Statistical estimation and limit theorems for Gaussian fields using the Malliavin calculus

Anthony Réveillac

Laboratoire Mathématiques, Image et Applications
Ecole doctorale SPIA (Université de La Rochelle)

Thèse

présentée pour l'obtention du titre de

Docteur de l'Université de La Rochelle

en Mathématiques

spécialité : Probabilités et Statistiques

par

Anthony Réveillac

dirigée par

Pr. Nicolas Privault

Estimation statistique et théorèmes limites pour les champs gaussiens par le calcul de Malliavin

Soutenue publiquement le 11 octobre 2008 à La Rochelle devant le jury composé de :

Dr. Jean-Christophe Breton	Examinateur	Université de La Rochelle
Pr. Laure Coutin	Examinatrice	Université Paris V
Pr. Marco Dozzi	Rapporteur	Université Nancy II
Pr. Valentine Genon-Catalot	Président du jury	Université Paris V
Pr. Peter Imkeller	Rapporteur	Humboldt Universität zu Berlin
Dr. Ivan Nourdin	Examinateur invité	Université Paris VI
Pr. David Nualart	Rapporteur	University of Kansas
Pr. Nicolas Privault	Directeur de thèse	City University of Hong Kong

Table des matières

Introduction **1**
Chapitre 1 : Estimation de Stein 1
Chapitre 2 : Estimation bayésienne pour les canaux de Poisson 13
Chapitre 3 : TLC pour les variations quadratiques à poids de champs gaussiens 16
Chapitre 4 : Méthode de Stein pour l'approximation gaussienne multidimensionnelle 25

1 Estimation de Stein **29**
1.1 Cas d'une martingale gaussienne à accroissements indépendants 29
1.1.1 Notation and preliminaries 29
1.1.2 Maximum likelihood and Bayes estimators 30
1.1.3 Superefficient drift estimators 31
1.1.4 Numerical application 33
1.2 Cas d'un processus gaussien 35
1.2.1 Introduction 35
1.2.2 Efficient drift estimator 40
1.2.3 Bayes estimators 45
1.2.4 Malliavin calculus on Gaussian space 49
1.2.5 Superefficient drift estimators 51
1.2.6 Superharmonic functionals 54
1.2.7 Numerical application 59
1.2.8 Appendix 64
1.3 Cas d'un processus de Poisson 67
1.3.1 Introduction 67
1.3.2 Preliminaries 69
1.3.3 Analysis on the Poisson space 73
1.3.4 Stein estimators 78
1.3.5 Examples 82

2 Estimation bayésienne pour les canaux de Poisson **87**
2.1 Introduction 87
2.2 Integer-valued Poisson channel 89
2.3 Analysis on the Poisson space 90
2.4 Conditional mean estimators for Poisson channels 92
2.5 A generalization to a class of non-Gaussian channels 96

2.6 Appendix 97

3 TLC pour les variations quadratiques à poids de champs gaussiens 99

3.1 Estimation de la variation quadratique de certains processus de diffusions à deux paramètres 99

3.1.1 Introduction 99

3.1.2 Stochastic analysis of two-parameter processes 101

3.1.3 Central limit theorem 103

3.1.4 Estimation of the quadratic variation and asymptotic normality of the estimator 112

3.1.5 Appendix 117

3.2 Théorèmes limites pour les variations quadratiques à poids de certains draps browniens fractionnaires 121

3.2.1 Introduction 121

3.2.2 Preliminaries and notations 122

3.2.3 Convergence of finite-dimensional laws 124

3.2.4 Useful lemmas 132

3.3 Variations quadratiques à poids du mouvement brownien fractionnaire: le cas limite $H = 1/4$ 141

3.3.1 Introduction 141

3.3.2 Preliminaries and notations 145

3.3.3 Proof of Theorem 3.3.1 147

3.3.4 Proof of Theorem 3.3.2 160

4 Méthode de Stein pour l'approximation gaussienne multidimensionnelle 169

4.1 Introduction 169

4.2 Preliminaries and notations 172

4.3 Stein's method and Gaussian vectors 174

4.4 Applications 182

4.4.1 Convergence of marginal distributions in the functional Breuer-Major CLT 182

4.4.2 Vector-valued functionals of finite Gaussian sequences 184

Bibliographie 187

Remerciements

Tout d'abord je remercie énormément Nicolas Privault de m'avoir fait confiance il y a trois ans et de m'avoir encouragé durant les moments difficiles. Je lui suis infiniment reconnaissant de sa patience et de l'énergie qu'il a investie dans cet encadrement.

Je remercie également beaucoup Jean-Christophe Breton d'avoir toujours pris le temps de répondre à mes questions parfois naïves et de m'avoir sans cesse encouragé jusqu'à maintenant.

Je remercie également beaucoup Marco Dozzi, Peter Imkeller et David Nualart qui me font l'honneur de rapporter cette thèse ainsi que Valentine Genon-Catalot qui me fait l'honneur de présider ce jury de thèse et Laure Coutin et Ivan Nourdin qui me font l'honneur de participer au jury de cette thèse.

Je remercie bien évidemment ma famille qui m'a épaulé et encouragé durant toutes mes études.

Je remercie bien évidemment Ivan Nourdin, Giovanni Peccati, Ciprian Tudor et Frederi Viens pour l'intérêt qu'ils portent à mon travail, pour leur disponibilité, pour leur gentillesse et pour leurs encouragements. Je pense par exemple à l'invitation à venir réaliser un exposé dans le groupe de travail "Aspects fractals" ou encore de m'avoir proposé de réaliser avec eux certains travaux présents dans cette thèse.

Je dis également un grand "Merci" à Stefan Ankirchner, Marc Arnaudon, Caroline Chaux, Laure Coutin, Marco Dozzi, Piotr Graczyk, Peter Imkeller, Aldéric Joulin et à Ivan Nourdin pour m'avoir invité à venir faire un exposé ou à venir travailler dans leurs universités respectives. Je remercie Antoine Ayache, Hermine Biermé, Serge Cohen, Djalil Chafaï, Laurent Decreusefond, Anne Estrade, Valentine Genon-Catalot, Mihai Gradinaru, Jean Jacod et Monique Pontier pour les conseils et commentaires concernant mon travail lors de nos rencontres durant des conférences (je présente mes excuses à celles et ceux que j'oublierais en écrivant ces lignes).

Je remercie Michel Berthier, Directeur du laboratoire Mathématiques Image et Applications de La Rochelle, de son soutien durant ces trois ans. Je remercie aussi Aldéric Joulin, Yutao Ma. Une pensée pour Delphine Féral, Nathalie Krell, Mathieu Rosenbaum, les doctorants et moniteurs de La Rochelle, mes amis et tous ceux et celles à qui j'ai déjà témoigné ma reconnaissance et qui se reconnaîtront. Je remercie les membres du département de mathématiques et les autres membres du laboratoire Mathématiques Image et Applications de La Rochelle.

Je remercie tout particulièrement mon ami de toujours Benjamin Laquerrière pour sa fidélité et pour son soutien.

A mon père et à mes grands-parents à qui la vie n'a pas laissé le temps de lire ne serait-ce que *le quart De la moitié du commencement d'une* de ces lignes...

Résumé

Dans cette thèse nous appliquons le calcul de Malliavin à l'estimation statistique de paramètres de certains processus stochastiques et à l'obtention de théorèmes de la limite centrale pour les variations quadratiques à poids de processus fractionnaires et/ou à deux paramètres ainsi qu'à l'approximation gaussienne de mesures de probabilités multidimensionnelles. Dans le Chapitre 1 nous construisons des estimateurs de type Stein pour la dérive de processus gaussiens et pour l'intensité de processus de Poisson. Dans le Chapitre 2 nous calculons l'estimateur bayésien du signal d'entrée d'un canal de Poisson et nous étendons notre résultat aux canaux dont le bruit est une martingale normale possédant la propriété de représentation chaotique. Dans le Chapitre 3 nous établissons des théorèmes de la limite centrale pour les variations quadratiques à poids du drap brownien standard (nous permettant de donner un estimateur asymptotiquement normal de la variation quadratique de certains processus de diffusion à deux paramètres) puis pour celles de certains draps browniens fractionnaires. Dans ce même chapitre nous établissons un théorème de la limite centrale pour les variations quadratiques à poids du mouvement brownien fractionnaire d'indice $H = 1/4$ nous permettant de donner le comportement asymptotique des sommes de Riemann à signe alterné associées au mouvement brownien fractionnaire d'indice $H = 1/4$. Enfin dans le Chapitre 4 nous appliquons la méthode de Stein et du calcul de Malliavin afin d'obtenir des bornes explicites pour l'approximation gaussienne multidimensionnelle de fonctionnelles de champs gaussiens. Nous appliquons en particulier nos résultats aux théorème de la limite centrale de Breuer et Major pour des champs associés à un mouvement brownien fractionnaire.

Abstract

In this thesis we apply the Malliavin calculus to statistical estimation of parameters of stochastic processes and to derive limit theorems for the weighted quadratic variations of one or two-parameter fractional processes and to multidimensional normal approximation of probability measures. In Chapter 1 we construct Stein type estimators for the drift of Gaussian processes and for the intensity of Poisson processes. In Chapter 2, we compute the Bayesian estimator of the input of a Poisson channel then extended to normal martingales with chaotic representation property channels. In Chapter 3 we derive central limit theorems for the weighted quadratic variations of the standard Brownian sheet (applied then to the obtaining of an asymptotically normal estimator of the quadratic variation of some two-parameter diffusion processes) and of some fractional Brownian sheets. Then in this chapter we establish a central limit theorem for the weighted quadratic variations of the fractional Brownian motion with Hurst index $H = 1/4$ leading to the study of the asymptotic behavior of the Riemann sums with alternating signs associated to the fractional brownian motion with Hurst index $H = 1/4$. Finally in Chapter 4 we apply Stein's method and the Malliavin calculus in order to obtain explicit bounds in the multidimensional normal approximation of functionals of gaussian fields. In particular we provide an application to a functional version of the Breuer-Major TCL for fields subordinated to a fractional Brownian motion.

Introduction

Chapitre 1 : Estimation de Stein

Le Chapitre 1 est constitué de travaux réalisés en collaboration avec Nicolas Privault.

Brève introduction aux estimateurs de Stein

Soit X un vecteur gaussien de $\mathbb{R}^d$ de moyenne inconnue $\mu = (\mu_1, \dots, \mu_d)$ et de matrice de covariance $\sigma^2 \mathbf{I}_d$ (où σ est une constante strictement positive supposée connue) sous une mesure de probabilité $\mathbb{P}_\mu$. Un problème classique en estimation paramétrique est de donner un estimateur de la moyenne inconnue μ.

Définition 1. *Un estimateur ξ de μ est une variable aléatoire à valeurs dans $\mathbb{R}^d$ mesurable par rapport à la tribu engendrée par X. De plus, ξ est dit sans biais si*

$$\mathbb{E}_{\mathbb{P}_\mu}[\xi] = \mu, \quad \forall \mu \in \mathbb{R}^d.$$

Par exemple le vecteur gaussien lui-même ($\xi := X$) est un estimateur sans biais de sa propre moyenne μ. C'est un exemple très important (bien que très simple) dont nous allons largement discuter dans la suite de cette introduction.

Afin de pouvoir comparer la performance des différents estimateurs proposés, on introduit la fonction de risque quadratique R.

Définition 2. *Le risque quadratique d'un estimateur $\xi = (\xi_1, \dots, \xi_d)$ au point μ est défini par*

$$R(\xi, \mu) := \mathbb{E}_{\mathbb{P}_\mu}[\|\xi - \mu\|^2_{\mathbb{R}^d}] = \sum_{i=1}^{d} \mathbb{E}_{\mathbb{P}_\mu}[|\xi_i - \mu_i|^2], \quad \mu \in \mathbb{R}^d.$$

Remarquons dès lors que l'estimateur $\xi := X$ est un estimateur de risque constant égal à $\sigma^2 d$. On rappelle également que l'on peut maximiser la vraisemblance (ici la densité gaussienne) et l'on obtient que l'estimateur du maximum de vraisemblance (noté par la suite MLE) est égal à X. Ainsi le vecteur X est un estimateur sans biais de risque constant et est le MLE de sa propre moyenne μ. Enfin, on peut montrer que le MLE X est un estimateur *minimax* ce qui signifie qu'il minimise sur la classe de tous les estimateurs le risque maximum en μ,

$$X = argmin_\xi \sup_{\mu \in \mathbb{R}^d} R(\xi, \mu).$$

Ce minimum n'est en général pas unique puisque tout estimateur de risque inférieur ou égal à un estimateur minimax est également minimax.

D'où cette question naturelle : "peut-on trouver un *meilleur* estimateur que l'estimateur du maximum de vraisemblance X au sens suivant ?"

Définition 3. *Un estimateur ξ_1 est dit meilleur qu'un estimateur ξ_2 si*
- $R(\xi_1, \mu) \leq R(\xi_2, \mu), \quad \forall \mu \in \mathbb{R}^d$,
- *Il existe μ_0 dans $\mathbb{R}^d$ tel que $R(\xi_1, \mu_0) < R(\xi_2, \mu_0)$.*

De plus ξ est dit admissible si aucun estimateur n'est meilleur que lui.

Une réponse partielle est donnée par l'inégalité de Cramer-Rao qui nous permet de répondre (par la négative) à la question intermédiaire : "existe-t-il un estimateur *sans biais* meilleur que le MLE X ?"

Proposition 4 (Inégalité de Cramer-Rao)**.** *Soit ξ un estimateur sans biais de μ. Alors, il existe une constante CRB appelée borne de Cramer-Rao telle que*

$$R(\xi, \mu) \geq CRB, \quad \forall \mu \in \mathbb{R}^d$$

et $R(X, \mu) = CRB = \sigma^2 d, \quad \forall \mu \in \mathbb{R}^d$.

Ainsi le risque de tout estimateur sans biais est minoré par la constante CRB qui est précisément le risque constant du MLE X. En ce sens le MLE X est un estimateur *efficace* puisque de risque minimum parmi la classe des estimateurs sans biais. Par conséquent, un estimateur meilleur que X est à chercher parmi les estimateurs biaisés.

En 1956, Stein donne dans [122] un estimateur de la forme

$$\xi_i = (1 - b(a + X_i^2)^{-1})X_i, \quad i \in \{1, \ldots, d\} \tag{1}$$

dont le risque est plus faible que celui du MLE X dès que la dimension d est supérieure ou égale à 3, montrant ainsi que dans ce cas le MLE X n'est pas un estimateur admissible. De plus Stein montre que pour $d \leq 2$ le MLE X est admissible. Cette dimension critique $d = 3$ est connue sous le nom de phénomène de Stein.

L'estimateur défini par (1) est un estimateur *suroptimal* puisqu'il a un risque plus faible que l'estimateur efficace X. Il devient donc intéressant de pouvoir donner une construction de tels estimateurs. La première étape en ce sens fut réalisée par James et Stein en 1961 dans [68] où les auteurs améliorent l'estimateur (1) en montrant que l'estimateur de James-Stein ξ_{JS} est suroptimal dès que $d \geq 3$, avec

$$\xi_{JS} := X \left(1 - \frac{d-2}{\|X\|_{\mathbb{R}^d}^2}\right). \tag{2}$$

La seconde étape est franchie vingt ans plus tard par Stein dans l'article [124] dans lequel l'auteur décrit une méthode de construction d'estimateurs suroptimaux contenant l'estimateur de James-Stein. Cette dernière repose sur la formule d'intégration par parties pour un vecteur gaussien et sur l'étude des fonctions surharmoniques positives sur $\mathbb{R}^d$. Voici les principaux arguments à la base de cette construction.

Considérons un estimateur ξ_g de la forme

$$\xi_g := X + g(X)$$

où $g : \mathbb{R}^d \to \mathbb{R}^d$ est une fonction suffisamment régulière. Le risque de ξ_g est donné par,

$$\begin{aligned} R(X + g(X), \mu) &= \mathbb{E}_{\mathbb{P}_\mu}\left[\|X + g(X) - \mu\|^2_{\mathbb{R}^d}\right] \\ &= \mathbb{E}_{\mathbb{P}_\mu}\left[\|X - \mu\|^2_{\mathbb{R}^d}\right] + \mathbb{E}_{\mathbb{P}_\mu}\left[\|g(X)\|^2_{\mathbb{R}^d}\right] + 2\sum_{i=1}^{d} \mathbb{E}_{\mathbb{P}_\mu}\left[(X_i - \mu_i) g_i(X)\right] \end{aligned}$$

où g_i est la i-ème coordonnée de g. Le premier terme de la partie droite de l'égalité précédente est exactement la borne de Cramer-Rao CRB. Stein calcule le troisième terme au moyen de la formule d'intégration par parties pour la loi gaussienne. On obtient donc,

$$\mathbb{E}_{\mathbb{P}_\mu}\left[\|X + g(X) - \mu\|^2_{\mathbb{R}^d}\right] = CRB + \mathbb{E}_{\mathbb{P}_\mu}\left[\|g(X)\|^2_{\mathbb{R}^d}\right] + 2\sigma^2 \sum_{i=1}^{d} \mathbb{E}_{\mathbb{P}_\mu}\left[\nabla g_i(X)\right]. \quad (3)$$

Supposons maintenant que g est choisie de la forme $g := \sigma^2 \nabla \log f$ où f est une fonction strictement positive $f : \mathbb{R}^d \to \mathbb{R}_+$. Le risque ci-dessus pour ce choix particulier de fonction g se simplifie et devient

$$\begin{aligned} \mathbb{E}_{\mathbb{P}_\mu}\left[\|X + g(X) - \mu\|^2_{\mathbb{R}^d}\right] &= CRB + 4\sigma^2 \sum_{i=1}^{d} \mathbb{E}_{\mathbb{P}_\mu}\left[\frac{\partial_i^2 \sqrt{f(X)}}{\sqrt{f(X)}}\right] \\ &= \sigma^2 d + 4\sigma^2\, \mathbb{E}_{\mathbb{P}_\mu}\left[\frac{\Delta\sqrt{f(X)}}{\sqrt{f(X)}}\right]. \end{aligned}$$

Ainsi si f est une fonction strictement positive sur $\mathbb{R}^d$ et surharmonique sur $\mathbb{R}^d$, *i.e.*, $\Delta\sqrt{f}(x) < 0,\ dx - p.p.$ alors l'estimateur $X + \sigma^2 \nabla \log f(X)$ (appelé estimateur de Stein) est un estimateur suroptimal.

Remarques 5.
- *Une fonction (non-constante) positive et surharmonique sur $\mathbb{R}^d$ ne peut exister que pour $d \geq 3$ respectant ainsi l'admissibilité du MLE X en dimension 1 et 2.*
- *L'estimateur de James-Stein est un cas particulier de cette construction avec $f(x) := \|x\|^{d-2}_{\mathbb{R}^d}$.*

On peut résumer les principaux "ingrédients" de cette construction ainsi.
1) La formule d'intégration par parties pour la gaussienne.
2) La propriété de dérivation du gradient sur $\mathbb{R}^d$, permettant d'écrire les deux derniers termes de (3) comme $4\sigma^2\, \mathbb{E}_{\mathbb{P}_\mu}\left[\frac{\Delta\sqrt{f(X)}}{\sqrt{f(X)}}\right]$ réduisant le problème de minimisation du risque à un problème d'analyse harmonique sur $\mathbb{R}^d$.

3) Une littérature abondante en analyse harmonique rendant possible la construction des estimateurs de Stein.

Notre construction d'estimateurs de Stein dans un contexte infini-dimensionnel dans le Chapitre 1 tiendra compte de ces contraintes et nous en expliquerons les conséquences dans la suite de cette introduction.

Averkamp et Houdré ont généralisé dans [6] la formule d'intégration par parties aux lois infiniment divisibles. A chaque loi est associé un opérateur différentiel (par exemple le gradient pour la loi normale). Cependant cet opérateur ne satisfait pas toujours la propriété de dérivation nous interdisant par conséquent une construction semblable à celle de Stein. Considérons par exemple le cas de la loi de Poisson. Soit Y une variable aléatoire de Poisson de paramètre inconnu λ sous une mesure de probabilité $\mathbb{P}_\lambda$. Soit $g : \mathbb{N} :\longrightarrow \mathbb{R}$. Un calcul direct donne,

$$\mathbb{E}_{\mathbb{P}_\lambda}[(Y-\lambda)g(Y)] = \lambda\, \mathbb{E}_{\mathbb{P}_\lambda}[\nabla_{df}\, g(Y)], \tag{4}$$

où l'opérateur ∇_{df} désigne l'opérateur aux différences défini par $\nabla_{df}\, g(n) := g(n+1) - g(n),\ n \in \mathbb{N}$ qui ne vérifie pas la propriété de dérivation. Nous détaillerons le cas de la loi et des processus de Poisson un peu plus loin dans cette introduction.

Une autre extension consiste à se placer dans un cadre gaussien infini-dimensionnel. Soit T un réel strictement positif fixé et soit $X = (X_t)_{t\in[0,T]}$ un processus gaussien centré défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ de fonction de covariance γ. Soit $u = (u_t)_{t\in[0,T]}$ une fonction continue supposée inconnue. Il existe une mesure de probabilité $\mathbb{P}_u$ sous laquelle le processus $X^u := (X_t^u := X_t - u_t)_{t\in[0,T]}$ est un processus gaussien de même loi que X sous $\mathbb{P}$. Autrement dit, sous une mesure $\mathbb{P}_u$ le processus observé X est un processus gaussien "drifté" de dérive u,

$$X = X^u + u.$$

Ce cadre constitue un problème de filtrage ou d'estimation non-paramétrique pour lequel la fonction u est alors un signal donné perturbé par un bruit gaussien centré X^u. Définissons tout d'abord une fonction de risque à travers une mesure finie et sans atome μ sur $[0,T]$. Le risque $R(\hat{u}, u)$ d'un estimateur $\hat{u}$ de u est défini par

$$R(\hat{u}, u) := \mathbb{E}_{\mathbb{P}_u}\left[\int_0^T |\hat{u}_t - u_t|^2 \mu(dt)\right].$$

Berger et Wolpert ont donné dans [14, 135] un estimateur de James-Stein pour ce problème de filtrage en ramenant l'estimation non-paramétrique de u à une estimation paramétrique à travers une décomposition de la fonction u et du processus observé X dans une base orthonormale de $L^2([0,T], d\mu)$. Berger et Wolpert choisissent un système orthonormal et total $(h_n)_{n\geq 1}$ de $L^2([0,T], d\mu)$ lié à la fonction de covariance du processus X^u sous $\mathbb{P}_u$ et obtiennent un *développement (ou une représentation) de type Karhunen-Loève* pour le processus X,

$$X = \sum_{n=1}^{+\infty} \langle X, h_n\rangle_{L^2([0,T],d\mu)} h_n$$

et

$$u = \sum_{n=1}^{+\infty} \langle u, h_n \rangle_{L^2([0,T],d\mu)} h_n.$$

Berger et Wolpert montrent ensuite que pour un élément N fixé de $\mathbb{N}^*$, le vecteur

$$\left(\langle X, h_1 \rangle_{L^2([0,T],d\mu)}, \ldots, \langle X, h_N \rangle_{L^2([0,T],d\mu)} \right)$$

est un vecteur gaussien de $\mathbb{R}^N$ de moyenne

$$\left(\langle u, h_1 \rangle_{L^2([0,T],d\mu)}, \ldots, \langle u, h_N \rangle_{L^2([0,T],d\mu)} \right)$$

et de matrice de covariance la matrice diagonale $diag(v_1, \ldots, v_N)$ où les réels $(v_k)_{k \geq 1}$ sont les valeurs propres d'un opérateur construit à partir de la fonction de covariance du processus X^u sous $\mathbb{P}_u$. Par conséquent si l'on peut estimer la moyenne du vecteur $\left(\langle X, h_i \rangle_{L^2([0,T],d\mu)} \right)_{i=1,\ldots,N}$ pour chaque N, l'on peut alors "reconstruire" u lorsque N tend vers l'infini. Notons que dans cette introduction nous avons présenté les estimateurs de Stein dans le cas d'une matrice de covariance de la forme $\sigma^2 \mathbf{I}_d$. Cependant les estimateurs de type James-Stein ont été étendu au cas où la matrice de covariance du vecteur considéré est diagonale. Berger et Wolpert montrent ensuite par un passage à la limite lorsque N tend vers l'infini que l'estimateur X est minimax et construisent un estimateur de type James-Stein défini comme limite d'estimateurs de James-Stein pour chaque dimension N de risque plus faible que celui de X. Nous renvoyons aux références [14, 135] pour les détails techniques liés à cette convergence et aux choix de la base $(h_n)_{n \geq 1}$.

Remarquons enfin que l'estimation réalisée ici ne dépend d'aucune hypothèse sur la différentiabilité de u nous plaçant dans le cas dit *non-informatif*.

Estimation de Stein de la dérive d'un processus gaussien

Dans cette introduction nous exposons la construction d'estimateurs de Stein de la dérive d'un mouvement brownien de variance σ^2. Cette construction est étendue au cas des martingales gaussiennes à accroissements indépendants dans la Section 1.1 puis dans la Section 1.2 aux processus gaussiens de la forme

$$X_t := \int_0^t K(t,s) dW_s, \quad t \in [0,T]$$

où W est un mouvement brownien standard et $K(\cdot,\cdot)$ est un noyau déterministe. Cette représentation contient en particulier le mouvement brownien fractionnaire.

Soit $X := (X_t)_{t \in [0,T]}$ un mouvement brownien de variance $\sigma^2 > 0$ défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P}^\sigma)$,

$$\mathbb{E}^\sigma[X_s X_t] = \sigma^2 (s \wedge t), \quad (s,t) \in [0,T]^2$$

où $\mathbb{E}^\sigma$ dénote l'espérance par rapport à la mesure de probabilité $\mathbb{P}^\sigma$. On note par $(\mathcal{F}_t)_{t\in[0,T]}$ la filtration engendrée par X. Soit H l'espace des fonctions de Cameron-Martin défini par

$$H := \{u : [0,T] \to \mathbb{R},\ u_t = \int_0^T \dot{u}_s ds,\ \dot{u} \in L^2([0,T], dt)\}$$

avec

$$\langle h, g\rangle_H := \sigma^2 \langle \dot{h}, \dot{g}\rangle_{L^2([0,T],dt)}, \quad (h,g) \in H^2. \tag{5}$$

Soit u un élément de H supposé inconnu. On suppose que u est un processus $(\mathcal{F}_t)_{t\in[0,T]}$-adapté. D'après le Théorème de Girsanov, il existe une mesure de probabilité $\mathbb{P}^\sigma_u$ absolument continue par rapport à la mesure de Wiener $\mathbb{P}^\sigma$ sous laquelle le processus $X^u := (X^u_t)_{t\in[0,T]}$ défini par

$$X^u_t := X_t - u_t, \quad t \in [0,T]$$

est un mouvement brownien de variance σ^2. Autrement dit, sous $\mathbb{P}^\sigma_u$ le processus X est un mouvement brownien "drifté" de dérive u et de variance σ^2. De plus la densité de Girsanov de $\mathbb{P}^\sigma_u$ par rapport à $\mathbb{P}^\sigma$ est donnée par,

$$\Lambda(u) := \frac{d\mathbb{P}^\sigma_u}{d\mathbb{P}^\sigma} = \exp\left(\int_0^T \frac{\dot{u}_s}{\sigma^2} dX_s - \frac{1}{2}\int_0^T \frac{\dot{u}_s^2}{\sigma^2} ds\right).$$

Remarque 6. *On sait que dans ce contexte l'information de dérive n'est accessible que si l'on observe un grand nombre de trajectoires de X puisque l'horizon de temps T est supposé fini. Notons également que l'hypothèse d'observation continue des trajectoires bien que non réaliste n'est pas inintéressante. En effet, les résultats obtenus dans ce cas sont les meilleurs que l'on peut espérer obtenir avec des observations discrètes.*

De même que Berger et Wolpert ont généralisé l'estimateur de James-Stein en dimension infinie, notre but est d'étendre les estimateurs de Stein pour ce cadre infini-dimensionnel. De plus, nous démontrons une inégalité de Cramer-Rao dans le cas des martingales gaussiennes à accroissements indépendants et nous redémontrons que l'estimateur X est minimax pour les processus gaussiens généraux.

Tout d'abord nous donnons une inégalité de Cramer-Rao. On note $\mathbb{E}^\sigma_u$ l'espérance sous la mesure de probabilité $\mathbb{P}^\sigma_u$.

Définition 7. *Un estimateur $\xi := (\xi_t)_{t\in[0,T]}$ est dit*

- *sans biais si,*

$$\mathbb{E}^\sigma_u[\xi_t] = \mathbb{E}^\sigma_u[u_t], \quad \forall t \in [0,T].$$

- *adapté s'il est $(\mathcal{F}_t)_{t\in[0,T]}$-adapté.*

Etant donnée une mesure finie et sans atome μ de $[0,T]$, on définit le risque de ξ à u par

$$R(\xi, u) := \mathbb{E}^\sigma_u\left[\int_0^T |\xi_t - u_t|^2 \mu(dt)\right].$$

Proposition 8 (Inégalité de Cramer-Rao)**.** *Soit ξ un estimateur sans biais et adapté de u. Alors,*

$$\mathbb{E}_u^\sigma\left[\int_0^T |\xi_t - u_t|^2\mu(dt)\right] \geq \mathrm{CRB}(\sigma,\mu),$$

où $\mathrm{CRB}(\sigma,\mu)$ *est la borne de Cramer-Rao égale au risque constant de l'estimateur X,*

$$CRB(\sigma,\mu) := \mathbb{E}_u^\sigma\left[\int_0^T |X_t - u_t|^2\mu(dt)\right] = \frac{\sigma^2T^2}{2}.$$

Cette proposition est généralisée au cas où X est une martingale gaussienne à accroissements indépendants (*c.f.* Propositions 1.1.1 et 1.2.3). Nous verrons également plus loin l'équivalent de cette inégalité pour un processus de Poisson (Proposition 1.3.4). L'inégalité de Cramer-Rao ci-dessus montre donc que dans le cas d'un processus à accroissements indépendants le processus observé est un estimateur efficace de sa propre dérive.

Nous prouvons également que l'estimateur efficace X est minimax. Cette propriété avait déjà été démontré par Berger et Wolpert. Cependant notre méthode (*c.f.* Section 1.2.3) est différente de celle de Berger et Wolpert puisque fondée sur le calcul d'un estimateur bayésien (voir la Section 1.1.2 pour le cas d'une martingale gaussienne à accroissements indépendants et la Section 1.2.3 pour le cas général d'un processus gaussien).

A ce stade nous avons donné les principales propriétés de l'estimateur X. Nous entrons désormais dans la construction des estimateurs de Stein. Nous énonçons une fois encore nos résultats pour X un mouvement brownien de variance σ^2 ; le cas le plus général que nous obtenons (X processus gaussien) est décrit dans la Section 1.2.5. Remarquons que dans la Section 1.2.1 nous exposons deux représentations des processus gaussiens :

- (A) Représentation comme un processus gaussien isonormal (suivant la construction d'Alòs, Mazet et Nualart dans [5]) et utilisation de la représentation de Paley-Wiener.
- (B) Développements de Karhunen-Loève comme cela est réalisé par Berger et Wolpert dans [14, 135].

On rappelle que sous la mesure $\mathbb{P}_u^\sigma$ le processus $X^u := X - u$ est un mouvement brownien non-drifté de variance σ^2. Pour un élément h de H, on note

$$X^u(h) := \int_0^T \dot{h}(s)dX_s^u.$$

En utilisant le calcul de Malliavin pour les processus gaussiens (détaillé dans les Sections 1.1.1 et 1.2.4) et en particulier la formule d'intégration par parties de Malliavin (*c.f.* (1.1.2) et (1.2.15)), on calcule le risque d'un estimateur de la forme $X + \xi$ où $\xi := (\xi_t)_{t\in[0,T]}$ est un processus stochastique suffisamment régulier (dans le sens du calcul de Malliavin). En particulier les calculs ci-dessous reposent sur la propriété de dérivation du gradient de Malliavin ; propriété qui était déjà essentielle pour le gradient usuel sur $\mathbb{R}^d$ dans la construction initiale de Stein. Dans les Lemmes 1.1.2 et 1.2.13 on montre que le risque de l'estimateur $X+\xi$ est égal à,

$$\begin{aligned} R(X+\xi,u) &= \mathbb{E}_u^\sigma\left[\|X+\xi-u\|^2_{L^2([0,T],\mu)}\right] \\ &= \mathrm{CRB}(\sigma,\mu)+\|\xi\|^2_{L^2(\Omega\times[0,T],\mathbb{P}_u^\sigma\otimes\mu)}+2\,\mathbb{E}_u^\sigma\left[\int_0^T D_t\xi_t\mu(dt)\right], \end{aligned}$$

où D désigne le gradient de Malliavin associé au processus X considéré. Choisissons maintenant une forme particulière pour ξ, $\xi := D\log F$ où F est une variable aléatoire $F:\Omega\to\mathbb{R}_+$ strictement positive. Le risque de l'estimateur $X+D\log F$ vaut

$$\mathbb{E}_u^\sigma\left[\|X+D\log F-u\|^2_{L^2([0,T],d\mu)}\right]=\mathrm{CRB}(\sigma,\mu)-\mathbb{E}_u^\sigma\left[\|D\log F\|^2_{L^2([0,T],\mu)}\right]+2\,\mathbb{E}_u^\sigma\left[\frac{\Delta F}{F}\right],$$

où Δ est un "laplacien de Malliavin" (*c.f.* Sections 1.1.1 et 1.2.4) défini par

$$\Delta F := \int_0^T D_tD_tF\mu(dt).$$

Ainsi, l'estimateur $X+D\log(F)$ est suroptimal si F est Δ-surharmonique sur Ω c'est à dire si

$$\Delta F<0, \quad \mathbb{P}-p.s.$$

Cette condition peut être affaiblie en considérant une variable aléatoire F de racine carrée Δ-surharmonique (comme cela est expliqué par Fourdrinier, Strawderman et Wells dans [42] pour le cas fini-dimensionnel).

Pour toute variable aléatoire F suffisamment régulière et strictement positive $\mathbb{P}$-p.s. on obtient (Propositions 1.1.3 et 1.2.14) que

$$\mathbb{E}_u^\sigma\left[\|X+D\log F-u\|^2_{L^2([0,T],d\mu)}\right]=\mathrm{CRB}(\sigma,\mu)+4\,\mathbb{E}_u^\sigma\left[\frac{\Delta\sqrt{F}}{\sqrt{F}}\right].$$

Il nous faut maintenant donner des fonctionnelles positives Δ-surharmoniques sur l'espace de Wiener. Dans la Section 1.2.6, nous effectuons une étude théorique des fonctionnelles Δ-surharmoniques en adaptant au cas infini-dimensionnel l'argument suivant. La théorie du potentiel donne qu'une fonction suffisamment régulière f est surharmonique par rapport au laplacien usuel sur $\mathbb{R}^d$ si

$$f(x)\geq\mathbb{E}^x[f(B_{\tau_r})],$$

où B est un mouvement brownien standard sur $\mathbb{R}^d$, $\mathbb{E}^x$ désigne l'espérance conditionnelle sachant que $B_0=x$ et τ_r est le premier temps d'atteinte de la boule de centre x et de rayon r par B pour $r>0$.

Dans les Sections 1.1.4, 1.2.6 et 1.2.7 nous considérons une classe particulière de fonctionnelles cylindriques. Pour $\mu(dt):=dt$, $(h_k)_{k\geq1}$ une famille orthonormale de $L^2([0,T])$

particulière, $(\lambda_k)_{k\geq 1}$ une famille de réels bien choisie et pour n fixé on s'intéresse aux fonctionnelles cylindriques F de la forme,

$$F := f_n(\lambda_1^{-1}X^u(h_1), \dots, \lambda_n^{-1}X^u(h_n)),$$

où $f_n : \mathbb{R}^n \to \mathbb{R}$ est une fonction $\mathscr{C}^2(\mathbb{R}^n)$ donnée. Par définition de l'opérateur Δ le laplacien ΔF de F vaut

$$\Delta F = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} f_n(\lambda_1^{-1}X^u(h_1), \dots, \lambda_n^{-1}X^u(h_n)).$$

Ainsi la variable F est Δ-surharmonique si et seulement si la fonction f_n est surharmonique pour le laplacien usuel sur $\mathbb{R}^n$. Notons que seule l'information X est disponible et que X^u n'est pas observable. Cependant dans le cas où la dérive u est déterministe, on peut remplacer dans ce qui précède $f_n(\lambda_1^{-1}X^u(h_1), \dots, \lambda_n^{-1}X^u(h_n))$ par $f_n(\lambda_1^{-1}X(h_1), \dots, \lambda_n^{-1}X(h_n))$ puisque dans ce cas on a que

$$DX^u(h) = DX(h), \quad h \in H.$$

Nous obtenons ainsi comme cas particulier de nos estimateurs de Stein, des estimateurs de type James-Stein. Dans la Section 1.2.7 nous étudions une classe d'estimateurs de type James-Stein et nous optimisons le risque de cet estimateur en fonction du nombre d'intégrales stochastiques choisies n. Dans cette même section, nous donnons explicitement pour un mouvement brownien de variance σ^2 la famille de fonctions $(h_k)_{k\geq 1}$, les réels $(\lambda_k)_{k\geq 1}$ ainsi que des simulations numériques fondées sur la décomposition de Paley-Wiener.

Estimation de Stein de l'intensité d'un processus de Poisson

Soit T un réel strictement positif fixé et $u := (u(t))_{t\in[0,T]}$ une fonction déterministe de la forme

$$u(t) := \lambda h(t), \quad t \in [0, T]$$

où h est une fonction positive supposée connue et λ est un réel strictement positif. Considérons $X := (X_t)_{t\in[0,T]}$ un processus de Poisson d'intensité u sous une mesure de probabilité $\mathbb{P}_\lambda$. On suppose que λ est un paramètre inconnu à estimer. On définit la fonction de risque quadratique R pour un estimateur ξ de λ par

$$R(\xi, \lambda) := \mathbb{E}_\lambda[|\xi - \lambda|^2], \quad \lambda > 0$$

où $\mathbb{E}_\lambda$ désigne l'espérance par rapport à la mesure $\mathbb{P}_\lambda$.

Pour ce problème d'estimation paramétrique on peut calculer l'estimateur du maximum de vraisemblance (*c.f.* [71, 75]) ; vraisemblance donnée par la densité de Girsanov

$$\lambda^{X_T} e^{-(\lambda h(T)-T)} \prod_{k=1}^{X_T} \dot{h}(T_k).$$

Un calcul immédiat donne que le MLE $\hat{\lambda}_T$ vaut

$$\hat{\lambda}_T := \frac{X_T}{h(T)}.$$

Comme pour le cas gaussien le MLE $\hat{\lambda}_T$ est un estimateur sans biais de risque

$$\mathbb{E}_\lambda \left[|\hat{\lambda}_T - \lambda|^2 \right] = \frac{\lambda}{h(T)}.$$

On peut également démontrer que le MLE $\hat{\lambda}_T$ est un estimateur efficace qui atteint la borne de Cramer-Rao $\frac{\lambda}{h(T)}$.

Dans la Section 1.3.4 nous construisons un estimateur suroptimal de type Stein pour le paramètre λ. Comme précisé plus haut, la formule d'intégration par parties pour la loi de Poisson (4) fait intervenir l'opérateur aux différences ∇_{df} ne possédant pas la propriété de dérivation. On ne peut donc pas reproduire directement le calcul de Stein menant à une simplification du risque. L'estimateur de type Stein que nous proposons est de la forme

$$\hat{\lambda}_T + \frac{\xi_T}{h(T)},$$

où ξ_T est la valeur au temps T d'un processus particulier ξ suffisamment régulier au sens du calcul de Malliavin.

Dans la Section 1.3.2 nous étudions tout d'abord une version non-paramétrique du problème précédent. C'est à dire que nous considérons un processus de Poisson X d'intensité u de la forme

$$u(t) = \int_0^t \dot{u}(s) ds, \quad \dot{u} : [0, T] \to [0, +\infty[, \ t \in [0, T]$$

sous une mesure de probabilité $\mathbb{P}_u$ (l'espérance par rapport à $\mathbb{P}_u$ est notée $\mathbb{E}_u$). Nous démontrons une inégalité de Cramer-Rao "temps par temps" (Proposition 1.3.4), *i.e.* pour tout estimateur $\xi = (\xi_t)_{t \in [0,T]}$ sans biais et adapté dans le sens de la Définition 1.3.2 on a que

$$\mathbb{E}_u \left[|\xi_t - u(t)|^2 \right] \geq CRB_t := \mathbb{E}_u \left[|X_t - u(t)|^2 \right] = u(t), \quad \forall t \in [0, T].$$

Le processus X est donc un estimateur efficace de sa propre intensité. Ce résultat généralise une inégalité de Cramer-Rao obtenue par Kutoyants dans [72] pour une intensité u constante par morceaux.

Nous donnons un estimateur non-paramétrique de type James-Stein. Soit $\xi = (\xi_t)_{t \in [0,T]}$ un processus suffisamment régulier au sens du calcul de Malliavin, nous montrons dans le Lemme 1.3.11 que le risque de l'estimateur $X + \xi$ est donné par,

$$\mathbb{E}_u \left[|X_t + \xi_t - u(t)|^2 \right] = u(t) + \mathbb{E}_u \left[|\xi_t|^2 \right] + 2\, \mathbb{E}_u \left[\nabla_t \xi_t \right],$$

où ∇ est le gradient de Malliavin initialement introduit par Carlen et Pardoux dans [23] puis étendu par Elliot et Tsoi dans [39] et par Privault dans [102]. Notons que cet opérateur satisfait la propriété de dérivation. Nous considérerons dans le Chapitre 2 l'opérateur au différences qui est un autre gradient de Malliavin sur l'espace de Poisson mais qui ne possède pas la propriété de dérivation. Dans la Section 1.3.3 nous rappelons la définition de ∇ et nous démontrons les principales propriétés de cet opérateur, en particulier nous donnons la formule d'intégration par parties (Proposition 1.3.7) permettant d'obtenir le calcul de risque ci-dessus. Comme dans le cas gaussien nous souhaitons considérer un estimateur de la forme $\xi := \nabla \log F$ où F est une fonctionnelle positive sur l'espace de Poisson. Malheureusement une condition technique nécessaire à l'obtention de la formule d'intégration par parties nous interdit de considérer ce processus car la fonctionnelle $\log F$ n'appartient pas au domaine du gradient ∇ même en choisissant F très régulière (comme une fonctionnelle cylindrique). Nous sommes obligés de considérer un processus de la forme,

$$\xi_t = c\frac{u(t)}{\dot{u}(T)}\mathbf{1}_{\{X_T=0\}} + \nabla_t \log F, \qquad t \in [0,T],$$

où c et F sont respectivement une constante et une fonctionnelle sur l'espace de Poisson bien choisies. Nous renvoyons le lecteur à la Section 1.3.4 pour les détails de ces résultats. Cet estimateur est un estimateur théorique puisque dépendant de la fonction inconnue u. Une possibilité consiste à remplacer u par un estimateur dans l'expression de ξ mais nous n'avons pas suivi cette voie. Cependant, dans le cas paramétrique où u est de la forme $u(t) = \lambda h(t),\ t \in [0,T]$ l'estimateur de Stein ci-dessus ne dépend pas du paramètre à estimer λ. On obtient par exemple dans ce cas un estimateur $\tilde{\lambda}_T$ donné par (*c.f.* Proposition 1.3.10),

$$\tilde{\lambda}_T := \hat{\lambda}_T - \frac{1}{\dot{h}(T)}\frac{f_1'(T)}{f_1(T)}\mathbf{1}_{\{X_T=0\}} + \frac{1}{h(T)}\nabla_T \log F,$$

où $\nabla_T F$ est donné par

$$\nabla_t F = -\sum_{n=1}^{\infty}\mathbf{1}_{\{X_T=n\}}\sum_{k=1}^{n}\frac{h(t\wedge T_k)}{\dot{h}(T_k)}\partial_k f_n(T_1,\dots,T_n)$$

avec F une fonction cylindrique

$$F = f_0\mathbf{1}_{X_T=0} + \sum_{n=1}^{\infty}\mathbf{1}_{X_T=n} f_n(T_1,\dots,T_n)$$

vérifiant les conditions "de continuité" données dans la Définition 1.3.4 et où les T_k sont les temps de sauts du processus X. L'estimateur $\tilde{\lambda}_T$ a pour risque

$$\mathbb{E}_u[|\tilde{\lambda}_T - \lambda|^2] = \frac{\lambda}{h(T)} + \frac{1}{\dot{h}^2(T)}\left|\frac{f_1'(T)}{f_1(T)}\right|^2 e^{-\lambda h(T)} + \frac{4}{h^2(T)}\mathbb{E}_{\lambda h}\left[\frac{\nabla_T\nabla_T\sqrt{F}}{\sqrt{F}}\right].$$

On retrouve la borne de Cramer-Rao et un terme du type laplacien constitué d'un gradient de Malliavin d'ordre de 2 et un terme additionnel dû au terme correctif $-\frac{1}{\dot{h}(T)}\frac{f_1'(T)}{f_1(T)}\mathbf{1}_{\{X_T=0\}}$

de l'estimateur. Ainsi contrairement au cas gaussien la recherche d'une fonctionnelle positive F surharmonique pour un laplacien de Malliavin n'est plus suffisante puisqu'il faut désormais que la somme des deux derniers termes du risque ci-dessus soit négative. Cette analyse est assez compliquée et nous obtenons dans la Section 1.3.5 quelques exemples et quelques simulations numériques d'estimateurs de Stein pour λ fondés sur des fonctionnelles (pseudo)-surharmoniques par rapport au laplacien de Malliavin considéré.

Chapitre 2 : Estimation bayésienne pour les canaux de Poisson

Dans le Chapitre 2 nous calculons l'estimateur bayésien d'un canal de Poisson (fini ou infini-dimensionnel) à l'aide du calcul de Malliavin sur l'espace de Poisson.

Un canal additif gaussien est la donnée d'un *signal d'entrée* X, d'un *signal de sortie* Y perturbé par addition d'un bruit gaussien w au signal d'entrée X,

$$Y = \rho X + w.$$

Plus précisément le signal d'entrée X est une variable aléatoire à valeurs dans un espace de Hilbert H de loi μ_X, w est une variable aléatoire à valeurs dans un espace de Banach W de loi μ_W et ρ est un réel appelé le *rapport signal sur bruit*. Le triplet (W, H, μ_W) est un *espace de Wiener abstrait* (*c.f.* [94]). Par exemple, considérons un processus stochastique $(Y_t)_{t\in[0,T]}$ défini par

$$dY_t = \rho\, X_t\, dt + dW_t,\ t \in [0, T]$$

où W est un mouvement brownien standard et X est un processus stochastique indépendant de W à valeurs dans $H := L^2([0, T])$. On note $\mathcal{Y}$ la tribu engendrée par Y. Dans [138], Zakai a démontré que l'estimateur bayésien $\mathbb{E}[X|\mathcal{Y}]$ vaut

$$\mathbb{E}[X|\mathcal{Y}] = \frac{1}{\rho}\nabla \log l(Y) \tag{6}$$

où l est le rapport de vraisemblance du modèle considéré c'est à dire la densité

$$l := \frac{d\mu_Y}{d\mu_W}$$

et ∇ est le gradient de Malliavin sur l'espace de Wiener. Zakai obtient les résultats précédents pour tout canal gaussien additif défini sur un espace de Wiener abstrait donnant un équivalent infini-dimensionnel à l'expression bien connue de l'estimateur (6) en dimension finie. Soit Y un vecteur gaussien de $\mathbb{R}^d$ de la forme,

$$Y = \rho X + N,$$

où X est un vecteur aléatoire dans $\mathbb{R}^d$ et N est un vecteur gaussien centré de variance $\mathbf{I}_d$ indépendant de X avec $d \geq 1$. On note $\mathcal{Y}$ la tribu engendrée par Y. L'estimateur bayésien de X est donné par

$$\mathbb{E}[X|\mathcal{Y}] = \frac{1}{\rho}\nabla \log \tilde{m}(Y),$$

avec

$$\tilde{m}(y) = \int_H \exp\left(-\frac{\rho^2 x^2 - 2\rho y x}{2}\right)\, \mu_X(dx)$$

et ∇ est le gradient usuel sur $\mathbb{R}^d$. Notons que cette expression et la relation (6) sont valables quelque soit la distribution de X.

Zakai applique l'identité (6) pour montrer que ([138, Proposition 5.1])

$$\frac{dI(X;Y)}{d\rho} = \rho \, \mathbb{E}\left[\|X - \mathbb{E}[X|\mathcal{Y}]\|_H^2\right], \tag{7}$$

où $I(X;Y)$ est l'information mutuelle entre X et Y définie par

$$I(X;Y) := \int_{H\times W} \log \frac{d\mu_{X,Y}}{d(\mu_X \times \mu_Y)} \mu_{X,Y}(dx, dy).$$

Une autre application de l'expression de l'estimateur bayésien (6) est donnée par Üstünel dans [133] où l'auteur relie l'erreur quadratique de l'estimateur bayésien au problème de transport optimal des mesures de Monge-Kantorovitch.

Dans le Chapitre 2, nous étendons la relation (6) pour des canaux de Poisson généraux (fini et infini-dimensionnels) en utilisant du calcul de Malliavin sur l'espace de Poisson dans les Sections 2.2 et 2.4. Puis nous étendons nos résultats dans la Section 2.5 pour des canaux plus généraux incluant les mélanges gaussiens-Poisson et des canaux dont le bruit est une martingale normale possédant la propriété de représentation chaotique telles que les martingales d'Azéma fournissant ainsi un exemple de martingales à accroissements non-indépendants.

Comme cela est précisé dans la Remarque 2 de [138], l'expression (6) (et son analogue pour le cas Poisson (2.4.8)) permet une approximation numérique de l'estimateur de X en évaluant la densité du signal de sortie Y par un schéma de Monte-Carlo.

On suppose que l'on observe une loi de Poisson Y d'intensité $\alpha X + \lambda_0$ où X est une variable aléatoire à valeurs dans $\mathbb{R}_+^*$,

$$Y \sim \mathcal{P}(\alpha X + \lambda_0).$$

Plus précisément la loi de Y sachant X est une loi de Poisson d'intensité $\alpha X + \lambda_0$. Ceci constitue un canal de Poisson fini-dimensionnel. Remarquons que contrairement au cas gaussien ce canal n'est pas "additif" puisque la sortie Y n'est pas la somme de l'entrée X et d'une variable de Poisson (une telle somme ne serait pas à valeurs entières). Ce type de canal est utilisé pour modéliser des cellules photo-sensibles (comme par exemple la diode p-i-n). Le paramètre λ_0 représente le *dark current noise* qui est un courant résiduel dans la cellule photo-sensible en l'absence de signal (lumière) tandis que le paramètre α est un paramètre d'échelle propre à la cellule. Un calcul élémentaire montre que (voir le Lemme 2.2.1) l'estimateur bayésien $\mathbb{E}[X|\mathcal{Y}]$ où $\mathcal{Y}$ est la tribu engendrée par Y est égal à

$$\mathbb{E}[X|\mathcal{Y}] = \frac{\lambda_0}{\alpha} \frac{m(Y+1) - m(Y)}{m(Y)}$$

avec

$$m(y) := \int_0^{+\infty} \frac{d\mu_{Y|X=x}}{d\mu_{\lambda_0}}(y)\,\mu_X(dx),\ y \in \mathbb{N},$$

et μ_{λ_0} la loi d'une variable de Poisson d'intensité λ_0. Il est fort probable que ce résultat (fini-dimensionnel) soit déjà connu cependant nous n'avons trouvé aucune référence dans la littérature. Nous étendons cette relation à canal de Poisson infini-dimensionnel en utilisant du calcul de Malliavin sur l'espace de Poisson.

Soit Y un processus ponctuel de Poisson sur un espace S d'intensité $1 + \alpha X$. Par exemple prenons $S = [0, T]$. Soit X une variable aléatoire à valeurs dans l'espace H des mesures positives sur $[0, T]$ absolument continues par rapport à la mesure de Lebesgue

$$H := \{h : [0, T] \to \mathbb{R}_+,\ h(t) := \int_0^t \dot{h}(s)ds,\ \dot{h} \in L^2([0, T])\}.$$

Le processus Y sachant X est un processus de Poisson sur $[0, T]$ d'intensité $1+\alpha X$ où α est un réel positif fixé. Les notations et définitions relatives à un processus ponctuel de Poisson sont données dans les Sections 2.3 et 2.4. Dans la Proposition 2.4.6 on calcule l'estimateur bayésien de l'entrée inconnue X. On obtient en particulier (Corollaire 2.4.8) l'expression suivante pour $\mathbb{E}[X|\mathcal{Y}]$

$$\mathbb{E}[X_t|\mathcal{Y}] = \frac{\nabla^d_{[0,t]} m(Y)}{\alpha\, m(Y)},\ t \in [0, T], \tag{8}$$

où m est la densité de l'observation Y et ∇ est un gradient de Malliavin sur l'espace de Poisson présenté dans la Section 2.3.

Remarque 9. *Contrairement au gradient de Malliavin de type Carlen-Pardoux décrit précédemment dans l'introduction, le gradient ∇^d considéré ici ne satisfait pas la propriété de dérivation. En particulier,*

$$\frac{\nabla^d m}{m} \neq \nabla^d \log m.$$

Une extension de ce résultat est obtenue dans la Section 2.5 en remplaçant le processus de Poisson par une martingale normale possédant la propriété de représentation chaotique. Les relations du type (8) que nous obtenons reposent sur une "compatibilité" du gradient de Malliavin avec la représentation chaotique représentée par la formule (2.5.3). Le "bon" gradient à considérer est donc celui apparaissant dans la formule (2.5.3). Dans le cas gaussien le gradient de Malliavin usuel regroupe à lui seul cette propriété de "compatibilité" et la propriété de dérivation. L'absence de cette dernière pour le gradient ∇^d ne nous a pas permis à l'heure actuelle d'adapter au cas Poisson le résultat de Zakai (7).

Chapitre 3 : TLC pour les variations quadratiques à poids de champs gaussiens

Brève introduction à l'étude du comportement asymptotique des variations à poids

Supposons que l'on observe une unique trajectoire d'un processus stochastique $Z := (Z_t)_{t\in[0,1]}$ aux temps $\left\{\frac{i}{n},\ 0 \leq i \leq n\right\}$ avec $n \geq 1$. On se place dans le cas d'observations dites "à hautes fréquences" c'est à dire avec n convergeant vers l'infini et avec un horizon de temps fini ici 1. De nombreuses propriétés statistiques du processus Z peuvent être obtenues en étudiant le comportement asymptotique des p-variations à poids du processus définies comme

$$\sum_{i=1}^{n} f\left(Z_{\frac{i-1}{n}}\right)(\Delta_i Z)^p, \quad \Delta_i Z := Z_{\frac{i}{n}} - Z_{\frac{i-1}{n}},$$

où la fonction de poids $f : \mathbb{R} \to \mathbb{R}$ est déterministe. Si $f \equiv 1$ on parle de p-variations du processus Z.

Par exemple l'étude des p-variations (et des bi-variations) a été utilisé par Barndorff-Nielsen et Shephard dans [10, 11] afin de résoudre certains problèmes d'économétrie et de finance (voir également l'article de Barndorff-Nielsen, Graversen, Jacod et Shephard [9]). Un objectif important en statistique inférentielle est l'obtention de théorèmes de la limite centrale (TLC) pour les p-variations de processus stochastiques. Citons par exemple les TLC établis par Aït Sahalia et Jacod dans [2] pour des processus de Lévy, par Gradinaru et Nourdin pour des processus de diffusions dans [50] et par Jacod dans [63, 64] pour des semi-martingales d'Itô. Ces TLC permettent de construire des estimateurs asymptotiquement normaux de la variation quadratique (*c.f.* [63]) ou des tests statistiques comme par exemple dans [50] où Gradinaru et Nourdin ont appliqué l'équivalent de la convergence (9) ci-dessous pour un mouvement brownien standard à la construction d'un test statistique non-paramétrique de la volatilité intégrée d'un processus de diffusion ou encore comme dans [1] où Aït Sahalia et Jacod construisent un test pour les sauts d'un processus observé à des temps discrets. Remarquons également que dans [63, 64] Jacod démontre des théorèmes limites pour des fonctions assez générales des accroissements généralisant le cas des accroissements à la puissance p.

Les variations à poids du mouvement brownien fractionnaire ont quant à elles été étudiées initialement pour l'obtention des taux de convergence exacts de certains schémas d'approximation de solutions d'équations différentielles stochastiques linéaires dirigées par un mouvement brownien fractionnaire par Gradinaru et Nourdin dans [49], par Neuenkirch et Nourdin dans [78] et par Nourdin dans [82]. L'étude des variations à poids du mouvement brownien fractionnaire a été complétée par Nourdin dans [81], par Nourdin et Nualart dans [83] et par Nourdin, Nualart et Tudor dans [84]. Nous reviendrons un plus loin à ces travaux puisque la Section 3.3 réalisée en collaboration avec Ivan Nourdin participe à cette étude.

Notons enfin que Nourdin et Peccati ont donné des résultats concernant les variations à poids du mouvement brownien itéré dans [88].

Nous donnons dans la Section 3.1 un théorème de la limite centrale pour les variations quadratiques à poids du drap brownien standard et nous appliquons ce résultat à l'obtention d'un estimateur asymptotiquement normal de la variation quadratique de certains processus de diffusions à deux paramètres. Puis dans la Section 3.2 nous prouvons un théorème de la limite centrale (au sens des lois fini-dimensionnelles) pour le processus des variations quadratiques à poids de certains draps browniens fractionnaires. Dans la Section 3.3 réalisée en collaboration avec Ivan Nourdin, nous complétons le cas manquant des variations quadratiques à poids du mouvement brownien fractionnaire d'indice $H = 1/4$. Ce résultat nous permet de démontrer une formule d'Itô en loi pour la limite de sommes de Riemann à signes alterné du mouvement brownien fractionnaire de paramètre de Hurst $H = 1/4$ conjecturée par Burdzy et Swanson comme cité dans [127].

Estimation de la variation quadratique de certains processus de diffusions à deux paramètres

Dans la Section 3.1 nous établissons un théorème de la limite centrale pour les variations quadratiques à poids d'un drap brownien standard. Ce résultat nous permet de donner un estimateur asymptotiquement normal de la variation quadratique de certains processus de diffusions à deux paramètres.

Récemment, les processus stochastiques à deux paramètres ont été utilisés pour modéliser des phénomènes naturels 2-dimensionnels. En particulier, de nombreuses références sont dédiées à l'analyse d'images médicales telles que les radiographies d'os : voir par exemple les travaux de Bonami et Estrade dans [19], de Cohen, Guyon, Perrin et Pontier dans [30] et de Léger dans [73] ou les mammographies, *c.f.* l'article [16] de Biermé et Richard. Dans ces applications, la texture de l'image est modélisée par un processus stochastique à deux paramètres, comme par exemple le champ brownien fractionnaire (voir [19]) ou le drap brownien fractionnaire ([7, 73]) dont les trajectoires sont auto-similaires. Il devient alors intéressant d'estimer l'ordre d'auto-similarité donné par respectivement un ou deux paramètres de Hurst. Cette estimation a permis dans les cas exposés plus haut de diagnostiquer la présence ou l'absence d'ostéoporose chez un patient (voir [7, 19, 30]) ou de mieux comprendre la structure de la partie du corps observée comme dans [16]. Une méthode d'estimation du (des) paramètre(s) repose sur l'étude du comportement asymptotique des variations quadratiques du processus observé comme cela a été réalisé initialement par Istas et Lang dans [61]. Des quantités similaires ont été étudié dans [16, 17, 19, 30, 73]. Aux vues de ces applications, il parait naturel d'étudier une généralisation des variations quadratiques en considérant un poids dans ces dernières pour des processus à deux paramètres et en particulier pour les draps browniens fractionnaires. Nous étudions tout d'abord le cas du drap brownien standard.

Soit $W := (W_{(s,t)})_{(s,t)\in[0,1]^2}$ un mouvement brownien à deux paramètres (ou drap brownien) sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, c'est à dire que W est un processus gaussien centré de fonction de covariance,

$$\mathbb{E}[W_{(s_1,t_1)} W_{(s_2,t_2)}] = \min(s_1, s_2) \min(t_1, t_2), \quad (s_1, t_1) \in [0,1]^2,\ (s_1, t_1) \in [0,1]^2.$$

La figure 1 présente une trajectoire d'un drap brownien standard. Cette simulation numérique a été réalisée en langage R.

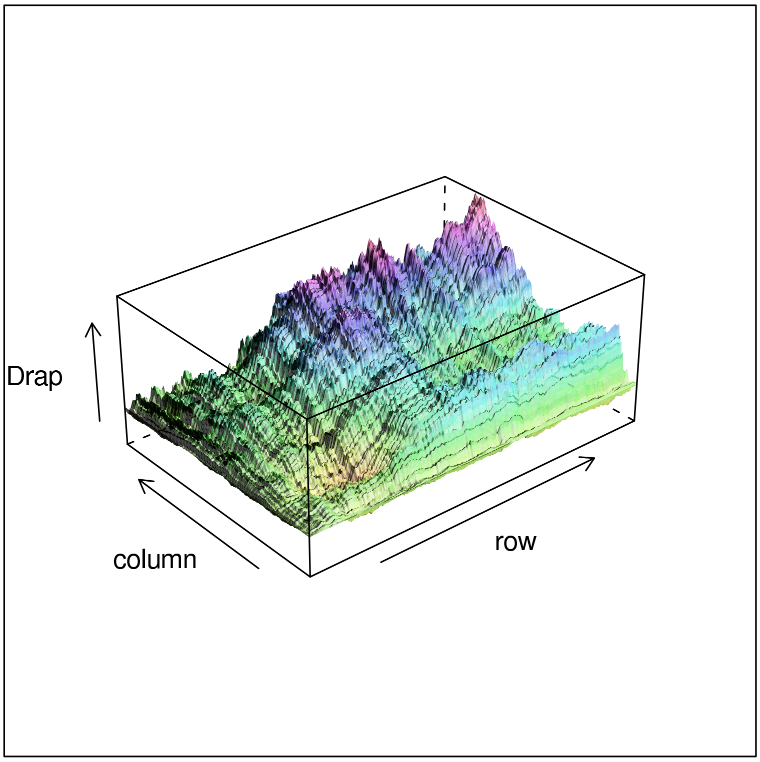

FIG. 1: Drap brownien standard.

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction déterministe. On s'intéresse au comportement lorsque n tend vers l'infini des variations quadratiques à poids (re-normalisées) du drap brownien standard données par

$$X^n := n \sum_{i=1}^{[n\cdot]} \sum_{j=1}^{[n\bullet]} f\left(W_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)}\right) \left(|\Delta_{i,j} W|^2 - \frac{1}{n^2}\right),$$

où $\Delta_{i,j} W$ est l'accroissement du processus W sur le rectangle $\left[\frac{i-1}{n}, \frac{i}{n}\right] \times \left[\frac{j-1}{n}, \frac{j}{n}\right]$ défini par

$$\Delta_{i,j} W := W_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)} + W_{\left(\frac{i}{n}, \frac{j}{n}\right)} - W_{\left(\frac{i-1}{n}, \frac{j}{n}\right)} - W_{\left(\frac{i}{n}, \frac{j-1}{n}\right)}.$$

Dans le Théorème 3.1.3 on montre la convergence suivante

$$X^n \xrightarrow[n\to\infty]{loi(\mathcal{S})} \sqrt{2} \int_{[0,\cdot]\times[0,\bullet]} f\left(W_{(u,v)}\right) \, dB_{(u,v)}. \tag{9}$$

La notation $loi(\mathcal{S})$ signifie que $(X^n)_{n\geq 1}$ converge stablement en loi entant que processus dans la topologie de Skorohod à deux paramètres vers le processus X décrit ci-dessous où B est un drap brownien standard indépendant de W. Le processus limite

$$X_{(s,t)} := \sqrt{2} \int_{[0,s]\times[0,t]} f\left(W_{(u,v)}\right)\, dB_{(u,v)}, \quad (s,t) \in [0,T]^2.$$

n'est pas gaussien bien que gaussien conditionnellement à W. De plus ce processus est défini sur une extension de l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ (*c.f.* Section 3.1.3). La convergence stable a été introduite par Rényi dans [114, 115], nous en donnons la définition formelle dans la preuve du Théorème 3.1.3 cependant de manière heuristique cette convergence correspond à la convergence en loi du couple $(X^n, B)_{n\geq 1}$ vers le couple (X, B). Nous décrivons ci-dessous une application du Théorème 3.1.3.

Soit le processus de diffusion $Y := (Y_{(s,t)})_{(s,t)\in[0,1]^2}$ défini comme suit :

$$Y_{(s,t)} := \int_{[0,s]\times[0,t]} \sigma(W_\rho)\, dW_\rho + \int_{[0,s]\times[0,t]} M_\rho\, d\rho, \quad (s,t) \in [0,1]^2$$

où $W = (W_{(s,t)})_{(s,t)\in[0,1]^2}$ est un drap brownien standard, $\sigma : \mathbb{R} \to \mathbb{R}$ une fonction déterministe suffisamment régulière et $M = (M_{(s,t)})_{(s,t)\in[0,1]^2}$ un processus continu et adapté. On suppose que l'on observe le processus Y sur une grille régulière $\left\{\left(\frac{i}{n}, \frac{j}{n}\right),\ 1 \leq i, j, \leq n\right\}$ avec $n \geq 1$. On pose

$$V^n_{(s,t)} := \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \left|\Delta_{i,j} Y\right|^2, \quad (s,t) \in [0,1]^2,\ n \geq 1,$$

et

$$C_{(s,t)} := \int_{[0,s]\times[0,t]} \sigma^2\left(W_{(u,v)}\right)\, dudv, \quad (s,t) \in [0,1]^2.$$

On pose également,

$$Y^n_{(s,t)} := n\left(V^n_{(s,t)} - C_{(s,t)}\right), \quad (s,t) \in [0,1]^2.$$

On montre dans le Théorème 3.1.7, en utilisant du calcul de Malliavin, que $(Y^n)_{n\geq 1}$ converge stablement en loi vers X avec $f(\cdot) := \sigma^2(\cdot)$. La preuve du Théorème 3.1.7 consiste à montrer que $Y^n = X^n + r_n$ où r_n est un terme négligeable lorsque n tend vers l'infini. Pour cela nous utilisons du calcul de Malliavin. On pourrait également utiliser du calcul stochastique pour les martingales à deux paramètres (développé dans [35, 60, 91, 92, 120]). Cependant on ne peut pas appliquer directement les résultats des références précédentes et c'est pourquoi on choisit d'utiliser le calcul de Malliavin qui permet de conclure plus rapidement la preuve au prix d'une régularité un peu plus élevée. On prouve ensuite dans la Proposition 3.1.9 que $(V^n)_{n\geq 1}$ est un estimateur consistent de la variation quadratique C de Y (on démontre même que $(V^n)_{n\geq 1}$ converge dans L^2 vers C). En utilisant le Théorème

3.1.7 on établit (dans le Corollaire 3.1.11) que l'estimateur $(V^n)_{n\geq 1}$ de C est asymptotiquement normal.

Pour (s,t) éléments de $[0,1]^2$, on définit $S^n_{(s,t)}$ par

$$S^n_{(s,t)} := n^2 \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} |\Delta_{i,j} Y|^4, \quad n \geq 1.$$

Soit (s,t) fixés dans $]0,1]^2$ tels que $S^n_{(s,t)}$ ne s'annule pas. Alors on a que

$$\left(S^n_{(s,t)}\right)^{-\frac{1}{2}} n \left(V^n_{(s,t)} - C_{(s,t)}\right) \xrightarrow[n\to\infty]{loi} \sqrt{\frac{2}{3}} N, \quad N \sim \mathcal{N}(0,1).$$

Théorèmes limites pour les variations quadratiques à poids de certains draps browniens fractionnaires

Dans la Section 3.2 nous prouvons un théorème de la limite centrale (au sens des lois fini-dimensionnelles) pour le processus des variations quadratiques à poids de certains draps browniens fractionnaires. Dans ce travail nous utilisons principalement une méthode développée par Nourdin et Nualart dans [83] pour l'étude des variations à poids du mouvement brownien fractionnaire fondée sur le calcul de Malliavin.

Soit $W^{\alpha,\beta} := (W^{\alpha,\beta}_{(s,t)})_{(s,t)\in[0,1]^2}$ un drap brownien fractionnaire de paramètres de Hurst α et β. Ce processus a été initialement introduit par Kamont dans [70] et complété par Ayache, Léger et Pontier dans [7]. $W^{\alpha,\beta}$ est un processus gaussien centré s'annulant sur les axes et de fonction de covariance $R^{\alpha,\beta}$,

$$\begin{aligned} R^{\alpha,\beta}((s_1,t_1),(s_2,t_2)) &:= \mathbb{E}\left[W^{\alpha,\beta}_{(s_1,t_1)} W^{\alpha,\beta}_{(s_2,t_2)}\right] \\ &= K^{\alpha}(s_1,s_2) K^{\beta}(t_1,t_2) \\ &= \frac{1}{2}\left(s_1^{2\alpha} + s_2^{2\alpha} - |s_1 - s_2|^{2\alpha}\right) \frac{1}{2}\left(t_1^{2\beta} + t_2^{2\beta} - |t_1 - t_2|^{2\beta}\right). \end{aligned}$$

En particulier, $W^{1/2,1/2}$ est le drap brownien standard décrit plus haut. Le résultat principal de la Section 3.2 est le Théorème 3.2.2 que nous rappelons ici.

Si α et β vérifient : $0 < \alpha, \beta < \frac{1}{2}$, avec $\alpha + \beta > \frac{1}{2}$ alors pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ suffisamment régulière on a que

$$n^{-1} \sum_{i=1}^{[n\cdot]} \sum_{j=1}^{[n\bullet]} f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) \left(n^{2(\alpha+\beta)} |\Delta_{i,j} W^{\alpha,\beta}|^2 - 1\right) \xrightarrow[n\to\infty]{fdd} \sigma_{\alpha,\beta} \int_0^{\cdot} \int_0^{\bullet} f\left(W^{\alpha,\beta}_{(u,v)}\right) dW_{(u,v)} \tag{10}$$

où W est un drap brownien standard indépendant de $W^{\alpha,\beta}$ et la notation fdd signifie que la convergence a lieu dans le sens de la convergence stable des lois fini-dimensionnelles (voir (3.2.8)). De plus la constante $\sigma_{\alpha,\beta}$ est une constante qui peut être exprimée de façon explicite en fonction de α et β (*c.f.* (3.2.7)). On rappelle la définition de l'accroissement $\Delta_{i,j}W^{\alpha,\beta}$ du processus $W^{\alpha,\beta}$ sur le rectangle $\left[\frac{i-1}{n}, \frac{j-1}{n}\right] \times \left[\frac{i}{n}, \frac{j}{n}\right]$:

$$\Delta_{i,j}W^{\alpha,\beta} := W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)} + W^{\alpha,\beta}_{\left(\frac{i}{n},\frac{j}{n}\right)} - W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j}{n}\right)} - W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j}{n}\right)}.$$

La convergence (10) complète partiellement le résultat obtenu dans la Section 3.1 pour les variations quadratiques à poids du drap brownien standard. Contrairement au cas du drap brownien standard, nous n'avons pas démontré la propriété de tension pour les draps browniens fractionnaires.

