 - \theta) \times \mathbf{E}' \left[ \varphi'' \left( (1 - \theta) \mathbf{F}(X) + \theta \mathbf{F}(X^{-a}, X'_a) \right) \left( \mathbf{F}(X) - \mathbf{F}(X^{-a}, X'_a) \right)^2 \right] d\theta. \quad (3.4.5)$$

According to (2.1.1) and to the definition of  $L$ ,

$$\begin{aligned} \mathbf{E} [\mathbf{F} \varphi(\mathbf{F})] &= \mathbf{E} [LL^{-1} \mathbf{F} \varphi(\mathbf{F})] = \mathbf{E} [-\delta(DL^{-1} \mathbf{F}) \varphi(\mathbf{F})] \\ &= \mathbf{E} [\langle D\varphi(\mathbf{F}), -DL^{-1} \mathbf{F} \rangle_{L^2(A)}]. \end{aligned} \quad (3.4.6)$$

Plug (3.4.4) into (3.4.6):

$$\begin{aligned} \mathbf{E} [\langle D\varphi(\mathbf{F}), -DL^{-1} \mathbf{F} \rangle_{L^2(A)}] &= - \sum_{a \in A} \mathbf{E} [D_a \varphi(\mathbf{F}) D_a (L^{-1} \mathbf{F})] \\ &= - \sum_{a \in A} \mathbf{E} [\varphi'(\mathbf{F}) D_a \mathbf{F} D_a (L^{-1} \mathbf{F})] + \sum_{a \in A} \mathbf{E} [\mathbf{R} D_a (L^{-1} \mathbf{F})] \\ &= \mathbf{E} [\varphi'(\mathbf{F}) \langle D\mathbf{F}, -DL^{-1} \mathbf{F} \rangle_{L^2(A)}] + \mathbf{E} [\langle \mathbf{R}, -DL^{-1} \mathbf{F} \rangle_{L^2(A)}]. \end{aligned}$$

Then,

$$\begin{aligned} & \left| \mathbf{E} \left[ \frac{1}{\ell} (\mathbf{F} + \frac{r}{\ell}) \varphi'(\mathbf{F}) - \mathbf{F} \varphi(\mathbf{F}) \right] \right| \\ & \leq \left| \mathbf{E} \left[ \varphi'(\mathbf{F}) \left( \frac{1}{\ell} (\mathbf{F} + \frac{r}{\ell}) - \langle \mathbf{D}\mathbf{F}, -\mathbf{D}\mathbf{L}^{-1}\mathbf{F} \rangle_{L^2(\mathbf{A})} \right) \right] \right| + \left| \mathbf{E} [\langle \mathbf{R}, -\mathbf{D}\mathbf{L}^{-1}\mathbf{F} \rangle_{L^2(\mathbf{A})}] \right| \\ & = B_1 + B_2. \end{aligned}$$

Since  $\varphi'$  is bounded, we get

$$B_1 \leq \|\varphi'\|_\infty \mathbf{E} \left[ \left| \frac{1}{\ell} (\mathbf{F} + \frac{r}{\ell}) - \langle \mathbf{D}\mathbf{F}, -\mathbf{D}\mathbf{L}^{-1}\mathbf{F} \rangle_{L^2(\mathbf{A})} \right| \right]$$

and from (3.4.5), we deduce that

$$B_2 \leq \|\varphi''\|_\infty \sum_{a \in \mathbf{A}} \mathbf{E} [|D_a \mathbf{F}|^2 |D_a \mathbf{L}^{-1} \mathbf{F}|].$$

The proof follows from (3.4.3) and (3.2.3). □

*Proof of Theorem 3.2.10.* For any  $a \in \mathbf{A}$ ,

$$D_a(\mathbf{X}_i \mathbf{X}_j) = \begin{cases} \mathbf{X}_a \mathbf{X}_j & \text{if } a = i \\ \mathbf{X}_i \mathbf{X}_a & \text{if } a = j \\ 0 & \text{otherwise.} \end{cases}$$

Then,

$$D_a \mathbf{F} = \sum_{(i,a) \in \mathbf{A}^\neq} f(i,a) \mathbf{X}_i \mathbf{X}_a + \sum_{(j,a) \in \mathbf{A}^\neq} f(a,j) \mathbf{X}_a \mathbf{X}_j = 2 \sum_{(i,a) \in \mathbf{A}^\neq} f(i,a) \mathbf{X}_i \mathbf{X}_a,$$

so that

$$\mathbf{L}\mathbf{F} = - \sum_{a \in \mathbf{A}} D_a \mathbf{F} = -2\mathbf{F} \quad \text{and} \quad \mathbf{L}^{-1}\mathbf{F} = -\frac{\mathbf{F}}{2}.$$

With our notations, the first term of the right-hand-side of (3.2.4) becomes

$$\mathbf{E} \left[ \left| 2\mathbf{F} + 2\nu - 2 \sum_{a \in \mathbf{A}} \sum_{(i,j) \in \mathbf{A}^2} f(i,a) f(j,a) \mathbf{X}_a^2 \mathbf{X}_i \mathbf{X}_j \right| \right] \leq \sum_{i=1}^2 A_i, \quad (3.4.7)$$

where

$$\begin{aligned} A_1 &= 2 \mathbf{E} \left[ \left| \sum_{(i,a) \in \mathbf{A}^2} f^2(i,a) (\mathbf{X}_a^2 \mathbf{X}_i^2 - 1) \right| \right], \\ A_2 &= 2 \mathbf{E} \left[ \left| \mathbf{F} - \sum_{a \in \mathbf{A}} \sum_{(i,j) \in \mathbf{A}^\neq} f(i,a) f(j,a) \mathbf{X}_a^2 \mathbf{X}_i \mathbf{X}_j \right| \right]. \end{aligned}$$

We first control  $A_1$ . According to the Cauchy-Schwarz inequality,

$$A_1^2 \leq 4 \mathbf{E} \left[ \sum_{(i,a) \in A^2} \sum_{(j,c) \in A^2} f^2(i,a) f^2(j,c) (X_a^2 X_i^2 - 1) (X_c^2 X_j^2 - 1) \right] \leq 4(A_{11} + A_{12}),$$

where

$$A_{11} = \mathbf{E} \left[ \sum_{(i,a) \in A^2} f^4(i,a) (X_a^2 X_i^2 - 1)^2 \right],$$

$$A_{12} = \mathbf{E} \left[ \sum_{a \in A} \sum_{(i,j) \in A^{\neq}} f^2(i,a) f^2(j,a) (X_a^2 X_i^2 - 1) (X_a^2 X_j^2 - 1) \right],$$

by orthogonality of the  $X_i$ 's. On the one hand,

$$A_{11} \leq \sum_{(i,a) \in A^2} f^4(i,a) \mathbf{E} \left[ (X_a^2 X_i^2 - 1)^2 \right] = (\mathbf{E} [X_1^4] - 1) \sum_{(i,a) \in A^2} f^4(i,a). \quad (3.4.8)$$

On the other hand,

$$\begin{aligned} A_{12} &= \mathbf{E} \left[ \sum_{(i,j) \in A^{\neq}} \sum_{a \in A} f^2(i,a) f^2(j,a) (X_a^2 X_i^2 - 1) (X_a^2 X_j^2 - 1) \right] \\ &\leq \sum_{(i,j) \in A^{\neq}} \sum_{a \in A} f^2(i,a) f^2(j,a) \mathbf{E} [(X_a^2 X_i^2 - 1) (X_a^2 X_j^2 - 1)] \\ &= (\mathbf{E} [X_1^4] - 1) \sum_{(i,a) \in A^2} f^2(i,a) \sum_{j \neq i} f^2(j,a) \\ &\leq (\mathbf{E} [X_1^4] - 1) \|f \star_2^1 f\|_{L^2(A)}^2. \end{aligned} \quad (3.4.9)$$

In a similar way,  $A_2 \leq A_{21} + A_{22}$ , where

$$A_{21} = 2 \mathbf{E} \left[ \left| \sum_{(i,j) \in A^{\neq}} f(i,j) X_i X_j - \sum_{(i,j) \in A^{\neq}} \sum_{a \in A} f(i,a) f(j,a) X_i X_j \right| \right],$$

$$A_{22} = 2 \mathbf{E} \left[ \left| \sum_{(i,j) \in A^{\neq}} \sum_{a \in A} f(i,a) f(j,a) X_i X_j (X_a^2 - \mathbf{E} [X_a^2]) \right| \right].$$

As above,

$$A_{21}^2 \leq 4 \mathbf{E} \left[ \left( \sum_{(i,j) \in A^{\neq}} (f(i,j) - \sum_{a \in A} f(i,a) f(j,a)) X_i X_j \right)^2 \right] = 4 \|f - f \star_1^1 f\|_2^2. \quad (3.4.10)$$



Furthermore, according to Cauchy-Schwarz inequality and by independence, we have

$$\begin{aligned}
A_{22} &\leq 2 \sum_{(i,j) \in A^\neq} \mathbf{E} \left[ |X_i X_j| \left| \sum_{a \in A} f(i, a) f(j, a) (X_a^2 - 1) \right| \right] \\
&\leq 2 \mathbf{E} \left[ \left( \sum_{(i,j) \in A^\neq} \sum_{a \in A} f(i, a) f(j, a) (X_a^2 - 1) \right)^2 \right]^{1/2} \\
&\leq 2 \left( \sum_{(i,j) \in A^\neq} \sum_{a \in A} f(i, a)^2 f(j, a)^2 \mathbf{E} [X_a^4 - 1] \right)^{1/2} \\
&\leq 2 (\mathbf{E} [X_1^4] - 1)^{1/2} \|f \star_{\frac{1}{2}} f\|_{L^2(A)}. \tag{3.4.11}
\end{aligned}$$

The remainder term is given by

$$A_3 = \sum_{a \in A} \mathbf{E} \left[ \int_{\mathbf{E}_A} \left( F(X_A) - F(X_{A-a}; x) \right)^2 d\mathbf{P}_a(x) |D_a L^{-1} F| \right].$$

Once again, using the orthogonality, we have

$$\begin{aligned}
G_a(X_A) &= \int_{\mathbf{E}_A} \left( F(X_A) - F(X_{A-a}; x) \right)^2 d\mathbf{P}_a(x) \\
&= 4 \mathbf{E}' \left[ \left( \sum_{i \in A} f(i, a) X_i X_a - \sum_{i \in A} f(i, a) X_i X'_a \right)^2 \right] \\
&= 4 \mathbf{E}' \left[ (X_a - X'_a)^2 \left( \sum_{i \in A} f(i, a) X_i \right)^2 \right] \\
&= 4 \left( \sum_{i \in A} f(i, a) X_i \right)^2 \mathbf{E}' [(X_a - X'_a)^2] \\
&= 4 \left( \sum_{i \in A} f(i, a) X_i \right)^2 (X_a^2 + 1).
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbf{E} \left[ \sum_{a \in A} G_a(X_A)^2 \right] &= 16 \mathbf{E} \left[ \sum_{a \in A} \left( \sum_{i \in A} f(i, a) X_i \right)^4 (X_a^2 + 1)^2 \right] \\
&= 16 (\mathbf{E} [X_1^4] + 3) \mathbf{E} [X_1^4] \sum_{a \in A} \sum_{i \in A} f^4(i, a) \\
&\quad + 96 (\mathbf{E} [X_1^4] + 3) \sum_{a \in A} \sum_{(i,j) \in A^\neq} f^2(i, a) f^2(j, a) \\
&\leq 16 (\mathbf{E} [X_1^4] + 3)^2 \sum_{a \in A} \sum_{i \in A} f^4(i, a) + 96 (\mathbf{E} [X_1^4] + 3) \|f \star_{\frac{1}{2}} f\|_{L^2(A)}^2. \tag{3.4.12}
\end{aligned}$$

Moreover,

$$\begin{aligned}
\sum_{a \in A} \mathbf{E} [ |D_a L^{-1} F|^2 ] &= \frac{1}{4} \sum_{a \in A} \mathbf{E} [ |D_a F|^2 ] \\
&= \sum_{a \in A} \mathbf{E} \left[ \left( \sum_{(i,a) \in A^\neq} f(i,a) X_i X_a \right)^2 \right] \\
&= \sum_{(i,a) \in A^\neq} f^2(i,a) = \nu.
\end{aligned} \tag{3.4.13}$$

Combine (3.4.8)–(3.4.13) to obtain (3.2.6).  $\square$

### 3.4.3 Proofs of Section 3.3

*Proof of Theorem 3.3.1.* As to construct the measure  $L_h$ , define the measure of probability  $\mathbf{P}_h$  on  $E^T$  by  $\mathbf{P}_h = \bigotimes_{t=1}^{N_T} \mathbf{P}_t^h$  where

$\mathbf{P}_t^h(U_t = 1) = p^* = p + \Delta_h p$ ,  $\mathbf{P}_t^h(U_t = -1) = q^* = q + \Delta_h q$ , and  $\mathbf{P}_t^h(U_t = 0) = s^* = s + \Delta_h s$  where  $s = 1 - p - q$  and with

$$\left\{ \begin{array}{ll} \Delta_h p = \frac{1}{2\lambda^3(\theta)\sigma} (\sqrt{\Delta t} \lambda(\theta) [\lambda(\theta) - \theta \lambda'(\theta)] + 2[\Delta t \theta \lambda(\theta) - \lambda'(\theta)(1 + \Delta t \theta^2)]) h & = [k(r) + g(r)]h \\ \Delta_h s = -\frac{\Delta t \theta \lambda(\theta) - \lambda'(\theta)(1 + \Delta t \theta^2)}{\lambda^3(\theta)\sigma} h & = -g(r)h \\ \Delta_h q = \frac{1}{2\lambda^3(\theta)\sigma} (-\sqrt{\Delta t} \lambda(\theta) [\lambda(\theta) - \theta \lambda'(\theta)] + 2[\Delta t \theta \lambda(\theta) - \lambda'(\theta)(1 + \Delta t \theta^2)]) h & = [-k(r) + g(r)]h, \end{array} \right.$$

since  $\Delta_h p + \Delta_h s + \Delta_h q = 0$ .

Define thus  $L_h = \prod_{t=1}^{N_T} L_t^h(U_t)$ , where  $L_t^h$  is the random variable defined on  $E_t$  by

$$L_t^h(U_t) = \frac{p^*}{p} \mathbf{1}_{\{U_t=1\}} + \frac{q^*}{q} \mathbf{1}_{\{U_t=-1\}} + \frac{s^*}{s} \mathbf{1}_{\{U_t=0\}}, \tag{3.4.14}$$

and such that the expectation under  $\mathbf{P}_h$  is defined by

$$\mathbf{E}_h [F] = \mathbf{E} [f(U_1, \dots, U_{N_T}) L_h] = \mathbf{E} \left[ f(U_1, \dots, U_{N_T}) \prod_{t=1}^{N_T} L_t^h(U_t) \right],$$

for any functional  $F$ .

Noting that  $\mathbf{E} [L_h] = 1$ , it follows from the Clark formula (2.2.3)

$$\begin{aligned}
L_h - 1 &= \sum_{t=1}^{N_T} \mathbf{E} [D_k L_h(U) | \mathcal{F}_t] \\
&= \sum_{t=1}^{N_T} \left\{ (L_t(U_t) - 1) \prod_{\ell=1}^{t-1} L_\ell(U_\ell) \right\},
\end{aligned}$$

since, for any  $r \in \mathbf{N}_T$ ,

$$\begin{aligned} \mathbf{E} [D_k L_h(U) | \mathcal{F}_r] &= \mathbf{E} \left[ \prod_{\ell=1}^{\mathbf{N}_T} L_\ell^h(U_\ell) - \mathbf{E}_h [L_h(U_t)] \times \prod_{\ell=1, \ell \neq t}^{\mathbf{N}_T} L_\ell^h(U_\ell) \middle| \mathcal{F}_k \right] \\ &= \left( L_t^h(U_t) - 1 \right) \prod_{\ell=1}^{t-1} L_\ell^h(U_\ell). \end{aligned}$$

Note thus that for any  $t \in \mathbf{N}_T$ ,

$$\begin{aligned} L_t^h(U_t) - 1 &= \frac{\Delta_h p}{p} \mathbf{1}_{(U_t=1)} + \frac{\Delta_h q}{q} \mathbf{1}_{(U_t=-1)} + \frac{\Delta_h s}{s} \mathbf{1}_{(U_t=0)} \\ &= \frac{[k(r) + g(r)]h}{p} \mathbf{1}_{(U_t=1)} + \frac{[k(r) - g(r)]h}{q} \mathbf{1}_{(U_t=-1)} - \frac{g(r)h}{s} \mathbf{1}_{(U_t=0)}, \end{aligned}$$

to state (3.3.5). Hence the result.  $\square$

## Part II

Malliavin calculus and enlargement of filtrations in discrete time: additional utility of an insider in a ternary model



The starting point for our second work arises in a discrete incomplete market, where we consider two agents with different information levels; the first one, called *insider*, possesses from the beginning extra information whereas the second one, the *ordinary agent*, bases his investment decisions on the public information flow. We ask the question, as the common thread of this second part:

How to quantify the additional utility of an insider in a incomplete discrete market model?  
(3.4.15)

The framework in which lies the insider problem is that of a financial market embodied by a probability space  $(\Omega, \mathbb{F}, \mathbf{P})$ . Its mathematical traduction boils down to the quantitative comparison of a regular agent whose information coincides with the natural filtration  $\mathbb{F}$  and an insider who possesses some extra information hidden in a random variable  $G$  from the beginning of the trading interval  $\mathbf{N}_T = [0, T] \cap \mathbf{N}$ . Two main approaches were developed to deal with this problem. Jorge A León, Reyla Navarro and David Nualart [84] stated that insider's portfolio satisfied an anticipative stochastic equation and used the tools of anticipating stochastic calculus to determine its maximal utility. The alternative approach, that we have chosen to follow is to address the problem from the point of view of martingale theory, as first envisaged by Karatzas and Pikovsky [115]. They made the insider evolve in a parallel probability space  $(\Omega, \mathbb{G}, \mathbf{P})$  where  $\mathbb{G}$  is an enrichment of the initial filtration  $\mathbb{F}$  and encompasses the additional information which benefits him. They converted the initial problem into that of a *grossissement de filtrations* one (see Jacod [70]), Yor ([141]) in which they exploited the powerful techniques. In the 90's the question of enlargement of filtrations aroused great interest and generated many works. Plenty applications in finance resulted from the related theories; to name but a few, the questions of arbitrage or indifference pricing can be addressed in terms of initial enlargement, whereas credit risk, or insurance issues can be modelised by a progressive enlargement situation.

In a continuous setting, the *Jacod's condition* indicates that the absolutely continuity of the conditional laws of  $G$  with respect to its law is a sufficient criterion to the preservation of semi-martingales. Transposing these works in the referred financial context, Imkeller highlighted the crucial role played by the *information drift*, i.e. the drift to eliminate in order to preserve the martingale property, and remarkably described it by the logarithmic Malliavin trace of the conditional laws of  $G$  with respect to  $\mathbb{F}$ . In discrete time, the question of enlargement of filtration seems easier. Indeed, Jacod's hypothesis is de facto satisfied in a discrete setting and we are directly provided with a "bridge formula" thanks to Doob's decomposition (see Blanchet-Scalliet, Jeanblanc and Romero [20]). This brings us to the question: can we provide an analogue of the information drift in a discrete context and express it in terms of Malliavin derivative?

In the very beginning we tried to fit our formalism to address the insider's problem in a the underlying trinomial model (see its definition in section 3.3). The difficulty came to our Clark-Ocone formula (2.2.3) from which we cannot derive a Karatzas-Ocone hedging formula. Indeed, the  $\mathcal{F}_k$ -measurability of the term  $D_k \mathbf{E} [F | \mathcal{F}_k]$  appearing in (2.2.3) prevented us from defining for the  $\mathbb{F}$ -predictable drift process we searched for. This observation was prone to replace the trinomial model with what we called a *ternary model*, equally distributed to the first one, but enjoying the properties of the jump processes on which it lies proved to be more conducive to state a predictable representation formula.

The second part consists of three chapters; the first one is an overview of enlargement of filtrations techniques and insider relative topics in familiar contexts. The new results are included in the two following chapters. The second chapter is devoted to the construction of Malliavin calculus for compound geometric processes. In the last one, we address the problem of insider trading in the so-called ternary market and apply our new formalism to provide a response to the problem 3.4.15.

## Chapter 4

# From a discrete insider model to Malliavin calculus for compound geometric processes

The chapter is organized as follows; in the first section we set up the question of insider's trading in a continuous setting and give some elements about enlargement of filtrations techniques. The second section is devoted to the problems of hedging and optimization of portfolios. In the following one, as to install our theory among what has been done, we introduce Malliavin calculus for jump processes. Last, we gather in the conclusion all the ingredients we need to solve the problem 3.4.15 we have chosen to investigate.

Troughout this chapter, we consider a simple *financial market* embodied by the couple of  $\mathbf{R}_+$ -valued processes, called *assets*,  $(A_t, S_t)_{t \in \mathbf{T}}$  defined on the same probability space  $(\Omega, \mathbb{F}, \mathbf{P})$  where  $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbf{T}}$  is a filtration (generally that generated by the canonical process) and such that :

- the process  $(A_t)_{t \in \mathbf{T}}$  is deterministic and modelises the risk-free asset,
- the process  $(S_t)_{t \in \mathbf{T}}$  is  $\mathbb{F}$ -adapted and modelises the risky asset, associated to its price  $S_t$  at time  $t \in \mathbf{T}$ .

The sequence of *discounted prices*  $(\bar{S}_t)_{t \in \mathbf{T}}$  is defined by  $\bar{S}_t = A_t^{-1} S_t$  ( $t \in \mathbf{T}$ ). On this market, operate different agents so that  $\mathbf{T}$  is called the *trading interval*; in a continuous setting it is equal to  $[0, T]$  ( $T \in \mathbf{R}$ ) and to  $\mathbf{N}_T = \mathbf{N} \cap [0, T]$  ( $T \in \mathbf{N}^*$ ) in a discrete one.

### 4.1 Insider's trading in a continuous setting and initial enlargement

Consider the market model described below, and assume in this section only that the sequence of discounted prices  $(\bar{S}_t)_{t \in \mathbf{T}}$  is a  $(\mathbb{F}, \mathbf{P})$ -martingale. The underlying question of *risk-neutral measure* will be addressed in the forthcoming section.

Until we transpose the problem into a discrete setting and that we mention it, let  $\mathbf{T} = [0, T]$  and consider the Wiener space  $(\Omega, \mathcal{A}, \mathbb{F}, \mathbf{P})$  equipped with the canonical process  $(B_t)_{t \in \mathbf{T}}$ , where



$\Omega$  is the set of continuous functions on  $\mathbf{T}$  starting at 0,  $\mathcal{A}$  the  $\sigma$ -algebra of Borel sets with respect to the uniform convergence of compact subsets of  $\mathbf{T}$ ,  $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbf{T}}$  the natural filtration generated by  $(B_t)_{t \in \mathbf{T}}$  (assume  $\mathcal{A} = \mathbb{F}$ ) and  $\mathbf{P}$  the Wiener measure. The financial market lying on  $(\Omega, \mathbb{F}, \mathbf{P})$  is defined through the couple of progressively measurable processes  $(r, \sigma)$ , where  $r$  stands for the mean rate of return and  $\sigma$  the volatility and satisfy the integrability conditions:

$$\int_0^T |r_t| dt < \infty \quad \text{and} \quad \int_0^T |\sigma_t|^2 dt < \infty, \quad \mathbf{P} - \text{a.s.}$$

The evolution of the stock process  $(S_t)_{t \in \mathbf{T}}$  is given by:

$$S_t = 1 + \int_0^t r_u S_u du + \int_0^t \sigma_u S_u dB_u, \quad t \in \mathbf{T}, \quad (4.1.1)$$

When the processes  $r$  et  $\sigma$  are deterministic and constant, the dynamics (4.1.1) of the underlying is a geometric Brownian motion and defines the so-called *Black-Scholes* model. Two agents act on the trading interval  $\mathbf{T}$ : the *ordinary agent* whose information level corresponds to the filtration  $\mathbb{F}$  i.e. whose knowledge at time  $t$  is given by  $\mathcal{F}_t$  and the *insider* who enjoys an information overload encoded by a  $\mathcal{F}_T$  or possibly  $\mathcal{F}_{T+\varepsilon}$ -measurable (for some small  $\varepsilon > 0$ ) random variable  $G$  with values in a Polish space  $(\Gamma, \mathcal{G})$ . Thus, the insider disposes at time  $t$  an information given by the  $\sigma$ -algebra  $\mathcal{G}_t$  defined by

$$\mathcal{G}_t = \mathcal{F}_t \vee \sigma(G).$$

We denote  $\mathbb{G} = (\mathcal{G}_t)_{t \in \mathbf{T}}$  (and  $\mathbb{G}^0 = (\mathcal{G}_t)_{t \in [0, T]}$ ) the insider's filtration. This depicts the easier case of enlargement of filtration called *initial enlargement*. The problem of *progressive enlargement* where  $\mathcal{G}_t = \mathcal{F}_t \vee \hat{\mathcal{F}}_t$  and  $\hat{\mathbb{F}} = (\hat{\mathcal{F}}_t)_{t \in \mathbf{T}}$  is an another filtration distinct of  $\mathbb{F}$  sounds to be more difficult, and won't be treated here.

The ordinary agent manages a  $\mathbb{F}$ -*portfolio* determined by the proportion of the wealth invested in stocks. It is encoded by a  $\mathbb{F}$ -progressively measurable process  $\psi$  called *strategy* such that

$$\int_0^T |\psi_t r_t| dt < \infty \quad \text{and} \quad \int_0^T |\psi_t r_t|^2 dt < \infty, \quad (\mathbf{P}, \mathbb{F}) - \text{a.s.},$$

which *value*  $V(\psi)$  satisfies the differential equation:

$$\frac{d\bar{V}_t(\psi)}{\bar{V}_t(\psi)} = \psi_t \frac{d\bar{S}_t}{\bar{S}_t}, \quad t \in [0, T]. \quad (4.1.2)$$

The maximal logarithmic utility of the ordinary agent whose initial wealth is  $x$  is defined via the maximization problem:

$$\Phi^{\mathbb{F}}(x) = \max_{\psi \in \mathbb{F}\text{-portfolio}} \mathbf{E} [\log(V_t^x(\psi))], \quad (4.1.3)$$

The insider's portfolio process, value and maximal logarithmic utility are analogously defined by systematically replacing  $\mathbb{F}$  by  $\mathbb{G}$ . Then, the additional expected logarithmic of the insider with initial wealth  $x$  is defined by:

$$\mathcal{U}(x) = \Phi^{\mathbb{G}}(x) - \Phi^{\mathbb{F}}(x) \quad (4.1.4)$$

where

$$\Phi^{\mathbb{G}} = \max_{\psi \in \mathbb{G}\text{-portfolio}} \mathbf{E} [\log(V_t^x(\psi^{\mathbb{G}}))]. \quad (4.1.5)$$

Introduce now the notion of *arbitrage*. From an economic point of view, an arbitrage designates a portfolio (i.e. the strategy it refers to) such that even if its initial value is zero, its terminal value is strictly positive, in other words, an "always winning strategy" called *free lunch*. Thus there exists an *arbitrage opportunity* in a financial market when making a profit without risk and without net investment of capital is possible. Mathematically, that translates as follows:

**Definition 4.1.1.** Given a filtration  $\mathbb{F}$ , consider a  $\mathbb{F}$ -semimartingale  $S$ . The model  $(S, \mathbb{F})$  has no *arbitrage* if there exists a positive  $\mathbb{F}$ -martingale  $L$  with  $L_0 = 1$  such that  $LS$  is a  $\mathbb{F}$ -martingale.

A financial market with no arbitrage opportunities is said to be *arbitrage free*. Two questions naturally arise: how to estimate the additional maximal logarithmic utility of the insider? Does the additional benefiting the insider provides him with an arbitrage opportunity?

For a given of  $\mathbb{F}$ -portfolio  $\psi$ , the solution of (4.1.2) expressed as

$$\bar{V}_t = \bar{V}_0 \exp \left( \int_0^t \psi_s \sigma_s dB_s - \frac{1}{2} \int_0^t \psi_s^2 \sigma_s^2 ds + \int_0^t \psi_s r_s ds \right)$$

such that the solution of (4.1.3) from the ordinary agent's point of view is provided by Merton's formula

$$\Phi^{\mathbb{F}}(x) = \frac{1}{2} \mathbf{E} \left[ \int_0^T \frac{r_s^2}{\sigma_s^2} ds \right],$$

which proof lies on the local martingale property of the stochastic integral  $\int_0^t \psi_s \sigma_s dB_s$ . Since the process  $(B_t)_{t \in \mathbf{T}}$  is not a  $\mathbb{G}$ -martingale, the stochastic integral  $\int_0^t \psi_s \sigma_s dB_s$  is not a  $\mathbb{G}$ -martingale any more, so that the previous computations no longer hold. This led to the conservation of martingale properties face to an enlargement of filtrations.

Main instigator of these questions, Jacod showed (Jacod Théorème 2.1 [70]) that every continuous local  $\mathbb{F}$ -martingale is a  $\mathbb{G}^0$ -martingale under the so-called *Jacod's hypothesis* :

**Assumption.** The regular conditional law of  $\mathbb{G}$  given  $\mathcal{F}_t$  is absolutely continuous with respect to the the law of  $\mathbb{G}$ ,  $\mathbf{P}$  almost everywhere, for all  $t \in [0, T)$ .

Its reinforcement, assuming of equivalence of the involved distributions (instead of simple absolute continuity) enables to define the Radon-Nikodym density process of the conditional laws of  $\mathbb{G}$  with respect to its law by

$$p_t^c(\omega) = \frac{d\mathbf{P}(\mathbb{G} \in \cdot | \mathcal{F}_t)(\omega)}{d\mathbf{P}(\mathbb{G} \in \cdot)}(c) ; c \in \Gamma, \omega \in \Omega, \quad (4.1.6)$$

that will be of relevance, and prone to exploit the techniques of Malliavin calculus. Provided Jacod's condition (4.1.6) holds, let the  $(\beta_t^{\mathbb{G}})_{t \in [0, T]}$  be the  $\mathbb{G}$ - progressively measurable process defined by

$$\beta_t^{\mathbb{G}} = \left. \frac{\frac{d}{dt} \langle p_t(\cdot, c), B_t \rangle_t}{p_t(\cdot, c)} \right|_{c=\mathbb{G}}$$

that we suppose to satisfy the additional integrability condition

$$\int_0^T |\beta_t^G| dt < \infty, \mathbf{P} - \text{a.s.}$$

Then the process

$$B = \tilde{B} + \int_0^\cdot |\beta_s^G| ds$$

is a  $\mathbb{G}$ -semimartingale with a  $\mathbb{G}$ -Brownian motion  $\tilde{B}$ . The maximal logarithmic utility of the insider whose initial wealth is  $x$  defined by (4.1.4) can be expressed as

$$\mathcal{U} = \frac{1}{2} \mathbf{E} \left[ \int_0^T (\beta_s^G)^2 ds \right],$$

so that the knowledge in advance of the information contained in the random variable  $G$  is quantifiable in terms of energy of the so-called *information drift*  $\beta^G$ . Extending the Clark-Ocone formula to measure valued martingales in order to apply it to the conditional laws of  $G$ ,

$$p_t^\ell = p_o^\ell + \int_0^t D_s p_s^\ell dB_s$$

P. Imkeller made clear the relationship between the random variable  $G$  and the information drift  $\beta$  that he identified - via the remarkable formula -

$$\beta_t^G = \frac{D_t p_t(\cdot, c)}{p_t(\cdot, c)} \Big|_{c=G} = D_t(\ln p_t(\cdot, c)) \Big|_{c=G},$$

with a logarithmic Malliavin derivative trace of the conditional density.