Dans [83], Nourdin et Nualart ont montré la convergence suivante

$$n^{-1/2}\sum_{i=1}^{n} f\left(B_{\frac{i-1}{n}}\right)\left(n^{2H}|\Delta_i B|^2 - 1\right) \xrightarrow[n\to\infty]{loi} \sigma_H \int_0^1 f(B_s)\, dW_s$$

où B un mouvement brownien fractionnaire de paramètre de Hurst H dans $]1/4, 1/2[$, W est un mouvement brownien standard indépendant de B et la convergence est stable. Nourdin et Nualart ont prouvé ce résultat en montrant (par le calcul de Malliavin) que la fonction caractéristique de la loi conditionnelle de $\sigma_H \int_0^1 f(B_s)\, dW_s$ sachant B vérifie l'équation différentielle dont la seule solution est la fonction caractéristique d'une loi gaussienne centrée et de variance $\sigma_H^2 \int_0^1 f^2(B_s)\, ds$. Dans la Section 3.2 nous étendons cet argument afin de montrer de la convergence fini-dimensionnelle en remarquant que la fonction caractéristique d'un vecteur gaussien centré de matrice covariance donnée est la seule solution d'un système d'équations aux dérivées partielles. Nous avons ensuite adapté au cas à deux paramètres les principaux calculs et estimées données dans l'article [83] de Nourdin et Nualart, ces calculs sont rassemblés dans la Section 3.2.4.

Variations quadratiques à poids du mouvement brownien fractionnaire : le cas limite $H = 1/4$

La Section 3.3 présente un travail en collaboration avec Ivan Nourdin.

Comme précisé plus haut dans cette introduction, le comportement asymptotique des variations à poids du mouvement brownien fractionnaire a été étudié par Nourdin dans [81], par Nourdin et Nualart dans [83] et par Nourdin, Nualart et Tudor dans [84]. Soit $B^H := (B_t^H)_{t\in[0,1]}$ un mouvement brownien de paramètre de Hurst $0 < H < 1$. Une compilation des résultats mentionnés plus haut permet d'obtenir les convergences ci-dessous pour les variations quadratiques à poids du mouvement brownien fractionnaire,

- si $H < \frac{1}{4}$ alors

$$n^{2H-1}\sum_{k=0}^{n-1} f(B^H_{k/n})\big[n^{2H}(B^H_{(k+1)/n} - B^H_{k/n})^2 - 1\big] \xrightarrow[n\to\infty]{L^2} \frac{1}{4}\int_0^1 f''(B^H_s)ds; \tag{11}$$

- si $\frac{1}{4} < H < \frac{3}{4}$ alors

$$\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1} f(B^H_{k/n})\big[n^{2H}(B^H_{(k+1)/n} - B^H_{k/n})^2 - 1\big] \xrightarrow[n\to\infty]{\text{loi}} C_H\int_0^1 f(B^H_s)dW_s \tag{12}$$

  où W est un mouvement brownien standard indépendant de B^H ;
- si $H = \frac{3}{4}$ alors

$$\frac{1}{\sqrt{n\,\log n}}\sum_{k=0}^{n-1} f(B^{3/4}_{k/n})\big[n^{3/2}(B^{3/4}_{(k+1)/n} - B^{3/4}_{k/n})^2 - 1\big] \xrightarrow[n\to\infty]{\text{loi}} C_{3/4}\int_0^1 f(B^{3/4}_s)dW_s \tag{13}$$

  où W est un mouvement brownien standard indépendant de $B^{3/4}$;
- si $H > \frac{3}{4}$ alors

$$n^{1-2H}\sum_{k=0}^{n-1} f(B^H_{k/n})\big[n^{2H}(B^H_{(k+1)/n} - B^H_{k/n})^2 - 1\big] \xrightarrow[n\to\infty]{L^2} \int_0^1 f(B^H_s)dZ_s \tag{14}$$

  où Z est le processus de Rosenblatt défini comme

$$Z_s = I_2^X(L_s), \tag{15}$$

  avec I_2^X l'intégrale stochastique double par rapport à un mouvement brownien X donné par l'équation de transfert (3.3.16) et, pour tout $s \in [0,1]$, L_s est le noyau symétrique de carré intégrable

$$L_s(y_1,y_2) = \frac{1}{2}\mathbf{1}_{[0,s]^2}(y_1,y_2)\int_{y_1\vee y_2}^s \frac{\partial K_H}{\partial u}(u,y_1)\frac{\partial K_H}{\partial u}(u,y_2)du.$$

On voit apparaître deux valeurs critiques $H = 1/4$ et $H = 3/4$. La dernière est une valeur connue pour ce problème puisqu'elle était la seule valeur critique pour les variations quadratiques sans poids, *i.e.* $f \equiv 1$ du mouvement brownien fractionnaire (nous renvoyons le lecteur aux équations (3.3.1)-(3.3.3) collectant différents résultats obtenus par Breuer et Major dans [20], par Dobrushin et Major dans [34], par Giraitis et Surgailis dans [46], et par Taqqu dans [128]). L'introduction d'une fonction de poids dans les variations quadratiques du mouvement brownien fractionnaire donne donc lieu à l'apparition d'une nouvelle valeur critique $H = 1/4$. Les convergences présentées plus haut ne couvrent pas la valeur critique

$H = 1/4$. Le premier résultat de la Section 3.3 est de compléter ce cas manquant et nous montrons dans le Théorème 3.3.1 que

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f(B^{1/4}_{k/n}) \big[\sqrt{n} (B^{1/4}_{(k+1)/n} - B^{1/4}_{k/n})^2 - 1 \big] \xrightarrow[n\to\infty]{\text{loi}} C_{1/4} \int_0^1 f(B^{1/4}_s) dW_s + \frac{1}{4} \int_0^1 f''(B^{1/4}_s) ds \tag{16}$$

où W est un mouvement brownien standard indépendant de $B^{1/4}$ et

$$C^2_{1/4} = \frac{1}{2} \sum_{p=-\infty}^{\infty} \left(\sqrt{|p+1|} + \sqrt{|p-1|} - 2\sqrt{|p|} \right)^2 < \infty.$$

Ce résultat est obtenu en utilisant une méthode développée par Nourdin et Nualart dans [83] et fondée sur le calcul de Malliavin.

Notons que le mouvement brownien fractionnaire de paramètre de Hurst $H = 1/4$ a une interprétation physique en terme de systèmes de particules. En effet plaçons initialement sur la droite réelle un nombre infini de particules (à l'aide d'une loi de Poisson) et supposons que chaque particule suit un mouvement brownien standard indépendant des autres browniens correspondant aux autres particules. On suppose que les collisions entre particules sont "élastiques", alors après renormalisation le mouvement d'une particule fixée suit un mouvement brownien fractionnaire de paramètre de Hurst $H = 1/4$. Ce phénomène a été présenté d'abord par Harris dans [56] puis rigoureusement démontré par Dürr, Goldstein et Lebowitz dans [37]. Notons également que Nourdin et Peccati ont étudié dans [88] les variations quadratiques à poids du mouvement brownien itéré qui n'est pas un processus gaussien mais qui est auto-similaire d'ordre d'auto-similarité 1/4. Ce contexte est un peu différent car la définition des variations quadratiques repose sur une partition aléatoire de l'espace liée à la représentation dite "structure intrinsèque squelettique" du mouvement brownien itéré.

Notre second résultat montre qu'une modification du Théorème 3.3.1 conduit à une première étape vers la construction d'un calcul stochastique par rapport au mouvement brownien fractionnaire de paramètre de Hurst $H = 1/4$ noté $B^{1/4}$. Comme cela est très bien expliqué par Swanson dans [127], on peut définir au moins deux types de sommes de Riemann-Stratonovich afin de définir $\int_0^1 f(B^{1/4}_s) \circ dB^{1/4}_s$ pour une fonction f suffisamment régulière. Le premier type de somme est donné par la "règle du trapèze",

$$S_n(f) = \sum_{k=0}^{n-1} \frac{f(B^{1/4}_{k/n}) + f(B^{1/4}_{(k+1)/n})}{2} \big(B^{1/4}_{(k+1)/n} - B^{1/4}_{k/n} \big).$$

La seconde correspond à la "règle du point milieu"

$$T_n(f) = \sum_{k=1}^{\lfloor n/2 \rfloor} f(B^{1/4}_{(2k-1)/n}) \big(B^{1/4}_{(2k)/n} - B^{1/4}_{(2k-2)/n} \big).$$

Par le Théorème 3 de [84] (voir également [28, 52, 53]), on a que

$$\int_0^1 f'(B_s^{1/4})d^\circ B_s^{1/4} := \lim_{n\to\infty} S_n(f') \quad \text{existe en probabilité}$$

et vérifie la formule classique de changement de variables

$$\int_0^1 f'(B_s^{1/4})d^\circ B_s^{1/4} = f(B_1^{1/4}) - f(0).$$

Concernant la somme T_n, Burdzy et Swanson ont conjecturé dans [127] que

$$\int_0^1 f'(B_s^{1/4})d^\star B_s^{1/4} := \lim_{n\to\infty} T_n(f') \quad \text{existe en loi}$$

et vérifie la formule de changement de variables "non-classique"

$$\int_0^1 f'(B_s^{1/4})d^\star B_s^{1/4} \stackrel{\text{loi}}{=} f(B_1^{1/4}) - f(0) - \frac{\kappa}{2}\int_0^1 f''(B_s^{1/4})dW_s \tag{17}$$

où κ est une constante explicite et W est un mouvement brownien standard indépendant de $B^{1/4}$. Notons qu'en réalité Burdzy et Swanson ont conjecturé (17) non pas pour $B^{1/4}$ mais pour la solution de l'équation de la chaleur stochastique notée F,

$$F_t = u(t,0),\ t \in [0,1],$$

avec

$$u_t = \frac{1}{2}u_{xx} + \dot{W}(t,x),\ t\in[0,1],\ x\in\mathbb{R}, \quad \text{avec la condition initiale } u(0,x)=0.$$

Ici $\dot{W}$ est un bruit blanc "espace-temps" sur $[0,1]\times\mathbb{R}$. On vérifie que F est un processus gaussien centré de fonction de covariance

$$E(F_sF_t) = \frac{1}{\sqrt{2\pi}}\big(\sqrt{t+s} - \sqrt{|t-s|}\big).$$

Ainsi, F est un processus bifractionnaire de paramètres $\frac{1}{2}$ et $\frac{1}{2}$ dans le sens d'Houdré et Villa ([57]). En utilisant le résultat principal de Lei et Nualart ([74]), on obtient que $B^{1/4}$ et F diffèrent en loi d'un processus à trajectoires absolument continues. Par conséquent, une transformation de type Girsanov nous permet de montrer qu'il est équivalent de résoudre la conjecture de Burdzy-Swanson énoncée plus haut pour $B^{1/4}$ ou pour F.

Notre second résultat (Théorème 3.3.2) montre que la conjecture de Burdzy-Swanson (17) est vraie.

Notons que cette conjecture a été résolue simultanément et indépendamment par Burdzy et Swanson dans [21] par une méthode différente de celle utilisée dans la Section 3.3. De plus Burdzy et Swanson obtiennent une convergence en tant que processus sous des hypothèses un peu plus faibles que les nôtres. Cependant les auteurs n'étudient pas le cas du mouvement brownien fractionnaire d'indice 1/4 comme nous le faisons dans la Section 3.3.

Chapitre 4 : Méthode de Stein pour l'approximation gaussienne multidimensionnelle

Dans le Chapitre 4 nous présentons un travail en collaboration avec Ivan Nourdin et Giovanni Peccati dans lequel nous obtenons des bornes explicites pour l'approximation (relative à la distance de Wasserstein) gaussienne multidimensionnelle de certaines fonctionnelles de champs gaussiens. Pour cela nous utilisons la méthode de Stein et du calcul de Malliavin.

La méthode de Stein est basée sur la caractérisation suivante d'une variable aléatoire gaussienne réelle.

Lemma 10. *Une variable aléatoire* Y *est telle que* $Y \overset{\text{loi}}{=} Z \sim \mathscr{N}(0,1)$ *si et seulement si pour toute fonction continue et continûment différentiable par morceaux, avec* $E\big|f'(Z)\big| < \infty$*, on a*

$$E[f'(Y) - Yf(Y)] = 0. \tag{18}$$

Ainsi la formule d'intégration par parties (18) caractérise la loi gaussienne. Heuristiquement on peut donc penser que la distance de Wasserstein entre Y et Z est faible si

$$E[f'(Y) - Yf(Y)] \simeq 0,$$

pour f appartenant à une classe de fonctions suffisamment large. Dans les travaux [123, 125], Stein formalise cet argument en utilisant l'équation de Stein.

Soit $g : \mathbb{R} \to \mathbb{R}$ appartenant à une classe de fonctions $\mathscr{H}$. On cherche une fonction $h : \mathbb{R} \to \mathbb{R}$ solution de l'équation de Stein,

$$g(x) - E[g(Z)] = h'(x) - xh(x), \quad \forall x \in \mathbb{R}, \tag{19}$$

où $Z \sim \mathscr{N}(0,1)$. Considérons $\mathscr{H}$ l'ensemble des fonctions lipschitziennes g avec $\|g\|_{Lip} \leq 1$. Stein a montré dans [125] que pour tout g dans $\mathscr{H}$ il existe une solution h de l'équation de Stein (19) vérifiant $\|h'\|_\infty \leq 1$ et $\|h''\|_\infty \leq 2$. On a :

$$g(Y) - E[g(Z)] = h'(Y) - Yh(Y)$$

et donc

$$E[g(Y)] - E[g(Z)] = E[h'(Y) - Yh(Y)].$$

Ainsi comme $d_{\mathrm{W}}(Y,X) = \sup_{g \in \mathrm{Lip}(1)} |E[g(Y)] - E[g(X)]|$, on a que

$$d_{\mathrm{W}}(Y,Z) \leq \sup_{f \in \mathscr{F}_{\mathrm{W}}} |E[f'(Y) - Yf(Y)]|, \tag{20}$$

où $\mathscr{F}_{\mathrm{W}} = \{f : \|f'\|_\infty \leq 1,\ \|f''\|_\infty \leq 2\}$.

Des méthodes très différentes pour majorer la partie droite de (20) sont proposées dans une littérature très abondante et nous renvoyons le lecteur par exemple à l'article (et aux références présentées dans ce dernier) de Reinert [111]. Le point de départ du travail présenté dans le Chapitre 4 est une majoration de la partie droite de l'inégalité (20) réalisée par Nourdin et Peccati dans [86] puis exploitée dans [87] lorsque Y est une fonctionnelle centrée d'un certain champ gaussien infini-dimensionnel. Dans ce contexte Nourdin et Peccati obtiennent que

$$d_{\mathrm{W}}(Y,Z) \leq E[(1-\langle DY,-DL^{-1}Y\rangle_{\mathfrak{H}})^2]^{1/2}, \tag{21}$$

où D est le gradient de Malliavin. Comme cela est démontré dans [86], si Y est une intégrale multiple de type Wiener-Itô, la borne présentée ci-dessus (21) est reliée aux théorèmes de la limite centrale donnés par Nualart et Peccati dans [96] et par Nualart et Ortiz-Latorre dans [95].

Dans le Chapitre 4 nous donnons (Théorème 4.3.5) un équivalent de la majoration (21) lorsque Z est remplacé par un vecteur d-dimensionnel gaussien centré, $F=(F_1,...,F_d)$ est un vecteur de fonctionnelles régulières d'un champ gaussien et d_{W} est la distance de Wasserstein entre des mesures de probabilités sur $\mathbb{R}^d$ (voir Définition 4.3.1). Nos résultats s'appliquent à l'approximation de vecteurs gaussiens de matrice de covariance définie positive donnée. Plus précisément, on rappelle ci-dessous l'énoncé du Théorème 4.3.5.

Soit $d\geq 2$, soit $C=\{C(i,j) : i,j=1,...,d\}$ une matrice définie positive. Supposons que $Z\sim\mathscr{N}_d(0,C)$ et que $F:=(F_1,\dots,F_d)$ est un vecteur aléatoire à valeurs dans $\mathbb{R}^d$ avec $E[F_i]=0$ et $F_i\in\mathbb{D}^{1,2}$ pour tout $i=1,\dots,d$. Alors,

$$\begin{aligned} d_{\mathrm{W}}(F,Z) &\leq \|C^{-1}\|_{op}\,\|C\|_{op}^{1/2}\sqrt{E\|C-\Phi(DF)\|_{H.S}^2} \\ &= \|C^{-1}\|_{op}\,\|C\|_{op}^{1/2}\sqrt{\sum_{i,j=1}^{d}E[(C(i,j)-\langle DF_i,-L^{-1}DF_j\rangle_{\mathfrak{H}})^2]}, \end{aligned}$$

où $\Phi(DF)$ est la matrice

$$\Phi(DF)=\{\langle DF_i,-L^{-1}DF_j\rangle_{\mathfrak{H}} : 1\leq i,j\leq d\}.$$

Dans le cas où F est un vecteur d'intégrales multiples de type Wiener-Itô on obtient le Corollaire 4.3.6 résumé comme suit.

Soit $d\geq 2$ et $1\leq q_1\leq\ldots\leq q_d$. On considère un vecteur $F:=(F_1,\dots,F_d)=(I_{q_1}(f_1),\dots,I_{q_d}(f_d))$ avec $f_i\in\mathfrak{H}^{\odot q_i}$ pour tout $i=1\dots,d$. Soit $Z\sim\mathscr{N}_d(0,C)$, avec C matrice définie positive. Alors,

$$d_{\mathrm{W}}(F,Z)\leq\|C^{-1}\|_{op}\,\|C\|_{op}^{1/2}\sqrt{\sum_{1\leq i,j\leq d}E\left[\left(C(i,j)-\frac{1}{q_j}\langle DF_i,DF_j\rangle_{\mathfrak{H}}\right)^2\right]}. \tag{22}$$

Ceci nous permet de généraliser dans la Proposition 4.3.10 les résultats multidimensionnels obtenus dans [95, 96] montrant que pour des vecteurs d'intégrales multiples de Wiener-Itô la distance de Wasserstein métrise la convergence en loi vers un vecteur gaussien de matrice de covariance définie positive donnée. Notons qu'une partie importante de nos calculs est directement inspirée de ceux réalisés par Nualart et Ortiz-Latorre dans [95].

Nous appliquons ensuite nos résultats, dans un premier temps, à l'obtention de bornes explicites pour le Théorème de la limite centrale de Breuer et Major pour le mouvement brownien fractionnaire (voir Théorème 4.4.1), puis dans un second temps, à l'approximation gaussienne (Proposition 4.4.3) de fonctionnelles de vecteurs gaussiens généralisant ainsi un résultat prouvé par Chatterjee dans [24].

Chapitre 1

Estimation de Stein

Ce chapitre est constitué des références [107, 105, 106, 108].

1.1 Cas d'une martingale gaussienne à accroissements indépendants

Cet article est réalisé en collaboration avec Nicolas Privault et publié en [107].

1.1.1 Notation and preliminaries

The maximum likelihood estimator $\hat{\mu}$ of the mean $\mu \in \mathbb{R}^d$ of a Gaussian random vector X in $\mathbb{R}^d$ with covariance $\sigma^2\mathrm{I}$ under a probability $\mathbb{P}_\mu$ is well-known to be equal to X itself. It is efficient in the sense that it attains the Cramer-Rao bound

$$\sigma^2 d = \mathbb{E}_\mu[\|X-\mu\|_d^2] = \inf_Z \mathbb{E}_\mu[\|Z-\mu\|_d^2],$$

over all unbiased estimators Z of $\mu \in \mathbb{R}^d$. In [124], Stein used integration by parts with respect to the Gaussian density to prove the identity

$$\mathbb{E}_\mu[\|X+\sigma^2\partial \log f(X)-\mu\|_d^2] = \sigma^2 d + 4\sigma^4 \sum_{i=1}^d \mathbb{E}_\mu\left[\frac{\partial_i^2\sqrt{f}(X)}{\sqrt{f}(X)}\right],$$

which shows that if $\sqrt{f}$ is superharmonic on $\mathbb{R}^d$, then $X+\sigma^2 \operatorname{grad}\log f$ is a superefficient estimator of μ.

In this Note we construct nonparametric superefficient estimators of the drift $(u_t)_{t\in[0,T]}$ of a Brownian motion X_t, using the integration by parts formula of the Malliavin calculus and harmonic analysis on the Wiener space. Let $(\Omega, H, \mathbb{P}^\sigma)$ denote the Wiener space, where $\Omega = \mathcal{C}_0([0,T])$ is the space of continuous functions on $[0,T]$ starting at 0,

$$H = \left\{ v : [0,T] \to \mathbb{R} \ : \ v(t) = \int_0^t \dot{v}(s)ds, \ t \in [0,T], \ \dot{v} \in L^2([0,T]) \right\}$$

is the Cameron-Martin space with inner product

$$\langle v_1, v_2\rangle_H = \int_0^T \dot{v}_1(s)\dot{v}_2(s)ds, \qquad v_1, v_2 \in H,$$

and $\mathbb{P}^\sigma$ is the Wiener measure with variance $\sigma^2 > 0$. Let $(\mathcal{F}_t)_{t\in[0,T]}$ denote the filtration generated by the canonical process $(X_t)_{t\in[0,T]}$, and for $u \in \Omega$, let $\mathbb{P}_u^\sigma$ denote the translation of the Wiener measure on Ω by u, i.e. $(X_t - u_t)_{t\in[0,T]}$ is a standard Brownian motion with variance $\sigma^2 > 0$ under $\mathbb{P}_u^\sigma$.

We fix $(h_n)_{n\geq 1}$ a total subset of H and let $\mathcal{S}$ denote the space of cylindrical functionals of the form

$$F = f_n\left(\int_0^T \dot{h}_1(s)dX_s, \dots, \int_0^T \dot{h}_n(s)dX_s\right), \tag{1.1.1}$$

where f_n is in the space of infinitely differentiable rapidly decreasing functions on $\mathbb{R}^n$, $n \geq 1$. The H-valued Malliavin derivative D, see [132] and references therein, is defined as

$$D_tF = \sum_{i=1}^n h_i(t)\partial_i f_n\left(\int_0^T \dot{h}_1(s)dX_s, \dots, \int_0^T \dot{h}_n(s)dX_s\right),$$

for $F \in \mathcal{S}$ of the form (1.1.1). Let $\delta : L_u^2(\Omega; H) \to L_u^2(\Omega)$ denote the divergence operator under $\mathbb{P}_u^\sigma$, which satisfies the integration by parts formula

$$\mathbb{E}_u^\sigma[F\delta(v)] = \sigma^2\, \mathbb{E}_u^\sigma[\langle v, DF\rangle_H], \qquad v \in \text{Dom}(\delta). \tag{1.1.2}$$

It is well-known that D and δ are closable, and their domains will be denoted by Dom(D) and Dom(δ). We define the Laplacian Δ by

$$\Delta F = \text{trace}_{L^2([0,T])^{\otimes 2}} DDF = \int_0^T D_sD_sFds, \qquad F \in \mathcal{S},$$

i.e.

$$\Delta F = \sum_{i,j=1}^n \int_0^T h_i(s)h_j(s)ds\ \partial_i\partial_j f_n\left(\int_0^T \dot{h}_1(s)dX_s, \dots, \int_0^T \dot{h}_n(s)dX_s\right),$$

if F has the form (1.1.1). Note that unlike the Gross Laplacian Δ_G defined by $\Delta_G F = \text{trace}_{H^{\otimes 2}} DDF$, the operator Δ is closable, and its domain will be denoted by Dom(Δ).

1.1.2 Maximum likelihood and Bayes estimators

An estimator ξ of u is called unbiased if

$$\mathbb{E}_u^\sigma[\xi_t] = u_t, \quad t \in [0,T], \quad u \in H,$$

and adapted if $(\xi_t)_{t\in[0,T]}$ is $\mathcal{F}_t$-adapted. Here, the process $\hat{u} = (X_t)_{t\in[0,T]}$ will be considered as an unbiased estimator of its own drift $(u_t)_{t\in[0,T]}$ under $\mathbb{P}_u^\sigma$. In particular, given N

independent samples $(X_t^1)_{t\in[0,T]}, \ldots, (X_t^N)_{t\in[0,T]}$ of $(X_t)_{t\in[0,T]}$, the process $\bar{X}_t := (X_t^1 + \cdots + X_t^N)/N$, $t \in [0,T]$, is an unbiased and consistent estimator of $(u_t)_{t\in[0,T]}$ as N goes to infinity. Moreover, $\hat{u} := (X_t)_{t\in[0,T]}$ is a maximum likelihood estimator (MLE) of $(u_t)_{t\in[0,T]}$ in the sense that it formally maximizes the Girsanov density

$$\frac{d\mathbb{P}_v^\sigma}{d\mathbb{P}^\sigma} = \exp\left(\int_0^T \frac{\dot{v}_s}{\sigma^2} dX_s - \frac{1}{2}\int_0^T \frac{\dot{v}_s^2}{\sigma^2} ds\right)$$

in $v \in H$. Moreover it attains the Cramer-Rao bound $\sigma^2 T^2/2$ presented in the next proposition.

Proposition 1.1.1. *For any unbiased and adapted estimator ξ of u we have*

$$\mathbb{E}_u^\sigma\left[\int_0^T |\xi_t - u_t|^2 dt\right] \geq \mathbb{E}_u^\sigma\left[\int_0^T |X_t - u_t|^2 dt\right] = \sigma^2 \frac{T^2}{2}, \qquad u \in H. \tag{1.1.3}$$

For any $\tau > 0$ and $v \in H$, the Bayes risk

$$\int_\Omega \mathbb{E}_z^\sigma\left[\int_0^T |\xi_t - z_t|^2 dt\right] d\mathbb{P}_v^\tau(z) \tag{1.1.4}$$

of any estimator $(\xi_t)_{t\in[0,T]}$ on Ω under the prior distribution $\mathbb{P}_v^\tau$ is uniquely minimized by

$$\xi_t^{\tau,v} := \frac{\sigma^2}{\tau^2+\sigma^2} v_t + \frac{\tau^2}{\tau^2+\sigma^2} X_t, \qquad t \in [0,T],$$

which has risk $\tau^2\sigma^2T^2/(2(\tau^2+\sigma^2))$. Clearly $\xi^{\tau,v}$ is unique in the sense that it is the only estimator to minimize (1.1.4), hence it is also admissible. However $\xi^{\tau,v}$ is not minimax since for all $u \in H$, its mean square error under $\mathbb{P}_u^\sigma$ is equal to

$$\mathbb{E}_u^\sigma\left[\int_0^T |\xi_t^{\tau,v} - u_t|^2 dt\right] = \frac{\sigma^2\tau^4}{(\tau^2+\sigma^2)^2}\frac{T^2}{2} + \frac{\sigma^4}{(\tau^2+\sigma^2)^2}\int_0^T \left|\int_0^t \dot{v}_s - \dot{u}_s ds\right|^2 dt.$$

Since the Bayes risk $\tau^2\sigma^2T^2/(2(\tau^2+\sigma^2))$ of $\xi^{\tau,v}$, $\tau \in \mathbb{R}$, converges as $\tau \to \infty$ to the Cramer-Rao bound $\sigma^2T^2/2$, it follows that the maximum likelihood estimator $\hat{u} = (X_t)_{t\in[0,T]}$ is minimax, i.e. for all $u \in \Omega$ we have

$$\sigma^2\frac{T^2}{2} = \mathbb{E}_u^\sigma\left[\int_0^T |X_t - u_t|^2 dt\right] = \inf_\xi \sup_{v\in\Omega} \mathbb{E}_v^\sigma\left[\int_0^T |\xi_t - v_t|^2 dt\right]. \tag{1.1.5}$$

1.1.3 Superefficient drift estimators

We aim to construct a superefficient estimator of u of the form $X + \xi$, with mean square error strictly smaller than the Cramer-Rao bound when ξ is a suitably chosen stochastic process. This estimator will be biased and possibly anticipating with respect to the Brownian filtration. For this we carry over Stein's argument to the Wiener space using the duality relation (1.1.2) between the gradient and divergence operators.

Lemma 1.1.2. *Unbiased risk estimate. For any* $\xi \in L^2_u(\Omega \times [0,T])$ *such that* $\xi_t \in \mathrm{Dom}(D)$, $t \in [0,T]$, *we have*

$$\mathbb{E}^\sigma_u \left[\|X + \xi - u\|^2_{L^2([0,T])} \right] = \sigma^2 \frac{T^2}{2} + \|\xi\|^2_{L^2_u(\Omega\times[0,T])} + 2\sigma^2 \, \mathbb{E}^\sigma_u \left[\int_0^T D_t \xi_t dt \right]. \qquad (1.1.6)$$

The next proposition specializes the above lemma to processes ξ of the form

$$\xi_t = \sigma^2 D_t \log F, \qquad t \in [0,T],$$

where F is an a.s. strictly positive and sufficiently smooth random variable.

Proposition 1.1.3. *Stein type estimator. Let* $F \in \mathcal{S}$ *be such that* $F > 0$, $\mathbb{P}^\sigma$*-a.s. and* $\sqrt{F} \in \mathrm{Dom}(\Delta)$. *We have*

$$\mathbb{E}^\sigma_u \left[\|X + \sigma^2 D \log F - u\|^2_{L^2([0,T])} \right] = \sigma^2 \frac{T^2}{2} + 4\sigma^4 \, \mathbb{E}^\sigma_u \left[\frac{\Delta\sqrt{F}}{\sqrt{F}} \right]. \qquad (1.1.7)$$

Relation (1.1.7) extends to any $F \in \mathrm{Dom}(D)$ such that $\sqrt{F} \in \mathrm{Dom}(\Delta)$, and $F > 0$, $\Delta\sqrt{F} \le 0$, $\mathbb{P}^\sigma$-a.s., and in particular, $X + \sigma^2 D \log F$ is a superefficient estimator of u if $\Delta\sqrt{F} < 0$ on a set of strictly positive $\mathbb{P}^\sigma$-measure. As in [124], the superefficient estimators constructed in this way are minimax in the sense that for all $u \in H$ we have

$$\mathbb{E}^\sigma_u \left[\|X + \sigma^2 D \log F - u\|^2_{L^2([0,T])} \right] < \sigma^2 \frac{T^2}{2} = \inf_\xi \sup_{v\in\Omega} \mathbb{E}^\sigma_v \left[\int_0^T |\xi_t - v_t|^2 dt \right],$$

thus showing that the MLE $\hat{u} = (X_t)_{t\in[0,T]}$ is inadmissible. In contrast to the MLE, such estimators are not only biased but also anticipating with respect to the Brownian filtration $(\mathcal{F}_t)_{t\in[0,T]}$. For all $v \in H$ the estimator

$$X_t + \sigma^2 \, \mathbb{E}^\sigma_v [D_t \log F \mid \mathcal{F}_t]$$

has risk

$$\begin{aligned} &\mathbb{E}^\sigma_u \left[\int_0^T |X_t + \sigma^2 \, \mathbb{E}^\sigma_v[D_t \log F \mid \mathcal{F}_t] - u_t|^2 dt \right] \\ &= \mathbb{E}^\sigma_u \left[\|X + \sigma^2 D \log F - u\|^2_{L^2([0,T])} \right] - \sigma^4 \, \mathbb{E}^\sigma_u \left[\int_0^T |D_t \log F - \mathbb{E}^\sigma_v[D_t \log F \mid \mathcal{F}_t]|^2 dt \right] \\ &< \sigma^2 \frac{T^2}{2} + 4\sigma^4 \, \mathbb{E}^\sigma_u \left[\frac{\Delta\sqrt{F}}{\sqrt{F}} \right] < \sigma^2 \frac{T^2}{2}, \end{aligned}$$

when $\sqrt{F}$ is Δ-superharmonic, thus showing that our Stein type estimators are not admissible since none of them is adapted. Actually, only adapted estimators of u can be admissible and as a consequence, none of the Bayes estimators $\xi^{\tau,v}$ can be written under the form of Stein type estimators.

1.1.4 Numerical application

Here we assume that $(h_k)_{k\geq 1}$ is orthonormal in $L^2([0,T])$, and additionally that $(h_k)_{k\geq 1}$ is orthogonal in H, for instance one can take

$$h_n(t) = \sqrt{\frac{2}{T}}\sin\left(\left(n-\frac{1}{2}\right)\frac{\pi t}{T}\right), \qquad t\in[0,T], \quad n\geq 1.$$

Superharmonic functionals on the Wiener space can be constructed as cylindrical functionals, by composition with finite-dimensional functions. Given $a\in\mathbb{R}$, let

$$F_{n,a} = \left(\left|\int_0^T \dot h_1(t)dX_t\right|^2 + \cdots + \left|\int_0^T \dot h_n(t)dX_t\right|^2\right)^{a/2}.$$

We have

$$\frac{\Delta\sqrt{F_{n,a}}}{\sqrt{F_{n,a}}} = \frac{a(n-2+a/2)/2}{\left|\int_0^T \dot h_1(t)dX_t\right|^2 + \cdots + \left|\int_0^T \dot h_n(t)dX_t\right|^2},$$

which is negative if $4-2n\leq a\leq 0$, and minimal for $a=2-n$. The estimator of u will be given by

$$D_t\log F_{n,2-n} = -(n-2)\sqrt{\frac{2}{T}}\sum_{k=1}^n \frac{\displaystyle\int_0^T \dot h_k(s)dX_s \sin\left(\left(k-\frac{1}{2}\right)\frac{\pi t}{T}\right)}{\left|\int_0^T \dot h_1(s)dX_s\right|^2 + \cdots + \left|\int_0^T \dot h_n(s)dX_s\right|^2},$$

with

$$\|D\log F_{n,2-n}\|^2_{L^2([0,T])} = \frac{(n-2)^2}{\left|\int_0^T \dot h_1(s)dX_s\right|^2 + \cdots + \left|\int_0^T \dot h_n(s)dX_s\right|^2}. \tag{1.1.8}$$

For simulation purposes we construct the (nondrifted) Brownian motion $(Y_t)_{t\in[0,T]}$ via the Paley-Wiener series

$$Y_t = \sigma\frac{\sqrt{2T}}{\pi}\sum_{n=1}^\infty \eta_n \frac{\sin\left(\left(n-\frac{1}{2}\right)\frac{\pi t}{T}\right)}{\left(n-\frac{1}{2}\right)}, \tag{1.1.9}$$

where $(\eta_n)_{n\geq 1}$ are independent standard Gaussian random variables with unit variance under $\mathbb{P}_u^\sigma$. Unlike in classical Stein estimation, n becomes a free parameter and there is some interest in determining optimal values of n. Next we present numerical simulations which allow us to measure the efficiency of our estimators. The gain of the superefficient estimator $X + D\log F_{n,2-n}$ compared to that of the MLE is given by

$$G(u,\sigma,T,n) := -\frac{8}{\sigma^2T^2}\,\mathbb{E}_u^\sigma\left[\frac{\Delta\sqrt{F_{n,2-n}}}{\sqrt{F_{n,2-n}}}\right] = \frac{2}{\sigma^2T^2}\,\mathbb{E}_u^\sigma\left[\|D\log F_{n,2-n}\|^2_{L^2([0,T])}\right].$$

Finally we choose $u_t = \alpha t$, $t \in [0, T]$, for some $\alpha \in \mathbb{R}$. In this case, $G(\alpha, \sigma, T, n)$ converges to

$$(n-2)^2 \frac{8}{\pi^2} \mathbb{E}\left[\left(\sum_{l=1}^{n} (2l-1)^2 \eta_l^2\right)^{-1}\right],$$

when $\alpha^{-2}\sigma^2/T$ tends to infinity, which yields 11.38% for $n = 4$, and is equivalent to $\alpha^{-2}(1-2/n)^2\sigma^2/T$ as $\alpha^{-2}\sigma^2/T$ tends to 0, and to $6/(n\pi^2)$ as n goes to infinity. See Figure 1.4 in Section 1.2.

1.2 Cas d'un processus gaussien

Cette section est une version détaillée de l'article écrit avec Nicolas Privault publié en [105] et de la prépublication [106] réalisée avec Nicolas Privault.

1.2.1 Introduction

The maximum likelihood estimator $\hat{\mu}$ of the mean $\mu \in \mathbb{R}^d$ of a Gaussian random vector X in $\mathbb{R}^d$ with covariance $\sigma^2 \mathrm{I}_{\mathbb{R}^d}$ under a probability $\mathbb{P}_\mu$ is well-known to be equal to X itself, and can be computed by maximizing the likelihood ratio

$$\frac{1}{(2\pi\sigma^2)^{d/2}} e^{-\frac{\|X-m\|_d^2}{2\sigma^2}}$$

with respect to m, where $\|\cdot\|_d$ denotes the Euclidean norm on $\mathbb{R}^d$. It is efficient in the sense that it attains the Cramer-Rao bound

$$\sigma^2 d = \mathbb{E}_\mu[\|X-\mu\|_d^2] = \inf_Z \mathbb{E}_\mu[\|Z-\mu\|_d^2], \qquad \mu \in \mathbb{R}^d,$$

over all unbiased estimators Z satisfying $\mathbb{E}_\mu[Z] = \mu$, for all $\mu \in \mathbb{R}^d$.

In [68], James and Stein have constructed superefficient estimators for the mean of $X \in \mathbb{R}^d$, of the form

$$\left(1 - \frac{d-2}{\|X\|_d^2}\right) X$$

whose risk is lower than the Cramer Rao bound $\sigma^2 d$ in dimension $d \geq 3$.

Drift estimation for Gaussian processes is of interest in several fields of application. For example in the decomposition

$$X_t = X_t^u + u_t, \qquad t \in [0,T],$$

the process $(X_t)_{t\in[0,T]}$ is interpreted as an observed output signal, the drift $(u_t)_{t\in[0,T]}$ is viewed as an input signal to be estimated and perturbed by a centered Gaussian noise $(X_t^u)_{t\in[0,T]}$, cf. e.g. [59], Ch. VII. Such results find applications in e.g. telecommunication (additive Gaussian channels) and finance (identification of market trends).

Berger and Wolpert [14], [135], have constructed estimators of James-Stein type for the drift of a Gaussian process $(X_t)_{t\in[0,T]}$ by applying the James-Stein procedure to the independent Gaussian random variables appearing in the Karhunen-Loève expansion of the process. In this context, $\hat{u} := (X_t)_{t\in\mathbb{R}_+}$ is seen as a minimax estimator of its own drift $(u_t)_{t\in\mathbb{R}_+}$.

Stein [124] has shown that the James-Stein estimators on $\mathbb{R}^d$ could be extended to a wider family of estimators, using integration by parts for Gaussian measures. Let us briefly recall Stein's argument, which relies on integration by parts with respect to the Gaussian density and on the properties of superharmonic functionals for the Laplacian on $\mathbb{R}^d$. Given an estimator of $\mu \in \mathbb{R}^d$ of the form $X + g(X)$, where $g : \mathbb{R}^d \to \mathbb{R}^d$ is sufficiently smooth, and applying the integration by parts formula

$$\mathbb{E}_\mu[(X_i - \mu_i) g_i(X)] = \sigma^2 \, \mathbb{E}_\mu[\partial_i g_i(X)], \tag{1.2.1}$$

$g = \sigma^2 \operatorname{grad} \log f = \sigma^2(\partial_1 \log f, \ldots, \partial_d \log f)$, one obtains

$$\mathbb{E}_\mu[\|X + \sigma^2 \operatorname{grad} \log f(X) - \mu\|_d^2] = \sigma^2 d + 4\sigma^4 \sum_{i=1}^d \mathbb{E}_\mu \left[\frac{\partial_i^2 \sqrt{f}(X)}{\sqrt{f}(X)} \right],$$

i.e. $X + \sigma^2 \operatorname{grad} \log f(X)$ is a superefficient estimator if

$$\sum_{i=1}^d \partial_i^2 \sqrt{f}(x) < 0, \qquad dx - a.e.,$$

which is possible if $d \geq 3$. In this case, $X + \sigma^2 \operatorname{grad} \log f(X)$ improves in the mean square sense over the efficient estimator $\hat{u}$ which attains the Cramer-Rao bound $\sigma^2 d$ on unbiased estimators of μ.

In this paper we present an extension of Stein's argument to an infinite-dimensional setting using the Malliavin integration by parts formula, with application to the construction of Stein type estimators for the drift of a Gaussian process $(X_t)_{t\in[0,T]}$. Our approach applies to Gaussian processes such as Volterra processes and fractional Brownian motions. It also extends the results of Berger and Wolpert [14] in the same way that the construction of Stein [124] extends that of James and Stein [68], and this allows us to recover the estimators of James-Stein type introduced by Berger and Wolpert [14] as particular cases. Here we replace the Stein equation (1.2.1) with the integration by parts formula of the Malliavin calculus on Gaussian space. Our estimators are given by processes of the form

$$X_t + D_t \log F, \qquad t \in [0, T],$$

where F is a positive superharmonic random variable on Gaussian space and D_t is the Malliavin derivative indexed by $t \in [0, T]$. In contrast to the minimax estimator $\hat{u}$, such estimators are not only biased but also anticipating with respect to the Brownian filtration $(\mathcal{F}_t)_{t\in[0,T]}$. This however poses no problem when one has access to complete paths from time 0 to T.

For large values of σ it can be shown that the percentage gain of this estimator is at least equal to the universal constant

$$\frac{16}{\pi^4} \int_{\mathbb{R}^4} e^{-\frac{x^2+y^2+z^2+r^2}{2}} \frac{dxdydzdr}{x^2 + 9y^2 + 25z^2 + 49r^2} \tag{1.2.2}$$

which approximately represents 11.38%, see (1.2.33) below.

We proceed as follows. In Section 1.2.2 we use stochastic calculus in the independent increment case to derive a Cramer-Rao bound over all unbiased drift estimators. This bound is attained by the process $\hat{u} := (X_t)_{t\in[0,T]}$, which will be considered as an efficient drift estimator. In Section 1.2.3 we compute the Bayes estimators obtained under prior Gaussian distributions. We show that these Bayes estimators are admissible, and use them to prove that the drift estimator $\hat{u}$ is minimax. The tools and results presented in Sections 1.2.2 and 1.2.3 are not surprising, but we did not find any source covering them in the literature. In Section 1.2.4 we recall the elements of analysis and integration by parts on Gaussian space which will be needed in Section 1.2.5 to construct superefficient drift estimators for Gaussian processes using superharmonic random functionals on Gaussian space. The superefficiency of these estimators will show, as in the classical case, that the minimax estimator $\hat{u}$ is not admissible. In Section 1.2.6 we give examples of nonnegative superharmonic functionals using cylindrical functionals and potential theory on Gaussian space. Examples are considered in Section 1.2.7 in case u is deterministic. We show that the James-Stein estimators of Berger and Wolpert [14] can be recovered as particular cases in our approach, and we provide numerical simulations for the gain of such estimators. It turns out that in those examples, the gain obtained in comparison with the minimax estimator $(X_t)_{t\in[0,T]}$ is a function of σ^2/T, thus making σ and T play inverse roles, unlike in the usual setting of Brownian rescaling.

This paper is an extended version of [105] and provides proofs of the results presented in [107].

Notation

Let $T > 0$. Consider a real-valued centered Gaussian process $(X_t)_{t\in[0,T]}$ with covariance function

$$\gamma(s,t) = \mathbb{E}[X_s X_t], \qquad s,t \in [0,T],$$

on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is the σ-algebra generated by X. Recall that $(X_t)_{t\in[0,T]}$ can be represented in different ways as an isonormal Gaussian process on a real separable Hilbert space H, i.e. as an isometry $X : H \to L^2(\Omega, \mathcal{F}, P)$ such that $\{X(h) : h \in H\}$ is a family of centered Gaussian random variables satisfying

$$\mathbb{E}[X(h)X(g)] = \langle h, g\rangle_H, \qquad h, g \in H,$$

where $\langle \cdot, \cdot \rangle_H$ and $\| \cdot \|_H$ denote the scalar product and norm on H.

One can distinguish two main types of such isonormal representations of X_t, see e.g. [5] and [14] respectively for details.

(A) Paley-Wiener expansions. In this case, H is the completion of the linear space generated by the functions $\chi_t(s) = \min(s,t)$, $s,t \in [0,T]$, with respect to the norm

$$\langle \chi_t, \chi_s \rangle_H := \gamma(s,t), \qquad s,t \in [0,T],$$

and $X(\cdot)$ is constructed on H from $X(\chi_t) := X_t$, $t \in [0,T]$, i.e. we have

$$X(\chi_t) = \sum_{k=0}^{\infty} \langle \chi_t, h_k(t) \rangle_H X(h_k), \qquad t \in [0,T],$$

for any orthonormal basis $(h_k)_{k\in\mathbb{N}}$ of H. Assume in addition $\gamma(s,t)$ has the form

$$\gamma(s,t) = \int_0^{s\wedge t} K(t,r)K(s,r)dr, \qquad s,t \in [0,T],$$

where $K(\cdot,\cdot)$ is a deterministic kernel and

$$(Kh)(t) := \int_0^t K(t,s)\dot{h}(s)ds$$

is differentiable in $t \in [0,T]$, and let K^* denote the adjoint of K with respect to

$$\langle h, g \rangle := \langle \dot{h}, \dot{g} \rangle_{L^2([0,T],dt)}.$$

The scalar product in H then satisfies

$$\langle h, g \rangle_H = \langle K^* h, K^* g \rangle = \langle h, \Gamma g \rangle,$$

where $\Gamma = KK^*$, and we have the decomposition

$$X_t = \sum_{k=0}^{\infty} \langle \chi_t, h_k \rangle_H X(h_k) = \sum_{k=0}^{\infty} \langle 1_{[0,t]}, \dot{\Gamma} h_k \rangle_{L^2([0,T],dt)} X(h_k) = \sum_{k=0}^{\infty} \Gamma h_k(t) X(h_k),$$

$t \in [0,T]$. In this case we also have the representation

$$X_t = \int_0^t K(t,s)dW_s, \qquad t \in [0,T],$$

where $(W_s)_{s\in[0,T]}$ is a standard Brownian motion, cf. [5].

(B) Karhunen-Loève expansions. This framework is used in [14]. In this case, μ is a finite Borel measure on $[0,T]$ and H is defined from

$$\langle h, g \rangle_H = \langle h, \Gamma g \rangle,$$

where

$$\langle h, g \rangle := \langle h, g \rangle_{L^2([0,T],d\mu)},$$

and

$$(\Gamma g)(t) = \int_0^T g(s)\gamma(s,t)\mu(ds), \qquad t \in [0,T],$$

with

$$X(h) = \int_0^T X_s h(s)\mu(ds), \qquad h \in H.$$

Given $(h_k)_{k\in\mathbb{N}}$ an orthonormal basis of $L^2([0,T], d\mu)$, we have the expansion

$$X_t = \sum_{k=0}^{\infty} h_k(t) X(h_k), \qquad t \in [0,T].$$

In the sequel we will use mainly the framework (A) with $\langle h, g\rangle = \langle \dot{h}, \dot{g}\rangle_{L^2([0,T],dt)}$, which is better adapted to our approach, although some results valid in the general framework of Gaussian processes will be valid for (B) as well.

The Girsanov theorem for Gaussian processes, cf. e.g. [94], states that X^u defined as

$$X^u(g) := X(g) - \langle g, u\rangle = X(g) - \langle g, \Gamma^{-1}u\rangle_H, \qquad g \in H,$$

where $u \in H$ is deterministic, has same law as X under the probability $\mathbb{P}_u$ defined by

$$\frac{d\mathbb{P}_u}{d\mathbb{P}} = \exp\left(X(\Gamma^{-1}u) - \frac{1}{2}\|\Gamma^{-1}u\|_H^2 \right).$$

In other terms, in case (A) we have

$$\begin{aligned} X_t - u(t) &= X_t - \Gamma\Gamma^{-1}u(t) \\ &= \sum_{k=0}^{\infty} \Gamma h_k(t)(X(h_k) - \langle h_k, \Gamma^{-1}u\rangle_H) \\ &= \sum_{k=0}^{\infty} \Gamma h_k(t) X^u(h_k), \qquad t \in [0,T], \end{aligned}$$

where $(h_k)_{k\in\mathbb{N}}$ is orthonormal basis of H, and in case (B),

$$\begin{aligned} X_t - u(t) &= \sum_{k=0}^{\infty} h_k(t)(X(h_k) - \langle h_k, u\rangle_{L^2([0,T],d\mu)}) \\ &= \sum_{k=0}^{\infty} h_k(t)(X(h_k) - \langle h_k, \Gamma^{-1}u\rangle_H) \\ &= \sum_{k=0}^{\infty} h_k(t) X^u(h_k), \qquad t \in [0,T], \end{aligned}$$

where $(h_k)_{k\in\mathbb{N}}$ is orthonormal basis of $L^2([0,T], d\mu)$.

1.2.2 Efficient drift estimator

Here we work in the framework of (A), in the particular case where $(X_t)_{t\in[0,T]}$ has independent increments, i.e.

$$\gamma(s,t) = \int_0^{s\wedge t} \sigma_u^2 du,$$

where $\sigma \in L^2([0,T], dt)$ is an a.e. non-vanishing function,

$$(\dot{K}h)(t) = (\dot{K}^* h)(t) = \sigma_t \dot{h}(t), \qquad t \in [0,T],$$

with $K(t,r) = 1_{[0,t]}(r)\sigma_r$ and

$$\Gamma h(t) = \int_0^t \dot{h}_s \sigma_s^2 ds, \qquad t \in [0,T].$$

In other terms, $(X_t)_{t\in[0,T]}$ is a continuous Gaussian martingale with quadratic variation $\sigma_t^2 dt$, which can be represented as the time change

$$X_t = W_{\int_0^t \sigma_s^2 ds}, \qquad t \in [0,T],$$

of the standard Brownian motion $(W_t)_{t\in\mathbb{R}_+}$, or as the stochastic integral process $X_t = \displaystyle\int_0^t \sigma_s dW_s$, $t \in [0,T]$, and we have $X(h) = \displaystyle\int_0^T \dot{h}(s) dX_s$, $h \in H$, where

$$H = \left\{ v : [0,T] \to \mathbb{R} \ : \ v(t) = \int_0^t \dot{v}(s) ds, \ t \in [0,T], \ \dot{v} \in L^2([0,T], \sigma_t^2 dt) \right\}$$

is the Cameron-Martin space with inner product

$$\langle v_1, v_2 \rangle_H = \int_0^T \dot{v}_1(s) \dot{v}_2(s) \sigma_s^2 ds, \qquad v_1, v_2 \in H.$$

Let $(\mathcal{F}_t)_{t\in[0,T]}$ denote the filtration generated by $(X_t)_{t\in[0,T]}$, and for u an $\mathcal{F}_t$-adapted process, let $\mathbb{P}_u^\sigma$ denote the translation of the Wiener measure on Ω by u, i.e. $\mathbb{P}_u^\sigma$ is the measure on Ω under which

$$X_t^u := X_t - u_t, \qquad t \in [0,T],$$

is a continuous Gaussian martingale with quadratic variation

$$d\langle X^u, X^u \rangle_t = \sigma_t^2 dt.$$

Consider u an $\mathcal{F}_t$-adapted processes of the form

$$u_t = \int_0^t \dot{u}_s ds, \qquad t \in [0,T],$$

with

$$\mathbb{E}^{\sigma}\left[\int_0^T \frac{\dot{u}_s^2}{\sigma_s^2} ds\right] < \infty.$$

By the Girsanov theorem, $\mathbb{P}_u^{\sigma}$ is absolutely continuous with respect to $\mathbb{P}^{\sigma}$, with

$$d\mathbb{P}_u^{\sigma} = \Lambda(u) d\mathbb{P}^{\sigma},$$

where

$$\Lambda(u) := \exp\left(\int_0^T \frac{\dot{u}_s}{\sigma_s^2} dX_s - \frac{1}{2}\int_0^T \frac{\dot{u}_s^2}{\sigma_s^2} ds\right)$$

denotes the Girsanov-Cameron-Martin density, the canonical process $(X_t)_{t\in[0,T]}$ becomes a continuous Gaussian semimartingale under $\mathbb{P}_u^{\sigma}$, with quadratic variation $\sigma_t^2 dt$ and drift $\dot{u}_t dt$. The expectation under $\mathbb{P}_u$ will be denoted by $\mathbb{E}_u$.

Definition 1.2.1. *A drift estimator ξ is called unbiased if*

$$\mathbb{E}_u[\xi_t] = \mathbb{E}_u[u_t], \quad t \in [0,T],$$

for all square-integrable $\mathcal{F}_t$-adapted process $(u_t)_{t\in[0,T]}$. It is called adapted if the process $(\xi_t)_{t\in[0,T]}$ is $\mathcal{F}_t$-adapted.

Here, the canonical process $(X_t)_{t\in[0,T]}$ will be considered as an unbiased estimator of its own drift $(u_t)_{t\in[0,T]}$ under $\mathbb{P}_u^{\sigma}$, with risk defined as

$$\mathbb{E}_u^{\sigma}\left[\|X-u\|^2_{L^2([0,T],d\mu)}\right] = \int_0^T \mathbb{E}_u^{\sigma}\left[|X_t^u|^2\right]\mu(dt) = \int_0^T \int_0^t \sigma_s^2 ds \mu(dt),$$

where μ is a finite Borel measure on $[0,T]$. Clearly this estimator is consistent as σ or T tend to 0: precisely, given N independent samples

$$(X_t^1)_{t\in[0,T]}, \ldots, (X_t^N)_{t\in[0,T]},$$

of $(X_t)_{t\in[0,T]}$, the process

$$\bar{X}_t := \frac{X_t^1 + \cdots + X_t^N}{N}, \qquad t \in [0,T], \tag{1.2.3}$$

is an unbiased estimator of $(u_t)_{t\in[0,T]}$ whose risk

$$\mathbb{E}_u^{\sigma/\sqrt{N}}\left[\|\bar{X}-u\|^2_{L^2([0,T],d\mu)}\right] = \frac{1}{N}\int_0^T \int_0^t \sigma_s^2 ds \mu(dt)$$

converges to zero as N goes to infinity.

The justification of the use of $\hat{u} = (X_t)_{t\in[0,T]}$ as an efficient estimator comes from the following proposition which allows us to compute a Cramer-Rao bound attained by $\hat{u}$. Here the parameter space is restricted to the space of adapted processes in $L^2(\Omega \times [0,T], \mathbb{P} \otimes \mu)$, which corresponds in a sense to a parametric estimation. First we need a technical lemma.

Lemma 1.2.2. *Assume that* $(\xi_t)_{t\in[0,T]}$ *is a square integrable estimator of* u *then for every deterministic* ζ *in* H *we have that*

$$\frac{d}{d\varepsilon}\,\mathbb{E}^{\sigma}_{u+\varepsilon\zeta}[\xi_t-u_t]_{|\varepsilon=0}=\mathbb{E}^{\sigma}\left[(\xi_t-u_t)\frac{d}{d\varepsilon}\Lambda(u+\varepsilon\zeta)_{|\varepsilon=0}\right].$$

Proof. For notation simplicity we assume that $\sigma^2=1$. Let $t\in[0,1]$ and $-1<\varepsilon<1$.

$$\begin{aligned}
&\frac{1}{\varepsilon}(\mathbb{E}_{u+\varepsilon\zeta}[\xi_t-u_t]-\mathbb{E}_u[\xi_t-u_t])\\
=\;&\frac{1}{\varepsilon}\,\mathbb{E}[(\xi_t-u_t)(\Lambda(u+\varepsilon\zeta)-\Lambda(u))]\\
=\;&\frac{1}{\varepsilon}\,\mathbb{E}\left[(\xi_t-u_t)\Lambda(u)\left(\exp\left(\varepsilon\int_0^T\dot\zeta_s dX_s-\varepsilon\int_0^T\dot u_s\dot\zeta_s ds-\varepsilon^2\frac{1}{2}\int_0^T\dot\zeta_s^2 ds\right)-1\right)\right]\\
=\;&\frac{1}{\varepsilon}\,\mathbb{E}_u\left[(\xi_t-u_t)f(\varepsilon)\right],
\end{aligned}$$

where

$$f(\varepsilon):=\exp\left(\varepsilon\int_0^T\dot\zeta_s dX_s-\varepsilon\int_0^T\dot u_s\dot\zeta_s ds-\varepsilon^2\frac{1}{2}\int_0^T\dot\zeta_s^2 ds\right)-1.$$

Now since we are interested by the behavior of the quantity above when ε is small we can use a pathwise Taylor expansion and we obtain that

$$f(\varepsilon)=\varepsilon\left(\int_0^T\dot\zeta_s dX_s-\int_0^T\dot u_s\dot\zeta_s ds\right)+\frac{1}{2}\varepsilon^2 f''(\varepsilon\theta),\quad \theta\in[0,1].$$

So for ε small enough we have that,

$$\begin{aligned}
&\frac{1}{\varepsilon}(\mathbb{E}_{u+\varepsilon\zeta}[\xi_t-u_t]-\mathbb{E}_u[\xi_t-u_t])\\
=\;&\mathbb{E}_u\left[(\xi_t-u_t)\left(\int_0^T\dot\zeta_s dX_s-\int_0^T\dot u_s\dot\zeta_s ds\right)\right]+\frac{\varepsilon^2}{2}\,\mathbb{E}_u\left[(\xi_t-u_t)f''(\varepsilon\theta)\right].
\end{aligned}$$

We show that

$$\lim_{\varepsilon\to0}\left|\mathbb{E}_u\left[(\xi_t-u_t)f''(\varepsilon\theta)\right]\right|<\infty.$$

Let t in $[0,1]$.

$$\begin{aligned}
\left|f''(t)\right| &= (f(t)+1)\left|\left(\int_0^T\dot\zeta_s dX_s^u-t\int_0^T\dot\zeta_s^2 ds\right)^2-\int_0^T\dot\zeta_s^2 ds\right|\\
&\le (f(t)+1)\left(\int_0^T\dot\zeta_s dX_s^u-t\int_0^T\dot\zeta_s^2 ds\right)^2+\int_0^T\dot\zeta_s^2 ds\\
&\le (f(t)+1)\left[2\left(\int_0^T\dot\zeta_s dX_s^u\right)^2+t^2\left(\int_0^1\dot\zeta_s^2 ds\right)^2+\int_0^T\dot\zeta_s^2 ds\right]\le (f(t)+1)g,
\end{aligned}$$

where $g := 2\left(\int_0^T \dot\zeta_s dX_s^u\right)^2 + 2\left(\int_0^T \dot\zeta_s^2 ds\right)^2 + \int_0^T \dot\zeta_s^2 ds$. In addition we have that,

$$\begin{aligned} f(t)+1 &= \exp\left(t\int_0^T \dot\zeta_s dX_s^u - \frac{t^2}{2}\int_0^T \dot\zeta_s^2 ds\right) \\ &\leq \exp\left(t\int_0^T \dot\zeta_s dX_s^u\right). \end{aligned}$$

Combining the two previous facts we obtain that

$$\begin{aligned} \left|\mathbb{E}_u\left[(\xi_t - u_t)f''(\varepsilon\theta)\right]\right| &\leq \mathbb{E}_u\left[|\xi_t - u_t|\ g\ \exp\left(\varepsilon\theta\int_0^1 \dot\zeta_s dX_s^u\right)\right] \\ &\leq \mathbb{E}_u\left[|\xi_t - u_t|^2\right]^{1/2} E_u\left[g^2 \exp\left(2\varepsilon\theta\int_0^T \dot\zeta_s dX_s^u\right)\right]^{1/2}. \end{aligned}$$

Assume that the family $(g^2 \exp\left(2\varepsilon\theta\int_0^1 \dot\zeta_s dX_s^u\right))_{-1<\varepsilon<1}$ is uniformly integrable under $\mathbb{P}_u$ then we have that

$$\lim_{\varepsilon\to 0}\left|E_u\left[\exp\left(2\varepsilon\theta\int_0^1 \dot\zeta_s dX_s^u\right)\right]\right| < \infty$$

which shows that

$$\lim_{\varepsilon\to 0}\frac{\varepsilon}{2}\left|\mathbb{E}_u\left[(\xi_t - u_t)f''(\varepsilon\theta)\right]\right| = 0.$$

So it remains to show that the family $(g^2 \exp\left(2\varepsilon\theta\int_0^T \dot\zeta_s dX_s^u\right))_{-1<\varepsilon<1}$ is uniformly integrable under $\mathbb{P}_u$. By [66, Theorem 27.2] it is enough to show that there exists $M > 0$ (independent of ε) such that for every $-1 < \varepsilon < 1$ we have that

$$\mathbb{E}_u\left[g^4 \exp\left(2\varepsilon\theta\int_0^T \dot\zeta_s dX_s^u\right)^2\right] \leq M.$$

Let ε in $(0,1)$. By Cauchy-Schwarz inequality we have that

$$\begin{aligned} \mathbb{E}_u\left[g^4 \exp\left(2\varepsilon\theta\int_0^T \dot\zeta_s dX_s^u\right)^2\right] &\leq \mathbb{E}_u\left[g^8\right]^{1/2}\mathbb{E}_u\left[\exp\left(4\varepsilon\theta\int_0^T \dot\zeta_s dX_s^u\right)\right]^{1/2} \\ &= \mathbb{E}_u\left[g^8\right]^{1/2}\mathbb{E}_u\left[\exp(Z)\right]^{1/2},\ Z \sim \mathcal{N}(0, 16(\varepsilon\theta)^2\|\dot\zeta\|^2_{L^2([0,1])}) \\ &= \mathbb{E}_u\left[g^8\right]^{1/2} e^{4(\varepsilon\theta)^2\|\dot\zeta\|^2_{L^2([0,1])}} \leq \mathbb{E}_u\left[g^8\right]^{1/2} e^{4\|\dot\zeta\|^4_{L^2([0,1])}} =: M. \end{aligned}$$

Since $\int_0^1 \dot\zeta_s dX_s^u \sim \mathcal{N}(0, \|\dot\zeta\|^2_{L^2([0,1])})$ under $\mathbb{P}_u$ we have that $\mathbb{E}_u[g^8] < \infty$. □

Proposition 1.2.3. *Cramer-Rao inequality. For any unbiased and adapted estimator ξ of u we have*

$$\mathbb{E}_u^\sigma\left[\int_0^T |\xi_t - u_t|^2 \mu(dt)\right] \geq \mathrm{CRB}(\sigma, \mu, \hat u), \tag{1.2.4}$$

where $u \in L^2(\Omega \times [0,T], \mathbb{P}_u^\sigma \otimes \mu)$ *is adapted and the Cramer-Rao type bound*

$$\mathrm{CRB}(\sigma, \mu, \hat{u}) := \int_0^T \int_0^t \sigma_s^2 ds \mu(dt)$$

is independent of u *and attained by the efficient estimator* $\hat{u} = X$.

Proof. Since ξ is unbiased, for all $\zeta \in H$ we have

$$\begin{aligned} \mathbb{E}_{u+\varepsilon\zeta}^\sigma[\xi_t] &= \mathbb{E}_{u+\varepsilon\zeta}^\sigma[u_t + \varepsilon\zeta_t] \\ &= \mathbb{E}_{u+\varepsilon\zeta}^\sigma[u_t] + \varepsilon\, \mathbb{E}_{u+\varepsilon\zeta}^\sigma[\zeta_t] \\ &= \mathbb{E}_{u+\varepsilon\zeta}^\sigma[u_t] + \varepsilon\zeta_t, \quad t \in [0,T], \quad \varepsilon \in \mathbb{R}, \end{aligned}$$

hence

$$\begin{aligned} \zeta_t &= \frac{d}{d\varepsilon}\, \mathbb{E}_{u+\varepsilon\zeta}^\sigma[\xi_t - u_t]_{|\varepsilon=0} \\ &= \frac{d}{d\varepsilon}\, \mathbb{E}^\sigma[(\xi_t - u_t)\Lambda(u + \varepsilon\zeta)]_{|\varepsilon=0} \\ &= \mathbb{E}^\sigma\left[(\xi_t - u_t)\frac{d}{d\varepsilon}\Lambda(u+\varepsilon\zeta)_{|\varepsilon=0}\right] \\ &= \mathbb{E}_u^\sigma\left[(\xi_t - u_t)\frac{d}{d\varepsilon}\log\Lambda(u+\varepsilon\zeta)_{|\varepsilon=0}\right] \\ &= \mathbb{E}_u^\sigma\left[(\xi_t - u_t)\left(\int_0^T \frac{\dot{\zeta}_s}{\sigma_s^2}dX_s - \int_0^T \frac{\dot{\zeta}_s\dot{u}_s}{\sigma_s^2}ds\right)\right] \\ &= \mathbb{E}_u^\sigma\left[(\xi_t - u_t)\int_0^T \frac{\dot{\zeta}_s}{\sigma_s^2}dX_s^u\right] \\ &= \mathbb{E}_u^\sigma\left[(\xi_t - u_t)\int_0^t \frac{\dot{\zeta}_s}{\sigma_s^2}dX_s^u\right], \end{aligned}$$

where the exchange between expectation and derivative is justified by Lemma 1.2.2. Thus, by the Cauchy-Schwarz inequality and the Itô isometry we have

$$\zeta_t^2 \leq \mathbb{E}_u^\sigma\left[\left(\int_0^t \frac{\dot{\zeta}_s}{\sigma_s^2}dX_s^u\right)^2\right]\mathbb{E}_u^\sigma[|\xi_t - u_t|^2] = \int_0^t \frac{\dot{\zeta}_s^2}{\sigma_s^2}ds\, \mathbb{E}_u^\sigma[|\xi_t - u_t|^2], \qquad t \in [0,T].$$

It then suffices to take

$$\zeta_t = \int_0^t \sigma_s^2 ds, \qquad t \in [0,T],$$

to get

$$\mathrm{Var}_u^\sigma[\xi_t] = \mathbb{E}_u^\sigma[|\xi_t - u_t|^2] \geq \int_0^t \sigma_s^2 ds, \qquad t \in [0,T], \tag{1.2.5}$$

which leads to (1.2.4) after integration with respect to $\mu(dt)$. As noted above, $\hat{u} = (X_t)_{t\in[0,T]}$ is clearly unbiased under $\mathbb{P}_u^\sigma$ and it attains the lower bound $\mathrm{CRB}(\sigma, \mu, \hat{u})$. □

Recall that the classical linear parametric estimation problem for the drift of a diffusion consists in estimating the coefficient θ appearing in

$$d\xi_t = \theta a_t(\xi_t)dt + dY_t, \qquad \xi_0 = 0,$$

with a maximum likelihood estimator $\hat{\theta}_T$ given by

$$\hat{\theta}_T = \frac{\int_0^T a_t(\xi_t)d\xi_t}{\int_0^T a_t^2(\xi_t)dt}, \tag{1.2.6}$$

cf. [75], [110] for Brownian motion and [131] for an extension to fractional Brownian motions.

Here we consider the nonparametric functional estimation of the drift of a one-dimensional drifted Brownian motion $(X_t)_{t\in\mathbb{R}_+}$ with decomposition

$$dX_t = \dot{u}_t dt + dX_t^u, \tag{1.2.7}$$

where $(\dot{u}_t)_{t\in[0,T]} \in L^2(\Omega \times [0,T])$ is an adapted process and $(X_t^u)_{t\in\mathbb{R}_+}$ is a standard Brownian motion with quadratic variation σ_t^2 under a probability $\mathbb{P}_u^\sigma$.

In case u is constrained to have the form $u_t = \theta t$, $t \in [0,T]$, $\theta \in \mathbb{R}$, our efficient estimator $\hat{u}$ satisfies $\hat{u}_T = \hat{\theta}_T T$, where $\hat{\theta}_T$ is given by (1.2.6), $T > 0$, with the asymptotics $\hat{\theta}_T \to \theta$ in probability as T tends to infinity. The asymptotics is not in large time since T can be a fixed parameter, but the efficient estimator $\hat{u} = X$ converges to u as σ tends to 0, or equivalently as T tends to 0 by rescaling.

To close this section we note that, at least informally, $\hat{u} = (X_t)_{t\in[0,T]}$ can be viewed as a maximum likelihood estimator of its own adapted drift $(u_t)_{t\in[0,T]}$ under $\mathbb{P}_u^\sigma$. Indeed the functional differentiation of the Cameron-Martin density

$$\frac{d}{d\varepsilon}\Lambda(\hat{u} + \varepsilon\zeta)_{|\varepsilon=0} = 0, \qquad \zeta \in H,$$

implies

$$\int_0^T \frac{\dot{\zeta}_s}{\sigma_s^2} dX_s - \int_0^T \frac{\dot{\zeta}_s}{\sigma_s^2} d\hat{u}_s = 0, \qquad \zeta \in H,$$

which leads to $X = \hat{u}$.

1.2.3 Bayes estimators

In this section we consider Bayes estimators which will be useful in proving the minimaxity of the estimator $\hat{u} = X$ in the framework of (A) for Gaussian processes with non-necessarily independent increments. We will make use of the next lemma which is classical in the framework of Gaussian filtering and is proved in the Appendix.

Lemma 1.2.4. *Let Z be a Gaussian process with covariance operator Γ_τ and drift $v \in H$, and assume that X is a Gaussian process with drift Z and covariance operator Γ given Z. Then, conditionally to X, Z has drift*

$$f \mapsto \langle f, (\Gamma + \Gamma_\tau)^{-1} \Gamma v \rangle + X((\Gamma + \Gamma_\tau)^{-1} \Gamma_\tau f) \quad \textit{and covariance} \quad \Gamma_\tau (\Gamma + \Gamma_\tau)^{-1} \Gamma.$$

Note that unlike in Proposition 1.2.3, no adaptedness or unbiasedness restriction is made on ξ in the infimum taken in (1.2.9) below.

Proposition 1.2.5. *Bayes estimator. Let P_v^τ denote the Gaussian distribution on Ω with covariance operator Γ_τ and drift $v \in H$. The Bayes risk*

$$\int_\Omega \mathbb{E}_z \left[\int_0^T |\xi_t - z_t|^2 \mu(dt) \right] d\mathbb{P}_v^\tau(z) \tag{1.2.8}$$

of any estimator $(\xi_t)_{t\in[0,T]}$ on Ω under the prior distribution $\mathbb{P}_v^\tau$ is uniquely minimized by

$$\xi_t^{\tau,v} := \langle \chi_t, (\Gamma_\tau + \Gamma)^{-1} \Gamma v \rangle + X((\Gamma_\tau + \Gamma)^{-1} \Gamma_\tau \, \chi_t), \qquad t \in [0, T],$$

which has risk

$$\int_0^T \langle \chi_t, \Gamma (\Gamma_\tau + \Gamma)^{-1} \Gamma_\tau \, \chi_t \rangle \mu(dt) = \inf_\xi \int_\Omega \mathbb{E}_z \left[\int_0^T |\xi_t - z_t|^2 \mu(dt) \right] d\mathbb{P}_v^\tau(z). \tag{1.2.9}$$

Proof. Let Z denote a Gaussian process with drift $v \in H$ and covariance Γ_τ. Recall (cf. Lemma 1.2.4) that if X has drift Z and covariance Γ then, conditionally to X, $(Z_t)_{t\in[0,T]}$ has drift

$$t \mapsto \langle \chi_t, (\Gamma_\tau + \Gamma)^{-1} \Gamma v \rangle + X((\Gamma_\tau + \Gamma)^{-1} \Gamma_\tau \, \chi_t)$$

and covariance $\Gamma_\tau (\Gamma_\tau + \Gamma)^{-1} \Gamma$. Hence the Bayes risk of an estimator ξ under the prior distribution $\mathbb{P}_v^\tau$ is given by

$$\begin{aligned}
&\int_\Omega \mathbb{E}_z \left[\int_0^T |\xi_t - z_t|^2 \mu(dt) \right] d\mathbb{P}_v^\tau(z) = \mathbb{E} \left[\mathbb{E} \left[\int_0^T |\xi_t - Z_t|^2 \mu(dt) \Big| X \right] \right] \\
&= \mathbb{E} \left[\int_0^T |\xi_t - \mathbb{E}[Z_t \mid X]|^2 \mu(dt) \right] + \mathbb{E} \left[\int_0^T \mathrm{Var}(Z_t | X) \mu(dt) \right] \\
&= \mathbb{E} \left[\int_0^T \left| \xi_t - \langle \chi_t, (\Gamma_\tau + \Gamma)^{-1} \Gamma v \rangle - X((\Gamma_\tau + \Gamma)^{-1} \Gamma_\tau \, \chi_t) \right|^2 \mu(dt) \right] \\
&\quad + \int_0^T \langle \chi_t, \Gamma (\Gamma_\tau + \Gamma)^{-1} \Gamma_\tau \, \chi_t \rangle \mu(dt),
\end{aligned}$$

which is minimized by

$$\xi_t^{\tau,v} := \mathbb{E}[Z_t \mid X] = \langle \chi_t, (\Gamma_\tau + \Gamma)^{-1} \Gamma v \rangle - X((\Gamma_\tau + \Gamma)^{-1} \Gamma_\tau \, \chi_t), \qquad t \in [0, T].$$

□

Clearly $\xi^{\tau,v}$ is unique in the sense that it is the only estimator to minimize the Bayes risk (1.2.8). This shows in particular that every $\xi^{\tau,v}$ is admissible in the sense that if an estimator ξ satisfies

$$\mathbb{E}_z\left[\|\xi - z\|^2_{L^2([0,T],d\mu)}\right] \leq \mathbb{E}_z\left[\|\xi^{\tau,v} - z\|^2_{L^2([0,T],d\mu)}\right], \qquad z \in \Omega,$$

then

$$\begin{aligned}\int_\Omega \mathbb{E}_z\left[\|\xi - z\|^2_{L^2([0,T],d\mu)}\right] d\mathbb{P}^\tau_v &\leq \int_\Omega \mathbb{E}_z\left[\|\xi^{\tau,v} - z\|^2_{L^2([0,T],d\mu)}\right] d\mathbb{P}^\tau_v \\ &= \int_0^T \langle \chi_t, \Gamma(\Gamma_\tau + \Gamma)^{-1}\Gamma_\tau\, \chi_t\rangle \mu(dt),\end{aligned}$$

hence

$$\int_\Omega \mathbb{E}_z\left[\|\xi - z\|^2_{L^2([0,T],d\mu)}\right] d\mathbb{P}^\tau_v = \int_0^T \langle \chi_t, \Gamma(\Gamma_\tau + \Gamma)^{-1}\Gamma_\tau\, \chi_t\rangle \mu(dt), \tag{1.2.10}$$

and $\xi = \xi^{\tau,v}$ by Proposition 1.2.5.

The Bayes estimator $\xi^{\tau,v}$ is biased in general, and for deterministic $u \in H$ its mean square error under $\mathbb{P}_u$ is equal to

$$\begin{aligned}&\mathbb{E}_u\left[\int_0^T |\xi^{\tau,v}_t - u_t|^2 \mu(dt)\right] \qquad\qquad (1.2.11)\\ &= \mathbb{E}_u\left[\int_0^T \left|\xi^{\tau,v}_t - \mathbb{E}_u[\xi^{\tau,v}_t]\right|^2 \mu(dt)\right] + \mathbb{E}_u\left[\int_0^T |\,\mathbb{E}_u[\xi^{\tau,v}_t] - u_t|^2 \mu(dt)\right] \\ &= \mathbb{E}_u\left[\int_0^T \left|X^u((\Gamma_\tau + \Gamma)^{-1}\Gamma_\tau\, \chi_t)\right|^2 \mu(dt)\right] + \int_0^T \left|\langle \chi_t, (\Gamma_\tau + \Gamma)^{-1}\Gamma(v-u)\rangle\right|^2 \mu(dt) \\ &= \int_0^T \langle (\Gamma_\tau + \Gamma)^{-1}\Gamma_\tau\, \chi_t, \Gamma(\Gamma_\tau + \Gamma)^{-1}\Gamma_\tau\, \chi_t\rangle \mu(dt) \\ &\quad + \int_0^T \left|\langle \chi_t, (\Gamma_\tau + \Gamma)^{-1}\Gamma(v-u)\rangle\right|^2 \mu(dt),\end{aligned}$$

which shows that

$$\sup_{u\in H} \mathbb{E}_u\left[\int_0^T |\xi^{\tau,v}_t - u_t|^2 \mu(dt)\right] = +\infty,$$

hence $\xi^{\tau,v}$ is not minimax.

In the independent increment case of Section 1.2.2 we have, if $\Gamma_\tau f(s) = \tau_s^2 f(s)$, $s \in [0,T]$:

$$\xi^{\tau,v}_t := \int_0^t \frac{\sigma_s^2 \dot{v}_s}{\tau_s^2 + \sigma_s^2} ds + \int_0^t \frac{\tau_s^2}{\tau_s^2 + \sigma_s^2} dX_s, \qquad t \in [0,T],$$

with risk

$$\int_0^T \int_0^t \frac{\tau_s^2 \sigma_s^2}{\tau_s^2 + \sigma_s^2} ds \mu(dt) = \inf_\xi \int_\Omega \mathbb{E}_z \left[\int_0^T |\xi_t - z_t|^2 \mu(dt) \right] d\mathbb{P}_v^\tau(z). \tag{1.2.12}$$

Assuming now that $\Gamma_\tau f(t) = \tau^2 f(t)$, $t \in [0, T]$, the Bayes risk

$$\int_0^T \langle \chi_t, \Gamma_\tau (\Gamma_\tau + \Gamma)^{-1} \Gamma \, \chi_t \rangle \mu(dt) = \int_0^T \langle \chi_t, (I + \Gamma/\tau^2)^{-1} \Gamma \, \chi_t \rangle \mu(dt),$$

of $\xi^{\tau,v}$, $\tau \in \mathbb{R}$, converges as $\tau \to \infty$ to the bound

$$\mathrm{CRB}(\sigma, \mu, \hat{u}) = \int_0^T \langle \chi_t, \Gamma \, \chi_t \rangle \mu(dt),$$

hence it follows in the next proposition that, as in the finite dimensional Gaussian case, the estimator $\hat{u} = (X_t)_{t \in [0,T]}$ is minimax. Note again that unlike in Proposition 1.2.3, no adaptedness condition is imposed on ξ in the infima (1.2.9) and (1.2.13).

Proposition 1.2.6. *The estimator $\hat{u} = X$ is minimax. For all $u \in \Omega$ we have*

$$\mathrm{CRB}(\gamma, \mu, \hat{u}) = \mathbb{E}_u \left[\int_0^T |X_t - u_t|^2 \mu(dt) \right] = \inf_\xi \sup_{v \in \Omega} \mathbb{E}_v \left[\int_0^T |\xi_t - v_t|^2 \mu(dt) \right]. \tag{1.2.13}$$

Proof. Clearly, taking $\xi = 0$ yields

$$\mathrm{CRB}(\gamma, \mu, \hat{u}) = \sup_{u \in \Omega} \mathbb{E}_u \left[\int_0^T |X_t - u_t|^2 \mu(dt) \right] \geq \inf_\xi \sup_{u \in \Omega} \mathbb{E}_u \left[\int_0^T |\xi_t - u_t|^2 \mu(dt) \right].$$

On the other hand, from Proposition 1.2.5, for all processes ξ we have

$$\begin{aligned} \sup_{u \in \Omega} \mathbb{E}_u \left[\int_0^T |\xi_t - u_t|^2 \mu(dt) \right] &\geq \int_\Omega \mathbb{E}_z \left[\int_0^T |\xi_t - z_t|^2 \mu(dt) \right] d\mathbb{P}_0^\tau(z) \\ &\geq \int_0^T \langle \chi_t, (I + \Gamma/\tau^2)^{-1} \Gamma \, \chi_t \rangle \mu(dt), \end{aligned}$$

for all $\tau > 0$, hence

$$\inf_\xi \sup_{u \in H} \mathbb{E}_u \left[\int_0^T |\xi_t - u_t|^2 \mu(dt) \right] \geq \int_0^T \langle \chi_t, \Gamma \, \chi_t \rangle \mu(dt) = \mathrm{CRB}(\gamma, \mu, \hat{u}).$$

□

1.2.4 Malliavin calculus on Gaussian space

Before proceeding to the construction of Stein type estimators, we need to introduce some elements of analysis on Gaussian space, see e.g. [94]. This construction is valid in both frameworks (A) and (B). Given $u \in H$, let

$$X^u = X - u.$$

We fix $(h_n)_{n\geq 1}$ a total subset of H and let $\mathcal{S}$ denote the space of cylindrical functionals of the form

$$F = f_n\left(X^u(h_1), \ldots, X^u(h_n)\right), \tag{1.2.14}$$

where f_n is in the space of infinitely differentiable rapidly decreasing functions on $\mathbb{R}^n$, $n \geq 1$.

Definition 1.2.7. *The H-valued Malliavin derivative is defined as*

$$\nabla_t F = \sum_{i=1}^{n} h_i(t)\partial_i f_n\left(X^u(h_1), \ldots, X^u(h_n)\right),$$

for $F \in \mathcal{S}$ of the form (1.2.14).

It is known that ∇ is closable, cf. Proposition 1.2.1 of [94], and its closed domain will be denoted by $\mathrm{Dom}(\nabla)$.

Definition 1.2.8. *Let D be defined on* $\mathrm{Dom}(\nabla)$ *as*

$$D_t F := (\Gamma \nabla F)(t), \quad t \in [0, T], \quad F \in \mathrm{Dom}(\nabla).$$

Let $\delta : L^2_u(\Omega; H) \to L^2(\Omega, \mathbb{P}_u)$ denote the closable adjoint of ∇, i.e. the divergence operator under $\mathbb{P}_u$, which satisfies the integration by parts formula

$$\mathbb{E}_u[F\delta(v)] = \mathbb{E}_u[\langle v, \nabla F\rangle_H], \qquad F \in \mathrm{Dom}(\nabla), \quad v \in \mathrm{Dom}(\delta), \tag{1.2.15}$$

with the relation

$$\delta(hF) = FX(h) - \langle h, \nabla F\rangle_H,$$

cf. [94], for $F \in \mathrm{Dom}(\nabla)$ and $h \in H$ such that $hF \in \mathrm{Dom}(\delta)$. Note that (1.2.15) is an infinite-dimensional version of the integration by parts (1.2.1), which can be proved e.g. using the countable Gaussian random variables constructed from X.

Lemma 1.2.9. *We have*

$$\mathbb{E}_u[FX^u_t] = \mathbb{E}_u[D_t F], \qquad t \in [0, T], \quad F \in \mathrm{Dom}(\nabla).$$

Proof.

(A) In the case of Paley-Wiener expansions we have

$$\begin{aligned} \mathbb{E}_u[FX_t^u] &= \mathbb{E}_u[FX^u(\chi_t)] \\ &= \mathbb{E}_u[F\delta(\chi_t)] \\ &= \mathbb{E}_u[\langle \chi_t, \nabla F\rangle_H] \\ &= \mathbb{E}_u[\langle 1_{[0,t]}, \dot{\Gamma}\nabla F\rangle_{L^2([0,T],dt)}] \\ &= \mathbb{E}_u[(\Gamma\nabla F)(t)], \qquad F \in \mathrm{Dom}(\nabla), \quad t \in [0,T]. \end{aligned}$$

(B) In the case of Karhunen-Loève expansions we have

$$\begin{aligned} \mathbb{E}_u[FX_t^u] &= \sum_{k=0}^{\infty} h_k(t)\, \mathbb{E}_u[FX^u(h_k)] \\ &= \sum_{k=0}^{\infty} h_k(t)\, \mathbb{E}_u[F\delta(h_k)] \\ &= \sum_{k=0}^{\infty} h_k(t)\, \mathbb{E}_u[\langle h_k, \nabla F\rangle_H] \\ &= \sum_{k=0}^{\infty} h_k(t)\, \mathbb{E}_u[\langle h_k, \Gamma\nabla F\rangle_{L^2([0,T],\mu)}] \\ &= \mathbb{E}_u[(\Gamma\nabla F)(t)], \qquad F \in \mathrm{Dom}(\nabla), \quad t \in [0,T]. \end{aligned}$$

□

Definition 1.2.10. *We define the Laplacian* Δ *by*

$$\Delta F = \mathrm{trace}_{L^2([0,T],d\mu)^{\otimes 2}} DDF = \int_0^T D_t D_t F \mu(dt)$$

on the space $\mathrm{Dom}(\Delta)$ *made of all* $F \in \mathrm{Dom}(\nabla)$ *such that* $D_t F \in \mathrm{Dom}(\nabla)$*,* $t \in [0,T]$*, and* $(D_t D_t F)_{t\in[0,T]} \in L^2([0,T],\mu)$*,* $\mathbb{P}$*-a.s.*

If $F \in \mathcal{S}$ has the form (1.2.14) we have

$$\Delta F = \sum_{i,j=1}^{n} \langle \Gamma h_i, \Gamma h_j\rangle_{L^2([0,T],\mu)} \partial_i \partial_j f_n \left(X^u(h_1), \ldots, X^u(h_n)\right).$$

Unlike the Gross Laplacian Δ_G defined by

$$\Delta_G F = \mathrm{trace}_{H^{\otimes 2}} \nabla\nabla F,$$

the operator Δ is closable, as shown in the following proposition.

Proposition 1.2.11. *Closability of Δ. For any sequence $(F_n)_{n\in\mathbb{N}}$ of random variables converging to* 0 *in $L^2(\Omega,\mathbb{P}_u)$ and such that $(\Delta F_n)_{n\in\mathbb{N}}$ converges in $L^2(\Omega,\mathbb{P}_u)$, we have*

$$\lim_{n\to\infty} \Delta F_n = 0.$$

Proof. Let $(G_n)_{n\in\mathbb{N}}$ a sequence in $\mathcal{S}$ converging to 0 in $L^2(\Omega,\mathbb{P}_u)$, and such that $(\Delta G_n)_{n\in\mathbb{N}}$ converges to F in $L^2(\Omega,\mathbb{P}_u)$. For all $G\in\mathcal{S}$ we have, in the notation of (A):

$$\begin{aligned}
|\langle \Delta G_n, G\rangle_{L^2(\Omega,\mathbb{P}_u)}| &= \left|\mathbb{E}_u\left[G\int_0^T D_t D_t G_n \mu(dt)\right]\right| \\
&= \left|\int_0^T \mathbb{E}_u[\langle \nabla D_t G_n, \chi_t\, G\rangle_H]\,\mu(dt)\right| \\
&= \left|\int_0^T \mathbb{E}_u[D_t G_n\,\delta(\chi_t\, G)]\,\mu(dt)\right| \\
&= \left|\int_0^T \mathbb{E}_u[\langle \nabla G_n, \chi_t\, \delta(\chi_t\, G)\rangle_H]\,\mu(dt)\right| \\
&= \left|\int_0^T \mathbb{E}_u[G_n\delta(\chi_t\, \delta(\chi_t\, G))]\,\mu(dt)\right| \\
&\le \|G_n\|_{L^2(\Omega,\mathbb{P}_u)}\int_0^T \|\delta(\chi_t\, \delta(\chi_t\, G))\|_{L^2(\Omega,\mathbb{P}_u)}\,\mu(dt),
\end{aligned}$$

hence $\langle F, G\rangle_{L^2(\Omega,\mathbb{P}_u)} = 0,\ G\in\mathcal{S}$, which implies $F=0$. □

We will say that a random variable F in $\mathrm{Dom}(\Delta)$ is Δ-superharmonic on Ω if

$$\Delta F(\omega) \le 0, \qquad \mathbb{P}(d\omega) - a.s. \tag{1.2.16}$$

Remark 1.2.12. *In the independent increment case where $\gamma(s,t)$ is given by*

$$\gamma(s,t) = \int_0^{s\wedge t} \sigma_u^2 dt, \qquad s,t\in[0,T],$$

we have

$$\delta(v) = \int_0^T \dot{v}_t dX_t^u, \tag{1.2.17}$$

for every $\mathcal{F}_t$-adapted process $v\in L^2(\Omega; H,\mathbb{P}_u)$.

1.2.5 Superefficient drift estimators

Our aim is to construct a superefficient estimator of u of the form $X+\xi$, whose mean square error is strictly smaller than the minimax risk $\mathrm{CRB}(\gamma,\mu,\hat{u})$ of Proposition 1.2.6 when $\xi\in L^2([0,T]\times\Omega,\mathbb{P}_u\otimes\mu)$ is a suitably chosen stochastic process. This estimator will be biased and anticipating with respect to the Brownian filtration. In the next lemma we follow Stein's argument which uses integration by parts but we replace (1.2.1) by the duality relation (1.2.15) between the gradient and divergence operators on Gaussian space. The results of this section are valid in both frameworks (A) and (B).

Lemma 1.2.13. *Unbiased risk estimate. For any* $\xi \in L^2(\Omega \times [0,T], \mathbb{P}_u \otimes \mu)$ *such that* $\xi_t \in \mathrm{Dom}(\nabla)$, $t \in [0,T]$, *and* $(D_t\xi_t)_{t\in[0,T]} \in L^1(\Omega \times [0,T], \mathbb{P}_u \otimes \mu)$, *we have*

$$\mathbb{E}_u\left[\|X+\xi-u\|^2_{L^2([0,T],\mu)}\right] = \mathrm{CRB}(\gamma,\mu,\hat{u}) + \|\xi\|^2_{L^2(\Omega\times[0,T],\mathbb{P}_u\otimes\mu)} + 2\,\mathbb{E}_u\left[\int_0^T D_t\xi_t\mu(dt)\right]. \tag{1.2.18}$$

Proof. We have

$$\begin{aligned}
\mathbb{E}_u\left[\|X+\xi-u\|^2_{L^2([0,T],d\mu)}\right] &= \mathbb{E}_u\left[\int_0^T \left|X_t^u+\xi_t\right|^2\mu(dt)\right] \\
&= \mathbb{E}_u\left[\int_0^T |X_t^u|^2\mu(dt)\right] + \|\xi\|^2_{L^2(\Omega\times[0,T],\mathbb{P}_u\otimes\mu)} + 2\,\mathbb{E}_u\left[\int_0^T X_t^u\xi_t\mu(dt)\right] \\
&= \mathrm{CRB}(\gamma,\mu,\hat{u}) + \|\xi\|^2_{L^2(\Omega\times[0,T],\mathbb{P}_u\otimes\mu)} + 2\,\mathbb{E}_u\left[\int_0^T X_t^u\xi_t\mu(dt)\right],
\end{aligned}$$

and apply Lemma 1.2.9 to obtain (1.2.18). □

The next proposition specializes the above lemma to processes ξ of the form

$$\xi_t = D_t \log F, \qquad t \in [0,T],$$

where F is an a.s. strictly positive and sufficiently smooth random variable.

Proposition 1.2.14. *Logarithmic gradient. Stein-type estimator. For any* $\mathbb{P}$*-a.s. positive random variable* $F \in \mathrm{Dom}(\nabla)$ *such that* $D_tF \in \mathrm{Dom}(\nabla)$, $t \in [0,T]$, *and* $(D_tD_tF)_{t\in[0,T]} \in L^1(\Omega \times [0,T], \mathbb{P}_u \otimes \mu)$, *we have*

$$\mathbb{E}_u\left[\|X+D\log F-u\|^2_{L^2([0,T],d\mu)}\right] = \mathrm{CRB}(\gamma,\mu,\hat{u}) - \mathbb{E}_u\left[\|D\log F\|^2_{L^2([0,T],\mu)}\right] + 2\,\mathbb{E}_u\left[\frac{\Delta F}{F}\right].$$

Proof. From (1.2.18) we have

$$\begin{aligned}
&\mathbb{E}_u\left[\|X+D\log F-u\|^2_{L^2([0,T],\mu)}\right] \\
&\quad= \mathrm{CRB}(\gamma,\mu,\hat{u}) + \|D\log F\|^2_{L^2(\Omega\times[0,T],\mathbb{P}_u\otimes\mu)} + 2\,\mathbb{E}_u\left[\int_0^T D_tD_t\log F\mu(dt)\right] \\
&\quad= \mathrm{CRB}(\gamma,\mu,\hat{u}) + \mathbb{E}_u\left[\int_0^T \left(\left|\frac{D_tF}{F}\right|^2 + 2D_tD_t\log F\right)\mu(dt)\right],
\end{aligned}$$

and we use the relation

$$\left|\frac{D_tF}{F}\right|^2 + 2D_tD_t\log F = 2\frac{D_tD_tF}{F} - \left|\frac{D_tF}{F}\right|^2, \qquad t \in [0,T].$$

□

From the above proposition it suffices that F be Δ-superharmonic for $X + D\log F$ to be superefficient. In this case we have

$$\mathbb{E}_u\left[\|X + D\log F - u\|^2_{L^2([0,T],d\mu)}\right] \leq \mathrm{CRB}(\gamma,\mu,\hat{u}) - \mathbb{E}_u\left[\|D\log F\|^2_{L^2([0,T],d\mu)}\right], \quad (1.2.19)$$

with equality in (1.2.19) when F is Δ-harmonic.

In the next proposition we show that the Δ-superharmonicity of F is not necessary for $X + D\log F$ to be superefficient, namely the Δ-superharmonicity of F can be replaced by the Δ-superharmonicity of $\sqrt{F}$, which is a weaker assumption, see [42] in the finite dimensional case. In particular, $X + D\log F$ is a superefficient estimator of u if $\Delta\sqrt{F} < 0$ on a set of strictly positive $\mathbb{P}$-measure.

Proposition 1.2.15. *Stein-type estimator. For any $\mathbb{P}$-a.s. positive random variable $F \in \mathrm{Dom}(\nabla)$ such that $D_tF \in \mathrm{Dom}(\nabla)$, $t \in [0,T]$, and $(D_tD_tF)_{t\in[0,T]} \in L^1(\Omega \times [0,T], \mathbb{P}_u \otimes \mu)$, we have*

$$\mathbb{E}_u\left[\|X + D\log F - u\|^2_{L^2([0,T],d\mu)}\right] = \mathrm{CRB}(\gamma,\mu,\hat{u}) + 4\,\mathbb{E}_u\left[\frac{\Delta\sqrt{F}}{\sqrt{F}}\right]. \quad (1.2.20)$$

Proof. For any $F \in \mathrm{Dom}(\nabla)$ such that $F > 0$, $\mathbb{P}$-a.s., and $\sqrt{F} \in \mathrm{Dom}(\Delta)$, we have

$$2\frac{D_tD_tF}{F} - \left|\frac{D_tF}{F}\right|^2 = \frac{2}{\sqrt{F}}D_t\left(\frac{D_tF}{\sqrt{F}}\right) = \frac{4}{\sqrt{F}}D_tD_t\sqrt{F}, \qquad t \in [0,T],$$

which implies

$$4\frac{\Delta\sqrt{F}}{\sqrt{F}} = 2\frac{\Delta F}{F} - \int_0^T |D_t \log F|^2 \mu(dt), \quad (1.2.21)$$

and allows us to conclude from Lemma 1.2.13. □

Relation (1.2.20) extends to any $F \in \mathrm{Dom}(\nabla)$ such that $\sqrt{F} \in \mathrm{Dom}(\Delta)$, and $F \geq 0$, $\Delta\sqrt{F} \leq 0$, $\mathbb{P}$-a.s.

In case $(X_t)_{t\in[0,T]}$ is a Brownian motion with constant variance $\sigma_t = \sigma$, $t \in [0,T]$, we have

$$\mathbb{E}_u\left[\|X + D\log F - u\|^2_{L^2([0,T])}\right] \leq \frac{\sigma^2T^2}{2} + 4\,\mathbb{E}_u\left[\frac{\Delta\sqrt{F}}{\sqrt{F}}\right]. \quad (1.2.22)$$

Given $(X_t^{\sigma,1})_{t\in[0,T]}, \ldots, (X_t^{\sigma,N})_{t\in[0,T]}$ N independent samples of $(X_t)_{t\in[0,T]}$, the process $\bar{X}$ defined in (1.2.3) satisfies

$$\mathbb{E}_u^{\sigma/N}\left[\|\bar{X} + D\log F - u\|^2_{L^2([0,T])}\right] = \frac{1}{N}\mathrm{CRB}(\sigma,\mu,\hat{u}) + \frac{4}{N^2}\,\mathbb{E}_u\left[\frac{\Delta\sqrt{F}}{\sqrt{F}}\right].$$

As in [124], the superefficient estimators constructed in this way are minimax in the sense that from Proposition 1.2.6 and Proposition 1.2.14, for all $u \in H$ we have

$$\mathbb{E}_u\left[\|X + D\log F - u\|^2_{L^2([0,T],\mu)}\right] < \mathrm{CRB}(\gamma, \mu, \hat{u}) = \inf_{\xi} \sup_{v \in \Omega} \mathbb{E}_v\left[\int_0^T |\xi_t - v_t|^2 dt\right],$$

provided $\Delta\sqrt{F} < 0$ on a set of strictly positive $\mathbb{P}$-measure, thus showing that the minimax estimator $\hat{u} = (X_t)_{t\in[0,T]}$ is inadmissible.

Both estimators $X_t + D_t \log F$ and $X_t + \mathbb{E}_u[D_t \log F \mid \mathcal{F}_t]$ have bias

$$b_t = \mathbb{E}_u[X_t + D_t \log F - u_t] = \mathbb{E}_u[D_t \log F], \qquad t \in [0, T],$$

which can be bounded as follows from (1.2.21):

$$\begin{aligned} \|b\|^2_{L^2([0,T],\mu)} &= \int_0^T |\,\mathbb{E}_u[D_t \log F]|^2 dt \\ &\le \mathbb{E}_u\left[\int_0^T |D_t \log F|^2 dt\right] \\ &= 2\,\mathbb{E}_u\left[\frac{\Delta F}{F}\right] - 4\,\mathbb{E}_u\left[\frac{\Delta\sqrt{F}}{\sqrt{F}}\right]. \end{aligned}$$

Remark 1.2.16. *In the independent increment case of Section 1.2.2, the formulas obtained in this section also hold for u an adapted process in $L^2(\Omega \times [0,T])$. However, in this case the computation of the gradient $D \log F$ requires in principle the knowledge of X^u, except when u is deterministic, in which case the knowledge of X is sufficient. Thus, assuming u to be deterministic will be necessary for the applications of Section 1.2.7.*

1.2.6 Superharmonic functionals

In this section we give examples of nonnegative superharmonic functionals with respect to the Laplacian Δ. We start by reviewing the construction of such functionals using potential theory on the Gaussian space $(\Omega, H, \mathbb{P})$, and next we turn to cylindrical functionals which will be used in the numerical applications of Section 1.2.7. We assume that $(\Gamma h_k)_{k\ge1}$ is orthogonal in $L^2([0,T], d\mu)$, and we let

$$\lambda_k = \|\Gamma h_k\|_{L^2([0,T],\mu)}, \qquad k \ge 1.$$

The sequence $(h_k)_{k\ge1}$ can be realized as the solution of the eigenvalue problem

$$\Gamma h_k = -\lambda_k^2 \ddot{h}_k, \qquad \dot{h}_k(T) = 0, \qquad k \ge 1, \tag{1.2.23}$$

in case (A), provided $\mu(dt) = dt$, and

$$\Gamma h_k = \lambda_k^2 h_k, \qquad k \ge 1,$$

in case (B) for general μ.

Potentials

We refer to [54] and [47] for the notion of harmonicity on the Wiener space with respect to the Gross Laplacian. From our orthonormality assumption on $(h_k)_{k\geq 1}$, the Laplacian Δ is written as

$$\Delta F = \sum_{i=1}^{n} \partial_i^2 f_n \left(\int_0^T \dot{h}_1(s) dX_s^u, \ldots, \int_0^T \dot{h}_n(s) dX_s^u \right)$$

on cylindrical functionals. Let $(W_t^\Omega)_{t\geq 0}$ denote the standard Ω-valued Wiener process with generator $\frac{1}{2}\Delta_G$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, represented as

$$W_t^\Omega = \sum_{n=1}^{\infty} \int_0^{\cdot} \dot{h}_n(s) \sigma_s^2 ds \frac{\beta_n(t)}{\|h_n\|_H}, \qquad t \in \mathbb{R}_+, \tag{1.2.24}$$

where $(\beta_n(t))_{t\in\mathbb{R}_+}$, $n \geq 1$, are independent standard Brownian motions on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, given as

$$\beta_n(t) = \int_0^T \frac{\dot{h}_n(r)}{\|h_n\|_H} dW_t^\Omega(r), \qquad t \in \mathbb{R}_+, \quad n \geq 1.$$

We have the covariance relation

$$\tilde{\mathbb{E}} \left[\int_0^T \dot{v}_1(r) dW_s^\Omega(r) \int_0^T \dot{v}_2(r) dW_t^\Omega(r) \right] = (s \wedge t) \langle v_1, v_2 \rangle_H, \qquad s, t \in \mathbb{R}_+, \quad v_1, v_2 \in H.$$

In other terms we have

$$\begin{aligned}
\tilde{\mathbb{E}}[W_t^\Omega(a) W_s^\Omega(b)] &= (s \wedge t) \sum_{n=1}^{\infty} \frac{\int_0^a \dot{h}_n(s) \sigma_s^2 ds \int_0^b \dot{h}_n(s) \sigma_s^2 ds}{\|h_n\|_H^2} \\
&= (s \wedge t) \left\langle \sum_{n=1}^{\infty} \frac{\dot{h}_n}{\|h_n\|_H^2} \int_0^a \dot{h}_n(s) \sigma_s^2 ds, \sum_{n=1}^{\infty} \frac{\dot{h}_n}{\|h_n\|_H^2} \int_0^b \dot{h}_n(s) \sigma_s^2 ds \right\rangle_{L^2([0,T], \sigma_t^2 dt)} \\
&= (s \wedge t) \langle \mathbf{1}_{[0,a]}, \mathbf{1}_{[0,b]} \rangle_{L^2([0,T], \sigma_t^2 dt)} \\
&= (s \wedge t) \int_0^{a \wedge b} \sigma_r^2 dr, \qquad 0 \leq a, b \leq T, \quad s, t \in \mathbb{R}_+,
\end{aligned}$$

which shows that $(W_t^\Omega(a))_{a\in[0,T]}$ is a continuous Gaussian martingale with quadratic variation $\sigma_a^2 da$ for fixed $t \in \mathbb{R}_+$.

Denote by $(B_t)_{t\in\mathbb{R}_+}$ the H-valued Wiener process represented as

$$B_t = \sum_{n=0}^{\infty} \frac{h_n}{\|h_n\|_H^2} \beta_n(t),$$

with $\beta_n(t) = \langle B_t, h_n \rangle_H$, $n \geq 1$, and covariance

$$\tilde{\mathbb{E}}[\langle B_s, h_n \rangle_H \langle B_t, h_m \rangle_H] = \mathbf{1}_{\{n=m\}} (s \wedge t), \qquad s, t \in \mathbb{R}_+,$$

i.e.

$$\tilde{\mathbb{E}}[\langle B_t, v_1\rangle_H \langle B_s, v_2\rangle_H] = (s\wedge t)\langle Qv_1, v_2\rangle_H, \qquad s,t\in\mathbb{R}_+, \quad v_1, v_2\in H,$$

where $Q : H \to H$ is the operator with eigenvalues $\{\|h_n\|_H^{-2} \ : \ n\geq 1\}$ in the Hilbert basis $(h_n)_{n\geq 1}$. Itô's formula for Hilbert-valued Wiener processes, cf. Theorem 4.17 of [33], shows that

$$F(B_t) = F(B_0) + \int_0^t \langle DF(B_s), dB_s\rangle_H + \frac{1}{2}\int_0^t \Delta F(B_s)ds, \qquad F\in\mathcal{S},$$

hence $(B_t)_{t\in\mathbb{R}_+}$ has generator $\frac{1}{2}\Delta$.

Dynkin's formula, cf. [38], Theorem 5.1, shows that for all stopping time τ such that $\tilde{\mathbb{E}}[\tau \mid B_0 = \omega] < \infty$ we have, $\mathbb{P}^\sigma(d\omega)$-a.s.:

$$\tilde{\mathbb{E}}[F(B_\tau) \mid B_0 = \omega] - F(\omega) = \frac{1}{2}\tilde{\mathbb{E}}\left[\int_0^\tau \Delta F(B_s)ds \mid B_0 = \omega\right],$$

hence $\Delta F \leq 0$ implies

$$F(\omega) \geq \tilde{\mathbb{E}}[F(B_\tau) \mid B_0 = \omega].$$

For $r > 0$, let

$$\tau_r = \inf\{t\in\mathbb{R}_+ \ : \ B_t \notin \mathfrak{B}_r(B_0)\}$$

denotes the first exit time of $(B_t)_{t\in[0,T]}$ from the open ball $\mathfrak{B}_r(\omega)$ of radius $r > 0$, centered at $B_0 = \omega \in \Omega$. We have the following converse.

Proposition 1.2.17. *Let $F \in \mathrm{Dom}(\Delta)$ be such that ΔF is continuous on Ω, and assume that there exists $r_0 > 0$ such that*

$$F(\omega) \geq \tilde{\mathbb{E}}[F(B_{\tau_r}) \mid B_0 = \omega], \qquad \mathbb{P}_u^\sigma(d\omega) - a.s., \quad 0 < r < r_0. \tag{1.2.25}$$

Then F is Δ-superharmonic on Ω in the sense of Relation (1.2.16).

Proof. From Remark 3, page 134 of [38], we have

$$\frac{1}{2}\Delta F(\omega) = \lim_{n\to\infty} \frac{\tilde{\mathbb{E}}[F(B_{\tau_{1/n}}) \mid B_0 = \omega] - F(\omega)}{\tilde{\mathbb{E}}[\tau_{1/n} \mid B_0 = \omega]}, \tag{1.2.26}$$

which shows that $\Delta F \leq 0$ when (1.2.25) is satisfied. □

This yields in particular the following class of Δ-superharmonic functionals.

Proposition 1.2.18. *Let the potential of $F \geq 0$ be defined by*

$$G(\omega) = \int_0^{+\infty} \tilde{\mathbb{E}}[F(B_t) \mid B_0 = \omega]dt, \qquad \mathbb{P}_u^\sigma(d\omega) - a.s., \tag{1.2.27}$$

assume that $G \in \mathrm{Dom}(\Delta)$ and that ΔG is continuous on Ω. Then G is a Δ-superharmonic on Ω.

Proof. For all $r > 0$ we have

$$\begin{aligned} G(\omega) &= \tilde{\mathbb{E}}\left[\int_0^{\tau_r} F(B_t)dt \Big| B_0 = \omega\right] + \tilde{\mathbb{E}}[G(B_{\tau_r}) \mid B_0 = \omega] \\ &\geq \tilde{\mathbb{E}}[G(B_{\tau_r}) \mid B_0 = \omega], \end{aligned}$$

which shows that G is Δ-superharmonic. □

Note that if F is bounded with bounded support in Ω then G is bounded on Ω, see e.g. Remark 3.5 of [54].

Convolution

Positive superharmonic functionals can also be obtained by convolution, i.e. if F is Δ-superharmonic and G is positive and sufficiently integrable, then

$$\omega \mapsto \int_\Omega G(\tilde{\omega})F(\omega - \tilde{\omega})\mathbb{P}^\sigma(d\tilde{\omega})$$

is positive and Δ-superharmonic.

Cylindrical functionals

Superharmonic functionals on Gaussian space can also be constructed as cylindrical functionals, by composition with finite-dimensional functions. Here we use the expansions of case (A). From the expression of Δ on cylindrical functionals

$$\Delta F = \sum_{i=1}^n \partial_i^2 f_n \left(\lambda_1^{-1} X^u(h_1), \ldots, \lambda_n^{-1} X^u(h_n)\right),$$

we check that

$$F = f_n \left(\lambda_1^{-1} X^u(h_1), \ldots, \lambda_n^{-1} X^u(h_n)\right)$$

is superharmonic on Ω if and only if f_n is superharmonic on $\mathbb{R}^n$. Given $a \in \mathbb{R}$ and $b \in \mathbb{R}^n$, let $f_{n,a,b} : \mathbb{R}^n \to \mathbb{R}$ be defined as

$$f_{n,a,b}(x_1, \ldots, x_n) = \|x + b\|^a = ((x_1 + b_1)^2 + \cdots + (x_n + b_n)^2)^{a/2},$$

then $\sqrt{f_{n,a,b}}$ is superharmonic on $\mathbb{R}^n$, $n \geq 3$, if and only if $a \in [4 - 2n, 0]$. Let

$$F_{n,a,b} = f_{n,a,b} \left(\lambda_1^{-1} X^u(h_1), \ldots, \lambda_n^{-1} X^u(h_n)\right).$$

We have

$$D_t \log F_{n,a,b} = a \sum_{i=1}^n \frac{\lambda_i^{-1} \Gamma h_i(t) \left(b_i + \lambda_i^{-1} X^u(h_i)\right)}{\left|b_1 + \lambda_1^{-1} X^u(h_1)\right|^2 + \cdots + \left|b_n + \lambda_n^{-1} X^u(h_n)\right|^2},$$

and

$$\Delta\sqrt{F_{n,a,b}} \;=\; \sum_{i=1}^{n} \partial_i^2 \sqrt{f_{n,a,b}}\left(\lambda_1^{-1}X^u(h_1),\ldots,\lambda_n^{-1}X^u(h_n)\right),$$

since $(\Gamma h_k)_{k\geq 1}$ is orthogonal in $L^2([0,T],dt)$, hence

$$\frac{\Delta\sqrt{F_{n,a,b}}}{\sqrt{F_{n,a,b}}} \;=\; \frac{a(n-2+a/2)/2}{\left|b_1+\lambda_1^{-1}X^u(h_1)\right|^2+\cdots+|b_n+\lambda_n^{-1}X^u(h_n)|^2},$$

is negative if $4-2n \leq a \leq 0$, which is minimal for $a = 2-n$. We also have

$$\frac{\Delta F_{n,a,b}}{F_{n,a,b}} \;=\; \frac{a(n+a-2)}{\left|b_1+\lambda_1^{-1}X^u(h_1)\right|^2+\cdots+|b_n+\lambda_n^{-1}X^u(h_n)|^2},$$

which is negative for $a \in [2-n,0]$ and vanishes for $a = 2-n$. In this case the estimator is given by

$$D_t \log F_{n,2-n,b} = -(n-2)\sum_{i=1}^{n} \frac{\lambda_i^{-1}\left(b_i+\lambda_i^{-1}X^u(h_i)\right)\Gamma h_i(t)}{\left|b_1+\lambda_1^{-1}X^u(h_1)\right|^2+\cdots+|b_n+\lambda_n^{-1}X^u(h_n)|^2},$$

and from Proposition 1.2.14, inequality (1.2.19) actually also holds as an equality:

$$\mathbb{E}_u\left[\|X+D\log F_{n,2-n,b}-u\|^2_{L^2([0,T],dt)}\right] = \mathrm{CRB}(\sigma,\mu,\hat{u}) - \mathbb{E}_u\left[\int_0^T |D_t \log F_{n,2-n,b}|^2 dt\right], \tag{1.2.28}$$

with

$$\|D\log F_{n,2-n,b}\|^2_{L^2([0,T],dt)} = \frac{(n-2)^2}{\left|b_1+\lambda_1^{-1}X^u(h_1)\right|^2+\cdots+|b_n+\lambda_n^{-1}X^u(h_n)|^2}. \tag{1.2.29}$$

Note that when u is deterministic, any superharmonic functional of the form

$$f_n\left(\lambda_1^{-1}X^u(h_1),\ldots,\lambda_n^{-1}X^u(h_n)\right),$$

can be replaced with

$$f_n\left(\lambda_1^{-1}X(h_1),\ldots,\lambda_n^{-1}X(h_n)\right),$$

which retains the same harmonicity property, and can be directly computed from an observation of X.

The Stein type estimator of u is given by

$$X_t + D_t \log F_{n,2-n,b}, \qquad t \in [0,T],$$

with

$$b_i = \lambda_i^{-1}\langle u, h_i\rangle, \qquad i = 1,\ldots,n,$$

i.e.

$$D_t \log F_{n,2-n,b} = -(n-2)\frac{[\Pi_n X]_t}{\|\Pi_n X\|^2_{L^2([0,T],dt)}},$$

where Π_n denotes the orthogonal projection

$$\Pi_n X(t) := \sum_{k=1}^{n} \lambda_k^{-1} X(h_k)\Gamma h_k(t) = \sum_{k=1}^{n} \lambda_k^{-1}\left(b_k + \lambda_k^{-1} X^u(h_k)\right)\Gamma h_k(t).$$

We have

$$\begin{aligned}
\|D \log F_{n,2-n,b}\|^2_{L^2([0,T]\times\Omega,\mathbb{P}_u\otimes dt)} &= -4\,\mathbb{E}_u\left[\frac{\Delta\sqrt{F_{n,2-n,b}}}{\sqrt{F_{n,2-n,b}}}\right] \\
&= (n-2)^2\,\mathbb{E}_u\left[\frac{1}{\left|\lambda_1^{-1}X(h_1)\right|^2 + \cdots + |\lambda_n^{-1}X(h_n)|^2}\right] \\
&= (n-2)^2\,\mathbb{E}_u\left[\|\Pi_n X\|^{-2}_{L^2([0,T],dt)}\right],
\end{aligned}$$

and

$$\mathbb{E}_u\left[\|X + D\log F_{n,2-n,b} - u\|^2_{L^2([0,T],dt)}\right] = \mathrm{CRB}(\sigma,\mu,\hat{u}) - (n-2)^2\,\mathbb{E}_u\left[\|\Pi_n X\|^{-2}_{L^2([0,T],dt)}\right].$$

Note that the estimator

$$X_t - (n-2)\frac{[\Pi_n X]_t}{\|\Pi_n X\|^2_{L^2([0,T],dt)}}, \qquad t \in [0,T],$$

is of James-Stein type, but it is not a shrinkage operator. Another difference with James-Stein estimators is that here the denominator consists in a sum of squared Gaussians with different variances.

Given $(X_t^1)_{t\in[0,T]}, \ldots, (X_t^N)_{t\in[0,T]}$, N independent samples of $(X_t)_{t\in[0,T]}$, the process

$$\bar{X}_t = \frac{1}{N}\left(X_t^1 + \cdots + X_t^N\right)$$

is a Brownian motion with drift u and quadratic variation $\sigma_t^2 dt/N$ under $\mathbb{P}_u$, and can be used for both efficient and Stein type estimation.

1.2.7 Numerical application

In this section we present numerical simulations which allow us to measure the efficiency of our estimators. We use the framework of case (A) and the superharmonic functionals constructed as cylindrical functionals in the previous section, and we assume that $u \in H$

is deterministic.

We work in the independent increment framework of Section 1.2.2 and we additionally assume that $\sigma_t = \sigma$ is constant, $t \in [0,T]$, i.e. $(X_t)_{t\in[0,T]}$ is a Brownian motion with variance σ^2, $\Gamma h(t) = \sigma^2 h(t)$, $t \in [0,T]$, and

$$\mathrm{CRB}(\sigma,\mu,\hat{u}) = \frac{\sigma^2}{2}T^2.$$

Letting

$$h_n(t) = \frac{\sqrt{2T}}{\sigma\pi\,(n-1/2)} \sin\left(\left(n-\frac{1}{2}\right)\frac{\pi t}{T}\right), \qquad t\in[0,T], \quad n\geq 1,$$

i.e.

$$\dot{h}_n(t) = \frac{1}{\sigma}\sqrt{\frac{2}{T}}\cos\left(\left(n-\frac{1}{2}\right)\frac{\pi t}{T}\right), \qquad t\in[0,T], \quad n\geq 1,$$

provides an orthonormal basis $(h_n)_{n\geq 1}$ of H such that $(\Gamma h_k)_{k\geq 1}$ is orthogonal in $L^2([0,T],dt)$, with

$$\lambda_n = \frac{\sigma T}{\pi(n-1/2)}, \qquad n\geq 1,$$

solution of (1.2.23). The estimator of u will be given by

$$D_t \log F_{n,2-n,b} = -(n-2)\sqrt{\frac{2}{T}}\sum_{k=1}^{n}\frac{X(h_k)}{\left|\lambda_1^{-1}X(h_1)\right|^2+\cdots+\left|\lambda_n^{-1}X(h_n)\right|^2}\sin\left(\left(k-\frac{1}{2}\right)\frac{\pi t}{T}\right),$$

For simulation purposes we will use $X + D\log F$, and construct the (nondrifted) Brownian motion $(X_t^u)_{t\in[0,T]}$ via the Paley-Wiener expansion

$$X_t^u = \sigma^2\sum_{n=1}^{\infty}\eta_n h_n(t) = \sigma\frac{\sqrt{2T}}{\pi}\sum_{n=1}^{\infty}\eta_n\frac{\sin\left(\left(n-\frac{1}{2}\right)\frac{\pi t}{T}\right)}{\left(n-\frac{1}{2}\right)}, \tag{1.2.30}$$

where $(\eta_n)_{n\geq 1}$ are independent standard Gaussian random variables with unit variance under $\mathbb{P}_u$ and

$$\eta_n = \int_0^T \dot{h}_n(s)dX_s^u, \qquad n\geq 1.$$

In this case we have

$$\begin{aligned} & D_t \log F_{n,2-n,b} \qquad\qquad (1.2.31) \\ &= -(n-2)\sqrt{\frac{2}{T}}\sum_{k=1}^{n}\frac{\eta_k + \langle u,h_k\rangle}{\sum_{l=1}^{n}\lambda_l^{-2}\left(\eta_l+\langle u,h_l\rangle\right)^2}\sin\left(\left(k-\frac{1}{2}\right)\frac{\pi t}{T}\right). \end{aligned}$$

Recall that the improvement obtained in comparison with the efficient estimator $\hat{u}$ is not obtained pathwise, but in expectation. The gain of the superefficient estimator $X + D \log F_{n,2-n,b}$ compared to the efficient estimator $\hat{u}$ is given by

$$G(u, \sigma, T, n) := -\frac{4}{\mathrm{CRB}(\sigma, \mu, \hat{u})} \mathbb{E}_u \left[\frac{\Delta \sqrt{F_{n,2-n,b}}}{\sqrt{F_{n,2-n,b}}} \right]$$

as a function of $n \geq 3$. From (1.2.28) and (1.2.31) we have

$$G(u, \sigma, T, n) = 2(n-2)^2 \, \mathbb{E} \left[\left(\sum_{l=1}^{n} \left(\pi \left(l - \frac{1}{2} \right) (\eta_l + \langle u, h_l \rangle) \right)^2 \right)^{-1} \right], \qquad (1.2.32)$$

hence $G(u, \sigma, T, n)$ converges to

$$(n-2)^2 \frac{8}{\pi^2} \, \mathbb{E} \left[\left(\sum_{l=1}^{n} (2l-1)^2 \, \eta_l^2 \right)^{-1} \right], \qquad (1.2.33)$$

as σ tends to infinity. The quantity (1.2.33) can be evaluated as a Gaussian integral to yield (1.2.2). Unlike in the classical Stein method, we stress that here n becomes a free parameter and there is some interest in determining the values of n which yield the best performance.

Proposition 1.2.19. *For all $\sigma, T > 0$, and $u \in H$ we have*

$$G(u, \sigma, T, n) \simeq \frac{6}{n\pi^2}$$

as n goes to infinity.

Proof. Let

$$S_n = \sum_{l=1}^{n} \left(\pi \left(n - l + \frac{1}{2} \right) (\eta_l + \langle u, h_l \rangle) \right)^2, \qquad n \geq 1.$$

We have

$$G(\alpha, \sigma, T, n) = 2(n-2)^2 \, \mathbb{E} \left[\frac{1}{S_n} \right],$$

and by the strong law of large numbers, $n(n-2)^2 {S_n}^{-1}$ converges to $3/\pi^2$ as n goes to infinity, since

$$\lim_{n \to \infty} \frac{\mathbb{E}[S_n]}{n^3} = \frac{\pi^2}{4} \lim_{n \to \infty} \frac{1}{n^3} \sum_{i=1}^{n} (2i-1)^2 = \frac{\pi^2}{3}.$$

Now for all $n > 10$ we have

$$\mathbb{E}_u \left[\left(\frac{(n-2)^3}{S_n} \right)^2 \right] \;\; = \;\; \mathbb{E} \left[\Lambda(u) \left(\frac{(n-2)^3}{S_n} \right)^2 \right]$$

$$
\begin{aligned}
&\leq n^2\pi^2\,\mathbb{E}\left[\Lambda(u)^2\right]^{1/2}\mathbb{E}\left[\left(\sum_{l=1}^{[n/2]}\left(1-\frac{l}{n}+\frac{1}{2n}\right)^2\eta_l^2\right)^{-4}\right]^{1/2}\\
&\leq n^2\frac{4}{\pi^4}\,\mathbb{E}\left[\Lambda(u)^2\right]^{1/2}\mathbb{E}\left[\left(\sum_{l=1}^{[n/2]}\eta_l^2\right)^{-4}\right]^{1/2}\\
&\leq \frac{4n^2}{\pi^4}\,\mathbb{E}\left[\Lambda(u)^2\right]^{1/2}\left(\prod_{k=1}^{4}([n/2]-2k)\right)^{-1/4},
\end{aligned}
$$

hence n^3/S_n is uniformly integrable in $n > 16$, where $[n/2]$ denotes the integer part of $n/2$. This concludes the proof. □

In the sequel we choose $u_t = \alpha t$, $t \in [0,T]$, $\alpha \in \mathbb{R}$. Figure 1.1 gives a sample path representation of the process $X + D\log F$.

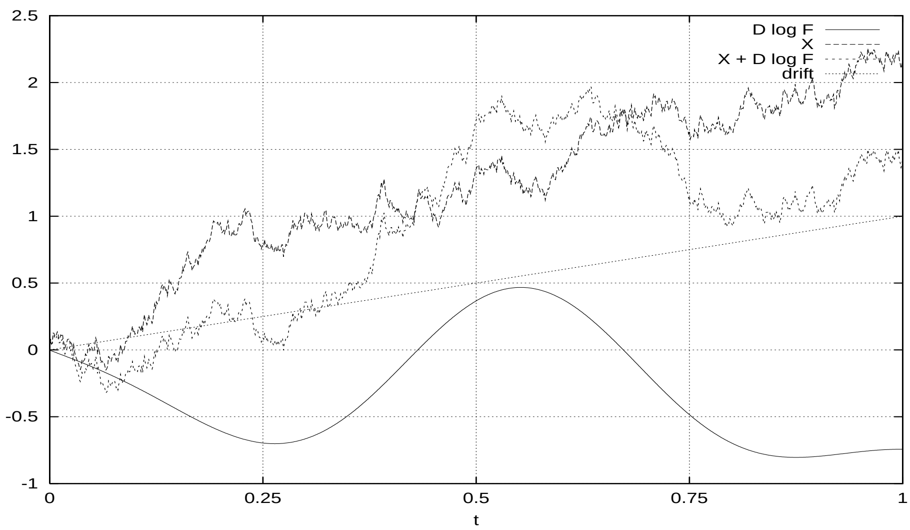

Figure 1.1: $u(t) = t$, $t \in [0,T]$; $n = 5$.

In this case, from (1.2.32) we have

$$
G(\alpha,\sigma,T,n) = 2(n-2)^2\,\mathbb{E}\left[\left(\sum_{l=1}^{n}\left(\pi\left(l-\frac{1}{2}\right)\eta_l - \alpha\frac{\sqrt{2T}}{\sigma}(-1)^l\right)^2\right)^{-1}\right],
$$

from which it follows that $G(\alpha,\sigma,T,n)$ converges to

$$
(n-2)^2\frac{8}{\pi^2}\,\mathbb{E}\left[\left(\sum_{l=1}^{n}(2l-1)^2\eta_l^2\right)^{-1}\right],
$$

when $\alpha^{-2}\sigma^2/T$ tends to infinity, and is equivalent to

$$\left(1 - \frac{2}{n}\right)^2 \frac{\sigma^2}{\alpha^2 T}$$

as $\alpha^{-2}\sigma^2/T$ tends to 0. Figure 1.2 represents the gain in percentage of the superefficient estimator $X + \sigma^2 D \log F_{n,2-n,b}$ compared to the efficient estimator $\hat{u}$ using Monte-Carlo simulations, i.e. we represent $100 \times G(\alpha, \sigma, T, n)$ as a function of $n \geq 3$.

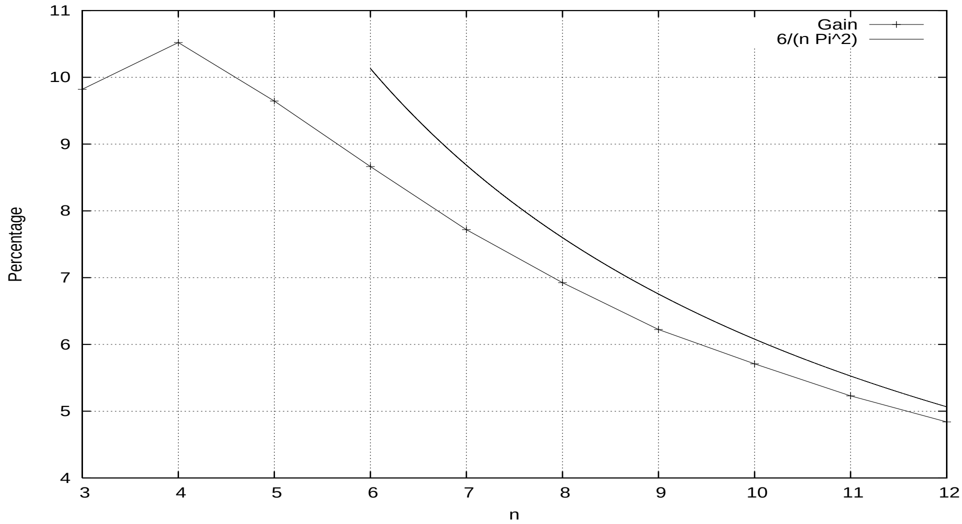

Figure 1.2: Percentage gain as a function of n for 10000 samples and $\alpha = \sigma = T = 1$.

An optimal value

$$n_{\text{opt}} = \text{argmax}\,\{G(\alpha, \sigma, T, n) \ : \ n \geq 3\}$$

of n exists in general and is equal to 4 when $\alpha = \sigma = T = 1$.

Figure 1.3 shows the variation of the gain as a function of n and T for $\alpha = \sigma = 1$.

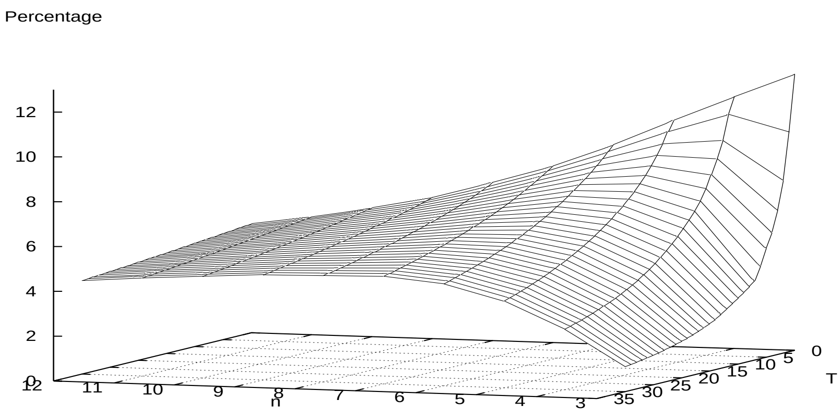

Figure 1.3: Gain as a function of n and T.

Figure 1.4 represents the variation of the gain as a function of n and σ.

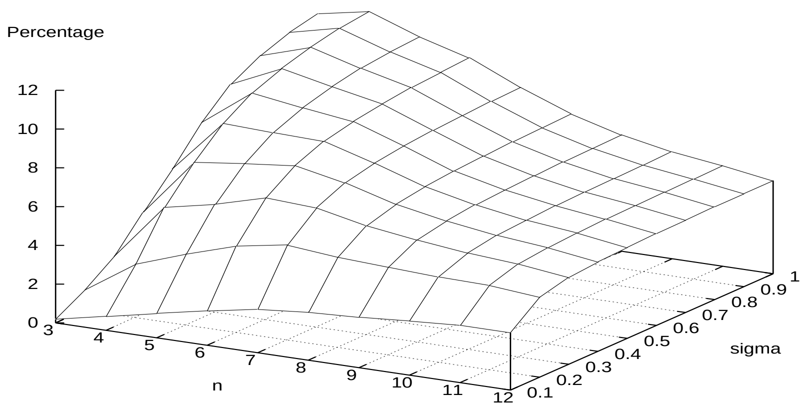

Figure 1.4: Gain as a function of n and σ.

1.2.8 Appendix

The next Proposition is classical in the framework of Gaussian filtering and is needed in Section 1.2.2 for Bayes estimation. Its proof is stated for completeness since we did not find it in the literature.

Proposition 1.2.20. *Let Z be a Gaussian process with covariance operator Γ_τ and drift $v \in H$, and assume that X is a Gaussian process with drift Z and quadratic covariance* *operator* Γ *given* Z. *Then, conditionally to* X, Z *has drift*

$$f \mapsto \langle \chi_t, (\Gamma + \Gamma_\tau)^{-1}\Gamma v\rangle + X((\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f\, \chi_t) \quad \textit{and covariance} \quad \Gamma_\tau(\Gamma + \Gamma_\tau)^{-1}\Gamma.$$

Proof. For convenience of notation, let

$$V(f) = \langle f, (\Gamma + \Gamma_\tau)^{-1}\Gamma v\rangle, \qquad f \in H.$$

For all $f, g \in H$ we have:

$$\begin{aligned} \mathbb{E}\left[\exp\left(iX(f)\right)\right] &= \mathbb{E}\left[\mathbb{E}\left[\exp\left(iX(f)\right)\Big|Z\right]\right] \\ &= \mathbb{E}\left[\exp\left(iZ(f) - \frac{1}{2}\langle f, \Gamma f\rangle\right)\right] \\ &= \exp\left(iV(f) - \frac{1}{2}\langle f, (\Gamma_\tau + \Gamma)f\rangle\right), \end{aligned}$$

and

$$\begin{aligned} &\mathbb{E}\left[\exp\left(iX(g)\right)\exp\left(iZ(f)\right)\right] \\ &\quad= \mathbb{E}\left[\exp\left(iZ(f)\right)\mathbb{E}\left[\exp\left(iX(g)\right)\Big|Z\right]\right] \\ &\quad= \mathbb{E}\left[\exp\left(iZ(f+g) - \frac{1}{2}\langle g, \Gamma g\rangle\right)\right] \\ &\quad= \exp\left(-\frac{1}{2}\langle f+g, \Gamma_\tau(f+g)\rangle - \frac{1}{2}\langle g, \Gamma g\rangle + iV(f+g)\right) \\ &\quad= \exp\left(iV((\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f) + iV(g + (\Gamma + \Gamma_\tau)^{-1}\Gamma f) - \frac{1}{2}\langle \Gamma_\tau f, (\Gamma + \Gamma_\tau)^{-1}\Gamma f\rangle \right. \\ &\qquad \left. -\frac{1}{2}\langle g + (\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f, (\Gamma + \Gamma_\tau)(g + (\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f)\rangle\right) \\ &\quad= \mathbb{E}\left[\exp\left(iX(g)\right)\right. \\ &\qquad \left.\exp\left(iX((\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f) + iV((\Gamma + \Gamma_\tau)^{-1}\Gamma f) - \frac{1}{2}\langle \Gamma_\tau f, (\Gamma_\tau + \Gamma)^{-1}\Gamma f\rangle\right)\right], \end{aligned}$$

which shows that

$$\begin{aligned} &\mathbb{E}\left[\exp\left(iZ(f)\right)\Big|X\right] \\ &\quad= \exp\left(iV((\Gamma + \Gamma_\tau)^{-1}\Gamma f) + iX((\Gamma + \Gamma_\tau)^{-1}\Gamma_\tau f) - \frac{1}{2}\langle \Gamma_\tau f, (\Gamma + \Gamma_\tau)^{-1}\Gamma f\rangle\right). \end{aligned}$$

□

In particular we get the following corollary which is classical in the framework of Gaussian filtering.

Proposition 1.2.21. *Let $(Z_t)_{t\in[0,T]}$ be a Brownian motion with quadratic variation $\tau_t^2 dt$, $\tau \in L^2([0,T],dt)$, and drift $(v_t)_{t\in[0,T]}$, $v \in H$, and let $(X_t)_{t\in[0,T]}$ have drift $(Z_t)_{t\in[0,T]}$ and quadratic variation $(\sigma_t^2)_{t\in[0,T]}$, given Z. Then, conditionally to X, the process $(Z_t)_{t\in[0,T]}$ has drift*

$$\int_0^t \frac{\sigma_s^2}{\tau_s^2+\sigma_s^2} dv_s + \int_0^t \frac{\tau_s^2}{\tau_s^2+\sigma_s^2} dX_s \quad \textit{and variance} \quad \int_0^t \frac{\tau_s^2\sigma_s^2}{\tau_s^2+\sigma_s^2} ds, \qquad t \in [0,T].$$

Acknowledgement

We thank the editors and referees for suggestions which led to several improvements of this paper, and in particular for communicating to us the references [14] and [135].

1.3 Cas d'un processus de Poisson

Cette section est un article écrit en collaboration avec Nicolas Privault et publié en [108].

1.3.1 Introduction

Consider a Poisson process $(X_t)_{t\in[0,T]}$ with intensity u of the form $u(t) = \lambda h(t)$, $t \in [0, T]$, under a probability $\mathbb{P}_u$, where $(h(t))_{t\in[0,T]}$ is a given deterministic function. As is well-known, cf. [71], or [75], p. 351, Example 2, Ch. XIX, the classical parametric maximum likelihood estimator (MLE)

$$\hat{\lambda}_T := \frac{X_T}{h(T)}$$

of λ on the time interval $[0, T]$ is obtained by maximization of the Girsanov density, i.e. under the condition:

$$\frac{d}{d\lambda}\left(\lambda^{X_T} e^{-(\lambda h(T)-T)} \prod_{k=1}^{X_T} \dot{h}(T_k)\right) = \left(\frac{X_T}{\lambda} - h(T)\right) \lambda^{X_T} e^{-(\lambda h(T)-T)} \prod_{k=1}^{X_T} \dot{h}(T_k) = 0.$$

The MLE $\hat{\lambda}_T$ is efficient in the sense that it attains the Cramer-Rao bound

$$\mathbb{E}_u\left[|\hat{\lambda}_T - \lambda|^2\right] = \frac{\lambda}{h(T)}$$

over all unbiased estimators ζ_T satisfying $\mathbb{E}_u[\zeta_T] = \lambda$, for all $\lambda > 0$, where $\mathbb{E}_u$ denotes expectation under $\mathbb{P}_u$.

In this paper we construct superefficient estimators for the intensity parameter $\lambda > 0$ when the intensity $(u(t))_{t\in[0,T]}$ of $(X_t)_{t\in[0,T]}$ is constrained to have the form $u(t) = \lambda h(t)$, $t \in [0, T]$.

We use integration by parts and harmonic analysis on the Poisson space, via the technique introduced by Stein [124] for the estimation of the mean of a standard Gaussian random vector Z in $\mathbb{R}^d$, and extended to drift estimation on the Wiener space in [107], [105]. Recall that Stein's argument relies on:

a) the integration by parts

$$\mathbb{E}_\mu[(Z_i - \mu_i)g_i(Z)] = \mathbb{E}_\mu[\partial_i g_i(Z)], \qquad (1.3.1)$$

where $\mathbb{E}_\mu$ denotes expectation under the standard Gaussian measure with mean $\mu \in \mathbb{R}^d$,

b) the chain rule of derivation for the partial derivative ∂_i on $\mathbb{R}^d$,

c) the existence and properties of non-negative superharmonic functions on $\mathbb{R}^d$ for $d \geq 3$.

Precisely, given an estimator of $\mu \in \mathbb{R}^d$ of the form $Z + \operatorname{grad} \log f(Z)$, where $f : \mathbb{R}^d \to \mathbb{R}^d$ is sufficiently smooth, one gets, using the chain rule of derivation,

$$\begin{aligned}
& \mathbb{E}_\mu[\|Z + \operatorname{grad} \log f(Z) - \mu\|^2_{\mathbb{R}^d}] \\
&= \mathbb{E}_\mu[\|Z - \mu\|^2_{\mathbb{R}^d}] + \mathbb{E}_\mu[\|\operatorname{grad} \log f(Z)\|^2_{\mathbb{R}^d}] + 2 \sum_{i=1}^d \mathbb{E}_\mu[(Z_i - \mu_i)\partial_i \log f(Z)] \\
&= d + \mathbb{E}_\mu[\|\operatorname{grad} \log f(Z)\|^2_{\mathbb{R}^d}] + 2\, \mathbb{E}_\mu \left[\sum_{i=1}^d \partial_i^2 \log f_i(Z) \right] \qquad (1.3.2) \\
&= d + 4 \sum_{i=1}^d \mathbb{E}_\mu \left[\frac{\partial_i^2 \sqrt{f}(Z)}{\sqrt{f}(Z)} \right],
\end{aligned}$$

i.e. $Z + \operatorname{grad} \log f(Z)$ improves in the mean square sense over the maximum likelihood estimator (MLE) Z if $d \geq 3$ and $\sqrt{f}$ is superharmonic on $\mathbb{R}^d$.

Integration by parts for $g : \mathbb{N} \to \mathbb{R}$ with respect to the discrete Poisson distribution $P(X = k) = e^{-\lambda}\lambda^k/k!$, $k \in \mathbb{N}$, can be written as

$$\mathbb{E}_\lambda[(X - \lambda)g(X)] = \lambda\, \mathbb{E}_\lambda[g(X+1) - g(X)].$$

where $\mathbb{E}_\lambda$ denotes expectation under the Poisson distribution with parameter $\lambda > 0$, and has been used to derive Stein identities for jump processes, such as

$$\begin{aligned}
\mathbb{E}_\lambda[|X - \lambda + g(X)|^2] &= \lambda + \mathbb{E}_\lambda[|g(X)|^2] + 2\, \mathbb{E}_\lambda[(X - \lambda)g(X)] \\
&= \lambda + \mathbb{E}_\lambda[|g(X)|^2] + 2\lambda\, \mathbb{E}_\lambda[g(X+1) - g(X)],
\end{aligned}$$

cf. [6], [27]. However the absence of chain rule for the finite difference operation $g \mapsto g(\cdot + 1) - g(\cdot)$ prevents us from continuing the calculation as in (1.3.2) above, and from using superharmonic functions as in the Gaussian case. On a more general level the derivation property requirement prevents us from using finite difference gradients on Poisson functionals cf. e.g. [97].

In this paper we apply Stein's argument on the Poisson space, and construct superefficient estimators for the discrete Poisson law, by replacing the Stein equation (1.3.1) with the integration by parts formula of [23], [39], extended to arbitrary intensity functions on the Poisson space as in [102], in which the gradient ∇ satisfies the chain rule of derivation. When $u(t)$ has the form $u(t) = \lambda h(t)$ we apply our result to the parametric estimation of the Poisson process intensity $\lambda > 0$ via estimators of the form

$$\hat{\lambda}_T + \frac{c}{h(T)} \mathbf{1}_{\{X_T = 0\}} + \frac{1}{h(T)} \nabla_T \log F,$$

where F is a positive superharmonic random variable on the Poisson space, $c \in \mathbb{R}$ is a suitably chosen constant, and ∇_T is a gradient operator on the Poisson space.

Unlike in the Gaussian case, the Laplacian considered here contains first order terms and is not the standard Laplacian on $\mathbb{R}^d$. As a consequence the $d \geq 3$ dimension condition imposed in the Gaussian case can be waived and superharmonic functionals can be constructed as functions of d jump times for $d \geq 1$.

We proceed as follows. In Section 1.3.2 we introduce the Poisson space and derive the Cramer-Rao bound for a non-parametric estimator of the intensity. Our proof uses stochastic calculus, and in this respect it differs from the ones usually found in the literature, cf. e.g. § 1.2 of [72]. In Section 1.3.3 we recall the elements of analysis and integration by parts on the Poisson space which will be needed in Section 1.3.4 to construct superefficient estimators for the intensity of a Poisson process. In case u has the form $u(t) = \lambda t$, numerical applications and simulations are given in Section 1.3.5 using simple examples of (pseudo) superharmonic functionals.