Until the end of the section, consider a discrete setting which consists of a probability space  $(\Omega, \mathbb{F}, \mathbf{P})$  where  $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbf{N}_T}$ . The evolution of the discounted stock process  $(\bar{S}_t)_{t \in \mathbf{N}_T}$  initially defined by  $S_0 = 1$  is dictated by  $(r_t)_{t \in \mathbf{N}_T}$  through the equation :

$$\frac{\Delta \bar{S}_t}{\bar{S}_t} = f_t(\alpha_t, X_t, \Delta X_t), \quad t \in \mathbf{N}_T.$$

where  $\Delta \bar{S}_t = \bar{S}_t - \bar{S}_{t-1}$  ( $t \in \mathbf{N}_T^0$ ),  $(X_t)_{t \in \mathbf{N}_T}$  is a  $\mathbb{F}$ -adapted process and  $f_t$  a function defined on  $\mathbf{R}_+ \times X_t(\Omega) \times \Delta X_t(\Omega)$ . For instance, in the so-called *Cox-Ross-Rubinstein* model,  $(X_t)_{t \in \mathbf{N}_T}$  is a Rademacher process and

$$f_t(r_t, X_t, \Delta X_t) = \frac{1}{1+r} \left( b \mathbf{1}_{\{X_t=1\}} + a \mathbf{1}_{\{X_t=-1\}} - 1 \right) \quad (4.1.7)$$

In other words, the dynamics of the stock price  $(S_t^{\text{bin}})_{t \in \mathbf{N}_T}$  satisfying the recurrent relation

$$S_t^{\text{bin}} = \begin{cases} (1+b)S_t^{\text{bin}} & \text{if } X_t = 1 \\ (1+a)S_t^{\text{bin}} & \text{if } X_t = -1 \end{cases}$$

The ordinary agent manages a  $\mathbb{F}$ -portfolio determined by the proportion of the wealth invested in stocks encoded by a  $\mathbb{F}$ -progressively measurable process  $\psi$  value  $V(\psi)$  (and discounted value  $\bar{V}(\psi)$ ) satisfies the self-financed condition given here by an equation in differences:

$$\frac{\Delta \bar{V}_t(\psi)}{\bar{V}_t(\psi)} = \psi_t \frac{\Delta \bar{S}_t}{\bar{S}_t}, \quad t \in \mathbf{N}_T^0. \quad (4.1.8)$$

The difficulties inherent with the preservation of semimartingales are directly lifted as a consequence of Doob's decomposition. Indeed it is clear (see Blanchet-Scalliet, Jeanblanc and Romero [20]) that any integrable process is a special semimartingale in any filtration with respect to which it is adapted.

The specific case where the  $\mathcal{F}_T$ -measurable  $G$  takes its values in a countable set  $\Gamma$  arouses interest since the enlargement of filtrations can be also interpreted in terms of drift information. Within this framework, define for any  $c \in \Gamma$ , the  $\mathbb{F}$ -martingale  $p^c$  by

$$p_t^c = \mathbf{P}(G = c \mid \mathcal{F}_t)$$

Then the process  $\bar{S}^G$  defined by

$$\bar{S}_t^G = \bar{S}_t - \sum_{s=1}^t \frac{\mathbf{E} [\Delta \bar{S}_s \Delta p_s^c \mid \mathcal{F}_{s-1}] \Big|_{c=G}}{p_{s-1}^G} \tag{4.1.9}$$

is a  $\mathbb{G}$ -martingale. We identify thus a drift information which one of the challenges of Chapter 3 will be to provide an expression in terms of Malliavin calculus (for a discrete setting to be precised).

**Theorem 4.1.2.** *If the random variable  $G$  is not  $\mathcal{F}_0$ -measurable, the model  $(S, \mathbb{G})$  is not arbitrage free.*

## 4.2 Hedging and portfolio's optimization in incomplete markets

Along with the computation of the maximal utilities of the two traders (the ordinary agent and the insider) as indicated in the previous section, the determination of the financial strategy to follow in order to realize the maximum is of interest; starting with that of the traders themselves. In other words, assuming that  $\Phi^{\mathbb{F}}(x)$  is known (for the ordinary agent), we wonder how to determine a  $\mathbb{F}$ -predictable process  $\psi$  which value  $V(\psi)$  satisfying (4.1.2) and such that (if  $\mathbf{T} = \mathbf{N}_T$  in a discrete setting)

$$V_0(\psi) = x, \quad V_T(\psi) = \Phi^{\mathbb{F}}(x) \quad \text{and} \quad V_t(\psi) \geq 0,$$

for any  $t \in \mathbf{N}_T$ . This turns out to estimate the quantities of the risk-free and risky assets with which the agent must compose his portfolio to maximize his utility. This leads us to tackle with the question of *hedging*. A *hedge* is basically an investment that protects the finances of a trading agent. Mathematically, this is the response to the problem: given a *claim*  $F$ , i.e. a  $\mathcal{F}_T$ -measurable random variable  $F$ , is there a  $\mathbb{F}$ -predictable process  $\psi$  such that (if  $\mathbf{T} = \mathbf{N}_T$  in a discrete setting)

$$V_T(\psi) = V_0(\psi) + \sum_{t=1}^T \psi_t \Delta S_t = F \quad \text{and} \quad V_t(\psi) \geq 0 \quad \text{for any } t \in \mathbf{N}_T ? \tag{4.2.1}$$

The section consists of two parts; in the first one, we indicate how a martingale approach (via martingale representation theorem) performs advantageously to solve the problem of hedging in *complete markets*. Focusing on incomplete markets in the following one, we drop this approach for technical reasons (as it will be explained) and tackle the problem from the standpoint of optimization issues.

The financial market is *complete* if all claims are reachable i.e. the answer to the question 4.2.1 is always positive. This is the case of the Black-Scholes model (see (4.1.1)) in a continuous paradigm or the Cox-Ross-Rubinstein (see (4.1.7)) model in a discrete setting. The *Fundamental Theorem of Asset Pricing* holds in both frameworks; we implicitly evoked when we assumed in section 4.1 that the sequence of discounted prices  $(\bar{S}_t)_{t \in \mathbf{T}}$  was a  $(\mathbf{P}, \mathbb{F})$ -martingale. In its most basic form it reads as follows:

**Theorem 4.2.1** (Fundamental Theorem of Asset Pricing). *The financial market lying on  $(\Omega, \mathbb{F}, \mathbf{P})$  is arbitrage-free if and only if there exists a probability measure  $\mathbf{P}^*$ , equivalent to  $\mathbf{P}$ , and such that the sequence  $(\bar{S}_t)_{t \in \mathbf{N}_T}$  is a  $(\mathbf{P}^*, \mathbb{F})$ -martingale. The measure  $\mathbf{P}^*$  is a risk-neutral measure. If the market is further assumed to be complete, there is a unique risk-neutral measure.*

The set of measures with respect to which the sequence of discounted prices  $(\bar{S}_t)_{t \in \mathbf{T}}$  is a  $(\cdot, \mathbb{F})$ -martingale will be called the set of *martingale measures* and denoted by  $\mathcal{C}^{\mathbb{F}}$ . The subset of  $\mathcal{C}^{\mathbb{F}}$  composed of the martingale measures equivalent to  $\mathbf{P}$  is denoted  $\mathcal{M}^{\mathbb{F}}$ . The Theorem 4.2.1 ensures that in a complete market, the set  $\mathcal{M}^{\mathbb{F}}$  is reduced to the singleton  $\{\mathbf{P}^*\}$ .

The result is of key importance to tackle with the problems of pricing or hedging. Indeed, by systematically reason under risk-neutral measure(s) we benefit from all tools the martingale theory is provided with.

In the Black-Scholes model, the value of the portfolio  $\psi = (\alpha, \varphi)$  denoted by the process  $V(\psi)$  is given at time  $t \in \mathbf{T}$  by

$$V_t(\psi) = \alpha_t A_t + \varphi_t S_t, \quad (4.2.2)$$

where  $\alpha_t$  and  $\varphi_t$  stand respectively for the amount of the riskless asset  $A_t$  and the amount of risky asset  $S_t$  composing the portfolio at time  $t$ . The self-financing condition(4.1.8) can be rewritten as

$$\alpha_t dA_t + \varphi_t dS_t = 0; \quad t \in \mathbf{T}. \quad (4.2.3)$$

In the Black-Scholes model lying a the Wiener space, the hedging formula is thus directly provided by the Clark formula under the unique neutral-risk probability measure  $\mathbf{P}^*$  (see prop 1.14.4 of chapter 1 in Privault [117]).

**Theorem 4.2.2** (Hedging formula in the Black-Scholes model). *Let a claim depicted by a random variable  $F \in L^2(\Omega)$ . Let the  $\mathbb{F}$ -adapted processes  $\alpha$  and  $\varphi$  defined by*

$$\varphi_t = \frac{(e^{-r(T-t)} \mathbf{E}^* [D_t F | \mathcal{F}_t])}{\sigma S_t} \quad \text{and} \quad \alpha_t = \frac{e^{-r(T-t)} \mathbf{E}^* [D_t F | \mathcal{F}_t] - \varphi_t S_t}{A_t}$$

for any  $t \in \mathbf{N}_T$ . The strategy  $\psi = (\alpha, \varphi)$  satisfies the self-financing (4.2.3) assumption and simulates  $F$ .

In the Cox-Ross-Rubinstein which underlying Rademacher space is also equipped with a Malliavin calculus (2.0.2). The value of the portfolio process  $V(\psi)$  is defined by (4.2.2) with the discrete time corresponding notations, and the self-financing assumption (4.1.8) boils down to: for any  $t \in \mathbf{N}_{T-1}$ ,

$$A_t(\alpha_{t+1} - \alpha_t) + S_t(\varphi_{t+1} - \varphi_t) = 0. \quad (4.2.4)$$

In discrete time models, a strategy  $\psi$  is *admissible* if it the self-financing (4.2.4) assumption and satisfies for any  $t \in \mathbf{N}_T$ ,

$$V_t(\psi) \geq 0.$$

Again, the hedging formula comes out by the application of the Clark formula to the claim  $F$ .

**Theorem 4.2.3** (Hedging formula in the Cox-Ross-Rubinstein model). *Let a claim depicted by a random variable  $F \in L^2(\Omega)$ . Let the  $\mathbb{F}$ -predictable processes  $\alpha$  and  $\varphi$  defined by  $\alpha_0 = (1+r)^{-T} \mathbf{E}^* [F]$ ,*

$$\varphi_t = \frac{(1+r)^{(T-t)} \mathbf{E}^* [D_t F | \mathcal{F}_{t-1}]}{\sqrt{pq}(b-a) S_{t-1}} \quad \text{and} \quad \alpha_t = \sum_{s=1}^t A_{s-1}^{-1} (S_{s-1}(\varphi_s - \varphi_{s-1}))$$

for any  $t \in \mathbf{N}_T$ . The strategy  $\psi = (\alpha, \varphi)$  is admissible and simulates  $F$ .

Up to now, let  $\mathbf{T} = \mathbf{N}_T$  and consider the simple discrete financial market described in section 4.1. When the market is incomplete, additional difficulties arise. All claims are no longer simulable. The hedging approach using Clark formula in the complete Cox-Ross-Rubinstein model for instance (Theorem 4.2.3), does not stand anymore.

The initial problem (4.1.5) can be addressed in terms of portfolio's optimization, and leads our steps towards the duality approach. A very substantiated expression and comprehensive look at arbitrage-related issues can be found in the book of Freddy Delbaen and Walter Schachermayer [42]; our work is greatly inspired by it. We yield thereafter some of the key points of the proof they used to deal with the optimization problem (4.1.3) in the trinomial model (see the definition (6.1.1)). This is a textbook case of incomplete models.

As explained by Wolfgang Runggaldier in [126] the trinomial model modelises an incomplete market, where the risk-neutral probability measure, with respect to which the sequence of discounted prices is a  $(\cdot, \mathbb{F})$ -martingale, is not unique. Let be a probability measure  $\mathbf{P}'$ , equivalent to  $\mathbf{P}$  and, for any  $t \in \mathbf{N}_{T-1}$  define

$$p_{t+1} = \mathbf{P}'(X_{t+1} = 1+b | \mathcal{F}_t), \quad q_{t+1} = \mathbf{P}'(X_{t+1} = 1+a | \mathcal{F}_t) \quad \text{and} \quad s_{t+1} = \mathbf{P}'(X_{t+1} = 1 | \mathcal{F}_t).$$

The measure  $\mathbf{P}'$  is a  $\mathbb{F}$ -martingale measure if and only if the triplet  $(p_{t+1}, q_{t+1}, s_{t+1})$  satisfies

$$(\mathcal{P}) : \begin{cases} \left( (1+b)p_{t+1} + (1+a)q_{t+1} + s_{t+1} \right) \bar{S}_t = (1+r)\bar{S}_t & ; \forall t \in \mathbf{N}_{T-2}. \\ p_{n+1} + q_{n+1} + s_{n+1} = 1 \end{cases}$$

The system  $(\mathcal{P})$  admits infinitely many solutions  $(p_t, q_t, s_t)_{T \in \mathbf{N}_{T-1}}$  such that any triplet  $(p_t, q_t, s_t)$  forms a convex  $\mathcal{M}^{\mathbb{F}}$  set (here a segment) characterized by its extremal points (independent of  $t$ ), i.e. the measures

$$\mathbf{P}_t^0 = \mathbf{P}^{0, \text{Tri}} = \left( \frac{r-a}{b-a}, \frac{b-r}{b-a}, 0 \right) \quad \text{and} \quad \mathbf{P}_t^1 = \mathbf{P}^{1, \text{Tri}} = \left( \frac{r}{b}, 0, \frac{b-r}{b} \right), \quad (4.2.5)$$

which are not equivalent to  $\mathbf{P}$  but such that any convex combination

$$\mathbf{P}^\gamma = \gamma \mathbf{P}^{0, \text{Tri}} + (1-\gamma) \mathbf{P}^{1, \text{Tri}},$$

is. Note that the extremal measures  $\mathbf{P}_n^0$  and  $\mathbf{P}_n^1$  are independent of  $t$  and  $\mathcal{F}_t$ , so that, if  $N = 2$  and  $\omega = (\omega_1, \omega_2) \in \Omega$ ,

$$\mathbf{P}'(\omega_1^i, \omega_2^j) = \mathbf{P}'(\omega_1^i) \mathbf{P}'(\omega_2^j | \omega_1^i) = \mathbf{P}'(\omega_1^i) \mathbf{P}'(\omega_2^j)$$

and by induction

$$\mathbf{P}'(\omega_1^{i_1}, \dots, \omega_t^{i_t}) = \prod_{s=1}^t \mathbf{P}'(\omega_s^{i_s}).$$

Then we can show that for any  $t \in \mathbf{N}_T$ , the set  $\mathbb{F}$ -martingale measures is described by a polyhedron which the  $J_t = 2^t$  vertices depict the extremal measure  $\mathbf{P}_t^j$  ( $j \in J_t$ ) that be can be written as

$$\mathbf{P}_t^j = \bigotimes_{s \in \mathbf{N}_t} (\mathbf{P}^{0, \text{Tri}})^{\gamma_s^j} (\mathbf{P}^{1, \text{Tri}})^{1-\gamma_s^j},$$

where  $(\gamma_s^j)_{s \in \mathbf{N}_t} \in \{0, 1\}^t$ .

As a reminder, the problem to solve is: given an economic agent disposing of  $x$  euros at date  $t = 0$  (initial budget constraint), determine the optimal admissible strategy allowing to maximize the value of the portfolio at time  $t \in \mathbf{N}_{T-1}$ . That is mathematically traduced by:

$$\Phi(x) = \max_{\psi \in \mathbb{F}\text{-portfolio}} \mathbf{E} [u(\bar{V}_{x,t}(\psi))] \quad (4.2.6)$$

where, for any  $t \in \mathbf{N}_{T-1}$ ,  $\bar{V}_{x,t}(\psi)$  stands for the discounted value at time  $t$  of a  $\mathbb{F}$ -portfolio corresponding to the strategy  $\psi = (\alpha_t, \varphi_t)_{t \in \mathbf{N}_{T-1}}$  and starting at  $x \in \mathbf{R}_+^*$  (the initial wealth), and  $u$  to an utility real function defined on  $\mathbf{R}_+^*$  or  $\mathbf{R}$  strictly increasing and strictly concave. Here consider the utility function

$$u : x \in \mathbf{R}_+^* \mapsto \log x,$$

denote  $I = (u')^{-1}$  its "inverse" function and  $u^*$  its Legendre-Fenchel transform, i.e. the convex function defined by

$$u^*(r) = \sup_{s \in \mathbf{R}_+^*} \{u(s) - sr\}. \quad (4.2.7)$$

Let  $\mathcal{S}^{\mathbb{F}}$  (resp.  $\mathcal{S}_t^{\mathbb{F}}$ ) the set of  $\mathbb{F}$ -predictable self-financing strategies on  $\mathbf{N}_T^0$  (resp. on  $\mathbf{N}_t^0$ ). For  $x \in \mathbf{R}_+^*$ , let

$$\Lambda_{x,t} = \{V : \exists \psi \in \mathcal{S}^{\mathbb{F}}, V = \bar{V}_t(\psi) \text{ and } V_0 = x\}.$$

As seen above, the trinomial market is not complete and there exists a convex set  $\mathcal{M}_t^{\mathbb{F}}$  of  $\mathbb{F}$ -martingales measures on  $\mathbf{N}_t^0$  with a finite number  $J_t$  of extremal values  $(\mathbf{P}_t^j, j \in J_t)$ . Denote  $\mathcal{C}_t^{\mathbb{F}}$  the convex subset of  $\mathcal{M}_t^{\mathbb{F}}$  formed by the  $\mathbb{F}$ -martingale measures equivalent to  $\mathbf{P}$ . Note in particular that for any strategy  $\psi = (\alpha, \varphi)$  with initial value  $x$  is written by

$$\bar{V}_{x,t}(\psi) = x + \sum_{s=1}^t \varphi_s \Delta \bar{S}_s,$$

and satisfies

$$\mathbf{E}_{\mathbf{P}_t^j}[\bar{V}_{x,t}(\psi)] = 0 ; j \in J_t. \quad (4.2.8)$$

As precisely developed by Freddy Delbaen and Walter Schachermayer in [42], this optimization problem can be solved by a duality approach decomposable in the following steps :

1. Rewrite (4.2.6) as a maximization problem under a finite number of linear constraints (4.2.9), called *primal problem*

$$\max_{V \in \Lambda_{x,t}} \mathbf{E} [u(V)] \quad \text{under the constraints} \quad \mathbf{E}_{\mathbf{P}_t^j}[V] \leq x ; \forall j \in J_t, \quad (4.2.9)$$

where

$$\Lambda_{x,t} = \left\{ V \in \mathcal{S}_t^{\mathbb{F}} : \mathbf{E}_{\mathbf{P}_t^j}[V] \leq x, \forall j \in J_t \right\}.$$

The equivalence of the problems (4.2.6) and (4.2.9) lies on the identities (4.2.8).

2. Determine its associated *dual problem* (4.2.10)

$$\sup_{V \in \Lambda_{x,t}} \Phi(V), \quad (4.2.10)$$

i.e. the optimization problem satisfied by the Lagrangian associated to the primal problem defined here by

$$\begin{aligned} \Lambda_{x,t} \times \mathbf{R}^{J_t} &\longrightarrow \mathbf{R}_+ \\ \mathcal{L} : (V, \lambda_1, \dots, \lambda_{J_t}) &\longmapsto \mathbf{E}[u(V)] + \sum_{j=1}^{J_t} \lambda_j \left( \mathbf{E}_{\mathbf{P}_t^j}[V] - x \right) \end{aligned}$$

3. Solve (4.2.10) by minimizing the value function

$$\psi(y) = \inf_{\mathbf{M} \in \mathcal{M}_t^{\mathbb{F}}} \mathbf{E} \left[ u^* \left( y \frac{d\mathbf{M}}{d\mathbf{P}} \right) \right] \quad (4.2.11)$$

where  $y = \sum_{j \in J_t} \lambda_j$ .

Thus, it can be stated that:

**Theorem 4.2.4** (Delbaen, Schachermayer [42]). *There exists unique optimizers  $\widehat{V}_{x,t}$  and  $\widehat{\mathbf{M}}_{\widehat{y}_{x,t}}$  of the problems (4.2.6) and (4.2.11) such that  $\widehat{\mathbf{M}}_{\widehat{y}_{x,t}} \in \mathcal{C}_t^{\mathbb{F}}$  and*

$$\widehat{V}_{x,t} = I \left( \widehat{y}_x \frac{d\widehat{\mathbf{M}}_{\widehat{y}_{x,t}}}{d\mathbf{P}} \right).$$

### 4.3 Malliavin calculus for jump processes

The impossibility to state a martingale representation theorem for any sequence of independent random variables as explained in the introduction of the part II, and thus to provide an hedging formula for the trinomial model within our formalism, enhanced us to investigate the option of an alternative financial model. This "model quest" led us to design a *ternary model* equivalent in distribution to the Cox-Ross-Rubinstein's (see Remark 6.1.1). This new model lying on what we have called a *geometric compound process* suggests to focus on jump models to understand its behavior. These models are additionally worthy of interest ; contrary to the Black-Scholes model, they integrate the possibility of the occurrence of rare events (sudden change in the international environment for instance, market crashes, gaps or opening jumps). For more details on the subject, see for instance the Chapter 1 of the book Rama Cont and Peter Tankov [30].

Up to know, consider a probability sapce  $(\Omega, \mathcal{A}, \mathbf{P})$  where are defined a standard Brownian motion  $(B_t)_{t \in \mathbf{R}_+}$ , a Poisson process  $(N_t)_{t \in \mathbf{R}_+}$  of intensity  $\lambda$  and a sequence of i.i.d.  $(-1, +\infty)$ -valued random variables  $\mathbf{V} = (V_t)_{t \in \mathbf{R}_+}$ . The  $\sigma$ -algebras generated by these three processes are supposed to be independent of each other. The filtration  $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbf{R}_+}$  is defined by

$$\mathcal{F}_t = \sigma(B_s, N_s, \mathbf{V} \cdot \mathbf{1}_{\{\cdot \leq N_s\}}), \quad s \leq t).$$

The riskless asset  $(A_t)_{t \in \mathbf{R}_+}$  is defined as in the Black-Scholes model (see 0.0.11) whereas the price of risky asset  $(S_t)_{t \in \mathbf{R}_+}$  is occasionally disjoint with significant jumps which values



and occurrences are respectively given by  $(V_t)_{t \in \mathbf{R}_+}$  and  $(T_t)_{t \in \mathbf{R}_+}$  (namely the times process associated to  $(N_t)_{t \in \mathbf{R}_+}$ ) and follows the Black-Scholes model between two jump times. Its dynamics characterizing a *jump-diffusion* process is provided in its differential form by:

$$dS_t = S_{t-} (r dt + \sigma dB_t + dZ_t) ; t \in \mathbf{R}_+,$$

where  $Z = (Z_t)_{t \in \mathbf{R}_+}$  defined by

$$Z_t = \sum_{s=1}^{N_t} V_s, \quad (4.3.1)$$

is a *compound Poisson process*.

When  $V$  is a deterministic constant process,  $Z$  coincides with the standard Poisson process on  $\mathbf{R}_+$ . Dean of the family of the jump processes, the latter is equipped with an extensive Malliavin theory, born shortly after that on Wiener space. Indeed, its associated compensated process is in fact a *normal martingale* so that most of fundamental properties such as martingale representation theorem, decomposition in chaos, Itô formula are immediately provided and make the theory powerful. For more details on Malliavin calculus for Poisson processes, see for instance chaos expansion (Last and Penrose [81], Privault [116]) for anticipative calculus (Nualart and Vives [103]).

When  $V$  denotes in a more general case an independent sequence of square-integrable random variables identically distributed with a probability distribution  $\nu$  defined on  $\mathbf{R}$ ,  $Z$  belongs to the family of compound Poisson processes, for which a stochastic calculus has already been designed. Extending what has been done before for the standard Poisson process, it is less rich in applications since the compound process does not enjoy all the properties of its illustrious elder. If the independence of increments they share makes possible a characterization in terms of Laplace's transform and to define stochastic integrals satisfying an Itô isometry property, it is however impossible to state a predictable representation property for compound Poisson processes in general. As suggested in the in the remark below, the associated compensated compound process  $(\bar{Y}_t)_{t \in \mathbf{R}_+}$  does not satisfy a structure equation when the process  $V$  is not deterministic constant.

**Remark 4.3.1.** As explained in the last item the remark p. 95 in Privault [117], the quadratic variation of the compensated compound  $\mathbb{F}$ -martingale  $(\bar{Y}_t)_{t \in \mathbf{R}_+}$  defined by

$$\bar{Y}_t = \frac{Y_t - \lambda t \mathbf{E}[V_1]}{\sqrt{\lambda \text{Var} V_1}}$$

satisfies

$$\begin{aligned} [\bar{Y}, \bar{Y}]_t &= \frac{1}{\sqrt{\lambda \text{Var}[V_1]}} \sum_{s=1}^{N_t} |V_s|^2 \\ &= \frac{1}{\sqrt{\lambda \text{Var}[V_1]}} \int_0^t \frac{|V_{1+N_{s-}}|^2}{V_{N_s}} d\bar{Y}_s + \frac{\mathbf{E}[V_1]}{\sqrt{\text{Var}[V_1]}} \int_0^t \frac{|V_{1+N_{s-}}|^2}{V_{N_s}} ds. \end{aligned}$$

If this last expression does not allow find a square integrable  $\mathbb{F}$ -adapted process  $(\phi_t)_{t \in \mathbf{R}_+}$  satisfying the *structure equation*

$$[\bar{Y}, \bar{Y}]_t = t + \int_0^t \phi_s d\bar{Y}_s, \quad (4.3.2)$$

in the general case, it seems to remain possible to state a kind of *modified structure equation*

$$[\bar{Y}, \bar{Y}]_t = \frac{1}{\sqrt{\lambda}} \sum_{s=1}^{N_t} 1 = \phi_t + \int_0^t \phi_s d\bar{Y}_s, \quad (4.3.3)$$

in the particular case where  $(|V|_t)_{t \in \mathbf{R}_+}$  is deterministic constant equal to 1

$$\Delta \bar{Y}_t = \frac{\phi_t}{2} \pm \sqrt{\left(\frac{\phi_t}{2}\right)^2 + \phi_t \Delta t}$$

(or under restrictive conditions  $|V_t| = 1$  and  $\mathbf{E}[V_t] = 0$  for any  $t \in \mathbf{R}_+$ ).

Despite of this intrinsic absence of predictable representation property, a Malliavin calculus can be elaborated for compound Poisson processes seen as a particular case a the larger class of Lévy processes to wich it belongs (see chapter 6 in [117], [109], [106]). The formalism thus conceived gives in particular birth to an analogue Clark-Ocone formula, inescapable tool to address hedging and portfolios optimization problems in jump-diffusion models. See for instance the works of Elisa Alòs, Jorge A. León, and Josep Vives [3], Ioannis Karatzas and Daniel Ocone [108].

## 4.4 Conclusion

In the same spirit as the conclusion of chapter 1, we close this one by listing what we need to answer the question: how can we quantify the additional utility of the insider trading on a financial market bore by a ternary model? The challenge is threefold: we have to understand the mechanics of the underlying compound geometric process, transpose one of the methods usually used for the optimization of dynamic portfolios to our framework, and transfer all these tools to an insider's point of view. In other words, we need to be equipped with:

- A Malliavin calculus for compound geometric processes and derived functional identities: martingale representation formula, Clark-Ocone type formula...
- A method for portfolio optimization in the incomplete market designed by a ternary model.
- A technique of enlargement of filtrations in this particular discrete setting.

If the chapter 5 will be charged the elaboration of a stochastic variational calculus for compound geometric process, the questions of enlargement of filtrations and portefolio optimization will be tackled in the last one to lead to the answer to question 3.4.15 of the second part.



## Chapter 5

# Malliavin calculus for compound geometric processes

As one of the different ways to construct it, we choose to develop a Malliavin calculus for geometric compound process by deriving it from a modified chaotic decomposition of square integrable functionals of that process.

### 5.1 Compound geometric processes

The general framework is that of a probability space  $(\Omega, \mathbb{F}, \mathbf{P})$  described in this first subsection. We borrow the notations and presentation from the framework introduced for filtered Poisson processes ([40]).

Let  $(\lambda, p) \in (0, 1)^2$ ,  $E = \{-1, 1\}$ ,  $T \in \mathbf{N}^*$ . Denote  $\mathbf{N}_t = \mathbf{N} \cap [1, t]$  and  $\mathbf{N}_t^0 = \mathbf{N}_t \cup \{0\}$  for any  $t \in \mathbf{N}_T$ . We consider  $(\mathbb{X}, \mathcal{T}, \nu)$  the measured space defined by  $\mathbb{X} = \mathbf{N}_T \times E$ ,  $\mathcal{T} = \mathcal{P}(\mathbb{X})$  and

$$\nu((t, k)) = \sum_{s \geq 1} \lambda \delta_s(\{t\}) \otimes (p \delta_1(\{k\}) + (1-p) \delta_{-1}(\{k\})),$$

for any  $(t, k) \in \mathbf{N}_T \times \{-1, 1\}$ ,  $\nu(0, \cdot) = 0$  and  $\eta$  its marginal distribution with respect to the time variable

$$\eta(t) = \delta_0(\{t\}) + \sum_{s \geq 1} \lambda \delta_s(\{t\}); \quad t \in \mathbf{N}_T.$$

For any  $n \in \mathbf{N}^*$  the tensor measure  $\nu^{\otimes n}$  is defined by:

$$\nu^{\otimes n}(\mathbf{t}_n, \mathbf{k}_n) = \mathbf{1}_{\Gamma_n}((\mathbf{t}_n, \mathbf{k}_n)) \prod_{i=1}^n \nu((t_i, k_i)), \quad (5.1.1)$$

where  $(\mathbf{t}_n, \mathbf{k}_n) = ((t_1, k_1), \dots, (t_n, k_n)) \in (\mathbf{N}_T \times E)^n$  and

$$\Gamma_n = \left\{ ((t_1, k_1), \dots, (t_n, k_n)) \in (\mathbf{N}_T \times E)^n : \forall i \neq j, t_i \neq t_j \right\}$$

The space of simple, integer-valued, locally finite measures on  $\mathbb{X}$  is denoted by  $\Omega$ .

**Remark 5.1.1.** The definition of the probability space  $(\Omega, \mathbb{F}, \mathbf{P})$  from the measured space  $(\mathbb{X}, \mathcal{T}, \nu)$  as described above is of key importance; it is through this underlying jump structure that we can afford to state a Karatzas-Ocone hedging formula for compound geometric

processes (6.3.5), and thus to circumvent the impossibility to do that for general independent random variables.

Let  $V_0 = 0$ ,  $T_0 = 0$ , and for any  $t \in \mathbf{N}_T$ , let  $T_t$  denote the time of the  $t$ -th jump,  $V_t$  the  $t$ -th mark,  $\xi_t$  the  $t$ -th inter-arrival and  $(N_t)_{t \in \mathbf{N}_T^0}$  the geometric process defined by  $N_0 = 0$  and

$$T_0 = 0, \quad T_t = \sum_{s=1}^t \xi_s, \quad N_t = \sum_{s=1}^{\infty} \mathbf{1}_{\{T_s \leq t\}},$$

where  $(V_t)_{t \in \mathbf{N}_T}$  and  $(\xi_t)_{t \in \mathbf{N}_T}$  are sequences of independent and identically distributed variables, and such that  $V_t$  is a  $\{-1, 1\}$ -Bernoulli variable of parameter  $p$ , whereas  $\xi_t$  is a geometric random variable of parameter  $\lambda$ . The canonical process  $\omega$  can be thus expressed by  $\omega(0, k) = 0$  and

$$\omega(t, k) = \sum_{s \in \mathbf{N}_T} \delta_{(T_s, V_s)}(t, k) = \sum_{s \in \mathbf{N}_T} \delta_{(N_t, W_s)}(s, k); \quad (t, k) \in \mathbf{N}_T \times E,$$

where, for any  $t \in \mathbf{N}_T$ , the random variable  $W_t$  is defined by

$$W_t(\omega) = \begin{cases} \sum_{k \in E} k \mathbf{1}_{\{1\}}(\omega(t, k)) & \text{if } t \leq T \\ \square & \text{if } t > T \end{cases}$$

and  $\square$  stands for a point at infinity. Note that any variable  $V_t$  ( $t \in \mathbf{N}_T$ ) is independent of the underlying jump process  $(N_t)_{t \in \mathbf{N}_T^0}$ . Define also the sequence of increments  $(\Delta N_t)_{t \in \mathbf{N}_T}$  by

$$\Delta N_t(\omega) = N_t(\omega) - N_{t-1}(\omega) = \begin{cases} \sum_{k \in E} \mathbf{1}_{\{1\}}(\omega(t, k)) & \text{if } t \leq T \\ \infty & \text{if } t > T \end{cases}$$

The variables  $\Delta N_t$  and  $W_t$  thus defined play a major part and we will often refer to them. Indeed,  $\Delta N_t$  indicates whether there is a jump at time  $t$ , and, if so, the variable  $W_t$  gives its value (+1 or -1) i.e. indicates if the process goes up or down.

**Remark 5.1.2.** To lighten the notations, we omit some parenthesis;  $\omega(t, k)$  and  $\delta.(t, k)$  stand respectively for  $\omega((t, k))$  and  $\delta.((t, k))$ .

From now on we consider geometric functionals  $F$  of the form

$$F = f_0 \mathbf{1}_{\{\omega(\infty)=0\}} + \sum_{n=1}^{\infty} \frac{1}{n!} f_n(\mathbf{t}_n, \mathbf{k}_n) \prod_{i=1}^n \mathbf{1}_{\{\omega(t_i, k_i)=1\}} \quad (5.1.2)$$

with  $f_n \in L^1(\mathcal{X}^n, \nu^{\otimes n})$  and  $(\mathbf{t}_n, \mathbf{k}_n) = ((t_1, k_1), \dots, (t_n, k_n)) \in (\mathbf{N}_T \times E)^{n, <}$ , where for any  $B \in \mathcal{P}(\mathbf{N}_T)$ , we denote

$$B^{n, <} = \{(t_1, \dots, t_n) : t_i \in B, i \in \mathbf{N}_T \text{ and } 1 \leq t_1 < \dots < t_n\}. \quad (5.1.3)$$

and for  $A = B \times E \in \mathcal{T}$ ,

$$(B \times E)^{n, <} = \{((t_1, k_1), \dots, (t_n, k_n)) : (t_i, k_i) \in B \times E, i \in \mathbf{N}_T \text{ and } t_1 < \dots < t_n\}.$$

For instance, if we consider the functional defined as

$$F(\omega) = \omega(A \times E),$$

for a given  $A = (\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times E)^{n, <}$  we have

$$f_n(\mathbf{t}_n, \mathbf{k}_n) = \sum_{i=1}^n \mathbf{1}_A((s_i, k_i)).$$

The canonical filtration  $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbf{N}_T^0}$  is the  $\sigma$ -field generated by the functionals of the form (5.1.2) and can be rewritten as

$$\mathcal{F}_0 = \{\{\emptyset\}, \Omega\} \quad \text{and} \quad \mathcal{F}_t = \sigma \left\{ \sum_{n=0}^t \sum_{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times E)^{n, <}} \omega(\mathbf{t}_n, \mathbf{k}_n) \right\}. \quad (5.1.4)$$

The compound geometric process  $Y$  of parameters  $(\lambda, p)$  is defined by

$$Y_t = \sum_{s \in \mathbf{N}_T} V_s \mathbf{1}_{\{T_s \leq t\}}. \quad (5.1.5)$$

**Remark 5.1.3.** The compound geometric process is the discrete analogue of the compound Poisson process. The underlying jump structure modelised by a Poisson process in this latter is replaced in our framework by a binomial/geometric process.

It is possible to define the space of square integrable functionals of the compound geometric process.

**Lemma 5.1.4 (★).** *The space  $L^2(\Omega)$  is made of the functions  $F : \Omega \rightarrow \mathbf{R}$  such that  $\int F^2 d\omega < \infty$  where*

$$\int_{\mathbb{X}} F^2 d\omega(t, k) = 1 + \sum_{n=1}^T \sum_{A \in (\mathbf{N}_T \times E)^{n, <}} F^2(A) \nu(A).$$

The set of simple processes, denoted by  $\mathcal{S}_0$  is the set of random variables of the form

$$u = \sum_{s \in \mathbf{N}_T} \sum_{k \in E} u_{s,k} \mathbf{1}_{(s,k)}.$$

## 5.2 Stochastic integrals

Throughout this subsection, assume the existence of a family  $\{\Delta R_{t,k}, t \in \mathbf{N}_T, k \in E\}$  of random variables satisfying the following hypotheses:

1. The process

$$\left( \sum_{s \in \mathbf{N}_t} \sum_{k \in E} \Delta R_{s,k} \right)_{t \in \mathbf{N}_T}$$

is a  $\mathbb{F}$ -martingale,

2. The family  $\mathcal{R} = \{\Delta R_{t,k}, t \in \mathbf{N}_T, k \in E\}$  is orthogonal for the scalar product  $(X, Y) \in L^2(\Omega) \mapsto \mathbf{E}[XY]$  and we denote  $\mathbf{E}[(\Delta R_{t,k})^2] = \kappa_k$  for any  $(t, k) \in \mathbf{N}_T \times E$ .

We draw our inspiration from the construction of Malliavin calculus for Gaussian fields, by embedding them in isonormal Gaussian processes. Similarly, and in order to encode the properties of family  $\{\Delta R_{t,k}, t \in \mathbf{N}_T, k \in \mathbf{E}\}$ , we introduce an Hilbert space  $\mathcal{H}$ , whose inner product is reproduced by the covariance structure of the family of random variables. Consider thus  $\mathcal{H}$  the space composed of  $\mathbf{R}^2$ -valued processes  $u = (u_{s,k}, k \in \mathbf{E})$  such that  $u$  is  $\mathbb{F}$ -predictable and

$$\|u\|_{\mathcal{H}}^2 = \mathbf{E} \left[ \sum_{k \in \mathbf{E}} \sum_{s=0}^T \kappa_k |u_{s,k}|^2 \right].$$

Endowed with the scalar product

$$\langle u, v \rangle_{\mathcal{H}} = \mathbf{E} \left[ \sum_{k \in \mathbf{E}} \sum_{s=0}^T \kappa_k u_{s,k} v_{s,k} \right],$$

the space  $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$  is hilbertian. The second Hilbert space is the set (of equivalence classes) of square-integrable  $\mathbb{F}$ -martingales over  $\mathbf{N}_T^0$  endowed by the scalar product

$$\langle M, L \rangle_{\mathcal{M}} = \mathbf{E} [\langle M_T, L_T \rangle_{\mathbf{E}}]$$

so that the space  $(\mathcal{M}, \|\cdot\|_{\mathcal{M}})$  is hilbertian. The Hilbert subspace of  $\mathcal{M}$  consisting of the 0-mean martingales of  $\mathcal{M}$  is denoted  $\mathcal{M}_0$ .

The simple stochastic integral can be defined with respect to this latter  $\mathbb{F}$ -martingale:

**Definition 5.2.1** (★). For any function  $f \in \ell^2(\mathbf{N}_T \times \mathbf{E})$ , the stochastic integral  $\tilde{J}_1(f)$  is defined by

$$\tilde{J}_1(f) = \sum_{s \in \mathbf{N}_T} \sum_{k \in \mathbf{E}} f(s, k) \Delta R_{s,k}.$$

In particular,  $\tilde{J}_1$  satisfies the identity:

$$\tilde{J}_1(\mathbf{1}_{(t,k)}) = \Delta R_{t,k}; (t, k) \in \mathbf{N}_T \times \mathbf{E}.$$

The isometry property verified by  $\tilde{J}_1$  enables to extend its definition to  $\mathcal{H}$ .

**Theorem 5.2.2** (★). *The stochastic integral defined as the application*

$$\begin{aligned} \tilde{J}_1 : \mathcal{H} &\longrightarrow \mathcal{M}_0 \\ u &\longmapsto \tilde{J}_1(u) = \sum_{s \in \mathbf{N}_T} \sum_{k \in \mathbf{E}} u_{s,k} \Delta R_{s,k} \end{aligned}$$

is an isometry. In other words,

$$\mathbf{E} \left[ \tilde{J}_1(u) \tilde{J}_1(v) \right] = \langle u, v \rangle_{\mathcal{H}}; u, v \in \mathcal{H}.$$

The stochastic integral  $\tilde{J}_1^t(f)$  of  $u \in \mathcal{H}$  is defined as

$$\tilde{J}_1^t(u) = \tilde{J}_1(u \mathbf{1}_{\llbracket 0, t \rrbracket}),$$

and satisfies the following lemma:

**Lemma 5.2.3** (★). For  $u \in \mathcal{H}$ ,

$$\mathbf{E} \left[ \tilde{J}_1(u) \mid \mathcal{F}_t \right] = \tilde{J}_1^t(u).$$

In order to define multiple stochastic integrals of random variables of the form (5.1.2), we can work in a space of symmetrical functions; indeed the occurrence of the jumps does not affect the value of  $F$ . Our construction follows closely that depicted by Nicolas Privault (see chapter 6 in [117]); in a certain sense we transpose it in our context. The space  $\ell^2(\mathbf{N}_T \times \mathbf{E})^{\circ 0}$  is by convention identified to  $\mathbf{R}$ ; let thus for any  $f \in \ell^2(\mathbf{N}_T \times \mathbf{E})^{\circ 0}$ ,

$$\tilde{J}_0(f_0) = f_0.$$

**Definition 5.2.4** (★). For  $n \in \mathbf{N}_T$ , let  $\ell^2(\mathbf{N}_T \times \mathbf{E})^{\circ n}$  denote the subspace of  $\ell^2(\mathbf{N}_T \times \mathbf{E})^{\otimes n} = \ell^2(\mathbf{N}_T \times \mathbf{E})^n$  composed of the functions  $f_n : (\mathbf{N}_T \times \mathbf{E})^n \rightarrow \mathbf{R}$  symmetric in their  $n$  variables, i.e. such that for any permutation  $\tau$  of  $\{1, \dots, n\}$ ,

$$f_n((t_{\tau(1)}, k_{\tau(1)}), \dots, (t_{\tau(n)}, k_{\tau(n)})) = f_n((t_1, k_1), \dots, (t_n, k_n)),$$

for any  $(t_1, k_1), \dots, (t_n, k_n) \in \mathbf{N}_T \times \mathbf{E}$ . The space  $\ell^2(\mathbf{N}_T \times \mathbf{E})^{\circ n}$  is endowed by the scalar product

$$\begin{aligned} \langle f_n, g_n \rangle_{\ell^2(\mathbf{N}_T \times \mathbf{E})^{\circ n}} &= n! \int_{(\mathbf{N}_T \times \mathbf{E})^n} f_n(\mathbf{t}_n, \mathbf{k}_n) g_n(\mathbf{t}_n, \mathbf{k}_n) d\nu^{\otimes n}(\mathbf{t}_n, \mathbf{k}_n) \\ &= n! \sum_{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times \mathbf{E})^{n, <}} f_n(\mathbf{t}_n, \mathbf{k}_n) g_n(\mathbf{t}_n, \mathbf{k}_n) \prod_{i=1}^n \kappa_{k_i} \end{aligned}$$

where we denote  $(\mathbf{t}_n, \mathbf{k}_n) = ((t_1, k_1), \dots, (t_n, k_n))$ .

**Theorem 5.2.5** (★ Multiple stochastic integral). The multiple stochastic integral  $\tilde{J}_n(f_n)$  of  $f_n \in \ell^2(\mathbf{N}_T \times \mathbf{E})^{\circ n}$  is defined as

$$\tilde{J}_n(f_n) = n! \sum_{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times \mathbf{E})^{n, <}} f_n((t_1, k_1), \dots, (t_n, k_n)) \prod_{i=1}^n \Delta R_{t_i, k_i}. \quad (5.2.1)$$

It satisfies the recurrence relation

$$\tilde{J}_n(f_n) = n \sum_{(t, k) \in \mathbf{N}_T \times \mathbf{E}} \tilde{J}_{n-1}(\pi_{t, k}^n f_n) \Delta R_{t, k} \quad (5.2.2)$$

where the function  $\pi_{t, \ell}^n$  is defined on  $\ell^2(\mathbf{N}_T \times \mathbf{E})^{\circ n}$  by

$$[\pi_{t, k}^n f_n]((t_1, k_1), \dots, (t_{n-1}, k_{n-1})) = f_n((t_1, k_1), \dots, (t_{n-1}, k_{n-1}), (t, k)) \mathbf{1}_{\llbracket 0, t \rrbracket^{n-1}}(t_1, \dots, t_{n-1})$$

**Remark 5.2.6.** The description of  $\mathbf{N}_T^{\leq}$  by (5.1.3) implies that the definition of the multiple stochastic integral  $\tilde{J}_n(f_n)$  holds if  $n \in \mathbf{N}_T$ .



Moreover, the application  $\tilde{J}_n$  satisfies an isometry property which enables to extend it to  $\mathcal{H}^{\otimes n}$ , the Hilbert space of  $(\mathbf{R}^2)^n$ -valued and  $\mathcal{F}$ -predictable processes  $u^{\otimes n}$  where  $u \otimes v$  denotes the element of  $L^2(\Omega \times \Omega, \mathbf{P} \otimes \mathbf{P})$  defined by

$$(u \otimes v)_{t,k}(\omega, \omega') = u_{t,k}(\omega) v_{t,k}(\omega'),$$

and endowed with the scalar product

$$\langle u, v \rangle_{\mathcal{H}^{\otimes n}} = n! \mathbf{E} [\langle u^{\otimes n}, v^{\otimes n} \rangle_{\ell^2(\mathbf{N}_T \times \mathbf{E})^{\circ n}}].$$

**Theorem 5.2.7 (★).** *The stochastic integral defined as the application*

$$\begin{aligned} \tilde{J}_n : \mathcal{H}^{\otimes n} &\longrightarrow \mathcal{H}_n \\ u^{\otimes n} &\longmapsto \tilde{J}_n(u^{\otimes n}) = n \sum_{t=1}^T \sum_{k \in \mathbf{E}} \tilde{J}_{n-1}(u_{t,k}^{\otimes n-1}) \Delta R_{t,k} \end{aligned}$$

satisfies the isometry formula

$$\mathbf{E} [\tilde{J}_n(u) \tilde{J}_n(v)] = \langle u, v \rangle_{\mathcal{H}^{\otimes n}} ; \forall u, v \in \mathcal{H}^{\otimes n}.$$

### 5.3 Modified chaos representation

Let  $\mathcal{H}_0 = \mathbf{R}$  and for  $n \in \mathbf{N}^*$ ,  $\mathcal{H}_n$  be the subspace of  $L^2(\Omega)$  made of integrals of order  $n \geq 1$ :

$$\mathcal{H}_n = \left\{ \tilde{J}_n(f_n) ; f_n \in \ell^2(\mathbf{N}_T^* \times \mathbf{E})^{\circ n} \right\},$$

and called *modified chaos of order n*. Let  $\mathcal{S}$  denote the linear space spanned by multiple stochastic integrals i.e.

$$\mathcal{S} = \text{Span} \left\{ \bigcup_{n=0}^T \mathcal{H}_n \right\}.$$

The completion of  $\mathcal{S}$  in  $L^2(\Omega)$  is denoted by the sum

$$\bigoplus_{n=0}^T \mathcal{H}_n.$$

**Lemma 5.3.1 (★).** *For  $t \in \mathbf{N}_T$ ,*

$$L^0(\Omega, \mathcal{F}_t) = (\mathcal{H}_0 \oplus \cdots \oplus \mathcal{H}_t) \cap L^0(\Omega, \mathcal{F}_t) \quad (5.3.1)$$

As a direct consequence of lemma 5.3.1, any random variable  $F \in L^0(\Omega, \mathcal{F}_t)$  can be expressed as

$$F = \mathbf{E}[F] + \sum_{n=1}^t \tilde{J}_n(f_n \mathbf{1}_{\llbracket 0, t \rrbracket^n})$$

As the sequence of jump times is bounded by  $T$  and the space  $L^2(\Omega)$  is characterized by Lemma 5.1.4, the chaotic decomposition naturally extends to  $L^2(\Omega)$  without any recourse to a limit procedure or a density argument.

**Theorem 5.3.2 (★).** *The space  $L^2(\Omega)$  is provided with the modified chaos decomposition property*

$$L^2(\Omega) = \bigoplus_{n=0}^T \mathcal{H}_n. \quad (5.3.2)$$

*In other words, any random variable  $F \in L^2(\Omega)$  can be expanded as*

$$F = \mathbf{E}[F] + \sum_{n=1}^T \tilde{J}_n(f_n). \quad (5.3.3)$$

In this subsection we aim at providing a decomposition of any square integrable random variable in terms of iterated integrals with respect to the family  $\mathcal{Z} = \{\Delta Z_{t,k}; (t,k) \in \mathbf{N}_T \times \mathbf{E}\}$  defined by

$$\Delta Z_{t,k} = \text{sgn}(k) (\mathbf{1}_{\{(1,k)\}} (\Delta N_t, V_{N_t}) - \lambda p_k) = \text{sgn}(k) (\mathbf{1}_{\{(1,k)\}} (\Delta N_t, W_t) - \lambda p_k).$$

The definition of  $\mathcal{Z}$  is quite natural since

$$\bar{Y}_t = \sum_{s \leq t} \sum_{k \in \mathbf{E}} \Delta Z_{s,k} \quad (5.3.4)$$

can be interpreted as the compensated  $\mathbb{F}$ -martingale associated to the compound geometric process  $Y$ . The family  $\mathcal{Z}$  is not orthogonal. The finite dimension of the relative spanned space, being equal to

$$1 + \sum_{s=1}^T 2^s \times \binom{T}{s} = 3^T,$$

we can derive from it an orthogonal family through the Gram-Schmidt process. The derived orthogonal family  $\mathcal{R} = \{\Delta R_{t,k}; (t,k) \in \mathbf{N}_T \times \mathbf{E}\}$  is defined by

$$R_0 = 1, \quad \Delta R_{1,1} = \Delta Z_{1,1} \quad \text{and} \quad \Delta R_{1,-1} = \Delta Z_{1,-1} - \frac{\lambda q}{1 - \lambda p} \Delta Z_{1,1}.$$

By induction, we have also for any  $t \in \mathbf{N}_T \setminus \{1\}$ ,

$$\Delta R_{t,1} = \Delta Z_{t,1} \quad \text{and} \quad \Delta R_{t,-1} = \Delta Z_{t,-1} - \frac{\lambda q}{1 - \lambda p} \Delta Z_{t,1} \quad (5.3.5)$$

by noting that for any  $s \in \mathbf{N}_{t-1}$ ,

$$\mathbf{E}[\Delta R_{s,k} \Delta Z_{t,1}] = \mathbf{E}[\Delta R_{s,k} \mathbf{E}[\Delta Z_{t,1} | \mathcal{F}_s]] = 0 \quad \text{and} \quad \mathbf{E}[\Delta R_{t,1} \Delta Z_{t,-1}] = \lambda^2 p q,$$

and where the first identity comes from the independence of  $\Delta N_t$  and  $\mathcal{F}_s$  for  $s \in \mathbf{N}_{t-1}$ .

**Remark 5.3.3.** Consider here the worthy of interest situations when the parameters  $\lambda$  and  $p$  take the extremal values  $\lambda = 1$  or  $p = 1$ . Note that the cases  $\lambda = 0$  or  $q = 1$  are sort of the negatives of the latter cases and thus left to the reader.

When  $\lambda$  is equal to 1, we retrieve the definition of the structure equation solution sequence  $(Y_t)_{t \in \mathbf{N}}$  defined by Nicolas Privault (chapter 1 - 1.4.6 in [117]). Up to a constant of normalization, the variable  $Y_t$  can be expressed with respect to the variables  $\Delta Z_t$ , by

$$Y_t = 2\sqrt{pq} (\Delta Z_{t,1} + \Delta Z_{t,-1}),$$

so that our underlying jump process coincides with a Rademacher sequence.

When  $p$  is equal to 1, the process  $(Y_t)_{t \in \mathbf{N}_T}$  defined by (5.3.4) boils down to a single compensated geometric process which properties could be soundly compared to these of a Poisson process on  $\mathbf{R}_+$  of intensity  $\lambda$ .

In what follows, we frequently refer to the Rademacher sequence (resp. the Poisson process on  $\mathbf{R}_+$  of intensity  $\lambda$ ) as a "guarantor" of our construction, and which we hope to retrieve the properties when letting  $\lambda$  equal to 1 (resp.  $p$  equal to 1).

The (multiple) stochastic integral with respect to  $\mathcal{Z}$  defined for any  $g_n \in \ell^2(\mathbf{N}_T \times \mathbf{E})^{\text{on}}$  by

$$J_n(g_n) = n! \sum_{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times \mathbf{E})^{n, <}} g_n((t_1, k_1), \dots, (t_n, k_n)) \prod_{i=1}^n \Delta Z_{t_i, k_i}$$

inherits of the properties of  $\tilde{J}_n$  by isomorphism between  $\text{Span } \mathcal{R}$  and  $\text{Span } \mathcal{Z}$ .

**Remark 5.3.4.** Let the application

$$\mathbf{1}^{<} : \begin{array}{ll} (\mathbf{N}_T \times \mathbf{E})^{n, <} & \longrightarrow \{0, 1\} \\ (\mathbf{t}_n, \mathbf{k}_n) & \longmapsto \mathbf{1}_{(\mathbf{t}_n, \mathbf{k}_n)} \end{array}$$

We retrieve the remarkable and usual identity

$$J_n(\mathbf{1}_{(\mathbf{t}_n, \mathbf{k}_n)}^{<}) = \prod_{i=1}^n \Delta Z_{t_i, k_i}; \quad n \in \mathbf{N}^*,$$

where  $(\mathbf{t}_n, \mathbf{k}_n) = ((t_1, k_1), \dots, (t_n, k_n))$ . This is of key importance; it basically means that we can reconstruct the signal  $Y$  (which connection with the family  $\mathcal{R}$  will be precised in section 5.3) by the means of the stochastic integral of elementary functions defined on  $(\mathbf{N}_T \times \mathbf{E})^n$ . In particular for  $n = 1$ ,

$$J_1(\mathbf{1}_{(t, k)}) = \Delta Z_{t, k}$$

appears as a reminiscence of (1.3.2).

We can thus state that:

**Theorem 5.3.5 (★).** *Any random variable  $F \in L^2(\Omega)$  can be expressed as*

$$F = \mathbf{E}[F] + \sum_{n=1}^T J_n(g_n).$$

where

$$g(\mathbf{t}_n, \mathbf{k}_n) = \sum_{m=1}^n \sum_{A \subset [n]^{m, <}} \prod_{i \in A} \rho^m \mathbf{1}_{\{k_i=1\}} f((\mathbf{t}_A, -\mathbf{k}_A), (\mathbf{t}_{[n] \setminus A}, \mathbf{k}_{[n] \setminus A}))$$

where  $(\mathbf{t}_n, \mathbf{k}_n) = ((t_1, k_1), \dots, (t_n, k_n))$ ,  $[n] = \{1, \dots, n\}$  and the functions  $f_n$  are the ones defined through (5.3.3) and by letting  $\rho = -\frac{\lambda q}{1 - \lambda}$ .

**Remark 5.3.6.** The latter representation formula can be put in perspective compared to remark 4.3.1. Our framework is close to that of a compound Poisson process where the process  $|V|$  is constant equal to one. It is thus possible to state a *modified structure equation* which interpretation (via Theorem 5.3.5) can be transposed in our context by: if it remains impossible to supply our compound geometric process with an usual chaotic decomposition for functionals of the form (5.1.2) (as for Brownian motion or in Poisson space), we can establish a *modified decomposition in chaos* in terms of multiple stochastic integrals of functions  $g_n$  where are the images of functions  $f_n$  (appearing in (5.1.2)), via a Gram-Schmidt process. The existence of this shift seems to be directly attributed to the presence of  $\phi_t$  in place of  $t$  in (4.3.2).

We introduce the Malliavin derivative as the *annihilation operator* acting on the space  $L^2(\Omega)$  seen in terms of its chaotic expansion (5.3.2).

**Definition 5.3.7.** Let the linear, unbounded, closable operator

$$D : L^2(\Omega) \longrightarrow L^2(\Omega \times \mathbf{N}_T^*, \mathbf{P})$$

defined for any element  $J_n(f_n)$  of  $\mathcal{H}_n$  by

$$D_{t,k}J_n(f_n) = n J_{n-1}(f_n(\star, (t, k))\mathbf{1}_{\llbracket 1, t-1 \rrbracket^{n, <}}). \quad (5.3.6)$$

where the notation  $\star$  will be used to indicate the first  $k-1$  variables  $((t_1, k_1), \dots, (t_{n-1}, k_{n-1}))$  of  $f_n((t_1, k_1), \dots, (t_{n-1}, k_{n-1}), (t, k))$ .

Let the application

$$\pi_t : \omega \longmapsto \sum_{s \neq t} \sum_{k \in \mathbf{E}} \omega(s, k)$$

i.e. the restriction of  $\omega$  to  $\mathcal{G}_t$ . The applications defined on  $\Omega \times (\mathbf{N}_t \times \mathbf{E})$ , and expressed for  $(\omega, (t, k)) \in \Omega \times (\mathbf{N}_t \times \mathbf{E})$  by

$$\omega \mapsto \pi_t(\omega) \cup (t, k) \quad \text{and} \quad \omega \mapsto \pi_t(\omega), \quad (5.3.7)$$

can be interpreted as the applications acting on  $\omega$  respectively by forcing a jump of height  $k$  at time  $t$  or forbidding any jump at time  $t$ .

**Remark 5.3.8.** In particular, the applications defined by (??) satisfy the remarkable identities:

- $W_s(\pi_t(\omega) \cup (t, k)) = k\mathbf{1}_{\{s\}}(t) + W_s(\omega)\mathbf{1}_{\{s\}^c}(t)$  and  $W_s(\pi_t(\omega)) = W_s(\omega)\mathbf{1}_{\{t\}^c}(s)$ ,
- $\Delta N_s(\pi_t(\omega) \cup (t, k)) = \mathbf{1}_{\{t\}}(s) + \Delta N_t(\omega)\mathbf{1}_{\{t\}^c}(s)$  and  $\Delta N_s(\pi_t(\omega)) = \Delta N_t(\omega)\mathbf{1}_{\{t\}^c}(s)$ .

We deduce that, for any  $(t, k) \in \mathbf{N}_T \times \mathbf{E}$  that on the one hand,

$$\Delta Z_{t,k}(\pi_t(\omega) \cup (t, k)) - \Delta Z_{t,k}(\pi_t(\omega)) = \text{sgn}(k)$$

and, on the other hand

$$\begin{aligned} \Delta Z_{t,k}(\pi_t(\omega) \cup (t, -k)) - \Delta Z_{t,k}(\pi_t(\omega)) &= \mathbf{1}_{\{(\Delta N_t, W_t) = (1, k)\}}(\pi_t(\omega) \cup (t, -k)) - \lambda p_k \\ &\quad - [\mathbf{1}_{\{(\Delta N_t, W_t) = (1, k)\}}(\pi_t(\omega)) - \lambda p_k] = 0, \end{aligned}$$

This entails in particular the remarkable identity:

$$\Delta R_{t,k}(\pi_t(\omega) \cup (t, k)) - \Delta R_{t,k}(\pi_t(\omega)) = \text{sgn}(k).$$

The above remark gives the tools to provide a more tractable expression of the Malliavin derivative, in terms of difference operator acting on  $L^2(\Omega)$ .

**Theorem 5.3.9** (★). *For any  $F \in L^2(\Omega)$ ,*

$$D_{t,k}F = \operatorname{sgn}(k) \left[ F(\pi_t(\omega) \cup (t, k)) - F(\pi_t(\omega)) \right]$$

**Remark 5.3.10.** By definition, given  $k \in E$ , the random variables  $\pi_t(\omega) \cup (t, k)$  and  $\pi_t(\omega)$  are  $\mathcal{G}_t$ -measurable; so does  $D_{t,k}F$ .

**Remark 5.3.11.** We can retrieve one of the specific identities existing in the Gaussian and Poisson spaces: for any process  $u \in \mathcal{H}$ ,

$$D_{t,k}J_1(u) = u(t, k).$$

The application of the definition of  $D_{t,k}$  to  $F = J_1(u)$  gives

$$\begin{aligned} D_{t,k}J_1(u) &= \operatorname{sgn}(k) \left[ \sum_{s \in \mathbf{N}_T} \sum_{\ell \in E} u(s, \ell) \Delta R_{s,\ell}(\pi_t(\omega) \cup (t, k)) - \sum_{s \in \mathbf{N}_T} \sum_{\ell \in E} u(s, \ell) \Delta R_{s,\ell}(\pi_t(\omega)) \right] \\ &= u(t, k) \Delta R_{t,k}(\pi_t(\omega) \cup (t, k)) + u(t, -k) \Delta R_{t,-k}(\pi_t(\omega) \cup (t, k)) \\ &\quad - u(t, k) \Delta R_{t,k}(\pi_t(\omega)) - u(t, -k) \Delta R_{t,-k}(\pi_t(\omega)) \\ &= \operatorname{sgn}(k)^2 u(t, k) = u(t, k), \end{aligned}$$

so that we get the desired result.

## 5.4 Functional identities

We provide our construction with the analogues of Itô formula, Clark formula, martingale representation theorem and Girsanov theorem.