1.3.2 Preliminaries

In this section we state some notation on the Poisson space and Poisson process, and derive the Cramer-Rao bound. Let $T > 0$ and consider $(X_t)_{t\in[0,T]}$ the canonical process on

$$\Omega = \left\{ \omega = \sum_{k=1}^{n} \delta_{t_k} \ : \ 0 \leq t_1 < \cdots < t_n \leq T, \quad n \in \mathbb{N} \cup \{\infty\} \right\},$$

defined as

$$X_t(\omega) = \omega([0,t]), \qquad t \in [0,T],$$

where δ_x denotes the Dirac measure at $x \in [0,T]$. Let $(T_k)_{k\geq 1}$ denote the jump times of $(X_t)_{t\in[0,T]}$, i.e. any $\omega \in \{X_T = n\}$ is written as

$$\omega = \sum_{k=1}^{n} \delta_{T_k}.$$

Let $\mathbb{P}$ denote the standard Poisson measure on Ω, under which $(X_t)_{t\in\mathbb{R}_+}$ is a standard Poisson process, and let $(\mathcal{F}_t)_{t\in[0,T]}$ denote the filtration generated by $(X_t)_{t\in[0,T]}$.

Definition 1.3.1. *Let $\mathcal{P}$ denote the set of functions of the form*

$$u(t) = \int_0^t \dot{u}(s)ds, \qquad t \in [0,T],$$

where $\dot{u} : [0,T] \to [0,\infty)$ is a non-negative function.

Let now $u \in \mathcal{P}$. By the Girsanov theorem, the measure $\mathbb{P}_u$ on Ω, under which the canonical process $(X_t)_{t\in[0,T]}$ is a Poisson process with intensity $\dot{u}(t)dt$, is absolutely continuous with respect to $\mathbb{P}$ with

$$d\mathbb{P}_u = \Lambda(u)d\mathbb{P},$$

where

$$\Lambda(u) = \exp\left(-\int_0^T (\dot{u}(s) - 1)ds\right) \prod_{k=1}^{X_T} \dot{u}(T_k)$$

denotes the Girsanov density. In the sequel we will denote by $\mathbb{E}_u$ the expectation under $\mathbb{P}_u$ and let $L^2_u(\Omega) = L^2(\Omega, \mathbb{P}_u)$.

We close this section with a derivation of the Cramer-Rao inequality using stochastic calculus, for non-parametric estimation of the intensity. In case the intensity is constrained to be constant on intervals, our bound can be recovered from the Cramer-Rao inequality for arbitrary finite dimensional estimators, cf. Theorem 1.5 of [72].

Definition 1.3.2. *An estimator ξ_t of $u \in \mathcal{P}$ is called unbiased if*

$$\mathbb{E}_u[\xi_t] = u(t), \qquad t \in [0, T],$$

and adapted if the process $(\xi_t)_{t\in[0,T]}$ is adapted to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ generated by $(X_t)_{t\in[0,T]}$.

Here, X_t can be considered as an unbiased maximum likelihood estimator of its own intensity $u(t)$ under $\mathbb{P}_u$, $t \in [0, T]$. From the next proposition, this estimator is efficient since its mean square error is equal to

$$\mathbb{E}_u\left[|X_t - u(t)|^2\right] = u(t), \qquad t \in [0, T]. \tag{1.3.3}$$

Lemma 1.3.3. *Let ξ be a square integrable estimator of u and let v in $\mathcal{P}$ such that $\dot{v} \le \dot{u}$ then we have that*

$$\frac{d^+}{d\varepsilon}\,\mathbb{E}[(\xi_t - u(t))\Lambda(u + \varepsilon v)]_{|\varepsilon=0} = \mathbb{E}\left[(\xi_t - u(t))\frac{d^+}{d\varepsilon}\Lambda(u + \varepsilon v)_{|\varepsilon=0}\right],$$

where $\frac{d^+}{d\varepsilon}$ denotes the right derivative.

Proof. Let $t \in [0, 1]$ and $0 < \varepsilon < 1$.

$$\begin{aligned} &\frac{1}{\varepsilon}(\mathbb{E}_{u+\varepsilon v}[\xi_t - u(t)] - \mathbb{E}_u[\xi_t - u(t)]) \\ = &\frac{1}{\varepsilon}\,\mathbb{E}[(\xi_t - u(t))(\Lambda(u + \varepsilon v) - \Lambda(u))] \\ = &\frac{1}{\varepsilon}\,\mathbb{E}_u[(\xi_t - u(t))f(\varepsilon)], \end{aligned}$$

where

$$f(\varepsilon) := \exp\left(-\varepsilon\int_0^T \dot{v}(s)ds + \int_0^T \log\left(1 + \varepsilon\mathbf{1}_{\{\dot{u}(s)\neq 0\}}\frac{\dot{v}(s)}{\dot{u}(s)}\right) dX_s - 1\right).$$

Let $X^u_t := X_t - \dot{u}(t),\ t \in [0,T]$. For ε small enough we have that

$$f(\varepsilon) = \varepsilon \int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s)} dX^u_s + \varepsilon^2 f''(\varepsilon\theta), \quad \theta \in (0,1).$$

Let t in $(0,1)$ we have that

$$\begin{aligned}
\left|f''(t)\right| &= (f(t)+1)\left|\left(-\int_0^T \dot{v}(s)ds + \int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}_s + t\dot{v}(s)} dX_s\right)^2 - \int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}^2(s)}{(\dot{u}_s + t\dot{v}(s))^2} ds\right| \\
&\leq (f(t)+1)\left[\left(-\int_0^T \dot{v}(s)ds + \int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}_s + t\dot{v}(s)} dX_s\right)^2 + \int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}^2(s)}{\dot{u}^2(s)} ds\right]
\end{aligned}$$

Since $t \in (0,1)$ and $\dot{u}, \dot{v} \geq 0$ the following inequality holds

$$\begin{aligned}
&\left|-\int_0^T \dot{v}(s)ds + \int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}_s + t\dot{v}(s)} dX_s\right| \\
\leq &\left|-\int_0^T \dot{v}(s)ds + \int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s) + \dot{v}(s)} dX_s\right| + \left|\int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s)} dX^u_s\right|.
\end{aligned}$$

As a consequence we have that

$$\begin{aligned}
&\left|-\int_0^T \dot{v}(s)ds + \int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}_s + t\dot{v}(s)} dX_s\right|^2 \\
\leq\ & 2\left|-\int_0^T \dot{v}(s)ds + \int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s) + \dot{v}(s)} dX_s\right|^2 + 2\left|\int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s)} dX^u_s\right|^2 \\
=\ & 2\left|\int_0^T \frac{\dot{v}(s)}{\dot{u}_s + \dot{v}(s)} - \dot{v}(s)ds + \int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s) + \dot{v}(s)} dX^u_s\right|^2 + 2\left|\int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s)} dX^u_s\right|^2 \\
\leq\ & 4\left|\int_0^T \frac{\dot{v}(s)}{\dot{u}_s + \dot{v}(s)} - \dot{v}(s)ds\right|^2 + 4\left|\int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s) + \dot{v}(s)} dX^u_s\right|^2 + 2\left|\int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s)} dX^u_s\right|^2 \\
=:\ & g_1 + g_2 + g_3.
\end{aligned}$$

Thus, we can deduce an estimate for f'' which is given by

$$\left|f''(t)\right| \leq \sum_{i=1}^4 (f(t)+1) g_i \leq \sum_{i=1}^4 \exp\left(\int_0^T \log\left(1 + \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s)}\right) dX_s\right) g_i$$

where $g_4 := \int_0^T \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}^2(s)}{\dot{u}^2(s)} ds$. Finally we have that

$$\left|\mathbb{E}_u\left[(\xi_t - u_t) f''(\varepsilon\theta)\right]\right| \leq \sum_{i=1}^4 \mathbb{E}_u\left[|\xi_t - u(t)| \exp\left(\int_0^T \log\left(1 + \mathbf{1}_{\{\dot{u}(s)\neq 0\}} \frac{\dot{v}(s)}{\dot{u}(s)}\right) dX_s\right) g_i\right]$$

$$
\begin{aligned}
&\le \sum_{i=1}^{4} \mathbb{E}_u\left[|\xi_t - u(t)|^2\right]^{1/2} \mathbb{E}_u\left[\exp\left(2\int_0^T \log\left(1+\mathbf{1}_{\{\dot{u}(s)\neq 0\}}\frac{\dot{v}(s)}{\dot{u}(s)}\right) dX_s\right) g_i^2\right]\\
&\le \sum_{i=1}^{4} \mathbb{E}_u\left[|\xi_t - u(t)|^2\right]^{1/2} \mathbb{E}_u\left[g_i^4\right]^{1/4} \mathbb{E}_u\left[\exp\left(4\int_0^T \log\left(1+\mathbf{1}_{\{\dot{u}(s)\neq 0\}}\frac{\dot{v}(s)}{\dot{u}(s)}\right) dX_s\right)\right]^{1/4}\\
&\le \sum_{i=1}^{4} \Bigg[\mathbb{E}_u\left[|\xi_t - u(t)|^2\right]^{1/2} \mathbb{E}_u\left[g_i^4\right]^{1/4} \exp\left(\int_0^T \log\left(1+\mathbf{1}_{\{\dot{u}(s)\neq 0\}}\frac{\dot{v}(s)}{\dot{u}(s)}\right) \dot{u}(s)ds\right)\\
&\qquad \times \mathbb{E}_u\left[\exp\left(4\int_0^T \log\left(1+\mathbf{1}_{\{\dot{u}(s)\neq 0\}}\frac{\dot{v}(s)}{\dot{u}(s)}\right) dX_s^u\right)\right]^{1/4}\Bigg]\\
&< \infty.
\end{aligned}
$$

We give some precision about the finiteness of the previous quantities. First note that

$$
\exp\left(\int_0^T \log\left(1+\mathbf{1}_{\{\dot{u}(s)\neq 0\}}\frac{\dot{v}(s)}{\dot{u}(s)}\right) \dot{u}(s)ds\right) \le \exp\left(\int_0^T \dot{u}(s)+\dot{v}(s)ds\right) < \infty.
$$

Using the Exponential formula presented in [15, p. 8] we have that

$$
\mathbb{E}_u\left[\exp\left(4\int_0^T \log\left(1+\mathbf{1}_{\{\dot{u}(s)\neq 0\}}\frac{\dot{v}(s)}{\dot{u}(s)}\right) dX_s^u\right)\right] = \exp\left(-T\int_0^T \left[1-\left(1+\mathbf{1}_{\{\dot{u}(s)\neq 0\}}\frac{\dot{v}(s)}{\dot{u}(s)}\right)^4\right] \dot{u(s)}ds\right)
$$

which is finite provided $\int_0^T \left|1-\left(1+\mathbf{1}_{\{\dot{u}(s)\neq 0\}}\frac{\dot{v}(s)}{\dot{u}(s)}\right)^4\right| \dot{u}(s)ds < \infty$ which is true since we have supposed that $\dot{v} \le \dot{u}$. □

Proposition 1.3.4. *Cramer-Rao inequality. Let $u \in \mathcal{P}$ and $t \in [0,T]$. For any unbiased and adapted estimator ξ_t of $u(t)$ we have*

$$
\mathbb{E}_u\left[|\xi_t - u(t)|^2\right] \ge u(t), \qquad u \in \mathcal{P}, \tag{1.3.4}
$$

where for all $u \in \mathcal{P}$ the lower bound $u(t)$ is attained by $\xi_t = X_t$.

Proof. Since ξ_t is unbiased, for all $v \in \mathcal{P}$ with $\dot{v} \le \dot{u}$ and $\varepsilon \in (-1,1)$ we have

$$
\mathbb{E}_{u+\varepsilon v}[\xi_t] = u(t) + \varepsilon v(t) = \mathbb{E}_{u+\varepsilon v}[u(t)] + \varepsilon v(t),
$$

hence

$$
\begin{aligned}
v(t) &= \frac{d^+}{d\varepsilon}\mathbb{E}_{u+\varepsilon v}[\xi_t - u(t)]_{|\varepsilon=0}\\
&= \frac{d^+}{d\varepsilon}\mathbb{E}[(\xi_t - u(t))\Lambda(u+\varepsilon v)]_{|\varepsilon=0}
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbb{E}\left[(\xi_t - u(t))\frac{d^+}{d\varepsilon}\Lambda(u+\varepsilon v)_{|\varepsilon=0}\right] \\
&= \mathbb{E}_u\left[(\xi_t - u(t))\frac{d^+}{d\varepsilon}\log\Lambda(u+\varepsilon v)_{|\varepsilon=0}\right] \\
&= \mathbb{E}_u\left[(\xi_t - u(t))\int_0^T \mathbf{1}_{\{\dot u(s)\neq 0\}}\frac{\dot v(s)}{\dot u(s)}(dX_s - \dot u(s)ds)\right] \\
&= \mathbb{E}_u\left[(\xi_t - u(t))\int_0^t \mathbf{1}_{\{\dot u(s)\neq 0\}}\frac{\dot v(s)}{\dot u(s)}(dX_s - \dot u(s)ds)\right].
\end{aligned}
$$

Note that the adaptedness hypothesis on the estimator ξ_t was used to get the last equality above, and that the exchange between expectation and derivative is due to Lemma 1.3.3. Thus, by the Cauchy-Schwarz inequality and the Itô isometry, we have

$$
\begin{aligned}
v^2(t) &\leq \mathbb{E}_u\left[\left(\int_0^t \mathbf{1}_{\{\dot u(s)\neq 0\}}\frac{\dot v(s)}{\dot u(s)}(dX_s - \dot u(s)ds)\right)^2\right]\mathbb{E}_u[|\xi_t - u(t)|^2] \\
&= \int_0^t \mathbf{1}_{\{\dot u(s)\neq 0\}}\frac{|\dot v(s)|^2}{\dot u(s)}ds\,\mathbb{E}_u[|\xi_t - u(t)|^2].
\end{aligned}
$$

It then suffices to take

$$\dot v(s) := \dot u(s), \qquad s \in [0,t],$$

to get

$$\mathrm{Var}_u[\xi_t] = \mathbb{E}_u[|\xi_t - u(t)|^2] \geq u(t), \tag{1.3.5}$$

which leads to (1.3.4). As noted in (1.3.3), $\hat u_t = X_t$ is clearly unbiased under $\mathbb{P}_u$ and attains the lower bound $u(t)$. □

1.3.3 Analysis on the Poisson space

In this section we recall the elements of analysis and integration by parts on the Poisson space which will be needed for the construction of Stein estimators.

Definition 1.3.5. *We denote by $\mathcal{S}$ the space of Poisson functionals of the form*

$$F = f_0 \mathbf{1}_{\{X_T=0\}} + \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} f_n(T_1,\ldots,T_n), \tag{1.3.6}$$

where $f_0 \in \mathbb{R}$ and f_n, $n \geq 1$, is $\mathcal{C}^1$ on $\{0 \leq t_1 \leq \cdots \leq t_n \leq T\}$, and satisfying the continuity condition

$$f_n(t_1,\ldots,t_n) = f_{n+1}(t_1,\ldots,t_n,T), \qquad 0 \leq t_1 \leq \cdots \leq t_n \leq T, \quad n \in \mathbb{N}. \tag{1.3.7}$$

Recall that for all $F \in \mathcal{S}$ of the form (1.3.6), letting

$$\tilde{f}_n(t_1, \ldots, t_n) := f_n\left(t_{(1)}, \ldots, t_{(n)}\right), \qquad n \geq 1,$$

where $(t_{(1)}, \ldots, t_{(n)})$ represents the arrangement of $(t_1, \ldots, t_n) \in [0, T]^n$ in increasing order, we have:

$$\begin{aligned}
&\mathbb{E}_u[F] && (1.3.8)\\
&= e^{-u(T)} f_0 + e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^T \cdots \int_0^T \tilde{f}_n(t_1, \ldots, t_n) \dot{u}(t_1) \cdots \dot{u}(t_n) dt_1 \cdots dt_n \\
&= e^{-u(T)} f_0 + e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^{u(T)} \cdots \int_0^{u(T)} \tilde{f}_n(u^{-1}(t_1), \ldots, u^{-1}(t_n)) dt_1 \cdots dt_n.
\end{aligned}$$

From now we assume that $u \in \mathcal{P}$ is such that $\dot{u}(t)$ is lower bounded by a (strictly) positive constant for all $t \in \mathbb{R}_+$ in order to satisfy the integrability conditions needed in the sequel.

Definition 1.3.6. *For $F \in \mathcal{S}$ of the form* (1.3.6), *let*

$$\dot{D}_t F = -\sum_{n=1}^{\infty} \mathbf{1}_{\{X_T = n\}} \sum_{k=1}^{n} \mathbf{1}_{[0,T_k]}(t) \frac{1}{\dot{u}(T_k)} \partial_k f_n(T_1, \ldots, T_n), \qquad t \in [0, T],$$

for F of the form (1.3.6), *where $\partial_k f_n$ denotes the partial derivative of f_n with respect to its k-th variable.*

Let

$$H = \left\{ v : [0, T] \to \mathbb{R} \ : \ v(t) := \int_0^t \dot{v}(s) ds, \ t \in [0, T], \ \dot{v} \in L^2([0, T], \dot{u}(t) dt) \right\}$$

denote the Cameron-Martin space with inner product

$$\langle v, w \rangle_H = \int_0^T \dot{v}(s) \dot{w}(s) \dot{u}(s) ds, \qquad v, w \in H.$$

We have

$$\langle DF, v \rangle_H = \int_0^T \dot{v}(t) \dot{D}_t F \dot{u}(t) dt, \qquad F \in \mathcal{S}, \quad v \in H.$$

Let $L^2_u(\Omega; H)$ denote the space of processes $(v(t))_{t \in [0,T]}$ of the form

$$v(t) = \int_0^t \dot{v}(s) ds, \qquad t \in [0, T],$$

such that

$$\mathbb{E}_u \left[\int_0^T |\dot{v}(s)|^2 \dot{u}(s) ds \right] < \infty.$$

We now turn to the definition of the operator δ adjoint of D. Note that as D has the derivation property, the operator δ is different from the Kabanov-Skorohod integral [69], whose adjoint is a finite difference operator [97]. See [101], [103] for a comparative study of these gradient and Skorohod type integral operators.

Proposition 1.3.7. *i) The operator D is closable and admits a closable adjoint δ : $L^2_u(\Omega; H) \to L^2_u(\Omega)$ under $\mathbb{P}_u$, which satisfies the integration by parts formula*

$$\mathbb{E}_u[F\delta(v)] = \mathbb{E}_u[\langle v, DF\rangle_H], \qquad F \in \mathrm{Dom}(D), \quad v \in \mathrm{Dom}(\delta). \tag{1.3.9}$$

ii) We have

$$\delta(v) = \int_0^T \dot{v}(t)(dX_t - \dot{u}(t)dt), \tag{1.3.10}$$

for every $\mathcal{F}_t$-adapted process $v \in L^2_u(\Omega; H)$.

Proof. By standard integration by parts we first prove (1.3.9) when $v \in H$:

$$\begin{aligned}
&\mathbb{E}_u[\langle DF, v\rangle_H] \\
&= -e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^{n} \int_0^T \cdots \int_0^T \int_0^{t_k} \frac{\dot{u}(s)}{\dot{u}(t_k)} \dot{v}(s) ds \partial_k \tilde{f}_n(t_1, \ldots, t_n) \dot{u}(t_1) \cdots \dot{u}(t_n) dt_1 \cdots dt_n \\
&= -e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^{n} \int_0^{u(T)} \cdots \int_0^{u(T)} \int_0^{u^{-1}(t_k)} \dot{v}(s)\dot{u}(s) ds \frac{\partial}{\partial t_k} \tilde{f}_n(u^{-1}(t_1), \ldots, u^{-1}(t_n)) dt_1 \cdots dt_n \\
&= -e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^{n} \int_0^{u(T)} \cdots \int_0^{u(T)} \int_0^{t_k} \dot{v}(u^{-1}(s)) ds \frac{\partial}{\partial t_k} \tilde{f}_n(u^{-1}(t_1), \ldots, u^{-1}(t_n)) dt_1 \cdots dt_n \\
&= e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^{n} \int_0^{u(T)} \cdots \int_0^{u(T)} \tilde{f}_n(u^{-1}(t_1), \ldots, u^{-1}(t_n)) \dot{v}(u^{-1}(t_k)) dt_1 \cdots dt_n \\
&\quad - e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int_0^{u(T)} \dot{v}(u^{-1}(s)) ds \int_0^{u(T)} \cdots \int_0^{u(T)} \tilde{f}_n(u^{-1}(t_1), \ldots, u^{-1}(t_{n-1}), T) dt_1 \cdots dt_{n-1} \\
&= e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k=1}^{n} \int_0^T \cdots \int_0^T \tilde{f}_n(t_1, \ldots, t_n) \dot{v}(t_k) \dot{u}(t_1) \cdots \dot{u}(t_n) dt_1 \cdots dt_n \\
&\quad - e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int_0^T \dot{v}(s)\dot{u}(s) ds \int_0^T \cdots \int_0^T \tilde{f}_n(t_1, \ldots, t_{n-1}, T) \dot{u}(t_1) \cdots \dot{u}(t_{n-1}) dt_1 \cdots dt_{n-1}.
\end{aligned}$$

The continuity condition (1.3.7), i.e.

$$\tilde{f}_{n-1}(t_1, \ldots, t_{n-1}) = \tilde{f}_n(t_1, \ldots, t_{n-1}, T), \qquad n \geq 1,$$

yields

$$\begin{aligned}
&\mathbb{E}_u[\langle DF, v\rangle_H] \\
&= e^{-u(T)} \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^T \cdots \int_0^T \tilde{f}_n(t_1, \ldots, t_n) \dot{u}(t_1) \cdots \dot{u}(t_n) \sum_{k=1}^{n} \dot{v}(t_k) dt_1 \cdots dt_n \\
&\quad - e^{-u(T)} \int_0^T \dot{v}(s)\dot{u}(s) ds \sum_{n=0}^{\infty} \frac{1}{n!} \int_0^T \cdots \int_0^T \tilde{f}_n(t_1, \ldots, t_n) \dot{u}(t_1) \cdots \dot{u}(t_n) dt_1 \cdots dt_n
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}_u\left[F\left(\sum_{k=1}^{X_T}\dot v(T_k)-\int_0^T \dot v(s)\dot u(s)ds\right)\right]\\
&= \mathbb{E}_u\left[F\left(\int_0^T \dot v(s)(dX(s)-\dot u(s)ds\right)\right].
\end{aligned}
$$

Next we define $\delta(Gv)$, $G\in\mathcal{S}$, $v\in H$, by

$$
\delta(Gv)=G\int_0^T \dot v(t)(dX_t-\dot u(t)dt)-\langle v,DG\rangle_H, \tag{1.3.11}
$$

with for all $G\in\mathcal{S}$:

$$
\begin{aligned}
\mathbb{E}_u[G\langle DF,v\rangle_H] &= \mathbb{E}_u[\langle D(FG),v\rangle_H - F\langle DG,v\rangle_H]\\
&= \mathbb{E}_u\left[F\left(G\int_0^T v(t)dX_t-\langle DG,v\rangle_H\right)\right]\\
&= \mathbb{E}_u[F\delta(Gv)],
\end{aligned}
$$

which proves (1.3.9). The closability of D then follows from the integration by parts formula (1.3.9): if $(F_n)_{n\in\mathbb{N}}\subset\mathcal{S}$ is such that $F_n\to 0$ in $L^2_u(\Omega)$ and $DF_n\to U$ in $L^2_u(\Omega;H)$, then (1.3.9) implies

$$
\begin{aligned}
|\,\mathbb{E}_u[\langle U,Gv\rangle_H]| &\le |\,\mathbb{E}_u[F_n\delta(Gv)]-\mathbb{E}_u[\langle U,Gv\rangle_H]|+|\,\mathbb{E}_u[F_n\delta(Gv)]|\\
&= |\,\mathbb{E}_u[\langle DF_n-U,Gv\rangle_H]|+|\,\mathbb{E}_u[F_n\delta(Gv)]|\\
&\le \|\langle DF_n,v\rangle_H-\langle U,v\rangle_H\|_{L^2_u(\Omega)}\|G\|_{L^2_u(\Omega)}+\|F_n\|_{L^2_u(\Omega)}\|\delta(Gv)\|_{L^2_u(\Omega)},
\end{aligned}
$$

$n\in\mathbb{N}$, hence $\mathbb{E}_u[\langle U,Gv\rangle_H]=0$, $G\in\mathcal{S}$, $v\in H$, i.e. $U=0$. The proof of the closability of δ is similar. Finally, by standard arguments we consider processes of the form $\dot v=G\mathbf{1}_{[t,T]}$ where $G\in\mathcal{S}$ is $\mathcal{F}_t$-measurable, $t\in[0,T]$, for which we have $\mathbf{1}_{[t,T]}(s)D_sG=0$, $s\in[0,T]$, which shows from (1.3.11) that

$$
\delta(v)=G\int_0^T \mathbf{1}_{[t,T]}(s)(dX_s-\dot u(s)ds)=\int_0^T \dot v_s(dX_s-\dot u(s)ds),
$$

hence δ extends the Itô integral on all square-integrable $\mathcal{F}_t$-adapted processes, and (1.3.10) is proved. □

For all $t\in[0,T]$ we let $\chi_t(s)=\min(s,t)$, $s\in[0,T]$.

Definition 1.3.8. *Let*

$$
\nabla_t F:=\langle DF,\chi_t\rangle_H=\int_0^t \dot u(s)\dot D_sF ds,\qquad F\in\mathrm{Dom}(D).
$$

For F of the form (1.3.6) we have:

$$\nabla_t F = \int_0^t \dot{D}_s F \dot{u}(s) ds = -\sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \sum_{k=1}^{n} \frac{u(t \wedge T_k)}{\dot{u}(T_k)} \partial_k f_n(T_1, \ldots, T_n).$$

In the parametric case $u(t) = \lambda h(t)$, $t \in [0,T]$, $h \in \mathcal{P}$, we have

$$\nabla_t F = -\sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \sum_{k=1}^{n} \frac{h(t \wedge T_k)}{\dot{h}(T_k)} \partial_k f_n(T_1, \ldots, T_n), \tag{1.3.12}$$

which is independent of λ.

We close this section by introducing a Laplacian on the Poisson space.

Definition 1.3.9. *We define the Laplacian* Δ_t *by*

$$\Delta_t F = \nabla_t \nabla_t F, \qquad F \in \mathcal{S}.$$

The operator Δ_t is easily shown to be closable, i.e. for any sequence $(F_n)_{n\in\mathbb{N}}$ of random variables converging to 0 in $L^2_u(\Omega)$ and such that $(\Delta_t F_n)_{n\in\mathbb{N}}$ converges in $L^2_u(\Omega)$, we have

$$\lim_{n\to\infty} \Delta_t F_n = 0.$$

This allows one to define the domain of Δ_t, denoted by $\mathrm{Dom}(\Delta_t)$, as the set of functionals F for which there exists a sequence of cylindrical functionals $(F_n)_{n\in\mathbb{N}}$, which converges in $L^2_u(\Omega)$ to F and such that the sequence $(\Delta_t F_n)_{n\in\mathbb{N}}$ converges in $L^2_u(\Omega)$. We will say that a random variable F in $\mathrm{Dom}(\Delta_t)$ is Δ_t-superharmonic on Ω if

$$\Delta_t F(\omega) \le 0, \qquad \omega \in \Omega. \tag{1.3.13}$$

For example if $u(t) = \lambda t$, then for any $F \in \mathcal{S}$ of the form (1.3.6) we have

$$\begin{aligned} \Delta_t F &= -\sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \nabla_t \sum_{k=1}^{n} \partial_k f_n(T_1, \ldots, T_n)(t \wedge T_k) \\ &= \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \sum_{k,l=1}^{n} (t \wedge T_l)(t \wedge T_k) \partial_k \partial_l f_n(T_1, \ldots, T_n) \\ &\quad + \sum_{n=1}^{\infty} \mathbf{1}_{\{X_T=n\}} \sum_{k=1}^{n} \mathbf{1}_{[0,t]}(T_k) T_k \partial_k f_n(T_1, \ldots, T_n), \end{aligned}$$

which is independent of λ. Note that due to the presence of first order terms this Laplacian differs from the canonical Laplacian used in the Gaussian case, as a consequence the existence of associated positive superharmonic functions is not conditioned by a lower bound (such as $n \ge 3$) on the number of variables, see the examples in Section 1.3.5.

1.3.4 Stein estimators

Our aim is to construct a superefficient estimator $\tilde{\lambda}_T$ of λ of the form

$$\hat{\lambda}_T + \frac{\xi_T}{h(T)},$$

whose mean square error will be strictly smaller than the Cramer-Rao bound when $\xi_T \in L^2_u(\Omega)$ is suitably chosen, where $\hat{\lambda}_T = X_T/h(T)$ is the MLE of λ. In agreement with Proposition 1.3.4, this estimator will be biased and anticipating with respect to the Poisson filtration.

The next proposition is our main result on estimation of the intensity parameter $\lambda > 0$.

Proposition 1.3.10. *In the parametric case $u(t) = \lambda h(t)$, $t \in [0,T]$, for any $F \in \mathcal{S}$ of the form* (1.3.6) *the estimator*

$$\tilde{\lambda}_T := \hat{\lambda}_T - \frac{1}{h(T)} \frac{f_1'(T)}{f_1(T)} \mathbf{1}_{\{X_T=0\}} + \frac{1}{h(T)} \nabla_T \log F,$$

of λ, where $\nabla_T F$ is given in (1.3.12), *has risk*

$$\mathbb{E}_{\lambda h}[|\tilde{\lambda}_T - \lambda|^2] = \frac{\lambda}{h(T)} + \frac{1}{h^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-\lambda h(T)} + \frac{4}{h^2(T)} \mathbb{E}_{\lambda h} \left[\frac{\nabla_T \nabla_T \sqrt{F}}{\sqrt{F}} \right]. \tag{1.3.14}$$

The proof of Proposition 1.3.10 will rely on the following two lemmas. First in the next lemma we construct an unbiased risk estimator by applying Stein's integration by parts argument in which we replace (1.3.1) by the duality relation (1.3.9) between the gradient and divergence operators on the Poisson space.

Lemma 1.3.11. *Let $t \in [0,T]$. For any $\xi_t \in \mathrm{Dom}(D)$ we have*

$$\mathbb{E}_u \left[|X_t + \xi_t - u(t)|^2 \right] = u(t) + \|\xi_t\|^2_{L^2_u(\Omega)} + 2\, \mathbb{E}_u \left[\nabla_t \xi_t \right]. \tag{1.3.15}$$

Proof. We have

$$\begin{aligned}
\mathbb{E}_u \left[|X_t - u(t) + \xi_t|^2 \right] &= \mathbb{E}_u \left[|X_t - u(t)|^2 \right] + \|\xi_t\|^2_{L^2_u(\Omega)} + 2\, \mathbb{E}_u \left[(X_t - u(t))\xi_t \right] \\
&= u(t) + \|\xi_t\|^2_{L^2_u(\Omega)} + 2\, \mathbb{E}_u \left[(X_t - u(t))\xi_t \right].
\end{aligned}$$

We now use the duality relation (1.3.9) and Relation (1.3.10) to get

$$\begin{aligned}
\mathbb{E}_u \left[(X_t - u(t))\xi_t \right] &= \mathbb{E}_u \left[\delta(\chi_t)\xi_t \right] \\
&= \mathbb{E}_u \left[\langle \chi_t, D\xi_t \rangle_H \right] \\
&= \mathbb{E}_u \left[\int_0^t \dot{u}(s) \dot{D}_s \xi_t ds \right] \\
&= \mathbb{E}_u \left[\nabla_t \xi_t \right],
\end{aligned}$$

which yields (1.3.15). □

The proof of Proposition 1.3.10 is then a consequence of the following result which applies Lemma 1.3.11 to processes $(\xi_t)_{t\in[0,T]}$ of the form

$$\xi_t = c\frac{u(t)}{\dot u(T)}\mathbf{1}_{\{X_T=0\}} + \nabla_t \log F, \qquad t \in [0,T],$$

where c is chosen in such a way that $\xi_t \in \mathrm{Dom}(D)$, $t \in [0,T]$, and $F \in \mathrm{Dom}(D)$ is such that $F > 0$ and $\sqrt{F} \in \mathrm{Dom}(\Delta_t)$.

Lemma 1.3.12. *Let $t \in [0,T]$ and let $F \in \mathcal{S}$ of the form* (1.3.6) *such that $F > 0$, $\mathbb{P}$-a.s., $F \in \mathrm{Dom}(\Delta_t)$, and*

$$\partial_n f_n(t_1,\ldots,t_{n-1},T) = 0 \quad \textit{and} \quad \partial_k f_n(t_1,\ldots,t_n) = \partial_k f_{n+1}(t_1,\ldots,t_n,T),$$

$0 \le t_1 \le \cdots \le t_n \le T$, $1 \le k < n$, $n \ge 2$. *Let also*

$$\xi_t := -\frac{u(t)}{\dot u(T)}\frac{f_1'(T)}{f_1(T)}\mathbf{1}_{\{X_T=0\}} + \nabla_t \log F. \tag{1.3.16}$$

Then $\xi_t \in \mathcal{S} \subset \mathrm{Dom}(D)$ and

$$\mathbb{E}_u\left[|X_t+\xi_t-u(t)|^2\right] = u(t) + \frac{u^2(t)}{\dot u^2(T)}\left|\frac{f_1'(T)}{f_1(T)}\right|^2 e^{-u(T)} + 4\,\mathbb{E}_u\left[\frac{\Delta_t\sqrt{F}}{\sqrt{F}}\right], \qquad t \in [0,T]. \tag{1.3.17}$$

Proof. By construction we have $\xi_t \in \mathcal{S}$, $t \in [0,T]$, and from Lemma 1.3.11:

$$\begin{aligned}
\mathbb{E}_u\left[|X_t+\xi_t-u(t)|^2\right] &= u(t) + \|\xi_t\|^2_{L^2_u(\Omega)} + 2\,\mathbb{E}_u\left[\nabla_t\xi_t\right] \\
&= u(t) + \frac{u^2(t)}{\dot u^2(T)}\left|\frac{f_1'(T)}{f_1(T)}\right|^2 e^{-u(T)} + \mathbb{E}_u\left[\left|\frac{\nabla_t F}{F}\right|^2 + 2\nabla_t\nabla_t \log F\right] \\
&= u(t) + \frac{u^2(t)}{\dot u^2(T)}\left|\frac{f_1'(T)}{f_1(T)}\right|^2 e^{-u(T)} + \mathbb{E}_u\left[2\frac{\nabla_t\nabla_t F}{F} - \left|\frac{\nabla_t F}{F}\right|^2\right] \\
&= u(t) + \frac{u^2(t)}{\dot u^2(T)}\left|\frac{f_1'(T)}{f_1(T)}\right|^2 e^{-u(T)} + 4\,\mathbb{E}_u\left[\frac{\nabla_t\nabla_t\sqrt{F}}{\sqrt{F}}\right].
\end{aligned} \tag{1.3.18}$$

□

Proof of Proposition 1.3.10. Apply (1.3.16) and (1.3.17) above at $t = T$ with $u(t) = \lambda h(t)$.
□

As a consequence, the Δ_t-superharmonicity of F may imply the superefficiency of $X+\xi$. Note also that $X_t+\xi_t$ may not be positive and replacing ξ_t with $\max(X_t+\xi_t,0)$ will yield a lower risk since the intensity $\dot{u}$ is known to be positive.

We close this section with some additional remarks.

Remarks

a) Relation (1.3.18) established in proof of Lemma 1.3.12 shows that the Δ_t-superharmonicity of F implies

$$\begin{aligned} & \mathbb{E}_u[|X_t+\xi_t-u(t)|^2] \qquad\qquad (1.3.19)\\ & \leq u(t)+\frac{u^2(t)}{\dot{u}^2(T)}\left|\frac{f_1'(T)}{f_1(T)}\right|^2 e^{-u(T)}-\mathbb{E}_u\left[|\nabla_t \log F|^2\right], \qquad t\in[0,T], \end{aligned}$$

with equality in (1.3.19) when F is Δ_t-harmonic. Nevertheless the Δ_t-superharmonicity of $\sqrt{F}$ is a weaker condition.

b) Note that the risk of any non-adapted estimator ζ_t of $u(t)$ can be lowered by adapted projection, indeed we have

$$\begin{aligned} \mathbb{E}_u[|\,\mathbb{E}_u[\zeta_t \mid \mathcal{F}_t]-u(t)|^2] &= \mathbb{E}_u\left[|\zeta_t-u(t)|^2\right]-\mathbb{E}_u\left[|\,\mathbb{E}_u[\zeta_t\mid\mathcal{F}_t]-\zeta_t|^2\right]\\ &< \mathbb{E}_u\left[|\zeta_t-u_t|^2\right], \qquad (1.3.20) \end{aligned}$$

for all $u\in\mathcal{P}$, and in particular

$$\mathbb{E}_u[X_t+\zeta_t\mid\mathcal{F}_t]=X_t-\frac{u(t)}{\dot{u}(T)}\frac{f_1'(T)}{f_1(T)}\mathbf{1}_{\{X_t=0\}}e^{-(u(T)-u(t))}+\mathbb{E}_u[\nabla_t\log F\mid\mathcal{F}_t],$$

$t\in[0,T]$, which is however dependent on the intensity u.

c) Both estimators $X_t+\xi_t$ and $X_t+\mathbb{E}_u[\xi_t\mid\mathcal{F}_t]$ have bias

$$b(t)=\mathbb{E}_u[X_t+\xi_t-u(t)]=\mathbb{E}_u[\xi_t], \qquad t\in[0,T],$$

which, using the relation

$$\left|\frac{\nabla_t F}{F}\right|^2=2\frac{\nabla_t\nabla_t F}{F}-\frac{4}{\sqrt{F}}\nabla_t\nabla_t\sqrt{F},$$

can be bounded as follows:

$$\begin{aligned} b^2(t) &= |\,\mathbb{E}_u[\xi_t]|^2\\ &\leq \mathbb{E}_u\left[|\xi_t|^2\right]\\ &= 2\,\mathbb{E}_u\left[\frac{\nabla_t\nabla_t F}{F}\right]+\frac{u^2(t)}{\dot{u}^2(T)}\left|\frac{f_1'(T)}{f_1(T)}\right|^2 e^{-u(T)}-4\,\mathbb{E}_u\left[\frac{\nabla_t\nabla_t\sqrt{F}}{\sqrt{F}}\right], \end{aligned}$$

hence when F is Δ_t-superharmonic we have

$$b^2(t) \le \frac{u^2(t)}{\dot{u}^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-u(T)} - 4\,\mathbb{E}_u \left[\frac{\nabla_t \nabla_t \sqrt{F}}{\sqrt{F}} \right], \qquad t \in [0,T].$$

d) By integration over $[0,T]$, Proposition 1.3.4 immediately yields, for any unbiased and adapted estimator ζ of $u \in \mathcal{P}$:

$$\mathbb{E}_u \left[\int_0^T |\zeta_t - u(t)|^2 dt \right] \ge \|u\|_{L^1([0,T])}, \qquad u \in \mathcal{P}, \tag{1.3.21}$$

where the lower bound $\|u\|_{L^1([0,T])}$ is attained by $\zeta = X$. This bound can be used to derive a nonparametric estimation result for the process $(u(t))_{t\in[0,T]}$.

On the other hand, formal maximization of the Girsanov density $\Lambda(u)$ gives

$$\frac{d}{d\varepsilon}\Lambda(u+\varepsilon v)_{|\varepsilon=0} = \Lambda(u) \int_0^T \frac{\dot{v}(s)}{\dot{u}(s)}(dX_s - du(s)) = 0$$

for all $v \in H$, i.e. $\hat{u}_t = X_t$. Hence the canonical process $\hat{u} = (X_t)_{t\in[0,T]}$ can be considered as an unbiased maximum likelihood estimator of its own intensity $(u(t))_{t\in[0,T]}$ under $\mathbb{P}_u$, which is efficient in the sense that it attains the Cramer-Rao bound

$$\mathbb{E}_u \left[\|X - u\|^2_{L^2([0,T])} \right] = \|u\|_{L^1([0,T])}. \tag{1.3.22}$$

Given $(X_t^{(1)})_{t\in[0,T]}, \ldots, (X_t^{(N)})_{t\in[0,T]}$, N independent samples of $(X_t)_{t\in[0,T]}$, the process

$$\bar{X}_t = \frac{1}{N}\left(X_t^{(1)} + \cdots + X_t^{(N)} \right)$$

is a point process with intensity u under $\mathbb{P}_u$, which is consistent as T tends to 0 and as N goes to infinity, since by independence we have

$$\mathbb{E}_u \left[\|\bar{X} - u\|^2_{L^2([0,T])} \right] = \frac{1}{N^2}\,\mathbb{E}_u \left[\sum_{i=1}^N \int_0^T |X_t^{(i)} - u(t)|^2 dt \right] = \frac{1}{N}\int_0^T u(t)dt.$$

Similarly to the above, integration over $[0,T]$ and Lemma 1.3.12 show that for ξ_t defined as in (1.3.16), $t \in [0,T]$, $\xi_t \in \mathcal{S}$, $t \in [0,T]$, and

$$\begin{aligned} \mathbb{E}_u \left[\|X + \xi - u\|^2_{L^2([0,T])} \right] &= \|u\|_{L^1([0,T])} + \frac{\|u\|^2_{L^2([0,T])}}{\dot{u}^2(T)} \left| \frac{f_1'(T)}{f_1(T)} \right|^2 e^{-u(T)} \\ &\quad + 4\,\mathbb{E}_u \left[\frac{1}{\sqrt{F}} \int_0^T \Delta_t \sqrt{F} dt \right], \end{aligned}$$

hence the Δ_t-superharmonicity of F, $t \in [0, T]$, may imply the superefficiency of $(X_t + \xi_t)_{t\in[0,T]}$. Note however that in the general non-parametric case, the estimator $(X_t + \xi_t)_{t\in[0,T]}$ of u is dependent on u.

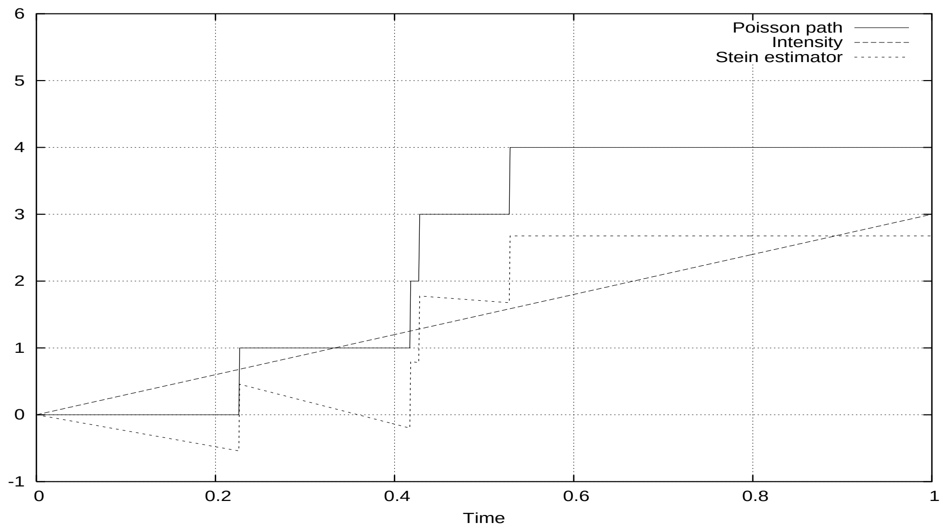

Figure 1.5: $u(t) = 3t,\ t \in [0, T];\ N = 5.$

Figure 1.5 represents a sample path of the process $X_t + \xi_t$, $t \in [0, T]$ when $u(t) = \lambda t$, $\lambda = 3$.

1.3.5 Examples

In this section we present some examples of estimators satisfying the hypotheses of the previous sections, and we test their superefficiency. In the parametric case $u(t) = \lambda h(t)$, $t \in [0, T]$, the percentage gain of an estimator $\tilde{\lambda}_T$ of λ over the MLE $\hat{\lambda}_T = X_T/h(T)$ is defined as

$$100 \times \frac{\mathbb{E}_u[|\hat{\lambda}_T - \lambda|^2] - \mathbb{E}_u[|\tilde{\lambda}_T - \lambda|^2]}{\mathbb{E}_u[|\hat{\lambda}_T - \lambda|^2]} = 100 \times \frac{\lambda/h(T) - \mathbb{E}_u[|\tilde{\lambda}_T - \lambda|^2]}{\lambda/h(T)}.$$

In the sequel we assume that $u(t) = \lambda t$, $t \in [0, T]$, hence (1.3.14) reads

$$\mathbb{E}_u[|\tilde{\lambda}_T - \lambda|^2] = \frac{\lambda}{T} + \left|\frac{f_1'(T)}{f_1(T)}\right|^2 e^{-u(T)} + \frac{4}{T^2}\,\mathbb{E}_u\left[\frac{\nabla_T\nabla_T\sqrt{F}}{\sqrt{F}}\right]$$

and $\tilde{\lambda}_T$ is superefficient, i.e. its gain is positive, provided $\sqrt{F}$ is Δ_T-superharmonic and $f_1'(T)/f_1(T)$ vanishes or is small enough.

The positive Δ_t-superharmonic functionals we consider are of the form

$$\sqrt{F} = \int_0^T g_{N_t}(t)dN_t = \sum_{n=1}^\infty \mathbf{1}_{\{X_T=n\}} \sum_{k=1}^n g_k(T_k),$$

where $g_k : [0,T] \to (0,\infty)$, $k \geq 1$, and

$$g_1(t_1) + \cdots + g_n(t_n) \geq 0, \qquad 0 \leq t_1 \leq \cdots \leq t_n \leq T, \quad n \geq 1. \tag{1.3.23}$$

Then, ξ_t defined from (1.3.16) as

$$\begin{aligned}
\xi_t &:= -2t\frac{g_1'(T)}{g_1(T)}\mathbf{1}_{\{X_T=0\}} + \nabla_t \log F \\
&= -2t\frac{g_1'(T)}{g_1(T)}\mathbf{1}_{\{X_T=0\}} - 2\sum_{n=1}^\infty \mathbf{1}_{\{X_T=n\}} \sum_{k=1}^n \frac{(t\wedge T_k)g_k'(T_k)}{g_1(T_1)+\cdots+g_n(T_n)} \\
&= -2t\frac{g_1'(T)}{g_1(T)}\mathbf{1}_{\{X_T=0\}} - \frac{2}{\sqrt{F}}\int_0^T (t\wedge s)g_{N_s}'(s)dN_s,
\end{aligned}$$

belongs to $\mathrm{Dom}(D)$ provided

$$g_k(T) = 0, \quad \text{and} \quad g_k'(T) = 0, \qquad k \geq 2, \tag{1.3.24}$$

and for the condition $\Delta_T\sqrt{F} \leq 0$ to hold it suffices that

$$g_k'(x) + xg_k''(x) \leq 0, \qquad x \in [0,T], \quad k \geq 1. \tag{1.3.25}$$

a) Let $g_1(x) = T(1+\beta) - x$ and $g_k = 0$, $k \geq 2$, i.e.

$$F = \mathbf{1}_{\{X_T\geq 1\}}\left(\beta T + T - T_1\right)^2,$$

with $\beta > 0$. We have from (1.3.16):

$$\xi_t = \mathbf{1}_{\{X_T=0\}}\frac{2}{\beta T}t + \mathbf{1}_{\{X_T\geq 1\}}\frac{2}{T+\beta T - T_1}(t\wedge T_1),$$

and

$$\nabla_t\nabla_t\sqrt{F} = -T_1\mathbf{1}_{[0,t]}(T_1)\mathbf{1}_{\{X_T\geq 1\}} \leq 0,$$

hence

$$\begin{aligned}
\mathbb{E}_{\lambda h}[|\tilde{\lambda}_T - \lambda|^2] &= \frac{\lambda}{T} + \frac{4}{\beta^2T^2}e^{-\lambda T} - \frac{4}{T^2}\,\mathbb{E}_{\lambda h}\left[\frac{T_1}{\beta T + T - T_1}\right] \\
&= \frac{\lambda}{T} + \frac{4}{\beta^2T^2}e^{-\lambda T} - \frac{4\lambda}{T^2}\int_0^T \frac{x}{\beta T + T - x}e^{-\lambda x}dx.
\end{aligned}$$

The gain of this estimator is equal to the function of λT:

$$\begin{aligned}\frac{4}{T}\int_0^T \frac{x}{\beta T+T-x}e^{-\lambda x}dx-\frac{4}{\lambda T\beta^2}e^{-\lambda T} &= 4\int_0^1 \frac{x}{1+\beta-x}e^{-\lambda T x}dx-\frac{4}{\lambda T\beta^2}e^{-\lambda T}\\ &\geq 4e^{-\lambda T}\left(\int_0^1 \frac{1-x}{1+\beta}e^{\lambda T x}dx-\frac{1}{\lambda T\beta^2}\right)\\ &= 4\frac{e^{-\lambda T}}{\beta\lambda T}\left(\frac{((\lambda T)^2-\lambda T+1)e^{\lambda T}-1}{(1+1/\beta)\lambda T}-\frac{1}{\beta}\right),\end{aligned}$$

which is strictly positive (i.e. $\tilde{\lambda}_T$ is superefficient) provided $\beta \geq 2\lambda^{-1}T^{-1}$. Figure 1.6 represents the gain of $\tilde{\lambda}_T$ as a function of β.

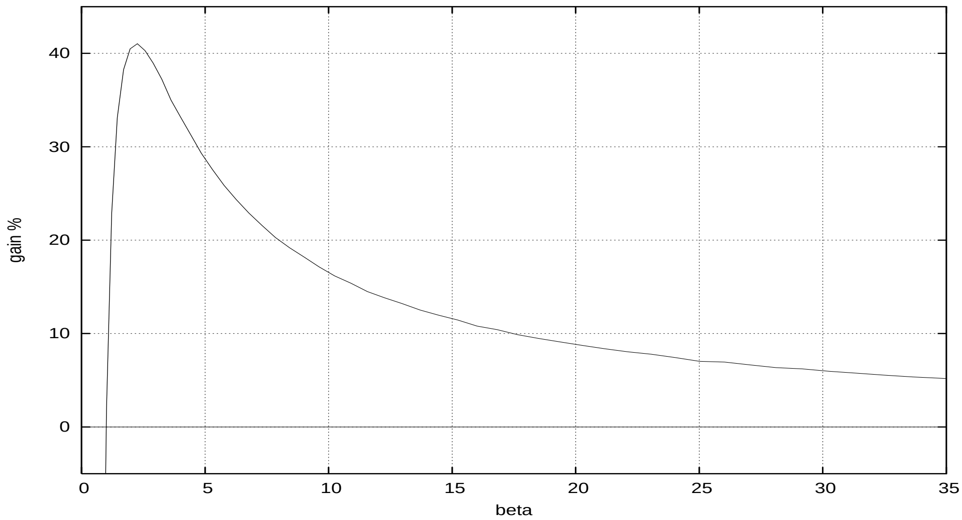

Figure 1.6: Gain as a function of β with $\lambda = 1$ and $N = 1$.

b) When $k \geq 2$, conditions (1.3.24) and (1.3.25) are not compatible and as a consequence, superefficiency of $\tilde{\lambda}_T$ will be dependent on the value of λ. We take

$$g_1(x) = C, \qquad g_k(x) = -(-\log((c+x)/(c+T)))^{\alpha_k}, \qquad 2 \leq k \leq N,$$

$g_k = 0$, $k > N$, with $C \geq \sum_{k=1}^N(-\log(c/(c+T)))^{\alpha_k}$ and $\alpha_k > 1$, $2 \leq k \leq N$. In this case, $\Delta_T g_k$ is not everywhere negative as shown in Figure 1.7, with $\alpha_2 = 2$, $k = 2$, $T = 1$, and $c = 0.01$, but this suffices to achieve superefficiency for most values of λ, see below.

Figure 1.8 represents the gain of $\tilde{\lambda}_T$ as a function of λ, with 10^6 samples, $\alpha_2 = \alpha_3 = 2$, $N = 3$, and $C = g_2(0) + g_3(0)$.

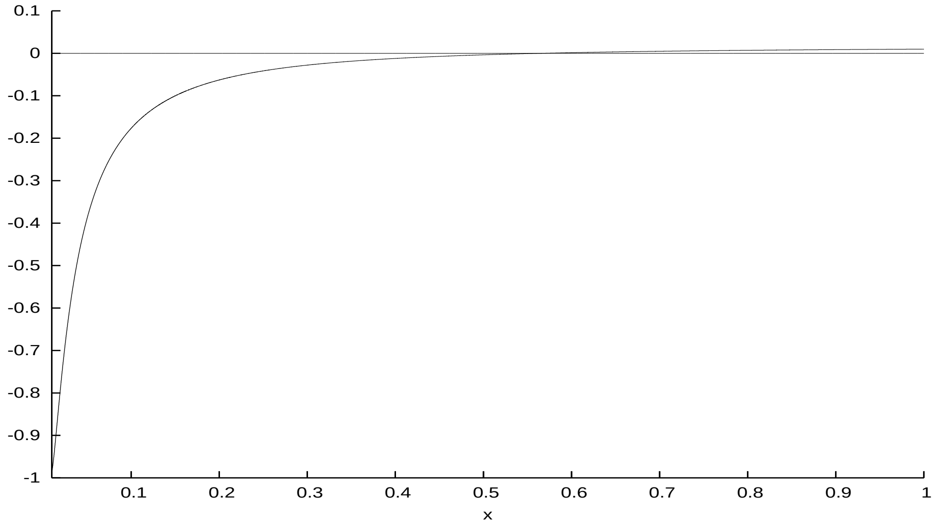

Figure 1.7: Graph of $\Delta_T g_2$

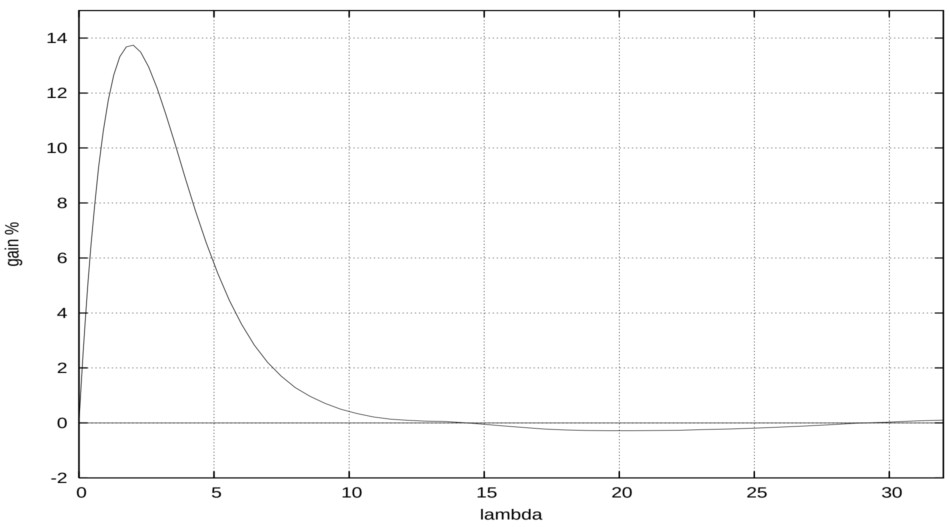

Figure 1.8: Gain as a function of λ.

Chapitre 2

Estimation bayésienne pour les canaux de Poisson

Ce chapitre correspond à la prépublication [118].

2.1 Introduction

Recently in [138], infinite-dimensional methods have been used to derive a new expression of the conditional mean estimator for infinite-dimensional additive Gaussian channels. More precisely the conditional mean estimator is obtained as the Malliavin derivative of the logarithm of the likelihood ratio. This relationship has been recently applied in [133] in order to link the quadratic risk of the conditional mean estimator to the Monge-Kantorovitch measure transportation theory. Let us give some details about these results.

In the general framework of *additive Gaussian channel*, an observed signal Y is decomposed into the sum of an *input* signal X plus an independent Gaussian noise W as

$$Y = \rho X + W, \tag{2.1.1}$$

where ρ is the "signal to noise ratio". In this context the signals "lie" in an *abstract Wiener space* (W, H, μ_W) (in particular the input X is an H-valued random variable). This setting contains the case of an observed continuous-time stochastic process $(Y_t)_{t\in[0,T]}$ related to an input stochastic process $(X_t)_{t\in[0,T]}$ (with values into the Hilbert space $H := L^2([0,T])$) by the following stochastic differential equation,

$$dY_t = \rho X_t \, dt + dW_t, \; t \in [0,T] \tag{2.1.2}$$

where $(W_t)_{t\in[0,T]}$ is a real valued standard Brownian motion independent of $(X_t)_{t\in[0,T]}$ and ρ denotes the "signal to noise ratio". In [138, Prop 4.1], Zakai has shown that

$$\mathbb{E}[X|\mathcal{Y}] = \frac{1}{\rho} \nabla \log l(Y), \tag{2.1.3}$$

where $\mathcal{Y}$ denotes the sigma field generated by Y, ∇ denotes the Malliavin gradient which is a infinite-dimensional counterpart of the usual derivative on $\mathbb{R}^n$ and l is the likelihood ratio associated to model (2.1.1) that is,

$$l := \frac{d\mu_Y}{d\mu_W}.$$

Relation (2.1.3) entails the following result ([138, Proposition 5.1]),

$$\frac{dI(X;Y)}{d\rho} = \rho\ \mathbb{E}\left[\|X - \mathbb{E}[X|\mathcal{Y}]\|_H^2\right],$$

where $I(X;Y)$ denotes the mutual information between X and Y, defined as,

$$I(X;Y) := \int_{H\times W} \log \frac{d\mu_{X,Y}}{d(\mu_X \times \mu_Y)} \mu_{X,Y}(dx, dy).$$

Note that (2.1.2) is an infinite-dimensional counterpart of a well known result for finite-dimensional additive Gaussian channels described as,

$$Y = \rho X + N, \tag{2.1.4}$$

where X is an $\mathbb{R}^n$-valued input signal and N is a standard Gaussian random variable on $\mathbb{R}^n$ independent of X. Let μ_X (respectively μ_Y) be the distribution of X (respectively Y) and $\mathcal{Y}$ be the sigma field generated by Y.
Using the fact that the law of Y given X is absolutely continuous with respect to the law of a standard Gaussian random vector on $\mathbb{R}^n$ denoted by Z, the unconditional law of Y is absolutely continuous with respect to the distribution of Z with density $\tilde{m}$,

$$\tilde{m}(y) = \int_H \exp\left(-\frac{\rho^2 x^2 - 2\rho y x}{2}\right)\ \mu_X(dx),\ \text{and}$$

$$\mathbb{E}[X|\mathcal{Y}] = \frac{1}{\rho}\nabla \log \tilde{m}(Y). \tag{2.1.5}$$

Another application of (2.1.3) is given by Üstünel in [133] since he has related the quadratic risk of the estimation to the Monge-Kantorovitch measure transportation problem.

Note that the main results of this Note are equation (2.4.8) in Proposition 2.4.6 and equation (2.4.9) in Corollary 2.4.8 which provide a method to compute the conditional mean estimator. In particular the nonlinear filter of the input can be numerically approximated by evaluating the unconditional density of the output thanks to a Monte-Carlo scheme (see Remark 2 of [138]).

We proceed as follows. First in Section 2.2 we extend Relation (2.1.5) to the setting of classical Poisson channels where ∇ in (2.1.5) will be replaced by a difference operator. Secondly, we will use infinite-dimensional stochastic analysis methods presented in

Section 2.3 to derive in Section 2.4 an equivalent of (2.1.3) for infinite-dimensional Poisson channels using a Malliavin gradient for Poisson processes. Finally in Section 2.5, we generalize the results obtained in the preceding sections to a class of normal martingales which contains the continuous time Poisson channel, the Gaussian one and a mixture of the both. Furthermore this class includes some martingales with jumps and non-independent increments.

2.2 Integer-valued Poisson channel

Let us briefly describe the integer-valued Poisson channel (see [134] for a survey on Poisson channels).

Poisson channels are different from Gaussian channels in the sense that the observed signal cannot be expressed as the sum of the input signal plus some additional noise; which means that it cannot be expressed in an "additive" way like in (2.1.1). Consider a positive input signal X with distribution μ_X. We assume the output Y is a Poisson random variable on $\mathbb{N}$ with intensity $\alpha X + \lambda_0$,

$$Y \sim \mathcal{P}(\alpha X + \lambda_0).$$

This setting is used for example in photo-detection problems where a photo-sensitive device (*e.g.* a p-i-n diode) is modelled by a Poisson channel. In this setting λ_0 is a residual current in the device called the "dark current noise" and α is some scale parameter. Note that contrary to the Gaussian channel λ_0 and α cannot be replaced by a single coefficient, the "signal to noise ratio".

Let μ_{λ_0} be the distribution of a Poisson random variable on $\mathbb{N}$ with intensity λ_0. Finally assume that the joint distribution of (X, Y) is absolutely continuous with respect to measure $\mu_X \times \mu_{\lambda_0}$. This last condition expresses the independence between the input signal and the noise introduced by the channel. Consequently, the distribution of Y given X is absolutely continuous with respect to μ_{λ_0} with,

$$\begin{aligned} &\frac{d\mu_{Y|X=x}}{d\mu_{\lambda_0}}(y) \\ = \;& \exp(-\alpha x)\left(\frac{\alpha x + \lambda_0}{\lambda_0}\right)^y, \; x \in \mathbb{R}_+, \; y \in \mathbb{N}, \end{aligned}$$

and the law of Y is absolutely continuous with respect to μ_{λ_0} with density m,

$$m(y) := \int_0^{+\infty} \frac{d\mu_{Y|X=x}}{d\mu_{\lambda_0}}(y)\, \mu_X(dx), \; y \in \mathbb{N}. \tag{2.2.1}$$

Now we can state the following lemma which will be extended in Section 2.4 as Proposition 2.4.6 and Corollary 2.4.8.

Lemma 2.2.1. *If the Bayesian risk* $\mathbb{E}\left[|X - \mathbb{E}[X|\mathcal{Y}]|^2\right]$ *is finite then*

$$\mathbb{E}[X|\mathcal{Y}] = \frac{\lambda_0}{\alpha}\,\frac{m(Y+1) - m(Y)}{m(Y)}. \tag{2.2.2}$$

Proof. Let y in $\mathbb{N}$.

$$\begin{aligned} & m(y+1) - m(y) \\ = \; & \frac{\alpha}{\lambda_0} \int_0^{+\infty} x \; \exp(-\alpha x) \left(\frac{\lambda_0 + \alpha x}{\lambda_0} \right)^y \mu_X(dx) \\ \overset{(*)}{=} \; & \frac{\alpha}{\lambda_0} \, m(y) \int_0^{+\infty} x \, \frac{d\mu_{X|Y=y}}{d\mu_X}(x) \, \mu_X(dx), \; \mathbb{P} - a.s. \\ = \; & \frac{\alpha}{\lambda_0} \, m(y) \int_0^{+\infty} x \; \mu_{X|Y=y}(dx) \\ = \; & \frac{\alpha}{\lambda_0} \, m(y) \; \mathbb{E}[X|Y=y] \end{aligned}$$

Equality $(*)$ is justified by a relation of the form *iii*) of Proposition 2.4.1. □

Remark 2.2.2. *The nonlinear filter of X given in (2.2.3) can be numerically approximated thanks to a Monte-Carlo scheme (see Remark 2.4.9).*

Remark 2.2.3. *The conditional distributions used in Lemma 2.2.1 are well defined in this context, one can refer to Propositions 2.4.1 and 2.4.4 for more details.*

We give a short and straightforward example by assuming X is a random variable with values in $\{a, b\}$ where a and b are positive with $\mathbb{P}(X = a) = \mathbb{P}(X = b) = 1/2$. Let Y to be a Poisson random variable with intensity $1 + X$. The density of $\mu_{Y|X}$ with respect to μ_1 is given by

$$\frac{d\mu_{Y|X=x}}{d\mu_1}(y) = e^{-x}(x+1)^y, \; (x, y) \in \mathbb{R}_+ \times \mathbb{N},$$

and one obtains,

$$\mathbb{E}[X|\mathcal{Y}] = \frac{a \, e^{-a} \, (1+a)^y + b \, e^{-b} \, (1+b)^y}{e^{-a} \, (1+a)^y + e^{-b} \, (1+b)^y},$$

which leads to

$$\mathbb{E}[X|\mathcal{Y}] = \frac{m(Y+1) - m(Y)}{m(Y)}. \tag{2.2.3}$$

To obtain results for more general Poisson channels we have first to recall some elements of analysis on the Poisson space.

2.3 Analysis on the Poisson space

In this Section we introduce some elements of analysis on the Poisson space in a general framework then, we will describe these elements in a concrete example.
Let $(S, \mathcal{B}(S), \nu)$ a measure space where ν is an intensity atomless σ-finite measure. Define the Poisson space Ω_S as

$$\Omega_S = \left\{ y = \sum_{k=0}^{n} \delta_{z_k}, \; n \in \bar{\mathbb{N}}, \; z_k \in S, \; 1 \le k \le n \right\},$$

with $\bar{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$ and let for $\sum_{k=0}^{n} \delta_{z_k}$,

$$\mathcal{C} = \left(\sum_{k=0}^{n} \delta_{z_k} \right) := \{z_1, \ldots, z_n\}. \tag{2.3.1}$$

Define the canonical process $(N_A)_{A \in \mathcal{B}(S)}$ on Ω_S as

$$N_A(y) := y(A), \; y \in \Omega_S.$$

We define the σ-field $\mathcal{F_S}$ on Ω_S with $\mathcal{F_S} = \sigma(\{y \mapsto y(B), \; B \in \mathcal{B}(S)\})$.
There exists a probability measure $\mathbb{P}_S$ on $(\Omega_S, \mathcal{F}_S)$ called the Poisson measure such that,

- $\forall B \in \mathcal{B}(S), \; \forall n \in \mathbb{N}$,
  $\mathbb{P}_S(\{y \,|\, y(B) = n\}) = \exp(-\nu(B)) \, \dfrac{\nu(B)^n}{n!}$
- For disjoints subsets $(B_1, \cdots, B_n)$ in $\mathcal{B}(S)$, $y(B_1), \; \ldots, y(B_n)$ are $\mathbb{P}_S$-independent.

Under $\mathbb{P}_S$ the canonical process $(N_A)_{A \in \mathcal{B}(S)}$ is a Poisson process with intensity ν.
Define $\mathcal{M}(S)$ the set of non-negative measure on $(S, \mathcal{B})$. Let H_S be the space

$$H_S = \left\{ \omega \in \mathcal{M}(S), \; \omega(\cdot) = \int_{\cdot} h \, d\nu, \; h \in L^2_+(S, \nu) \right\},$$

where $L^2_+(S, \nu)$ denotes the set of positive function of $L^2(S, \nu)$.
H_S is equipped with an inner product $\langle \cdot, \cdot \rangle_{H_S}$ given by

$$\langle \omega_1, \omega_2 \rangle_{H_S} = \langle h_1, h_2 \rangle_{L^2(S, d\nu)}, \; \omega_1 \in H_S, \; \omega_2 \in H_S.$$

The setting describe above contains the canonical Poisson space as a particular case.
Let $\Omega_{[0,T]}$ the canonical Poisson space on $(S, \mathcal{B}(S), \nu) = ([0,T], \mathcal{B}([0,T]), ds)$,

$$\Omega_{[0,T]} = \left\{ y = \sum_{k=0}^{n} \delta_{t_k}, \; n \in \bar{\mathbb{N}}, \; 0 \le t_1 < \ldots < t_n \le T \right\}.$$

In this case $\mathcal{C}(y)$ given by (2.3.1) is the set of the jump times of the path y and under $\mathbb{P}_{[0,T]}$, $(N_{[0,t]})_{t \in [0,T]}$ is a Poisson process with intensity dt, that is, the stochastic process $(N_t - t)_{t \in [0,T]}$ is a $\mathbb{P}_{[0,T]}$-martingale.
In this case $H_{[0,T]}$ can be defined in a more tractable way by,

$$H_{[0,T]} = \left\{ \quad v : [0,T] \to \mathbb{R}, \; v(t) = \int_0^t \dot{v}_s \, ds, \; \dot{v} \in L^2([0,T]) \right\},$$

equipped with

$$\langle h_1, h_2 \rangle_{H_{[0,T]}} := \langle \dot{h}_1, \dot{h}_2 \rangle_{L^2([0,T])}, \; h_1, h_2 \in H_{[0,T]}.$$

The Malliavin operator ∇^d we introduce will be of interest in Section 2.4.
Let $L^0(\Omega_S, \mathcal{F}_S, \mathbb{P}_S)$ be the space of measurable mapping from $(\Omega_S, \mathcal{F}_S, \mathbb{P}_S)$ to $\mathbb{R}$.
Define first the operator D^d by,

$$\begin{array}{rcl} L^0(\Omega_S, \mathcal{F}_S, \mathbb{P}_S) & \to & L^0(\Omega_S \times S, \mathcal{F}_S \otimes \mathcal{B}(S), \mathbb{P}_S \otimes \nu) \\ F & \mapsto & D^d_z F(y) := F(y + \delta_z) - F(y). \end{array}$$

Technical justifications about the measurability of the previous map can be found in [137] and references therein. This allows us to define the operator ∇^d.

Definition 2.3.1. *For $F : \Omega \to \mathbb{R}$ we define ∇F as the H_S-valued random variable*

$$\nabla^d_A F := \int_A D^d_z F \, \nu(dz), \; A \in \mathcal{B}(S).$$

Finally in the case of the classical Poisson space the Malliavin derivative ∇ can be expressed in a different way, for $F : \Omega_{[0,T]} \to \mathbb{R}$, ∇F is a $H_{[0,T]}$-valued random variable and

$$\nabla^d_{[0,t]} F := \int_0^t D^d_s F \, ds, \; t \in [0, T].$$

2.4 Conditional mean estimators for Poisson channels

We introduce in this section the Bayesian framework and we compute the conditional mean estimator in the setting of Poisson point process. Let X be an input signal with values in a space $(H, \sigma(H))$ with distribution μ_X. Consider $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space and assume the output Y lies in Ω. Until the end of this paper we will denote by $\mathcal{Y}$ the σ-field generated by Y. We make the following assumptions,

(H1) The "noise" and the "signal" are independent, *i.e.* the law of the pair $(output, input) = (Y, X)$ is absolutely continuous with respect to $\mathbb{P} \times \mu_X$.

(H2) For all x in H, $\mu_{Y|X=x}$ (the distribution of Y given $X = x$) is absolutely continuous with respect to $\mathbb{P}$ and we denote by L the corresponding Radon-Nikodym density.

(H3) L is $(\sigma(H) \otimes \mathcal{F})$-measurable.

(H4) The Bayesian risk $\mathbb{E}\left[\|X - \mathbb{E}[X|\mathcal{Y}]\|_H^2\right]$ with respect to μ_X is finite.