The compound process  $(Y_t)_{t \in \mathbf{N}_T}$  defined by (5.1.5) is equipped with the pathwise Itô formula provided by the following theorem:

**Theorem 5.4.1** (★ Itô formula). *Let  $(Y_t)_{t \in \mathbf{N}_T}$  be the compound Poisson process defined by (5.1.5) and  $f$  be a  $\mathbf{R}$ -valued function defined on  $\mathbf{N}$ . Then, for any  $t \in \mathbf{N}_T$ ,*

$$f(Y_t) = f(0) + \sum_{s=1}^t (f(Y_s) - f(Y_{s-1})) Z_{N_s} \Delta Y_s \quad (5.4.1)$$

The face to face of theorems 5.4.2 and 5.4.5 highlights clearly their connection to each other (as for Brownian motion): while the latter one means that a martingale adapted to the compound geometric process filtration is in fact a stochastic integral, the first one gives the expression of the integrand of this stochastic integral in terms of the Malliavin derivative of the terminal value of the martingale.

**Theorem 5.4.2** (★ Clark formula). *For any  $F \in L^2(\Omega)$ ,*

$$F = \mathbf{E}[F] + \sum_{t \in \mathbf{N}_T} \sum_{k \in E} \mathbf{E}[D_{t,k}F | \mathcal{F}_{t-1}] \Delta R_{t,k} \quad (5.4.2)$$

As a direct consequence of Lemma 5.2.3 and Clark formula (5.4.2) we get:

**Corollary 5.4.3** (★). *For any  $t \in \mathbf{N}_T$  and  $F$  element of  $L^2(\Omega)$ ,*

$$F = \mathbf{E}[F | \mathcal{F}_t] = \sum_{s=t+1}^T \sum_{k \in E} \mathbf{E}[D_{s,k} F | \mathcal{F}_{s-1}] \Delta R_{s,k} \quad (5.4.3)$$

**Remark 5.4.4.** The transposition of the Clark formula with respect to the family  $\mathcal{Z}$  is written: for any  $F \in L^2(\Omega)$ ,

$$F = \mathbf{E}[F | \mathcal{F}_t] = \sum_{s=t+1}^T \sum_{k \in E} \left( b_{k,k} \mathbf{E}[D_{s,k} F | \mathcal{F}_{s-1}] + b_{k,-k} \mathbf{E}[D_{s,-k} F | \mathcal{F}_{s-1}] \right) \Delta Z_{s,k}$$

where the family  $\{b_{k,\ell}, (k, \ell) \in E^2\}$  defined by

$$b_{1,1} = 1, \quad b_{1,-1} = -\frac{\lambda q}{1 - \lambda p}, \quad b_{-1,1} = 0 \quad \text{and} \quad b_{-1,-1} = 1$$

as provided by (5.3.5).

From Corollary 5.4.3 we derive a predictable representation formula for geometric discrete-time martingales.

**Theorem 5.4.5** (★ Martingale representation theorem). *Let  $(X_t)_{t \in \mathbf{N}_T}$  be a  $\mathbb{F}$ -martingale. There exists a  $\mathbb{F}$ -predictable process  $u = (u_{\cdot,k}, k \in E)$  such that*

$$X_t = X_0 + \sum_{s=1}^t \sum_{k \in E} u_{s,k} \Delta R_{s,k}; \quad t \in \mathbf{N}_T. \quad (5.4.4)$$

In particular, for any  $(t, k) \in \mathbf{N}_T \times E$ ,

$$u_{t,k} = \mathbf{E}[D_{t,k} X_t | \mathcal{F}_{t-1}]$$

**Theorem 5.4.6** (★ Girsanov theorem). *Let  $(\tilde{\lambda}, \tilde{p}) \in (0, 1)^2$  and let the function  $\varphi$  defined on  $\{-1, 1\}$  such that*

$$\varphi = \frac{\tilde{\lambda}(1 - \lambda) \tilde{p} \mathbf{1}_{\{1\}} + \tilde{q} \mathbf{1}_{\{-1\}}}{\lambda(1 - \tilde{\lambda}) p \mathbf{1}_{\{1\}} + q \mathbf{1}_{\{-1\}}} - 1, \quad (5.4.5)$$

Then, under the probability measure

$$d\tilde{\mathbf{P}}_{\tilde{\lambda}, \tilde{p}} = \left( \frac{1 - \tilde{\lambda}}{1 - \lambda} \right)^t \prod_{s=1}^{N_t} (1 + \varphi(V_s)) d\mathbf{P}$$

the process

$$Y_t = \sum_{s=1}^{N_t} V_s$$

is a compound geometric process of parameters  $(\tilde{\lambda}, \tilde{p})$ .

The probability measure  $\mathbf{P}_{\tilde{\lambda}, \tilde{p}}$  can be equivalently expressed by

$$d\tilde{\mathbf{P}}_{\tilde{\lambda}, \tilde{p}} = \left( \frac{1 - \tilde{\lambda}}{1 - \lambda} \right)^t \prod_{s=1}^t (1 + \psi(\Delta N_s, W_s)) d\mathbf{P},$$

where  $\psi$  is the function defined on  $\mathbf{N}_T \times \mathbb{E}$  such that

$$\psi(\Delta N_t, W_t) = \Delta N_t \varphi(V_t).$$

**Remark 5.4.7.** Note that for any  $t \in \mathbf{N}_T$ ,

$$\begin{aligned} \psi(\Delta N_t, W_t) &= \Delta N_t \left( \frac{\tilde{\lambda}(1 - \lambda) \tilde{p}}{\lambda(1 - \tilde{\lambda}) p} \mathbf{1}_{\{(\Delta N_t, W_t) = (1, 1)\}} + \frac{\tilde{\lambda}(1 - \lambda) \tilde{q}}{\lambda(1 - \tilde{\lambda}) q} \mathbf{1}_{\{(\Delta N_t, W_t) = (1, -1)\}} - 1 \right) \\ &= \frac{\tilde{\lambda}(1 - \lambda)}{\lambda(1 - \tilde{\lambda})} \left( \frac{\tilde{p}}{p} \Delta Z_{t,1} - \frac{\tilde{q}}{q} \Delta Z_{t,-1} + \lambda \right). \end{aligned}$$

**Remark 5.4.8.** As in the Poisson case, we retrieve that the *shift space* appearing in Gaussian analysis (Cameron-Martin space for the Brownian motion in particular) is replaced here by a perturbation on what characterizes the jumps: their occurrence and their height, respectively parameterized by  $\lambda$  and  $p$ .

## 5.5 Link with the Malliavin calculus in Rademacher and in Poisson spaces

In the same vein as the remark 5.3.3, we devote this subsection to the comparison of our formalism to the Malliavin calculus in the Rademacher space (resp. the Poisson and Rademacher spaces) which we retrieve the structure when letting  $\lambda$  be equal to 1 (resp.  $p$  equal to 1).

Let first  $\lambda = 1$ . Basically, that means that the underlying geometric process jumps every time step. A Rademacher process  $(X_t)_{t \in \mathbf{N}_T^0}$  can be defined by letting  $X_t = V_t$  and the variables  $Y_t$  by

$$Y_t = \frac{\Delta Z_{t,1} + \Delta Z_{t,-1}}{\sqrt{2pq}} = \frac{X_t - p + q}{\sqrt{2pq}},$$

so that we can verify that  $(Y_t)_{t \in \mathbf{N}_T^0}$  defines a  $\mathbb{F}$ -normal martingale. If we let  $\bar{D}$  the derivative

$$\bar{D}_t F(X_1, \dots, X_T) = D_{t,1} G \left( \sum_{s=1}^T \sum_{k \in \mathbb{E}} \Delta Z_{s,k} \right) - D_{t,-1} G \left( \sum_{s=1}^T \sum_{k \in \mathbb{E}} \Delta Z_{s,k} \right)$$

where  $F$  and  $G$  as defined to satisfy  $F(X_1, \dots, X_T) = G \left( \sum_{s=1}^T \sum_{k \in \mathbb{E}} \Delta Z_{s,k} \right)$ , we get

$$\bar{D}_t Y_s = \frac{2}{\sqrt{2pq}} \mathbf{1}_{\{t\}}(s) = \sqrt{\frac{2}{pq}} \hat{D}_t Y_s$$

i.e. - up to a constant - the expression of the gradient  $\hat{D}$  defined on the Rademacher space (2.0.2). All identities and formulas, such as the Clark formula and the predictable representation (see Privault chapter 1), are inherited by construction.

Consider now the case  $p = 1$ . From a first point of view, that means that all variables  $\{V_t, t \in \mathbf{N}_T\}$  are deterministic (w.l.o.g. supposed equal to 1), and that the process  $(\bar{Y}_t)_{t \in \mathbf{N}_T}$  defined by (5.3.4) matches with a single compensated geometric process. The construction developed still holds and, as expected, enables to state stronger identities that are reminiscent of these holding for Poisson point processes on the real line. To take notice of it, consider the following framework:  $\mathcal{H}^P$  designates the Hilbert space of  $\mathbf{R}$ -valued and  $\mathbb{F}$ -predictable processes  $u$  such that

$$\|u\|_{\mathcal{H}^P}^2 = \mathbf{E} \left[ \sum_{t=0}^T \lambda(1-\lambda) |u_t|^2 \right].$$

Consider thus the space of simple, locally finite on  $\mathbf{N}_T$  integer-valued measures still denoted by  $\Omega^P$ . In particular, the family  $Z^P = \{\Delta Z_t^P; t \in \mathbf{N}_T\}$  defined by

$$\Delta Z_t^P = \mathbf{1}_{\{\Delta N_t=1\}} - \lambda,$$

is orthogonal for the canonical product on  $\Omega^P$  so that the stochastic integral defined as the application

$$\begin{aligned} I_1 : \mathcal{H}^P &\longrightarrow \mathcal{M}_0^P \\ u &\longmapsto I_1(u) = \sum_{t \in \mathbf{N}_T} u_t \Delta Z_t^P \end{aligned}$$

is an isometry. We can establish that for any  $F \in L^2(\Omega^P)$ ,

$$F = \mathbf{E}[F] + \sum_{t \in \mathbf{N}_T^0} \mathbf{E}[D_t^P F | \mathcal{F}_{t-1}] \Delta Z_t^P$$

where

$$D_t^P F = F(\omega \hat{+} \{t\}) - F(\omega \hat{-} \{t\}) \quad (5.5.1)$$

which is - up to a constant - a reminiscent of the gradient used by Laurent Decreusefond and Ian Flint [37] on the Poisson space

$$\nabla_t F = F(\omega \cup \{t\}) - F(\omega \setminus \{t\}).$$

From another point of view, consider the family of  $\{-1, 1\}$ -random variables  $\{X_t, t \in \mathbf{N}_T\}$  such that

$$X_t = 2\Delta Z_t^P + 2\lambda - 1 = 2\Delta N_t - 1$$

so that  $\lambda = \mathbf{P}(\{X_t = 1\}) = \mathbf{P}(\{\Delta N_t = 1\})$ . Define the random walk  $(S_t)_{t \in \mathbf{N}_T^0}$  by  $S_0 = 0$  and

$$S_t = \sum_{s=1}^t \frac{1 + X_s}{2} = N_t; \quad t \in \mathbf{N}_T.$$

The application of  $D$  to  $N = (N_t)_{t \in \mathbf{N}_T^0}$  gives

$$\begin{aligned} D_t F(N) &= \left[ F(\pi_t N \cup \{t\}) - F(\pi_t N) \right] \mathbf{1}_{\llbracket t, T \rrbracket}(\cdot) \\ &= F(S \cdot + \mathbf{1}_{\{X_t=-1\}} \mathbf{1}_{\llbracket t, T \rrbracket}(\cdot)) - F(S \cdot - \mathbf{1}_{\{X_t=1\}} \mathbf{1}_{\llbracket t, T \rrbracket}(\cdot)), \end{aligned}$$

by noting that  $\{\Delta N_t = 1\} = \{X_t = 1\}$  and  $\{\Delta N_t = 0\} = \{X_t = -1\}$ . We retrieve - up to a constant - the expression of Malliavin derivative  $\hat{D}$  defined on the Rademacher space (2.0.2)



(by taking  $p_k = \lambda$  in Corollary 1.6.3. in Privault [117]) applied to a functional  $F$  of the form  $F = f(X_1, \dots, X_T) = g(S_T)$ ,  $f : \{-1, 1\}^T \rightarrow \mathbf{R}$  and  $g : \mathbf{R} \rightarrow \mathbf{R}$  smooth functions. Besides, the Clark formula (5.5.1) is exactly the same of the eponymous one on the Rademacher space (se Proposition 1.7.1. in Privault [117]) noting that the definition of the process  $(\Delta Z_t^P)_{t \in \mathbf{N}_T}$  coincides with this of normal martingales through structure equations. All results arise from these definitions.

## 5.6 Proofs of chapter 5

### 5.6.1 Proofs of Section 5.2

*Proof of Proposition 5.2.2.* For any  $(u, v) \in \mathcal{H}^2$ ,

$$\begin{aligned} \mathbf{E} \left[ \tilde{J}_1(u) \tilde{J}_1(v) \right] &= \sum_{(t,k) \in \mathbf{N}_T \times \mathbf{E}} \sum_{(s,\ell) \in \mathbf{N}_T \times \mathbf{E}} u_{t,k} v_{s,\ell} \mathbf{E} [\Delta R_{t,k} \Delta R_{s,\ell}] \\ &= \sum_{t \in \mathbf{N}_T} \sum_{(k,\ell) \in \mathbf{E}^2} u_{t,k} v_{t,\ell} \mathbf{E} [\Delta R_{t,k} \Delta R_{t,\ell}] \\ &= \sum_{t \in \mathbf{N}_T} \sum_{k \in \mathbf{E}} \kappa_k u_{t,k} v_{t,k} \end{aligned}$$

where we have conditioning with respect to  $\mathcal{F}_{t \wedge s}$  to get the second line and we use the orthogonality of  $\Delta R_{t,k}$  and  $\Delta R_{t,\ell}$  (with  $\ell \neq k$ ) in the last one.  $\square$

*Proof of Lemma 5.2.3.* For any  $u \in \mathcal{H}$ ,

$$\begin{aligned} \mathbf{E} \left[ \tilde{J}_n(u) \mid \mathcal{F}_t \right] &= n! \sum_{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times \mathbf{E})^{n, <}} f_n(\mathbf{t}_n, \mathbf{k}_n) \mathbf{E} \left[ \prod_{i=1}^n \Delta R_{t_i, k_i} \mid \mathcal{F}_t \right] \\ &= n! \sum_{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_t \times \mathbf{E})^{n, <}} f_n(\mathbf{t}_n, \mathbf{k}_n) \mathbf{E} \left[ \prod_{i=1}^n \Delta R_{t_i, k_i} \mid \mathcal{F}_t \right] \\ &= \tilde{J}_n(u \mathbf{1}_{\llbracket 0, t \rrbracket}) \end{aligned}$$

since the independence of the centered variables  $\{\Delta R_{t_i, k_i}, (t_i, k_i) \in \mathcal{T}, i \in \mathbf{N}_n\}$  implies that  $\mathbf{E} \left[ \prod_{i=1}^n \Delta R_{t_i, k_i} \mid \mathcal{F}_t \right] = 0$  if there exists  $i_0 \in \mathbf{N}_T$  such that  $t_{i_0} > t$ .  $\square$

*Proof of Proposition 5.2.5.* For any  $f_n \in \ell^2(\mathbf{N}_T \times \mathbf{E})^{o_n}$ ,

$$\begin{aligned} \tilde{J}_n(f_n) &= n! \sum_{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times \mathbf{E})^{n, <}} f_n((t_1, k_1), \dots, (t_n, k_n)) \prod_{i=1}^n \Delta R_{t_i, k_i} \\ &= n! \sum_{(t,k)} \sum_{(\mathbf{t}_{n-1}, \mathbf{k}_{n-1}) \in (\mathbf{N}_T \times \mathbf{E})^{n, <}} f_n((t_1, k_1), \dots, (t_{n-1}, k_{n-1}), (t, k)) \Delta R_{t,k} \prod_{i=1}^{n-1} \Delta R_{t_i, k_i} \\ &= n \sum_{(t,k) \in \mathbf{N}_T \times \mathbf{E}} \tilde{J}_{n-1}(\pi_{t,k}^n f_n) \Delta R_{t,k} \end{aligned}$$

Hence the result.  $\square$

*Proof of Proposition 5.2.7.* Assume with no loss of generality that  $n > m$ .

For any  $(\mathbf{t}_n, \mathbf{s}_m) \in (\mathbf{N}_T)^{n, <} \times (\mathbf{N}_T)^{m, <}$ , there exists  $i_0 \in \mathbf{N}_n$  such that  $t_{i_0} \in \mathbf{t}_n \setminus \mathbf{s}_m$ . By independence of the random variable  $\Delta R_{t_{i_0}, k_{i_0}}$  with respect to the  $\sigma$ -algebra  $\mathcal{F}_{t_{i_0}-1}$ , for any  $(u, v) \in \mathcal{H}^{\otimes n} \times \mathcal{H}^{\otimes m}$ ,

$$\begin{aligned} \mathbf{E} \left[ \tilde{J}_n(u^{\otimes n}) \tilde{J}_m(v^{\otimes m}) \right] &= n! m! \sum_{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times \mathbf{E})^{n, <}} \sum_{(\mathbf{s}_m, \mathbf{l}_m) \in (\mathbf{N}_T \times \mathbf{E})^{m, <}} u^{\otimes n}(\mathbf{t}_n, \mathbf{k}_n) v^{\otimes m}(\mathbf{s}_m, \mathbf{l}_m) \\ &\quad \times \mathbf{E} \left[ \prod_{i=1}^n \prod_{j=1}^m \Delta R_{t_i, k_i} \Delta R_{s_j, \ell_j} \right] \\ &= n! m! \sum_{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times \mathbf{E})^{n, <}} \sum_{(\mathbf{s}_m, \mathbf{l}_m) \in (\mathbf{N}_T \times \mathbf{E})^{m, <}} u^{\otimes n}(\mathbf{t}_n, \mathbf{k}_n) v^{\otimes m}(\mathbf{s}_m, \mathbf{l}_m) \\ &\quad \times \mathbf{E} \left[ \Delta R_{t_{i_0}, k_{i_0}} \right] \mathbf{E} \left[ \prod_{\substack{i=1 \\ i \neq i_0}}^n \prod_{j=1}^m \Delta R_{t_i, k_i} \Delta R_{s_j, \ell_j} \middle| \mathcal{F}_{t_{i_0}-1} \right] \prod_{i=1}^{i_0-1} \prod_{j=1}^{i_0-1} \Delta R_{t_i, k_i} \Delta R_{s_j, \ell_j} \\ &= 0, \end{aligned}$$

and for  $m = n$ ,

$$\begin{aligned} \mathbf{E} \left[ \tilde{J}_n(u^{\otimes n}) \tilde{J}_n(v^{\otimes n}) \right] &= (n!)^2 \sum_{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times \mathbf{E})^{n, <}} \sum_{(\mathbf{s}_n, \mathbf{l}_n) \in (\mathbf{N}_T \times \mathbf{E})^{n, <}} u^{\otimes n}(\mathbf{t}_n, \mathbf{k}_n) v^{\otimes n}(\mathbf{s}_n, \mathbf{l}_n) \\ &\quad \times \mathbf{E} \left[ \prod_{i,j=1}^n \Delta R_{t_i, k_i} \Delta R_{s_j, \ell_j} \right] \\ &= (n!)^2 \sum_{\substack{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times \mathbf{E})^{n, <} \\ \mathbf{l}_n \in \mathbf{E}^{n, <}}} f_n(\mathbf{t}_n, \mathbf{k}_n) g_n(\mathbf{t}_n, \mathbf{l}_n) \mathbf{E} \left[ \prod_{i,j=1}^n \Delta R_{t_i, k_i} \Delta R_{t_i, \ell_i} \right] \\ &= n! \langle u, v \rangle_{\mathcal{H}^{\otimes n}}, \end{aligned}$$

since  $\mathbf{E} [\Delta R_{t,k} \Delta R_{t,\ell}] = \kappa_k \mathbf{1}_{\{k\}}(\ell)$ . Hence the result.  $\square$

### 5.6.2 Proofs of Section 5.3

*Proof of Lemma 5.3.1.* It suffices to note that  $\mathcal{H}_s \cap L^0(\Omega, \mathcal{F}_t)$  is generated by the orthogonal basis

$$\{1\} \cup \left\{ \prod_{i=1}^r \Delta R_{t_i, k_i}, 1 \leq t_1 < \dots < t_r \leq s, (k_1, \dots, k_r) \in \mathbf{E}^r \right\} \quad (5.6.1)$$

Indeed it appears that the family can be expressed in terms of multiple integrals as

$$\prod_{i=1}^s \Delta R_{t_i, k_i} = \tilde{J}_n \left( \mathbf{1}_{\{(t_1, k_1), \dots, (t_s, k_s)\}}^< \right),$$

and moreover, by equality of the dimensions of the two spaces at stake is equal to  $2 \binom{t}{s}$ . We conclude by remarking that the dimensions of the two spaces appearing in (5.3.1) are both

equal to

$$1 + \sum_{s=1}^t 2^s \times \binom{t}{s} = 3^t.$$

The proof is thus complete.  $\square$

*Proof of Theorem 5.3.5.* By induction on  $n$ ; in what follows  $[n]$  denotes the set  $\{1, \dots, n\}$ . For  $n = 1$  there exists  $f : \{1\} \times \mathbf{E} \rightarrow \mathbf{R}$  such that

$$\mathbf{F} = \mathbf{E}[\mathbf{F}] + \sum_{k \in \mathbf{E}} f((1, k)) \Delta \mathbf{R}_{1,k}.$$

Since  $\Delta \mathbf{Z}_{1,k} = \Delta \mathbf{R}_{1,k}$  and  $\Delta \mathbf{R}_{-1,k} = \Delta \mathbf{Z}_{-1,k} + \rho \Delta \mathbf{Z}_{1,k}$ , we get,

$$g((1, 1)) = f((1, 1)) + \rho f((1, -1)) \quad \text{and} \quad g((1, -1)) = f((1, -1))$$

Assume there is  $n \in \mathbf{N}_T$  such that

$$\mathbf{F} = \mathbf{E}[\mathbf{F}] + \sum_{m=1}^n J_m(g_m) + \sum_{m=n+1}^T \tilde{J}_m(f_m),$$

and where for  $m \in \mathbf{N}_n$ ,

$$g(\mathbf{t}_m, \mathbf{k}_m) = \sum_{\ell=1}^m \sum_{\mathbf{A} \subset [m]^{\ell, <}} \prod_{i \in \mathbf{A}} \rho^\ell \mathbf{1}_{\{k_i=1\}} f((\mathbf{t}_\mathbf{A}, -\mathbf{k}_\mathbf{A}), (\mathbf{t}_{[m] \setminus \mathbf{A}}, \mathbf{k}_{[m] \setminus \mathbf{A}})).$$

By definition, and applying induction hypothesis

$$\tilde{J}_n(f_n) = n \sum_{(t,k) \in \mathbf{N}_T \times \mathbf{E}} \tilde{J}_{n-1}(\pi_{t,k}^n f_n) \Delta \mathbf{R}_{t,k} = n \sum_{(t,k) \in \mathbf{N}_T \times \mathbf{E}} J_{n-1}(\pi_{t,k}^n g_n) \Delta \mathbf{Z}_{t,k}$$

Then

$$(\pi_{t_j, k_j}^n g_n((\mathbf{t}_n, \mathbf{k}_n))) = g((\mathbf{t}_n, \mathbf{k}_n)^{- (t_j, k_j)}) = \sum_{\ell=1}^{n-1} \sum_{\mathbf{A} \in ([n]^{-j})^{\ell, <}} \prod_{i \in \mathbf{A}} \rho^\ell \mathbf{1}_{\{k_i=1\}} f((\mathbf{t}_\mathbf{A}, -\mathbf{k}_\mathbf{A}), (\mathbf{t}_{[n]^{-j} \setminus \mathbf{A}}, \mathbf{k}_{[n]^{-j} \setminus \mathbf{A}}))$$

Noting that

$$g_n((\mathbf{t}_{n-1}, \mathbf{k}_{n-1}), (t_j, 1)) = f_n((\mathbf{t}_{n-1}, \mathbf{k}_{n-1}), (t_j, 1)) + \rho f_n((\mathbf{t}_{n-1}, \mathbf{k}_{n-1}), (t_j, -1)),$$

where  $[n]^{-j} = [n] \setminus \{j\}$ , we get

$$\begin{aligned} g_n((\mathbf{t}_{n-1}, \mathbf{k}_{n-1})^{- (t_j, k_j)}, (t_j, 1)) &= \sum_{\ell=1}^{n-1} \sum_{\mathbf{A} \in ([n]^{-j})^{\ell, <}} \prod_{i \in \mathbf{A}} \rho^\ell \mathbf{1}_{\{k_i=1\}} f((\mathbf{t}_\mathbf{A}, -\mathbf{k}_\mathbf{A}), (t_j, 1)), (\mathbf{t}_{[n]^{-j} \setminus \mathbf{A}}, \mathbf{k}_{[n]^{-j} \setminus \mathbf{A}})) \\ &+ \sum_{\ell=1}^{n-1} \sum_{\mathbf{A} \in ([n]^{-j})^{\ell, <}} \prod_{i \in \mathbf{A}} \rho^{\ell+1} \mathbf{1}_{\{k_i=1\}} f((\mathbf{t}_\mathbf{A}, -\mathbf{k}_\mathbf{A}), (\mathbf{t}_{[n]^{-j} \setminus \mathbf{A}}, \mathbf{k}_{[n]^{-j} \setminus \mathbf{A}}), (t_j, -1))) \\ &= \sum_{\ell=1}^{n-1} \sum_{\mathbf{A} \in ([n]^{-j})^{\ell, <}} \prod_{i \in \mathbf{A}} \rho^\ell \mathbf{1}_{\{k_i=1\}} f((\mathbf{t}_\mathbf{A}, -\mathbf{k}_\mathbf{A}), (t_j, 1)), (\mathbf{t}_{[n]^{-j} \setminus \mathbf{A}}, \mathbf{k}_{[n]^{-j} \setminus \mathbf{A}})) \\ &+ \sum_{\ell=2}^n \sum_{\mathbf{A} \in ([n]^{-j})^{\ell, <}} \prod_{i \in \mathbf{A}} \rho^\ell \mathbf{1}_{\{k_i=1\}} f((\mathbf{t}_\mathbf{A}, -\mathbf{k}_\mathbf{A}), (\mathbf{t}_{[n]^{-j} \setminus \mathbf{A}}, \mathbf{k}_{[n]^{-j} \setminus \mathbf{A}}), (t_j, -1))), \end{aligned}$$

and

$$g_n((\mathbf{t}_{n-1}, \mathbf{k}_{n-1})^{-\langle t_j, k_j \rangle}, (t_j, -1)) = \sum_{\ell=1}^{n-1} \sum_{A \in ([n]^{-j})^{\ell, <}} \prod_{i \in A} \rho^\ell \mathbf{1}_{\{k_i=1\}} f((\mathbf{t}_A, -\mathbf{k}_A, (t_j, -1)), (\mathbf{t}_{[n]^{-j} \setminus A}, \mathbf{k}_{[n]^{-j} \setminus A})).$$

Last write

$$\begin{aligned} g_n(\mathbf{t}_n, \mathbf{k}_n) &= g_n((\mathbf{t}_{n-1}, \mathbf{k}_{n-1})^{-\langle t_j, k_j \rangle}, (t_j, 1)) \mathbf{1}_{(k_j=1)} + g_n((\mathbf{t}_{n-1}, \mathbf{k}_{n-1})^{-\langle t_j, k_j \rangle}, (t_j, -1)) \mathbf{1}_{(k_j=-1)} \\ &= \sum_{\ell=1}^n \sum_{A \in ([n]^\ell)^{<}} \prod_{i \in A} \rho^\ell \mathbf{1}_{\{k_i=1\}} f((\mathbf{t}_A, -\mathbf{k}_A, (\mathbf{t}_{[n] \setminus A}, \mathbf{k}_{[n] \setminus A})) \end{aligned}$$

to complete the proof.  $\square$

*Proof of Theorem 5.3.9.* For any  $(t, k) \in \mathbf{N}_T \times \mathbf{E}$ , consider

$$\nabla_{t,k} \mathbf{F} = \text{sgn}(k) [\mathbf{F}(\pi_t(\omega) \cup (t, k)) - \mathbf{F}(\pi_t(\omega))]$$

$$\begin{aligned} \nabla_{t,k} \tilde{J}_n(f_n) &= n! \sum_{(\mathbf{t}_n, \mathbf{k}_n) \in (\mathbf{N}_T \times \mathbf{E})^{n, <}} f_n((t_1, k_1), \dots, (t_n, k_n)) \prod_{i=1}^n \nabla_{t_i, k_i} \Delta \mathbf{R}_{t_i, k_i} \\ &= n! \sum_{(\mathbf{t}_n^{-t}, \mathbf{k}_n^{-k}) \in (\mathbf{N}_T \times \mathbf{E})^{n-1, <}} f_n((t_1, k_1), \dots, (t, k), \dots, (t_n, k_n)) \prod_{\substack{i=1 \\ t_i \neq t}}^n \Delta \mathbf{R}_{t_i, k_i} \\ &= n! \sum_{(\mathbf{t}_{n-1}, \mathbf{k}_{n-1}) \in (\mathbf{N}_T^{-t} \times \mathbf{E})^{n, <}} f_n((\mathbf{t}_{n-1}, \mathbf{k}_{n-1}), (t, k)) \prod_{\substack{i=1 \\ t_i \neq t}}^n \Delta \mathbf{R}_{t_i, k_i} \\ &= n \tilde{J}_{n-1}(f_n(\star, (t, k))) \mathbf{1}_{\Delta_n^<} \end{aligned}$$

where we have used in the second line the remark 5.3.8 stating that

$$\nabla_{t,k} \Delta \mathbf{R}_{s, \ell} = \mathbf{1}_{\{(t, k)\}}((s, \ell))$$

so that

$$\nabla_{t,k} \left( \prod_{i=1}^n \Delta \mathbf{R}_{t_i, \ell_i} \right) = \left( \sum_{i=1}^n \mathbf{1}_{\{(t, k)\}}(t_i, k_i) \right) \prod_{\substack{i=1 \\ t_i \neq t}}^n \Delta \mathbf{R}_{t_i, \ell_i}.$$

Thus, for any  $\mathbf{F} \in \mathcal{S}$ ,  $\mathbf{D}_{t,k} \mathbf{F} = \text{sgn}(k) [\mathbf{F}(\pi_t(\omega) \cup (t, k)) - \mathbf{F}(\pi_t(\omega))]$ . By a limit procedure we extend this result to  $L^2(\Omega)$ .  $\square$

### 5.6.3 Proofs of Section 5.4

*Proof of Theorem 5.4.1.* The proof follows closely that of (see Proposition 19.12 in Privault [122]) For any  $t \in \mathbf{N}_T$

$$\begin{aligned}
f(Y_t) &= f(0) + \sum_{n=1}^{N_t} f(Y_{T_n}) - f(Y_{T_{n-1}}) \\
&= f(0) + \sum_{n=1}^{N_t} f(Y_{T_{n-1}} + V_n) - f(Y_{T_{n-1}}) \\
&= f(0) + \sum_{n=1}^{N_t} f(Y_{T_{n-1}} + V_{N_{T_n}}) - f(Y_{T_{n-1}}) \\
&= f(0) + \sum_{s=1}^t (f(Y_{s-1} + V_{N_s}) - f(Y_{s-1})) \Delta N_s \\
&= f(0) + \sum_{s=1}^t (f(Y_s) - f(Y_{s-1})) \Delta N_s \\
&= f(0) + \sum_{s=1}^t (f(Y_s) - f(Y_{s-1})) V_{N_s} \Delta Y_s
\end{aligned}$$

by noting that  $V_{N_s} \Delta Y_s = V_{N_s}^2 \Delta N_s = \Delta N_s$ . Hence the result.  $\square$

*Proof of Theorem 5.4.2.* Let  $F \in \mathcal{S}$  ; follows from the chaotic decomposition of  $F$  together with the definition of the operator gradient that

$$\begin{aligned}
F &= \mathbf{E}[F] + \sum_{n \geq 1} \tilde{J}_n(f_n \mathbf{1}_{\mathbf{N}^{n, <}}) \\
&= \mathbf{E}[F] + n \sum_{n \geq 1} \sum_{(t,k) \in \mathbf{N}_T \times \mathbf{E}} \tilde{J}_{n-1}(f_n(\star, (t, k)) \mathbf{1}_{\llbracket 0, t-1 \rrbracket^{n-1, <}}) \Delta R_{t,k} \\
&= \mathbf{E}[F] + n \sum_{n \geq 1} \sum_{(t,k) \in \mathbf{N}_T \times \mathbf{E}} \mathbf{E} \left[ \tilde{J}_{n-1}(f_n(\star, (t, k)) | \mathcal{F}_{t-1} \right] \Delta R_{t,k} \\
&= \mathbf{E}[F] + \sum_{(t,k) \in \mathbf{N}_T \times \mathbf{E}} \mathbf{E} [D_{t,x} F | \mathcal{F}_{t-1}] \Delta R_{t,k},
\end{aligned}$$

where we have used lemma 5.2.3 to get the third line. Hence the result. The result to any random variable  $F \in L^2(\Omega)$  is extended by noting that

$$\left| \mathbf{E} \left[ \sum_{(t,k) \in \mathbf{N}_T \times \mathbf{E}} u(t, k) \mathbf{E} [D_{t,x} F | \mathcal{F}_{t-1}] \right] \right| \leq \|F\|_{L^2(\Omega)} \|u\|_{L^2(\Omega \times \mathbb{X})}$$

and using the continuity of the operator  $\mathbf{E} [D_{t,k}(\cdot) | \mathcal{F}_{t-1}]$ .  $\square$

*Proof of Theorem 5.4.4.* Let  $(X_t)_{t \in \mathbf{N}_T}$  be a  $\mathbb{F}$ -martingale. Follows from Corollaries 5.3.9 and

5.4.3 applied to  $t = s - 1$  and  $F = X_s$  that

$$\begin{aligned} X_s &= \mathbf{E}[X_s | \mathcal{F}_{s-1}] + \sum_{k \in E} \mathbf{E}[D_{s,k} X_t | \mathcal{F}_{s-1}] \\ &= X_{s-1} + \sum_{k \in E} \mathbf{E}[D_{s,k} X_s | \mathcal{F}_{s-1}] \end{aligned}$$

Then by letting

$$u_{s,k} = \mathbf{E}[D_{s,k} X_s | \mathcal{F}_{s-1}],$$

we get

$$X_t = X_0 + \sum_{s=1}^t X_s - X_{s-1} = X_0 + \sum_{s=1}^t \sum_{k \in E} u_{s,k} \Delta R_{s,k}.$$

Hence the result.  $\square$

*Proof of Theorem 5.4.6.* The proof follows closely the proof of Theorem 19.16. in Privault [122] for compound Poisson processes. Let  $\varphi$  and  $\tilde{\mathbf{P}}_{\tilde{\lambda}, \tilde{p}}$  as defined in the theorem. For any  $s \in \mathbf{R}^*$ ,

$$\begin{aligned} \mathbf{E}_{\tilde{\lambda}, \tilde{p}}[s^{Y_t}] &= \left( \frac{1 - \tilde{\lambda}}{1 - \lambda} \right)^t \sum_{n=0}^t \mathbf{E} \left[ s^{Y_t} \prod_{k=1}^n (1 + \varphi(V_k)) \middle| N_t = n \right] \mathbf{P}(N_t = n) \\ &= \left( \frac{1 - \tilde{\lambda}}{1 - \lambda} \right)^t \sum_{n=0}^t \binom{t}{n} \lambda^n (1 - \lambda)^{t-n} \mathbf{E} \left[ \prod_{k=1}^n (1 + \varphi(V_k)) s^{V_k} \middle| N_t = n \right] \\ &= (1 - \tilde{\lambda})^t \sum_{n=0}^t \binom{t}{n} \left( \frac{\lambda}{1 - \lambda} \right)^n \left( \frac{\tilde{\lambda}(1 - \lambda)}{\lambda(1 - \tilde{\lambda})} \right)^n \prod_{k=1}^n \left( \frac{\tilde{p}}{p} \cdot ps + \frac{1 - \tilde{p}}{1 - p} \cdot \frac{1 - p}{s} \right) \\ &= \sum_{n=0}^t \binom{t}{n} \tilde{\lambda}^n (1 - \tilde{\lambda})^{t-n} \left( \tilde{p}s + \frac{1 - \tilde{p}}{s} \right)^n \\ &= \left( 1 - \tilde{\lambda} + \tilde{\lambda} \left( \tilde{p}s + \frac{1 - \tilde{p}}{s} \right) \right)^t. \end{aligned}$$

Hence the result.  $\square$



## Chapter 6

# The insider problem in the ternary model

### 6.1 Framework: the ternary model

In the perspective of exploiting our theoretical results to focus on the question of the additional utility of an insider in what we call the *ternary model*, we adopt now the following notations; our model is driven by an underlying compound geometric process  $Z = (Z_t)_{t \in \mathbf{N}_T}$  ( $\mathbf{N}_T = \mathbf{N} \cap [0, T]$ ) defined by

$$Z_t = \sum_{s=1}^{N_t} V_s. \quad (6.1.1)$$

As a reminder (see Section 5.1 for more details),  $(N_t)_{t \in \mathbf{N}_T^0}$  is the geometric process defined by  $N_0 = 0$  and

$$T_0 = 0, \quad T_t = \sum_{s=1}^t \xi_s, \quad N_t = \sum_{s=1}^{\infty} \mathbf{1}_{\{T_s \leq t\}},$$

where  $T_0 = 0$ ,  $(\xi_t)_{t \in \mathbf{N}_T}$  is a sequence of independent geometric random variables of parameter  $\lambda$  and  $(V_t)_{t \in \mathbf{N}_T}$  is a sequence of independent  $\{-1, 1\}$ -Bernoulli variables of parameter  $p \in (0, 1)$ . The canonical process  $\omega$  can be thus expressed by  $\omega(0, k) = 0$  and

$$\omega(t, k) = \sum_{s \in \mathbf{N}_T} \delta_{(T_s, V_s)}(t, k) = \sum_{s \in \mathbf{N}_T} \delta_{(N_t, W_s)}(s, k); \quad (t, k) \in \mathbf{N}_T \times \mathbf{E},$$

where, for any  $t \in \mathbf{N}_T$ , the random variables  $\Delta N_t$  and  $W_t$  are defined by

$$\Delta N_t(\omega) = \begin{cases} \sum_{k \in \mathbf{E}} \mathbf{1}_{\{1\}}(\omega(t, k)) & \text{if } t \leq T \\ \infty & \text{if } t > T \end{cases}$$

and

$$W_t(\omega) = \begin{cases} \sum_{k \in \mathbf{E}} k \mathbf{1}_{\{1\}}(\omega(t, k)) & \text{if } t \leq T \\ \square & \text{if } t > T \end{cases}$$

The financial market lying on the probability space denoted  $(\Omega, \mathbb{F}, \mathbf{P})$  is the one defined in 4.1. The riskless asset  $(A_t)_{t \in \mathbf{N}_T}$  with initial value  $A_0 = a_0$  is defined by

$$A_t = a_0(1 + r)^t, \quad (6.1.2)$$



whereas the stock price  $(S_t)_{t \in \mathbf{N}_T}$  with (deterministic) initial value  $S_0 = 1$  satisfies the equation:

$$\Delta S_t = \eta_t S_{t-1} \Delta N_t, \quad (6.1.3)$$

where  $\eta_t = b\mathbf{1}_{\{W_t=1\}} + a\mathbf{1}_{\{W_t=-1\}}$ ,  $a$  and  $b$  are real numbers such that  $-1 < a < r < b$ . The two assets evolve on a given finite time period  $\mathbf{N}_T^0$ . The sequence  $(\bar{S}_t)_{t \in \mathbf{N}_T}$  of discounted stock prices is given by

$$\bar{S}_t = \frac{1}{(1+r)^t} S_t, \quad ; \quad t \in \mathbf{N}_T.$$

**Remark 6.1.1.** The price process defined in our ternary model is identically distributed with the one of a well-chosen trinomial model. As a reminder, stock price  $(T_t)_{t \in \mathbf{N}_T}$  is defined in this latter model by  $T_0 = 1$  and verifies the recurrent relation:

$$T_t = \begin{cases} (1+b)T_{t-1} & \text{if } X_t = 1 \\ T_{t-1} & \text{if } X_t = 0 \\ (1+a)T_{t-1} & \text{if } X_t = -1 \end{cases},$$

where the process  $(X_t)_{t \in \mathbf{N}_T}$  is distributed according to the measure  $\mathbf{P}$  such that:

$$\mathbf{P}(X_t = 1) = \bar{p}, \quad \mathbf{P}(X_t = -1) = \bar{q} \quad \text{and} \quad \mathbf{P}(X_k = 0) = 1 - \bar{p} - \bar{q}.$$

with  $(\bar{p}, \bar{q}) \in (0, 1)^2$ . Let  $\bar{p} = \lambda p$  and  $\bar{q} = \lambda(1-p)$  such that  $1 - \bar{p} - \bar{q} = 1 - \lambda$ .

$$\begin{aligned} \mathbf{E} \left[ s^{\frac{S_t}{S_{t-1}}} \right] &= \mathbf{E} [s^{\eta_t \Delta N_t + 1}] \\ &= s^{b+1} \lambda p + s^{a+1} \lambda(1-p) + s(1-\lambda) \\ &= s^{1+b} \bar{p} + s^{1+a} \bar{q} + s(1-\lambda) \\ &= \mathbf{E} \left[ s^{\frac{T_t}{T_{t-1}}} \right], \end{aligned}$$

and  $S_0 = T_0$ . Thus the trinomial and the ternary models are equivalent in law. The second one, based on a jump process, lends itself more easily to the statement of an hedging formula, directly derived from Clark-Ocone formula (5.4.2). This motivated us to substitute it to the trinomial model for our further investigations.

**Remark 6.1.2.** It could be interesting to extend this ternary model to a "stochastic volatility-type" ternary model by allowing the up and down factors  $(1+b)$  and  $(1+a)$  to be stochastic. We could thus imagine to define the processes  $(b_{kt/n}, k = 1, \dots, n)$  and  $(a_{kt/n}, k = 1, \dots, n)$  as well-defined functions of the (stochastic volatility) Cox-Ingersoll-Ross process  $(\nu_t)_{t \in \mathbf{N}_T}$  defined by

$$\nu_t = \nu_0 + \int_0^t \kappa(\theta - \nu_s) ds + \int_0^t \sigma \sqrt{\nu_s} dB_s^\nu, \quad t \in \mathbf{T},$$

and so that, by letting  $n$  go to infinity, the "stochastic volatility ternary" model tends (at time  $t$ ) to the classical Heston model lying on the stochastic process  $H$

$$H_t = 1 + \int_0^t r H_s ds + \int_0^t \sqrt{\nu_s} H_s dB_s^H, \quad t \in \mathbf{T},$$

where  $r$  is the rate of return asset,  $\kappa$  denotes the speed of reversion,  $\sigma$  is the volatility of the volatility, and  $B^\nu, B^H$  are Wiener processes with covariance  $\rho dt$ .

### 6.1.1 Martingales measures in the ternary market model

Before exhibiting an hedging strategy for any for any simulating claim, we wonder about the completeness of the market. As explained by Runggaldier in *Portfolio optimization in discrete time*, like the trinomial model, our ternary model is an incomplete market. Indeed, the measure with respect to which the sequence of discounted prices is a  $\mathbb{F}$ -martingale, is not unique. Given the process  $(S_t)_{t \in \mathbf{N}_T}$  defined by (6.1.3) is identically distributed to the one of the trinomial model, we expect to reach the same incompleteness result. By writing for any  $t \in \mathbf{N}_{T-1}$ ,

$$\begin{aligned} \Delta \bar{S}_t &= \frac{S_t - (1+r)S_{t-1}}{(1+r)^t} \\ &= \frac{\eta_t S_{t-1} \Delta N_t - r S_{t-1}}{(1+r)^t} \\ &= \frac{[b\mathbf{1}_{\{W_t=1\}} + a\mathbf{1}_{\{W_t=-1\}}] \Delta N_t - r}{(1+r)^t} \times S_{t-1}, \end{aligned}$$

it appears that the discounted prices sequence is a  $\mathbb{F}$ -martingale if the condition

$$\lambda_t(bp_t + aq_t) - r = 0,$$

holds for any  $t \in \mathbf{N}_T$ . As expected, the system

$$\begin{cases} \lambda(bp_t + aq_t) &= r \\ p_t + q_t &= 1 \end{cases}$$

admits infinitely many solutions  $(\lambda_t, p_t, q_t) \in (0, 1)^3$  such that any triplet  $(\lambda_t, p_t, q_t)$  forms a convex  $\mathcal{M}^{\mathbb{F}}$  set (here a segment) characterized by its extremal points (independent of  $t$ ), i.e. the measures

$$\mathbf{P}_t^0 = \mathbf{P}^0 = \left(1, \frac{r-a}{b-a}, \frac{b-r}{b-a}\right) \quad \text{and} \quad \mathbf{P}_t^1 = \mathbf{P}^1 = \left(\frac{r}{b}, 1, 0\right), \quad (6.1.4)$$

which are not equivalent to  $\mathbf{P}$  but such that any convex combination

$$\mathbf{P}^\gamma = \gamma \mathbf{P}^0 + (1-\gamma) \mathbf{P}^1, \quad (6.1.5)$$

is. Any measure defined on  $\Omega$  and with respect to which the sequence  $\bar{S}$  is a  $\mathbb{F}$ -martingale is called a  $\mathbb{F}$ -martingale measure.

Note that the extremal measures  $\mathbf{P}_t^0$  and  $\mathbf{P}_t^1$  are independent of  $t$  and  $\mathcal{F}_t$ , so that, if  $T = 2$  and  $\omega = (\omega_1, \omega_2) \in \Omega$ ,

$$\mathbf{P}^\gamma(\omega_1^i, \omega_2^j) = \mathbf{P}^\gamma(\omega_1^i) \mathbf{P}^\gamma(\omega_2^j | \omega_1^i) = \mathbf{P}^\gamma(\omega_1^i) \mathbf{P}^\gamma(\omega_2^j)$$

and by induction

$$\mathbf{P}^\gamma(\omega_1^{i_1}, \dots, \omega_n^{i_n}) = \prod_{k=1}^n \mathbf{P}^\gamma(\omega_k^{i_k}).$$

If  $J_t = 2^t$  denotes the number of extremal points  $\mathbf{P}_t^j$  of the polyhedron of  $\mathbb{F}$ -martingale measures, for any  $j \in J_t$  there exists  $(\gamma_k^j)_{k \in \mathbf{N}_t} \in \{0, 1\}^t$  such that

$$\mathbf{P}_t^j = \bigotimes_{s \in \mathbf{N}_t} (\mathbf{P}^0)^{\gamma_s^j} (\mathbf{P}^1)^{1-\gamma_s^j} \quad (6.1.6)$$

For any  $t \in \mathbf{N}_T$ ,  $\mathcal{M}_t^{\mathbb{F}}$  denotes the set of  $\mathbb{F}$ -martingales measures on  $\mathbf{N}_t^0$ , whereas the set  $\mathcal{C}_t^{\mathbb{F}}$  composed of the convex combinations of extremal measures (6.1.6) matches with the convex subset of  $\mathcal{M}_t^{\mathbb{F}}$  formed by the  $\mathbb{F}$ -martingale measures equivalent to  $\mathbf{P}$ .

**Remark 6.1.3.** We can give interpretation to the definitions of the measures  $\mathbf{P}^0 = (\lambda^0, p^0, q^0)$  and  $\mathbf{P}^1 = (\lambda^1, p^1, q^1)$  introduced above. These stand for the "limit models". Actually when the probability parameters tend towards extreme values, we fall back on well-known models and complete markets.

Under  $\mathbf{P}^0$ , the process  $(\Delta N_t)_{t \in \mathbf{N}_T}$  is deterministic, constant equal to one. In other words, the process  $S$  jumps at each time so that it coincides with the price process defined in the Cox-Ross-Rubinstein model (4.1.7). To make sure, keep the parameters  $a, b, r$  and define the underlying Rademacher process  $X^{\text{bin}}$  by  $X^{\text{bin}} = \mathbf{1}_{\{X \neq 0\}} X$ . We can check this is well-defined since:

$$\mathbf{P}(X_t = 1) = p^0, \quad \mathbf{P}(X_t = -1) = q^0 \quad \text{and} \quad \mathbf{P}(X_t = 0) = 1 - p^0 - q^0,$$

so that the equalities  $\mathbf{P}(X_t^{\text{bin}} = 1) + \mathbf{P}(X_t^{\text{bin}} = -1) = 1$  and  $\mathbf{P}(X_t = 0) = 0$  hold for any  $t \in \mathbf{N}_T$ . The definition of  $(p^0, q^0) = ((r - a)/(b - a), (b - r)/(b - a))$  corresponds exactly to the risk-neutral measure in the binomial model.

Under  $\mathbf{P}^1$ , the process  $(V_t)_{t \in \mathbf{N}_T}$  is deterministic, constant equal to one, so that the compound process  $(Z_t)_{t \in \mathbf{N}_T}$  coincides with a single geometric process of intensity  $\lambda$ . Once again, the resulting market is complete and the risk-neutral measure is obtained by letting  $\hat{\lambda} = r/b$ .

## 6.2 Enlargement of filtration

As mentioned in chapter 4, trying to quantify the additional information provided by the known in advance of a  $\mathcal{F}_T$ -measurable random variable  $G$  leads to tackle with an enlargement of the initial filtration. In a discrete setting, it seems to boil down to the exploitation of the *Doob decomposition* (see Section 4.1 and [20])

In what follows, the random variable  $G$  is supposed to take its values in a measurable space  $(\Gamma, \mathcal{G})$  where  $\Gamma$  is a countable set. We define a particular process  $p^G$  which is of key importance in many ways. First, it naturally appears in the Doob decomposition and enables to connect the eponymous *information drift* (as reminiscence with the continuous case) with the variable  $G$ . Besides, to study the martingales with respect to the enlarged filtration, since, as it will be proved, the  $t$ -th term of the process  $1/p^G$  preserves the martingale property of any  $(\mathbb{F}, \mathbf{P})$ -martingale on  $\mathbf{N}_t^0$ . As such, optimizing the insider's portfolio (see subsection 6.4.1) requires to determine the crucial set of  $\mathbb{G}$ -martingale measures on  $\mathbf{N}_t^0$  equivalent to  $\mathbf{P}$ .

Up to now, let  $\mathbb{F}^0$  (resp.  $\mathbb{G}^0$ ) be the filtrations  $\mathbb{F}^0 = (\mathcal{F}_t)_{t \in \mathbf{N}_T^0}$  (resp.  $\mathbb{G}^0 = (\mathcal{G}_t)_{t \in \mathbf{N}_T^0}$ ), where the  $\sigma$ -algebras  $(\mathcal{F}_t)_{t \in \mathbf{N}_T^0}$  are defined in (5.1.4) (resp.  $\mathcal{G}_t = \mathcal{F}_t \vee \mathbb{G}$ ) and which distribution given  $\mathcal{F}_t$  for some  $t \in \mathbf{N}_T$  is given by a family of  $\mathcal{F}_t$ -measurable random variables  $\{\mathbf{P}(G = c | \mathcal{F}_t), c \in \Gamma\}$ .

### 6.2.1 The information drift process

As the set  $\Gamma$  is countable, the absolute continuity of law type condition (inescapable in the continuous case) is de facto verified in our discrete context. Indeed, any set  $C \in \mathcal{G}$  is of the

form  $C = \bigcup_{c \in C} \{G = c\}$  and, for any  $t \in \mathbf{N}_{T-1}^0$ ,

$$\mathbf{P}(G \in C | \mathcal{F}_t) = \sum_{c \in C} \mathbf{P}(G = c | \mathcal{F}_t) = \sum_{c \in C} \frac{\mathbf{P}(G = c | \mathcal{F}_t)}{\mathbf{P}(G = c)} \mathbf{P}(G = c) = \mathbf{E} [p_t^G \mathbf{1}_A],$$

where the random variable  $p_t^G$  is defined by

$$p_t^G(\omega) = \frac{\mathbf{P}(G \in \cdot | \mathcal{F}_t)(\omega)}{\mathbf{P}(G \in \cdot)} \quad \text{such that} \quad p_t^c(\omega) = \frac{\mathbf{P}(G = c | \mathcal{F}_t)(\omega)}{\mathbf{P}(G = c)}, \quad (6.2.1)$$

for any  $\omega \in \Omega, c \in \Gamma$ . Define also the process  $(\nu_t^G)_{t \in \mathbf{N}_T}$  by

$$\nu_t^G = \mathbf{P}(G \in \cdot | \mathcal{F}_t) = \mathbf{P}(G \in \cdot) p_t^G \quad (6.2.2)$$

The combination of (6.2.2) and (4.1.9) (see section 2.2 in [20]) applied to the  $(\mathbf{P}, \mathbb{F})$ -martingale  $(\bar{Y})_{t \in \mathbf{N}_T}$  ensures that the process  $(Y_t^G)_{t \in \mathbf{N}_T}$  defined by

$$\bar{Y}_t^G = \bar{Y}_t - \sum_{s=1}^t \frac{\langle Y, p_t^c \rangle_s |_{c=G}}{p_{s-1}^G} = Y_t - \mu_t^G \quad (6.2.3)$$

is a  $(\mathbf{P}, \mathbb{G})$ -martingale. The process  $\mu^G$  thus defined is called the *information drift*, namely the drift to absorb in passing to the insider paradigm in order to preserve the martingale property of the initial compound geometric process. As designed by Peter Imkeller in [68], we can traduct its connection to the random variable  $G$  thanks to the Malliavin derivative.

**Theorem 6.2.1 (★).** *The information drift  $\mu^G$  defined in (6.2.3) can be written as*

$$\mu_t^G = \sum_{k \in E} \sum_{\ell \in E} \frac{a_{k,\ell} \mathbf{E} [D_{t,\ell} p_t^c] |_{c=G}}{p_{t-1}^G}$$

for any  $t \in \mathbf{N}_T$ , where the family  $\{a_{k,\ell}, (k, \ell) \in E^2\}$  is defined by  $a_{k,\ell} = \mathbf{E} [\Delta Z_{t,k} \Delta R_{t,\ell}]$ , i.e.

$$a_{1,1} = \lambda p(1 - \lambda p), \quad a_{1,-1} = 0, \quad a_{-1,-1} = \lambda^2 p q \quad \text{and} \quad a_{-1,1} = \frac{\lambda q(1 - \lambda)}{1 - \lambda p}.$$

The following result is incomplete. We chose to present it anyway for the interest it could arouse in possible future investigations. Actually, it *would seem* possible to interpret  $\mu^G$  in terms of drift on the initial process  $X$  at the cost of an additional assumption:

**Assumption.** For any  $c \in \Gamma$ , assume the existence of a couple  $(\tilde{\lambda}^c, \tilde{p}^c) \in (0, 1)^2$  such that

$$p_t^c = \frac{d\tilde{\mathbf{P}}_{\tilde{\lambda}, \tilde{p}}}{d\mathbf{P}} \Big|_{\mathcal{F}_t}$$

that is, satisfying:

$$\frac{p_t^c}{p_{t-1}^c} = 1 + \frac{1 - \tilde{\lambda}^c}{1 - \lambda} \left( \psi(\Delta N_t, W_t) - \frac{\lambda - \tilde{\lambda}^c}{1 - \tilde{\lambda}^c} \right). \quad (6.2.4)$$

Then by Girsanov's theorem 5.4.6, the process  $X$  defined by

$$X_t = \sum_{s=1}^{N_t} V_s$$

is geometric compound process of parameters  $(\tilde{\lambda}^G, \tilde{p}^G)$  under the probability measure  $\tilde{\mathbf{P}}_{\tilde{\lambda}, \tilde{p}}$ . Recall moreover (see [20]):

**Theorem 6.2.2** (Blanchet-Scalliet, Jeanblanc, Roméro). *Let  $(X_t)_{t \in \mathbf{N}_T}$  be a  $(\mathbf{P}, \mathbb{F})$ -martingale and  $\mathbf{Q}$  a probability measure supposed to be equivalent to  $\mathbf{P}$  on  $\mathcal{F}_t$  for any  $t \in \mathbf{N}_T$ .*

*Let  $L_t = \frac{d\mathbf{Q}}{d\mathbf{P}} \Big|_{\mathcal{F}_t}$ . Then the process  $X^{\mathbf{Q}}$  defined by*

$$X_t^{\mathbf{Q}} = X_t - \sum_{s=1}^t \frac{\Delta \langle X, L \rangle_s}{L_{s-1}^{\mathbf{G}}},$$

*is a  $(\mathbf{Q}, \mathbb{F})$ -martingale.*

The application of Theorem 6.2.2 to  $L = p^{\mathbf{G}}$  provides:

**Theorem 6.2.3** (★). *Assume Assumption 6.2.4 holds. Let  $X$  be a geometric compound process of parameters  $(\lambda, p)$ . Then the process  $X^{\mathbf{G}}$*

$$X_t^{\mathbf{G}} = X_t - \sum_{s=1}^t \frac{\Delta \langle X, p^c \rangle_s |_{c=\mathbf{G}}}{p_{s-1}^{\mathbf{G}}}$$

*is a geometric compound process of parameters  $(\lambda^{\mathbf{G}}, \tilde{p}^{\mathbf{G}})$  where  $\tilde{p}^c$  is defined by (6.2.3) for any  $c \in \Gamma$ .*

**Remark 6.2.4.** This result sounds interesting because it does the counterpart with the Brownian case. However, it remains incomplete in its current form; indeed we still don't know what meaning to give to a compound geometric process of *random* parameters  $(\lambda^{\mathbf{G}}, p^{\mathbf{G}})$ . We could imagine associating to the process  $X$  a sequence of parameters  $(\lambda_t^{\mathbf{G}}, p_t^{\mathbf{G}})_{t \in \mathbf{N}_T}$  defined in terms of conditional probabilities  $\mathbf{P}(G \in \cdot | \mathcal{F}_t)$  and in a Markov chain-like way... To be continued !

## 6.2.2 The martingale preserving measure

In this section, we focus on the  $\mathbb{F}$ -adapted process  $1/p^{\mathbf{G}}$ , of key importance afterwards. This is well defined; indeed, resulting from the countability of  $\Gamma$ , for any  $(t, c) \in \mathbf{N}_{T-1}^0 \times \Gamma$ , the random variable  $p_t^c$  is not null  $\mathbf{P}$ -a.s.

Note that for any  $t \in \mathbf{N}_T$ , the  $\sigma$ -algebra  $\mathcal{G}_t$  is generated by the set

$$\{A \cap B ; A \in \mathcal{F}_t, B \in \mathcal{E}\}.$$

**Theorem 6.2.5** (★). *1. The process  $1/p^{\mathbf{G}}$  is a  $(\mathbf{P}, \mathbb{G}^0)$  martingale.*

*2. For  $t \in \mathbf{N}_T^0$ , the  $\sigma$ -algebras  $\mathcal{F}_t$  and  $\sigma(\mathbf{G})$  are independent under the probability measure defined for any  $A \in \mathcal{G}_t$  by*

$$\mathbf{Q}_t(A) = \mathbf{E} \left[ \frac{1}{p_t^{\mathbf{G}}} \mathbf{1}_A \right].$$

*3. For any  $t \in \mathbf{N}_{T-1}^0$ , the probability measure  $\mathbf{Q}_t$  coincides with  $\mathbf{P}$  on  $\mathcal{F}_t$ .*

**Remark 6.2.6.** The process  $(\mathbf{Q}_t)_{t \in \mathbf{N}_T^0}$  thus defined is called *the martingale preserving measure*.

It is thus possible to establish the following result.

**Theorem 6.2.7** (★). *For fixed  $t \in \mathbf{N}_{T-1}^0$ , any  $(\mathbf{P}, \mathbb{F})$ -martingale on  $\mathbf{N}_t^0$  is a  $(\mathbf{Q}_t, \mathbb{G})$ -martingale and also a  $(\mathbf{Q}_t, \mathbb{F})$ -martingale on  $\mathbf{N}_t^0$ .*

**Remark 6.2.8.** The process  $1/p^{\mathbf{G}}$  defines a probability measure, which density is absolutely continuous with respect to  $\mathbf{P}$ .

### 6.3 Hedging formula in a ternary model

The value of the portfolio at time  $t \in \mathbf{N}_T^0$  is given by the random variable

$$V_t = \alpha_t A_t + \varphi_t S_t,$$

where  $(\alpha_t, \varphi_t)_{t \in \mathbf{N}_T}$  is a couple of predictable processes modeling respectively the amounts of riskless and risky assets held in the portfolio, and its discounted value by  $\bar{V}_0 = V_0$  and

$$\bar{V}_t = \bar{V}_0 \prod_{s=1}^t \frac{1}{1+r_s} V_s; \quad t \in \mathbf{N}_T.$$

The aim of this subsection is to exhibit an hedging formula, i.e. given a nonnegative  $\mathcal{F}_T$ -measurable random variable  $F$  (called *claim*), to determine a portfolio modelised by  $(\alpha_t, \varphi_t)_{t \in \mathbf{N}_T^0}$  such that

$$V_0 > 0, \quad V_t \geq 0 \quad (t \in \mathbf{N}_{T-1}), \quad \text{and} \quad V_T = F.$$

In an incomplete market, all claims are not reachable; they have an *intrinsic risk*. Face to the impossibility to perform a perfect hedge in the general case, we can only hope to reduce the a priori risk to this minimal component. The question of hedging in an incomplete market has been widely investigated for years (see for instance [34],[53] in continuous time, [130] in discrete time). The ternary model, as the trinomial one, is not complete; we choose to deal with the optimization problem

$$\min_{(x, \varphi) \in \mathbf{R}_+^* \times \Theta} \mathbf{E} [(F - x - V_T(\varphi))^2], \quad (6.3.1)$$

where  $x$  is the initial capital and  $\mathcal{S}^{\mathbb{F}}$  is the set of  $\mathbb{F}$ -predictable admissible strategies. The *mean-variance tradeoff* process  $(K_t)_{t \in \mathbf{N}_T}$  is defined by

$$K_t = \sum_{s=1}^t \frac{(\mathbf{E}[(\Delta S_s) | \mathcal{F}_{s-1}])^2}{\text{var}[\Delta S_s | \mathcal{F}_{s-1}]}; \quad t \in \mathbf{N}_T.$$

Define also the discrete analogue of the *minimal martingale measure* (see Föllmer and Schweizer [53]), i.e. the signed measure  $\hat{\mathbf{P}}$  defined on  $(\Omega, \mathbb{F})$  such that

$$\frac{d\hat{\mathbf{P}}}{d\mathbf{P}} = \prod_{t=1}^{\mathbf{N}_T} \frac{1 - \theta_t \Delta S_t}{1 - \theta_t \mathbf{E}[\Delta S_t | \mathcal{F}_{t-1}]}, \quad (6.3.2)$$

where, for any  $t \in \mathbf{N}_T$ ,

$$\theta_t = \frac{\mathbf{E}[\Delta S_t | \mathcal{F}_{t-1}]}{\mathbf{E}[(\Delta S_t)^2 | \mathcal{F}_{t-1}]}.$$

Last, consider the Kunita-Watanabe decomposition of  $F$  (see Metivier [87] or Schweizer [130]) i.e. the unique couple of processes  $(\xi^{\mathbb{F}}, L^{\mathbb{F}})$  where  $\xi^{\mathbb{F}}$  is a square-integrable admissible strategy and  $L^{\mathbb{F}}$  a  $\mathbb{F}$ -martingale, strongly orthogonal to  $S$ , with null initial value such that

$$F = F_0 + \sum_{t \in \mathbf{N}_T} \xi_t^{\mathbb{F}} \Delta S_t + L_T^{\mathbb{F}} \quad \mathbf{P}\text{-a.s.}$$

Within previous notations, Martin Schweizer gives an expression of the quadratic-loss minimizing strategy.