Then, the following function

$$\begin{array}{l} H \times \mathcal{F} \to [0, 1] \\ (x, B) \mapsto \mu_{Y|X=x}(B) \end{array}$$

is a transition probability in the sense of [80, Definition III-2-1 p. 69]. Moreover from [80, Proposition III-2-1 p. 69-70] there exists a probability measure μ on $(H \times \Omega, \sigma(H) \otimes \mathcal{F})$ such that,

$$\mu(A \times B) = \int_A \mu_{Y|X=x}(B) \, \mu_X(dx), \; A \in \sigma(H), \; B \in \mathcal{F},$$

and μ is the joint distribution of (X, Y). Denote by M the marginal distribution of μ on H defined by,

$$M(B) := \mu(H \times B),\ B \in \mathcal{F}. \tag{2.4.1}$$

Proposition 2.4.1 is mainly devoted to show the existence of the following transition probability

$$\begin{aligned} &\Omega \times \sigma(H) \to [0,1] \\ &(y, A) \mapsto \mu_{X|Y=y}(A) \end{aligned} \tag{2.4.2}$$

and that the couple $(M, (\mu_{X|Y}(\cdot, y))_{y\in\Omega})$ allows us to recover μ as

$$\mu(A \times B) = \int_B \mu_{X|Y=y}(A)\, M(dy),\ A \in \sigma(H),\ B \in \mathcal{F}. \tag{2.4.3}$$

Proposition 2.4.1. *If* (H1), (H2), *and* (H3) *are satisfied then*

i) μ is absolutely continuous with respect to $\mu_X \times \mathbb{P}$ and the corresponding Girsanov-Radon-Nikodym density is L.

ii) M is absolutely continuous with respect to $\mathbb{P}$ with m as Radon-Nikodym density.

iii) For M almost all y in Ω, $\mu_{X|Y=y}$ is absolutely continuous with respect to μ_X and for y such that $m(y) \neq 0$, the Radon-Nikodym density is given by,

$$\frac{d\mu_{X|Y=y}}{d\mu_X}(x) = \frac{L(y,x)}{m(y)}.$$

iv) For a $(\sigma(H) \times \mathcal{F})$-measurable function $f : H \times \Omega \to \mathbb{R}$,

$$\int_\Omega \int_H f(x,y)\mu_{X|Y=y}(dx)\, M(dy) = \int_H \int_\Omega f(x,y)\mu_{Y|X=x}(dy)\, \mu_X(dx).$$

Proof. See [41, Theorem 1.8]. □

As we have defined in Section 2.2 by equation (2.2.1) the unconditional density of the observed signal for Poison channels on $\mathbb{N}$, we introduce the unconditional Radon-Nikodym density of Y with respect to $\mathbb{P}$.

Definition 2.4.2. *Let m be the unconditional Radon-Nikodym density of Y with respect to $\mathbb{P}$,*

$$m(y) = \int_{H_S} L(y,x)\, \mu_X(dx),\ y \in \Omega. \tag{2.4.4}$$

Remark 2.4.3. *With the definition of the posterior distribution given in iii), (2.4.2) is a transition probability in the sense of Definition III-2-1 in [80] p. 69.*

Now we will make use of the general Bayesian framework described above.

Let $(S, \mathcal{B}(S), \nu)$ and H_S as in Section 2.3 and suppose Y is a Poisson point process with intensity measure denoted "$1 + \alpha X$". That is for every x in H_S, $\mu_{Y|X=x}$ is a Poisson point process with intensity measure ν_x absolutely continuous with respect to the measure ν with $1 + \alpha \dot{x}$ as Radon-Nikodym density where $\dot{x}$ denotes the Radon-Nikodym density of x with respect to ν.
Denote by $(\Omega_S, \mathcal{F}_S, \mathbb{P}_S)$ the Poisson space (see Section 2.3). Assume hypothesis (H1) is satisfied. By Girsanov theorem, $\mu_{Y|X}$ the conditional probability on Ω given X is absolutely continuous with respect to $\mathbb{P}_S$ and the Girsanov-Radon-Nikodym density denoted L is given by

$$\begin{aligned} L(y,x) &:= \frac{d\mu_{Y|X=x}}{d\mathbb{P}_S}(y) \qquad (2.4.5) \\ &= \exp\left(-\alpha \int_S \dot{x}_z \nu(dz)\right) \prod_{k=0}^{y(S)} (1 + \alpha\, \dot{x}(z_k)), \end{aligned}$$

where $y = \sum_{k=1}^{n} \delta_{z_k}$. So hypothesis (H2) is satisfied. Finally assume (H4) holds. We recall a result about Bayesian estimator under quadratic loss.

Proposition 2.4.4. *The Bayesian estimator* $(\mathcal{B}_A)_{A \in \mathcal{B}(S)}$ *is*

$$\mathcal{B}_A(Y) = \mathbb{E}[X(A)|\mathcal{Y}] = \int_{H_S} x(A)\, \mu_{X|Y}(dx),\ M - a.e. \qquad (2.4.6)$$

Remark 2.4.5. *Note that the expression (2.4.6) is theoretical and cannot be used in practice. In contradistinction, relation (2.4.8) obtained below enables a numerical approximation of the Bayesian estimator as mentioned in Remark 2.4.9.*

In fact it is more tractable to estimate the densities rather than the intensity measures. So we denote by $\dot{X}$ *the* $L^2(S, d\nu)$ *valued random variable associated to* X*. For* $z \in S$ *(2.4.6) can be rewritten as*

$$\dot{\mathcal{B}}_z(y) = \mathbb{E}[\dot{X}_z|Y = y] = \int_{H_S} \dot{x}_z\, \mu_{X|Y}(dx, y),\ M - a.e. \qquad (2.4.7)$$

We can state the main result of this Note. It allows us to express the Bayesian estimator of the input as a discrete logarithmic Malliavin gradient of the likelihood ratio m.

Proposition 2.4.6.

$$\mathbb{E}[X_A|\mathcal{Y}] = \frac{\nabla_A^d m(Y)}{\alpha\, m(Y)},\ A \in \mathcal{B}(S). \qquad (2.4.8)$$

Proof. Let z in S and y in Ω_S, recall $\mathcal{C} = \left(\sum_{k=1}^{n} \delta_{z_k}\right) = \{z_1, \ldots, z_n\}$.

$$D_z^d m(y) = m(y + \delta_z) - m(y)$$

$$= \alpha \int_H L(y,x)\, \mathbf{1}_{z \notin JT(y)}\, \dot{x}_z\, \mu_X(dx).$$

So

$$\begin{aligned}
\nabla^d_A m(y) &= \int_A D^d_z m(y)\, \nu(dz) \\
&= \alpha \int_{H_S} L(y,x) \int_A \mathbf{1}_{z \notin \mathcal{C}(y)}\, \dot{x}_z\, \nu(dz)\, \mu_X(dx) \\
&= \alpha \int_{H_S} L(y,x) \int_A \dot{x}_z\, \nu(dz)\, \mu_X(dx), \text{ as } \nu \text{ is atomless,} \\
&= \alpha \int_{H_S} L(y,x)\, x(A)\, \mu_X(dx) \\
&= \alpha \int_{H_S} x(A)\, m(y)\, \mu_{X|Y}(dx,y), \text{by } \textit{iii)} \text{ of Proposition 2.4.1.}
\end{aligned}$$

By Proposition 2.4.4, this leads to

$$\mathbb{E}[X_A|\mathcal{Y}] = \frac{\nabla^d_A m(Y)}{\alpha\, m(Y)}, \; A \in \mathcal{B}(S).$$

□

Remark 2.4.7. *Neither ∇^d nor D^d satisfy the chain rule of derivation, and consequently*

$$\frac{\nabla^d_A F}{F} \neq \nabla_A \log F.$$

We conclude this Section by a more explicit case, that is the classical Poisson process on a time interval $[0,T]$ equipped with the Lebesgue measure dt. More precisely, let $(X_t)_{t\in[0,T]}$ be an input signal with values in $H_{[0,T]}$ (see Section 2.3). The output $(Y_t)_{t\in[0,T]}$ is supposed to be a Poisson process with intensity $1+\alpha X$ where α is some fixed scale parameter.
The likelihood denoted by L is given by

$$L(y,x) = \exp\left(-\alpha \int_S \dot{x}_s ds\right) \prod_{k=0}^{y([0,T])} (1+\alpha\, \dot{x}_{z_k}),$$

where $z_k \in \mathcal{C}(y)$.
Proposition 2.4.6 becomes the following corollary,

Corollary 2.4.8.

$$\mathbb{E}[X_t|\mathcal{Y}] = \frac{\nabla^d_{[0,t]} m(Y)}{\alpha\, m(Y)}, \; t \in [0,T]. \tag{2.4.9}$$

Remark 2.4.9. *The nonlinear filter given by equations (2.4.6) and (2.4.9) can be numerically approximated by evaluating m in (2.4.4) by a Monte-Carlo scheme. This computation is really tractable since the Malliavin derivative ∇ is a difference operator.*

2.5 A generalization to a class of non-Gaussian channels

In this Section we give a generalization of results from Section 2.4. We use some notations and definitions presented in Section 2.6.
Let $(M_t)_{t\in[0,T]}$ a normal martingale which satisfy a structure equation of the form (2.6.1) and which has the chaos representation property (see Definition 2.6.4). Let $(X_t)_{t\in[0,T]}$ a real-valued input process with $X_t = \displaystyle\int_0^t \dot{X}_s\, ds,\ t \in [0,T]$. Assume the output signal $(Y_t)_{t\in[0,T]}$ is a normal martingale such that the measure of the output given the input $\mu_{Y|X}$ is absolutely continuous with respect to $\mathbb{P}$ with likelihood given by

$$\begin{aligned} L(y,x) &= \frac{d\mu_{Y|X=x}}{d\mathbb{P}} \\ &= \exp\left(\int_0^T \dot{x}_s\, dy_s - \frac{1}{2}\int_0^T \dot{x}_s^2 \mathbf{1}_{\{\phi_s=0\}}\, ds\right) \\ &\times \prod_{s\le T}(1+\dot{x}_s\phi(s))e^{-\dot{x}_s\phi(s)}. \end{aligned} \tag{2.5.1}$$

We give a brief explanation of the previous formula (see [109, Theorem 36, p. 77]).

- The continuous martingale parts of Y given X and Y have the same quadratic variations.
- The random measure which define the pure jump martingale part of Y given X has intensity $(1+\dot{X}_t)\nu(dt)$ where ν denotes the random measure associated to the pure jump martingale part of Y.

Lemma 2.5.1. *With notations of Definition 2.6.3 we have,*

$$L(y,x) = \sum_{n=0}^{\infty} \frac{1}{n!} I_n^y(\dot{x}^{\otimes n}).$$

Proof. We give a sketch of the proof. This identity is proved by noting that the process

$$Z_t := 1 + \sum_{n=1}^{\infty} \frac{1}{n!} I_n^y\left(\dot{x}^{\otimes n}\mathbf{1}_{[0,t]^n}\right)$$

is solution to the stochastic differential equation

$$dZ_t = \dot{x}_t Z_t dy_t, \quad t \in [0,T], \tag{2.5.2}$$

Cf. [104]. Furthermore the process defined in (2.5.1) is also solution of the SDE (2.5.2) which ends the proof. □

This formulation of L and the Definition 2.6.5 of the Malliavin derivative in this context give

$$D_t L(y,x) = \dot{x}_t L(y,x),\ t \in [0,T]. \qquad (2.5.3)$$

By using the general Bayesian results presented in Section 2.4 we have the following Proposition.

Proposition 2.5.2. $E[X_t|\mathcal{Y}] = \dfrac{\nabla_t m(Y)}{m(Y)},\ t \in [0,T].$

Proof. Propositions 2.4.1 is valid in this context. □

We conclude this section by giving two important examples of normal martingales considered above.

- Assume $(\phi_t)_{t\in[0,T]}$ is deterministic.
  Then $(M_t)_{t\in[0,T]}$ has the chaos representation property see [40], and $(M_t)_{t\in[0,T]}$ can be represented as

  $$dM_t = i_t dB_t + j_t(dN_t - \lambda_t dt),\ M_0 = 0,\ t \in [0,T]$$

  where $(B_t)_{t\in[0,T]}$ is a standard Brownian motion, $i_t = \mathbf{1}_{\{\phi_t=0\}}$, $j_t = \mathbf{1}_{\{\phi_t=1\}}$, and $(N_t)_{t\in[0,T]}$ is a Poisson process independent of $(B_t)_{t\in[0,T]}$ with intensity $\nu_t = \displaystyle\int_0^t \frac{j_s}{\phi_s^2}\,ds$.
  Consequently,

  - For $\phi_t = 1,\ t \in [0,T]$; $(M_t)_{t\in[0,T]}$ is a Poisson process with intensity $\nu_t = \displaystyle\int_0^t \frac{1}{\phi_s^2}\,ds$.
  - For $\phi_t = 0,\ t \in [0,T]$; then $(M_t)_{t\in[0,T]}$ is a standard Brownian motion.

- Consider $\phi_t = \beta M_t,\ \beta \in [-2,0)$. Then $(M_t)_{t\in[0,T]}$ is an Azéma martingale. This process has the chaos decomposition property but its increments are not independent contrary to the previous example.

2.6 Appendix

In this Appendix we give some further elements of stochastic analysis in the framework of normal martingales.

Definition 2.6.1. *A stochastic process $(M_t)_{t\in[0,T]}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a right continuous filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is a normal martingale in $L^2(\Omega)$ if it is a martingale, that is, $\mathbb{E}[M_t^2] < \infty,\ t \in [0,T]$ and*

$$\mathbb{E}[M_t|\mathcal{F}_s] = M_s,\ 0 \le s < t \le T,$$

such that,

$$\mathbb{E}[(M_t - M_s)^2|\mathcal{F}_s] = t - s,\ 0 \le s < t \le T.$$

Let $(M_t)_{t\in[0,T]}$ a normal martingale on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with right continuous filtration $(\mathcal{F}_t)_{t\in[0,T]}$.

Definition 2.6.2. *$(M_t)_{t\in[0,T]}$ satisfies a structure equation if there exists an adapted process $(\phi_t)_{t\in[0,T]}$ such that*

$$[M,M]_t = t + \int_0^t \phi_s \, dM_s, \; t \in [0,T]. \tag{2.6.1}$$

Definition 2.6.3. *For $n \geq 1$, let $L^2([0,T])^{\circ n}$ be the space of symmetric functions f_n in n variables. For, f_n in $L^2([0,T])^{\circ n}$ define the iterated stochastic integral $I_n^M(f_n)$ by*

$$\begin{aligned} & I_n^M(f_n) \\ := \; & n! \int_0^T \int_0^{t_n} \dots \int_0^{t_2} f_n(t_1, \cdots, t_n) \, dM_{t_1} \dots dM_{t_n}. \end{aligned}$$

For f_0 in $\mathbb{R}$ we let $I_0(f_0) := f_0$.

With the notations of the previous Definition one can show that

$$I_n^M(f_n) = n \int_0^T I_{n-1}^M(f_n(*, t)\mathbf{1}_{[0,t]^{n-1}}(*)) \, dM_t, \; n \geq 1.$$

Definition 2.6.4.

Denote for $n \geq 1$,

$$\mathcal{H}_n = \{I_n^M(f_n), \; f_n \in L^2([0,T])^{\circ n}\}.$$

We say that $(M_t)_{t\in[0,T]}$ has the chaos representation property if

$$L^2(\Omega) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n,$$

that is, for every F in $L^2(\Omega)$ there exists $(f_n)_{n\in\mathbb{N}}$ such that $f_n \in L^2([0,T])^{\circ n}, \; n \geq 1$ and

$$F = \sum_{n=0}^{\infty} I_n^M(f_n).$$

We introduce the Malliavin derivative with respect to $(M_t)_{t\in[0,T]}$.

Definition 2.6.5. *Let*

$$\mathcal{S} = \left\{ \sum_{k=0}^{n} I_k^M(f_k), \; f_k \in L^2([0,T])^{\circ k}, 0 \leq k \leq n, n \in \mathbb{N} \right\}.$$

We define the Malliavin derivative D as the linear operator from $\mathcal{S}$ to $L^2(\Omega \times [0,T])$ by

$$D_t I_n^M(f_n) = n I_{n-1}^M(f_n(*, t)), \; d\mathbb{P} \times dt - a.e.$$

For t in $[0,T]$ define ∇_t as

$$\nabla_t F = \int_0^t D_s F \, ds, \; F \in Dom(D).$$

Chapitre 3

TLC pour les variations quadratiques à poids de champs gaussiens

Ce chapitre est composé des références [117, 116, 90].

3.1 Estimation de la variation quadratique de certains processus de diffusions à deux paramètres

Cette section est publiée [117].

3.1.1 Introduction

In recent years, two-parameter stochastic processes have been used to model and study 2-dimensional spatial phenomena. In particular a great interest was given to medical images such as X-ray pictures of bones (*c.f.* [19, 30, 73]) or digital mammograms ([16]) where one would like to test presence or absence of anomalies from these data. In these applications, the texture of the image is modeled by a two-parameter stochastic process, as for example the fractional Brownian field (see [19]) or the fractional Brownian sheet ([7, 73]) whose sample paths present fractal (or autosimilar) properties (determined respectively by one or two parameter, called the Hurst index(es)) representing the roughness features of the phenomenon itself. The estimation of the Hurst index(es) from the data leads to diagnosis of the target anomalies as for example osteoporosis ([7, 19, 30]) or leads to a best understanding of the observed structure as *e.g.* the "Full-Filled Digital Mammograms" texture ([16]). A major issue is consequently to obtain an accurate estimation of the Hurst index(es). One way to achieve this goal is to study the asymptotic behavior of the quadratic variations of the discretely observed process as it was initially realized in ([61]). Similar quantities have been studied in [17, 16, 19, 30, 73]. Regarding these applications, a natural next step in order to refine the statistical models and statistical methods in this area is the study of a generalization of the quadratic variations which is the weighted quadratic variations of two-parameter processes and in particular of the fractional Brownian sheet.

In this paper, we study the asymptotic behavior of the weighted quadratic variations process of the standard Brownian sheet since conclusions obtained in this case can be thought as the most optimistic one can expect for a general fractional Brownian sheet. First we give a central limit theorem for the weighted quadratic variations process of a two-parameter Brownian motion. More precisely if $W = (W_{(s,t)})_{(s,t)\in[0,1]^2}$ is a two-parameter Brownian motion and if $f : \mathbb{R} \to \mathbb{R}$ is a deterministic and regular enough function, we show in Theorem 3.1.3 that

$$n \sum_{i=1}^{[n\cdot]} \sum_{j=1}^{[n\bullet]} f\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)\left(|\Delta_{i,j}W|^2 - \frac{1}{n^2}\right) \xrightarrow[n\to\infty]{law(\mathcal{S})} \sqrt{2}\int_{[0,\cdot]\times[0,\bullet]} f\left(W_{(u,v)}\right)\, dB_{(u,v)}, \quad (3.1.1)$$

where B is a two-parameter Brownian motion independent of W and $\Delta_{i,j}W$ denotes the increment of the process W on the subset $\Delta_{i,j} := \left[\frac{i-1}{n}, \frac{i}{n}\right] \times \left[\frac{j-1}{n}, \frac{j}{n}\right]$ of $[0,1]^2$ defined by

$$\Delta_{i,j}W := W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)} + W_{\left(\frac{i}{n},\frac{j}{n}\right)} - W_{\left(\frac{i-1}{n},\frac{j}{n}\right)} - W_{\left(\frac{i}{n},\frac{j-1}{n}\right)}. \quad (3.1.2)$$

The notation $law(\mathcal{S})$ used above in (3.1.1) means that the convergence is in the sense of stable convergence in law in the two-parameter Skorohod space. In addition we stress that the limiting process is defined on an extension of the considered probability basis. Secondly, as an application, we deduce a central limit theorem (Theorem 3.1.7) for the quadratic variations process of a two-parameter diffusion $Y = (Y_{(s,t)})_{(s,t)\in[0,1]^2}$ observed on a regular grid which allows us to construct an asymptotically normal consistent estimator of the quadratic variation process defined below by (3.1.5). Indeed, consider a two-parameter stochastic process $(Y_{(s,t)})_{(s,t)\in[0,1]^2}$ defined by

$$Y_{(s,t)} := \int_{[0,s]\times[0,t]} \sigma\left(W_{(u,v)}\right)\, dW_{(u,v)} + \int_{[0,s]\times[0,t]} M_{(u,v)}\, dudv,\ (s,t)\in[0,1]^2, \quad (3.1.3)$$

where $(W_{(s,t)})_{(s,t)\in[0,1]^2}$ is a two-parameter Brownian motion, $\sigma : \mathbb{R} \to \mathbb{R}$ is a sufficiently smooth deterministic function and $(M_{(s,t)})_{(s,t)\in[0,1]^2}$ is a continuous adapted process. Assume Y is observed on the regular grid $G_n := \{(i/n, j/n)\ 1 \le i,j \le n\}$. Let

$$V^n_{(s,t)} := \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} |\Delta_{i,j}Y|^2, \quad (s,t)\in[0,1]^2,\ n\ge 1, \quad (3.1.4)$$

and

$$C_{(s,t)} := \int_{[0,s]\times[0,t]} \sigma^2\left(W_{(u,v)}\right)\, dudv, \quad (s,t)\in[0,1]^2. \quad (3.1.5)$$

Using (3.1.1) we show in Theorem 3.1.7 that

$$n\left(\sum_{i=1}^{[n\cdot]} \sum_{j=1}^{[n\bullet]} |\Delta_{i,j}Y|^2 - \int_{[0,\cdot]\times[0,\bullet]} \sigma^2\left(W_{(u,v)}\right)\, dudv\right) \quad (3.1.6)$$

$$\xrightarrow[n\to\infty]{law(\mathcal{S})} \quad \sqrt{2}\int_{[0,\cdot]\times[0,\bullet]} \sigma^2\left(W_{(u,v)}\right)\ dB_{(u,v)}.$$

Then (3.1.6) is used to prove that the consistent estimator V^n of C is asymptotically normal (*cf.* Corollary 3.1.11).

Similar results have been recently established in the one-parameter setting [2, 9, 50, 64, 63, 43, 44] and applied to the estimation of the integrated volatility (see for example [2] and references therein), testing for jumps of a process observed at discrete times like for example in [1] or to construct a goodness-of-fit test for the integrated volatility ([50]). The reader is also referred to [9, 10, 11] where the study of the power variations process has been used to solve some financial econometric problems.
Note that in [64, 63], functional limit theorems for quite general functions of the variations of an Itô semimartingales are proved. We also mention that the asymptotic behavior of the weighted power variations of a one-parameter fractional Brownian motion was studied in [81, 84] (see also [88] for a similar questions about the iterated Brownian motion).

We proceed as follows. First we recall in Section 3.1.2 some elements of stochastic analysis of two-parameter processes. Actually we present some definitions concerning stochastic calculus of two-parameter processes and the definition of the two-parameter Skorohod space initially introduced in [79] and in [126]. Secondly, in Section 3.1.3 we establish the central limit theorem (Theorem 3.1.3) for the weighted quadratic variations process of the two-parameter Brownian motion briefly presented in (3.1.1). As an application we prove in Section 3.1.4 that the consistent estimator V^n (3.1.4) of the quadratic variation C (3.1.5) is asymptotically normal (Corollary 3.1.11). Finally we present in an appendix (Section 3.1.5) some background on set-indexed processes and on the Malliavin calculus for the two-parameter Brownian motion which are used in Sections 3.1.3 and 3.1.4.

3.1.2 Stochastic analysis of two-parameter processes

In this section we recall some definitions about two-parameter stochastic analysis (we refer to [35, 60, 91, 92, 120] for complete explanation about this topic) and about two-parameter Skorohod space (introduced in [13, 79, 126]) which will be used in Sections 3.1.3 and 3.1.4.

Some elements of two-parameter stochastic calculus

Let $(\Omega, \mathcal{F}, (\mathcal{F}_z)_{z\in[0,1]^2}, \mathbb{P})$ be a filtered probability space.
We denote the partial order relation $\preceq$ on $[0,1]^2$ defined by,

$$z' \preceq z \Leftrightarrow (s' \leq s \text{ and } t' \leq t), \quad z' = (s', t'),\ z = (s, t).$$

We also define the *strong past information* filtration on $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 3.1.1. *Let $z = (s,t)$ in $[0,1]^2$.*

$$\mathcal{F}^*_{(s,t)} := \bigvee_{s' \leq s \text{ or } t' \leq t} \mathcal{F}_{(s',t')}.$$

Until the end of this paper we assume that the following *commutation condition* hold. This property is a *conditional independence property* (CI in short) and corresponds to the condition (F4) of [22].

Assumption (CI):
The filtration $(\mathcal{F}_z)_{z\in[0,1]^2}$ is supposed to satisfy the (CI) *condition i.e. for all $z = (s,t)$ and $z' = (s',t')$ in $[0,1]^2$*

$$\mathbb{E}\left[\mathbb{E}\left[\cdot|\mathcal{F}_z\right]|\mathcal{F}_{[0,z]\cap[0,z']}\right] = \mathbb{E}\left[\cdot|\mathcal{F}_{(s\wedge s',t\wedge t')}\right].$$

Definition 3.1.2. *An $(\mathcal{F}_z)_{z\in[0,1]^2}$-adapted process $(Y_z)_{z\in[0,1]^2}$ is said to be*

i) a martingale if for every z and z' in $[0,1]^2$ such that $z \preceq z'$

$$\mathbb{E}[Y_{z'}|\mathcal{F}_z] = Y_z,$$

ii) a strong martingale if for all z and z' in $[0,1]^2$ such that $z \preceq z'$

$$\mathbb{E}[Y_{[z,z']}|\mathcal{F}^*_z] = 0,$$

where $Y_{[z,z']}$ denotes the increments of Y on the interval $[z,z']$.

As an example, we mention the two-parameter Brownian motion $(W_z)_{z\in[0,1]^2}$ is a strong martingale with respect to its natural filtration and a centered Gaussian process with covariance function,

$$\mathbb{E}[W_{(s,t)}W_{(s',t')}] = (s\wedge s')(t\wedge t'), \quad (s,t),(s',t') \in [0,1]^2.$$

Skorohod space $\mathcal{D}([0,1]^2)$

In the one-parameter setting, Skorohod introduced in [121] four topologies known as J_1, J_2, M_1 and M_2. The topology M_2 is the weakest of theses topologies in the sense that convergence of a sequence $(x_n)_n$ of functions on $[0,1]$ to x for J_1, J_2 or M_1 consists in the convergence of $(x_n)_n$ to x in M_2 plus some additional conditions. The M_2 topology has been extended to the general setting of set-indexed functions by Bass and Pyke in [13] whereas the J_1 topology has been extended to multiparameter functions by Neuhaus and Straf respectively in [79] and [126]. The two-parameter Skorohod space (relative to J_1) introduced in [79] and [126] is denoted by $\mathcal{D}([0,1]^2)$ and give an equivalent to two-parameter functions of the notion of càdlàg functions on $[0,1]$. The set $\mathcal{D}([0,1]^2)$ can be equipped with a metric d which makes it a Polish space and we denote by $\mathcal{L}_2$ the Borel σ-algebra on $(\mathcal{D}([0,1]^2), d)$. Note that as in the one-parameter setting the J_1 topology is stronger than the M_2 topology. Furthermore compact sets (relative to J_1) on $(\mathcal{D}([0,1]^2), d, \mathcal{L}_2)$ can be described thanks to a modulus of continuity w which enables us to use techniques described in [18] for one-parameter functions. We conclude this section by giving the definition of w. Let $f : [0,1]^2 \to \mathbb{R}$ be an element of $\mathcal{D}([0,1]^2)$ and $\delta > 0$ we define $w(f,\delta)$ as,

$$w(f,\delta) := \sup_{\|(s,t)-(s',t')\|<\delta} |f(s,t) - f(s',t')|, \tag{3.1.7}$$

where $\|(s,t) - (s',t')\| := \max\{|s-s'|; |t-t'|\}$ for $(s,t), (s',t') \in [0,1]^2$.

3.1.3 Central limit theorem

In this section we state and prove the functional limit theorem (Theorem 3.1.3) which will allow us to show in Section 3.1.4 that the consistent estimator V^n (3.1.4) of the quadratic variation C (3.1.5) is asymptotically normal (Corollary 3.1.11).

Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable deterministic function. Let a two-parameter Brownian motion $W = (W_{(s,t)})_{(s,t)\in[0,1]^2}$ defined on a probability basis $\mathcal{B} := \left(\Omega, \mathcal{F}, (\mathcal{F}_{(s,t)})_{(s,t)\in[0,1]^2}, \mathbb{P}\right)$. Let also

$$\xi_{i,j} := n\, f\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) \left(|\Delta_{i,j}W|^2 - \frac{1}{n^2}\right), \quad 1 \le i,j \le n, \quad n \ge 1.$$

The re-normalized weighted quadratic variations process $X^n = (X^n_{(s,t)})_{(s,t)\in[0,1]^2}$ is defined as

$$X^n_{(s,t)} := \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \xi_{i,j}, \quad (s,t) \in [0,1]^2. \tag{3.1.8}$$

Stable convergence in law has been introduced in [114, 115]. It requires some particular care, since the limiting process X (see below) is not defined on the probability basis $\mathcal{B} = (\Omega, \mathcal{F}, (\mathcal{F}^*_z)_{z\in[0,1]^2}, \mathbb{P})$ on which the X^n, $n \ge 1$ are defined but on an extension $\tilde{\mathcal{B}} := (\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_z)_{z\in[0,1]^2}, \tilde{\mathbb{P}})$ of $\mathcal{B}$. We set the hypothesis we will need in the following theorem.

Hypothesis (**H**):
$f : \mathbb{R} \to \mathbb{R}$ is a measurable function such that

$$\sup_{(s,t)\in[0,1]^2} \mathbb{E}\left[\left|f\left(W_{(s,t)}\right)\right|^p\right] < \infty, \quad \forall p \in (0,4].$$

Theorem 3.1.3. *Assume that the deterministic function $f : \mathbb{R} \to \mathbb{R}$ satisfies hypothesis (**H**). Then $(X^n)_{n\ge1}$ defined by (3.1.8) converges $\mathcal{F}$-stably in law in the Skorohod space $(\mathcal{D}([0,1]^2), d, \mathcal{L}_2)$ to a non-Gaussian continuous process X defined below in the proof by (3.1.9). Moreover, X is defined on an extension of the probability basis $\mathcal{B}$.*

Remark 3.1.4. *The definition of $\mathcal{F}$-stable convergence in law will be given in the proof of Theorem 3.1.3 but heuristically this convergence can be thought as the convergence in law of $(X^n, W)_n$ to (X, W). Note also that the limiting process X is not a Gaussian process but it is conditionally Gaussian given the filtration generated by W, that is $\mathcal{F}$.*

Proof. Let us first describe the extension of $\mathcal{B}$ on which the limiting process X is defined. We denote by $\mathcal{B}' := (\Omega', \mathcal{F}', (\mathcal{F}'_z)_{z\in[0,1]^2}, \mathbb{P}')$ the two-parameter Wiener space, that is $\Omega' := \mathcal{C}^0([0,1]^2)$ is the space of real-valued continuous functions on $[0,1]^2$ vanishing on the set $\{(s,t) \in [0,1]^2,\ s = 0 \text{ or } t = 0\}$. Set $\mathbb{P}'$ the unique measure on $(\Omega', \mathcal{F}')$ under which the canonical process $(B_z)_{z\in[0,1]^2}$ on Ω' defined by

$$B_z(\omega') := \omega'(z), \quad \omega' \in \Omega',\ z \in [0,1]^2,$$

is a standard two-parameter Brownian motion. Let the extension $\tilde{\mathcal{B}} := (\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}^*_z)_{z\in[0,1]^2}, \tilde{\mathbb{P}})$ defined as

$$\begin{cases} \tilde{\Omega} := \Omega \times \Omega', \\ \tilde{\mathcal{F}} := \mathcal{F} \otimes \mathcal{F}', \\ (\tilde{\mathcal{F}}_z)_{z\in[0,1]^2} := (\cap_{\rho > z} \mathcal{F}^*_\rho \otimes \mathcal{F}'_\rho)_{z\in[0,1]^2}, \\ \tilde{\mathbb{P}}(d\omega, dy) := \mathbb{P}(d\omega)\mathbb{P}'(dy). \end{cases}$$

We will denote by $\mathbb{E}$ (respectively $\tilde{\mathbb{E}}$) the expectation under $\mathbb{P}$ (respectively under $\tilde{\mathbb{P}}$). On $\tilde{\mathcal{B}}$ we define the stochastic process $(X_z)_{z\in[0,1]^2}$ as

$$X_z(\omega, \omega') := \sqrt{2} \left(\int_{[0,z]} f(W_\rho(\omega))\, dB_\rho \right)(\omega'), \quad z \in [0,1]^2 \tag{3.1.9}$$

which is a $\mathcal{F}$-progressive conditional Gaussian martingale with independent increments on $\tilde{\mathcal{B}}$; that is, X is an $(\tilde{\mathcal{F}}_z)_{z\in[0,1]^2}$- adapted process such that for $\mathbb{P}$ almost ω in Ω, $X(\omega, \cdot)$ is a Gaussian process on $\mathcal{B}'$ with covariance function

$$\begin{aligned} & \mathbb{E}_{\mathbb{P}'}\left[X_{(s_1,t_1)}(\omega,\cdot)X_{(s_2,t_2)}(\omega,\cdot)\right] \\ = \ & 2\int_{[s_1\wedge s_2, s_1\vee s_2]\times[t_1\wedge t_2, t_1\vee t_2]} f^2\left(W_\rho\right)(\omega)\, d\rho, \quad (s_1,t_1),(s_2,t_2) \in [0,1]^2. \end{aligned}$$

Note that $\tilde{\mathcal{B}}$ is clearly a *very good extension* of $\mathcal{B}$ in the sense of a two-parameter counterpart of [67, Definition II.7.1]. Since $(\mathcal{D}([0,1]^2), d, \mathcal{L}_2)$ is a Polish space, by [67, Proposition VIII.5.33], $\mathcal{F}$-stable convergence in law holds if for every random variable Z on $(\Omega, \mathcal{F}, \mathbb{P})$ the couple $(Z, X^n)_n$ converges in law. Adapting an argument presented in the proof of [67, Theorem VIII.5.7 b)], the convergence in law of a such couple $(Z, X^n)_n$ will be obtained by first proving in ***Step 1)*** that the sequence $(X^n)_n$ is tight (relative to the Skorohod space $(\mathcal{D}([0,1]^2), d, \mathcal{L}_2)$) and then by making in ***Step 2)*** an "identification of the limit" *via* $\mathcal{F}$-stable finite-dimensional convergence in law to X. Recall that the latter property means that for every integer $m \geq 0$, for every continuous and bounded function $\psi : \mathbb{R}^{m+1} \to \mathbb{R}$ and every elements $z_0, \ldots, z_m$ in a dense subset of $[0,1]^2$,

$$\mathbb{E}\left[Z\psi\left(X^n(z_0), X^n(z_1), \ldots, X^n(z_m)\right)\right] \underset{n\to\infty}{\longrightarrow} \tilde{\mathbb{E}}\left[Z\psi\left(X(z_0), X(z_1), \ldots, X(z_m)\right)\right]. \tag{3.1.10}$$

Step 1)

We show the sequence $(X^n)_n$ is tight in the Skorohod space $(\mathcal{D}([0,1]^2), d, \mathcal{L}_2)$.

A complete description of $(\mathcal{D}([0,1]^2), d, \mathcal{L}_2)$ can be found in [79]. In particular it is shown in [79] that the set of conditions (3.1.11) and (3.1.12) is necessary and sufficient for the sequence $(X^n)_n$ to be tight in $(\mathcal{D}([0,1]^2), d, \mathcal{L}_2)$,

$$(X_0^n)_n \text{ converges in distribution,} \tag{3.1.11}$$

$$\lim_{\delta\to 0}\limsup_{n\to\infty} \mathbb{P}[w(X^n,\delta)\geq \varepsilon] = 0, \quad \varepsilon > 0, \tag{3.1.12}$$

where w is defined in Section 3.1.2. Property (3.1.11) is clear since for every $n \geq 1$ $X_0^n = X_0 = 0$, $\mathbb{P}$-a.s.. We will show (3.1.12) using a method from [18, p. 89].
Let $\varepsilon > 0$, $\delta > 0$ and $n \geq 1$. Let $m := [n\delta]$ and $v := \left[\frac{n}{m}\right]$. We consider on $[0,1]^2$ the rectangles $R_{i,j} := \left[\frac{m_{i-1}}{n}, \frac{m_i}{n}\right] \times \left[\frac{m_{j-1}}{n}, \frac{m_j}{n}\right]$, $(i,j) \in \{1,\cdots,v\}^2$ where $m_i := im$, $1 < i < v$ and $m_v = n$. With this notation the length of the shortest side of the rectangles $R_{i,j}$ is greater than δ and $v \leq 2/\delta$. We can adapt the proof of [18, Theorem 7.4] to our case and we have

$$\mathbb{P}[w(X^n,\delta)\geq 3\varepsilon] \leq \sum_{i=1}^{v}\sum_{j=1}^{v} \mathbb{P}\left[\sup_{z\in R_{i,j}}\left|X_z^n - X^n_{\left(\frac{m_{i-1}}{n},\frac{m_{j-1}}{n}\right)}\right| \geq \varepsilon\right]. \tag{3.1.13}$$

Let us give some notations. For $(k,j) \in \{1,\ldots,n\}^2$ let

$$S_{k,l} := \sum_{i=1}^{k}\sum_{j=1}^{l} f\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)\left(|\Delta_{i,j}W|^2 - 1/n^2\right),$$

that is $X^n_{(k,l)} = nS_{k,l}$. For z in $R_{i,j}$ we write $\hat{S}^{i,j}_{k,l} := S_{k,l} - S_{\frac{m_{i-1}}{n},\frac{m_{j-1}}{n}}$. Using these notations we can write (3.1.13) as,

$$\mathbb{P}\left[w(X^n,\delta)\geq 3\varepsilon\right] \;\leq\; \sum_{i=1}^{v}\sum_{j=1}^{v}\mathbb{P}\left[\sup_{m_{i-1}\leq k\leq m_i,\; m_{j-1}\leq l\leq m_j}\left|\hat{S}^{i,j}_{k,l}\right| \geq \frac{\varepsilon}{n}\right].$$

We use [18, Section 10] which provides maximal inequalities for partial sums of non-independent and non-stationary random variables. For i,j fixed as above, we re-index the random variables appearing in $\hat{S}^{i,j}_{k,l}$ to obtain

$$\hat{S}^{i,j}_{k,l} = \sum_{p=1}^{\eta(i,j,k,l)} \tau_p,$$

with τ_p equal to some $\xi_{\cdot,\cdot}$ divided by n and $\eta(i,j,k,l)$ is an integer. With this notation we can show that

$$\mathbb{E}[|\tau_p|^4] \leq \frac{R}{n^8}, \quad \text{where } R \text{ is a constant.} \tag{3.1.14}$$

Actually, $\tau_p = f\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)(|\Delta_{i,j}W|^2 - 1/n^2)$ for some $1 \leq i,j \leq n$ and

$$\begin{aligned}
\mathbb{E}[|\tau_p|^4] &= \mathbb{E}\left[f^4\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)\mathbb{E}\left[\left(|\Delta_{i,j}W|^2 - 1/n^2\right)^4 \left|\mathcal{F}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right.\right]\right] \\
&\leq \sup_{(s,t)\in[0,1]^2} \mathbb{E}\left[f^4\left(W_{(s,t)}\right)\right]\mathbb{E}\left[\left(|\Delta_{i,j}W|^2 - 1/n^2\right)^4\right] \\
&= \frac{R}{n^8}, \text{ by hypothesis } (\mathbf{H}).
\end{aligned}$$

Let two integers $\alpha \leq \beta$. Let K denote a constant which can differ from one line to another.

$$\begin{aligned}
\mathbb{P}\left[\left|\sum_{p=\alpha+1}^{\beta}\tau_p\right| \geq \lambda\right] &\leq \frac{1}{\lambda^4}\mathbb{E}\left[\left|\sum_{p=\alpha+1}^{\beta}\tau_p\right|^4\right] \\
&\leq \frac{1}{\lambda^4}\sum_{p=\alpha+1}^{\beta}\mathbb{E}[|\tau_p|^4] + \frac{1}{\lambda^4}\sum_{p,q=\alpha+1}^{\beta}\mathbb{E}[|\tau_p|^2|\tau_q|^2] \\
&\leq \frac{K}{\lambda^4 n^8}(\beta-\alpha)^2, \text{ by (3.1.14).} \qquad (3.1.15)
\end{aligned}$$

Using [18, Theorem 10.2] and (3.1.15) we obtain

$$\mathbb{P}\left[\max_{i\leq k\leq m, j\leq l\leq m}\hat{S}_{k,l}^{i,j} \geq \lambda\right] \leq \frac{Km^4}{n^4\lambda^4}. \qquad (3.1.16)$$

Now injecting inequality (3.1.16) in (3.1.13) we have,

$$\begin{aligned}
\mathbb{P}\left[w(X^n,\delta) \geq 3\varepsilon\right] &\leq \frac{v^2 K m^4}{n^4\varepsilon^4} \\
&\leq \frac{Km^4}{\varepsilon^4 n^4 \delta^2}, \quad \text{since } v \leq 2/\delta, \\
&\leq \frac{K}{\varepsilon^4}\delta^2, \quad \text{since } m = [n\delta],
\end{aligned}$$

which leads to (3.1.12).

Step 2)

Here we choose to consider processes X^n and X as set-indexed processes and we use all the notations and definitions of Subsection 3.1.5. Consequently the $\mathcal{F}$-stable finite-dimensional convergence in law property (3.1.10) can be rewritten as follows: for every continuous and bounded function ψ, for every elements $C_0, \ldots, C_m$ in a dense subset of $\mathcal{A}$ (see Subsection 3.1.5 for definitions and notations) and for every random variable Z on $(\Omega, \mathcal{F}, \mathbb{P})$,

$$\mathbb{E}\left[Z\psi\left(X^n(C_0), X^n(C_1), \ldots, X^n(C_m)\right)\right] \underset{n\to\infty}{\longrightarrow} \tilde{\mathbb{E}}\left[Z\psi\left(X(C_0), X(C_1), \ldots, X(C_m)\right)\right]. \qquad (3.1.17)$$

To obtain (3.1.17) we adapt [62, Proposition 7.3.7] which allows us to replace $\mathcal{F}$-stable finite-dimensional convergence in law with *$\mathcal{F}$-stable semi-functional convergence in law* that is, for every *simple flow* φ (see Subsection 3.1.5) the sequence of one-parameter processes $(X^n\circ\varphi)_n$ converges $\mathcal{F}$-stably in law to the one-parameter process $X\circ\varphi$. Let us make precise this argument. Assume that *stable semi-functional convergence in law* holds. We aim at showing (3.1.17). As in [62, Proposition 7.3.7] since for every $n\geq 1$, X^n is an additive process (*c.f.* Subsection 3.1.5) it is enough to prove (3.1.17) for elements $C_0,\cdots,C_m$ such that for every $i\in\{1,\cdots,m\}$, $C_i=\varphi(i/m)-\varphi((i-1)/m)$ where φ is an arbitrary simple flow (see Subsection 3.1.5). Since the sequence of one-parameter càdlàg processes $(X^n\circ\varphi)_n$ is supposed to converge $\mathcal{F}$-stably in law to $X\circ\varphi$, and since one can choose a continuous version of $X\circ\varphi$, the projection $\pi_{(0,1/m,\ldots,1)}:\mathcal{D}([0,1])\longrightarrow\mathbb{R}^{m+1}$ is continuous and by mapping Theorem,

$$\begin{aligned}
&\mathbb{E}\left[Z\psi\left((X^n\circ\varphi)(0),(X^n\circ\varphi)(1/m),\cdots,(X^n\circ\varphi)(1)\right)\right]\\
\underset{n\to\infty}{\longrightarrow}\ &\int_{\Omega\times\Omega'_\varphi} Z(\omega)\psi\left((X\circ\varphi)(\omega,x)(0),\cdots,(X\circ\varphi)(\omega,x)(1)\right)\mathbb{P}'(dx)\mathbb{P}(d\omega)\\
=\ &\tilde{E}\left[Z\psi\left((X\circ\varphi)(0),(X\circ\varphi)(1/m),\cdots,(X\circ\varphi)(1)\right)\right],
\end{aligned}\tag{3.1.18}$$

where Z and ψ are like in (3.1.17). Consequently relation (3.1.17) holds since

$$X^n(C_i)=(X^n\circ\varphi)(i/m)-(X^n\circ\varphi)((i-1)/m),\quad 1\leq i\leq m.$$

Using the argument presented above we will now prove $\mathcal{F}$-stable semi-functional convergence in law to establish $\mathcal{F}$-stable finite-dimensional convergence in law.
Let φ be a simple flow (we write φ as $\varphi=(\varphi_1,\varphi_2)$). We have to show that the sequence of one-parameter càdlàg processes $(X^n\circ\varphi)_n$ converges $\mathcal{F}$-stably in law to the one-parameter process $X\circ\varphi$. We give some precisions about the extension of probability basis we use. We set $\mathcal{B}_\varphi:=(\Omega,\mathcal{F},(\mathcal{F}_{\varphi(t)})_{t\in[0,1]},\mathbb{P})$ and $\mathcal{B}'_\varphi:=(\Omega',\mathcal{F}',(\mathcal{F}'_{\varphi(t)})_{t\in[0,1]},\mathbb{P}')$. From $\mathcal{B}_\varphi$ and $\mathcal{B}'_\varphi$ we define the probability basis $\tilde{\mathcal{B}}_\varphi:=(\tilde{\Omega},\tilde{\mathcal{F}},(\tilde{\mathcal{F}}_{\varphi(t)})_{t\in[0,1]},\tilde{\mathbb{P}})$, with,

$$(\tilde{\mathcal{F}}_{\varphi(t)})_{t\in[0,1]}:=(\cap_{s>t}\mathcal{F}_{\varphi(s)}\otimes\mathcal{F}'_{\varphi(s)})_{t\in[0,1]}.$$

Let $n\geq 1$, by Lemma 3.1.14 and Lemma 3.1.5 the one-parameter processes $((X^n\circ\varphi)_t)_{t\in[0,1]}$ are martingales on the probability basis $\mathcal{B}^n_\varphi:=(\Omega,\mathcal{F},(\mathcal{F}^{n,\varphi}_t)_{t\in[0,1]},\mathbb{P})$ where,

$$\mathcal{F}^{n,\varphi}_t:=\mathcal{F}_{([n\varphi_1(t)]n^{-1},[n\varphi_2(t)]n^{-1})},\quad t\in[0,1].$$

We also define

$$\left\{\begin{array}{l}
\mathcal{F}'^{n,\varphi}_t:=\mathcal{F}'_{([n\varphi_1(t)]n^{-1},[n\varphi_2(t)]n^{-1})}\\
\tilde{\mathcal{F}}^{n,\varphi}_t:=\cap_{s>t}\mathcal{F}^{n,\varphi}_s\otimes\mathcal{F}'^{n,\varphi}_s,
\end{array}\right.$$

and the following probability bases,

$$\left\{\begin{array}{l}
\mathcal{B}'^n_\varphi:=(\Omega',\mathcal{F}',(\mathcal{F}'^{n,\varphi}_t)_{t\in[0,1]},\mathbb{P}'),\\
\tilde{\mathcal{B}}^n_\varphi:=(\tilde{\Omega},\tilde{\mathcal{F}},(\tilde{\mathcal{F}}^{n,\varphi}_t)_{t\in[0,1]},\tilde{\mathbb{P}}).
\end{array}\right.$$

We will now apply [67, Theorem IX.7.3] (or its triangular array formulation [67, Theorem IX.7.28]). This result gives conditions insuring $\mathcal{F}$-stable finite-dimensional convergence in law for a sequence of one-parameters martingales to a continuous conditional martingale with independent increments by identifying the caracteristics of these martingales plus some additional conditions. This identification is realized in Lemma 3.1.6 in which convergences (3.1.19) and (3.1.20) can be thought as identification of the characteristics whereas properties (3.1.21) and (3.1.22) ensure the $\mathcal{F}$-stable feature of the convergence. Consequently from [67, Theorem IX.7.3] and Lemma 3.1.6, the sequence $(X^n \circ \varphi)_n$ of one-parameter martingales on $\mathcal{B}^n_f$ converges $\mathcal{F}$-stably in law to $X \circ \varphi$ on the extension $\tilde{\mathcal{B}}_\varphi$ of $\mathcal{B}_\varphi$ which ends the proof. Note that $\tilde{\mathcal{B}}_\varphi$ is a very good extension of $\mathcal{B}_\varphi$ since $\tilde{\mathcal{B}}$ is a very good extension of $\mathcal{B}$. □

Before turning to estimation results in Section 3.1.4 we state and prove Lemma 3.1.5 and Lemma 3.1.6 which were used in the proof of Theorem 3.1.3.

Lemma 3.1.5. *We use notations of Theorem 3.1.3 and of its proof. X^n is a strong martingale.*

Proof. Let y and x in $[0,1]^2$ such that $y = (y_1, y_2) \preceq x = (x_1, x_2)$. Let also, $\mathcal{F}^{n,*}_y := \mathcal{F}^*_{(n^{-1}[ny_1], n^{-1}[ny_2])}$. We have

$$\begin{aligned}
& \mathbb{E}[X^n([y,x])|\mathcal{F}^{n,*}_y] \\
= \; & n \sum_{i=[ny_1]}^{[nx_1]} \sum_{j=[ny_2]}^{[nx_2]} \mathbb{E}\left[f\left(W_{(\frac{i-1}{n},\frac{j-1}{n})}\right) \left(|\Delta_{i,j} W|^2 - \frac{1}{n^2} \right) \middle| \mathcal{F}^{n,*}_y \right] \\
= \; & n \sum_{i=[ny_1]}^{[nx_1]} \sum_{j=[ny_2]}^{[nx_2]} \mathbb{E}\left[f\left(W_{(\frac{i-1}{n},\frac{j-1}{n})}\right) \mathbb{E}\left[\left(|\Delta_{i,j} W|^2 - \frac{1}{n^2} \right) \middle| \mathcal{F}^*_{((i-1)/n,(j-1)/n)} \right] \middle| \mathcal{F}^{n,*}_y \right] \\
= \; & 0.
\end{aligned}$$

□

Lemma 3.1.6. *We use notations of Theorem 3.1.3 and of its proof. In particular we denote a flow φ as $\varphi = (\varphi_1, \varphi_2)$. For every $n \geq 1$, $X^n \circ \varphi$ is a one-parameter martingale with modified second caracteristics $(0, \tilde{C}_{X^n \circ \varphi}, \nu_{X^n \circ \varphi})$ on $\mathcal{B}^n_\varphi$ such that*

$$\nu_{X^n \circ \varphi}([0,t] \times \{|x| > \varepsilon\}) \xrightarrow[n\to\infty]{P} 0, \quad \forall t \in [0,1],\ \varepsilon > 0, \tag{3.1.19}$$

$$\tilde{C}_{X^n \circ \varphi}(t) \xrightarrow[n\to\infty]{P} 2 \int_{[0,\varphi_1(t)]\times[0,\varphi_2(t)]} f^2(W_\rho)\, d\rho, \quad \forall t \in [0,1]. \tag{3.1.20}$$

Furthermore for every bounded martingale $\tilde{N}$ orthogonal to $W \circ \varphi$,

$$\langle X^n \circ \varphi, \tilde{N} \rangle_t \xrightarrow[n\to\infty]{P} 0, \quad \forall t \in [0,1]. \tag{3.1.21}$$

We also have

$$\langle X^n \circ \varphi, W \circ \varphi \rangle_t \xrightarrow[n\to\infty]{P} 0, \quad \forall t \in [0,1]. \tag{3.1.22}$$

Proof. Let $n \geq 1$.

Proof of (3.1.19):
In the following $\mu_{X^n \circ \varphi}$ denote the jump measure of $X^n \circ \varphi$. For $\omega \in \Omega$,

$$\begin{aligned}
\mu_{X^n \circ \varphi}(\omega, dt, dx) &= \sum_{k=1}^{n} \delta_{\{\varphi_1^{-1}(k/n)\} \cap (\varphi_2^{-1}(\{l/n,\, 1 \leq l \leq n\}))^c, H^{1,k,n}(\omega)}(dt, dx) \\
&+ \sum_{l=1}^{n} \delta_{\{\varphi_2^{-1}(k/n)\} \cap (\varphi_2^{-1}(\{k/n,\, 1 \leq k \leq n\}))^c, H^{2,l,n}(\omega)}(dt, dx) \\
&+ \sum_{k=1}^{n} \sum_{l=1}^{n} \delta_{\varphi_1^{-1}(k/n), \varphi_2^{-1}(l/n), H^{3,k,l,n}(\omega)}(dt, dx),
\end{aligned}$$

$$\begin{cases}
H^{1,k,n}(\omega) = \sum_{j=1}^{[n\varphi_2(\varphi_1^{-1}(k/n))]} \xi_{k,j}(\omega), \\
H^{2,l,n}(\omega) = \sum_{i=1}^{[n\varphi_1(\varphi_2^{-1}(l/n))]} \xi_{i,l}(\omega), \\
H^{3,k,l,n}(\omega) = H^{1,k,n} + H^{2,l,n} + \xi_{k,l}.
\end{cases}$$

We denote by $\nu_{X^n \circ \varphi}$ the compensator of the measure $\mu_{X^n \circ \varphi}$. Let A a Borel set in $\mathbb{R}$. We have

$$\begin{aligned}
\nu_{X^n \circ \varphi}(\omega, [0,t] \times A) &= \sum_{k=1}^{[n\varphi_1(t)]} \mathbb{E}\left[\mathbf{1}_{H^{1,k,n} \in A} | \mathcal{F}_{(k-1)/n, ([n\varphi_2(t)]-1)/n}\right] \\
&+ \sum_{l=1}^{[n\varphi_2(t)]} \mathbb{E}\left[\mathbf{1}_{H^{2,l,n} \in A} | \mathcal{F}_{(([n\varphi_1(t)]-1)/n, (l-1)/n)}\right] \\
&+ \sum_{k=1}^{[n\varphi_1(t)]} \sum_{l=1}^{[n\varphi_2(t)]} \mathbb{E}\left[\mathbf{1}_{H^{3,k,l,n} \in A} | \mathcal{F}_{([n\varphi_1(t)]-1)/n, ([n\varphi_2(t)]-1)/n}\right].
\end{aligned}$$

Let $\varepsilon > 0$ and k, l, n, t as above. Denote by C a constant which can differ from one line to another.

$$\begin{aligned}
\mathbb{P}\left[|H^{1,k,n}| > \varepsilon | \mathcal{F}_{((k-1)/n, [n\varphi_2(t)]/n)}\right] &\leq \sum_{j=1}^{[n\varphi_2(t)]} \mathbb{P}\left[|\xi_{k,j}| > \frac{\varepsilon}{n} | \mathcal{F}_{((k-1)/n, [n\varphi_2(t)]/n)}\right] \\
&\leq \frac{n^4}{\varepsilon^4} \sum_{j=1}^{[n\varphi_2(t)]} \mathbb{E}\left[|\xi_{k,j}|^4 | \mathcal{F}_{((k-1)/n, [n\varphi_2(t)]/n)}\right] \\
&= \frac{n^4}{\varepsilon^4} \sum_{j=1}^{[n\varphi_2(t)]} f^4\left(W_{\left(\frac{k-1}{n}, \frac{j-1}{n}\right)}\right) \mathbb{E}\left[\left|\left(|\Delta_{k,j} W|^2 - 1/n^2\right)\right|^4\right] \\
&\leq C \frac{1}{\varepsilon^4 n^3} \sup_{s \in [0,1]} \frac{1}{n} \sum_{j=1}^{[n\varphi_2(t)]} f^4\left(W_{\left(s, \frac{j-1}{n}\right)}\right)
\end{aligned}$$

$$\underset{n\to\infty}{\longrightarrow} 0, \quad \mathbb{P}-\text{a.s.}.$$

Indeed, by hypothesis (**H**), we have that $sup_{s\in[0,1]}\int_0^{\varphi_2(t)} f^4(W_{(s,t)})dt < \infty$ $\mathbb{P}$-a.s.. We deduce of the preceding inequalities that

$$\begin{aligned}
& \mathbb{P}\left[|H^{3,k,l,n}| > \varepsilon \middle| \mathcal{F}_{\left(\frac{[n\varphi_1(t)]-1}{n},\frac{[n\varphi_2(t)]-1}{n}\right)}\right] \\
\leq\ & \mathbb{P}\left[|H^{1,k,n}| > \frac{\varepsilon}{3} \middle| \mathcal{F}_{\left(\frac{[n\varphi_1(t)]-1}{n},\frac{[n\varphi_2(t)]-1}{n}\right)}\right] + \mathbb{P}\left[|H^{2,l,n}| > \frac{\varepsilon}{3} \middle| \mathcal{F}_{\left(\frac{[n\varphi_1(t)]-1}{n},\frac{[n\varphi_2(t)]-1}{n}\right)}\right] \\
+\ & \mathbb{P}\left[|\xi_{k,l}| > \frac{\varepsilon}{3} \middle| \mathcal{F}_{\left(\frac{[n\varphi_1(t)]-1}{n},\frac{[n\varphi_2(t)]-1}{n}\right)}\right] \\
\underset{n\to\infty}{\longrightarrow}\ & 0, \quad \mathbb{P}-\text{a.s.}.
\end{aligned}$$

This leads to (3.1.19). ***Proof of (3.1.20):***
$\tilde{C}_{X^n\circ\varphi}(t) = \langle X^n\circ\varphi, X^n\circ\varphi\rangle$ is the compensator of $[X^n\circ\varphi, X^n\circ\varphi]$ with respect to $(\mathcal{F}_t^{n,\varphi})_{t\in[0,1]}$ (see for example [67, Proof of Proposition II.2.17 b)]) and we have (*cf.* [67, (I.4.53)])

$$[X^n\circ\varphi, X^n\circ\varphi]_t = \sum_{0\leq s\leq t} (X^n\circ\varphi)(s) - (X^n\circ\varphi)(s_-) = \sum_{i=1}^{[n\varphi_1(t)]}\sum_{j=1}^{[n\varphi_2(t)]} \xi_{i,j}^2, \quad t\in[0,1].$$

Consequently,

$$\begin{aligned}
\tilde{C}_{X^n\circ\varphi}(t) &= \sum_{k=1}^{[n\varphi_1(t)]} \mathbb{E}\left[(H^{1,k,n})^2|\mathcal{F}_{(k-1)/n,([n\varphi_2(t)]-1)/n}\right] \\
&+ \sum_{l=1}^{[n\varphi_2(t)]} \mathbb{E}\left[(H^{2,l,n})^2|\mathcal{F}_{(([n\varphi_1(t)]-1)/n,(l-1)/n)}\right] \\
&+ \sum_{k=1}^{[n\varphi_1(t)]}\sum_{l=1}^{[n\varphi_2(t)]} \mathbb{E}\left[(H^{3,k,l,n})^2|\mathcal{F}_{([n\varphi_1(t)]-1)/n,([n\varphi_2(t)]-1)/n}\right].
\end{aligned}$$

We can show that this sum is equal to

$$\tilde{C}_{X^n\circ\varphi}(t) = \frac{2}{n^2}\sum_{i=1}^{[n\varphi_1(t)]}\sum_{j=1}^{[n\varphi_2(t)]} f^2\left(W_{\left(\frac{k-1}{n},\frac{j-1}{n}\right)}\right), \quad t\in[0,1], \tag{3.1.23}$$

since terms of the form $\mathbb{E}\left[\xi_{k,j}\xi_{k,l}\middle|\mathcal{F}_{\left(\frac{k-1}{n},\frac{[n\varphi_2(t)]-1}{n}\right)}\right]$ vanish for $j < l \leq [n\varphi_2(t)]$ using the same type an argument described in the proof of (3.1.19). Furthermore terms of the form $\mathbb{E}\left[\xi_{k,j}^2|\mathcal{F}_{\left(\frac{k-1}{n},\frac{[n\varphi_2(t)]-1}{n}\right)}\right]$ are given by

$$\mathbb{E}\left[\xi_{k,j}^2|\mathcal{F}_{\left(\frac{k-1}{n},\frac{[n\varphi_2(t)]-1}{n}\right)}\right] = n^2 f^2\left(W_{\left(\frac{k-1}{n},\frac{j-1}{n}\right)}\right)\mathbb{E}\left[\left|\left|\Delta_{k,j}^n W\right|^2 - 1/n^2\right|^2\right]$$

$$= \frac{2}{n^2} f^2 \left(W_{\left(\frac{k-1}{n}, \frac{j-1}{n}\right)} \right).$$

We deduce (3.1.20) from (3.1.23).

Proof of (3.1.21):

Let $\tilde{N}$ be a martingale orthogonal to $W \circ \varphi$. Without loss of generality we can assume there exists a strong martingale N on $\mathcal{B}$ orthogonal to W such that NW is a strong martingale and such that $\tilde{N} = N \circ \varphi$. Let $n \geq 1$ and t in $[0,1]$. We have

$$\begin{aligned} \langle X^n, N \rangle_t &= \sum_{i=1}^{[n\varphi_1(t)]} \sum_{j=1}^{[n\varphi_2(t)]} \mathbb{E}\left[\xi_{i,j} \Delta_{i,j} N \middle| \mathcal{F}_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)} \right] \\ &= \sum_{i=1}^{[n\varphi_1(t)]} \sum_{j=1}^{[n\varphi_2(t)]} f\left(W_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)} \right) \mathbb{E}\left[\left(|\Delta_{i,j} W|^2 - 1/n^2 \right) \Delta_{i,j} N \middle| \mathcal{F}_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)} \right]. \end{aligned}$$

Let $1 \leq i,j \leq n$. We use a technique presented in [63, Lemma 6.8]. For z in $[0,1]^2$ we define $U_z := \mathbb{E}\left[|\Delta_{i,j} W|^2 - 1/n^2 | \mathcal{F}_z \right]$. $(U_z)_{z \succeq ((i-1)/n, (j-1)/n)}$ is a martingale for the filtration generated by $U_z - U_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)}$. Using the representation $|\Delta_{i,j} W|^2 - \frac{1}{n^2} = I_2\left(\mathbf{1}_{\Delta_{i,j}}^{\otimes 2} \right)$ as a multiple stochastic integral (see for example [94, Section 1.1.2] or [98]) we have by [94, Lemma 1.2.5] that for $z \succeq ((i-1)/n, (j-1)/n)$, $U_z = I_2\left(\mathbf{1}_{\Delta_{i,j}}^{\otimes 2} \mathbf{1}_{[0,z]} \right)$. From [77] there exists an adapted process $(\Phi_{(s,t)})_{(s,t) \in [0,1]^2}$ such that

$$U_{(s,t)} = U_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)} + \int_{[(i-1)/n, s] \times [(j-1)/n, t]} \Phi_\rho \, dW_\rho.$$

Define $N'_z = N_z - N_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)}$, $z \succeq ((i-1)/n, (j-1)/n)$. This process is orthogonal to $(U_z)_{z \geq ((i-1)/n, (j-1)/n)}$. Consequently using a characterization of orthogonal two-parameter martingales given in [22, Proposition 1.6] we have that

$$\mathbb{E}\left[\Delta_{i,j} N' \Delta_{i,j} U \middle| \mathcal{F}_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)} \right] = 0.$$

A straightforward computation gives that

$$\mathbb{E}\left[\left(|\Delta_{i,j} W|^2 - 1/n^2 \right) \Delta_{i,j} N \middle| \mathcal{F}_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)} \right] = \mathbb{E}\left[\Delta_{i,j} N' \Delta_{i,j} U \middle| \mathcal{F}_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)} \right].$$

Proof of (3.1.22):

Let $n \geq 1$ and $t \in [0,1]$. We have

$$\begin{aligned} \langle X^n \circ \varphi, W \circ f \rangle_t &= \sum_{i=1}^{[n\varphi_1(t)]} \sum_{j=1}^{[n\varphi_2(t)]} \mathbb{E}\left[\xi_{i,j} \Delta_{i,j} W \middle| \mathcal{F}_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)} \right] \\ &= \sum_{i=1}^{[n\varphi_1(t)]} \sum_{j=1}^{[n\varphi_2(t)]} f\left(W_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)} \right) \mathbb{E}\left[\left(|\Delta_{i,j} W|^2 - \frac{1}{n^2} \right) \Delta_{i,j} W \middle| \mathcal{F}_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)} \right] \\ &= 0. \end{aligned}$$

□

3.1.4 Estimation of the quadratic variation and asymptotic normality of the estimator

In this section we prove an asymptotic normality property (Corollary 3.1.11) for the consistent estimator V^n (see (3.1.4)) of the quadratic variation C (defined in (3.1.5)).

Consider the following two-parameter stochastic process

$$Y_{(s,t)} := \int_{[0,s]\times[0,t]} \sigma(W_\rho)\, dW_\rho + \int_{[0,s]\times[0,t]} M_\rho\, d\rho, \quad (s,t) \in [0,1]^2 \tag{3.1.24}$$

defined on a probability basis $\left(\Omega, \mathcal{F}, (\mathcal{F}_{(s,t)})_{(s,t)\in[0,1]^2}, \mathbb{P}\right)$. Until the end of this paper we assume that $(M_{(s,t)})_{(s,t)\in[0,1]^2}$ is a continuous and $(\mathcal{F}_{(s,t)})_{(s,t)\in[0,1]^2}$-adapted process. The following assumption will be used in the results presented below.

Hypothesis $(\mathbf{H}_q)$:
$\sigma : \mathbb{R} \to \mathbb{R}$ is in $\mathscr{C}^2(\mathbb{R})$ with,

$$\sup_{(s,t)\in[0,1]^2} \mathbb{E}\left[\left|\sigma^{(a)}\left(W_{(s,t)}\right)\right|^p\right] < \infty,\ p \in (0, 4q],\ a = 0,1,2, \quad \text{with the convention } \sigma^{(0)} := \sigma.$$

For $n \geq 1$, let

$$\begin{aligned} Y^n_{(s,t)} &:= n\left(V^n_{(s,t)} - C_{(s,t)}\right) \\ &= n\left(\sum_{i=1}^{[ns]}\sum_{j=1}^{[nt]} |\Delta_{i,j}Y|^2 - \int_{[0,s]\times[0,t]} \sigma^2\left(W_{(u,v)}\right)\, dudv\right). \end{aligned} \tag{3.1.25}$$

In order to prove the asymptotic normality of V^n in Corollary 3.1.11 we need the following theorem.