**Theorem 6.3.1** (Schweizer, 1992). *Provided  $(K_t)_{t \in \mathbf{N}_T}$  is deterministic, the solution of (6.3.1) is given by*

$$x^* = \widehat{\mathbf{E}}[\mathbf{F}] \quad \text{and} \quad \varphi_t^* = \xi_t^{\mathbf{F}} + \frac{\mathbf{E}[\Delta S_t | \mathcal{F}_{t-1}]}{\mathbf{E}[(\Delta S_t)^2 | \mathcal{F}_{t-1}]} (\widehat{\mathbf{E}}[\mathbf{F} | \mathcal{F}_t] - x^* - V_{t-1}(\varphi^*)) \quad (6.3.3)$$

and  $\widehat{\mathbf{P}}$  is the minimal martingale measure defined by (6.3.2). Moreover, the quota of the riskless asset  $(A_t)_{t \in \mathbf{N}_T}$  is given by  $\alpha_0 = (1+r)^{-T} \widehat{\mathbf{E}}[\mathbf{F}] / S_0$  and for any  $t \in \mathbf{N}_T$ ,

$$\alpha_t = \alpha_{t-1} - \frac{(\varphi_t - \varphi_{t-1})S_{t-1}}{A_{t-1}}.$$

**Remark 6.3.2.** If the contingent claim  $\mathbf{F}$  is reachable, then  $\varphi^* = \xi^{\mathbf{F}}$ . The term  $\xi^{\mathbf{F}}$  in (6.3.3) can be interpreted as a pure hedging demand, whereas the second one can be viewed as a demand for mean-variance purposes (see Schweizer [129]).

We slot these results to our formalism to solve (6.3.3) in the ternary model.

**Lemma 6.3.3.** *The mean-variance tradeoff process of the ternary model is deterministic.*

*Proof.* For any  $t \in \mathbf{N}_T$ ,

$$\frac{(\mathbf{E}[\Delta S_t | \mathcal{F}_{t-1}])^2}{\text{var}[\Delta S_t | \mathcal{F}_{t-1}]} = \frac{(\mathbf{E}[1 + \eta_t \Delta N_t | \mathcal{F}_{t-1}])^2}{\text{var}[1 + \eta_t \Delta N_t | \mathcal{F}_{t-1}]} = \frac{(1 + \lambda(2p-1))^2}{\lambda - \lambda^2(2p-1)^2}$$

is a deterministic constant. □

**Lemma 6.3.4** ((★) Kunita-Watanabe decomposition in ternary model). *For any claim  $\mathbf{F} \in L^2(\Omega)$  there exist a square-integrable admissible strategy  $\xi^{\mathbf{F}}$  and a  $\mathbb{F}$ -martingale  $L^{\mathbf{F}}$ , strongly orthogonal to  $S$ , with null initial value such that*

$$\mathbf{F} = F_0 + \sum_{t \in \mathbf{N}_T} \xi_t^{\mathbf{F}} \Delta S_t + L_T^{\mathbf{F}} \quad \mathbf{P}\text{-a.s.}$$

Moreover, for any  $t \in \mathbf{N}_T$ ,

$$\xi_t^{\mathbf{F}} = \frac{\sum_{k \in \mathbf{E}} c_k \widehat{\mathbf{E}}[D_{t,k} \mathbf{F} | \mathcal{F}_{t-1}]}{\|c\|_2^2},$$

and

$$L_t^{\mathbf{F}} = \sum_{k \in \mathbf{E}} c_k (\widehat{\mathbf{E}}[D_{t,k} \mathbf{F} | \mathcal{F}_{t-1}] - \xi_t),$$

where  $c = (c_1, c_{-1}) \in \mathbf{R}^2$  with

$$c_1 = (b + a\rho)\kappa_1 \quad \text{and} \quad c_{-1} = -a\kappa_{-1}.$$

Define the minimal martingale measure  $\widehat{\mathbf{P}}$

$$\frac{d\widehat{\mathbf{P}}}{d\mathbf{P}} = \prod_{t=1}^{\mathbf{N}_T} \frac{1 - \theta \Delta S_t}{1 - \theta \mathbf{E}[\Delta S_t | \mathcal{F}_{t-1}]}, \quad (6.3.4)$$

where  $\theta = (1 + \lambda(bp - aq)) / \|c\|_2^2$ .

**Theorem 6.3.5** (★ Loss quadratic minimizing strategy in ternary model). *Let  $\widehat{\mathbf{P}}$  be the minimal martingale measure defined by (6.3.4) and let a claim  $F \in L^2(\Omega, \mathbb{F})$ . The quadratic loss minimizing hedge  $(x^*, \varphi^*)$  is given by*

$$x^* = \widehat{\mathbf{E}}[F] \text{ and } \varphi_t^* = \xi_t^F + \theta(\widehat{\mathbf{E}}[F|\mathcal{F}_t] - x^* - V_{t-1}(\varphi^*))$$

where  $\xi^F \in \Theta$  is given by the Kunita-Watanabe decomposition

*Proof.* Since the mean-variance process is deterministic by Lemma 6.3.3, it suffices to incorporate the result of Lemma 6.3.4 to Theorem 6.3.1. The process  $(\alpha_t)_{t \in \mathbf{N}_T}$  is defined by the self-financing condition (6.5.1).  $\square$

## 6.4 Additional utility of the insider in a ternary model

### 6.4.1 Theoretical results

Given an economic agent (resp. an insider) disposing of  $x$  ( $x > 0$ ) euros at date  $t = 0$  (initial budget constraint), we want to determine the optimal admissible strategy allowing to maximize his portfolio value at time  $t \in \mathbf{N}_T$ . This leads to consider the optimization problem from the agent's point of view:

$$\Phi_t^{\text{ag}}(x) = \max_{\psi \in \mathbb{F}\text{-portfolio}} \mathbf{E}[u(V_{x,t}(\psi))], \quad (6.4.1)$$

and respectively from the insider's one:

$$\Phi_t^{\text{ins}}(x) = \max_{\psi \in \mathbb{G}\text{-portfolio}} \mathbf{E}[u(V_{x,t}(\psi))] \quad (6.4.2)$$

where  $u = \log$ .

### Utility of the economic agent and the insider before the deadline $T$

In this subsection we tackle problems (6.4.1) and (6.4.2) when  $t \in \mathbf{N}_{T-1}$  ( $T \in \mathbf{N}^*$ ). Based on the equivalence in law of the trinomial model and ours (see remark 6.1.1), there is a bijection between  $\mathcal{M}^{\mathbb{F}}$ , the convex set of  $\mathbb{F}$ -martingale measures equivalent to  $\mathbf{P}$  in our model, and the one existing in the trinomial. Indeed, for  $T = 1$ , the measures  $\mathbf{P}^\gamma = (\lambda_\gamma, p_\gamma, q_\gamma)$  (defined by 6.1.5) and  $\mathbf{P}^{\gamma, \text{Tri}} = (p_\gamma^{\text{Tri}}, q_\gamma^{\text{Tri}}, s_\gamma^{\text{Tri}})$  ( $\gamma \in [0, 1]$ ) are equivalent by letting  $\lambda_\gamma = p_\gamma^{\text{Tri}} + q_\gamma^{\text{Tri}}$  and  $p_\gamma = \lambda_\gamma^{-1} p_\gamma^{\text{Tri}}$ . For a given  $T \in \mathbf{N}^*$ , the set  $\mathcal{M}^{\mathbb{F}}$  is the polyhedron characterized by its extremal points; these are constructed by tensorization of the measures (independent of  $t$ )

$$\mathbf{P}_t^0 = \mathbf{P}^0 = \left(1, \frac{r-a}{b-a}, \frac{b-r}{b-a}\right) \quad \text{and} \quad \mathbf{P}_t^1 = \mathbf{P}^1 = \left(\frac{r}{b}, 1, 0\right).$$

By induction, the convex set of  $\mathbb{F}$ -martingale measures equivalent to  $\mathbf{P}$  in our model, and the one existing in the trinomial are equivalent. The following portfolio optimization result is directly deduced. Its proof which turns out to be the rewrite of that of [42] for the trinomial model (see more details in section 4.2) and won't be detailed.

**Theorem 6.4.1** (★). *There exists unique optimizers  $\widehat{V}_{x,t}^{\mathbb{F}}$  and  $\widehat{M}_{\widehat{y}_x,t}$  of the problems (4.2.6) and (4.2.11) such that  $\widehat{M}_{\widehat{y}_x} \in \mathcal{C}_t^{\mathbb{F}}$  and*

$$\widehat{V}_{x,t}^{\mathbb{F}} = I \left( \widehat{y}_x \frac{d\widehat{M}_{\widehat{y}_x,t}}{d\mathbf{P}} \right).$$



**Remark 6.4.2.** Some of the computations rely on the following Theorem 3.2.1 in [42]: within the previous notations, the optimizers  $\widehat{V}_{x,t}$  and  $\widehat{\mathbf{M}}_{\widehat{y}_{x,t}}$  exist and are unique. The measure  $\widehat{\mathbf{M}}_{\widehat{y}_x}$  belongs to  $\mathcal{C}_t^{\mathbb{F}}$  and satisfies the remarkable relation:

$$\frac{d\widehat{\mathbf{M}}_{\widehat{y}_{x,t}}}{d\mathbf{P}} = \frac{1}{\widehat{y}_x} u'(\widehat{V}_{x,t}) \quad \text{where} \quad \widehat{y}_x = (\Phi_t^{\text{ag}})'(x).$$

Further computations will rely on. Denote by  $\{\lambda_j \mid j \in J_t\}$  the coordinates family of  $\widehat{\mathbf{M}}_{\widehat{y}_{x,t}}$  i.e. such that

$$\widehat{\mathbf{M}}_{\widehat{y}_{x,t}} = \sum_{j \in J_t} \lambda_j \mathbf{P}_t^j. \quad (6.4.3)$$

The optimization problem from the insider point of view (6.4.2) seems to be solved by transposing that of the agent to the enriched filtration  $\mathbb{G}$ . That lies on the possibility to rewrite (6.4.2) as a primal problem as in (4.2.9). With this edict in mind, we have to identify the set of measures which respect to which the sequence  $(\bar{S}_s)_{s \in \mathbf{N}_t}$  is a  $(\cdot, \mathbb{G})$ -martingale, namely the  $\mathbb{G}$ -martingale measures. Denote thus by  $\mathcal{M}_t^{\mathbb{G}}$  and  $\mathcal{C}_t^{\mathbb{G}}$  ( $t \in \mathbf{N}_T$ ), respectively the set of  $\mathbb{G}$ -martingale measures on  $\mathbf{N}_t^0$ , and the convex subset of  $\mathcal{M}_t^{\mathbb{G}}$  formed by the  $\mathbb{G}$ -martingale measures equivalent to  $\mathbf{P}$  on  $\mathbf{N}_t^0$ .

In the first subsection we established that for any  $t \in \mathbf{N}_{T-1}^0$ , any  $(\mathbf{P}, \mathbb{F}^0)$ -martingale on  $\mathbf{N}_t^0$  is a  $(\mathbf{Q}_t, \mathbb{G}^0)$ -martingale where  $\mathbf{Q}_t$  is the measure defined by

$$\mathbf{Q}_t(A) = \int_A \frac{1}{p_t^{\mathbb{G}}} d\mathbf{P} ; A \in \mathcal{G}_t.$$

For the sake of simplicity let  $\mathbf{Q} = \mathbf{Q}_T$ . The aim of this subsection is to describe  $\mathcal{M}_t^{\mathbb{G}}$ , the set of  $\mathbb{G}$ -martingale measures on  $\mathbf{N}_t^0$ . As seen in subsection 4.2, the set of  $\mathbb{F}$ -martingale measures on  $\mathbf{N}_t^0$  consists of the convex combinations of  $\{\mathbf{P}_t^j, j \in J_t\}$ . For any  $j \in J_t$ , define the sequence of measures  $(\mathbf{Q}_t^j)_{t \in \mathbf{N}_{T-1}}$  by

$$\mathbf{Q}_t^j(A) = \int_A \frac{1}{p_t^{\mathbb{G}}} d\mathbf{P}^j ; A \in \mathcal{G}_t,$$

where we let  $\mathbf{P}^j = \mathbf{P}_T^j$ . Furthermore, follows from theorem 6.2.7 that a  $(\mathbf{P}^j, \mathbb{F})$ -martingale on  $\mathbf{N}_t^0$  is a  $(\mathbf{Q}_t^j, \mathbb{G})$ -martingale on  $\mathbf{N}_t^0$ . In particular, the discounted prices process  $\bar{S}$  is a  $(\mathbf{Q}_t^j, \mathbb{G})$ -martingale on  $\mathbf{N}_t^0$ ; so does (on  $\mathbf{N}_t^0$ )  $V_{x,\cdot}^{\mathbb{G}}(\psi)$  viewed as the  $\mathbb{G}$ -martingale transform

$$V_{x,t}^{\mathbb{G}}(\psi) = x + \sum_{s=1}^t \varphi_s^{\mathbb{G}} \Delta \bar{S}_s,$$

where  $\psi = (\alpha, \varphi^{\mathbb{G}})$  belongs to  $\mathcal{S}^{\mathbb{G}}$ , the set of  $\mathbb{G}$ -predictable admissible strategies. Note besides that

$$\mathbf{E}_{\mathbf{Q}_t^j}[V_{x,t}^{\mathbb{G}}(\psi)] = x ; j \in J_t, \quad (6.4.4)$$

so that we can address the optimization problem as a constrained one. Define for  $(x, t) \in \mathbf{R}_+^* \times \mathbf{N}_{T-1}^0$ ,

$$\Lambda_{x,t}^{\mathbb{G}} = \left\{ V \in \mathcal{S}^{\mathbb{G}} : \mathbf{E}_{\mathbf{Q}_t^j}[V] \leq x, \forall j \in J_t \right\}.$$

As a consequence once again of the surreplication theorem, (6.4.2) can be rewritten as the constrained optimization problem

$$\max_{V \in \Lambda_{x,t}^{\mathbb{G}}} \mathbf{E}[u(V)] \quad \text{under the constraints} \quad \mathbf{E}_{\mathbf{Q}_t^j}[V] \leq x ; \forall j \in J_t. \quad (6.4.5)$$

**Theorem 6.4.3** (★). *There exists unique optimizers  $\widehat{V}_{x,t}^{\mathbb{G}}$  and  $\widehat{\mathbf{M}}_{\widehat{y}_x^{\mathbb{G}},t}^{\mathbb{G}}$  of the problems (4.2.6) and (4.2.11) such that  $\widehat{\mathbf{M}}_{\widehat{y}_x^{\mathbb{G}},t}^{\mathbb{G}} \in \mathcal{C}_t^{\mathbb{G}}$  and*

$$\widehat{V}_{x,t}^{\mathbb{G}} = I\left(\widehat{y}_x^{\mathbb{G}} \frac{d\widehat{\mathbf{M}}_{\widehat{y}_x^{\mathbb{G}},t}^{\mathbb{G}}}{d\mathbf{P}}\right).$$

Moreover,

$$\widehat{\mathbf{M}}_{\widehat{y}_x^{\mathbb{G}},t}^{\mathbb{G}} = \sum_{j \in J_t} \lambda_j \mathbf{Q}_t^j, \quad (6.4.6)$$

and  $\{\lambda_j, j \in J_t\}$  is exactly the family defined by (6.4.3).

### Additional utility before the deadline

The insider's additional expected logarithmic utility up to time  $t \in \mathbf{N}_{T-1}$  is defined by

$$\mathcal{U}_t(x) = \max_{\psi \in \mathbb{G}\text{-portfolio}} \mathbf{E}[u(V_{x,t}(\psi))] - \max_{\psi \in \mathbb{F}\text{-portfolio}} \mathbf{E}[u(V_{x,t}(\psi))].$$

Given two probabilities measures defined on the same space  $(\Omega, \mathcal{F})$ ,  $\mathfrak{D}_{\mathcal{F}}(\mathbf{P}||\mathbf{Q})$  designates the relative entropy of  $\mathbf{P}$  with respect to  $\mathbf{Q}$  on  $\mathcal{F}$  and is defined by

$$\mathfrak{D}_{\mathcal{F}}(\mathbf{P}||\mathbf{Q}) = \begin{cases} \mathbf{E}\left[\log\left(\frac{d\mathbf{P}}{d\mathbf{Q}}\right)\Big|_{\mathcal{F}}\right] & \text{if } \mathbf{P} \ll \mathbf{Q} \text{ on } \mathcal{F}, \\ +\infty & \text{otherwise.} \end{cases}$$

**Theorem 6.4.4** (★). *The insider's additional expected logarithmic utility up to time  $t \in \mathbf{N}_{T-1}$  is given by*

$$\mathcal{U}_t = \mathfrak{D}_{\mathcal{G}_t}(\mathbf{P}||\mathbf{Q}_t).$$

**Remark 6.4.5.** From a philosophical point of view, it is not very surprising to see the entropy appear when quantifying an additional... information.

**Remark 6.4.6.** Thus we recover the result already established by Jürgen Amendinger, Peter Imkeller and Martin Schweizer ([7]) in the continuous case; the additional expected logarithmic utility of the insider can be expressed in terms of relative entropy. They get more: by letting  $t$  go to  $T$ , they underscore that the result still holds at the deadline  $T$ . Obviously we cannot transpose this in a discrete setting. Since there is absolutely no reason why  $(\mathbf{P}, \mathbb{F})$ -martingale  $\mathbf{N}_t^0$  should be a  $(\mathbf{Q}, \mathbb{G}^0)$ -martingale, we need to address the problem at the deadline  $T$  from another perspective. We chose to dedicate the next section to this case.

### Utility of the insider at expiry

In order to remove the difficulty previously evokated (see remark 6.4.6), we link the optimal strategy utility final value with this at the instant just before.

**Lemma 6.4.7** (★). *For any  $x \in \mathbf{R}_+^*$  define  $\Phi_T^{\text{ins}}(x)$  by considering (6.4.2) at the deadline  $T$ . Then,*

$$\Phi_T^{\text{ins}}(x) = \max_{\psi \in \mathbb{G}\text{-portfolio}} \mathbf{E}[u(V_{\tilde{x},1}(\psi))] = \Phi_1^{\text{ins}}(\tilde{x}),$$

where  $\tilde{x} = \Phi_{T-1}^{\text{ins}}(x)$ .

### 6.4.2 Comparison with the binomial and the Black-Scholes models

As mentioned in Remark 6.4.6, the formula giving the additional utility of the insider before the deadline is exactly the same in the continuous case (Black-Scholes model) and in the discrete ones (binomial, trinomial or our ternary model). As a matter of fact, this lies on the enlargement of filtrations tools that are comparable in continuous and discrete time. The study of the insider problem from the additional expected utility point of view is similar in all frameworks.

The substantial difference we have met between our model and the classical ones (Black-Scholes and binomial one) concern the computation of the argmax admissible strategy appearing in (6.4.2). This is not strictly speaking linked with the insider matter, but rather with the impossibility to perform perfect hedges, because of the incompleteness of our ternary model.

### 6.4.3 Computations in the case $G = \mathbf{1}_{\{S_T \in [0, S_0]\}}$

In this last subsection we chose to investigate a specific case:  $G = \mathbf{1}_{\{S_T \in [0, S_0]\}}$ . Morally, this means the insider knows whether the stock price will have increased or not at the end of the trading period. Let  $x \in \mathbf{R}_+^*$  be the initial budget. As described in subsection 6.4.1 to cover all cases we need to compute:

- $\Phi_1^{\text{ag}}(x)$  and  $\Phi_1^{\text{ins}}(x)$ ,
- $\Phi_T^{\text{ag}}(x)$  (for  $T \geq 2$ ),
- $\mathcal{U}_t(x)$  for any  $t \in \mathbf{N}_{T-1}$ .

**Remark 6.4.8.** In fact the computation of  $\Phi_T^{\text{ag}}(x)$  does not require to invoke Lemma 6.4.7 and can be done directly.

#### Computations of $\Phi_1^{\text{ag}}(x)$ and $\Phi_1^{\text{ins}}(x)$

We provide a method to compute explicitly  $\Phi_1^{\text{ins}}(x)$ , namely to solve the optimization problem when  $T = 1$ . As evoked in subsection 6.1.1, the set  $\mathcal{M}_1^{\mathbb{F}}$  of  $\mathbb{F}$ -martingale measures equivalent to  $\mathbf{P}$  consists of the convex combination of the form

$$\mathbf{P}^\gamma = \gamma \mathbf{P}^0 + (1 - \gamma) \mathbf{P}^1,$$

where the two extremal measures are defined by (6.1.4). In other terms,

$$\mathbf{P}^\gamma = \left( \gamma + \frac{(1 - \gamma)r}{b}, 1 - \gamma + \frac{\gamma(r - a)}{b - a}, \frac{\gamma(b - r)}{b - a} \right) = (1 - \lambda_\gamma, \lambda_\gamma p_\gamma, \lambda_\gamma (1 - p_\gamma))$$

In that specific case, it seems easier to directly solve the primal problem, i.e. to compute:

$$\max_{V \in \Lambda_{x,1}} \mathbf{E}[\log(V)] \quad \text{under the constraints} \quad \mathbf{E}_{\mathbf{P}^j}[V] \leq x; \quad \forall j \in \{0, 1\},$$

where

$$V = x + \varphi \Delta S_1 = x + \varphi (b \mathbf{1}_{\{(\Delta N_1, W_1) = (1, 1)\}} + a \mathbf{1}_{\{(\Delta N_1, W_1) = (1, -1)\}}),$$

and  $\varphi$  is a  $\mathcal{F}_0$ -measurable random variable (since  $T = 1$ ). The ternary model described in subsection 6.1 can be thus treated as a stochastic volatility model: the closer  $\lambda_\gamma$  is to 1 the

more volatile is the model. In the extreme case (when  $\lambda_\gamma$  is equal to one), the model coincides with the Cox-Ross-Rubinstein one for which there exists a unique optimizer (see proposition 3.3.2. in [42]). On the other hand, when  $\lambda_\gamma$  is near to zero, the choice of the strategy  $\varphi$  has little influence on  $\varphi \Delta S_1$  and thus on the computation of  $\mathbf{E}[\log(V)]$ . This leads us to speculate that the optimal strategy in the ternary model coincides with that of the binomial model. To be convinced that the maxima are realized for the same strategy, note that

$$\mathbf{E}_\gamma \left[ \log(V(\varphi)) \right] = \lambda_\gamma \mathbf{E}_0 \left[ \log(V_1^{\text{bin}}(\varphi)) \right] + (1 - \lambda_\gamma) \log(x), \quad (6.4.7)$$

where  $\mathbf{E}_0$  designates the expectation under  $\mathbf{P}_0$ , namely the unique risk-neutral probability measure in the Cox-Ross-Rubinstein model, and  $V_1^{\text{bin}}(\varphi)$  designates the value of the strategy  $\varphi$  at time 1 i.e.

$$V_1^{\text{bin}}(\varphi) = x + \varphi \Delta S_1^{\text{bin}}.$$

The sequence of prices  $(S_t^{\text{bin}})_{t \in \mathbf{N}_T}$  is defined in the "associated" binomial model (by letting  $\lambda = 1$ ) by:

$$(\mathcal{S}^{\text{bin}}) \begin{cases} S_{t+1}^{\text{bin}} = (1+b)S_t^{\text{bin}} & \text{with probability } p \\ S_{t+1}^{\text{bin}} = (1+a)S_t^{\text{bin}} & \text{with probability } 1-p \end{cases}$$

In the latter one, picking up the results established by Freddy Delbaen et Walter Schachermayer ([42] example 3.3.2.) to solve the optimization problem from the agent's point of view, we can state:

**Theorem 6.4.9 (★).** *Consider the ternary model with  $T = 1$  and let the numbers  $p^*, q^* \in (0, 1)$  be equal to:*

$$p^* = \frac{r-a}{b-a} \quad \text{and} \quad q^* = \frac{b-r}{b-a}.$$

*The maximal expected utility of the ordinary agent is given by*

$$\Phi_1^{\text{ag}}(x) = \log(x) + \lambda_\gamma \left[ p \log\left(\frac{p}{p^*}\right) + q \log\left(\frac{q}{q^*}\right) \right],$$

*and realized by the strategy*

$$\varphi^{\text{ag}} = \frac{x(1+r)}{b-a} \times \frac{pq^* - qp^*}{p^*q^*}.$$

*The corresponding optimal investment is given by*

$$V_{x,1}^{\text{ag}} = x \mathbf{1}_{\{\Delta N_1=0\}} + \frac{xp}{p^*} \mathbf{1}_{\{(\Delta N_1, W_1)=(1,1)\}} + \frac{xq}{q^*} \mathbf{1}_{\{(\Delta N_1, W_1)=(1,-1)\}}.$$

**Corollary 6.4.10 (★).** *The optimal couple  $(\hat{y}_x, \widehat{\mathbf{M}}_{\hat{y}_x}) \in \mathbf{R} \times \mathcal{C}_1^{\mathbb{F}}$  defined in Theorem 6.4.7 are such that*

$$\hat{y}_x = \frac{1}{x} \quad \text{and} \quad \widehat{\mathbf{M}}_{\hat{y}_x} = \left( 1 - \lambda_\gamma, \frac{r - a\lambda_\gamma}{b-a}, \frac{b\lambda_\gamma - r}{b-a} \right)$$

**Remark 6.4.11.** Considering the specific case  $p = q = \frac{1}{2}$  and  $\lambda = 1$ , we exactly retrieve the result stated by Freddy Delbaen et Walter Schachermayer (see example 3.3.2. in [42]):

$$\pi_t^* = -(1+r) \frac{p(b-r) + q(a-r)}{(b-r)(a-r)} = \frac{1}{2} \frac{(1+r)(b+a-2r)}{(b-r)(r-a)}.$$

We can get the result using a variational argument:

**Theorem 6.4.12** (★). *The maximal expected logarithmic utility of the agent is given at maturity by*

$$\Phi_1^{\text{ag}}(x) = \log(V_0) + \mathbf{E} [\log ([r + \pi^* (b\Delta Z_{t,1} - a\Delta Z_{t,-1})])],$$

where, for any  $t \in \mathbf{N}_T$ ,

$$\pi^* = \frac{1+r}{b-a} \times \frac{pq^* - qp^*}{p^*q^*}.$$

The maximal logarithmic of the agent is performed when following the strategy  $\psi^{\text{ag}} = (\alpha_1^{\text{ag}}, \varphi_1^{\text{ag}})$  with

$$\varphi_1^{\text{ag}} = x\pi^* = \frac{x(1+r)}{b-a} \times \frac{pq^* - qp^*}{p^*q^*} \quad \text{and} \quad \alpha_1^{\text{ag}} = \frac{x - S_0}{\varphi_1^{\text{ag}}}.$$

**Remark 6.4.13.** As expected, the strategy  $\psi$  which components are  $\mathcal{F}_0$ -measurable, does not depend of the expression of  $G$ .

The situation is different from the point of view of the insider who knows wether the stock price will have increased or not at the end of the trading period  $T = 1$ . In other words, he knows from the start wether it is worth investing in the "risky" asset; this appears in fact as a riskless one for him since he knows the outcome. One of two things must be true: either  $G = 1$  namely the stock price does not increase and it may be better to invest the entire capital in the asset A, or  $G = 0$  and investing in S is more profitable. As a reminder the budget constraint can be written as

$$x = V_0(\psi) = \alpha_0^{\text{ins}} + \varphi_0^{\text{ins}}S_0,$$

and be transposed at time  $T = 1$  to  $\alpha_1^{\text{ins}} + \varphi_1^{\text{ins}}S_0 = x$ , by readjusting the portfolio under the self-financing condition. We get clearly:

$$\alpha_1^{\text{ins}} = x\mathbf{1}_{\{G=1\}} \quad \text{and} \quad \varphi_1^{\text{ins}} = \frac{x}{S_0}\mathbf{1}_{\{G=0\}}.$$

### Computation of $\Phi_T^{\text{ag}}(x)$ , $T \geq 2$

As suggested in the Example 3.3.5. depicted by Freddy Delbaen and Walter Schachermayer in [42], the case  $T \in \mathbf{N}^*$  can be treated by extending the previous results via the principle of dynamic programming. In a certain sense, we make of "a multiplicative concatenation" of what has been done in the case  $T = 1$ . The choice of this procedure can be justified by the independence of increments of the underlying jump process  $(N_t)_{t \in \mathbf{N}_T}$ . Conserving the previous notations, define

$$\Theta_t^{\text{ag}}(x) = \sup_{\psi \in \mathbb{H}_t\text{-portfolio}} \mathbf{E} \left[ u \left( x + \sum_{s=t+1}^T \varphi_s \Delta S_s \right) \middle| \mathcal{F}_t \right]$$

where  $\mathbb{H}_t = (\mathcal{F}_s)_{t \leq s \leq T}$  and  $\psi = (\alpha, \varphi)$ . From the i.i.d. property of the variables  $\{\Delta S_t, t \in \mathbf{N}_T\}$  we infer that

$$\Theta_t^{\text{ag}}(x) = \Phi_t^{\text{ag}}(x),$$

so that we can state its expression by the following induction system :

$$\begin{cases} \Phi_T^{\text{ag}}(x) &= \log(x) \\ \Phi_{t-1}^{\text{ag}}(x) &= \sup_{\psi \in \mathbb{H}_1\text{-portfolio}} \mathbf{E} [\Theta_t^{\text{ag}}(x + \varphi \Delta S_t)] \end{cases}$$

Indeed, for  $t = T - 1$ ,

$$\begin{aligned} \Phi_{T-1}^{\text{ag}}(x) &= \Theta_{T-1}^{\text{ag}}(x) = \sup_{\psi \in \mathbb{H}_{T-1}\text{-portfolio}} \mathbf{E} \left[ u(x + \varphi_T \Delta S_T) \mid \mathcal{F}_{T-1} \right] \\ &= \sup_{\psi \in \mathbb{H}_1\text{-portfolio}} \mathbf{E} [u(\tilde{x} + \varphi_1 \Delta S_1)] \Big|_{(\tilde{x}=x+\varphi_{T-1} \Delta S_{T-1})} \end{aligned}$$

with and so on by downward induction. The iteration of (6.5.2) provides :

$$\Phi_T^{\text{bin}}(x) = \log(x) + T k^{\text{bin}}$$

where  $k^{\text{bin}} = -p \log(p^*) - q \log(q^*) + p \log(p) + q \log(q)$  and then,

$$\Phi_T^{\text{ag}}(x) = \lambda_\gamma (\log(x) + T k^{\text{bin}}) + (1 - \lambda_\gamma) \log(x).$$

We have also  $\Phi_t^{\text{ag}}(x) = \lambda_\gamma (\log(x) + t k^{\text{bin}}) + (1 - \lambda_\gamma) \log(x)$  for any  $t \in \mathbf{N}_{T-1}$ . Moreover,  $\hat{y}_{x,t} = (\Phi_t^{\text{ag}})'(x) = 1/x$  and

$$\widehat{\mathbf{M}}_{\hat{y}_x}^t = \left( 1 - \lambda_\gamma, \frac{r - a\lambda_\gamma}{b - a}, \frac{b\lambda_\gamma - r}{b - a} \right)^{\otimes t}. \quad (6.4.8)$$

Indeed,  $\widehat{\mathbf{M}}_{\hat{y}_x}^t$  is an equivalent measure to  $\mathbf{P}|_{\mathcal{F}_t}$  such that  $\mathbf{E}_{\widehat{\mathbf{M}}_{\hat{y}_x}^t} [\bar{S}_t - \bar{S}_{t-1} \mid \mathcal{F}_{t-1}] = 0$ . From the i.i.d. property of the variables  $\{\Delta S_t, t \in \mathbf{N}_T\}$ , we deduce that  $\widehat{\mathbf{M}}_{\hat{y}_x}^t |_{\mathcal{F}_{t-1}} = \widehat{\mathbf{M}}_{\hat{y}_x}^1$  where  $\widehat{\mathbf{M}}_{\hat{y}_x}^1$  is solution of  $(\mathcal{S}^{\mathbf{M}})$ .