Theorem 3.1.7. *Let X be the process defined by (3.1.9) where f in (3.1.9) is replaced by σ^2. Under assumption $(\mathbf{H}_1)$, $(Y^n)_{n\geq1}$ defined by (3.1.25) converges $\mathcal{F}$-stably in law in the Skorohod space $(\mathcal{D}([0,1]^2), d, \mathcal{L}_2)$ to the non-Gaussian continuous process X defined on the extension $\tilde{\mathcal{B}}$ described in the proof of Theorem 3.1.3.*

Remark 3.1.8. *Note that X above is not a Gaussian process but it is conditionally Gaussian given the filtration generated by W.*

Proof. Using a localization argument the finite variation part of Y has no contribution in the limit. So we can assume that $M = 0$.
From Theorem 3.1.3 with $f = \sigma^2$, the process $(X^n)_n$ converges $\mathcal{F}$-stably in law to X with

$$X^n_{(s,t)} = n\sum_{i=1}^{[ns]}\sum_{j=1}^{[nt]} \sigma^2\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)\left(|\Delta_{i,j}W|^2 - \frac{1}{n^2}\right), \quad (s,t) \in [0,1]^2.$$

To conclude the proof we show that Y^n is equal to X^n plus a term r_n which becomes negligible when n to infinity. More precisely using the notations

$$\begin{cases} \eta_{i,j} := \left(\int_{\Delta_{i,j}} \sigma\left(W_u\right) dW_u\right)^2 - \sigma^2\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) |\Delta_{i,j}W|^2 \\ \eta'_{i,j} := -\int_{\Delta_{i,j}} \sigma^2\left(W_\rho\right) - \sigma^2\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) d\rho \\ 1 \le i,j \le n, \ n \ge 1, \end{cases}$$

r_n can be decomposed as

$$r_n(s,t) := r_n^{(1)}(s,t) + r_n^{(2)}(s,t) \tag{3.1.26}$$

where

$$\begin{cases} r_n^{(1)}(s,t) := \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \eta_{i,j} + \eta'_{i,j}, \\ r_n^{(2)}(s,t) := -\int_{[ns]/n}^{s} \int_{[nt]/n}^{t} \sigma^2\left(W_\rho\right) d\rho. \end{cases}$$

Using a standard argument of the form [18, Theorem 3.1] and Theorem 3.1.3 the proof is finished if we show that $n \sup_{(s,t)\in[0,1]^2} |r_n(s,t)|$ converges in probability to zero. We use the decomposition (3.1.26) and we show

$$n \sup_{(s,t)\in[0,1]^2} |r_n^{(1)}(s,t)| \xrightarrow{\mathbb{P}} 0, \tag{3.1.27}$$

$$n \sup_{(s,t)\in[0,1]^2} |r_n^{(2)}(s,t)| \xrightarrow{\mathbb{P}} 0. \tag{3.1.28}$$

Proof of (3.1.27):
Using Burkholder's inequality for two-parameter martingales (see Remark 2 of [91]) it is enough to show that

$$\mathbb{E}\left[\left|\eta_{i,j} + \eta'_{i,j}\right|^2\right] \le \frac{C}{n^5}, \tag{3.1.29}$$

where C is a constant.
The tool used here is the Malliavin calculus (see Appendix 3.1.5) and especially the Malliavin integration by parts formula (3.1.35). The main problem comes from the computation of $\mathbb{E}\left[|\eta_{i,j}|^2\right]$. First we express $\eta_{i,j}$ as,

$$\begin{aligned} \eta_{i,j} &= \left(\int_{\Delta_{i,j}} \sigma\left(W_\rho\right) - \sigma\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) dW_\rho\right) \left(\int_{\Delta_{i,j}} \sigma\left(W_\rho\right) + \sigma\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) dW_\rho\right) \\ &= \delta(u)\delta(v), \end{aligned}$$

where

$$\begin{cases} u_{(s,t)} := \mathbf{1}_{\Delta_{i,j}}(s,t)\left(\sigma\left(W_{(s,t)}\right) - \sigma\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)\right) \\ v_{(s,t)} := \mathbf{1}_{\Delta_{i,j}}(s,t)\left(\sigma\left(W_{(s,t)}\right) + \sigma\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)\right) \end{cases}, \quad (s,t) \in [0,1]^2.$$

$\delta(u)$ denotes the Skorohod integral of the process $(u_{(s,t)})_{(s,t)\in[0,1]^2}$ which coincides since u is adapted with the Itô stochastic integral. Consequently,

$$\begin{aligned}
& \mathbb{E}\left[|\eta_{i,j}|^2\right] \\
= \;& \mathbb{E}[\delta(u)\left(\delta(u)\,\delta(v)^2\right)] \\
= \;& \mathbb{E}\left[\langle u,\, D(\delta(u)\delta(v)^2)\rangle_{L^2([0,1]^2)}\right], \quad \text{by (3.1.35)} \\
= \;& \int_{\Delta_{i,j}} \mathbb{E}\left[u_{(s,t)}\, D_{(s,t)}(\delta(u)\,\delta(v)^2)\right] dsdt \\
= \;& \int_{\Delta_{i,j}} \mathbb{E}\left[u_{(s,t)}\,\delta(v)^2\, D_{(s,t)}(\delta(u))\right] dsdt + 2\int_{\Delta_{i,j}} \mathbb{E}\left[u_{(s,t)}\,\delta(u)\,\delta(v)\, D_{(s,t)}(\delta(v))\right] dsdt, \\
= \;& \int_{\Delta_{i,j}} \mathbb{E}\left[u^2_{(s,t)}\delta(v)^2\right] dsdt + \int_{\Delta_{i,j}} \mathbb{E}\left[\delta(D_{(s,t)}u)\delta(v)^2\right] dsdt \\
& +2\int_{\Delta_{i,j}} \mathbb{E}\left[u_{(s,t)}v_{(s,t)}\,\delta(u)\,\delta(v)\right] dsdt \\
& +2\int_{\Delta_{i,j}} \mathbb{E}\left[u_{(s,t)}\,\delta(u)\,\delta(v)\,\delta(D_{(s,t)}v)\right] dsdt \qquad (3.1.30)
\end{aligned}$$

where the last equality is deduced from the chain rule formula (3.1.37) and from the "Heisenberg commutativity relationship"(3.1.38). We compute the first term of (3.1.30) since the estimates for the other terms can be derived from mimicking the following computations. Using Malliavin calculus arguments already mentioned above we obtain that,

$$\begin{aligned}
& \int_{\Delta_{i,j}} \mathbb{E}\left[u^2_{(s,t)}\delta(v)^2\right] dsdt \\
= \;& \int_{\Delta_{i,j}}\int_{\Delta_{i,j}} \mathbb{E}\left[v_{(a,b)}D_{(a,b)}(u^2_{(s,t)}\delta(v))\right] dadb\, dsdt \\
= \;& 2\int_{\Delta_{i,j}}\int_{\Delta_{i,j}} \mathbb{E}\left[v_{(a,b)}u_{(s,t)}\delta(v)D_{(a,b)}(u_{(s,t)})\right] dadb\, dsdt + \int_{\Delta_{i,j}}\int_{\Delta_{i,j}} \mathbb{E}\left[v^2_{(a,b)}u^2_{(s,t)}\right] dadb\, dsdt \\
& + \int_{\Delta_{i,j}}\int_{\Delta_{i,j}} \mathbb{E}\left[v_{(a,b)}u^2_{(s,t)}\delta(D_{(a,b)}v)\right] dadb\, dsdt. \qquad (3.1.31)
\end{aligned}$$

Using Cauchy-Schwarz inequality we have that,

$$\left|\mathbb{E}\left[v_{(a,b)}u_{(s,t)}\delta(v)D_{(a,b)}(u_{(s,t)})\right]\right| \le \mathbb{E}\left[|v_{(a,b)}u_{(s,t)}D_{(a,b)}(u_{(s,t)})|^2\right]^{1/2} \mathbb{E}\left[\delta(v)^2\right]^{1/2}$$

which leads to

$$\left|\int_{\Delta_{i,j}}\int_{\Delta_{i,j}} \mathbb{E}\left[v_{(a,b)}u_{(s,t)}\delta(v)D_{(a,b)}(u_{(s,t)})\right] dadb\, dsdt\right| \le \frac{C}{n^5},$$

by Itô isometry (3.1.36) and by assumption ($\mathbf{H}_2$). The same method is valid for the two remaining terms in (3.1.31). Thus we obtain

$$\mathbb{E}\left[|\eta_{i,j}|^2\right] \le \frac{C}{n^5}.$$

Proof of (3.1.28):
Recall that $|s-[ns]| \leq 1/n$ and use assumption $(\mathbf{H}_2)$. □

Now we can show that V^n is a consistent estimator of C.

Proposition 3.1.9. *Under hypothesis ($\boldsymbol{H}_1$) the estimator V^n defined in (3.1.4) of the quadratic variation C (3.1.5) is consistent. That is, for every (s,t) in $[0,1]^2$,*

$$V^n_{(s,t)} \xrightarrow[n\to\infty]{L^2} C_{(s,t)}.$$

Proof. Once again in this section C denotes a constant which can differ from one line to another. Fix (s,t) in $[0,1]^2$. We show $(V^n_{(s,t)})_n$ converges to $C_{(s,t)}$ in $L^2(\Omega,\mathcal{F},\mathbb{P})$. Regarding the approximations realized in the proof of Theorem 3.1.7 it is enough to show that

$$\sum_{i=1}^{[ns]}\sum_{j=1}^{[nt]} \sigma^2\left(W_{(\frac{i-1}{n},\frac{j-1}{n})}\right)\left(|\Delta_{i,j}W|^2 - \frac{1}{n^2}\right)$$

tends to zero in $L^2(\Omega,\mathcal{F},\mathbb{P})$ as n goes to infinity. We have,

$$\begin{aligned}
&\mathbb{E}\left[\left|\sum_{i=1}^{[ns]}\sum_{j=1}^{[nt]} \sigma^2\left(W_{(\frac{i-1}{n},\frac{j-1}{n})}\right)\left(|\Delta_{i,j}W|^2 - \frac{1}{n^2}\right)\right|^2\right]\\
=\;& \sum_{i=1}^{[ns]}\sum_{j=1}^{[nt]} \mathbb{E}\left[\left|\sigma^2\left(W_{(\frac{i-1}{n},\frac{j-1}{n})}\right)\left(|\Delta_{i,j}W|^2 - \frac{1}{n^2}\right)\right|^2\right]\\
=\;& \sum_{i=1}^{[ns]}\sum_{j=1}^{[nt]} \mathbb{E}\left[\sigma^4\left(W_{(\frac{i-1}{n},\frac{j-1}{n})}\right)\mathbb{E}\left[\left||\Delta_{i,j}W|^2 - \frac{1}{n^2}\right|^2 \Big| \mathcal{F}_{(\frac{i-1}{n},\frac{j-1}{n})}\right]\right]\\
=\;& \sum_{i=1}^{[ns]}\sum_{j=1}^{[nt]} \frac{1}{n^4}\,\mathbb{E}\left[\sigma^4\left(W_{(\frac{i-1}{n},\frac{j-1}{n})}\right)\right]\mathbb{E}\left[\left||n\Delta_{i,j}W|^2 - 1\right|^2\right]\\
\leq\;& \frac{C}{n^2}, \quad \text{by using hypothesis } (\mathbf{H}_1).
\end{aligned}$$

□

Remark 3.1.10. *We have chosen to use the Malliavin calculus which leads to a short proof due to the specific form (3.1.24) of the process Y. We could also have used the two-parameter stochastic calculus techniques developed in [60, 91, 92, 120] which are valid for more general processes however with a longer argument.*

We state and prove that the estimator V^n of C is asymptotically normal.

Corollary 3.1.11 (Asymptotic normality)**.** *For (s,t) in $[0,1]^2$, let $S^n_{(s,t)}$ and $S_{(s,t)}$ be*

$$S^n_{(s,t)} := n^2 \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} |\Delta_{i,j} Y|^4 , \quad n \geq 1 \text{ and } S_{(s,t)} := 3 \int_{[0,s]\times[0,t]} \sigma^4 (W_\rho)\, d\rho.$$

Let (s,t) fixed in $(0,1]^2$ such that $S^n_{(s,t)}$ and $S_{(s,t)}$ don't vanish. Then, under hypothesis $(\mathbf{H}_2)$ *we have,*

$$\left(S^n_{(s,t)}\right)^{-\frac{1}{2}} n\left(V^n_{(s,t)} - C_{(s,t)}\right) \xrightarrow[n\to\infty]{law} \sqrt{\frac{2}{3}} N, \quad N \sim \mathcal{N}(0,1).$$

Proof. Using a localization argument, only the stochastic integral part of Y gives a contribution to the limit so we assume $M = 0$ in (3.1.24). The main argument of the proof is the convergence in law of $(S^n, Y^n)_n$ to (S, X) where X is defined by (3.1.9) with f replaced by σ^2. Actually assume this convergence holds. Since $(x,y) \mapsto x^{-\frac{1}{2}} y$ is continuous on $\mathbb{R}_+^* \times \mathbb{R}$ we have for every (s,t) in $(0,1]^2$,

$$\left(S^n_{(s,t)}\right)^{-\frac{1}{2}} Y^n_{(s,t)} \xrightarrow[n\to\infty]{law} \sqrt{\frac{2}{3}} \frac{\int_{[0,s]\times[0,t]} \sigma^2 (W_\rho)\, dB_\rho}{\left(\int_{[0,s]\times[0,t]} \sigma^4 (W_\rho)\, d\rho\right)^{\frac{1}{2}}}.$$

Computing the characteristic function with respect to the probability measure $\tilde{\mathbb{P}}$ we can show that

$$\sqrt{\frac{2}{3}} \frac{\int_{[0,s]\times[0,t]} \sigma^2 (W_\rho)\, dB_\rho}{\left(\int_{[0,s]\times[0,t]} \sigma^4 (W_\rho)\, d\rho\right)^{\frac{1}{2}}} \overset{law}{=} \sqrt{\frac{2}{3}} N, \quad N \sim \mathcal{N}(0,1).$$

We have now to show that $(S^n, Y^n)_n$ converges in law to (S, X). The key point is the $\mathcal{F}$-stable convergence in law of $(Y^n)_n$ to X obtained in Theorem 3.1.7 stated and proved at the end of this section. Using a result of Aldous and Eagleson (presented in [4]) concerning stable convergence in law if $(S^n)_n$ converges in $\mathbb{P}$-probability to S then $(S^n, Y^n)_n$ converges in law to (S, X) (and the convergence is even $\mathcal{F}$-stable convergence in law).
Let us finally show that $(S^n)_n$ converges in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ to S.
First we show that

$$\mathbb{E}\left[\left| n^2 \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \sigma^4 \left(W_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)}\right) \left(|\Delta_{i,j} W|^4 - \frac{3}{n^4}\right)\right|^2\right] \xrightarrow[n\to\infty]{} 0.$$

Actually for every (s,t) in $[0,1]^2$,

$$\mathbb{E}\left[\left| n^2 \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \sigma^4 \left(W_{\left(\frac{i-1}{n}, \frac{j-1}{n}\right)}\right) \left(|\Delta_{i,j} W|^4 - \frac{3}{n^4}\right)\right|^2\right]$$

$$
\begin{aligned}
&= n^4 \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \mathbb{E}\left[\sigma^8\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)\left|\left|\Delta_{i,j}W\right|^4 - \frac{3}{n^4}\right|^2\right] \\
&= n^4 \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \mathbb{E}\left[\sigma^8\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) E\left[\left|\left|\Delta_{i,j}W\right|^4 - \frac{3}{n^4}\right|^2 \middle| \mathcal{F}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right]\right] \\
&\leq n^4 \sup_{(a,b)\in[0,1]^2} \mathbb{E}\left[\sigma^8(W_{(a,b)})\right] \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} E\left[\left|\left|\Delta_{i,j}W\right|^4 - \frac{3}{n^4}\right|^2\right] \\
&\leq \frac{C}{n^2}, \quad \text{by assumption } (\mathbf{H}_2).
\end{aligned}
$$

Using a Riemann approximation for integrals we have that

$$
3n^2 \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \frac{\sigma^4\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)}{n^4} \underset{n\to\infty}{\overset{L^2}{\longrightarrow}} 3\int_{[0,s]\times[0,t]} \sigma^4\left(W_\rho\right) d\rho, \quad \forall (s,t) \in [0,1]^2.
$$

The proof is finished if we can show the estimate (3.1.32).

$$
\mathbb{E}\left[\left|n^2 \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \left(\left|\Delta_{i,j}Y\right|^4 - \frac{\sigma^4\left(W_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)}{n^4}\right)\right|^2\right] \underset{n\to\infty}{\longrightarrow} 0. \tag{3.1.32}
$$

This goal is achieved by adapting arguments of Theorem 3.1.7. Since we have to mimic some estimates done in Theorem 3.1.7 with σ^2 replaced with σ^4 we have to assume that σ fulfill hypothesis $(\mathbf{H}_2)$ and not only assumption $(\mathbf{H}_1)$.

□

Using classical techniques the asymptotic normality property of V^n enable construction of confidence interval for $C_{(s,t)}$ for every (s,t) in $[0,1]^2$.

3.1.5 Appendix

In this section we present some definitions and results used in Sections 3.1.3 and 3.1.4. First we provide some background on set-indexed processes. Then we briefly present the Malliavin calculus for two-parameter Brownian motion, including the Malliavin integration by parts formula which has been used to obtain the estimates of the previous section.

Set-indexed processes

In the following definition it will be convenient to think of two-parameter processes on $[0,1]^2$ as set-indexed processes on $\mathcal{A} := \{[0,z],\ z \in [0,1]^2\}$ where $Y_{[0,z]} := Y_z$ and $\mathcal{F}_{[0,z]} := \mathcal{F}_z$, z in $[0,1]^2$. We will use indifferently one of these two points of view.

Definition 3.1.12. *([62, Definition 7.3.1]) Let $\mathcal{A}(u)$ be the set of all finite unions of sets from $\mathcal{A}$. A simple flow φ is an application $\varphi : [0,1] \to \mathcal{A}(u)$ which satisfies the following properties,*

i) φ *is increasing,*

ii) φ *is continuous,*

iii) $\varphi(0) = \emptyset$,

iv) *for some k in $\mathbb{N}$ there exists some increasing functions $\varphi_i : [0,1] \to \mathcal{A}$, $i = 1, \ldots, k$ such that for $\frac{i-1}{k} \leq s \leq \frac{i}{k}$, $1 \leq i \leq k$*

$$\varphi(s) = \varphi\left(\frac{i-1}{k}\right) \cup \varphi_i(s).$$

Definition 3.1.13. *([62, Definition 1.4.5]) Let $\mathcal{C}$ denote the set of elements of the form $A \setminus B$ with A in $\mathcal{A}$ and B in $\mathcal{A}(u)$ (the set of all finite unions of elements of $\mathcal{A}$).*
A process $(X_z)_{z \in [0,1]^2}$ identified with $(X_A)_{A \in \mathcal{A}}$ is called additive if for every C, C_1, C_2 in $\mathcal{C}$ with $C = C_1 \cup C_2$ and $C_1 \cap C_2 = \emptyset$ we have,

$$X_C = X_{C_1} + X_{C_2}.$$

We recall a particular case of [62, Lemma 5.1.2].

Lemma 3.1.14. *Let X be a strong martingale and let φ be a $\mathcal{A}(u)$-valued simple flow defined on $[0,1]$. Then the one-parameter process $(X_{\varphi(s)})_{s \in [0,1]}$ is a martingale with respect to the filtration $(\mathcal{F}_{\varphi(s)})_{s \in [0,1]}$.*

Extension of a probability basis

We give the definition of a *very good extension* of a probability basis which has been introduced in [65]. The following definition is an adaptation of [67, Definition II.7.1].

Definition 3.1.15. *A probability basis $\tilde{\mathcal{B}} = (\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_z)_{z \in [0,1]^2}, \tilde{\mathbb{P}})$ is an extension of the probability basis $\mathcal{B} = (\Omega, \mathcal{F}, (\mathcal{F}_z)_{z \in [0,1]^2}, \mathbb{P})$ if there exists an auxiliary probability basis $\mathcal{B}' = (\Omega', \mathcal{F}', (\mathcal{F}'_z)_{z \in [0,1]^2}, \mathbb{P}')$ such that*

$$\begin{cases} \tilde{\Omega} = \Omega \times \Omega', \\ \tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}', \\ \tilde{\mathcal{F}}_z = \cap_{\rho \preceq z} \mathcal{F}_\rho \otimes \tilde{\mathcal{F}}'_z, \ z \in [0,1]^2, \\ \tilde{\mathbb{P}}(d\omega, d\omega') = \mathbb{P}(d\omega) \mathbb{Q}_\omega(d\omega'), \end{cases}$$

where $\mathbb{Q}_\omega(d\omega')$ is a transition probability from $(\Omega, \mathcal{F})$ into $(\Omega', \mathcal{F}')$.
This extension is called very good is for every z in $[0,1]^2$ and for all element A' in $\mathcal{F}'_z$ $\omega \mapsto \mathbb{Q}_\omega(A')$ is equal $\mathbb{P}$-a.s. to an $\mathcal{F}_z$-measurable random variable.

Notation:
If $(X_z)_{z\in[0,1]^2}$ *is a stochastic process on* $\mathcal{B}$ *we will denote by* $(X_z)_{z\in[0,1]^2}$ *again the stochastic process defined on* $\tilde{\mathcal{B}}$ *by*

$$X_z(\omega,\omega') := X_z(\omega), \quad z\in[0,1]^2,\ (\omega,\omega')\in\tilde{\Omega}.$$

Lemma 3.1.16. *As in [67, Lemma II.7.3], under* ***Assumption (CI)*** *an extension* $\tilde{\mathcal{B}}$ *is very good if and only if every martingale on* $\mathcal{B}$ *is a martingale on* $\tilde{\mathcal{B}}$.

The proof of this Lemma is similar to its one-parameter counterpart [67, Lemma II.7.3].

Malliavin calculus for two-parameter Brownian motion

Let $(W_{(s,t)})_{(s,t)\in[0,1]^2}$ be a two-parameter Brownian motion defined on a probability basis $\left(\Omega,\mathcal{F},(\mathcal{F}_z)_{z\in[0,1]^2},\mathbb{P}\right)$. This process is a centered Gaussian process whose covariance given by

$$\mathbb{E}\left[W_{(s,t)}W_{(s',t')}\right] = (s\wedge s')\,(t\wedge t'), \quad (s,t)\in[0,1]^2,\ (s',t')\in[0,1]^2.$$

The Malliavin calculus for general Gaussian processes has been described in [94] and the reader can refer to it for a complete explanation about this topic. Here we give the definition of the Malliavin derivative and we present the integration by parts formula which is extensively used in Section 3.1.3.

Definition 3.1.17. *Let* $\mathcal{S}$ *be the space of random variable* F *of the form*

$$F = f(W(h_1),\dots,W(h_n)), \tag{3.1.33}$$

where h_i *is an element of* $L^2([0,1]^2,dz)$ *and* $W(h_i)$ *denotes the stochastic integral* $W(h_i) := \int_{[0,1]^2} h_i(z)\,dz$ *for* $i=1,\cdots,n$ *and* $f:\mathbb{R}^n\to\mathbb{R}$ *is infinitely continuously differentiable.*
For F *of the form (3.1.33) we define the Malliavin derivative* DF *of* F *as the following* $L^2([0,1]^2,dz)$*-valued random variable*

$$DF := \sum_{i=1}^n \partial_i f(W(h_i))\,h_i. \tag{3.1.34}$$

Here we give the Malliavin integration by parts formula ([94, Lemma 1.2.1]).

Lemma 3.1.18. *Let* F *in* $\mathcal{S}$ *and* $(u_{(s,t)})_{(s,t)\in[0,1]^2}$ *in the domain of the divergence operator* δ. *We have*

$$\mathbb{E}\left[F\delta(u)\right] = \mathbb{E}\left[\langle DF,\delta(u)\rangle_{L^2([0,1]^2,dz)}\right]. \tag{3.1.35}$$

δ *is called the divergence operator (or the Skorohod integral of* u*) and it extends the Itô integral since when* u *is adapted to the filtration generated by the two-parameter Brownian motion* W,

$$\delta(u) = \int_{[0,1]^2} u_{(s,t)}\,dW_{(s,t)}.$$

Moreover we recall the Itô isometry

$$\mathbb{E}[\delta(u)^2] = \|u\|^2_{L^2(\Omega\times[0,1])}. \tag{3.1.36}$$

Note also that the Malliavin derivative D is a closable operator (see [94, Proposition 1.2.1]) and we denote by $Dom(D)$ its domain. Furthermore D satisfies a chain rule property that is, for every F and G elements of $Dom(D)$ such that $F\,G$ belongs to $Dom(D)$ we have

$$D(FG) = F\,DG + DF\,G. \tag{3.1.37}$$

We end this section with the "Heisenberg commutativity relationship"which enables the computation of the gradient of a Itô stochastic integral, more precisely we have for a process u such that the right hand side of (3.1.38) is well-defined (more details about assumption on u can be found in [94, (1.46)]),

$$D_{(s,t)}\delta(u) = u_{(s,t)} + \delta(D_{(s,t)}u), \quad (s,t) \in [0,1]^2. \tag{3.1.38}$$

Acknowledgement

I would like to thank the two referees for several helpful comments and remarks which led to majors improvements of this paper.
Financial support of the *European Social Fund (ESF)* is acknowledged.

3.2 Théorèmes limites pour les variations quadratiques à poids de certains draps browniens fractionnaires

Cette section est publiée en [116].

3.2.1 Introduction

Recently in several works, the asymptotic behavior of the weighted p-power variations of stochastic processes has been investigated. For a one-parameter stochastic process $(Z_t)_{t\in[0,1]}$ observed at times $\{i/n,\ 0 \le i \le n\}$ it is defined as

$$\sum_{i=1}^{n} f\left(Z_{\frac{i-1}{n}}\right)\left(Z_{\frac{i}{n}} - Z_{\frac{i-1}{n}}\right)^{p}, \tag{3.2.1}$$

and one is interested by the convergence of this quantity as n tends to infinity and by the nature of this convergence. Power variations play an important role in the recovering of statistical properties of discretely observed stochastic processes and as an example, the results obtained in this area have been applied in some financial econometric problems (see [9]) and in mathematical finance (*e.g.* in [2, 64, 63]). In these settings the considered stochastic processes were one-parameter Itô semimartingales.

In this paper we consider a different situation since we consider two-parameter processes which are not semimartingales. Actually, we state and prove a central limit theorem (Theorem 3.2.2) for the finite-dimensional laws of the weighted quadratic variations process of certain fractional Brownian sheets. More precisely, let $(W^{\alpha,\beta}_{(s,t)})_{(s,t)\in[0,1]^2}$ be a fractional Brownian sheet with Hurst indices α and β such that $0 < \alpha < \frac{1}{2}$, $0 < \beta < \frac{1}{2}$ with $\alpha+\beta > \frac{1}{2}$ then for a weight function $f : \mathbb{R} \to \mathbb{R}$ satisfying hypothesis **(H)** presented below, we have,

$$n^{-1}\sum_{i=1}^{[n\cdot]}\sum_{j=1}^{[n\bullet]} f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)\left(n^{2(\alpha+\beta)}|\Delta_{i,j}W^{\alpha,\beta}|^2 - 1\right) \xrightarrow[n\to\infty]{fdd} \sigma_{\alpha,\beta}\int_0^{\cdot}\int_0^{\bullet} f\left(W^{\alpha,\beta}_{(u,v)}\right) dW_{(u,v)} \tag{3.2.2}$$

where W is a standard Brownian sheet independent of $W^{\alpha,\beta}$ and the notation fdd means that the convergence is in the sense of stable convergence of the finite-dimensional laws (see (3.2.8)). Note that the constant $\sigma_{\alpha,\beta}$ appearing in the limiting process can be expressed explicitly with respect to α and β.

The study of p-power variations for fractional Brownian motion has been initiated by Gradinaru and Nourdin in [50], Neuenkirch and Nourdin in [78] and by Nourdin in [82] in view of obtaining exact rate of convergence of some approximating scheme of scalar stochastic differential equations driven by a fractional Brownian motion. This study has been recently pursued by Nourdin in [81], by Nourdin and Nualart in [83] and by Nourdin, Nualart and Tudor in [84] where more references about this topic can be found. A complete description of the nature of the convergence of weighted p-power variations of the form (3.2.1) for a fractional Brownian motion B with Hurst index H is given in [84, Theorem 1]. More precisely central and non-central limit theorems are derived depending on the values of p and H. Concerning the particular case of weighted quadratic variations ($p = 2$) it is shown in [84, Theorem 1] that for $\frac{1}{4} < H < \frac{3}{4}$,

$$n^{-1/2} \sum_{i=1}^{n} f\left(B_{\frac{i-1}{n}}\right) \left(n^{2H} |B_{\frac{i}{n}} - B_{\frac{i-1}{n}}|^2 - 1\right) \xrightarrow[n\to\infty]{\mathcal{L}} \sigma_H \int_0^1 f(B_s) dW_s,$$

where σ_H is an explicit constant depending on H and W is a standard Brownian motion independent of B. Note this result was also obtained in [83] when $\frac{1}{4} < H < \frac{1}{2}$. For the two-parameter case a central limit theorem has been obtained in [117] for the weighted quadratic variations process of a standard Brownian sheet and applied to the construction of an asymptotically normal estimator of the quadratic variation of a two-parameter diffusion process. However, in [117], Réveillac obtained the stable finite-dimensional convergence in law by using a result based on some semimartingales techniques developed by Jacod and Shiryaev in [67]. Consequently this method is useless for fractional Brownian sheets, that is why we propose to replace it by adapting an argument presented in [83] based on the Malliavin calculus which is valid in general Gaussian context. In order to make this paper self-contained we have adapted and reproduced the main computations collected in Section 3.2.4 originally realized by Nourdin and Nualart in [83].

We proceed as follows. In Section 3.2.2 we recall some definitions and properties of the fractional Brownian sheet and some elements of the Malliavin calculus relative to this process. Then in Section 3.2.3 we state and prove the convergence of the finite-dimensional laws of the weighted quadratic variations of certain fractional Brownian sheets (Theorem 3.2.2). Technical arguments used in Section 3.2.3 are collected in Section 3.2.4.

3.2.2 Preliminaries and notations

In this section we recall the definition of the fractional Brownian sheet and we present some elements of Malliavin calculus.

Several extensions of the fractional Brownian motion have been proposed in the literature as for example the *fractional Brownian field* (*c.f.* [76, 19]), the *Lévy's fractional Brownian field* (*c.f.* [29]) and the *fractional Brownian sheet* (*c.f.* [70, 7]) we consider in this paper. The definitions and properties of this section can be found in [5, 129].

Definition 3.2.1 (Fractional Brownian sheet). *A fractional Brownian sheet* $(W^{\alpha,\beta}_{(s,t)})_{(s,t)\in[0,1]^2}$ *with Hurst indices* $(\alpha, \beta) \in (0, 1)^2$ *is a centered two-parameter Gaussian process equal to zero on the set*

$$\{(s, t) \in [0, 1]^2,\ s = 0 \text{ or } t = 0\}$$

whose covariance function is given by,

$$\begin{aligned}
R^{\alpha,\beta}((s_1,t_1),(s_2,t_2)) &:= \mathbb{E}\left[W^{\alpha,\beta}_{(s_1,t_1)}W^{\alpha,\beta}_{(s_2,t_2)}\right] \\
&= K^{\alpha}(s_1,s_2)K^{\beta}(t_1,t_2) \\
&= \frac{1}{2}\big(s_1^{2\alpha}+s_2^{2\alpha}-|s_1-s_2|^{2\alpha}\big)\frac{1}{2}\big(t_1^{2\beta}+t_2^{2\beta}-|t_1-t_2|^{2\beta}\big).
\end{aligned}$$

We assume that $(W^{\alpha,\beta}_{(s,t)})_{(s,t)\in[0,1]^2}$ is defined on a complete probability space $(\Omega,\mathcal{F},\mathbb{P})$ where $\mathcal{F}$ is generated by $W^{\alpha,\beta}$. Let $\mathcal{H}$ be the closure of the linear span generated by indicator functions on $[0,1]^2$ with respect to the scalar product

$$\langle \mathbf{1}_{[0,s_1]\times[0,t_1]},\mathbf{1}_{[0,s_2]\times[0,t_2]}\rangle_{\mathcal{H}} = R^{\alpha,\beta}((s_1,t_1),(s_2,t_2)).$$

The mapping $\mathbf{1}_{[0,s]\times[0,t]}\mapsto W^{\alpha,\beta}_{(s,t)}$ provides an isometry between $\mathcal{H}$ and the first chaos $H_1^{\alpha,\beta}$. For an element φ of $\mathcal{H}$ we denote by $W^{\alpha,\beta}(\varphi)$ the image of φ in $H_1^{\alpha,\beta}$.
Note that we can also give a representation of $(W^{\alpha,\beta}_{(s,t)})_{(s,t)\in[0,1]^2}$ as a stochastic integral of kernels $\mathcal{K}^\alpha$ and $\mathcal{K}^\beta$ with respect to a standard Brownian sheet $(W_{(s,t)})_{(s,t)\in[0,1]^2}$:

$$W^{\alpha,\beta}_{(s,t)} = \int_0^s\int_0^t \mathcal{K}^\alpha(s,u)\mathcal{K}^\beta(t,v)\,dW_{(u,v)},\quad (s,t)\in[0,1]^2.$$

Using this representation, Tudor and Viens in [129, 130] have developed a Malliavin calculus with respect to $W^{\alpha,\beta}$. Note that in [130] Tudor and Viens have also given an extension of the divergence integral. Now we present some elements of Malliavin calculus with respect to fractional Brownian sheets and especially the Malliavin integration by parts formula (3.2.4).

We recall some definitions and properties of the Malliavin calculus for the fractional Brownian sheet. These elements are contained in the general framework described in [94] for Gaussian processes.

For a cylindrical functional F of the form

$$F = f\left(W^{\alpha,\beta}(\varphi_1),\ldots,W^{\alpha,\beta}(\varphi_n)\right),\quad n\geq 1,\ \varphi_1,\ldots,\varphi_n\in\mathcal{H},\ f\in\mathcal{C}_b^\infty(\mathbb{R}^n), \tag{3.2.3}$$

we define the Malliavin derivative DF of F as,

$$DF := \sum_{i=1}^n \partial_i f\left(W^{\alpha,\beta}(\varphi_1),\ldots,W^{\alpha,\beta}(\varphi_n)\right)\varphi_i.$$

Furthermore $D: L^2(\Omega,\mathcal{F},\mathbb{P})\to L^2(\Omega,\mathcal{F},\mathbb{P};\mathcal{H})$ is a closable operator and it can be extended to the closure of Sobolev space $\mathbb{D}^{1,2}$ defined by functional F whose norm $\|F\|_{1,2}$ is finite with,

$$\|F\|_{1,2}^2 := \mathbb{E}\left[F^2\right] + \|DF\|^2_{L^2(\Omega,\mathcal{F},\mathbb{P};\mathcal{H})}.$$

The adjoint operator I_1 of D is defined by the following duality relationship

$$\mathbb{E}\left[FI_1(u)\right] = \mathbb{E}\left[\langle DF, u\rangle_{\mathcal{H}}\right],$$

for F in $\mathbb{D}^{1,2}$ and for u in $\mathcal{H}$ such that there exists $c_u > 0$ verifying

$$|\mathbb{E}\left[\langle DG, u\rangle_{\mathcal{H}}\right]| \leq c_u \|G\|_{L^2(\Omega,\mathcal{F},\mathbb{P})}, \qquad \text{for every functional } G \text{ of the form (3.2.3).}$$

Let $n \geq 1$. The nth Wiener chaos $\mathfrak{H}_n$ of $W^{\alpha,\beta}$ is the closed linear subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$ generated by the random variables $\left\{H_n\left(W^{\alpha,\beta}(\varphi)\right),\ \varphi \in \mathcal{H},\ \|\varphi\|_{\mathcal{H}} = 1\right\}$ where H_n denotes the nth Hermite polynomial. A linear isometry between the symmetric tensor product $\mathcal{H}^{\odot n}$ and $\mathfrak{H}_n$ is defined as,

$$I_n\left(\varphi^{\otimes n}\right) := n! H_n\left(W^{\alpha,\beta}(\varphi)\right).$$

We conclude this section by the following integration by parts formula:

$$\mathbb{E}\left[FI_n(h)\right] = \mathbb{E}\left[\langle D^n F, h\rangle_{\mathcal{H}^{\otimes n}}\right], \quad h \in \mathcal{H}^{\odot n},\ F \in \mathbb{D}^{n,2}, \tag{3.2.4}$$

where $\mathbb{D}^{n,2}$ is the space of functionals F such that $\|F\|_{n,2}$ is finite with

$$\|F\|_{n,2}^2 := \mathbb{E}\left[F^2\right] + \sum_{i=1}^{n} \|D^i F\|^2_{L^2(\Omega,\mathcal{F},\mathbb{P};\mathcal{H})}.$$

3.2.3 Convergence of finite-dimensional laws

In this section we state and prove the central limit theorem (Theorem 3.2.2) for the finite-dimensional laws of the weighted quadratic variations process of a fractional Brownian sheet $W^{\alpha,\beta}$ whose Hurst indices satisfy $0 < \alpha < \frac{1}{2}$, $0 < \beta < \frac{1}{2}$ with $\alpha + \beta > \frac{1}{2}$. The main part of this result consists in adapting a method developed by Nourdin and Nualart in [83] relying on the Malliavin calculus.

Let $(W^{\alpha,\beta}_{(s,t)})_{(s,t)\in[0,1]^2}$ be a two-parameter fractional Brownian motion with Hurst indices α and β. Let $f : \mathbb{R} \to \mathbb{R}$ a function satisfying hypothesis **(H)** presented below. Let

$$X^n_{(s,t)} := n^{-1} \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) \left(n^{2(\alpha+\beta)}|\Delta_{i,j}W^{\alpha,\beta}|^2 - 1\right), \quad (s,t) \in [0,1]^2 \tag{3.2.5}$$

be the re-normalized weighted quadratic variations where the increments $\Delta_{i,j}W^{\alpha,\beta}$ are defined as

$$\Delta_{i,j}W^{\alpha,\beta} := W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)} + W^{\alpha,\beta}_{\left(\frac{i}{n},\frac{j}{n}\right)} - W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j}{n}\right)} - W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j}{n}\right)}.$$

Recall that using the product formula (*c.f.* [94, Proposition 1.1.2]) X^n can be expressed as,

$$X^n_{(s,t)} := n^{-1} \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) n^{2(\alpha+\beta)} I_2(\mathbf{1}^{\otimes 2}_{\Delta_{i,j}}), \quad (s,t) \in [0,1]^2.$$

We aim at showing that X^n converges stably in finite-dimensional law to a process X defined as

$$X_{(s,t)} := \sigma_{\alpha,\beta} \int_0^s \int_0^t f\left(W^{\alpha,\beta}_{(u,v)}\right) dB_{(u,v)}, \quad (s,t) \in [0,1]^2, \tag{3.2.6}$$

where B is a standard Brownian sheet independent of $W^{\alpha,\beta}$ and

$$\sigma_{\alpha,\beta} := \sqrt{\frac{1}{8} \sum_{c,d=-\infty}^{\infty} \left([|c+1|^{2\alpha} + |c-1|^{2\alpha} - 2|c|^{2\alpha}]\,[|d+1|^{2\beta} + |d-1|^{2\beta} - 2|d|^{2\beta}]\right)^2}. \tag{3.2.7}$$

Note that in the standard Brownian sheet case ($(\alpha,\beta) = (\frac{1}{2},\frac{1}{2})$), X coincide with the limiting process obtained in [117] since $\sigma_{\frac{1}{2},\frac{1}{2}} := \sqrt{2}$.
Stable convergence in law has been introduced by Rényi in [114, 115]. A full explanation about this subject can be found in [67, Section VIII.5.c]. Morally stable convergence in law consists in the convergence of the couple $(X^n, W)_n$ to (X, W) as n tends to infinity. The definition of the limiting process X in (3.2.6) requires some justification. A way to define X is to consider an extension of the probability basis say $\mathcal{B} := (\Omega, \mathcal{F}, (\mathcal{F}_{(s,t)})_{(s,t)\in[0,1]^2}, \mathbb{P})$ on which $W^{\alpha,\beta}$ is defined. We introduce an auxiliary probability basis $\mathcal{B}' := \mathcal{B} := (\Omega', \mathcal{F}', (\mathcal{F}'_{(s,t)})_{(s,t)\in[0,1]^2}, \mathbb{P}')$ as follows. The set Ω' is defined as the space of continuous functions on $[0,1]^2$ which vanish on $\{(s,t) \in [0,1]^2,\ s = 0 \text{ or } t = 0\}$. The process B is then the canonical process on Ω' that is

$$B_{(s,t)}(\omega') := \omega'(s,t), \quad (s,t) \in [0,1]^2$$

and $\mathcal{F}'$ is defined as the filtration generated by B. According to [94, Section 2.4.1] there exists a probability measure $\mathbb{P}'$ on $(\Omega', \mathcal{F}')$ under which B is a standard Brownian sheet. We finally define $\mathcal{F}'_{(s,t)}$ as the σ-field generated by $\{B_{(u,v)},\ (u,v) \preceq (s,t)\}$ with

$$(s,t) \preceq (u,v) \Leftrightarrow s \le u \text{ and } t \le v.$$

We now describe the extension $\tilde{\mathcal{B}} := (\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_{(s,t)})_{(s,t)\in[0,1]^2}, \tilde{\mathbb{P}})$ of $\mathcal{B}$ on which X will be defined.

$$\begin{cases} \tilde{\Omega} := \Omega \times \Omega', \\ \tilde{\mathcal{F}} := \mathcal{F} \otimes \mathcal{F}', \\ (\tilde{\mathcal{F}}_{(s,t)\in[0,1]^2}) := (\cap_{(u,v)>(s,t)} \mathcal{F}_{(u,v)} \otimes \mathcal{F}'_{(s,t)})_{(s,t)\in[0,1]^2}, \\ \tilde{\mathbb{P}}(d\omega, d\omega') := \mathbb{P}(d\omega)\tilde{\mathbb{P}}(d\omega'). \end{cases}$$

On $\tilde{\mathcal{B}}$, $(X_{(s,t)})_{(s,t)\in[0,1]^2}$ is defined by

$$X_{(s,t)}(\omega, \omega') := \sigma_{\alpha,\beta} \left(\int_0^s \int_0^t f\left(W^{\alpha,\beta}_{(u,v)}\right)(\omega)\, dB_{(u,v)} \right)(\omega'), \quad (s,t) \in [0,1]^2.$$

We will denote by $\mathbb{E}$ (respectively $\tilde{\mathbb{E}}$) the expectation under $\mathbb{P}$ (respectively $\tilde{\mathbb{P}}$).
Before stating and proving the first result we introduce the hypothesis (**H**) on the weight

function f.
Hypothesis (H):
$f : \mathbb{R} \to \mathbb{R}$ belongs to $\mathcal{C}^4(\mathbb{R})$ and for every $i = 1, \ldots, 4$

$$\sup_{(s,t)\in[0,1]^2} \mathbb{E}\left[\left|f^{(i)}\left(W^{\alpha,\beta}_{(s,t)}\right)\right|^p\right] < \infty, \quad p \in (0, \infty).$$

Theorem 3.2.2. *Let $f : \mathbb{R} \to \mathbb{R}$ satisfying hypothesis* (**H**)*. If one of the following situations hold*

(i) $\frac{1}{4} < \alpha < \frac{1}{2}$ *and* $\frac{1}{4} < \beta < \frac{1}{2}$,

(ii) $0 < \alpha < \frac{1}{4}$ *and* $\frac{1}{4} < \beta < \frac{1}{2}$ *with* $\alpha + \beta > \frac{1}{2}$,

(ii') $0 < \beta < \frac{1}{4}$ *and* $\frac{1}{4} < \alpha < \frac{1}{2}$ *with* $\alpha + \beta > \frac{1}{2}$.

then the sequence $(X^n)_n$ defined in (3.2.5) converges stably in finite-dimensional law to X (see (3.2.6)), that is for every $(t_1, \ldots, t_m) \in ([0,1]^2)^m$, for every bounded measurable function $g : \mathbb{R}^m \to \mathbb{R}$ and for every $\mathcal{F}$-measurable random variable Z, the following convergence holds

$$\mathbb{E}\left[g(X^n_{t_1}, \ldots, X^n_{t_m})Z\right] \underset{n\to\infty}{\longrightarrow} \tilde{\mathbb{E}}\left[g(X_{t_1}, \ldots, X_{t_m})Z\right]. \tag{3.2.8}$$

Recall that $\mathcal{F}$ denotes the σ-field generated by $(W^{\alpha,\beta}_{(s,t)})_{(s,t)\in[0,1]^2}$.
Before turning to the proof of this result, let us make the following remarks.

Remarks 3.2.3. *1) The set of conditions (i), (ii) and (ii') in Theorem 3.2.2 is equivalent to the condition*

$$0 < \alpha < \frac{1}{2} \;\; and \;\; 0 < \beta < \frac{1}{2} \;\; with \;\; \alpha + \beta > \frac{1}{2}.$$

This last condition will be preferred in the computations of the proof.

2) Condition (ii) (and its symmetric counterpart (ii')) in Theorem 3.2.2 states that one of the two components of the sheet can have a regularity lower than $\frac{1}{4}$ which is the critical lower bound in the one-parameter case (see [84, Theorem 1]) provided the other one has a regularity such that $\alpha + \beta > \frac{1}{2}$.

3) The standard Brownian case corresponding to $(\alpha, \beta) = (\frac{1}{2}, \frac{1}{2})$ has been obtained in [117, Theorem 3.1].

Now we prove below Theorem 3.2.2.

Proof. Let (α, β) such that $0 < \alpha < \frac{1}{2}$, $0 < \beta < \frac{1}{2}$ with $\alpha + \beta > \frac{1}{2}$. Let $m \geq 1$ and $t_1, \ldots, t_m$ in $([0,1]^2)^m$. For $1 \leq i \leq m$, t_i is denoted as $t_i := (t_{i,1}, t_{i,2})$. By Lemma 3.2.5 the sequence $((X^n_{t_1}, \ldots, X^n_{t_m}), (W^{\alpha,\beta}_z)_{z\in[0,1]^2})_n$ is tight in $\mathbb{R}^m \times \mathcal{C}([0,1]^2)$, in other words there exists a subsequence of $((X^n_{t_1}, \ldots, X^n_{t_m}), (W^{\alpha,\beta}_z)_{z\in[0,1]^2})_n$ which converges in law to some limit denoted by $((X_{t_1}^\infty,\dots,X_{t_m}^\infty),(W_z^{\alpha,\beta})_{z\in[0,1]^2})_n$. This subsequence and the sequence itself will be denoted by the same notation. We have to show that

$$\mathbb{E}\left[X_{t_1}^\infty,\dots,X_{t_m}^\infty|(W_z^{\alpha,\beta})_{z\in[0,1]^2}\right] \stackrel{\mathcal{L}}{=} \mathbb{E}\left[X_{t_1},\dots,X_{t_m}|(W_z^{\alpha,\beta})_{z\in[0,1]^2}\right].$$

Let us consider the conditional characteristic functions of respectively $(X_{t_1}^n,\dots,X_{t_m}^n)$ and $(X_{t_1},\dots,X_{t_m})$ given $W^{\alpha,\beta}$

$$\begin{aligned}\Phi^{m,n}(\lambda_1,\dots,\lambda_m) &:= \mathbb{E}\left[e^{i\left\langle(\lambda_1,\dots,\lambda_m),(X_{t_1}^n,\dots,X_{t_m}^n)\right\rangle}\big|(W_z^{\alpha,\beta})_{z\in[0,1]^2}\right]\\ \Phi^{m}(\lambda_1,\dots,\lambda_m) &:= \mathbb{E}\left[e^{i\left\langle(\lambda_1,\dots,\lambda_m),(X_{t_1},\dots,X_{t_m})\right\rangle}\big|(W_z^{\alpha,\beta})_{z\in[0,1]^2}\right].\end{aligned} \tag{3.2.9}$$

The proof consists to show that

$$\Phi^{m,n}(\lambda_1,\dots,\lambda_m) \underset{n\to\infty}{\longrightarrow} \Phi^{m}(\lambda_1,\dots,\lambda_m), \quad \tilde{\mathbb{P}}-a.s.$$

Furthermore $\Phi^m(\lambda_1,\dots,\lambda_m)$ can be computed as follows,

$$\Phi^{m}(\lambda_1,\dots,\lambda_m) = e^{-\frac{1}{2}\langle(\lambda_1,\dots,\lambda_m),\mathbb{Q}_{t_1,\dots,t_m}(\lambda_1,\dots,\lambda_m)\rangle_{\mathbb{R}^m}},$$

where $\mathbb{Q}_{t_1,\dots,t_m}$ is the symmetric matrix $\mathbb{Q}_{t_1,\dots,t_m} := (C_{i,j})_{1\le i,j,\le m}$ defined as,

$$C_{i,j} := \sigma_{\alpha,\beta}^2 \int_{[0,t_{i,1}\wedge t_{j,1}]\times[0,t_{i,2}\wedge t_{j,2}]} f^2\left(W_{(u,v)}^{\alpha,\beta}\right)\, dudv,$$

which leads to,

$$\Phi^{m}(\lambda_1,\dots,\lambda_m) = \exp\left(-\frac{1}{2}\sum_{k=1}^m \lambda_k^2 C_{k,k} - \sum_{i=1}^{m-1}\lambda_i \sum_{k=i+1}^m \lambda_k C_{i,k}\right).$$

The key point is that Φ^m is the unique solution of the system of partial differential equations (3.2.10). Actually let the system of PDE's (3.2.10) given by functions $\varphi:\mathbb{R}^m\to\mathbb{R}$ satisfying,

$$\left\{\begin{array}{lcl} \frac{\partial\varphi}{\partial\lambda_1}(\lambda_1,\dots,\lambda_m) &=& -\varphi(\lambda_1,\dots,\lambda_m)\left(\lambda_1 C_{1,1}+\sum_{k=2}^m \lambda_k C_{1,k}\right)\\ \vdots & \vdots & \vdots\\ \frac{\partial\varphi}{\partial\lambda_i}(\lambda_1,\dots,\lambda_m) &=& -\varphi(\lambda_1,\dots,\lambda_m)\left(\lambda_i C_{i,i}+\sum_{k=1;k\neq i}^m \lambda_k C_{i,k}\right)\\ \vdots & \vdots & \vdots\\ \frac{\partial\varphi}{\partial\lambda_m}(\lambda_1,\dots,\lambda_m) &=& -\varphi(\lambda_1,\dots,\lambda_m)\left(\lambda_m C_{m,m}+\sum_{k=1}^{m-1} \lambda_k C_{m,k}\right)\end{array}\right. , \tag{3.2.10}$$

with the condition $\varphi(0,\dots,0)=1$. Using standard techniques for PDE's systems that we briefly describe now, it can be shown that the unique solution to (3.2.10) is Φ^m. Indeed, let φ be a solution to (3.2.10). Using the first equation we have that

$$\varphi(\lambda_1,\dots,\lambda_m)=\exp\left(-\left(\frac{C_{11}}{2}\lambda_1^2+\lambda_1\sum_{k=2}^m C_{1,k}\lambda_k\right)\right)\exp\left(g_1(\lambda_2,\dots,\lambda_m)\right),$$

where $g_1:\mathbb{R}^{m-1}\to\mathbb{R}$ is regular enough. Since φ satisfies the second equation of the system we have,

$$\frac{\partial g_1}{\partial\lambda_2}(\lambda_2,\dots,\lambda_m)=-\left(\lambda_2C_{2,2}+\sum_{k=3}^m\lambda_kC_{2,k}\right).$$

Consequently

$$\begin{aligned}\varphi(\lambda_1,\dots,\lambda_m) &= \exp\left(-\frac{1}{2}\sum_{i=1}^2\lambda_i^2C_{i,i}-\sum_{i=1}^2\lambda_i\sum_{k=i+1}^m\lambda_kC_{i,k}\right)\\ &\quad\times\exp\left(g_2(\lambda_3,\dots,\lambda_m)\right),\end{aligned}$$

where $g_2:\mathbb{R}^{m-2}\to\mathbb{R}$ is regular enough. Continuing this procedure we obtain that

$$\varphi(\lambda_1,\dots,\lambda_m)=\exp\left(-\frac{1}{2}\sum_{i=1}^m\lambda_i^2C_{i,i}-\sum_{i=1}^{m-1}\lambda_i\sum_{k=i+1}^m C_{i,k}\lambda_k\right)\exp(g_m),$$

where g_m is a constant. Using the condition $\varphi(0,\dots,0)=1$ we conclude that $g_m=0$ and consequently, $\varphi=\Phi^m$.
We proceed as follows: we show that the conditional characteristic function of $(X^\infty_{t_1},\dots,X^\infty_{t_m})$ given $W^{\alpha,\beta}$ is solution to the system (3.2.10). This will be realized if we can show that for every random variable H of the form $H:=\psi\left(W^{\alpha,\beta}_{s_1},\dots,W^{\alpha,\beta}_{s_r}\right)$ with $\psi:\mathbb{R}^r\to\mathbb{R}$ in $\mathcal{C}^\infty_b(\mathbb{R}^r)$, $s_1,\dots,s_r\in[0,1]^2$, we have

$$\begin{aligned}&\frac{\partial}{\partial\lambda_p}\mathbb{E}\left[e^{i\langle(\lambda_1,\dots,\lambda_m),(X^\infty_{t_1},\dots,X^\infty_{t_m})\rangle_{\mathbb{R}^m}}H\right]\\ =&\ -\sigma^2_{\alpha,\beta}\lambda_p\int_{[0,t_p]}\mathbb{E}\left[f^2\left(W_{(s,t)}\right)He^{i\langle\Lambda,\mathbb{X}^\infty\rangle}\right]dsdt\\ -&\ \sigma^2_{\alpha,\beta}\sum_{a=1,a\neq p}^m\lambda_a\int_{[0,t_{p,1}\wedge t_{a,1}]\times[0,t_{p,2}\wedge t_{a,2}]}\mathbb{E}\left[f^2\left(W_{(s,t)}\right)He^{i\langle\Lambda,\mathbb{X}^\infty\rangle}\right]dsdt,\end{aligned}\qquad(3.2.11)$$

where we use the notation $\langle\Lambda,\mathbb{X}^n\rangle:=\sum_{k=1}^m\lambda_kX^n_{t_k}$. Since

$$\frac{\partial}{\partial\lambda_p}\mathbb{E}\left[e^{i\langle(\lambda_1,\dots,\lambda_m),(X^\infty_{t_1},\dots,X^\infty_{t_m})\rangle_{\mathbb{R}^m}}H\right]=\lim_{n\to\infty}\frac{\partial\Phi^{m,n}_H}{\partial\lambda_p}(\lambda_1,\dots,\lambda_m),\quad p=1,\dots,m,$$

where $\Phi_H^{m,n}$ is defined as

$$\Phi_H^{m,n} := \mathbb{E}\left[e^{\left(i\left\langle(\lambda_1,\ldots,\lambda_m),(X^n_{t_1},\ldots,X^n_{t_m})\right\rangle_{\mathbb{R}^m}\right)}H\right],$$

we have to compute $\frac{\partial\Phi_H^{m,n}}{\partial\lambda_p}(\lambda_1,\ldots,\lambda_m)$. The calculations presented below enable us to obtain the key expression (3.2.16).
Let $p\in\{1,\ldots,m\}$. We have,

$$\begin{aligned}
&\frac{\partial\Phi_H^{m,n}}{\partial\lambda_p}(\lambda_1,\ldots,\lambda_m)\\
&= i\,\mathbb{E}\left[X^n_{t_p}e^{i\langle\Lambda,\mathbb{X}^n\rangle}H\right]\\
&= in^{2(\alpha+\beta)-1}\sum_{i=1}^{[nt_{p,1}]}\sum_{j=1}^{[nt_{p,2}]}\mathbb{E}\left[I_2(\mathbf{1}^{\otimes 2}_{\Delta_{i,j}})f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)e^{i\langle\Lambda,\mathbb{X}^n\rangle}H\right]\\
&= in^{2(\alpha+\beta)-1}\sum_{i=1}^{[nt_{p,1}]}\sum_{j=1}^{[nt_{p,2}]}\mathbb{E}\left[\left\langle\mathbf{1}^{\otimes 2}_{\Delta_{i,j}},D^2\left(f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)e^{i\langle\Lambda,\mathbb{X}^n\rangle}H\right)\right\rangle_{\mathcal{H}^{\otimes 2}}\right]. \quad (3.2.12)
\end{aligned}$$

Let $1\le i\le[nt_{p,1}]$ and $1\le j\le[nt_{p,2}]$ and let $\delta_{i,j}:=\mathbf{1}_{[0,\frac{i-1}{n}]\times[0,\frac{j-1}{n}]}$. We have,

$$\begin{aligned}
&\mathbb{E}\left[\left\langle\mathbf{1}^{\otimes 2}_{\Delta_{i,j}},D^2\left(f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)e^{i\langle\Lambda,\mathbb{X}^n\rangle}H\right)\right\rangle_{\mathcal{H}^{\otimes 2}}\right]\\
&= \mathbb{E}\left[f''\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\right]\left\langle\delta_{i,j},\mathbf{1}_{\Delta_{i,j}}\right\rangle^2_{\mathcal{H}}\\
&+ 2\,\mathbb{E}\left[f'\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)e^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle DH,\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\right]\left\langle\delta_{i,j},\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\\
&+ 2i\,\mathbb{E}\left[f'\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle D\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\right]\left\langle\delta_{i,j},\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\\
&+ \mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)e^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle D^2H,\mathbf{1}^{\otimes 2}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right]\\
&+ 2i\,\mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)e^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle D\left\langle\Lambda,\mathbb{X}^n\right\rangle,\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\left\langle DH,\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\right]\\
&+ i\,\mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle D^2\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}^{\otimes 2}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right]\\
&- \mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle D\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}_{\Delta_{i,j}}\right\rangle^2_{\mathcal{H}}\right]. \quad (3.2.13)
\end{aligned}$$

Now we compute $\left\langle D\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}$ and $\left\langle D^2\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}^{\otimes 2}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}^{\otimes 2}}$.

$$\left\langle D\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}$$

$$
\begin{aligned}
&= n^{2(\alpha+\beta)-1}\sum_{a=1}^{m}\lambda_a\sum_{k=1}^{[nt_{a,1}]}\sum_{l=1}^{[nt_{a,2}]}\left[f'\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right)I_2\left(\mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right)\left\langle \delta_{k,l},\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\right.\\
&+ \left. 2f\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)\Delta_{i,j}W^{\alpha,\beta}\left\langle \mathbf{1}_{\Delta_{k,l}},\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\right]
\end{aligned}
$$

$$
\begin{aligned}
&\left\langle D^2\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}^{\otimes 2}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}^{\otimes 2}}\\
&= n^{2(\alpha+\beta)-1}\sum_{a=1}^{m}\lambda_a\sum_{k=1}^{[nt_{a,1}]}\sum_{l=1}^{[nt_{a,2}]}\left[f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right)I_2\left(\mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right)\left\langle \delta_{k,l},\mathbf{1}_{\Delta_{i,j}}\right\rangle^2_{\mathcal{H}}\right.\\
&+ 4f'\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)\Delta_{k,l}W^{\alpha,\beta}\left\langle \delta_{k,l},\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\left\langle \mathbf{1}_{\Delta_{k,l}},\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\\
&+ \left. 2f\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)\left\langle \mathbf{1}_{\Delta_{k,l}},\mathbf{1}_{\Delta_{i,j}}\right\rangle^2_{\mathcal{H}}\right]
\end{aligned}
$$

In the following we use the following notation $[nt_a] = ([nt_{a,1}],[nt_{a,2}])$. Equation (3.2.13) can be rewritten as

$$
\begin{aligned}
&\mathbb{E}\left[\left\langle \mathbf{1}^{\otimes 2}_{\Delta_{i,j}},D^2\left(f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)e^{i\langle\Lambda,\mathbb{X}^n\rangle}H\right)\right\rangle_{\mathcal{H}^{\otimes 2}}\right]\\
&= 2in^{2(\alpha+\beta)-1}\sum_{a=1}^{m}\lambda_a\sum_{k,l=1}^{[nt_a]}\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)f\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle \mathbf{1}_{\Delta_{k,l}},\mathbf{1}_{\Delta_{i,j}}\right\rangle^2_{\mathcal{H}}\right]\\
&+ r_{i,j,n}, \qquad (3.2.14)
\end{aligned}
$$

where

$$
\begin{aligned}
r_{i,j,n} &= \mathbb{E}\left[f''\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\right]\left\langle \delta_{i,j},\mathbf{1}_{\Delta_{i,j}}\right\rangle^2_{\mathcal{H}}\\
&+ 2\,\mathbb{E}\left[f'\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)e^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle DH,\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\right]\left\langle \delta_{i,j},\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\\
&+ 2i\,\mathbb{E}\left[f'\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle D\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\right]\left\langle \delta_{i,j},\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\\
&+ \mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)e^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle D^2H,\mathbf{1}^{\otimes 2}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right]\\
&+ 2i\,\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)e^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle D\left\langle\Lambda,\mathbb{X}^n\right\rangle,\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\left\langle DH,\mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\right]\\
&- \mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle D\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}_{\Delta_{i,j}}\right\rangle^2_{\mathcal{H}}\right]\\
&+ in^{2(\alpha+\beta)-1}\,\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\sum_{a=1}^{m}\lambda_a\sum_{k,l=1}^{[nt_a]}\right.
\end{aligned}
$$

$$
\begin{aligned}
&\quad \times f''\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{j-1}{n}\right)}\right) I_2\left(\mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right)\left\langle \delta_{k,l}, \mathbf{1}_{\Delta_{i,j}}\right\rangle^2_{\mathcal{H}}\Bigg] \\
&+ \quad 4in^{2(\alpha+\beta)-1}\,\mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle}\sum_{a=1}^{m}\lambda_a\right. \\
&\quad \left.\times \sum_{k,l=1}^{[nt_a]} f'\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)\Delta_{k,l}W^{\alpha,\beta}\left\langle \delta_{k,l}, \mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\left\langle \mathbf{1}_{\Delta_{k,l}}, \mathbf{1}_{\Delta_{i,j}}\right\rangle_{\mathcal{H}}\right] \\
&= \quad \sum_{k=1}^{8} r^{(k)}_{i,j,n}.
\end{aligned}
\tag{3.2.15}
$$

At this stage of the proof we have shown that,

$$
\begin{aligned}
&\frac{\partial \Phi^{m,n}_H}{\partial \lambda_p}(\lambda_1,\ldots,\lambda_p) \\
&= \quad -2n^{4(\alpha+\beta)-2}\sum_{a=1}^{m}\sum_{i,j=1}^{[nt_p]}\sum_{k,l=1}^{[nt_a]}\mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle \mathbf{1}_{\Delta_{i,j}}, \mathbf{1}_{\Delta_{k,l}}\right\rangle^2_{\mathcal{H}}\right] \\
&+ \quad in^{2(\alpha+\beta)-1}\sum_{i,j=1}^{[nt_p]} r_{i,j,n}.
\end{aligned}
\tag{3.2.16}
$$

From Lemma 3.2.6 we have

$$
\sup_{1\le i,j\le n}|r_{i,j,n}| \le Cn^{-4(\alpha+\beta)}, \quad n\ge 1. \tag{3.2.17}
$$

Using the estimate (3.2.17) we obtain,

$$
\begin{aligned}
&\lim_{n\to\infty}\frac{\partial \Phi^{m,n}_H}{\partial \lambda_p}(\lambda_1,\ldots,\lambda_m) \\
&= \quad -2\sum_{a=1}^{m}\lambda_a\lim_{n\to\infty} n^{4(\alpha+\beta)-2}\sum_{i,j=1}^{[nt_p]}\sum_{k,l=1}^{[nt_a]}\mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle \mathbf{1}_{\Delta_{k,l}}, \mathbf{1}_{\Delta_{i,j}}\right\rangle^2_{\mathcal{H}}\right] \\
&= \quad -\lim_{n\to\infty}\frac{1}{8n^2}\sum_{a=1}^{m}\lambda_a\sum_{i,j=1}^{[nt_p]}\sum_{k,l=1}^{[nt_a]}\mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle}\right. \\
&\quad \left.\times\left(\left[|k-i+1|^{2\alpha}+|k-i-1|^{2\alpha}-2|k-i|^{2\alpha}\right]\left[|l-j+1|^{2\alpha}+|l-j-1|^{2\alpha}-2|l-j|^{2\alpha}\right]\right)^2\right] \\
&= \quad -\lim_{n\to\infty}\frac{1}{8n^2}\lambda_p\sum_{i,j=1}^{[nt_p]}\sum_{k,l=1}^{[nt_p]}\mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle}\right. \\
&\quad \left.\times\left(\left[|k-i+1|^{2\alpha}+|k-i-1|^{2\alpha}-2|k-i|^{2\alpha}\right]\left[|l-j+1|^{2\alpha}+|l-j-1|^{2\alpha}-2|l-j|^{2\alpha}\right]\right)^2\right] \\
&- \quad \lim_{n\to\infty}\frac{1}{8n^2}\sum_{a=1,a\neq p}^{m}\lambda_a\sum_{i,j=1}^{[nt_p]}\sum_{k,l=1}^{[nt_a]}\mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle}\right.
\end{aligned}
$$

$$\times \left(\left[|k-i+1|^{2\alpha}+|k-i-1|^{2\alpha}-2|k-i|^{2\alpha}\right]\left[|l-j+1|^{2\alpha}+|l-j-1|^{2\alpha}-2|l-j|^{2\alpha}\right]\right)^2\Big]$$

$$= -\lim_{n\to\infty}\frac{1}{8n^2}\lambda_p\sum_{c=-\infty}^{\infty}\sum_{i=1\vee(1-c)}^{([nt_{p,1}]-c)\wedge[nt_{p,1}]}\sum_{d=-\infty}^{\infty}\sum_{j=1\vee(1-d)}^{([nt_{p,2}]-d)\wedge[nt_{p,2}]}\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)f\left(W^{\alpha,\beta}_{(\frac{i+c-1}{n},\frac{j+d-1}{n})}\right)\right.$$

$$\left.\times He^{i\langle\Lambda,\mathbb{X}^n\rangle}\left(\left[|c+1|^{2\alpha}+|c-1|^{2\alpha}-2|c|^{2\alpha}\right]\left[|d+1|^{2\alpha}+|d-1|^{2\alpha}-2|d|^{2\alpha}\right]\right)^2\right]$$

$$- \lim_{n\to\infty}\frac{1}{8n^2}\sum_{a=1,a\neq p}^{m}\lambda_a\sum_{c,d=-\infty}^{\infty}\sum_{i=1\vee(1-c)}^{([nt_{a,1}]-c)\wedge[nt_{p,1}]}\sum_{j=1\vee(1-d)}^{([nt_{a,2}]-d)\wedge[nt_{p,2}]}\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)f\left(W^{\alpha,\beta}_{(\frac{i+c-1}{n},\frac{j+d-1}{n})}\right)\right.$$

$$\left.\times He^{i\langle\Lambda,\mathbb{X}^n\rangle}\left(\left[|c+1|^{2\alpha}+|c-1|^{2\alpha}-2|c|^{2\alpha}\right]\left[|d+1|^{2\alpha}+|d-1|^{2\alpha}-2|d|^{2\alpha}\right]\right)^2\right]$$

$$= -\sigma^2_{\alpha,\beta}\lambda_p\int_{[0,t_p]}\mathbb{E}\left[f^2\left(W_{(s,t)}\right)He^{i\langle\Lambda,\mathbb{X}^\infty\rangle}\right]dsdt$$

$$- \sigma^2_{\alpha,\beta}\sum_{a=1,a\neq p}^{m}\lambda_a\int_{[0,t_{p,1}\wedge t_{a,1}]\times[0,t_{p,2}\wedge t_{a,2}]}\mathbb{E}\left[f^2\left(W_{(s,t)}\right)He^{i\langle\Lambda,\mathbb{X}^\infty\rangle}\right]dsdt,$$

which leads to (3.2.11). As a consequence,

$$\frac{\partial}{\partial\lambda_p}\mathbb{E}\left[e^{i\langle\Lambda,\mathbb{X}^\infty\rangle}|(W^{\alpha,\beta}_z)_{z\in[0,1]^2}\right]$$

$$= -\sigma^2_{\alpha,\beta}\lambda_p\int_{[0,t_p]}f^2\left(W_{(s,t)}\right)\mathbb{E}\left[e^{i\langle\Lambda,\mathbb{X}^\infty\rangle}|(W^{\alpha,\beta}_z)_{z\in[0,1]^2}\right]dsdt$$

$$- \sigma^2_{\alpha,\beta}\sum_{a=1,a\neq p}^{m}\lambda_a\int_{[0,t_{p,1}\wedge t_{a,1}]\times[0,t_{p,2}\wedge t_{a,2}]}f^2\left(W_{(s,t)}\right)\mathbb{E}\left[e^{i\langle\Lambda,\mathbb{X}^\infty\rangle}|(W^{\alpha,\beta}_z)_{z\in[0,1]^2}\right]dsdt$$

which is exactly the pth equation of the system (3.2.10). □

3.2.4 Useful lemmas

All the results of this section mimic those obtained by Nourdin and Nualart in [83]. We stress that we adapt these computations to the two-parameter case in order to make this paper self-contained. Furthermore the conditions imposed on α and β in Theorem 3.2.2 can be easily deduced from the calculations realized in this section. Note that in this section C will denote a generic constant which can differ from one line to another.