#### Computation of $\mathcal{U}_t(x)$ , $t \in \mathbf{N}_{T-1}$ ( $T \geq 2$ )

Consider first the general case where  $G$  is a discrete random variable in values in  $\Gamma$ . In what follows,  $\text{Ent}(G)$  designates the entropy of  $G$  namely the quantity

$$\text{Ent}(G) = - \sum_{c \in \Gamma} \log(\mathbf{P}(G = c)) \mathbf{P}(G = c),$$

and  $\text{Ent}(G \mid \mathcal{F}_t)$  the *conditional entropy*

$$\text{Ent}(G \mid \mathcal{F}_t) = \log(\mathbf{P}(G = c \mid \mathcal{F}_t)) \mathbf{P}(G = c \mid \mathcal{F}_t).$$

Recalling the definition of the process  $1/p^G$ ,

$$\begin{aligned} \mathcal{U}_t(x) &= \mathbf{E} [\log(p_t^G)] \\ &= \mathbf{E} \left[ \sum_{c \in \Gamma} \log(p_t^c) \mathbf{P}(G = c \mid \mathcal{F}_t) \right] \\ &= \mathbf{E} \left[ \sum_{c \in \Gamma} \log(\mathbf{P}(G = c \mid \mathcal{F}_t)) \mathbf{P}(G = c \mid \mathcal{F}_t) \right] - \sum_{c \in \Gamma} \log(\mathbf{P}(G = c)) \mathbf{E} [\mathbf{E}[\mathbf{1}_{(G=c)} \mid \mathcal{F}_t]] \\ &= \mathbf{E} \left[ \sum_{c \in \Gamma} \log(\mathbf{P}(G = c \mid \mathcal{F}_t)) \mathbf{P}(G = c \mid \mathcal{F}_t) \right] - \sum_{c \in \Gamma} \log(\mathbf{P}(G = c)) \mathbf{P}(G = c) \\ &= \text{Ent}(G) - \text{Ent}(G \mid \mathcal{F}_t). \end{aligned}$$

where we get the second line by conditioning on  $\mathcal{F}_t$ . In particular, for  $G = \mathbf{1}_{\{S_T \in [0, S_0]\}}$

$$\mathbf{P}(G = 1 \mid \mathcal{F}_T) = \mathbf{1}_{\{G=1\}} = \mathbf{1}_{\{S_T \in [0, S_0]\}}$$

and for any  $t \in \mathbf{N}_{T-1}^0$ ,

$$\begin{aligned} \mathbf{P}(G = 1 \mid \mathcal{F}_t) &= \mathbf{P}(S_T \in [0, S_0] \mid \mathcal{F}_t) \\ &= \mathbf{P}\left(\frac{S_T}{S_t} \times S_t \in [0, S_0] \mid \mathcal{F}_t\right) \\ &= \mathbf{P}\left(\frac{S_T}{S_t} \in \left[0, \frac{S_0}{c}\right] \mid \mathcal{F}_t\right) \Big|_{c=S_t} \\ &= \mathbf{P}\left(T_{T-t} \in \left[0, \frac{S_0}{c}\right] \mid \mathcal{F}_t\right) \Big|_{c=S_t} \end{aligned}$$

where we have used in the last line the variable  $S_T/S_t$  has the same law as  $S_{T-t}$  and  $T_{T-t}$  by equivalence of our model and the trinomial model (see remark 6.1.1). The distribution of  $T_t$  is given by

$$\mathbf{P}(T_t = c) = \sum_{\substack{(n,m,\ell) \in \mathbf{N}_t^3 \\ n+m+\ell=t, n-m=c}} \frac{t!}{n! m! \ell!} (\lambda p)^n (\lambda q)^m (1-\lambda)^{t-(n+m)}$$

for any  $c \in S_t(\Omega)$ .

## 6.5 Proofs of chapter 6

### 6.5.1 Proofs of Section 6.2

*Proof of Theorem 6.2.1.* The corollary 5.4.3 applied to  $p^G$  at time  $s = t - 1$  gives

$$\Delta p_t^G = \sum_{k \in E} \mathbf{E}[D_{t,\ell} p_t^G \mid \mathcal{F}_{t-1}] \Delta R_{t,\ell},$$

Then we get, for any  $c \in \Gamma$ ,

$$\begin{aligned} \Delta \langle \bar{Y}, p^c \rangle_t &= \mathbf{E} \left[ \sum_{k \in E} \Delta Z_{t,k} \sum_{\ell \in E} \mathbf{E}[D_{t,\ell} p_t^c \mid \mathcal{F}_{t-1}] \Delta R_{t,\ell} \right] \\ &= \sum_{k \in E} \sum_{\ell \in E} \mathbf{E}[\mathbf{E}[D_{t,\ell} p_t^c \mid \mathcal{F}_{t-1}] \mathbf{E}[\Delta Z_{t,k} \Delta R_{t,\ell}]] \\ &= \sum_{k \in E} \sum_{\ell \in E} a_{k,\ell} \mathbf{E}[D_{t,\ell} p_t^c], \end{aligned}$$

where we have got the second line by conditioning with respect to  $\mathcal{F}_{t-1}$  and by defining the family  $\{a_{k,\ell}, (k, \ell) \in E^2\}$  by  $a_{k,\ell} = \mathbf{E}[\Delta Z_{t,k} \Delta R_{t,\ell}]$ , i.e.

$$a_{1,1} = \lambda p(1 - \lambda p), \quad a_{1,-1} = 0, \quad a_{-1,1} = \lambda^2 p q \quad \text{and} \quad a_{-1,-1} = \frac{\lambda q(1 - \lambda)}{1 - \lambda p}.$$

Hence the result. □

*Proof of Theorem 6.2.5.* The proof follows closely the one of Proposition 2.3 ([7]). For any  $A_t \in \mathcal{F}_t$ ,  $B \in \mathcal{E}$ ,

$$\begin{aligned} \mathbf{E} \left[ \mathbf{1}_{A_t \cap \{G \in B\}} \frac{1}{p_t^G} \right] &= \mathbf{E} \left[ \mathbf{1}_{A_t} \mathbf{E} \left[ \mathbf{1}_{\{G \in B\}} \frac{1}{p_t^G} \mid \mathcal{F}_t \right] \right] \\ &= \mathbf{E} [\mathbf{1}_{A_t}] \mathbf{P}(G \in B) \\ &= \mathbf{P}(A_t) \mathbf{P}(G \in B), \end{aligned}$$

where we used:  $\mathbf{E} \left[ \mathbf{1}_{\{G \in B\}} \frac{1}{p_t^G} \mid \mathcal{F}_t \right] = \sum_{c \in \Gamma \cap B} \frac{1}{p_t^c(\omega)} \cdot p_t^c(\omega) \cdot \mathbf{P}(G = c) = \mathbf{P}(G \in B)$ .

This yields

$$\mathbf{Q}_t(A_t \cap \{G \in B\}) = \mathbf{P}(G \in B) \mathbf{P}(A_t),$$

for any  $t \in \mathbf{N}_{T-1}^0$ . Taking  $A_t = \Omega$ , then  $B = \Gamma$  provides

$$\mathbf{Q}_t(A_t \cap \{G \in B\}) = \mathbf{Q}_t(G \in B) \mathbf{Q}_t(A_t),$$

and enables to establish 2. and 3. Let  $A = A_s \cap \{G \in B\}$  an element of  $\mathcal{G}_s$  and  $t > s$ . By noting that  $A_s \in \mathcal{F}_t$  and using what precedes,

$$\begin{aligned} \mathbf{E} \left[ \mathbf{1}_A \frac{1}{p_t^G} \right] &= \mathbf{P}(A_s) \mathbf{P}(G \in B) \\ &= \mathbf{Q}_s(A_s \cap \{G \in B\}) \\ &= \mathbf{E} \left[ \mathbf{1}_A \frac{1}{p_s^G} \right], \end{aligned}$$

so that the process  $1/p^G$  is a  $(\mathbf{P}, \mathbb{G}^0)$ -martingale.  $\square$

*Proof of Theorem 6.2.7.* Let  $L = (L_s)_{1 \leq s \leq t}$  a  $(\mathbf{P}, \mathbb{F})$ -martingale on  $\mathbf{N}_t$ . For  $r < t$  and  $s \in \llbracket r+1, t \rrbracket$  let  $A = A_r \cap \{G \in B\}$  an element of  $\mathcal{G}_r$ . Let  $\mathbf{E}_{\mathbf{Q}_t}$  denote the expectation taken with respect to  $\mathbf{Q}_t$ .

$$\begin{aligned} \mathbf{E}_{\mathbf{Q}_t} [\mathbf{1}_A L_s] &= \mathbf{E}_{\mathbf{Q}_t} [\mathbf{1}_{A_r} L_s] \mathbf{E}_{\mathbf{Q}_t} [\mathbf{1}_{\{G \in B\}}] \\ &= \mathbf{E} [\mathbf{1}_{A_r} L_s] \mathbf{E}_{\mathbf{Q}_t} [\mathbf{1}_{\{G \in B\}}] \\ &= \mathbf{E} [\mathbf{1}_{A_r} \mathbf{E} [L_s \mid \mathcal{F}_r]] \mathbf{E}_{\mathbf{Q}_t} [\mathbf{1}_{\{G \in B\}}] \\ &= \mathbf{E} [\mathbf{1}_{A_r} L_r] \mathbf{E}_{\mathbf{Q}_t} [\mathbf{1}_{\{G \in B\}}] \\ &= \mathbf{E}_{\mathbf{Q}_t} [\mathbf{1}_{A_r} L_r] \mathbf{E}_{\mathbf{Q}_t} [\mathbf{1}_{\{G \in B\}}] \\ &= \mathbf{E}_{\mathbf{Q}_t} [\mathbf{1}_A L_r], \end{aligned}$$

where we have used that the  $\sigma$ -algebras  $\mathcal{F}_t$  and  $\sigma(G)$  are independent under  $\mathbf{Q}_t$  in the first line, that  $\mathbf{P}$  coincide with  $\mathbf{Q}_t$  on  $\mathcal{F}_t$  in the second one, and that  $(L_s)_{1 \leq s \leq t}$  is a  $(\mathbf{P}, \mathbb{F})$ -martingale on  $\mathbf{N}_t$  in the third one.  $\square$



### 6.5.2 Proofs of Section 6.3

*Proof of Lemma 6.3.4.* The Kunita-Watanabe decomposition (see [43], theorem VIII.51.) yields

$$\begin{aligned}\xi_t^F &= \frac{\mathbf{E} \left[ \widehat{\mathbf{E}} [F | \mathcal{F}_t] \Delta S_t | \mathcal{F}_{t-1} \right]}{\mathbf{E} [(\Delta S_t)^2 | \mathcal{F}_{t-1}]} \\ &= \frac{\sum_{s=1}^t \sum_{k \in \mathbf{E}} \mathbf{E} \left[ \widehat{\mathbf{E}} [D_{s,k} F | \mathcal{F}_{s-1}] \Delta R_{s,k} ((b + a\rho)\Delta R_{t,1} - a\Delta R_{t,-1}) | \mathcal{F}_{t-1} \right]}{\|c\|_2^2} \\ &= \frac{\sum_{k \in \mathbf{E}} c_k \widehat{\mathbf{E}} [D_{t,k} F | \mathcal{F}_{t-1}]}{\|c\|_2^2},\end{aligned}$$

where we applied the Clark decomposition to  $F$  in the second line, and by letting  $c = (c_1, c_{-1})$ ,

$$c_1 = (b + a\rho)\kappa_1 \quad \text{and} \quad c_{-1} = -a\kappa_{-1}.$$

Hence the result.  $\square$

*Proof of Theorem 6.3.5.* As a reminder, the strategy  $\psi = (\alpha, \varphi)$  is self-financed and only if for any  $t \in \mathbf{N}_T$ ,

$$A_t(\alpha_{t+1} - \alpha_t) = S_t(\varphi_{t+1} - \varphi_t), \quad (6.5.1)$$

so that  $V_{t-1}(\psi) = \alpha_t A_{t-1} + \varphi_t S_{t-1}$ . Let  $\varphi_0 = 0$ . Assume the existence of an admissible strategy  $(\alpha, \varphi)$  which final value satisfies

$$V_T(\psi) = \alpha_T A_T + \varphi_T S_T.$$

Let  $\pi$  be the  $\mathbb{F}$ -predictable process such that  $\pi_t = \frac{\varphi_t S_{t-1}}{V_{t-1}(\psi)}$  for any  $t \in \mathbf{N}_T$ . By definition,

$$\begin{aligned}\Delta V_t(\psi) &= \alpha_t \Delta A_t + \varphi_t \Delta S_t \\ &= V_{t-1}(\psi) \left[ \frac{\alpha_t \Delta A_t}{V_{t-1}(\psi)} + \pi_t \left( b \mathbf{1}_{\{(\Delta N_t, W_t) = (1,1)\}} + a \mathbf{1}_{\{(\Delta N_t, W_t) = (1,-1)\}} \right) \right] \\ &= V_{t-1}(\psi) \left( \frac{\alpha_t r A_{t-1}}{V_{t-1}(\psi)} + \pi_t (b \Delta Z_{t,1} - a \Delta Z_{t,-1} + r) \right) \\ &= V_{t-1}(\psi) (r(1 - \pi_t) + \pi_t (b \Delta Z_{t,1} - a \Delta Z_{t,-1} + r)) \\ &= r V_{t-1}(\psi) + V_{t-1}(\psi) \pi_t (b \Delta Z_{t,1} - a \Delta Z_{t,-1}),\end{aligned}$$

where we used in the third line that  $\frac{\alpha_t A_{t-1}}{V_{t-1}(\psi)} + \pi_t = 1$ . Then,

$$\Delta \bar{V}_t(\psi) = \frac{\bar{V}_{t-1}(\psi) \pi_t}{1+r} \left( \left( b - \frac{a\lambda q}{1-\lambda p} \right) \Delta R_{t,1} + a \Delta R_{t,-1} \right)$$

so that

$$\bar{V}_T(\psi) = 1 + \sum_{t=1}^T \frac{\bar{V}_{t-1}(\psi) \pi_t}{1+r} \left( \left( b - \frac{a\lambda q}{1-\lambda p} \right) \Delta R_{t,1} + a \Delta R_{t,-1} \right)$$

by assuming (even if it means dividing by  $\bar{V}_0$ ),  $\bar{V}_0 = 1$ . We get the equal expression

$$\bar{V}_T(\psi) = \prod_{t=1}^T \left[ 1 + \frac{\pi_t}{1+r} \left( \left( b - \frac{a\lambda q}{1-\lambda p} \right) \Delta R_{t,1} + a \Delta R_{t,-1} \right) \right]$$

The application of the derivative in both directions gives

$$D_{t,1} \bar{V}_T(\psi) = \bar{V}_T^{-t}(\psi) \left( b + \frac{a\lambda q}{1-\lambda p} \right) \pi_t$$

and

$$D_{t,-1} \bar{V}_T(\psi) = -a \pi_t \bar{V}_T^{-t}(\psi),$$

by using  $D_{t,k} \Delta Z_{t,k} = \text{sgn}(k)$ . This entails

$$\hat{\mathbf{E}} [D_{t,1} \bar{V}_T | \mathcal{F}_{t-1}] = \left( b + \frac{a\lambda q}{1-\lambda p} \right) \pi_t \bar{V}_{t-1}(\psi) \quad \text{and} \quad \hat{\mathbf{E}} [D_{t,-1} \bar{V}_T | \mathcal{F}_{t-1}] = -a \pi_t \bar{V}_{t-1}(\psi)$$

so that

$$\hat{\mathbf{E}} [D_{t,1} \bar{V}_T(\psi) + D_{t,-1} \bar{V}_T(\psi) | \mathcal{F}_{t-1}] = \left( b - a + \frac{a\lambda q}{1-\lambda p} \right) \pi_t \bar{V}_{t-1}(\psi)$$

Then, by letting on the one hand  $\varphi_0 = 0$  and

$$\varphi_t = V_{t-1}(\psi) \pi_t = (1+r)^{t-1} \frac{\sum_k \hat{\mathbf{E}} [D_{t,k} \bar{V}_T(\psi) | \mathcal{F}_{t-1}]}{b - a + \frac{a\lambda q}{1-\lambda p}} = (1+r)^{t-1} \frac{\sum_k \hat{\mathbf{E}} [D_{t,k} F | \mathcal{F}_{t-1}]}{b - a + \frac{a\lambda q}{1-\lambda p}},$$

and, on the other one,  $\alpha_0 = (1+r)^{-T} \hat{\mathbf{E}} [F] / S_0$  and for any  $t \in \mathbf{N}_T$ ,

$$\alpha_t = \alpha_{t-1} - \frac{(\varphi_t - \varphi_{t-1}) S_{t-1}}{A_{t-1}}$$

we get a couple of  $\mathbb{F}$ -predictable processes  $\psi = (\alpha, \varphi)$  such that satisfies the self-financing condition and of terminal value  $F$ .  $\square$

### 6.5.3 Proofs of Section 6.4

*Proof of Theorem 6.4.3.* Transposing the result of Theorem 6.4.1 to solve the problem (6.4.5) we get

$$\Phi^{\mathbb{G}}(x) = \mathbf{E} \left[ u(\hat{V}_{x,t}^{\mathbb{G}}) \right] = \mathcal{L}(\hat{V}_{x,t}^{\mathbb{G}}, \hat{y}_x^{\mathbb{G}}, \hat{\mathbf{M}}_{\hat{y}_x^{\mathbb{G}}}^{\mathbb{G}}),$$

where  $\mathcal{L}$  is the Lagrangian associated to the primal problem 6.4.5,  $\hat{\mathbf{M}}_{\hat{y}_x^{\mathbb{G}}}^{\mathbb{G}}$  is the element of  $\mathcal{C}_t^{\mathbb{G}}$  unique minimiser of the value function (for  $y = \sum_{j \in J_t} \eta_j$  fixed)

$$\Psi(y) : \mathbf{M} \in \mathcal{M}_t^{\mathbb{G}} \mapsto \left( \sup_{V \in \Lambda_{x,t}} \{ \mathbf{E} [u(V)] - y \mathbf{E}_{\mathbf{M}} [V - x] \} \right),$$

We get  $\hat{y}_x^{\mathbb{G}} = 1/x$  and  $\widehat{\mathbf{M}}_{\hat{y}_x^{\mathbb{G}},t}^{\mathbb{G}} = \sum_{j \in J_t} (\eta_j / \hat{y}_x^{\mathbb{G}}) \mathbf{Q}_t^j$  where  $(\eta_j)_{j \in J_t} \in [0, 1]^{J_t}$  is a  $J_t$ -uplet summoned to one. Thus, for any  $j \in J_t$ ,

$$\begin{aligned} \mathbf{E}_{\mathbf{Q}_t^j}[\widehat{\mathbf{V}}_{x,t}^{\mathbb{G}}] &= \mathbf{E}_{\mathbf{Q}_t^j} \left[ I \left( \hat{y}_x^{\mathbb{G}} \frac{d\widehat{\mathbf{M}}_{\hat{y}_x^{\mathbb{G}},t}^{\mathbb{G}}}{d\mathbf{P}} \right) \right] \\ &= \mathbf{E}_{\mathbf{Q}_t^j} \left[ \frac{d\mathbf{P}}{d\widehat{\mathbf{M}}_{\hat{y}_x^{\mathbb{G}},t}^{\mathbb{G}}} \right] \\ &= \mathbf{E}_{\mathbf{Q}_t^j} \left[ \frac{d\mathbf{P}}{d\mathbf{Q}_t^j} \frac{d\mathbf{Q}_t^j}{d\widehat{\mathbf{M}}_{\hat{y}_x^{\mathbb{G}},t}^{\mathbb{G}}} \right] \\ &= \mathbf{E} \left[ \frac{d\mathbf{Q}_t^j}{\sum_{j \in J_t} \eta_j d\mathbf{Q}_t^j} \right]. \end{aligned}$$

It is thus possible to compare the logarithmic utilities up to time  $t$  of the two agents and to quantify the insider's additional one. As for the economic agent, the sequence  $(\eta_j)_{j \in J_t}$  is determined thanks to the initial budget constraints, as to say

$$\begin{aligned} x &= \mathbf{E} \left[ \frac{d\mathbf{Q}_t^j}{\sum_{j \in J_t} \eta_j d\mathbf{Q}_t^j} \right] \\ &= \mathbf{E} \left[ \frac{L_t d\mathbf{Q}_t^j}{\sum_{j \in J_t} \eta_j L_t d\mathbf{Q}_t^j} \right] \\ &= \mathbf{E} \left[ \frac{d\mathbf{P}_t^j}{\sum_{j \in J_t} \eta_j d\mathbf{P}_t^j} \right]. \end{aligned}$$

The sequence  $(\eta_j)_{j \in J_t}$  is solution of the system satisfied by the sequence  $(\lambda_j)_{j \in J_t}$ ; by uniqueness of the solutions, for any  $j \in J_t$ ,  $\eta_j = \lambda_j$ . Thus,

$$\begin{aligned} \mathcal{U}_t &= \mathbf{E} \left[ \log \left( \frac{d\mathbf{P}}{\sum_{j \in J_t} \lambda_j d\mathbf{P}_t^j} \right) \right] - \mathbf{E} \left[ \log \left( \frac{d\mathbf{P}}{\sum_{j \in J_t} \lambda_j d\mathbf{Q}_t^j} \right) \right] \\ &= \mathbf{E} \left[ \log \left( \frac{\sum_{j \in J_t} \lambda_j d\mathbf{Q}_t^j}{\sum_{j \in J_t} \lambda_j d\mathbf{P}_t^j} \right) \right] \\ &= \mathbf{E} \left[ \log \left( \frac{\sum_{j \in J_t} L_t \lambda_j d\mathbf{P}_t^j}{\sum_{j \in J_t} \lambda_j d\mathbf{P}_t^j} \right) \right] \\ &= \mathbf{E} [\log(L_t)] = \mathfrak{D}_{\mathcal{G}_t}(\mathbf{P} || \mathbf{Q}_t) \end{aligned}$$

where we have used that  $L = 1/p^{\mathbb{G}}$  is a  $(\mathbf{P}, \mathbb{G}^0)$ -martingale.  $\square$

*Proof of Lemma 6.4.7.* Consider a  $\mathbb{G}$ -predictable admissible strategy  $\theta$ . By definition of  $\Phi$ ,  $\Phi_{T-1}^{\text{ins}}(x) > \mathbf{E}[u(V_{T-1}(\theta))]$ . Transposing once again the result of Theorem 3.2.1. of [42] to our model, we get  $(\Phi_T^{\text{ins}})'(x) > 0$ . In particular for  $T = 1$  and with  $x_\theta = \mathbf{E}[u(V_{T-1}(\theta))]$ , we get

$$\Phi_1^{\text{ins}}(\tilde{x}) > \Phi_1^{\text{ins}}(x_\theta).$$

This holds for any  $\mathbb{G}$ -predictable admissible strategy  $\theta$ . Assume moreover the existence of a  $\mathbb{G}$ -predictable admissible strategy  $\tilde{\theta}$  such that

$$\mathbf{E} \left[ u(V_{x,T}(\tilde{\theta})) \right] = \Phi_T^{\text{ins}}(x),$$

and let  $x_{\tilde{\theta}} = \mathbf{E} \left[ u(V_{T-1}(\tilde{\theta})) \right]$ . Then,

$$\Phi_T^{\text{ins}}(x) = \mathbf{E} \left[ u(V_{x,T}(\tilde{\theta})) \right] \leq \max_{\psi \in \mathbb{G}\text{-portfolio}} \mathbf{E} \left[ u(V_{x_{\tilde{\theta}},1}(\psi)) \right] \leq \Phi_1^{\text{ins}}(\tilde{x}).$$

Then,  $\Phi_1^{\text{ins}}(\tilde{x}) \leq \Phi_T^{\text{ins}}(x)$ . The result follows.  $\square$

*Proof of 6.4.9.* By transposing the results established by (example 3.3.2 in [42]) for the binomial model to our "associated" one ( $\mathcal{S}^{\text{bin}}$ ), we get the optimal strategy

$$\varphi^{\text{ag}} = \frac{x(1+r)}{b-a} \times \frac{pq^* - qp^*}{p^*q^*}$$

where

$$p^* = \frac{r-a}{b-a} \quad \text{and} \quad q^* = \frac{b-r}{b-a}$$

define the well-known risk-neutral probability measure in the Cox-Ross-Rubinstein model. The corresponding optimal investment from the agent's point of view (in the ternary model) is thus

$$\begin{aligned} V_{x,1}^{\text{ag}} &= x \mathbf{1}_{\{\Delta N_1=0\}} + \frac{x(b-a)}{p(r-a)} \mathbf{1}_{\{(\Delta N_1, W_1)=(1,1)\}} + \frac{x(b-a)}{(1-p)(b-r)} \mathbf{1}_{\{(\Delta N_1, W_1)=(1,-1)\}} \\ &= x \mathbf{1}_{\{\Delta N_1=0\}} + \frac{xp}{p^*} \mathbf{1}_{\{(\Delta N_1, W_1)=(1,1)\}} + \frac{xq}{q^*} \mathbf{1}_{\{(\Delta N_1, W_1)=(1,-1)\}} \end{aligned}$$

We get thus

$$\Phi_1^{\text{bin}}(x) = \log(x) + k^{\text{bin}} \tag{6.5.2}$$

where  $k^{\text{bin}} = -p \log(p^*) - q \log(q^*) + p \log(p) + q \log(q)$ . Follows from (6.4.7) that,

$$\log(x) + k^{\text{ag}} = \lambda_\gamma (\log(x) + k^{\text{bin}}) + (1 - \lambda_\gamma) \log(x)$$

that is  $k^{\text{ag}} = \lambda_\gamma k^{\text{bi}}$ . Thus, from

$$\Phi_1^{\text{ag}}(x) = \lambda_\gamma \Phi_1^{\text{bin}}(x) + (1 - \lambda_\gamma) \log(x)$$

so that

$$\Phi_1^{\text{ag}}(x) = \log(x) + \lambda_\gamma \left[ p \log \left( \frac{p}{p^*} \right) + q \log \left( \frac{q}{q^*} \right) \right].$$

Hence the result.  $\square$

*Proof.* The optimal couple  $(\hat{y}_x, \widehat{\mathbf{M}}_{\hat{y}_x}) \in \mathbf{R} \times \mathcal{C}_1^{\mathbb{F}}$  defined in Theorem 6.4.7 is such that

$$\hat{y}_x = \frac{1}{x} \quad \text{and} \quad \widehat{\mathbf{M}}_{\hat{y}_x} = \left( 1 - \lambda_\gamma, \frac{r - a\lambda_\gamma}{b-a}, \frac{b\lambda_\gamma - r}{b-a} \right)$$

Indeed, we have

$$\frac{d\widehat{\mathbf{M}}_{\widehat{y}_x}}{d\mathbf{P}}(\mathbf{0}) = (\widehat{y}_x)^{-1} \times u' \left( V_{x,1}^{\text{ag}}(\mathbf{0}) \right) = 1,$$

so that

$$\widehat{\mathbf{M}}_{\widehat{y}_x}(\mathbf{0}) = \frac{d\widehat{\mathbf{M}}_{\widehat{y}_x}}{d\mathbf{P}}(\mathbf{0}) \times \mathbf{P}(\mathbf{0}) = 1 - \lambda_\gamma.$$

Since  $\widehat{\mathbf{M}}_{\widehat{y}_x}$  is an equivalent measure to  $\mathbf{P}$  such that  $\mathbf{E}_{\widehat{\mathbf{M}}_{\widehat{y}_x}}[\overline{\mathbf{S}}_1 - \overline{\mathbf{S}}_0] = 0$  we have

$$(\mathcal{S}^{\mathbf{M}}) \begin{cases} \widehat{\mathbf{M}}_{\widehat{y}_x}(\mathbf{0}) & = 1 - \lambda_\gamma \\ b\widehat{\mathbf{M}}_{\widehat{y}_x}(\omega(1,1)) + a\widehat{\mathbf{M}}_{\widehat{y}_x}(\omega(1,-1)) & = r \\ \widehat{\mathbf{M}}_{\widehat{y}_x}(\omega(1,1)) + \widehat{\mathbf{M}}_{\widehat{y}_x}(\omega(1,-1)) + \widehat{\mathbf{M}}_{\widehat{y}_x}(\mathbf{0}) & = 1 \end{cases}$$

and thus

$$\widehat{\mathbf{M}}_{\widehat{y}_x}(\omega(1,1)) = \frac{r - a\lambda_\gamma}{b - a} \quad \text{and} \quad \widehat{\mathbf{M}}_{\widehat{y}_x}(\omega(1,-1)) = \frac{b\lambda_\gamma - r}{b - a}.$$

In other words,

$$\widehat{\mathbf{M}}_{\widehat{y}_x} = \lambda_0 \mathbf{P}_0 + \lambda_1 \mathbf{P}_1 = \frac{b\lambda_\gamma - r}{b - r} \mathbf{P}_0 + \frac{r(1 - \lambda_\gamma)}{b - r} \mathbf{P}_1. \quad (6.5.3)$$

Hence the result.  $\square$

# Bibliography

- [1] S. Albeverio, Y. G. Kondratiev, and M. Röckner. Analysis and geometry on configuration spaces. J. Funct. Anal., 154(2):444–500, 1998.
- [2] E. Alòs, J. A. León, and J. Vives. On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. Finance and Stochastics, 11(4):571–589, 2007.
- [3] E. Alòs, J. A. León, and J. Vives. An anticipating Itô formula for Lévy processes. Latin American Journal of Probability and Mathematical Statistics, 4:285–330, 2008.
- [4] E. Alòs and D. Nualart. An extension of Itô’s formula for anticipating processes. Journal of Theoretical Probability, 11(2):493–514, 1998.
- [5] J. Amendinger. Martingale representation theorems for initially enlarged filtrations. Stochastic Processes and their Applications, 89(1):101–116, 2000.
- [6] J. Amendinger, D. Becherer, and M. Schweizer. A monetary value for initial information in portfolio optimization. Finance and Stochastics, 7:29–46, 12 2002.
- [7] J. Amendinger, P. Imkeller, and M. Schweizer. Additional logarithmic utility of an insider. Stochastic Processes and their Applications, 75(2):263–286, 1998.
- [8] B. Arras and Y. Swan. A stroll along the gamma. Stochastic Processes and their Applications, 127(11):3661–3688, nov 2017.
- [9] R. Arratia, A. D. Barbour, and S. Tavaré. Poisson process approximations for the Ewens sampling formula. Ann. Appl. Probab., 2(3):519–535, 1992.
- [10] R. Arratia, A. D. Barbour, and S. Tavaré. Logarithmic combinatorial structures: a probabilistic approach. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2003.
- [11] E. Azmoodeh, S. Campese, and G. Poly. Fourth moment theorems for Markov diffusion generators. Journal of Functional Analysis, 266(4):2341–2359, Feb 2014.
- [12] D. Bakry, I. Gentil, and M. Ledoux Michel. Analysis and Geometry of Markov Diffusion operators. Grundlehren der mathematischen Wissenschaften, Vol. 348. Springer, January 2014.
- [13] V. Bally and D. Talay. The Euler scheme for stochastic differential equations: Error analysis with Malliavin calculus. Math. Comput. Simul., 38(1–3):35–41, May 1995.