Lemma 3.2.4. *Let $l=(l_1,l_2)\in[0,1]^2$, $s=(s_1,s_2)\preceq t=(t_1,t_2)$. If $\alpha<\frac{1}{2}$ and $\beta<\frac{1}{2}$ then,*

$$\left|\mathbb{E}\left[W^{\alpha,\beta}_{(l_1,l_2)}\left(W^{\alpha,\beta}_{(s_1,s_2)}+W^{\alpha,\beta}_{(t_1,t_2)}-W^{\alpha,\beta}_{(s_1,t_2)}-W^{\alpha,\beta}_{t_1,s_2}\right)\right]\right|\leq|t_1-s_1|^{2\alpha}|t_2-s_2|^{2\beta}.$$

Proof. We refer to [83, Lemma 1](or to [84, Lemmas 4 and 5]) where the following estimate is shown, for $a<b$ we have,

$$|b^{2\gamma}-a^{2\gamma}|\leq|b-a|^{2\gamma},\quad\forall\,0<\gamma<\frac{1}{2}.$$

□

Lemma 3.2.5. *Let $m \geq 1$ and $t_1, \ldots, t_m$ in $[0,1]^2$. Under hypotheses of Theorem 3.2.2, the sequence $(X^n_{t_1}, \ldots, X^n_{t_m})_{n\geq 1}$ is tight in $\mathbb{R}^m$.*

Proof. We show that

$$\lim_{n\to\infty} \mathbb{E}\left[(X^n_{t_1}, \ldots, X^n_{t_m})\right] = 0, \tag{3.2.18}$$

$$\lim_{n\to\infty} \mathbb{E}\left[\|(X^n_{t_1}, \ldots, X^n_{t_m})\|^2\right] = \sigma^2_{\alpha,\beta} \sum_{i=1}^{m} \int_{[0,t_i]} \mathbb{E}\left[f^2\left(W^{\alpha,\beta}_{(s,t)}\right)\right] dsdt. \tag{3.2.19}$$

Proof of (3.2.18):
Since, $\Delta_{i,j} W^{\alpha,\beta} = I_1(\mathbf{1}_{\Delta_{i,j}})$, using the definition $I_n(h^{\otimes n}) := n! H_n(W^{\alpha,\beta}(h))$ and relation on Hermite polynomials [94, (1.3) p. 5]

$$n^{2(\alpha+\beta)} |\Delta_{i,j} W^{\alpha,\beta}|^2 - 1 = n^{2(\alpha+\beta)} I_2(\mathbf{1}^{\otimes 2}_{\Delta_{i,j}}).$$

Consequently,

$$\begin{aligned} \mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right) I_2(\mathbf{1}^{\otimes 2}_{\Delta_{k,l}})\right] &= \mathbb{E}\left[\left\langle D^2\left(f\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)\right), \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right] \\ &= \mathbb{E}\left[f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)\right] \left\langle \delta^{\otimes 2}_{k,l}, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}} \end{aligned}$$

where $\delta_{k,l} = \mathbf{1}_{[0,\frac{k-1}{n}]\times[0,\frac{l-1}{n}]}$.
Let for p in $\{1, \ldots, n\}$ $t_p = (t_{p,1}, t_{p,2})$. We have that

$$\left|\mathbb{E}\left[X^n_{t_p}\right]\right| \leq n^{2(\alpha+\beta)-1} \sum_{k=1}^{[nt_{p,1}]} \sum_{l=1}^{[nt_{p,2}]} \left|\mathbb{E}\left[f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)\right]\right| \left\langle \delta_{k,l}, \mathbf{1}_{\Delta_{k,l}}\right\rangle^2_{\mathcal{H}}.$$

By Lemma 3.2.4 we have that

$$\left|\mathbb{E}\left[X^n_{t_p}\right]\right| \leq C\, n^{1-2(\alpha+\beta)}.$$

Proof of (3.2.19):
Let p in $\{1, \ldots, n\}$.

$$|X^n_{t_p}|^2 = n^{4(\alpha+\beta)-2} \sum_{i,k=1}^{[nt_{p,1}]} \sum_{j,l=1}^{[nt_{p,2}]} f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) f\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right) I_2\left(\mathbf{1}^{\otimes 2}_{\Delta_{i,j}}\right) I_2\left(\mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right) \tag{3.2.20}$$

Using the product formula ([94, Proposition 1.1.2]) we have,

$$I_2\left(\mathbf{1}^{\otimes 2}_{\Delta_{i,j}}\right) I_2\left(\mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right) = \sum_{r=0}^{2} r! \binom{2}{r}\binom{2}{r} I_{4-2r}\left(\mathbf{1}^{\otimes 2}_{\Delta_{i,j}} \otimes_r \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right)$$

$$\begin{aligned} &= I_4\left(\mathbf{1}_{\Delta_{i,j}}^{\otimes 2}\tilde{\otimes}\mathbf{1}_{\Delta_{k,l}}^{\otimes 2}\right) + 4I_2\left(\mathbf{1}_{\Delta_{i,j}}\otimes\mathbf{1}_{\Delta_{k,l}}\right)\langle\mathbf{1}_{\Delta_{i,j}},\mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}} \\ &+ 2\langle\mathbf{1}_{\Delta_{i,j}}^{\otimes 2},\mathbf{1}_{\Delta_{k,l}}^{\otimes 2}\rangle_{\mathcal{H}^{\otimes 2}}. \end{aligned} \tag{3.2.21}$$

From relations (3.2.21) and (3.2.20) we obtain that,

$$\mathbb{E}\left[|X_{t_p}^n|^2\right] \tag{3.2.22}$$

$$\begin{aligned} &= n^{4(\alpha+\beta)-2}\sum_{i,k=1}^{[nt_{p,1}]}\sum_{j,l=1}^{[nt_{p,2}]}\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)f\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)I_4\left(\mathbf{1}_{\Delta_{i,j}}^{\otimes 2}\tilde{\otimes}\mathbf{1}_{\Delta_{k,l}}^{\otimes 2}\right)\right] \\ &+ 4n^{4(\alpha+\beta)-2}\sum_{i,k=1}^{[nt_{\alpha,1}]}\sum_{j,l=1}^{[nt_{\alpha,2}]}\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)f\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)I_2\left(\mathbf{1}_{\Delta_{i,j}}\otimes\mathbf{1}_{\Delta_{k,l}}\right)\langle\mathbf{1}_{\Delta_{i,j}},\mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}}\right] \\ &+ 2n^{4(\alpha+\beta)-2}\sum_{i,k=1}^{[nt_{\alpha,1}]}\sum_{j,l=1}^{[nt_{\alpha,2}]}\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)f\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)\langle\mathbf{1}_{\Delta_{i,j}}^{\otimes 2},\mathbf{1}_{\Delta_{k,l}}^{\otimes 2}\rangle_{\mathcal{H}^{\otimes 2}}\right] \\ &= T_1+T_2+T_3. \end{aligned} \tag{3.2.23}$$

We end the proof by considering independently each of the three terms in the right hand side of expression (3.2.22).
First note that using Malliavin integration by parts formula the first term can be written as,

$$n^{4(\alpha+\beta)-2}\sum_{i,k=1}^{[nt_{p,1}]}\sum_{j,l=1}^{[nt_{p,2}]}\mathbb{E}\left[\left\langle D^4\left(f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)f\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)\right),\mathbf{1}_{\Delta_{i,j}}^{\otimes 2}\tilde{\otimes}\mathbf{1}_{\Delta_{k,l}}^{\otimes 2}\right\rangle_{\mathcal{H}^{\otimes 4}}\right]$$

which from Lemma 3.2.4 shows that $|T_1|\le Cn^{2-4(\alpha+\beta)}$.
Then we consider the term T_2 in (3.2.22). We also integrate by part in order to obtain,

$$\begin{aligned} T_2 &= 4n^{4(\alpha+\beta)-2}\sum_{i,k=1}^{[nt_{p,1}]}\sum_{j,l=1}^{[nt_{p,2}]}\mathbb{E}\left[\left\langle D^2\left(f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)f\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{l-1}{n})}\right)\right),\mathbf{1}_{\Delta_{i,j}}\otimes\mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right] \\ &\times \langle\mathbf{1}_{\Delta_{i,j}},\mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}} \end{aligned}$$

As a consequence we obtain that,

$$\begin{aligned} |T_2| &\le Cn^{-2}\sum_{i,k=1}^{[nt_{p,1}]}\sum_{j,l=1}^{[nt_{p,2}]}|\langle\mathbf{1}_{\Delta_{i,j}},\mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}}| \\ &= Cn^{-2-2(\alpha+\beta)}\sum_{i,k=1}^{[nt_{p,1}]}\sum_{j,l=1}^{[nt_{p,2}]}\Big[\,\big||k-i+1|^{2\alpha}+|k-i-1|^{2\alpha}-2|k-i|^{2\alpha}\big| \end{aligned}$$

$$\times \left| |l-j+1|^{2\beta} + |l-j-1|^{2\beta} - 2|l-j|^{2\beta} \right| \Big]$$

$$\leq \; Cn^{-2(\alpha+\beta)} \sum_{c=-\infty}^{\infty} \sum_{d=-\infty}^{\infty} \left| |c+1|^{2\alpha} + |c-1|^{2\alpha} - 2|c|^{2\alpha} \right| \left| |d+1|^{2\beta} + |d-1|^{2\beta} - 2|d|^{2\beta} \right|.$$

The serie presented above converges since $\alpha < \frac{1}{2}$ and $\beta < \frac{1}{2}$. Consider now T_3.

$$\begin{aligned}
T_3 \quad &= \quad \frac{1}{8} n^{-2} \sum_{i,k=1}^{[nt_{p,1}]} \sum_{j,l=1}^{[nt_{p,2}]} \left[\mathbb{E} \left[f \left(W^{\alpha,\beta}_{(\frac{i-1}{n}, \frac{j-1}{n})} \right) f \left(W^{\alpha,\beta}_{(\frac{k-1}{n}, \frac{l-1}{n})} \right) \right] \right. \\
&\quad \left. \times \left(\left| |k-i+1|^{2\alpha} + |k-i-1|^{2\alpha} - 2|k-i|^{2\alpha} \right| \left| |l-j+1|^{2\beta} + |l-j-1|^{2\beta} - 2|l-j|^{2\beta} \right| \right)^2 \right] \\
&= \quad \frac{1}{8} n^{-2} \sum_{c=-\infty}^{\infty} \sum_{i=1\vee(1-c)}^{([nt_{p,1}]-c)\wedge[nt_{p,1}]} \sum_{d=-\infty}^{\infty} \sum_{j=1\vee(1-d)}^{([nt_{p,2}]-d)\wedge[nt_{p,2}]} \mathbb{E} \left[f \left(W^{\alpha,\beta}_{(\frac{i-1}{n}, \frac{j-1}{n})} \right) f \left(W^{\alpha,\beta}_{(\frac{i+c-1}{n}, \frac{j+d-1}{n})} \right) \right. \\
&\quad \left. \times \left(\left[|c+1|^{2\alpha} + |c-1|^{2\alpha} - 2|c|^{2\alpha} \right] \left[|d+1|^{2\beta} + |d-1|^{2\beta} - 2|d|^{2\beta} \right] \right)^2 \right] \\
&\underset{n\to\infty}{\longrightarrow} \quad \sigma^2_{\alpha,\beta} \int_{[0,t_p]} \mathbb{E} \left[f^2 \left(W^{\alpha,\beta}_{(s,t)} \right) \right] dsdt.
\end{aligned}$$

Note the series appearing in the term T_3 converges since $\alpha < \frac{3}{4}$ and $\beta < \frac{3}{4}$. □

Lemma 3.2.6. *Let the notations of Theorem 3.2.2 and of its proof prevail. Under hypotheses of Theorem 3.2.2, we have*

$$\sup_{1 \leq i,j \leq n} |r_{i,j,n}| \leq Cn^{-4(\alpha+\beta)}, \quad n \geq 1.$$

Proof. Recall that in (3.2.15) $r_{i,j,n}$ is decomposed into eight terms,

$$r_{i,j,n} = \sum_{k=1}^{8} r^{(k)}_{i,j,n}.$$

We show each of them is less or equal to $Cn^{-4(\alpha+\beta)}$. From the definition of H we deduce the following expressions

$$\begin{cases} DH = \sum_{l=1}^{r} \partial_l \psi(W^{\alpha,\beta}_{s_1}, \ldots, W^{\alpha,\beta}_{s_r}) \mathbf{1}_{[0,s_l]} \\ D^2H = \sum_{l_2=1}^{r} \sum_{l_1=1}^{r} \partial_{l_2} \partial_{l_1} \psi(W^{\alpha,\beta}_{s_1}, \ldots, W^{\alpha,\beta}_{s_r}) \mathbf{1}_{[0,s_{l_2}]} \otimes \mathbf{1}_{[0,s_{l_1}]}. \end{cases}$$

By Lemma 3.2.4, for $k = 1, 2, 4$,

$$|r^{(k)}_{i,j,n}| \leq Cn^{-4(\alpha+\beta)}.$$

Estimate of $r^{(6)}_{i,j,n}$:
First note that

$$|r^{(6)}_{i,j,n}| \leq C \, \mathbb{E} \left[\left(1 + f^2 \left(W^{\alpha,\beta}_{(\frac{i-1}{n}, \frac{j-1}{n})} \right) \right) \left\langle D \langle \Lambda, \mathbb{X}^n \rangle, \mathbf{1}_{\Delta_{i,j}} \right\rangle^2_{\mathcal{H}} \right].$$

Let $F_{i,j,n} := 1 + f^2\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)$.

$$\begin{aligned}
& \mathbb{E}\left[F_{i,j,n}\left\langle D\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}_{\Delta_{i,j}}\right\rangle^2_{\mathcal{H}}\right] \\
\leq\ & n^{4(\alpha+\beta)-2}\sum_{a=1}^{m}\sum_{b=1}^{m}\lambda_a\lambda_b\sum_{k,l=1}^{[nt_a]}\sum_{\tilde{k},\tilde{l}=1}^{[nt_b]}\left[A^{(1)}_{i,j,n}\langle\delta_{k,l},\mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}}\langle\delta_{\tilde{k},\tilde{l}},\mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}}\right. \\
& \left.+A^{(2)}_{i,j,n}\langle\mathbf{1}_{\Delta_{k,l}},\mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}}\langle\mathbf{1}_{\Delta_{\tilde{k},\tilde{l}}},\mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}}\right]
\end{aligned}$$

where

$$\begin{cases}
A^{(1)}_{i,j,n} = 2\,\mathbb{E}\left[F_{i,j,n}f'\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)f'\left(W^{\alpha,\beta}_{\left(\frac{\tilde{k}-1}{n},\frac{\tilde{l}-1}{n}\right)}\right)I_2\left(\mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right)I_2\left(\mathbf{1}^{\otimes 2}_{\Delta_{\tilde{k},\tilde{l}}}\right)\right], \\
A^{(2)}_{i,j,n} = 8\,\mathbb{E}\left[F_{i,j,n}f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)f\left(W^{\alpha,\beta}_{\left(\frac{\tilde{k}-1}{n},\frac{\tilde{l}-1}{n}\right)}\right)\Delta_{k,l}W^{\alpha,\beta}\Delta_{\tilde{k},\tilde{l}}W^{\alpha,\beta}\right].
\end{cases}$$

The term $n^{4(\alpha+\beta)-2}\sum_{a=1}^{m}\sum_{b=1}^{m}\lambda_a\lambda_b\sum_{k,l=1}^{[nt_a]}\sum_{\tilde{k},\tilde{l}=1}^{[nt_b]}A^{(1)}_{i,j,n}$ is very similar to the term $\mathbb{E}[|X^n_{t_p}|^2]$ computed in Lemma 3.2.5. Furthermore by Lemma 3.2.4 we have

$$\left|\langle\delta_{k,l},\mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}}\langle\delta_{\tilde{k},\tilde{l}},\mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}}\right| \leq Cn^{-4(\alpha+\beta)}$$

which leads to,

$$\left|n^{4(\alpha+\beta)-2}\sum_{a=1}^{m}\sum_{b=1}^{m}\lambda_a\lambda_b\sum_{k,l=1}^{[nt_a]}\sum_{\tilde{k},\tilde{l}=1}^{[nt_b]}\left[A^{(1)}_{i,j,n}\langle\delta_{k,l},\mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}}\langle\delta_{\tilde{k},\tilde{l}},\mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}}\right]\right| \leq Cn^{-4(\alpha+\beta)}.$$

We consider the term $A^{(2)}_{i,j,n}$.

$$\begin{aligned}
& \left|\mathbb{E}\left[F_{i,j,n}f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)f\left(W^{\alpha,\beta}_{\left(\frac{\tilde{k}-1}{n},\frac{\tilde{l}-1}{n}\right)}\right)\Delta_{k,l}W^{\alpha,\beta}\Delta_{\tilde{k},\tilde{l}}W^{\alpha,\beta}\right]\right| \\
=\ & \left|\mathbb{E}\left[F_{i,j,n}f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)f\left(W^{\alpha,\beta}_{\left(\frac{\tilde{k}-1}{n},\frac{\tilde{l}-1}{n}\right)}\right)I_1\left(\mathbf{1}_{\Delta_{k,l}}\right)I_1\left(\mathbf{1}_{\Delta_{\tilde{k},\tilde{l}}}\right)\right]\right| \\
\leq\ & \left|\mathbb{E}\left[F_{i,j,n}f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)f\left(W^{\alpha,\beta}_{\left(\frac{\tilde{k}-1}{n},\frac{\tilde{l}-1}{n}\right)}\right)I_2\left(\mathbf{1}_{\Delta_{k,l}}\otimes\mathbf{1}_{\Delta_{\tilde{k},\tilde{l}}}\right)\right]\right| \\
& +\left|\mathbb{E}\left[F_{i,j,n}f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)f\left(W^{\alpha,\beta}_{\left(\frac{\tilde{k}-1}{n},\frac{\tilde{l}-1}{n}\right)}\right)\left\langle\mathbf{1}_{\Delta_{k,l}},\mathbf{1}_{\Delta_{\tilde{k},\tilde{l}}}\right\rangle_{\mathcal{H}}\right]\right|, \text{by [94, Proposition 1.1.2]} \\
=\ & \left|\mathbb{E}\left[\left\langle D^2\left(F_{i,j,n}f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)f\left(W^{\alpha,\beta}_{\left(\frac{\tilde{k}-1}{n},\frac{\tilde{l}-1}{n}\right)}\right)\right),\mathbf{1}_{\Delta_{k,l}}\otimes\mathbf{1}_{\Delta_{\tilde{k},\tilde{l}}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right]\right| \\
& +\left|\mathbb{E}\left[F_{i,j,n}f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)f\left(W^{\alpha,\beta}_{\left(\frac{\tilde{k}-1}{n},\frac{\tilde{l}-1}{n}\right)}\right)\left\langle\mathbf{1}_{\Delta_{k,l}},\mathbf{1}_{\Delta_{\tilde{k},\tilde{l}}}\right\rangle_{\mathcal{H}}\right]\right| \\
\leq\ & Cn^{-2(\alpha+\beta)}.
\end{aligned}$$

We don't give the details of the computations since the term

$$\mathbb{E}\left[\left\langle D^2\left(F_{i,j,n} f\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right) f\left(W^{\alpha,\beta}_{\left(\frac{\tilde{k}-1}{n},\frac{\tilde{l}-1}{n}\right)}\right)\right), \mathbf{1}_{\Delta_{k,l}} \otimes \mathbf{1}_{\Delta_{\tilde{k},\tilde{l}}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right]$$

can be written as a sum of

$$\mathbb{E}\left[F_{i,j,n} f^{(a)}\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right) f^{(b)}\left(W^{\alpha,\beta}_{\left(\frac{\tilde{k}-1}{n},\frac{\tilde{l}-1}{n}\right)}\right) f^{(c)}\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)\left\langle \delta_{u_1,v_1} \otimes \delta_{u_2,v_2}, \mathbf{1}_{\Delta_{k,l}} \otimes \mathbf{1}_{\Delta_{\tilde{k},\tilde{l}}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right],$$

where a, b, c are integer between zero and two and (u_i, v_i) are for (k, l) or $(\tilde{k}, \tilde{l})$, $i = 1, 2$. Consequently,

$$\begin{aligned}
&\left| n^{4(\alpha+\beta)-2} \sum_{a=1}^{m}\sum_{b=1}^{m} \lambda_a \lambda_b \sum_{k,l=1}^{[nt_a]} \sum_{\tilde{k},\tilde{l}=1}^{[nt_b]} A^{(2)}_{i,j,n} \langle \mathbf{1}_{\Delta_{k,l}}, \mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}} \langle \mathbf{1}_{\Delta_{\tilde{k},\tilde{l}}}, \mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}} \right| \\
\leq\ & n^{2(\alpha+\beta-1)} \sum_{a=1}^{m}\sum_{b=1}^{m} \lambda_a \lambda_b \sum_{k,l=1}^{[nt_a]} \sum_{\tilde{k},\tilde{l}=1}^{[nt_b]} A^{(2)}_{i,j,n} \left| \langle \mathbf{1}_{\Delta_{k,l}}, \mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}} \langle \mathbf{1}_{\Delta_{\tilde{k},\tilde{l}}}, \mathbf{1}_{\Delta_{i,j}}\rangle_{\mathcal{H}} \right| \\
=\ & C n^{-2(\alpha+\beta)-2} \left(\sum_{c=-\infty}^{\infty} \left| |c+1|^{2\alpha} + |c-1|^{2\alpha} - 2|c|^{2\alpha} \right| \right)^2 \\
& \times \left(\sum_{d=-\infty}^{\infty} \left| |d+1|^{2\beta} + |d-1|^{2\beta} - 2|d|^{2\beta} \right| \right)^2 \\
\leq\ & C n^{-4(\alpha+\beta)},
\end{aligned}$$

where the series above converge since $0 < \alpha < \frac{1}{2}$ and $0 < \beta < \frac{1}{2}$.

Estimate of $r^{(3)}_{i,j,n}$ and $r^{(5)}_{i,j,n}$:
Note that

$$|r^{(3)}_{i,j,n}| + |r^{(5)}_{i,j,n}| \leq C n^{-2(\alpha+\beta)} \left\| \left\langle D\langle \Lambda, \mathbb{X}^n\rangle, \mathbf{1}_{\Delta_{i,j}} \right\rangle_{\mathcal{H}} \right\|_{L^2(\Omega,\mathcal{F},\mathbb{P})},$$

which shows that for $k = 3, 5$,

$$|r^{(k)}_{i,j,n}| \leq C n^{-4(\alpha+\beta)},$$

using estimates obtained previously in this proof.

Estimate of $r^{(7)}_{i,j,n}$: We have,

$$\begin{aligned}
& |r^{(7)}_{i,j,n}| \\
\leq\ & C n^{-2(\alpha+\beta)-1} \sum_{a=1}^{m} \lambda_a \sum_{k,l=1}^{[nt_a]} \left| \mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right) He^{i\langle \Lambda, \mathbb{X}^n\rangle} f''\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{j-1}{n}\right)}\right) I_2\left(\mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right)\right]\right|.
\end{aligned}$$

We use the integration by parts formula,

$$\begin{aligned}
& \mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle} f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right) I_2\left(\mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right)\right] \\
= \; & \mathbb{E}\left[\left\langle D^2\left(f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle} f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right)\right), \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right] \\
= \; & \mathbb{E}\left[f''\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle} f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right)\right] \left\langle \delta_{k,l}\otimes\delta_{i,j}, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}} \\
& +2\,\mathbb{E}\left[f'\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) e^{i\langle\Lambda,\mathbb{X}^n\rangle} f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right) \left\langle DH\tilde{\otimes}\delta_{i,j}, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right] \\
& +2i\,\mathbb{E}\left[f'\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle} f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right) \left\langle D\langle\Lambda,\mathbb{X}^n\rangle\tilde{\otimes}\delta_{i,j}, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right] \\
& +2\,\mathbb{E}\left[f'\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle} f^{(3)}\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right)\right] \left\langle \delta_{k,l}\tilde{\otimes}\delta_{i,j}, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}} \\
& +\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) e^{i\langle\Lambda,\mathbb{X}^n\rangle} f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right) \left\langle D^2H, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right] \\
& +2i\,\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) e^{i\langle\Lambda,\mathbb{X}^n\rangle} f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right) \left\langle D\langle\Lambda,\mathbb{X}^n\rangle\tilde{\otimes}DH, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right] \\
& +2\,\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) e^{i\langle\Lambda,\mathbb{X}^n\rangle} f^{(3)}\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right) \left\langle \delta_{k,l}\tilde{\otimes}DH, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right] \\
& -\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle} f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right) \left\langle D\langle\Lambda,\mathbb{X}^n\rangle\otimes D\langle\Lambda,\mathbb{X}^n\rangle, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right] \\
& +i\,\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle} f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right) \left\langle D^2\langle\Lambda,\mathbb{X}^n\rangle, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right] \\
& +2i\,\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle} f^{(3)}\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right) \left\langle D\langle\Lambda,\mathbb{X}^n\rangle\tilde{\otimes}\delta_{k,l}, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right] \\
& +\mathbb{E}\left[f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) He^{i\langle\Lambda,\mathbb{X}^n\rangle} f^{(4)}\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right)\right] \left\langle \delta^{\otimes 2}_{k,l}, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}.
\end{aligned}$$

By Lemma 3.2.4 and the estimates shown above in this proof,

$$\left|\langle\delta_{s,t}, \mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}}\right| \le Cn^{-2(\alpha+\beta)}, \quad (s,t)\in\{(k,l),(i,j)\},$$

$$\left\|\left\langle DH, \mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}}\right\|_{L^2(\Omega,\mathcal{F},\mathbb{P})} \le Cn^{-2(\alpha+\beta)},$$

and

$$\left\|\left\langle D^2H, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right\|_{L^2(\Omega,\mathcal{F},\mathbb{P})} \le Cn^{-4(\alpha+\beta)}.$$

Furthermore as done in the part **Estimate of** $r^{(6)}_{i,j,n}$,

$$\mathbb{E}\left[\left(1+\left(f\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right) f''\left(W^{\alpha,\beta}_{(\frac{k-1}{n},\frac{j-1}{n})}\right)\right)^2\right) \left\langle D\langle\Lambda,\mathbb{X}^n\rangle, \mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}}\right] \le Cn^{-2(\alpha+\beta)}.$$

Consequently,

$$|r^{(7)}_{i,j,n}| \le Cn^{-2(\alpha+\beta)-1}\sum_{a=1}^{m}\lambda_a\left[n^{-4(\alpha+\beta)+2} + \sum_{k,l=1}^{[nt_a]}\left\|\left\langle D^2\langle\Lambda,\mathbb{X}^n\rangle, \mathbf{1}^{\otimes 2}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}^{\otimes 2}}\right\|_{L^2(\Omega,\mathcal{F},\mathbb{P})}\right].$$

$$\mathbb{E}\left[\left\langle D^2\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}_{\Delta_{k,l}}^{\otimes 2}\right\rangle_{\mathcal{H}^{\otimes 2}}^2\right]\leq n^{4(\alpha+\beta)-2}\sum_{a,b=1}^{m}\lambda_a\lambda_b\sum_{i,j=1}^{[nt_a]}\sum_{\tilde{i},\tilde{j}=1}^{[nt_b]}\mathbb{E}[J_1+J_2+J_3],$$

where

$$\begin{cases} J_1 := 3f''\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)f''\left(W^{\alpha,\beta}_{(\frac{\tilde{i}-1}{n},\frac{\tilde{j}-1}{n})}\right)I_2\left(\mathbf{1}_{\Delta_{i,j}}^{\otimes 2}\right)I_2\left(\mathbf{1}_{\Delta_{\tilde{i},\tilde{j}}}^{\otimes 2}\right)\langle\delta_{i,j},\mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}}^2\left\langle\delta_{\tilde{i},\tilde{j}},\mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}}^2, \\ J_2 := 48f'\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)f'\left(W^{\alpha,\beta}_{(\frac{\tilde{i}-1}{n},\frac{\tilde{j}-1}{n})}\right)\Delta_{i,j}W\Delta_{\tilde{i},\tilde{j}}W\left\langle\mathbf{1}_{\Delta_{i,j}},\mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}}\langle\delta_{i,j},\mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}} \\ \quad\times\left\langle\mathbf{1}_{\Delta_{\tilde{i},\tilde{j}}},\mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}}\left\langle\delta_{\tilde{i},\tilde{j}},\mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}}, \\ J_3 := 12f'\left(W^{\alpha,\beta}_{(\frac{i-1}{n},\frac{j-1}{n})}\right)f'\left(W^{\alpha,\beta}_{(\frac{\tilde{i}-1}{n},\frac{\tilde{j}-1}{n})}\right)\langle\mathbf{1}_{\Delta_{i,j}},\mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}}^2\left\langle\mathbf{1}_{\Delta_{\tilde{i},\tilde{j}}},\mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}}^2. \end{cases}$$

From computations made in the proof of Lemma 3.2.5 we have that

$$n^{4(\alpha+\beta)-2}\sum_{a,b=1}^{m}\lambda_a\lambda_b\sum_{i,j=1}^{[nt_a]}\sum_{\tilde{i},\tilde{j}=1}^{[nt_b]}\mathbb{E}[J_1]\leq Cn^{-8(\alpha+\beta)}.$$

Using an estimate obtained for the term $A^{(2)}_{i,j,n}$ obtained above in this proof we have that,

$$\begin{aligned} &\left|n^{4(\alpha+\beta)-2}\sum_{a,b=1}^{m}\lambda_a\lambda_b\sum_{i,j=1}^{[nt_a]}\sum_{\tilde{i},\tilde{j}=1}^{[nt_b]}\mathbb{E}[J_2]\right| \\ \leq\ & Cn^{-2(\alpha+\beta)-2}\sum_{i,j=1}^{[nt_a]}\sum_{\tilde{i},\tilde{j}=1}^{[nt_b]}\left|\langle\mathbf{1}_{\Delta_{i,j}},\mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}}\right|\left|\langle\mathbf{1}_{\Delta_{i,j}},\mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}}\right| \\ \leq\ & Cn^{-2(\alpha+\beta)-2}\left(\sum_{i,j=1}^{n}\left|\langle\mathbf{1}_{\Delta_{i,j}},\mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}}\right|\right)^2 \\ =\ & Cn^{-6(\alpha+\beta)-2}\left(\sum_{c,d=-\infty}^{\infty}\left||c+1|^{2\alpha}+|c-1|^{2\alpha}-2|c|^{2\alpha}\right|\left||d+1|^{2\beta}+|d-1|^{2\beta}-2|d|^{2\beta}\right|\right)^2. \end{aligned}$$

Note that the serie above is convergent since $0<\alpha<\frac{1}{2}$ and $0<\beta<\frac{1}{2}$. For the last term we have,

$$\begin{aligned} & n^{4(\alpha+\beta)-2}\sum_{a,b=1}^{m}\lambda_a\lambda_b\sum_{i,j=1}^{[nt_a]}\sum_{\tilde{i},\tilde{j}=1}^{[nt_b]}\mathbb{E}[J_3] \\ \leq\ & n^{4(\alpha+\beta)-2}\left(\sum_{i,j=1}^{n}\langle\mathbf{1}_{\Delta_{i,j}},\mathbf{1}_{\Delta_{k,l}}\rangle_{\mathcal{H}}^2\right)^2 \\ \leq\ & n^{-4(\alpha+\beta)-2}\left(\sum_{c,d=-\infty}^{\infty}\left||c+1|^{2\alpha}+|c-1|^{2\alpha}-2|c|^{2\alpha}\right|^2\left||d+1|^{2\beta}+|d-1|^{2\beta}-2|d|^{2\beta}\right|^2\right)^2. \end{aligned}$$

which leads to the result.

Estimate of $r^{(8)}_{i,j,n}$:

$$\begin{aligned}
& \mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)f'\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\Delta_{k,l}W^{\alpha,\beta}\right] \\
= \; & \mathbb{E}\left[f'\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)f'\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\right]\left\langle\delta_{i,j},\mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}} \\
+ \; & \mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)f''\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\right]\left\langle\delta_{k,j},\mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}} \\
+ \; & \mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)f'\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)e^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle DH,\mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}}\right] \\
+ \; & i\,\mathbb{E}\left[f\left(W^{\alpha,\beta}_{\left(\frac{i-1}{n},\frac{j-1}{n}\right)}\right)f'\left(W^{\alpha,\beta}_{\left(\frac{k-1}{n},\frac{l-1}{n}\right)}\right)He^{i\langle\Lambda,\mathbb{X}^n\rangle}\left\langle D\langle\Lambda,\mathbb{X}^n\rangle,\mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}}\right]
\end{aligned}$$

From preceding computations we have that

$$\left|r^{(8)}_{i,j,n}\right| \leq Cn^{-2(\alpha+\beta)-1}\sum_{a=1}^{m}\sum_{k,l=1}^{[nt_a]}\left|\left\langle\mathbf{1}_{\Delta_{i,j}},\mathbf{1}_{\Delta_{k,l}}\right\rangle_{\mathcal{H}}\right| \leq Cn^{-4(\alpha+\beta)-1}$$

which concludes the proof. ☐

Acknowledgement

I thank Ivan Nourdin for helpful discussions.
We thank the financial support of the *European Social Fund (ESF)*.

3.3 Variations quadratiques à poids du mouvement brownien fractionnaire: le cas limite $H = 1/4$

Cette section correspond à la référence [90] réalisée en collaboration avec Ivan Nourdin.

3.3.1 Introduction

Let B^H be a fractional Brownian motion with Hurst index $H \in (0,1)$. Since the seminal works by Breuer and Major [20], Dobrushin and Major [34], Giraitis and Surgailis [46] or Taqqu [128], it is well-known that

- if $H \in (0, \frac{3}{4})$ then

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \left[n^{2H} (B^H_{(k+1)/n} - B^H_{k/n})^2 - 1 \right] \xrightarrow[n\to\infty]{\text{Law}} \mathscr{N}(0, C_H^2); \tag{3.3.1}$$

- if $H = \frac{3}{4}$ then

$$\frac{1}{\sqrt{n \, \log n}} \sum_{k=0}^{n-1} \left[n^{3/2} (B^{3/4}_{(k+1)/n} - B^{3/4}_{k/n})^2 - 1 \right] \xrightarrow[n\to\infty]{\text{Law}} \mathscr{N}(0, C_{\frac{3}{4}}^2); \tag{3.3.2}$$

- if $H \in (\frac{3}{4}, 1)$ then

$$n^{1-2H} \sum_{k=0}^{n-1} \left[n^{2H} (B^H_{(k+1)/n} - B^H_{k/n})^2 - 1 \right] \xrightarrow[n\to\infty]{\text{Law}} \text{“Rosenblatt r.v.”}. \tag{3.3.3}$$

Here, $C_H > 0$ denotes a constant depending only on H and which can be computed explicitly. Moreover, the term “Rosenblatt r.v.” denotes a random variable whose distribution is the same as that of the Rosenblatt process Z at time one, see (3.3.9) below.

Now, let f be a real function assumed to be regular enough. Very recently, the asymptotic behavior of

$$\sum_{k=0}^{n-1} f(B^H_{k/n}) \left[n^{2H} (B^H_{(k+1)/n} - B^H_{k/n})^2 - 1 \right] \tag{3.3.4}$$

received a lot of attention, see [49, 78, 82, 81, 84] (see also the related works [88, 117, 116, 127]). The initial motivation of such a study was to derive the exact rates of convergence of some approximation schemes associated to scalar stochastic differential equations driven by B^H, see [49, 78, 82] for precise statements. But it turned out that it was also interesting for itself because it highlighted new phenomena with respect to (3.3.1)-(3.3.3). Indeed, in the study of the asymptotic behavior of (3.3.4), a new critical value ($H = \frac{1}{4}$) has appeared. More precisely:

- if $H < \frac{1}{4}$ then

$$n^{2H-1}\sum_{k=0}^{n-1} f(B^H_{k/n})\big[n^{2H}(B^H_{(k+1)/n}-B^H_{k/n})^2-1\big] \xrightarrow[n\to\infty]{L^2} \frac{1}{4}\int_0^1 f''(B^H_s)ds; \qquad (3.3.5)$$

- if $\frac{1}{4} < H < \frac{3}{4}$ then

$$\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1} f(B^H_{k/n})\big[n^{2H}(B^H_{(k+1)/n}-B^H_{k/n})^2-1\big] \xrightarrow[n\to\infty]{\text{Law}} C_H\int_0^1 f(B^H_s)dW_s \qquad (3.3.6)$$

  for W a standard Brownian motion independent of B^H;

- if $H = \frac{3}{4}$ then

$$\frac{1}{\sqrt{n\,\log n}}\sum_{k=0}^{n-1} f(B^{3/4}_{k/n})\big[n^{3/2}(B^{3/4}_{(k+1)/n}-B^{3/4}_{k/n})^2-1\big] \xrightarrow[n\to\infty]{\text{Law}} C_{3/4}\int_0^1 f(B^{3/4}_s)dW_s \quad (3.3.7)$$

  for W a standard Brownian motion independent of $B^{3/4}$;

- if $H > \frac{3}{4}$ then

$$n^{1-2H}\sum_{k=0}^{n-1} f(B^H_{k/n})\big[n^{2H}(B^H_{(k+1)/n}-B^H_{k/n})^2-1\big] \xrightarrow[n\to\infty]{L^2} \int_0^1 f(B^H_s)dZ_s \qquad (3.3.8)$$

  for Z the Rosenblatt process defined by

$$Z_s = I_2^X(L_s), \qquad (3.3.9)$$

  where I_2^X denotes the double stochastic integral with respect to the Wiener process X given by the transfer equation (3.3.16) and where, for every $s \in [0,1]$, L_s is the symmetric square integrable kernel given by

$$L_s(y_1,y_2) = \frac{1}{2}\mathbf{1}_{[0,s]^2}(y_1,y_2)\int_{y_1\vee y_2}^{s} \frac{\partial K_H}{\partial u}(u,y_1)\frac{\partial K_H}{\partial u}(u,y_2)du.$$

Even if it is not completely obvious at first glance, convergences (3.3.1) and (3.3.5) well agree. Indeed, since $2H-1 < -\frac{1}{2}$ if and only if $H < \frac{1}{4}$, (3.3.5) is actually a particular case of (3.3.1) when $f \equiv 1$. The convergence (3.3.5) is proved in [81] while the other cases (3.3.6)-(3.3.8) are proved in [84]. On the other hand, notice that the relations (3.3.5)-(3.3.8) do not cover the critical case $H = \frac{1}{4}$. Our first main result completes this important (see why just below) missing case:

Theorem 3.3.1. *If $H = \frac{1}{4}$ then*

$$\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1} f(B^{1/4}_{k/n})\big[\sqrt{n}(B^{1/4}_{(k+1)/n} - B^{1/4}_{k/n})^2 - 1\big] \xrightarrow[n\to\infty]{\text{Law}} C_{1/4}\int_0^1 f(B^{1/4}_s)dW_s + \frac{1}{4}\int_0^1 f''(B^{1/4}_s)ds \tag{3.3.10}$$

for W a standard Brownian motion independent of $B^{1/4}$ and where

$$C^2_{1/4} = \frac{1}{2}\sum_{p=-\infty}^{\infty}\Big(\sqrt{|p+1|} + \sqrt{|p-1|} - 2\sqrt{|p|}\Big)^2 < \infty.$$

Here, it is interesting to compare the obtained limit in (3.3.10) with those obtained in the recent work [88]. In [88], the authors also studied the asymptotic behavior of (3.3.4) but when the fractional Brownian motion B^H is replaced by an *iterated Brownian motion* Z, that is the process defined by $Z_t = X(Y_t)$, $t \in [0,1]$, with X and Y two independent standard Brownian motions. Iterated Brownian motion Z is self-similar of index $\frac{1}{4}$ and has stationary increments. Thus, although if it is not Gaussian, Z is "close" to the fractional Brownian motion $B^{1/4}$. For Z instead of $B^{1/4}$, it is proved in [88] that the correctly renormalized weighted quadratic variation (which is not exactly defined as in (3.3.4), but rather by means of a *random* partition composed of Brownian hitting times) converges in law towards the so-called *weighted Brownian motion in random scenery* at time one, defined as

$$\sqrt{2}\int_{-\infty}^{+\infty} f(X_x)L^x_1(Y)dW_x,$$

compare with the right-hand side of (3.3.10). Here, $\{L^x_t(Y)\}_{x\in\mathbb{R},t\in[0,1]}$ stands for the jointly continuous version of the local time process of Y, while W denotes a two-sided standard Brownian motion independent of X and Y.

For now, we take $B^H = B^{1/4}$ to be a fractional Brownian motion with Hurst index $H = \frac{1}{4}$. This particular value of H is important because the fractional Brownian motion with Hurst index $H = \frac{1}{4}$ has a remarkable physical interpretation in terms of particle systems. Indeed, if one consider an infinite number of particles, initially placed on the real line according to a Poisson distribution, performing independent Brownian motions and undergoing "elastic" collisions, then the trajectory of a fixed particle (after rescaling) converges to a fractional Brownian motion with Hurst index $H = \frac{1}{4}$. This striking fact has been first pointed out by Harris in [56], and then rigorously proven in [37] (see also references therein).

Now, let us explain an interesting consequence of a slight modification of Theorem 3.3.1 towards a first step in the construction of a stochastic calculus with respect to $B^{1/4}$. As it is nicely explained by Swanson in [127], there are at least two kinds of Stratonovitch-type Riemann sums that one can consider in order to define $\int_0^1 f(B^{1/4}_s)\circ dB^{1/4}_s$ when f is a real smooth function. The first corresponds to the so-called "trapezoid rule" and is given by

$$S_n(f) = \sum_{k=0}^{n-1}\frac{f(B^{1/4}_{k/n}) + f(B^{1/4}_{(k+1)/n})}{2}\big(B^{1/4}_{(k+1)/n} - B^{1/4}_{k/n}\big).$$

The second corresponds to the so-called "midpoint rule" and is given by

$$T_n(f) = \sum_{k=1}^{\lfloor n/2 \rfloor} f(B^{1/4}_{(2k-1)/n}) \big(B^{1/4}_{(2k)/n} - B^{1/4}_{(2k-2)/n}\big).$$

By Theorem 3 in [84] (see also [28, 52, 53]), we have that

$$\int_0^1 f'(B^{1/4}_s) d^\circ B^{1/4}_s := \lim_{n\to\infty} S_n(f') \quad \text{exists in probability}$$

and verifies the following *classical* change of variable formula:

$$\int_0^1 f'(B^{1/4}_s) d^\circ B^{1/4}_s = f(B^{1/4}_1) - f(0). \tag{3.3.11}$$

On the other hand, it is quoted in [127] that Burdzy and Swanson *conjectured*[1] that

$$\int_0^1 f'(B^{1/4}_s) d^\star B^{1/4}_s := \lim_{n\to\infty} T_n(f') \quad \text{exists in law}$$

and verifies, this time, the following *non classical* change of variable formula:

$$\int_0^1 f'(B^{1/4}_s) d^\star B^{1/4}_s \overset{\text{Law}}{=} f(B^{1/4}_1) - f(0) - \frac{\kappa}{2} \int_0^1 f''(B^{1/4}_s) dW_s, \tag{3.3.13}$$

where κ is an explicit universal constant and W denotes a standard Brownian motion independent of $B^{1/4}$. Our second main result is the following:

Theorem 3.3.2. *The conjecture of Burdzy and Swanson is true. More precisely, (3.3.13) holds for any real function* $f : \mathbb{R} \to \mathbb{R}$ *verifying* $(\mathbf{H}_9)$ *(see Section 3 below).*

[1]In reality, Burdzy and Swanson conjectured (3.3.13) not for the fractional Brownian motion $B^{1/4}$ but for process F defined by

$$F_t = u(t, 0),\ t \in [0, 1], \tag{3.3.12}$$

where

$$u_t = \frac{1}{2} u_{xx} + \dot{W}(t, x),\ t \in [0, 1],\ x \in \mathbb{R}, \quad \text{with initial condition } u(0, x) = 0.$$

(Here, as usual, $\dot{W}$ denotes the space-time white noise on $[0, 1] \times \mathbb{R}$). It is immediately checked that F is a centered Gaussian process with covariance function

$$E(F_s F_t) = \frac{1}{\sqrt{2\pi}} \big(\sqrt{t+s} - \sqrt{|t-s|}\big).$$

so that F is actually a *bifractional Brownian motion* of indices $\frac{1}{2}$ and $\frac{1}{2}$ in the sense of Houdré and Villa [57]. Using the main result of [74], we have that $B^{1/4}$ and F actually differ only from a process with absolutely continuous trajectories. As a direct consequence, using a Girsanov type transformation, we immediately see that it is equivalent to prove (3.3.13) either for $B^{1/4}$ or for F.

As quoted in [119, Remark 12], notice finally that the change of variable formula (3.3.11) also holds for F.

Finally, we would mention that the strategy we used in this paper can also be derived in order to obtain the following analogue of Theorem 3.3.2, that we propose (in order to keep the length of the present paper within limit) to prove in a forthcoming paper:

Theorem 3.3.3. *Let $f:\mathbb{R}\to\mathbb{R}$ be smooth enough. Then $\int_0^1 f'(B_s^{1/6})d^\circ B_s^{1/6}$ exists in law and verifies*

$$\int_0^1 f'(B_s^{1/6})d^\circ B_s^{1/6} \overset{\text{Law}}{=} f(B_1^{1/6}) - f(0) - \frac{\widetilde{\kappa}}{6}\int_0^1 f'''(B_s^{1/6})dW_s,$$

where $\widetilde{\kappa}$ is an explicit universal constant and W denotes a standard Brownian motion independent of $B^{1/6}$.

The rest of the paper is organized as follows. In Section 2, we recall some notion concerning fractional Brownian motion. In Section 3, we prove Theorem 3.3.1. In Section 4, we prove Theorem 3.3.2.

3.3.2 Preliminaries and notations

We begin by briefly recalling some basic facts about stochastic calculus with respect to a fractional Brownian motion. We refer to [93] for further details. Let $B^H = (B_t^H)_{t\in[0,1]}$ be a fractional Brownian motion with Hurst parameter $H \in (0,\frac{1}{2})$ defined on a probability space $(\Omega, \mathcal{A}, P)$. We mean that B^H is a centered Gaussian process with the covariance function

$$R_H(s,t) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right). \tag{3.3.14}$$

We denote by $\mathscr{E}$ the set of step $\mathbb{R}-$valued functions on $[0,1]$. Let $\mathfrak{H}$ be the Hilbert space defined as the closure of $\mathscr{E}$ with respect to the scalar product

$$\left\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]}\right\rangle_{\mathfrak{H}} = R_H(t,s).$$

The covariance kernel $R_H(t,s)$ introduced in (3.3.14) can be written as

$$R_H(t,s) = \int_0^{s\wedge t} K_H(s,u)K_H(t,u)du,$$

where $K_H(t,s)$ is the square integrable kernel defined by

$$K_H(t,s) = c_H\left[\left(\frac{t}{s}\right)^{H-\frac{1}{2}}(t-s)^{H-\frac{1}{2}} - (H-\frac{1}{2})\,s^{\frac{1}{2}-H}\int_s^t u^{H-\frac{3}{2}}(u-s)^{H-\frac{1}{2}}du\right], \quad 0<s<t, \tag{3.3.15}$$

where ${c_H}^2 = 2H(1-2H)^{-1}\beta(1-2H, H+1/2)^{-1}$ and β denotes the Beta function. By convention, we set $K_H(t,s) = 0$ if $s \geq t$.

Let $\mathcal{K}_H^*: \mathscr{E} \to \mathrm{L}^2([0,1])$ be the linear operator defined by:

$$\mathcal{K}_H^*\left(\mathbf{1}_{[0,t]}\right) = K_H(t,\cdot).$$

The following equality holds for any $s, t \in [0,1]$:

$$\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \rangle_{\mathfrak{H}} = \langle \mathcal{K}_H^* \mathbf{1}_{[0,t]}, \mathcal{K}_H^* \mathbf{1}_{[0,s]} \rangle_{\mathrm{L}^2([0,1])} = \mathrm{E}\left(B_t^H B_s^H\right)$$

and then $\mathcal{K}_H^*$ provides an isometry between the Hilbert spaces $\mathfrak{H}$ and a closed subspace of $\mathrm{L}^2([0,1])$. The process $X = (X_t)_{t\in[0,1]}$ defined by

$$X_t = B^H\big((\mathcal{K}_H^*)^{-1}(\mathbf{1}_{[0,t]})\big) \tag{3.3.16}$$

is a Wiener process, and the process B^H has an integral representation of the form

$$B_t^H = \int_0^t K_H(t,s) dX_s.$$

Let $\mathscr{S}$ be the set of all smooth cylindrical random variables, *i.e.* of the form

$$F = \psi(B_{t_1}^H, \ldots, B_{t_m}^H) \tag{3.3.17}$$

where $m \geq 1$, $\psi : \mathbb{R}^m \to \mathbb{R} \in \mathscr{C}_b^\infty$ and $0 \leq t_1 < \ldots < t_m \leq 1$. The Malliavin derivative of F with respect to B^H is the element of $\mathrm{L}^2(\Omega, \mathfrak{H})$ defined by

$$D_s F \;=\; \sum_{i=1}^m \frac{\partial \psi}{\partial x_i}(B_{t_1}^H, \ldots, B_{t_m}^H) \mathbf{1}_{[0,t_i]}(s), \quad s \in [0,1].$$

In particular $D_s B_t^H = \mathbf{1}_{[0,t]}(s)$. For any integer $k \geq 1$, we denote by $\mathbb{D}^{k,2}$ the closure of the set of smooth random variables with respect to the norm

$$\|F\|_{k,2}^2 \;=\; \mathrm{E}\left[F^2\right] + \sum_{j=1}^k \mathrm{E}\left[|D^j F|_{\mathfrak{H}^{\otimes j}}^2\right].$$

The Malliavin derivative D verifies the chain rule: if $\varphi : \mathbb{R}^n \to \mathbb{R}$ is $\mathscr{C}_b^1$ and if $(F_i)_{i=1,\ldots,n}$ is a sequence of elements of $\mathbb{D}^{1,2}$ then $\varphi(F_1, \ldots, F_n) \in \mathbb{D}^{1,2}$ and we have, for any $s \in [0,1]$:

$$D_s\, \varphi(F_1, \ldots, F_n) = \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i}(F_1, \ldots, F_n) D_s F_i.$$

The divergence operator I is the adjoint of the derivative operator D. If a random variable $u \in \mathrm{L}^2(\Omega, \mathfrak{H})$ belongs to the domain of the divergence operator, that is if it verifies

$$|\mathrm{E}\langle DF, u\rangle_{\mathfrak{H}}| \leq c_u \, \|F\|_{\mathrm{L}^2} \quad \text{for any } F \in \mathscr{S},$$

then $I(u)$ is defined by the duality relationship

$$\mathrm{E}\big(F I(u)\big) = \mathrm{E}\big(\langle DF, u\rangle_{\mathfrak{H}}\big),$$

for every $F \in \mathbb{D}^{1,2}$.

For every $n \geq 1$, let $\mathcal{H}_n$ be the nth Wiener chaos of B^H, that is, the closed linear subspace of $L^2(\Omega, \mathcal{A}, P)$ generated by the random variables $\{H_n\left(B^H(h)\right), h \in H, |h|_{\mathfrak{H}} = 1\}$, where H_n is the nth Hermite polynomial. The mapping $I_n(h^{\otimes n}) = n! H_n\left(B^H(h)\right)$ provides a linear isometry between the symmetric tensor product $\mathfrak{H}^{\odot n}$ and $\mathcal{H}_n$. For $H = \frac{1}{2}$, I_n coincides with the multiple stochastic integral. The following duality formula holds

$$E\left(F I_n(h)\right) = E\left(\langle D^n F, h\rangle_{\mathfrak{H}^{\otimes n}}\right), \tag{3.3.18}$$

for any element $h \in \mathfrak{H}^{\odot n}$ and any random variable $F \in \mathbb{D}^{n,2}$. Let $\{e_k,\, k \geq 1\}$ be a complete orthonormal system in $\mathfrak{H}$. Given $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$, for every $r = 0, \ldots, p \wedge q$, the rth contraction of f and g is the element of $\mathfrak{H}^{\otimes(p+q-2r)}$ defined as

$$f \otimes_r g = \sum_{i_1,\ldots,i_r=1}^{\infty} \langle f, e_{i_1} \otimes \ldots \otimes e_{i_r}\rangle_{\mathfrak{H}^{\otimes r}} \otimes \langle g, e_{i_1} \otimes \ldots \otimes e_{i_r}\rangle_{\mathfrak{H}^{\otimes r}}.$$

Note that $f \otimes_0 g = f \otimes g$ equals the tensor product of f and g while, for $p = q$, $f \otimes_p g = \langle f, g\rangle_{\mathfrak{H}^{\otimes p}}$. Finally, we mention the useful following multiplication formula: if $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$, then

$$I_p(f) I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}(f \otimes_r g). \tag{3.3.19}$$

3.3.3 Proof of Theorem 3.3.1

In all this section, $B = B^{1/4}$ denotes a fractional Brownian motion with Hurst index $H = 1/4$.

Let

$$G_n := \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f(B_{k/n})[\sqrt{n}(B_{(k+1)/n} - B_{k/n})^2 - 1], \quad n \geq 1.$$

For $k = 0, \ldots, n-1$ and $t \in [0, 1]$, we set

$$\delta_{k/n} := \mathbf{1}_{[k/n,(k+1)/n]} \quad \text{and} \quad \varepsilon_t := \mathbf{1}_{[0,t]}.$$

The relations between Hermite polynomials and multiple stochastic integrals (see Section 3.3.2) allow to write

$$\sqrt{n}(B_{(k+1)/n} - B_{k/n})^2 - 1 = \sqrt{n}\, I_2(\delta_{k/n}^{\otimes 2}).$$

As a consequence:

$$G_n = \sum_{k=0}^{n-1} f(B_{k/n}) I_2(\delta_{k/n}^{\otimes 2}).$$

In the sequel, for $f : \mathbb{R} \to \mathbb{R}$, we will need assumption of the type:

Hypothesis ($\mathbf{H}_\mathrm{q}$):
The function $f:\mathbb{R}\to\mathbb{R}$ belongs to $\mathscr{C}^q$ and is such that

$$\sup_{t\in[0,1]} E\left(|f^{(i)}(B_t)|^p\right) < \infty$$

for any $p \geq 1$ and $i \in \{0,\dots,q\}$.

We begin by the following technical lemma:

Lemma 3.3.4. *Let $n\geq 1$ and $k=0,\dots,n-1$. We have*

(i) $|E(B_r(B_t-B_s))| \leq \sqrt{t-s}$ *for any $r\in[0,1]$ and $0\leq s<t\leq 1$,*

(ii) $\displaystyle\sup_{t\in[0,1]}\sum_{k=0}^{n-1}\left|\left\langle \varepsilon_t, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \underset{n\to\infty}{=} O(1)$,

(iii) $\displaystyle\sum_{k,j=0}^{n-1}\left|\left\langle \varepsilon_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \underset{n\to\infty}{=} O(n)$,

(iv) $\displaystyle\left|\left\langle \varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2 - \frac{1}{4n}\right| \leq \frac{\sqrt{k+1}-\sqrt{k}}{4n}$*; consequently* $\displaystyle\sum_{k=0}^{n-1}\left|\left\langle \varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2 - \frac{1}{4n}\right| \underset{n\to\infty}{\longrightarrow} 0.$

Proof of Lemma 3.3.4.

(i) We have

$$E(B_r(B_t-B_s)) = \frac{1}{2}(\sqrt{t}-\sqrt{s}) + \frac{1}{2}\left(\sqrt{|s-r|}-\sqrt{|t-r|}\right).$$

Using the classical inequality $\left|\sqrt{|b|}-\sqrt{|a|}\right| \leq \sqrt{|b-a|}$, the desired result follows.

(ii) Observe that

$$\left\langle \varepsilon_t, \delta_{k/n}\right\rangle_{\mathfrak{H}} = \frac{1}{2\sqrt{n}}\left(\sqrt{k+1}-\sqrt{k}-\sqrt{|k+1-nt|}+\sqrt{|k-nt|}\right).$$

Consequently, we have

$$\sum_{k=0}^{n-1}\left|\left\langle \varepsilon_t, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \leq \frac{1}{2}+\frac{1}{2\sqrt{n}}\left(\sum_{k=0}^{\lfloor nt\rfloor-1}\sqrt{nt-k}-\sqrt{nt-k-1} \right.$$
$$\left. +\sqrt{\lfloor nt\rfloor+1-nt}-\sqrt{nt-\lfloor nt\rfloor}+\sum_{k=\lfloor nt\rfloor+1}^{n-1}\sqrt{nt-k}-\sqrt{nt-k-1}\right).$$

The desired conclusion follows easily.

(iii) It is a direct consequence of *(ii)*:

$$\begin{aligned}\sum_{k,j=0}^{n-1}\left|\left\langle\varepsilon_{j/n},\delta_{k/n}\right\rangle_{\mathfrak{H}}\right| &\leq n\sup_{j=0,\ldots,n-1}\sum_{k=0}^{n-1}\left|\left\langle\varepsilon_{j/n},\delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \\ &\underset{n\to\infty}{=} O(n).\end{aligned}$$

(iv) We have

$$\left|\left\langle\varepsilon_{k/n},\delta_{k/n}\right\rangle_{\mathfrak{H}}^2-\frac{1}{4n}\right|=\frac{1}{4n}\left(\sqrt{k+1}-\sqrt{k}\right)\left|\sqrt{k+1}-\sqrt{k}-2\right|.$$

Thus, the desired bound is immediately checked by using $0\leq\sqrt{x+1}-\sqrt{x}\leq 1$ available for $x\geq 0$.

□

The main result of the current section is the following:

Theorem 3.3.5. *Under Hypothesis* $(\mathbf{H}_4)$*, we have*

$$G_n\underset{n\to\infty}{\overset{Law}{\longrightarrow}}C_{1/4}\int_0^1 f(B_s)dW_s+\frac{1}{4}\int_0^1 f''(B_s)ds,$$

where $W=(W_t)_{t\in[0,1]}$ *is a standard Brownian motion independent of* B *and*

$$C_{1/4}:=\sqrt{\frac{1}{2}\sum_{p=-\infty}^{\infty}\left(\sqrt{|p+1|}+\sqrt{|p-1|}-2\sqrt{|p|}\right)^2}<\infty.$$

Proof. This proof is mainly inspired by the first draft of [83]. During all the proof, C will denote a constant depending only on $\|f^{(a)}\|_\infty$, $a=0,1,2,3,4$, which can differ from one line to another.

Step 1.- We begin the proof by showing the following limits:

$$\lim_{n\to\infty}E\left(G_n\right)=\frac{1}{4}\int_0^1 E\left(f''(B_s)\right)ds, \tag{3.3.20}$$

and

$$\lim_{n\to\infty}E\left(G_n^2\right)=C_{1/4}^2\int_0^1 E\left(f^2(B_s)\right)ds+\frac{1}{16}E\left(\int_0^1 f''(B_s)ds\right)^2. \tag{3.3.21}$$

Proof of (3.3.20): we can write

$$E\left(G_n\right) = \sum_{k=0}^{n-1}E\left(f(B_{k/n})I_2(\delta_{k/n}^{\otimes 2})\right)$$

$$
\begin{aligned}
&= \sum_{k=0}^{n-1} E\left(\left\langle D^2(f(B_{k/n})), \delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}}\right) \\
&= \sum_{k=0}^{n-1} E\left(f''(B_{k/n})\right)\left\langle \varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2 \\
&= \frac{1}{4n}\sum_{k=0}^{n-1} E\left(f''(B_{k/n})\right) + \sum_{k=0}^{n-1} E\left(f''(B_{k/n})\right)\left(\left\langle \varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2 - \frac{1}{4n}\right) \\
&\underset{n\to\infty}{\longrightarrow} \frac{1}{4}\int_0^1 E\big(f''(B_s)\big)ds, \text{ by Lemma 3.3.4}(iv) \text{ and under } (\mathbf{H}_4).
\end{aligned}
$$

Proof of (3.3.21):
By the multiplication formula (3.3.19), we have

$$
I_2(\delta_{j/n}^{\otimes 2})I_2(\delta_{k/n}^{\otimes 2}) = I_4(\delta_{j/n}^{\otimes 2}\otimes\delta_{k/n}^{\otimes 2}) + 4\, I_2(\delta_{j/n}\otimes\delta_{k/n})\left\langle \delta_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}} + 2\left\langle \delta_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2. \quad (3.3.22)
$$

Thus

$$
\begin{aligned}
E\left(G_n^2\right) &= \sum_{j,k=0}^{n-1} E\Big(f(B_{j/n})f(B_{k/n})I_2(\delta_{j/n}^{\otimes 2})I_2(\delta_{k/n}^{\otimes 2})\Big) \\
&= \sum_{j,k=0}^{n-1} E\Big(f(B_{j/n})f(B_{k/n})I_4(\delta_{j/n}^{\otimes 2}\otimes\delta_{k/n}^{\otimes 2})\Big) \\
&+ 4\sum_{j,k=0}^{n-1} E\left(f(B_{j/n})f(B_{k/n})I_2(\delta_{j/n}\otimes\delta_{k/n})\right)\left\langle \delta_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}} \\
&+ 2\sum_{j,k=0}^{n-1} E\left(f(B_{j/n})f(B_{k/n})\right)\left\langle \delta_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2 \\
&= A_n + B_n + C_n.
\end{aligned}
$$

Using Malliavin integration by parts formula (3.3.18), A_n can be expressed as follows:

$$
\begin{aligned}
A_n &= \sum_{j,k=0}^{n-1} E\left(\left\langle D^4(f(B_{j/n})f(B_{k/n})), \delta_{j/n}^{\otimes 2}\otimes\delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 4}}\right) \\
&= 24\sum_{j,k=0}^{n-1}\sum_{a+b=4} E\left(f^{(a)}(B_{j/n})f^{(b)}(B_{k/n})\right)\left\langle \varepsilon_{j/n}^{\otimes a}\widetilde{\otimes}\varepsilon_{k/n}^{\otimes b}, \delta_{j/n}^{\otimes 2}\otimes\delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 4}}.
\end{aligned}
$$

In fact, in the previous sum, each term is negligible except

$$
\sum_{j,k=0}^{n-1} E\left(f''(B_{j/n})f''(B_{k/n})\right)\left\langle \varepsilon_{j/n}, \delta_{j/n}\right\rangle_{\mathfrak{H}}^2\left\langle \varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2
$$

$$
\begin{aligned}
&= E\left(\left[\sum_{k=0}^{n-1} f''(B_{k/n}) \left\langle \varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2\right]^2\right) \\
&= E\left(\left[\frac{1}{4n}\sum_{k=0}^{n-1} f''(B_{k/n}) + \sum_{k=0}^{n-1} f''(B_{k/n})\big(\left\langle \varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2 - \frac{1}{4n}\big)\right]^2\right) \\
&\underset{n\to\infty}{\longrightarrow} E\left(\left[\frac{1}{4}\int_0^1 f''(B_s)ds\right]^2\right), \text{by Lemma 3.3.4}(iv) \text{ and under } (\mathbf{H}_4).
\end{aligned}
$$

The other terms appearing in A_n make no contribution to the limit. Indeed, they have the form

$$
\sum_{j,k=0}^{n-1} E\left(f^{(a)}(B_{j/n})f^{(b)}(B_{k/n})\right) \left\langle \varepsilon_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}} \prod_{i=1}^{3} \left\langle \varepsilon_{x_i/n}, \delta_{y_i/n}\right\rangle_{\mathfrak{H}}
$$

(where x_i and y_i are for j or k) and, from Lemma 3.3.4 *(i) (iii)*, we have that

$$
\begin{cases}
\sup_{j,k=0,\dots,n-1} \prod_{i=1}^3 \left|\left\langle \varepsilon_{x_i/n}, \delta_{y_i/n}\right\rangle_{\mathfrak{H}}\right| \underset{n\to\infty}{=} O(n^{-3/2}), \\
\sum_{j,k=0}^{n-1} \left|\left\langle \varepsilon_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \underset{n\to\infty}{=} O(n).
\end{cases}
$$

Still using Malliavin integration by parts formula (3.3.18), we can bound B_n as follows:

$$
\begin{aligned}
|B_n| &\leq 8\sum_{j,k=0}^{n-1}\sum_{a+b=2} \left|E\left(f^{(a)}(B_{j/n})f^{(b)}(B_{k/n})\right)\left\langle \varepsilon_{j/n}^{\otimes a}\widetilde{\otimes}\varepsilon_{k/n}^{\otimes b}, \delta_{j/n}\otimes\delta_{k/n}\right\rangle_{\mathfrak{H}^{\otimes 2}} \left\langle \delta_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \\
&\leq Cn^{-1}\sum_{j,k=0}^{n-1} \left|\left\langle \delta_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right|, \text{ by Lemma 3.3.4 } (i) \text{ and under } (\mathbf{H}_4) \\
&= Cn^{-3/2}\sum_{j,k=0}^{n-1} |\rho(j-k)| \leq Cn^{-1/2}\sum_{r=-\infty}^{\infty} |\rho(r)| \underset{n\to\infty}{=} O(n^{-1/2}),
\end{aligned}
$$

where

$$
\rho(r) := \sqrt{|r+1|} + \sqrt{|r-1|} - 2\sqrt{|r|}, \quad r \in \mathbb{Z}. \tag{3.3.23}
$$

Observe that the serie $\sum_{r=-\infty}^{\infty} |\rho(r)|$ is convergent since $|\rho(r)| \underset{|r|\to\infty}{\sim} \frac{1}{2}|r|^{-\frac{3}{2}}$.

Finally, we consider the term C_n:

$$
\begin{aligned}
C_n &= \frac{1}{2n}\sum_{j,k=0}^{n-1} E\left(f(B_{j/n})f(B_{k/n})\right)\rho^2(j-k) \\
&= \frac{1}{2n}\sum_{r=-\infty}^{\infty}\sum_{j=0\vee -r}^{(n-1)\wedge(n-1-r)} E\left(f(B_{j/n})f(B_{(j+r)/n})\right)\rho^2(r)
\end{aligned}
$$

$$\underset{n\to\infty}{\longrightarrow} \frac{1}{2}\int_0^1 E\left(f^2(B_s)\right) ds \sum_{r=-\infty}^{\infty} \rho^2(r) = C_{1/4}^2 \int_0^1 E\left(f^2(B_s)\right) ds.$$

The desired convergence (3.3.21) follows.

Step 2.- Since the sequence (G_n) is bounded in L^1, the sequence $\big(G_n, (B_t)_{t\in[0,1]}\big)$ is tight in $\mathbb{R}\times\mathscr{C}([0,1])$. Assume that $\big(G_\infty, (B_t)_{t\in[0,1]}\big)$ denotes the limit in law of a certain subsequence of $\big(G_n, (B_t)_{t\in[0,1]}\big)$, denoted again by $\big(G_n, (B_t)_{t\in[0,1]}\big)$.

We have to prove that

$$G_\infty \overset{Law}{=} C_{1/4}\int_0^1 f(B_s)dW_s + \frac{1}{4}\int_0^1 f''(B_s)ds,$$

where W denotes a standard Brownian motion independent of B, or equivalently that

$$E\left(e^{i\lambda G_\infty} \,|\, (B_t)_{t\in[0,1]}\right) = \exp\left\{ i\frac{\lambda}{4}\int_0^1 f''(B_s)ds - \frac{\lambda^2}{2} C_{1/4}^2 \int_0^1 f^2(B_s)ds \right\}. \qquad (3.3.24)$$

This will be done by showing that for every random variable ξ of the form (3.3.17) and every real number λ, we have

$$\lim_{n\to\infty} \phi_n'(\lambda) = E\left\{ e^{i\lambda G_\infty}\xi \left(\frac{i}{4}\int_0^1 f''(B_s)ds - \lambda C_{1/4}^2 \int_0^1 f^2(B_s)ds \right) \right\} \qquad (3.3.25)$$

where

$$\phi_n'(\lambda) := \frac{d}{d\lambda} E\left(e^{i\lambda G_n}\xi\right) = iE\left(G_n e^{i\lambda G_n}\xi\right), \quad n \geq 1.$$

Let us make precise this argument. Because $\big(G_\infty, (B_t)_{t\in[0,1]}\big)$ is the limit in law of $\big(G_n, (B_t)_{t\in[0,1]}\big)$ and (G_n) is bounded in L^1, we have that

$$E(G_\infty\,\xi\, e^{i\lambda G_\infty}) = \lim_{n\to\infty} E\left(G_n\,\xi\, e^{i\lambda G_n}\right), \quad \forall \lambda\in\mathbb{R},$$

for every ξ of the form (3.3.17). Furthermore, because convergence (3.3.25) holds for every ξ of the form (3.3.17), the conditional characteristic function $\lambda \mapsto E\left(e^{i\lambda G_\infty}|(B_t)_{t\in[0,1]}\right)$ satisfies the following linear ordinary differential equation:

$$\frac{d}{d\lambda}E\left(e^{i\lambda G_\infty}|(B_t)_{t\in[0,1]}\right) = E\left(e^{i\lambda G_\infty}|(B_t)_{t\in[0,1]}\right)\left[\frac{i}{4}\int_0^1 f''(B_s)ds - \lambda C_{1/4}^2\int_0^1 f^2(B_s)ds\right].$$

By solving it, we obtain (3.3.24), which yields the desired conclusion.

Thus, it remains to show (3.3.25). By the duality between the derivative and divergence operators, we have

$$E\left(f(B_{k/n})I_2(\delta_{k/n}^{\otimes 2})e^{i\lambda G_n}\xi\right) = E\left(\left\langle D^2\left(f(B_{k/n})e^{i\lambda G_n}\xi\right), \delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right). \qquad (3.3.26)$$

The first and second derivatives of $f(B_{k/n})e^{i\lambda G_n}\xi$ are given by

$$D\left(f(B_{k/n})e^{i\lambda G_n}\xi\right) = f'(B_{k/n})e^{i\lambda G_n}\xi\,\varepsilon_{k/n} + i\lambda f(B_{k/n})e^{i\lambda G_n}\xi\,DG_n + f(B_{k/n})e^{i\lambda G_n}D\xi$$

and

$$\begin{aligned} D^2\left(f(B_{k/n})e^{i\lambda G_n}\xi\right) &= f''(B_{k/n})e^{i\lambda G_n}\xi\,\varepsilon_{k/n}^{\otimes 2} + 2i\lambda f'(B_{k/n})e^{i\lambda G_n}\xi\left(\varepsilon_{k/n}\widetilde{\otimes}DG_n\right) \\ &\quad +2f'(B_{k/n})e^{i\lambda G_n}\left(\varepsilon_{k/n}\widetilde{\otimes}D\xi\right) - \lambda^2 f(B_{k/n})e^{i\lambda G_n}\xi\,DG_n^{\otimes 2} \\ &\quad +2i\lambda f(B_{k/n})e^{i\lambda G_n}\left(DG_n\widetilde{\otimes}D\xi\right) \\ &\quad +i\lambda f(B_{k/n})e^{i\lambda G_n}\xi D^2G_n + f(B_{k/n})e^{i\lambda G_n}D^2\xi. \end{aligned}$$

Hence, taking expectation and multiplying by $\delta_{k/n}^{\otimes 2}$ yields

$$\begin{aligned} & E\left(\left\langle D^2\left(f(B_{k/n})e^{i\lambda G_n}\xi\right), \delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right) \\ = \;& E\left(f''(B_{k/n})e^{i\lambda G_n}\xi\right)\left\langle \varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2 + 2i\lambda E\left(f'(B_{k/n})e^{i\lambda G_n}\xi\left\langle DG_n, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right)\left\langle \varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}} \\ & +2E\left(f'(B_{k/n})e^{i\lambda G_n}\left\langle D\xi, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right)\left\langle \varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}} - \lambda^2 E\left(f(B_{k/n})e^{i\lambda G_n}\xi\left\langle DG_n, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2\right) \\ & +2i\lambda E\left(f(B_{k/n})e^{i\lambda G_n}\left\langle D\xi, \delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle DG_n, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right) \\ & +i\lambda E\left(f(B_{k/n})e^{i\lambda G_n}\xi\left\langle D^2G_n, \delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right) + E\left(f(B_{k/n})e^{i\lambda G_n}\left\langle D^2\xi, \delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right). \end{aligned} \tag{3.3.27}$$

We also need explicit expressions for $\left\langle DG_n, \delta_{k/n}\right\rangle_{\mathfrak{H}}$ and for $\left\langle D^2G_n, \delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}$. Differentiating G_n we obtain

$$DG_n = \sum_{l=0}^{n-1}\left[f'(B_{l/n})I_2(\delta_{l/n}^{\otimes 2})\varepsilon_{l/n} + 2f(B_{l/n})\Delta B_{l/n}\delta_{l/n}\right] \tag{3.3.28}$$

and, as a consequence,

$$\left\langle DG_n, \delta_{k/n}\right\rangle_{\mathfrak{H}} = \sum_{l=0}^{n-1} f'(B_{l/n})I_2(\delta_{l/n}^{\otimes 2})\left\langle \varepsilon_{l/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}} + 2\sum_{l=0}^{n-1} f(B_{l/n})\Delta B_{l/n}\left\langle \delta_{l/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}. \tag{3.3.29}$$

Also

$$D^2G_n = \sum_{l=0}^{n-1}\left[f''(B_{l/n})I_2(\delta_{l/n}^{\otimes 2})\varepsilon_{l/n}^{\otimes 2} + 4f'(B_{l/n})\,\Delta B_{l/n}\left(\varepsilon_{l/n}\widetilde{\otimes}\delta_{l/n}\right) + 2f(B_{l/n})\delta_{l/n}^{\otimes 2}\right],$$

and, as a consequence,

$$\left\langle D^2G_n, \delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}} = \sum_{l=0}^{n-1}\left[f''(B_{l/n})I_2(\delta_{l/n}^{\otimes 2})\left\langle \varepsilon_{l/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}^2\right.$$

$$+4f'(B_{l/n})\ \Delta B_{l/n} \left\langle \varepsilon_{l/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}} \left\langle \delta_{l/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}} + 2f(B_{l/n}) \left\langle \delta_{l/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}}^2 \Big] . \tag{3.3.30}$$

Substituting (3.3.30) into (3.3.27) yields the following decomposition for $\phi_n'(\lambda) = i\,E(G_n e^{i\lambda G_n}\xi)$:

$$\begin{aligned} \phi_n'(\lambda) &= -2\lambda \sum_{k,l=0}^{n-1} E\left(f(B_{k/n})f(B_{l/n})e^{i\lambda G_n}\xi\right) \left\langle \delta_{l/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}}^2 \\ &+ i\sum_{k=0}^{n-1} E\left(f''(B_{k/n})e^{i\lambda G_n}\xi\right) \left\langle \varepsilon_{k/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}}^2 + i\sum_{k=0}^{n-1} r_{k,n} \end{aligned} \tag{3.3.31}$$

where $r_{k,n}$ is given by

$$\begin{aligned} r_{k,n} &= 2i\lambda E\left(f'(B_{k/n})e^{i\lambda G_n}\xi \left\langle DG_n, \delta_{k/n} \right\rangle_{\mathfrak{H}}\right) \left\langle \varepsilon_{k/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}} \\ &+2E\left(f'(B_{k/n})e^{i\lambda G_n}\ \left\langle D\xi, \delta_{k/n} \right\rangle_{\mathfrak{H}}\right) \left\langle \varepsilon_{k/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}} - \lambda^2 E\left(f(B_{k/n})e^{i\lambda G_n}\xi \left\langle DG_n, \delta_{k/n} \right\rangle_{\mathfrak{H}}^2\right) \\ &+2i\lambda E\left(f(B_{k/n})e^{i\lambda G_n}\ \left\langle D\xi, \delta_{k/n} \right\rangle_{\mathfrak{H}} \left\langle DG_n, \delta_{k/n} \right\rangle_{\mathfrak{H}}\right) \\ &+i\lambda \sum_{l=0}^{n-1} E\left(f(B_{k/n})e^{i\lambda G_n}\xi\ \ f''(B_{l/n})I_2(\delta_{l/n}^{\otimes 2})\ \right) \left\langle \varepsilon_{l/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}}^2 \\ &+4i\lambda \sum_{l=0}^{n-1} E\left(f(B_{k/n})e^{i\lambda G_n}\xi f'(B_{l/n})\ \Delta B_{l/n}\right) \left\langle \varepsilon_{l/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}} \left\langle \delta_{l/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}} \\ &+E\left(f(B_{k/n})e^{i\lambda G_n} \left\langle D^2\xi, \delta_{k/n}^{\otimes 2} \right\rangle_{\mathfrak{H}^{\otimes 2}}\right) = \sum_{j=1}^{7} R_{k,n}^{(j)}. \end{aligned} \tag{3.3.32}$$

Remark that the first sum in the right hand side of (3.3.31) is very similar to C_n presented in Step 1. In fact, similar computations give

$$\begin{aligned} &\lim_{n\to\infty} -2\lambda \sum_{k,l=0}^{n-1} \mathbb{E}[f(B_{k/n})f(B_{l/n})e^{i\lambda G_n}\xi] \left\langle \delta_{l/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}}^2 \\ &\qquad = -C_{1/4}^2 \lambda \int_0^1 E\left(f^2(B_s)e^{i\lambda G_\infty}\xi\right) ds. \end{aligned} \tag{3.3.33}$$

Furthermore, the second term of (3.3.31) is very similar to $E(G_n)$. In fact, using the arguments presented in Step 1, we obtain here that

$$\lim_{n\to\infty} i\sum_{k=0}^{n-1} E\left(f''(B_{k/n})e^{i\lambda G_n}\xi\right) \left\langle \varepsilon_{k/n}, \delta_{k/n} \right\rangle_{\mathfrak{H}}^2 = \frac{i}{4}\int_0^1 E\left(f''(B_s)e^{i\lambda G_\infty}\xi\right) ds. \tag{3.3.34}$$

Consequently, (3.3.25) will be shown as soon as we will prove that $\lim_{n\to\infty}\sum_{k=0}^{n-1} r_{k,n} = 0$. This will be done in several steps.