- [14] A. D. Barbour and L. H. Y. Chen. An introduction to Stein's method, volume 4 of Lecture Notes Series. National University of Singapore, 2005.
- [15] M.-P. Bavouzet-Morel and M. Messaoud. Computation of greeks using malliavin's calculus in jump type market models. Electronic Journal of Probability [electronic only], 11:276–300, 2006.
- [16] E. Benhamou. A Generalisation of Malliavin Weighted Scheme for Fast Computation of the Greeks. FMG Discussion Papers dp350, Financial Markets Group, April 2000.
- [17] M. A. Berger. A malliavin-type anticipative stochastic calculus. Ann. Probab., 16(1):231–245, 01 1988.
- [18] K. Bichteler. Malliavin calculus for processes with jumps. Stochastics Monographs, 1987.
- [19] J.-M. Bismut. Martingales, the Malliavin calculus and hypoellipticity under general Hormander's condition. Z fur Wahrsche. verw. Gebiete, 56:469–505, 1981.
- [20] C. Blanchet-Scalliet, M. Jeanblanc, and R. Romo Roméro. Enlargement of filtration in discrete time. In Pauline Barrieu, editor, Risk And Stochastics: Ragnar Norberg, pages 99–126. Worl Scientific, 2019.
- [21] S. Boucheron, G. Lugosi, and P. Massart. Concentration inequalities. Oxford University Press, Oxford, 2013. DOI: 10.1093/acprof:oso/9780199535255.001.0001.
- [22] N. Bouleau. Théorème de Donsker et formes de Dirichlet. Bulletin des sciences mathématiques, 129:63–73, 2005. arXiv: math/0610392.
- [23] N. Bouleau and F. Hirsch. Dirichlet forms and analysis on Wiener space, volume 14. Walter de Gruyter, 2010.
- [24] S. Chatterjee. A new method of normal approximation. Ann. Probab., 36(4):1584–1610, 07 2008.
- [25] S. Chatterjee, J. Fulman, and A. Rollin. Exponential approximation by stein's method and spectral graph theory, 2006.
- [26] L. H. Y. Chen. On the convergence of Poisson binomial to Poisson distributions. Ann. Probability, 2(1):178–180, 1974.
- [27] L. H. Y. Chen, L. Goldstein, and Q.-M. Shao. Normal approximation by Stein's method, pages xii+405. Probability and its Applications (New York). Springer, Heidelberg, 2011.
- [28] L. HY 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, pages 107–130. World Scientific, 2015.
- [29] F. Comte and N. Marie. Nonparametric Estimation in Fractional SDE. Statistical Inference for Stochastic Processes, 2019.
- [30] R. Cont and P. Tankov. Nonparametric calibration of jump-diffusion option pricing models. 2004.

- [31] J. M. Corcuera and A. Kohatsu-Higa. Statistical inference and malliavin calculus. In Seminar on Stochastic Analysis, Random Fields and Applications VI, pages 59–82. Springer, 2011.
- [32] L. Coutin and L. Decreusefond. Stein’s method for Brownian approximations. Communications on stochastic analysis, 7(5):349–372, 2013.
- [33] L. Coutin and L. Decreusefond. Higher order expansions via Stein’s method. Communications on Stochastic Analysis, 8(2):155–168, 2014.
- [34] R. Dalang, A. Morton, and W. Willinger. Equivalent martingale measures and no-arbitrage in stochastic securities market models. Stochastics: An International Journal of Probability and Stochastic Processes, 29(2):185–201, 1990.
- [35] A. Dasgupta and G. Kallianpur. Multiple fractional integrals. Probability Theory and Related Fields, 115:505–525, 01 1999.
- [36] L. Decreusefond. The Stein-Dirichlet-Malliavin method. ESAIM: Proc., 51:49–59, 2015.
- [37] L. Decreusefond and I. Flint. Moment formulae for general point processes. Comptes Rendus Mathématique, 352(4):357 – 361, 2014.
- [38] L. Decreusefond and H. Halconruy. Malliavin and dirichlet structures for independent random variables. Stochastic Processes and their Applications, 129(8):2611–2653, Aug 2019.
- [39] L. Decreusefond and P. Moyal. Modélisation et analyse stochastiques des réseaux de télécommunication. Hermès, 2011.
- [40] L. Decreusefond and N. Savy. Anticipative calculus with respect to filtered Poisson processes. Annales de l’Institut Henri Poincaré (B) Probability and Statistics, 42(3):343 – 372, 2006.
- [41] L. Decreusefond, M. Schulte, and C. Thäle. Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry. Annals of probability, 44(3):2147–2197, June 2015.
- [42] F. Delbaen and W. Schachermayer. The Mathematics of Arbitrage. Springer-Verlag Berlin Heidelberg, 2006.
- [43] C. Dellacherie and P.-A. Meyer. Probabilities and potential b, chapter v to viii, 1982.
- [44] C. Döbler and G. Peccati. The gamma stein equation and non-central de jong theorems.
- [45] R. M. Dudley. Real analysis and probability. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1989.
- [46] N. T. Dung. Explicit rates of convergence in the multivariate clt for nonlinear statistics. Acta Mathematica Hungarica, 158(1):173–201, 2019.
- [47] C. Döbler and G. Peccati. Fourth moment theorems on the Poisson space: analytic statements via product formulae. Electronic Communications in Probability, 23, 2018.



- [48] C. Döbler and G. Peccati. The gamma stein equation and noncentral de jong theorems. Bernoulli, 24(4B):3384–3421, 11 2018.
- [49] M. Eddahbi and J. Vives. Chaotic expansion and smoothness of some functionals of the fractional brownian motion. J. Math. Kyoto Univ., 43(2):349–368, 2003.
- [50] Y. El-Khatib and N. Privault. Computations of greeks in a market with jumps via the malliavin calculus. Finance and Stochastics, 8, 09 2003.
- [51] M. Émery. On the Azéma martingales. Séminaire de probabilités de Strasbourg, 23:66–87, 1989.
- [52] M. Émery. A discrete approach to the chaotic representation property. working paper or preprint, 2000.
- [53] H. Föllmer and M. Schweizer. Hedging of contingent claims. Applied stochastic analysis, 5:389, 1991.
- [54] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi. Applications of malliavin calculus to monte carlo methods in finance. Finance and Stochastics, 3:391–412, 1999.
- [55] E. Fournié, J.-M. Lasry, P.-L. Lions, J. Lebuchoux, and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. Finance and Stochastics, 3(4):391–412, 1999.
- [56] M. Fukushima, Y. Oshima, and M. Takeda. Dirichlet Forms and Symmetric Markov Processes. De Gruyter, Berlin, Boston, 2010.
- [57] L. Goldstein. Bounds on the constant in the mean central limit theorem. The Annals of Probability, 38(4):1672–1689, 2010.
- [58] L. Goldstein and G. Reinert. Stein’s method and the zero bias transformation with application to simple random sampling. Ann. Appl. Probab., 7(4):935–952, 1997.
- [59] F. Götze and A. Tikhomirov. Asymptotic distribution of quadratic forms and applications. Journal of Theoretical Probability, 15(2):423–475, 2002.
- [60] P. Graczyk, J.-J. Loeb, I. A. López, A. Nowak, and R. W. O. Urbina. Higher order Riesz transforms, fractional derivatives, and Sobolev spaces for Laguerre expansions. Journal de Mathématiques Pures et Appliquées, 84(3):375–405, mar 2005.
- [61] A. Grorud and M. Pontier. Insider trading in a continuous time market model. International Journal of Theoretical and Applied Finance, 01(03):331–347, 1998.
- [62] L. Gross. Logarithmic sobolev inequalities. American Journal of Mathematics, 97(4):1061–1083, 1975.
- [63] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. The review of financial studies, 6(2):327–343, 1993.

- [64] M. Hitsuda. Formula for brownian partial derivatives, the second japan-ussr symp. on probab. theory.-tbilisi, 1972.
- [65] Y. Hu, J. Huang, K.Lê, D. Nualart, and Samy S. Tindel. Stochastic heat equation with rough dependence in space. The Annals of Probability, 45(6B):4561–4616, 2017.
- [66] Y. Hu, D. Nualart, and X. Song. Malliavin calculus for backward stochastic differential equations and application to numerical solutions. Ann. Appl. Probab., 21(6):2379–2423, 12 2011.
- [67] J. Hull and A. White. The pricing of options on assets with stochastic volatilities. The journal of finance, 42(2):281–300, 1987.
- [68] P. Imkeller. Malliavin’s calculus in insider models: Additional utility and free lunches. Mathematical Finance. An International Journal of Mathematics, Statistics and Financial Economics, 13(1):153–169, 2003. Conference on Applications of Malliavin Calculus in Finance (Rocquencourt, 2001).
- [69] P. Imkeller. Malliavin’s calculus and applications in stochastic control and finance, volume 1 of IMPAN Lecture Notes. Polish Academy of Sciences, Institute of Mathematics, Warsaw, 2008.
- [70] J. Jacod. Grossissement initial, hypothèse h et théorème de girsanov. In Grossissements de filtrations: exemples et applications, pages 15–35. Springer, 1985.
- [71] S. Janson. Gaussian Hilbert Spaces. Cambridge Tracts in Mathematics. Cambridge University Press, 1997.
- [72] O. Kallenberg. Random Measures. Academic Press, 3rd edition, 1983.
- [73] I. Karatzas, D. L. Ocone, and J. L. Jinlu. An extension of Clark’s formula. Stochastics: An International Journal of Probability and Stochastic Processes, 37(3):127–131, 1991.
- [74] S. Kerov, G. Olshanski, and A. Vershik. Harmonic analysis on the infinite symmetric group. Inventiones mathematicae, 158(3):551–642, dec 2004.
- [75] A. Kohatsu-Higa and M. Montero. Malliavin calculus in finance. In Handbook of computational and numerical methods in finance, pages 111–174. Springer, 2004.
- [76] V. S. Koroljuk and Y. V. Borovskich. Theory of  $U$ -statistics, volume 273 of Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht, 1994. Translated from the 1989 Russian original by P. V. Malyshev and D. V. Malyshev and revised by the authors.
- [77] H. Kunita. Stochastic flows and stochastic differential equations. Cambridge Univ. Press, 1988.
- [78] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus, part i. In Kiyosi Itô, editor, Stochastic Analysis, volume 32 of North-Holland Mathematical Library, pages 271 – 306. Elsevier, 1984.

- [79] 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, December 2013.
- [80] G. Last and M. Penrose. Lectures on the Poisson Process. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2017.
- [81] G. Last and M. D. Penrose. Martingale representation for poisson processes with applications to minimal variance hedging. Stochastic Processes and their Applications, 121(7):1588 – 1606, 2011.
- [82] M. Ledoux. Concentration of measure and logarithmic sobolev inequalities. In Seminaire de probabilités XXXIII, pages 120–216. Springer, 1999.
- [83] J. A León, R. Navarro, and D. Nualart. An anticipating calculus approach to the utility maximization of an insider. Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics, 13(1):171–185, 2003.
- [84] J. A. León, R. Navarro, and D. Nualart. An anticipating calculus approach to the utility maximization of an insider. Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics, 13(1):171–185, 2003.
- [85] P. Malliavin. Stochastic calculus of variations and hypoelliptic operators. Proceedings of the International Symposium on Stochastic Differential Equations, page 195–263, 1978.
- [86] P. Malliavin and A. Thalmaier. Stochastic Calculus of Variations in Mathematical Finance. Springer-Verlag Berlin Heidelberg, 1 edition, 2006.
- [87] M. Métivier. Semimartingales, volume 2 of de gruyter studies in mathematics, 1982.
- [88] P.-A. Meyer. Transformations de riesz pour les lois gaussiennes. Séminaire de probabilités de Strasbourg, 18:179–193, 1984.
- [89] P.-A. Meyer. Quantum probability for probabilists. Springer, 2006.
- [90] G. Mokobodzki. Sur l’algebre contenue dans le domaine etendu d’un generateur infinitesimal. In Séminaire de Théorie du Potentiel Paris, No. 3, pages 168–187. Springer, 1978.
- [91] I. Nourdin, Giovanni G. Peccati, and X. Yang. Restricted hypercontractivity on the poisson space. Proceedings of the American Mathematical Society, page 1, Dec 2019.
- [92] I. Nourdin and G. Peccati. Noncentral convergence of multiple integrals. The Annals of Probability, 37(4):1412–1426, 2009.
- [93] I. Nourdin and G. Peccati. Stein’s method on wiener chaos. Probability Theory and Related Fields, 145(1-2):75–118, 2009.
- [94] I. Nourdin and G. Peccati. Normal Approximations with Malliavin Calculus: From Stein’s Method to Universality. Cambridge University Press, 2012.

- [95] I. Nourdin, G. Peccati, G. Poly, and S. Rosaria. Classical and free fourth moment theorems: Universality and thresholds. Journal of Theoretical Probability, 29(2):653–680, Jan 2015.
- [96] I. Nourdin, G. Peccati, and G. Reinert. Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos. The Annals of Probability, 38(5):1947–1985, 2010.
- [97] I. Nourdin, G. Peccati, and A. Réveillac. Multivariate normal approximation using stein’s method and malliavin calculus. Ann. Inst. H. Poincaré Probab. Statist., 46(1):45–58, 02 2010.
- [98] D. Nualart. The Malliavin Calculus and Related Topics. Springer-Verlag, 1995.
- [99] D. Nualart and É. Pardoux. Stochastic calculus with anticipating integrands. Probability Theory and Related Fields, 78(4):535–581, 1988.
- [100] D. Nualart and G. Peccati. Central limit theorems for sequences of multiple stochastic integrals. Ann. Probab., 33(1):177–193, 01 2005.
- [101] D. Nualart and B. Saussereau. Malliavin calculus for stochastic differential equations driven by a fractional Brownian motion. Stochastic Processes and their Applications, 119(2):391 – 409, 2009.
- [102] D. Nualart and W. Schoutens. Chaotic and predictable representations for lévy processes. Stochastic Processes and their Applications, 90(1):109 – 122, 2000.
- [103] D. Nualart and J. Vives. Anticipative calculus for the Poisson process based on the Fock space. In Séminaire de probabilités XXIV, pages 154–165. Springer-Verlag, 1988.
- [104] G. Di Nunno. Stochastic integral representations, stochastic derivatives and minimal variance hedging. 2002.
- [105] G. Di Nunno, T. Meyer-Brandis, B. Oksendal, and F. Proske. Malliavin calculus and anticipative Itô formulae for Lévy processes. Infinite Dimensional Analysis, Quantum Probability and Related Topics, 08(02):235–258, 2005.
- [106] G. Di Nunno, B. Oksendal, and F. Proske. White noise analysis for lévy processes. Journal of Functional Analysis, 206:109–148, 2004.
- [107] N. Obata. White noise calculus and Fock space. Springer, 2006.
- [108] D. L. Ocone and I. Karatzas. A generalized clark representation formula, with application to optimal portfolios. Stochastics and Stochastic Reports, 34(3-4):187–220, 1991.
- [109] B. Oksendal and F. Proske. White noise of poisson random measures. Potential Analysis, 21:375–403, 2004.
- [110] G. Peccati and M. Reitzner. Stochastic analysis for Poisson point processes: Malliavin calculus, Wiener-Itô chaos expansions and stochastic geometry, volume 7. Springer, 2016.

- [111] G. Peccati, J. L. Solé, M. S. Taqqu, and F. Utzet. Stein's method and Normal approximation of Poisson functionals. Ann. Probab., 38(2):443–478, 03 2010.
- [112] E. A. Peköz and Adrian A. Röllin. New rates for exponential approximation and the theorems of rényi and yaglom. Ann. Probab., 39(2):587–608, 03 2011.
- [113] E. A. Peköz, A. Röllin, and N. Ross. Total variation error bounds for geometric approximation. Bernoulli, 19(2):610–632, 05 2013.
- [114] E. A. Peköz, A. Röllin, and N. Ross. Generalized gamma approximation with rates for urns, walks and trees. Ann. Probab., 44(3):1776–1816, 05 2016.
- [115] I. Pikovsky and I. Karatzas. Anticipative portfolio optimization. Advances in Applied Probability, 28(4):1095–1122, 1996.
- [116] N. Privault. Chaotic and variational calculus in discrete and continuous time for the Poisson process. Stoch. & Stoch. Rep., 51:83–109, 1994.
- [117] N. Privault. Stochastic Analysis in Discrete and Continuous Settings. 1982. Springer-Verlag Berlin Heidelberg, 2009.
- [118] N. Privault and A. Réveillac. Stein estimation for the drift of gaussian processes using the malliavin calculus. The Annals of Statistics, 36(5):2531–2550, 2008.
- [119] N. Privault and W. Schoutens. Discrete chaotic calculus and covariance identities. Stochastics and Stochastic Reports, 72(3-4):289–316, 2002.
- [120] N. Privault, J. L. Solé, and J. Vives. Chaotic kabanov formula for the azéma martingales. Bernoulli, 6(4):633–651, 2000.
- [121] N. Privault and X. Wei. A Malliavin calculus approach to sensitivity analysis in insurance. Insurance: Mathematics and Economics, 35(3):679–690, 2004.
- [122] Nicolas Privault. Chapter 19 stochastic calculus for jump processes. 2018.
- [123] M. Reitzner and M. Schulte. Central limit theorems for  $U$ -statistics of Poisson point processes. The Annals of Probability, 41(6):3879–3909, 2013.
- [124] W. T. Rhee and M. Talagrand. Martingale inequalities and the jackknife estimate of variance. Statist. Probab. Lett., 4(1):5–6, 1986.
- [125] N. Ross. Fundamentals of Stein's method. Probability Surveys, 8(0):210–293, 2011.
- [126] W. Runggaldier. Portfolio optimization in discrete time. Accademia delle Scienze dell'Istituto di Bologna, 2006.
- [127] F. Russo and P. Vallois. Forward, backward and symmetric stochastic integration. Probability theory and related fields, 97(3):403–421, 1993.
- [128] M. Sanz-Solé. Malliavin calculus: With applications to stochastic partial differential equations. 01 2005.
- [129] M. Schweizer. Mean-variance hedging for general claims. The Annals of Applied Probability, pages 171–179, 1992.

- [130] M. Schweizer. Variance-optimal hedging in discrete time. Mathematics of Operations Research, 20(1):1–32, 1995.
- [131] A. V. Skorohod. On a generalization of a stochastic integral. Theory Probab. Appl., 20:219–233, 1975.
- [132] 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, pages 583–602, Berkeley, Calif., 1972. University of California Press.
- [133] D. W. Stroock. The malliavin calculus, a functional analytic approach. Journal of Functional Analysis, 44:212–257, 1981.
- [134] M. Talagrand. Transportation cost for gaussian and other product measures. Geometric and Functional Analysis GAFA, 6(3):587–600, 1996.
- [135] C. A. Tudor and F. G. Viens. Variations and estimators for self-similarity parameters via malliavin calculus. The Annals of Probability, 37(6):2093–2134, 2009.
- [136] A. S. Üstünel and M. Zakai. The construction of filtrations on abstract Wiener space. J. Funct. Anal., 143(1):10–32, 1997.
- [137] A.S. Üstünel and M. Zakai. Transformations of measure on Wiener spaces. Monographs in Mathematics. Springer-Verlag, 2000.
- [138] S. Watanabe, M. G. Nair, and B. Rajeev. Lectures on stochastic differential equations and Malliavin calculus, volume 73. Springer-Verlag for the Tata Institute of Fundamental Research, 1984.
- [139] D. Williams. Probability with martingales. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.
- [140] L. Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. Probability Theory and Related Fields, 118(3):427–438, 2000.
- [141] M. Yor. Entropie d’une partition, et grossissement initial d’une filtration. In Grossissements de filtrations: exemples et applications, pages 45–58. Springer, 1985.
- [142] K. Yosida. Functional analysis. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.
- [143] G. Zheng. Normal approximation and almost sure central limit theorem for non-symmetric rademacher functionals. 2017.

### Résumé en français

Le calcul de Malliavin est un calcul des variations stochastiques en dimension infinie initialement élaboré sur l'espace de Wiener et étendu par la suite à d'autres familles de processus tels que les processus de Poisson, de Rademacher, de Lévy. La multiplicité des approches et la variété des espaces canoniques sur lesquels il opère semblent dissuader l'éventuel projet d'un travail unificateur. Il demeure néanmoins possible d'identifier une terminologie commune à tous ces formalismes autour des notions d'opérateurs de Malliavin (gradient, divergence, semi-groupe d'Ornstein-Uhlenbeck) et de la relation fondamentale liant l'opérateur gradient et la divergence (défini comme adjoint du gradient) : *la formule d'intégration par parties*.

On élabore dans cette thèse un calcul de Malliavin pour deux classes de processus discrets : les suites de variables aléatoires indépendantes (non nécessairement de même loi) et les processus géométriques composés.

Dans le premier cadre, on équipe tout produit dénombrable d'espaces de probabilité d'une structure de Dirichlet-Malliavin discrète échafaudée sur une famille d'opérateurs de Malliavin (gradient discret, divergence, opérateur nombre), une formule d'intégration par parties, et les formes de Dirichlet naturellement induites dans ce contexte. On obtient les analogues discrets des principales identités fonctionnelles (formule de Clark-Ocone, identité de covariance, inégalité de log-Sobolev, de déviation...) établies pour les processus Brownien et de Poisson et dont on retrouve par ailleurs les structures de Dirichlet usuelles comme limites de celles induites par notre formalisme.

En exploitant la combinaison de la formule d'intégration par parties et de la méthode de Stein, comme l'ont conceptualisée I. Nourdin et G. Peccati, on établit des critères de Stein-Malliavin pour les approximations Normale et Gamma par toute fonctionnelle de variables aléatoires indépendantes. Ce calcul de Malliavin discret généralise ce qui avait été théorisé dans le cas Rademacher et donne un cadre unificateur à de nombreux résultats disséminés dans la littérature suggérant également que des résultats sans lien apparent (Efron-Stein, paires échangeables etc.) sont en réalité les faces d'une même pièce.

La motivation du second travail provient de l'impossibilité constatée d'établir (via le précédent formalisme) une formule de Ocone-Karatzas pour le modèle trinomial. On lui substitue un modèle ternaire, sous-tendu par un processus géométrique composé (défini par une famille de variables i.i.d. à valeurs dans  $\{-1, 0, 1\}$ ) et qui lui est équivalent en loi. Pour des raisons techniques énoncées par M. Émery, il n'existe pas de décomposition chaotique. On en donne une version modifiée au moyen d'intégrales multiples relatives à une famille orthogonalisable de variables aléatoires. Les opérateurs gradient et divergence sont alors définis comme les opérateurs d'annihilation et de création agissant sur cette décomposition et vérifient une formule d'intégration par parties. Des analogues géométriques à la formule d'Itô, au théorème de Girsanov, et à la formule de représentation de martingale, et la formule de Ocone-Karatzas sont établis.

Ce formalisme est appliqué au calcul d'utilité dans un modèle ternaire où agissent un agent ordinaire et un initié disposant d'un surplus d'information dès le début de la période de trading. En incorporant les techniques d'élargissement de filtration en temps discret développées par C. Blanchet-Scalliett, M. Jeanblanc et R. Romero à notre construction, on définit l'équivalent du drift d'information dans ce cadre et on montre

alors que l'utilité logarithmique espérée additionnelle de l'initié au temps  $t$  peut être exprimée comme l'entropie relative de la mesure initiale par rapport à la mesure préservant les martingales sur  $\llbracket 0, t \rrbracket$ . On retrouve ainsi l'exact analogue du résultat établi par J. Amendinger, P. Imkeller et M. Schweizer dans le cas continu.

### Summary in english

Malliavin calculus is an infinite-dimensional variational calculus initially developed on the Wiener space. Further extended to other classes of processes, such as Poisson, Rademacher, or Lévy processes, it has reached a certain maturity. Cause of the multiplicity of approaches and the variety of canonical spaces, it seems to be difficult to propose a unifying work. We can nevertheless identify a terminology common to all these formalisms; this lies around the notions of Malliavin operators (gradient, divergence, Ornstein-Uhlenbeck semi-group) and the fundamental relationship between the operator gradient and the divergence (defined as the adjoint of the gradient): *the integration by parts formula*.

In this thesis work, we propose to develop a Malliavin calculus for two classes of discrete processes: sequences of independent random variables (not necessarily identically distributed) and compound geometric processes. Both are wholly or partly motivated from an application perspective.

In the first mentioned framework, we equip any countable product of probability spaces with a discrete Dirichlet-Malliavin structure. This lies on a family of Malliavin operators (discrete gradient, divergence, number operator), an integration by parts formula, and the induced Dirichlet forms. We obtain the discrete analogues of the classical functional identities (Clark-Ocone formula, covariance identity, deviation inequality...) established for Brownian and Poisson processes. Besides, we retrieve the usual Poisson and Brownian Dirichlet structures associated to their respective gradient, as limits of the structures induced by our formalism. Re-using the combination of Malliavin calculus and Stein's method designed by I. Nourdin and G. Peccati, we provide Stein-Malliavin criterions for the Normal and the Gamma approximations by functionals of independent random variables.

This discrete Dirichlet-Malliavin structure seems to give a unified framework for many results scattered in the literature so that we hope to bring new insights on why these apparently disjoint results (Efron-Stein, exchangeable pairs, etc.) are in fact multiple sides of the same coin.

Faced with the impossibility to state (via the discrete Malliavin calculation) an Ocone-Karatzas formula for the trinomial model, we replace it by a *ternary model*; equivalent in law to the initial one, it is supported by a compound geometric process (with  $\{-1, 0, 1\}$  valued i.i.d. random variables). For technical reasons outlined by M. Émery, there is no chaotic decomposition. We state a *modified* decomposition in terms of multiple integrals with respect to an orthogonalizable family of random variables. The gradient and divergence operators are then defined as the annihilation and creation operators acting on this decomposition and verify an integration formula by parts. Our approach is close to the one developed for (compound) Poisson processes, with which the (compound) geometric process shares many properties (as a jump process). We give



the analogues of Itô's formula, Girsanov's theorem, martingale representation formula, and Ocone-Karatzas formula for compound geometric processes.

Then, we make use of this new formalism to compute the additional utility of an insider in the new ternary model. The insider benefits from an extra information hidden in a random variable from the start of trading period. Using the toolbox of C. Blanchet-Scalliet, M. Jeanblanc and R. Romero as for the enlargement of filtrations in discrete settings, we define an analogue of the *information drift* and express it in terms of our geometric gradient. We show that the insider's additional expected logarithmic utility at time  $t$  can be expressed as the relative entropy of the initial measure with respect to the *martingale preserving measure* on  $[[0, t]]$ . We retrieve exactly the result of J. Amendinger, P. Imkeller and M. Schweizer stated in the continuous case.

**Titre :** Calcul de Malliavin et structures de Dirichlet pour des variables aléatoires indépendantes.

**Mots clés :** Calcul de Malliavin, structures de Dirichlet, méthode de Stein, modèles financiers discrets, variables aléatoires indépendantes, processus géométriques composés.

**Résumé :** Cette thèse porte sur le calcul de Malliavin dont on munit deux cadres discrets. On équipe tout produit dénombrable d'espaces de probabilités d'une structure de Dirichlet-Malliavin au moyen d'opérateurs (gradient, divergence, opérateur nombre), d'une formule d'intégration par parties, et des formes de Dirichlet induites. On obtient les analogues discrets aux identités fonctionnelles classiques des processus Brownien et Poisson dont les structures de Dirichlet s'écrivent comme limites des structures induites par notre formalisme. Des critères de Stein-Malliavin discrets sont établis pour les approximations Normale et Gamma. Le second cadre est celui d'un

modèle financier ternaire sous-tendu par un processus géométrique composé à trois points, et équivalent en loi au modèle trinomial. Toute fonctionnelle de ce processus géométrique composé de carré intégrable possède un développement en chaos "modifié" sur lequel agissent des opérateurs d'annihilation/gradient et de création/divergence vérifiant en outre une formule de commutation généralisée. S'ensuit de la formule de Clark "géométrique" qu'il est alors possible d'établir, une formule de hedging pour l'initié dont l'utilité additionnelle espérée s'exprime en termes d'entropie relative, comme dans le cas continu.

**Title :** Malliavin calculus and Dirichlet structures for independent random variables.

**Keywords :** Malliavin calculus, Dirichlet structures, Stein's method, discrete financial models, independent random variables, compound geometric processes.

**Abstract :** Malliavin calculus was initially developed to provide an infinite-dimensional variational calculus on the Wiener space and further extended to other spaces. In this work, we develop such one in two discrete frameworks. First, we equip any countable product of probability spaces with a discrete Dirichlet-Malliavin structure, consisting of a family of Malliavin operators (gradient, divergence, number operator), an integration by parts formula, and the induced Dirichlet forms. We get the analogues of the classical functional identities and retrieve the usual Poisson and Brownian Dirichlet structures as limits of our induced struc-

tures. We provide discrete Stein-Malliavin criterions for the Normal and the Gamma approximations. Second, we study insider's trading in a ternary model, lying on a three-points compound geometric process. We state a modified chaotic decomposition and define the geometric gradient and divergence operators as the annihilation and creation operators acting on it. We state a geometric Ocone-Karatzas formula. We express the insider's additional expected logarithmic utility in terms of relative entropy as in the continuous case.