Step 3.- In this step, we state and prove some estimates which will be crucial in the rest of the proof. First, we will show that

$$\left|E\left(f'(B_{k/n})f'(B_{l/n})e^{i\lambda G_n}\xi\, I_2(\delta_{l/n}^{\otimes 2})\right)\right| \leq \frac{C}{n} \quad \text{for any } 0\leq k,l\leq n-1. \tag{3.3.35}$$

Then we will prove that

$$\left|E\left(f(B_{k/n})f'(B_{j/n})f'(B_{l/n})e^{i\lambda G_n}\xi\, I_4(\delta_{j/n}^{\otimes 2}\otimes\delta_{l/n}^{\otimes 2})\right)\right| \leq \frac{C}{n^2} \quad \text{for any } 0\leq k,j,l\leq n-1. \tag{3.3.36}$$

Proof of (3.3.35):
Let $\zeta_{\xi,k,n}$ denotes any random variable of the form $f^{(a)}(B_{k/n})f^{(b)}(B_{l/n})e^{i\lambda G_n}\xi$ with a and b two positive integers less or equal to four. From the Malliavin integration by parts formula (3.3.18) we have

$$E\left(f'(B_{k/n})f'(B_{l/n})e^{i\lambda G_n}\xi\, I_2(\delta_{l/n}^{\otimes 2})\right) = E\left(\left\langle D^2\left(f'(B_{k/n})f'(B_{l/n})e^{i\lambda G_n}\xi\right), \delta_{l/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right).$$

When computing the RHS, three types of terms appear. First, we have some terms of the form:

$$\begin{cases} E\left(\zeta_{\xi,k,n}\right)\left\langle \varepsilon_{k/n}, \delta_{l/n}\right\rangle_{\mathfrak{H}}^2, \text{ or} \\ E\left(\zeta_{\xi,k,n}\left\langle D\xi, \delta_{l/n}\right\rangle_{\mathfrak{H}}\right)\left\langle \varepsilon_{k/n}, \delta_{l/n}\right\rangle_{\mathfrak{H}}, \text{ or} \\ E\left(\zeta_{\xi,k,n}\left\langle D^2\xi, \delta_{l/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right), \end{cases} \tag{3.3.37}$$

where $D\xi$ and $D^2\xi$ are given by:

$$\begin{cases} D\xi = \sum_{i=1}^m \frac{\partial\psi}{\partial x_i}(B_{t_1},\dots,B_{t_m})\,\varepsilon_{t_i}, \\ D^2\xi = \sum_{i,j=1}^m \frac{\partial^2\psi}{\partial x_j\partial x_i}(B_{t_1},\dots,B_{t_m})\,\varepsilon_{t_j}\otimes\varepsilon_{t_i}. \end{cases}$$

From Lemma 3.3.4 *(i)* and under $(\mathbf{H}_4)$, we have that each of the three terms in (3.3.37) is less or equal to Cn^{-1}. The second type of terms we have to deal with is

$$\begin{cases} E\left(\zeta_{\xi,k,n}\left\langle DG_n, \delta_{l/n}\right\rangle_{\mathfrak{H}}\right)\left\langle \varepsilon_{k/n}, \delta_{l/n}\right\rangle_{\mathfrak{H}}, \text{ or} \\ E\left(\zeta_{\xi,k,n}\left\langle DG_n, \delta_{l/n}\right\rangle_{\mathfrak{H}}\left\langle D\xi, \delta_{l/n}\right\rangle_{\mathfrak{H}}\right). \end{cases} \tag{3.3.38}$$

By Cauchy-Schwarz inequality, under $(\mathbf{H}_4)$ and by using (4.20) in [83], that is

$$E\left(\left\langle DG_n, \delta_{l/n}\right\rangle_{\mathfrak{H}}^2\right) \leq Cn^{-1},$$

we have that both expressions in (3.3.38) are also less or equal to Cn^{-1}.

The last type of terms which has to be taken into account is the term

$$-\lambda^2 E\left(f'(B_{k/n})f'(B_{l/n})e^{i\lambda G_n}\xi\left\langle D^2G_n, \delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right).$$

Again, by using Cauchy-Schwarz inequality and the estimate

$$E\left(\left\langle D^2G_n, \delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}^2\right) \leq Cn^{-2}$$

(which can be obtained by mimicing the proof of (4.20) in [83]), we can conclude that

$$\left|-\lambda^2 E\left(f'(B_{k/n})f'(B_{l/n})e^{i\lambda G_n}\xi\left\langle D^2G_n, \delta_{k/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right)\right| \leq \frac{C}{n}.$$

As a consequence (3.3.35) is shown.

Proof of (3.3.36):
By the Malliavin integration by parts formula (3.3.18), we have

$$E\left(\zeta_{\xi,k,n}f'(B_{j/n})f'(B_{l/n})I_4(\delta_{j/n}^{\otimes 2}\otimes\delta_{l/n}^{\otimes 2})\right) = E\left(\left\langle D^4\left(\zeta_{\xi,k,n}f'(B_{j/n})f'(B_{l/n})\right), \delta_{j/n}^{\otimes 2}\otimes\delta_{l/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 4}}\right).$$

When computing the RHS, we have to deal with the same type of terms as in the proof of (3.3.35) plus two additional types of terms containing

$$E\left(\left\langle D^3G_n, \delta_{j/n}^{\otimes 2}\otimes\delta_{l/n}\right\rangle_{\mathfrak{H}^{\otimes 3}}^2\right) \quad \text{and} \quad E\left(\left\langle D^4G_n, \delta_{j/n}^{\otimes 2}\otimes\delta_{l/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 4}}^2\right).$$

In fact, by mimicing the proof of (4.20) in [83], we can obtain the following bounds:

$$E\left(\left\langle D^3G_n, \delta_{j/n}^{\otimes 2}\otimes\delta_{l/n}\right\rangle_{\mathfrak{H}^{\otimes 3}}^2\right) \leq Cn^{-3} \quad \text{and} \quad E\left(\left\langle D^4G_n, \delta_{j/n}^{\otimes 2}\otimes\delta_{l/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 4}}^2\right) \leq Cn^{-4}.$$

This allows us to obtain (3.3.36).
Step 4.- We compute the terms corresponding to $R_{k,n}^{(1)}$, $R_{k,n}^{(4)}$ and $R_{k,n}^{(6)}$ in (3.3.32). The derivative DG_n is given by (3.3.28), so that

$$\begin{aligned}
\sum_{k=0}^{n-1} R_{k,n}^{(1)} &= 2i\lambda\sum_{k,l=0}^{n-1} E\left(f'(B_{k/n})f'(B_{l/n})e^{i\lambda G_n}\xi I_2(\delta_{l/n}^{\otimes 2})\right)\left\langle\varepsilon_{l/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle\varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}} \\
&+ 2\sum_{k,l=0}^{n-1} E\left(f'(B_{k/n})f(B_{l/n})e^{i\lambda G_n}\xi\Delta B_{l/n}\right)\left\langle\delta_{l/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle\varepsilon_{k/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}} \\
&= T_1^{(1)} + T_2^{(1)}.
\end{aligned}$$

From (3.3.35), Lemma 3.3.4 *(i)*, *(iii)* and under $(\mathbf{H}_4)$, we have that

$$\left|T_1^{(1)}\right| \leq Cn^{-3/2}\sum_{k,l=0}^{n-1}\left|\left\langle\varepsilon_{l/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \leq Cn^{-1/2}.$$

For $T_2^{(1)}$, remark first that Cauchy-Schwarz inequality and hypothesis $(\mathbf{H}_4)$ yield

$$\left|E\left(f'(B_{k/n})e^{i\lambda G_n}\xi f(B_{l/n})\Delta B_{l/n}\right)\right| \leq Cn^{-1/4}. \tag{3.3.39}$$

Thus, by Lemma 3.3.4 (i),

$$\begin{aligned}\left|T_2^{(1)}\right| &\leq Cn^{-3/4}\sum_{k,l=0}^{n-1}\left|\left\langle \delta_{l/n},\delta_{k/n}\right\rangle_{\mathfrak{H}}\right| = Cn^{-5/4}\sum_{k,l=0}^{n-1}|\rho(k-l)| \\ &\leq Cn^{-1/4}\sum_{r=-\infty}^{\infty}|\rho(r)| = Cn^{-1/4},\end{aligned}$$

where ρ has been defined in (3.3.23).

The term corresponding to $R_{k,n}^{(4)}$ is very similar to $R_{k,n}^{(1)}$. Indeed, by (3.3.28), we have

$$\begin{aligned}\sum_{k=0}^{n-1}R_{k,n}^{(4)} &= 2i\lambda\sum_{i=1}^{m}\sum_{k,l=0}^{n-1}E\left(f(B_{k/n})f'(B_{l/n})e^{i\lambda G_n}\frac{\partial\psi}{\partial x_i}(B_{t_1},\ldots,B_{t_m})I_2(\delta_{l/n}^{\otimes 2})\right) \\ &\qquad\qquad \times\left\langle \varepsilon_{l/n},\delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle \varepsilon_{t_i},\delta_{k/n}\right\rangle_{\mathfrak{H}} \\ &+ 4i\lambda\sum_{i=1}^{m}\sum_{k,l=0}^{n-1}E\left(f(B_{k/n})f(B_{l/n})e^{i\lambda G_n}\Delta B_{l/n}\frac{\partial\psi}{\partial x_i}(B_{t_1},\ldots,B_{t_m})\right) \\ &\qquad\qquad \times\left\langle \delta_{l/n},\delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle \varepsilon_{t_i},\delta_{k/n}\right\rangle_{\mathfrak{H}} \\ &= T_1^{(4)}+T_2^{(4)}\end{aligned}$$

and we can proceed for $T_i^{(4)}$ as for $T_i^{(1)}$.

The term corresponding to $R_{k,n}^{(6)}$ is very similar to $T_2^{(1)}$. More precisely, we have

$$\begin{aligned}\left|\sum_{k=0}^{n-1}R_{k,n}^{(6)}\right| &\leq Cn^{-3/4}\sum_{k,l=0}^{n-1}\left|\left\langle \delta_{l/n},\delta_{k/n}\right\rangle_{\mathfrak{H}}\right| = Cn^{-5/4}\sum_{k,l=0}^{n-1}|\rho(k-l)| \\ &\leq Cn^{-1/4}\sum_{r=-\infty}^{\infty}|\rho(r)| = Cn^{-1/4}.\end{aligned}$$

Step 5.- Estimation of $R_{k,n}^{(3)}$. Let $\zeta_{\xi,k,n} := \lambda^2 f(B_{k/n})e^{i\lambda G_n}\xi$. Using (3.3.28), we have

$$\begin{aligned}\langle DG_n,\delta_{k/n}\rangle_{\mathfrak{H}}^2 &= \sum_{j,l=0}^{n-1}f'(B_{l/n})f'(B_{j/n})I_2(\delta_{l/n}^{\otimes 2})I_2(\delta_{j/n}^{\otimes 2})\langle\varepsilon_{j/n},\delta_{k/n}\rangle_{\mathfrak{H}}\langle\varepsilon_{l/n},\delta_{k/n}\rangle_{\mathfrak{H}} \\ &\quad+\sum_{j,l=0}^{n-1}f(B_{j/n})f(B_{l/n})\Delta B_{j/n}\Delta B_{l/n}\langle\delta_{j/n},\delta_{k/n}\rangle_{\mathfrak{H}}\langle\delta_{l/n},\delta_{k/n}\rangle_{\mathfrak{H}}\end{aligned}$$

and, consequently:

$$\begin{aligned}\left|\sum_{k=0}^{n-1} R_{k,n}^3\right| &\leq \sum_{k=0}^{n-1} \left|E\left(\zeta_{\xi,k,n}\langle DG_n, \delta_{k/n}\rangle_{\mathfrak{H}}^2\right)\right| \\ &\leq 2\sum_{k,j,l=0}^{n-1} \left|E\left(\zeta_{\xi,k,n} f'(B_{j/n}) f'(B_{l/n}) I_2(\delta_{j/n}^{\otimes 2}) I_2(\delta_{l/n}^{\otimes 2})\right)\langle\varepsilon_{j/n}, \delta_{k/n}\rangle_{\mathfrak{H}}\langle\varepsilon_{l/n}, \delta_{k/n}\rangle_{\mathfrak{H}}\right| \\ &+ 8\sum_{k,j,l=0}^{n-1} \left|E\left(\zeta_{\xi,k,n} f(B_{j/n}) f(B_{l/n})\Delta B_{j/n}\Delta B_{l/n}\right)\langle\delta_{j/n}, \delta_{k/n}\rangle_{\mathfrak{H}}\langle\delta_{l/n}, \delta_{k/n}\rangle_{\mathfrak{H}}\right|.\end{aligned}$$

Using the product formula (3.3.22), we have

$$\begin{aligned}\left|\sum_{k=0}^{n-1} R_{k,n}^3\right| &\leq 2\sum_{k,j,l=0}^{n-1} \left|E\left(\zeta_{\xi,k,n} f'(B_{j/n}) f'(B_{l/n}) I_4(\delta_{j/n}^{\otimes 2}\otimes\delta_{l/n}^{\otimes 2})\right)\right| \left|\left\langle\varepsilon_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle\varepsilon_{l/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \\ &+ 8\sum_{k,j,l=0}^{n-1} \left|E\left(\zeta_{\xi,k,n} f'(B_{j/n}) f'(B_{l/n}) I_2(\delta_{j/n}\otimes\delta_{l/n})\right)\right| \left|\left\langle\delta_{j/n}, \delta_{l/n}\right\rangle_{\mathfrak{H}}\left\langle\varepsilon_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle\varepsilon_{l/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \\ &+ 4\sum_{k,j,l=0}^{n-1} \left|E\left(\zeta_{\xi,k,n} f'(B_{j/n}) f'(B_{l/n})\right)\right| \left|\left\langle\delta_{j/n}, \delta_{l/n}\right\rangle_{\mathfrak{H}}^2\left\langle\varepsilon_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle\varepsilon_{l/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \\ &+ 8\sum_{k,j,l=0}^{n-1} \left|E\left(\zeta_{\xi,k,n} f(B_{j/n}) f(B_{l/n})\Delta B_{j/n}\Delta B_{l/n}\right)\right| \left|\left\langle\delta_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle\delta_{l/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \\ &= \sum_{i=1}^{4} T_i^{(3)}.\end{aligned}$$

From (3.3.36), we have

$$\begin{aligned}|T_1^{(3)}| &\leq Cn^{-1/2}\sum_{k,j,l=0}^{n-1} \left|E\left(\zeta_{\xi,k,n} f'(B_{j/n}) f'(B_{l/n}) I_4(\delta_{j/n}^{\otimes 2}\otimes\delta_{l/n}^{\otimes 2})\right)\right| \left|\left\langle\varepsilon_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \\ &\leq Cn^{-5/2}n^2 \sup_{j=0,\dots n-1}\sum_{k=0}^{n-1}\left|\left\langle\varepsilon_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \leq Cn^{-1/2} \text{ by Lemma 3.3.4}\,(ii).\end{aligned}$$

Now, let us consider $T_2^{(3)}$. Using (3.3.35) and Lemma 3.3.4 *(ii)*, we deduce that

$$\begin{aligned}\left|T_2^{(3)}\right| &\leq Cn^{-3/2}\sum_{j,l=0}^{n-1}\left|\left\langle\delta_{j/n}, \delta_{l/n}\right\rangle_{\mathfrak{H}}\right| \sup_{j=0,\dots,n-1}\sum_{k=0}^{n-1}\left|\left\langle\varepsilon_{j/n}, \delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \\ &\leq Cn^{-1/2}\sum_{r=-\infty}^{\infty}|\rho(r)| = Cn^{-1/2}.\end{aligned}$$

For $T_3^{(3)}$, we have

$$\begin{aligned}\left|T_3^{(3)}\right| &\leq Cn^{-1/2}\sum_{j,l=0}^{n-1}\left\langle\delta_{j/n},\delta_{l/n}\right\rangle_{\mathfrak{H}}^2 \sup_{j=0,\ldots,n-1}\sum_{k=0}^{n-1}\left|\left\langle\varepsilon_{j/n},\delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \\ &\leq Cn^{-1/2}\sum_{r=-\infty}^{\infty}\rho^2(r) = Cn^{-1/2}.\end{aligned}$$

Finally, by Cauchy-Schwarz inequality and under $(\mathbf{H}_4)$, we have

$$\left|E\left(\zeta_{\xi,k,n}f(B_{j/n})f(B_{l/n})\Delta B_{j/n}\Delta B_{l/n}\right)\right| \leq Cn^{-1/2}.$$

Consequently:

$$\begin{aligned}\left|T_4^{(3)}\right| &\leq Cn^{-1/2}\sum_{k,j,l=0}^{n-1}\left|\left\langle\delta_{j/n},\delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle\delta_{k/n},\delta_{l/n}\right\rangle_{\mathfrak{H}}\right| \\ &\leq Cn^{-3/2}\sum_{k,j,l=0}^{n-1}\left|\rho(k-l)\rho(k-j)\right| \leq Cn^{-1/2}\left(\sum_{r=-\infty}^{\infty}|\rho(r)|\right)^2 = Cn^{-1/2}.\end{aligned}$$

Step 6.- Estimation of $R_{k,n}^{(5)}$. From (3.3.35) and Lemma 3.3.4 *(iii)*, we have,

$$\left|\sum_{k=0}^{n-1}R_{k,n}^{(5)}\right| \leq Cn^{-3/2}\sum_{k,l=0}^{n-1}\left|\left\langle\varepsilon_{l/n},\delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \leq Cn^{-1/2}.$$

Step 7.- Estimation of $R_{k,n}^{(2)}$ and $R_{k,n}^{(7)}$. We recall that

$$0 \leq \sqrt{x+1}-\sqrt{x} \leq 1 \quad \text{for any } x \geq 0.$$

Thus, under $(\mathbf{H}_4)$ and using Lemma 3.3.4, we have:

$$\begin{aligned}\left|\sum_{k=0}^{n-1}R_{k,n}^{(2)}\right| &\leq 2\sum_{i=1}^{m}\sum_{k=0}^{n-1}\left|E\left(f'(B_{k/n})e^{i\lambda G_n}\frac{\partial\psi}{\partial x_i}(B_{t_1},\ldots,B_{t_m})\right)\left\langle\varepsilon_{t_i},\delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle\varepsilon_{k/n},\delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \\ &\leq C(f,\psi)\,n^{-\frac{1}{2}}\sup_{t\in[0,1]}\sum_{k=0}^{n-1}\left|\left\langle\varepsilon_t,\delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \leq Cn^{-1/2}.\end{aligned}$$

Similarly, the following bound holds:

$$\begin{aligned}\left|\sum_{k=0}^{n-1}R_{k,n}^{(7)}\right| &\leq \sum_{i,j=1}^{m}\sum_{k=0}^{n-1}\left|E\left(f(B_{k/n})e^{i\lambda G_n}\frac{\partial^2\psi}{\partial x_j\partial x_i}(B_{t_1},\ldots,B_{t_m})\right)\left\langle\varepsilon_{t_i},\delta_{k/n}\right\rangle_{\mathfrak{H}}\left\langle\varepsilon_{t_j},\delta_{k/n}\right\rangle_{\mathfrak{H}}\right| \\ &\leq Cn^{-1/2}.\end{aligned}$$

The proof of Theorem 3.3.5 is done. □

3.3.4 Proof of Theorem 3.3.2

Once again, $B = B^{1/4}$ denotes a fractional Brownian motion with Hurst index $H = 1/4$. Moreover, we recall that we note $\Delta B_{k/n}$ (resp. $\delta_{k/n}$; $\varepsilon_{k/n}$) instead of $B_{(k+1)/n} - B_{k/n}$ (resp. $\mathbf{1}_{[k/n,(k+1)/n]}$; $\mathbf{1}_{[0,k/n]}$). The aim of this section is to prove Theorem 3.3.2, or equivalently:

Theorem 3.3.6. *(Itô's formula) Let $f : \mathbb{R} \to \mathbb{R}$ verifying ($\mathbf{H}_9$). Then*

$$\int_0^1 f'(B_s)d^\star B_s := \lim_{n\to\infty} \sum_{k=1}^{\lfloor n/2 \rfloor} f'(B_{(2k-1)/n}) \left(B_{(2k)/n} - B_{(2k-2)/n}\right) \quad \textit{exists in law}$$

and we have

$$\int_0^1 f'(B_s)d^\star B_s \overset{\text{Law}}{=} f(B_1) - f(0) - \frac{\kappa}{2}\int_0^1 f''(B_s)dW_s,$$

with κ defined by

$$\kappa = \sqrt{2 + \sum_{r=1}^{\infty}(-1)^r \rho^2(r)} = 1,290\ldots. \tag{3.3.40}$$

and where W denotes a standard Brownian motion independent of B.

Proof. In [127], identity (1.6), it is proved that

$$\begin{aligned}
&\sum_{k=1}^{\lfloor n/2 \rfloor} f'(B_{(2k-1)/n}) \left(B_{(2k)/n} - B_{(2k-2)/n}\right) \\
\approx\ & f(B_1) - f(0) - \frac{1}{2}\sum_{k=1}^{\lfloor n/2 \rfloor} f''(B_{(2k-1)/n}) \left[(\Delta B_{(2k-1)/n})^2 - (\Delta B_{(2k-2)/n)})^2\right] \\
&-\frac{1}{6}\sum_{j=1}^{\lfloor n/2 \rfloor} f'''(B_{(2j-1)/n})\left[(\Delta B_{(2j-2)/n})^3 + (\Delta B_{(2j-1)/n})^3\right]
\end{aligned}$$

where "≈" means the difference goes to zero in L^2. Therefore, Theorem 3.3.6 is a direct consequence of Lemmas 3.3.7 and 3.3.8 below. □

Lemma 3.3.7. *Let $f : \mathbb{R} \to \mathbb{R}$ verifying ($\mathbf{H}_6$). Then*

$$\sum_{j=1}^{\lfloor n/2 \rfloor} f(B_{(2j-1)/n})\left[(\Delta B_{(2j-2)/n})^3 + (\Delta B_{(2j-1)/n})^3\right] \xrightarrow[n\to\infty]{L^2} 0. \tag{3.3.41}$$

Proof. Let $H_3(x) = x^3 - 3x$ be the third Hermite polynomial. Using the relation between Hermite polynomial and multiple integral (see Section 2), remark that

$$(\Delta B_{(2j-2)/n})^3 + (\Delta B_{(2j-1)/n})^3 \quad = \quad n^{-\frac{3}{4}}\left[H_3\left(n^{\frac{1}{4}}\Delta B_{(2j-2)/n}\right) + H_3\left(n^{\frac{1}{4}}\Delta B_{(2j-1)/n}\right)\right.$$

$$+\frac{3}{\sqrt{n}}\big(B_{(2j-2)/n}-B_{(2j)/n}\big)\Big]$$
$$= I_3\big(\delta_{(2j-2)/n}^{\otimes 3}\big)+I_3\big(\delta_{(2j-1)/n}^{\otimes 3}\big)+\frac{3}{\sqrt{n}}\,I_1\big(\mathbf{1}_{[(2j-2)/n,(2j)/n]}\big)$$

so that (3.3.41) can be shown by successively proving that

$$E\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor n/2\rfloor} f(B_{(2j-1)/n})\,I_1\big(\mathbf{1}_{[(2j-2)/n,(2j)/n]}\big)\right|^2 \underset{n\to+\infty}{\longrightarrow} 0; \qquad (3.3.42)$$

$$E\left|\sum_{j=1}^{\lfloor n/2\rfloor} f(B_{(2j-1)/n})\,I_3\big(\delta_{(2j-2)/n}^{\otimes 3}\big)\right|^2 \underset{n\to+\infty}{\longrightarrow} 0; \qquad (3.3.43)$$

$$E\left|\sum_{j=1}^{\lfloor n/2\rfloor} f(B_{(2j-1)/n})\,I_3\big(\delta_{(2j-1)/n}^{\otimes 3}\big)\right|^2 \underset{n\to+\infty}{\longrightarrow} 0. \qquad (3.3.44)$$

Let us first proceed with the proof of (3.3.42). We can write, using in particular (3.3.19):

$$E\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor n/2\rfloor} f(B_{(2j-1)/n})\,I_1\big(\mathbf{1}_{[(2j-2)/n,(2j)/n]}\big)\right|^2$$
$$= \frac{1}{n}\left|\sum_{j,k=1}^{\lfloor n/2\rfloor} E\big\{f(B_{(2j-1)/n})f(B_{(2k-1)/n})\,I_1\big(\mathbf{1}_{[(2j-2)/n,(2j)/n]}\big)\,I_1\big(\mathbf{1}_{[(2k-2)/n,(2k)/n]}\big)\big\}\right|$$
$$\le \frac{1}{n}\sum_{j,k=1}^{\lfloor n/2\rfloor}\big|E\big\{f(B_{(2j-1)/n})f(B_{(2k-1)/n})\,I_2\big(\mathbf{1}_{[(2j-2)/n,(2j)/n]}\otimes\mathbf{1}_{[(2k-2)/n,(2k)/n]}\big)\big\}\big|$$
$$+\frac{1}{n\sqrt{n}}\sum_{j,k=1}^{\lfloor n/2\rfloor}\big|E\big\{f(B_{(2j-1)/n})f(B_{(2k-1)/n})\rho(2j-2k)\big\}\big|$$
$$= \frac{2}{n}\sum_{a+b=2}\sum_{j,k=1}^{\lfloor n/2\rfloor}\big|E\big\{f^{(a)}(B_{(2j-1)/n})f^{(b)}(B_{(2k-1)/n})\big\}\big|$$
$$\times\left|\left\langle \varepsilon_{(2j-1)/n}^{\otimes a}\otimes\varepsilon_{(2k-1)/n}^{\otimes b},\mathbf{1}_{[(2j-2)/n,(2j)/n]}\otimes\mathbf{1}_{[(2k-2)/n,(2k)/n]}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right|$$
$$+\frac{1}{n\sqrt{n}}\sum_{j,k=1}^{\lfloor n/2\rfloor}\big|E\big\{f(B_{(2j-1)/n})f(B_{(2k-1)/n})\rho(2j-2k)\big\}\big|.$$

But, by Lemma 3.3.4 *(i)*, we have

$$\left|\left\langle \varepsilon_{(2j-1)/n}^{\otimes a}\otimes\varepsilon_{(2k-1)/n}^{\otimes b},\mathbf{1}_{[(2j-2)/n,(2j)/n]}\otimes\mathbf{1}_{[(2k-2)/n,(2k)/n]}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right|$$

$$\leq \frac{1}{\sqrt{n}}\Big(\big|\left\langle\varepsilon_{(2j-1)/n},\mathbf{1}_{[(2j-2)/n,(2j)/n]}\right\rangle_{\mathfrak{H}}\big|+\big|\left\langle\varepsilon_{(2k-1)/n},\mathbf{1}_{[(2k-2)/n,(2k)/n]}\right\rangle_{\mathfrak{H}}\big|\Big)$$
$$= \frac{1}{n}\big(\sqrt{2j}-\sqrt{2j-2}+\sqrt{2k}-\sqrt{2k-2}\big).$$

Thus, under ($\mathbf{H}_6$):

$$\sum_{a+b=2}\sum_{j,k=1}^{\lfloor n/2\rfloor}\big|E\{f^{(a)}(B_{(2j-1)/n})f^{(b)}(B_{(2k-1)/n})\}\big|$$
$$\times\left|\left\langle\varepsilon_{(2j-1)/n}^{\otimes a}\otimes\varepsilon_{(2k-1)/n}^{\otimes b},\mathbf{1}_{[(2j-2)/n,(2j)/n]}\otimes\mathbf{1}_{[(2k-2)/n,(2k)/n]}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right|=O(\sqrt{n}).$$

Moreover

$$\sum_{j,k=1}^{\lfloor n/2\rfloor}\big|E\{f(B_{(2j-1)/n})f(B_{(2k-1)/n})\rho(2j-2k)\}\big|\leq C\sum_{j,k=1}^{\lfloor n/2\rfloor}\big|\rho(2j-2k)\big|=O(n).$$

Finally, convergence (3.3.42) holds.

Now, let us only proceed with the proof of (3.3.43), the proof of (3.3.44) being similar. We have

$$E\left|\sum_{j=1}^{\lfloor n/2\rfloor}f(B_{(2j-1)/n})\,I_3\big(\delta_{(2j-2)/n}^{\otimes 3}\big)\right|^2$$
$$=\sum_{j,k=1}^{\lfloor n/2\rfloor}E\{f(B_{(2j-1)/n})f(B_{(2k-1)/n})I_3\big(\delta_{(2j-2)/n}^{\otimes 3}\big)I_3\big(\delta_{(2k-2)/n}^{\otimes 3}\big)\}$$
$$=\sum_{r=0}^{3}r!\binom{3}{r}^2 n^{-\frac{3-r}{2}}\sum_{j,k=1}^{\lfloor n/2\rfloor}E\{f(B_{(2j-1)/n})f(B_{(2k-1)/n})I_{2r}\big(\delta_{(2j-2)/n}^{\otimes r}\otimes\delta_{(2k-2)/n}^{\otimes r}\big)\}\rho^{3-r}(2j-2k).$$

To obtain (3.3.43), it is then sufficient to prove that, for every fixed $r\in\{0,1,2,3\}$, the quantities

$$R_n^{(r)}=n^{-\frac{3-r}{2}}\sum_{j,k=1}^{\lfloor n/2\rfloor}E\{f(B_{(2j-1)/n})f(B_{(2k-1)/n})I_{2r}\big(\delta_{(2j-2)/n}^{\otimes r}\otimes\delta_{(2k-2)/n}^{\otimes r}\big)\}\rho^{3-r}(2j-2k)$$

tend to zero as $n\to\infty$. We have, by Lemma 3.3.4 *(i)* and under ($\mathbf{H}_6$):

$$\sup_{j,k=1,\ldots,\lfloor n/2\rfloor}\left|E\{f(B_{(2j-1)/n})f(B_{(2k-1)/n})I_{2r}\big(\delta_{(2j-2)/n}^{\otimes r}\otimes\delta_{(2k-2)/n}^{\otimes r}\big)\}\right|$$
$$=\sup_{j,k=1,\ldots,\lfloor n/2\rfloor}(2r)!\left|\sum_{a+b=2r}E\left\{f^{(a)}(B_{(2j-1)/n})f^{(b)}(B_{(2k-1)/n})\right\}\right.$$

$$\times \left\langle \varepsilon_{(2j-1)/n}^{\otimes a} \widetilde{\otimes} \varepsilon_{(2j-1)/n}^{\otimes b}, \mathbf{1}_{[(2j-2)/n,(2j/n)]}^{\otimes r} \otimes \mathbf{1}_{[(2k-2)/n,(2k/n)]}^{\otimes r} \right\rangle_{\mathfrak{H}^{\otimes 2}} \Big|$$

$$\leq C \sup_{j,k=1,\dots,\lfloor n/2 \rfloor} \sup_{a+b=2r} \left| \left\langle \varepsilon_{(2j-1)/n}^{\otimes a} \widetilde{\otimes} \varepsilon_{(2j-1)/n}^{\otimes b}, \mathbf{1}_{[(2j-2)/n,(2j/n)]}^{\otimes r} \otimes \mathbf{1}_{[(2k-2)/n,(2k/n)]}^{\otimes r} \right\rangle_{\mathfrak{H}^{\otimes 2}} \right|$$

$$= O(n^{-r}).$$

Consequently, when $r \neq 3$, we deduce

$$\left|R_n^{(r)}\right| \leq C\, n^{-\frac{r+3}{2}} \sum_{j,k=1}^{\lfloor n/2 \rfloor} \left|\rho(2j-2k)\right| = O(n^{-\frac{r+1}{2}}) \underset{n\to+\infty}{\longrightarrow} 0$$

while, when $r = 3$, we deduce

$$\left|R_n^{(3)}\right| \leq C\, n^{-1} \underset{n\to+\infty}{\longrightarrow} 0.$$

The proof of (3.3.43) is done. Since the proof of (3.3.44) follows the same lines, we finally proved (3.3.41). □

Lemma 3.3.8. *Let $f : \mathbb{R} \to \mathbb{R}$ verifying $(\mathbf{H}_4)$. Set*

$$F_n = \sum_{k=1}^{\lfloor n/2 \rfloor} f(B_{(2k-1)/n}) \left[(\Delta B_{(2k-1)/n})^2 - (\Delta B_{(2k-2)/n})^2\right].$$

Then

$$F_n \quad \xrightarrow[n\to\infty]{\text{stably}} \quad \kappa \int_0^1 f(B_s) dW_s, \tag{3.3.45}$$

with κ defined by (3.3.40), and where W denotes a standard Brownian motion independent of B. Here, the stable convergence (3.3.45) has to be understood in the following sense: for any real number λ and any $\sigma\{B\}$-measurable and integrable random variable ξ, we have that

$$E\left(e^{i\lambda F_n}\,\xi\right) \underset{n\to\infty}{\longrightarrow} E\left(e^{-\frac{\lambda^2\kappa^2}{2}\int_0^1 f^2(B_s)ds}\,\xi\right).$$

Proof. Since we follow exactly the proof of Theorem 3.3.5, we only describe the main ideas. First, observe that

$$F_n = \sum_{k=1}^{\lfloor n/2 \rfloor} f\left(B_{(2k-1)/n}\right)\left(I_2(\delta_{(2k-1)/n}^{\otimes 2}) - I_2(\delta_{(2k-2)/n}^{\otimes 2})\right).$$

Here, the analogue of Lemma 3.3.4 is:

$$\sup_{t\in[0,1]} \sum_{k=1}^{\lfloor n/2 \rfloor} \left|\left\langle \varepsilon_t, \delta_{(2k-1)/n} \right\rangle_{\mathfrak{H}}\right| \underset{n\to\infty}{=} O(1), \quad \sup_{t\in[0,1]} \sum_{k=1}^{\lfloor n/2 \rfloor} \left|\left\langle \varepsilon_t, \delta_{(2k-2)/n} \right\rangle_{\mathfrak{H}}\right| \underset{n\to\infty}{=} O(1), \tag{3.3.46}$$

$$\left| \left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-1)/n} \right\rangle_{\mathfrak{H}}^2 - \frac{1}{4n} \right| \leq \frac{\sqrt{2k} - \sqrt{2k-1}}{4n} \tag{3.3.47}$$

and

$$\left| \left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-2)/n} \right\rangle_{\mathfrak{H}}^2 - \frac{1}{4n} \right| \leq \frac{\sqrt{2k-1} - \sqrt{2k-2}}{2n}. \tag{3.3.48}$$

In fact, the bounds (3.3.46) are obtained by following the arguments presented in the proof of Lemma 3.3.4. The only difference is that, in order to bound sums of the type $\sum_{k=1}^{\lfloor n/2 \rfloor} \sqrt{2k} - \sqrt{2k-1}$ (which are no more telescopic), we use

$$\sum_{k=1}^{\lfloor n/2 \rfloor} \sqrt{2k} - \sqrt{2k-1} \leq \sum_{k=1}^{\lfloor n/2 \rfloor} \sqrt{2k} - \sqrt{2k-2} = \sqrt{2\lfloor n/2 \rfloor} \leq \sqrt{n}.$$

As in Step 1 of the proof of Theorem 3.3.5, here we also have that (F_n) is bounded in L^2. Consequently the sequence $\left(F_n, (B_t)_{t\in[0,1]}\right)$ is tight in $\mathbb{R} \times \mathscr{C}([0,1])$. Assume that $\left(F_\infty, (B_t)_{t\in[0,1]}\right)$ denotes the limit in law of a certain subsequence of $\left(F_n, (B_t)_{t\in[0,1]}\right)$, denoted again by $\left(F_n, (B_t)_{t\in[0,1]}\right)$.

We have to prove that

$$E\left(e^{i\lambda F_\infty} \middle| (B_t)_{t\in[0,1]}\right) = \exp\left\{ -\frac{\lambda^2}{2} \kappa^2 \int_0^1 f^2(B_s) ds \right\}. \tag{3.3.49}$$

We proceed as in Step 2 of the proof of Theorem 3.3.5. That is, (3.3.49) will be obtained by showing that for every random variable ξ of the form (3.3.17) and every real number λ, we have

$$\lim_{n\to\infty} \phi_n'(\lambda) = -\lambda \kappa^2 E\left(e^{i\lambda F_\infty} \xi \int_0^1 f^2(B_s) ds \right)$$

where

$$\phi_n'(\lambda) := \frac{d}{d\lambda} E\left(e^{i\lambda F_n} \xi\right) = iE\left(F_n e^{i\lambda F_n} \xi\right), \quad n \geq 1.$$

By the duality formula (3.3.18) we have that

$$\phi_n'(\lambda) = \sum_{k=1}^{\lfloor n/2 \rfloor} E\left(\left\langle D^2\left(f(B_{(2k-1)/n}) e^{i\lambda F_n} \xi\right), \delta_{(2k-1)/n}^{\otimes 2} - \delta_{(2k-2)/n}^{\otimes 2} \right\rangle_{\mathfrak{H}^{\otimes 2}} \right).$$

The analogue of (3.3.27) is here:

$$\begin{aligned} &\left\langle E\left(D^2\left(f(B_{(2k-1)/n}) e^{i\lambda F_n} \xi\right)\right), \delta_{(2k-1)/n}^{\otimes 2} - \delta_{(2k-2)/n}^{\otimes 2} \right\rangle_{\mathfrak{H}^{\otimes 2}} \\ = \;& E\left(f''(B_{(2k-1)/n}) e^{i\lambda F_n} \xi\right) \left[\left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-1)/n} \right\rangle_{\mathfrak{H}}^2 - \left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-2)/n} \right\rangle_{\mathfrak{H}}^2 \right] \\ &+ 2i\lambda E\left(f'(B_{(2k-1)/n}) e^{i\lambda F_n} \xi \left\langle DF_n, \delta_{(2k-1)/n} \right\rangle_{\mathfrak{H}} \right) \left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-1)/n} \right\rangle_{\mathfrak{H}} \\ &- 2i\lambda E\left(f'(B_{(2k-1)/n}) e^{i\lambda F_n} \xi \left\langle DF_n, \delta_{(2k-2)/n} \right\rangle_{\mathfrak{H}} \right) \left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-2)/n} \right\rangle_{\mathfrak{H}} \end{aligned}$$

$$
\begin{aligned}
&+2E\left(f'(B_{(2k-1)/n})e^{i\lambda F_n}\ \left\langle D\xi,\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}\right)\left\langle\varepsilon_{(2k-1)/n},\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}\\
&-2E\left(f'(B_{(2k-1)/n})e^{i\lambda F_n}\ \left\langle D\xi,\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}\right)\left\langle\varepsilon_{(2k-1)/n},\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}\\
&-\lambda^2E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\xi\left\langle DF_n,\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}^2\right)\\
&+\lambda^2E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\xi\left\langle DF_n,\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}^2\right)\\
&+2i\lambda E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\ \left\langle D\xi,\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}\left\langle DF_n,\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}\right)\\
&-2i\lambda E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\ \left\langle D\xi,\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}\left\langle DF_n,\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}\right)\\
&+i\lambda E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\xi\left\langle D^2F_n,\delta_{(2k-1)/n}^{\otimes 2}-\delta_{(2k-2)/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right)\\
&+E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\left\langle D^2\xi,\delta_{(2k-1)/n}^{\otimes 2}-\delta_{(2k-2)/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right).
\end{aligned}
$$

As a consequence,

$$
\begin{aligned}
\phi_n'(\lambda) &= -2\lambda\sum_{k,l=0}^{n-1}E\left(f(B_{(2k-1)/n})f(B_{(2l-1)/n})e^{i\lambda F_n}\xi\right)\\
&\quad\times\left\langle\delta_{(2l-1)/n}^{\otimes 2}-\delta_{(2l-2)/n}^{\otimes 2},\delta_{(2k-1)/n}^{\otimes 2}-\delta_{(2k-2)/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}+i\sum_{k=0}^{n-1}r_{k,n} \qquad (3.3.50)
\end{aligned}
$$

where $r_{k,n}$ is given by

$$
\begin{aligned}
r_{k,n} &= E[f''(B_{(2k-1)/n})e^{i\lambda F_n}\xi]\left[\left\langle\varepsilon_{(2k-1)/n},\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}^2-\left\langle\varepsilon_{(2k-1)/n},\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}^2\right]\\
&+2i\lambda E\left(f'(B_{(2k-1)/n})e^{i\lambda F_n}\xi\left\langle DF_n,\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}\right)\left\langle\varepsilon_{(2k-1)/n},\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}\\
&-2i\lambda E\left(f'(B_{(2k-1)/n})e^{i\lambda F_n}\xi\left\langle DF_n,\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}\right)\left\langle\varepsilon_{(2k-1)/n},\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}\\
&+2E\left(f'(B_{(2k-1)/n})e^{i\lambda F_n}\ \left\langle D\xi,\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}\right)\left\langle\varepsilon_{(2k-1)/n},\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}\\
&-2E\left(f'(B_{(2k-1)/n})e^{i\lambda F_n}\ \left\langle D\xi,\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}\right)\left\langle\varepsilon_{(2k-1)/n},\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}\\
&-\lambda^2E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\xi\left\langle DF_n,\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}^2\right)\\
&+\lambda^2E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\xi\left\langle DF_n,\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}^2\right)\\
&+2i\lambda E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\ \left\langle D\xi,\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}\left\langle DF_n,\delta_{(2k-1)/n}\right\rangle_{\mathfrak{H}}\right)\\
&-2i\lambda E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\ \left\langle D\xi,\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}\left\langle DF_n,\delta_{(2k-2)/n}\right\rangle_{\mathfrak{H}}\right)\\
&+E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\left\langle D^2\xi,\delta_{(2k-1)/n}^{\otimes 2}-\delta_{(2k-2)/n}^{\otimes 2}\right\rangle_{\mathfrak{H}^{\otimes 2}}\right)\\
&+i\lambda\sum_{l=1}^{\lfloor n/2\rfloor}E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\xi\ \ f''(B_{(2l-1)/n})(I_2(\delta_{(2l-1)/n}^{\otimes 2})-I_2(\delta_{(2l-2)/n}^{\otimes 2})\ \right)
\end{aligned}
$$

$$\times \left\langle \varepsilon^{\otimes 2}_{(2l-1)/n}, \delta^{\otimes 2}_{(2k-1)/n} - \delta^{\otimes 2}_{(2k-2)/n} \right\rangle_{\mathfrak{H}^{\otimes 2}}$$

$$+4i\lambda \sum_{l=1}^{\lfloor n/2 \rfloor} E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\xi f'(B_{(2l-1)/n})\Delta B_{(2l-1)/n}\right)$$

$$\times \left\langle \varepsilon_{(2l-1)/n}\widetilde{\otimes}\delta_{(2l-1)/n}, \delta^{\otimes 2}_{(2k-1)/n} - \delta^{\otimes 2}_{(2k-2)/n} \right\rangle_{\mathfrak{H}^{\otimes 2}}$$

$$-4i\lambda \sum_{l=1}^{\lfloor n/2 \rfloor} E\left(f(B_{(2k-1)/n})e^{i\lambda F_n}\xi f'(B_{(2l-1)/n})\Delta B_{(2l-2)/n}\right)$$

$$\times \left\langle \varepsilon_{(2l-2)/n}\widetilde{\otimes}\delta_{(2l-2)/n}, \delta^{\otimes 2}_{(2k-1)/n} - \delta^{\otimes 2}_{(2k-2)/n} \right\rangle_{\mathfrak{H}^{\otimes 2}}$$

$$= \sum_{j=1}^{13} R^j_{k,n}. \tag{3.3.51}$$

The only difference with respect to (3.3.31) is that, this time, the term

$$i\sum_{k=0}^{n-1} E[f''(B_{(2k-1)/n})e^{i\lambda F_n}\xi] \left[\left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-1)/n}\right\rangle^2_{\mathfrak{H}} - \left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-2)/n}\right\rangle^2_{\mathfrak{H}}\right]$$

corresponding to (3.3.34) is negligible. Indeed, we can write

$$\begin{aligned}
&\sum_{k=0}^{n-1} E[f''(B_{(2k-1)/n})e^{i\lambda F_n}\xi] \left[\left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-1)/n}\right\rangle^2_{\mathfrak{H}} - \left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-2)/n}\right\rangle^2_{\mathfrak{H}}\right] \\
= &\sum_{k=0}^{n-1} E\left(f''(B_{(2k-1)/n})e^{i\lambda F_n}\xi\right) \left[\left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-1)/n}\right\rangle^2_{\mathfrak{H}} - \frac{1}{4n}\right] \\
&\qquad -\sum_{k=0}^{n-1} E\left(f''(B_{(2k-1)/n})e^{i\lambda F_n}\xi\right) \left[\left\langle \varepsilon_{(2k-1)/n}, \delta_{(2k-2)/n}\right\rangle^2_{\mathfrak{H}} - \frac{1}{4n}\right] \\
\underset{n\to\infty}{\longrightarrow}\ & 0 \quad \text{by (3.3.47)-(3.3.48), under } (\mathbf{H}_4).
\end{aligned}$$

Moreover, exactly as in the proof of Theorem 3.3.5, we can show that $\lim_{n\to\infty}\sum_{k=1}^{\lfloor n/2\rfloor} r_{k,n} = 0$. Consequently, we have

$$\begin{aligned}
&\lim_{n\to\infty} \phi'_n(\lambda) \\
= &-2\lambda \lim_{n\to\infty} \sum_{k,l=1}^{\lfloor n/2 \rfloor} E\left(f(B_{(2k-1)/n})f(B_{(2l-1)/n})e^{i\lambda F_n}\xi\right) \left\langle \delta^{\otimes 2}_{(2l-1)/n} - \delta^{\otimes 2}_{(2l-2)/n}, \delta^{\otimes 2}_{(2k-1)/n} - \delta^{\otimes 2}_{(2k-2)/n}\right\rangle_{\mathfrak{H}^{\otimes 2}} \\
= &-\frac{\lambda}{2} \lim_{n\to\infty} \frac{1}{n} \sum_{k,l=1}^{\lfloor n/2 \rfloor} E\left(f(B_{(2k-1)/n})f(B_{(2l-1)/n})e^{i\lambda F_n}\xi\right) \\
&\qquad \times\left(2\rho^2(2k-2l) - \rho^2(2l-2k+1) - \rho^2(2l-2k-1)\right)
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{\lambda}{4} \sum_{r=-\infty}^{\infty} \big(2\,\rho^2(2r) - \rho^2(2r+1) - \rho^2(2r-1)\big) \\
&\qquad \times \lim_{n\to\infty} \frac{2}{n} \sum_{k=1\vee(1-r)}^{\lfloor n/2\rfloor \wedge (\lfloor n/2\rfloor - r)} E\left(f(B_{(2k-1)/n}) f(B_{(2k-1-2r)/n}) e^{i\lambda F_n} \xi\right) \\
&= -\lambda\,\kappa^2 \int_0^1 E\left(f^2(B_s) e^{i\lambda F_\infty} \xi\right) ds,
\end{aligned}
$$

where κ is defined by (3.3.40). In other words, (3.3.49) is shown and the proof of Lemma 3.3.8 is done. □

Acknowledgments

Some of our computations are inspired by the first draft of [83]. We are grateful to David Nualart for letting us use them freely.

Chapitre 4

Méthode de Stein pour l'approximation gaussienne multidimensionnelle

Ce chapitre réalisé en collaboration avec Ivan Nourdin et Giovanni Peccati est publié en [89] .

4.1 Introduction

Let $Z \sim \mathscr{N}(0,1)$ be a standard Gaussian random variable on some probability space $(\Omega, \mathcal{F}, P)$, and let F be a real-valued functional of an infinite-dimensional Gaussian field. In the papers [86, 87] it is shown that one can combine Stein's method (see e.g. [26], [123] or [125]) with Malliavin calculus (see e.g. [94]), in order to deduce explicit (and, sometimes, optimal) bounds for quantities of the type $d(F, Z)$, where d stands for some distance between the law of F and the law of Z (e.g., d can be the Kolmogorov or the Wasserstein distance). The aim of this paper is to extend the results of [86, 87] to the framework of the *multidimensional* Gaussian approximation in the Wasserstein distance. Once again, our techniques hinge upon the use of infinite-dimensional operators on Gaussian spaces (like the *divergence operator* or the *Ornstein-Uhlenbeck generator*) and upon an appropriate multidimensional version of Stein's method (in a form close to Chatterjee and Meckes [25], but see also Reinert and Röllin [112]). As a result, we will obtain explicit bounds, both in terms of Malliavin derivatives and contraction operators, thus providing a substantial refinement of the main findings by Nualart and Ortiz-Latorre [95] and Peccati and Tudor [100]. Note that an important part of our computations (see e.g. Lemma 4.3.7) are directly inspired by those contained in [95]: we shall indeed stress that this last reference contains a fundamental methodological breakthrough, showing that one can deal with (possibly multidimensional) weak convergence on a Gaussian space, by means of Malliavin-type operators and "characterizing" differential equations. See [85] for an application of these techniques to non-central limit theorems. Incidentally, observe that the paper [99], which is mainly based on martingale-type techniques, also uses distances between probability measures (such as the Prokhorov distance) to deal with multidimensional Gaussian approximations on Wiener space, but without giving explicit bounds.

The rationale behind Stein's method is better understood in dimension one. In this framework, the starting point is the following crucial result, proved e.g. in [123].

Lemma 4.1.1 (Stein's Lemma)**.** *A random variable Y is such that $Y \stackrel{\text{Law}}{=} Z \sim \mathscr{N}(0,1)$ if, and only if, for every continuous and piecewise continuously differentiable function such that $E\left|f'(Z)\right| < \infty$, one has*

$$E[f'(Y) - Yf(Y)] = 0. \quad (4.1.1)$$

The fact that a random variable Y satisfying (4.1.1) is necessarily Gaussian can be proved by several routes: for instance, by taking f to be a complex exponential, one can show that the characteristic function of Y, say $\psi(t)$, is necessarily a solution to the differential equation $\psi'(t) + t\psi(t) = 0$, and therefore $\psi(t) = \exp(-t^2/2)$; alternatively, one can set $f(x) = x^n$, $n = 1, 2, ...$, and observe that (4.1.1) implies that, for every n, one must have $E(Y^n) = E(Z^n)$, where $Z \sim \mathscr{N}(0,1)$ (note that the law of Z is determined by its moments).

Heuristically, Lemma 4.1.1 suggests that the distance $d(Y, Z)$, between the law of a random variable Y and that of $Z \sim \mathscr{N}(0,1)$, must be "small" whenever

$$E[f'(Y) - Yf(Y)] \simeq 0,$$

for a sufficiently large class of functions f. In the seminal works [123, 125], Stein proved that this somewhat imprecise argument can be made rigorous by means of the use of differential equations. To see this, for a given function $g : \mathbb{R} \to \mathbb{R}$, define the *Stein equation* associated with g as

$$g(x) - E[g(Z)] = h'(x) - xh(x), \quad \forall x \in \mathbb{R}, \quad (4.1.2)$$

(we recall that $Z \sim \mathscr{N}(0,1)$). A solution to (4.1.2) is a function h which is Lebesgue-almost everywhere differentiable, and such that there exists a version of h' satisfying (4.1.2) for every $x \in \mathbb{R}$. If one assumes that $g \in \text{Lip}(1)$ (that is, if $\|g\|_{Lip} \leq 1$, where $\|\cdot\|_{Lip}$ stands for the usual Lipschitz norm), then a standard result (see e.g. [125]) yields that (4.1.2) admits a solution h such that $\|h'\|_\infty \leq 1$ and $\|h''\|_\infty \leq 2$. Now recall that the *Wasserstein distance* between the laws of two real-valued random variables Y and X is defined as

$$d_{\mathrm{W}}(Y, X) = \sup_{g \in \text{Lip}(1)} |E[g(Y)] - E[g(X)]|,$$

and introduce the notation $\mathscr{F}_{\mathrm{W}} = \{f : \|f'\|_\infty \leq 1, \|f''\|_\infty \leq 2\}$. By taking expectations on the two sides of (4.1.2), one obtains finally that, for $Z \sim \mathscr{N}(0,1)$ and for a generic random variable Y, we have:

$$d_{\mathrm{W}}(Y, Z) \leq \sup_{f \in \mathscr{F}_{\mathrm{W}}} |E[f'(Y) - Yf(Y)]|, \quad (4.1.3)$$

thus giving a precise meaning to the heuristic argument sketched above (note that an analogous conclusion can be obtained for other distances, such as the total variation distance or the Kolmogorov distance – see e.g. [26] for a discussion of this point). We stress that the topology induced by d_W, on probability measures on $\mathbb{R}$, is stronger than the topology induced by weak convergence.

The starting point of [86, 87] is that a relation such as (4.1.3) can be very effectively combined with Malliavin calculus, whenever Y is a centered regular functional of some infinite dimensional Gaussian field. To see this, denote by DY the Malliavin derivative of Y (observe that DY is a random element with values in some adequate Hilbert space $\mathfrak{H}$), and write L to indicate the (infinite-dimensional) Ornstein-Uhlenbeck generator (see Section 4.2 below for precise definitions). One crucial relation proved in [86], and then further exploited in [87], is the upper bound

$$d_W(Y, Z) \leq E[(1 - \langle DY, -DL^{-1}Y\rangle_{\mathfrak{H}})^2]^{1/2}. \tag{4.1.4}$$

As shown in [86], when specialized to the case of Y being equal to a multiple Wiener-Itô integral, the relation (4.1.4) yields bounds that are intimately related with the CLTs proved in [95] and [96]. See [87] for a characterization of the optimality of these bounds; see again [86] for extensions to non-Gaussian approximations and for applications to the Breuer-Major CLT (stated and proved in [20]) for functionals of a fractional Brownian motion.

The principal contribution of the present paper (see e.g. the statement of Theorem 4.3.5 below), consists in showing that a relation similar to (4.1.4) continues to hold when Z is replaced by a d-dimensional Gaussian vector ($d \geq 2$), $F = (F_1, ..., F_d)$ is a vector of smooth functionals of a Gaussian field, and d_W is the Wasserstein distance between probability laws on $\mathbb{R}^d$ (see Definition 4.3.1 below). Our results apply to Gaussian approximations by means of Gaussian vectors with arbitrary positive definite covariance matrices. The proofs rely on a multidimensional version of the Stein equation (4.1.2), that we combine with standard integration by parts formulae on an infinite-dimensional Gaussian space. Our approach bears some connections with the paper by Hsu [58], where the author proves an hybrid Stein/semimartingale characterization of Brownian motions on manifolds, via Malliavin-type operators.

The paper is organized as follows. In Section 4.2 we provide some preliminaries on Malliavin calculus. Section 4.3 contains our main results, concerning Gaussian approximations by means of vectors of Gaussian random variables with positive definite covariance matrices. Finally, Section 4.4 deals with two applications: (i) to a functional version of the Breuer-Major CLT (see [20]), and (ii) to Gaussian approximations of functionals of finite Normal vectors, providing a generalization of a technical result proved by Chatterjee in [24].

4.2 Preliminaries and notations

In this section, we recall some basic elements of Malliavin calculus for Gaussian processes. The reader is referred to [94] for a complete discussion of this subject. Let $X = \{X(h),\ h \in \mathfrak{H}\}$ be an *isonormal Gaussian process* on a probability space $(\Omega, \mathcal{F}, P)$. This means that X is a centered Gaussian family indexed by the elements of an Hilbert space $\mathfrak{H}$, such that, for every pair $h, g \in \mathfrak{H}$ one has that

$$E[X(h)X(g)] = \langle h, g \rangle_{\mathfrak{H}}.$$

We let $L^2(X)$ be shorthand for the space $L^2(\Omega, \sigma(X), P)$. It is well known that every random variable $F \in L^2(X)$ admits the chaotic expansion

$$F = E(F) + \sum_{n=1}^{\infty} I_n(f_n)$$

where the deterministic kernels f_n, $n \geq 1$, belong to $\mathfrak{H}^{\odot n}$ and the convergence of the series holds in $L^2(X)$. One sometimes uses the notation $I_0(f_0) = E[F]$. In the particular case where $\mathfrak{H} := L^2(T, \mathcal{A}, \mu)$, with $(T, \mathcal{A})$ a measurable space, the random variable $I_n(f_n)$ coincides with the *multiple Wiener-Itô integral* (of order n) of f_n with respect to X (see [94, Section 1.1.2.]).

Let $f \in \mathfrak{H}^{\odot p}$, $g \in \mathfrak{H}^{\odot q}$ and $0 \leq r \leq p \wedge q$. We define the *contraction* $f \otimes_r g$ of order r of f and g as the element of $\mathfrak{H}^{\otimes(p+q-2r)}$ given by,

$$f \otimes_r g := \sum_{i_1,\ldots,i_r=1}^{\infty} \langle f, e_{i_1} \otimes \ldots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}} \otimes \langle g, e_{i_1} \otimes \ldots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}},$$

where $\{e_k,\ k \geq 1\}$ is a complete orthonormal system in $\mathfrak{H}$. Note that $f \otimes_0 g = f \otimes g$; also, if $p = q$, then $f \otimes_p g = \langle f, g \rangle_{\mathfrak{H}^{\otimes p}}$. Note that, in general, $f \otimes_r g$ is not a symmetric element of $\mathfrak{H}^{\otimes(p+q-2r)}$; the canonical symmetrization of $f \otimes_r g$ is denoted by $f \widetilde{\otimes}_r g$. We recall the product formula for multiple stochastic integrals:

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

Now let $\mathscr{S}$ be the set of cylindrical functionals F of the form

$$F = \varphi(X(h_1), \ldots, X(h_n)), \tag{4.2.1}$$

where $n \geq 1$, $h_i \in \mathfrak{H}$ and the function $\varphi \in \mathscr{C}^{\infty}(\mathbb{R}^n)$ is such that its partial derivatives have polynomial growth. The *Malliavin derivative* DF of a functional F of the form (4.2.1) is the square integrable $\mathfrak{H}$-valued random variable defined as

$$DF = \sum_{i=1}^{n} \partial_i \varphi(X(h_1), \ldots, X(h_n)) h_i, \tag{4.2.2}$$

where $\partial_i\varphi$ denotes the ith partial derivative of φ. In particular, one has that $DX(h) = h$ for every h in $\mathfrak{H}$. By iteration, one can define the mth derivative $D^m F$ of $F \in \mathscr{S}$, which is an element of $L^2(\Omega; \mathfrak{H}^{\otimes m})$, for $m \geq 2$. As usual $\mathbb{D}^{m,2}$ denotes the closure of $\mathscr{S}$ with respect to the norm $\|\cdot\|_{m,2}$ defined by the relation

$$\|F\|_{m,2}^2 = E[F^2] + \sum_{i=1}^{m} E[\|D^i F\|_{\mathfrak{H}^{\otimes i}}^2].$$

Note that every finite sum of Wiener-Itô integrals always belongs to $\mathbb{D}^{m,2}$ ($\forall m \geq 1$). The Malliavin derivative D satisfies the following *chain rule formula*: if $\varphi : \mathbb{R}^n \to \mathbb{R}$ is in $\mathscr{C}_b^1$ (defined as the set of continuously differentiable functions with bounded derivatives) and if $(F_1, \ldots, F_n)$ is a random vector such that each component belongs to $\mathbb{D}^{1,2}$, then $\varphi(F_1, \ldots, F_n)$ is itself an element of $\mathbb{D}^{1,2}$, and moreover

$$D\varphi(F_1, \ldots, F_n) = \sum_{i=1}^{n} \partial_i \varphi(F_1, \ldots, F_n) DF_i. \tag{4.2.3}$$

The *divergence operator* δ is defined as the dual operator of D. A random element u of $L^2(\Omega; \mathfrak{H})$ belongs to the domain of δ (denoted by Domδ) if there exists a constant c_u satisfying

$$|E[\langle DF, u\rangle_{\mathfrak{H}}]| \leq c_u \|F\|_{L^2(\Omega)}, \qquad \text{for every } F \in \mathscr{S};$$

in this case, the divergence of u, written $\delta(u)$, is defined by the following duality property:

$$E[F\delta(u)] = E[\langle DF, u\rangle_{\mathfrak{H}}], \quad \forall F \in \mathbb{D}^{1,2}. \tag{4.2.4}$$

The crucial relation (4.2.4) is custumarily called the (Malliavin) *integration by parts formula*.

In what follows, we shall denote by $T = \{T_t : t \geq 0\}$ the *Ornstein-Uhlenbeck semigroup*. We recall that, for every $t \geq 0$ and every $F \in L^2(X)$,

$$T_t(F) = \sum_{n=0}^{\infty} e^{-nt} J_n(F), \tag{4.2.5}$$

where, for every $n \geq 0$ and for the rest of the paper, the symbol J_n denotes the projection operator onto the nth Wiener chaos. Note that T is indeed the semigroup associated with an infinite-dimensional stationary Gaussian process with values in $\mathbb{R}^{\mathfrak{H}}$, having the law of X as an invariant distribution (see e.g. [94, Section 1.4] for a more detailed discussion of the Ornstein-Uhlenbeck semigroup in the context of Malliavin calculus; see Barbour [8] for a version of Stein method involving Ornstein-Uhlenbeck semigroups on infinite-dimensional spaces; see Götze [48] for a version of Stein's method based on multi-dimensional Ornstein-Uhlenbeck semigroups). The *infinitesimal generator of the Ornstein-Uhlenbeck semigroup* is noted L. A square integrable random variable F is in the domain of L (noted Dom$L$) if F belongs to the domain of δD (that is, if F is in $\mathbb{D}^{1,2}$ and $DF \in$ Domδ) and, in this case,

$$LF = -\delta DF.$$

One can prove that LF is such that

$$LF = -\sum_{n=0}^{\infty} nJ_n(F).$$

As an example, if $F = I_q(f_q)$, with $f_q \in \mathfrak{H}^{\odot q}$, then $LF = -qF$. Note that, for every $F \in \mathrm{Dom}L$, one has $E[F] = 0$. The inverse L^{-1} of the operator L acts on mean-zero random variables F as

$$L^{-1}F = -\sum_{n=1}^{\infty} \frac{1}{n} J_n(F).$$

In particular, for every $q \geq 1$ and every $F = I_q(f_q)$ with $f_q \in \mathfrak{H}^{\odot q}$, one has that $L^{-1}F = -\frac{1}{q}F$.

We conclude this section by recording two important characterizations of the Ornstein-Uhlenbeck semigroup and its generator.

i) Mehler's formula. Let F be an element of $\mathrm{L}^2(X)$, so that F can be represented as an application from $\mathbb{R}^{\mathfrak{H}}$ into $\mathbb{R}$. Then, an alternative representation (due to Mehler) of the action of the Ornstein-Uhlenbeck semigroup T (as defined in (4.2.5)) on F, is the following:

$$T_t(F) = E[F(e^{-t}a + \sqrt{1-e^{-2t}}X)]\mid_{a=X}, \quad t \geq 0, \tag{4.2.6}$$

where a designs a generic element of $\mathbb{R}^{\mathfrak{H}}$. See Nualart [94, Section 1.4.1] for more details on this and other characterizations of T.

ii) Differential characterization of L. Let $F \in \mathrm{L}^2(X)$ have the form

$$F = f(X(h_1), ..., X(h_d)),$$

where $f \in \mathscr{C}^2(\mathbb{R}^d)$, and $h_i \in \mathfrak{H}$, $i = 1, ..., d$. Then,

$$LF = \sum_{i,j=1}^{d} \frac{\partial^2 f}{\partial x_i \partial x_j}(X(h_1), ..., X(h_d))\langle h_i, h_j\rangle_{\mathfrak{H}} - \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(X(h_1), ..., X(h_d))X(h_i). \tag{4.2.7}$$

See Propositions 1.4.4 and 1.4.5 in [94] for a proof and some generalizations of (4.2.7).

4.3 Stein's method and Gaussian vectors

We start by giving a definition of the Wasserstein distance, as well as by introducing some useful norms over classes of real-valued matrices.

Definition 4.3.1. (i) The *Wasserstein distance* between the laws of two $\mathbb{R}^d$-valued random vectors X and Y, noted $d_{\mathrm{W}}(X, Y)$, is given by

$$d_{\mathrm{W}}(X, Y) := \sup_{g \in \mathscr{H};\|g\|_{Lip} \leq 1} \big|E[g(X)] - E[g(Y)]\big|,$$

where $\mathscr{H}$ indicates the class of Lipschitz functions, that is, the collection of all functions $g : \mathbb{R}^d \to \mathbb{R}$ such that $\|g\|_{Lip} := \sup_{x \neq y} \frac{|g(x) - g(y)|}{\|x - y\|_{\mathbb{R}^d}} < \infty$ (with $\| \cdot \|_{\mathbb{R}^d}$ the usual Euclidian norm on $\mathbb{R}^d$).

(ii) The *Hilbert-Schmidt inner product* and the *Hilbert-Schmidt norm* on the class of $d \times d$ real matrices, denoted respectively by $\langle \cdot, \cdot \rangle_{H.S.}$ and $\| \cdot \|_{H.S.}$, are defined as follows: for every pair of matrices A and B,

$$\langle A, B \rangle_{H.S.} := \mathrm{Tr}(AB^T), \text{ and } \|A\|_{H.S.} := \sqrt{\langle A, A \rangle_{H.S.}}.$$

(iii) The *operator norm* of a $d \times d$ matrix A over $\mathbb{R}$ is given by

$$\|A\|_{op} := \sup_{x \in \mathbb{R}^d; \|x\|_{\mathbb{R}^d} = 1} \|Ax\|_{\mathbb{R}^d}.$$

Remark 4.3.2. 1. For every $d \geq 1$ the topology induced by d_{W}, on the class of all probability measures on $\mathbb{R}^d$, is strictly stronger than the topology induced by weak convergence (see e.g. Dudley [36, Chapter 11]).

2. According to the notation introduced in Definition 4.3.1(ii), relation (4.2.7) can be rewritten as

$$LF = \langle C, \mathrm{Hess} f(Y) \rangle_{H.S.} - \langle Y, \nabla f(Y) \rangle_{\mathbb{R}^d}, \tag{4.3.1}$$

where $Y = (X(h_1), ..., X(h_d))$, and $C = \{C(i,j) : i,j = 1, ..., d\}$ is the $d \times d$ covariance matrix such that $C(i,j) = E(X(h_i)X(h_j)) = \langle h_i, h_j \rangle_{\mathfrak{H}}$.

Given a $d \times d$ positive definite symmetric matrix C, we use the notation $\mathscr{N}_d(0, C)$ to indicate the law of a d-dimensional Gaussian vector with zero mean and covariance C. The following result, which is basically known (see e.g. [25] or [112]), is the d-dimensional counterpart of Stein's Lemma 4.1.1. In what follows, we provide a new proof which is almost exclusively based on the use of Malliavin operators.

Lemma 4.3.3. *Fix an integer $d \geq 2$ and let $C = \{C(i,j) : i,j = 1, ..., d\}$ be a $d \times d$ positive definite symmetric real matrix.*

(i) Let Y be a random variable with values in $\mathbb{R}^d$. Then $Y \sim \mathscr{N}_d(0, C)$ if, and only if, for every twice differentiable function $f : \mathbb{R}^d \to \mathbb{R}$ such that $E|\langle C, \mathrm{Hess} f(Y) \rangle_{H.S.} - \langle Y, \nabla f(Y) \rangle_{\mathbb{R}^d}| < \infty$, it holds that

$$E[\langle C, \mathrm{Hess} f(Y) \rangle_{H.S.} - \langle Y, \nabla f(Y) \rangle_{\mathbb{R}^d}] = 0. \tag{4.3.2}$$

(ii) Consider a Gaussian random vector $Z \sim \mathscr{N}_d(0, C)$. Let $g : \mathbb{R}^d \to \mathbb{R}$ belong to $\mathscr{C}^2(\mathbb{R}^d)$ and be such that g and its partial derivatives have subexponential growth at infinity. Then, the function $U_0(g)$ defined by

$$U_0 g(x) := \int_0^1 \frac{1}{2t} E[g(\sqrt{t}x + \sqrt{1-t}Z) - g(Z)]dt$$

is a solution to the following differential equation (with unknown function f):

$$g(x) - E[g(Z)] = \langle C, \text{Hess} f(x)\rangle_{H.S.} - \langle x, \nabla f(x)\rangle_{\mathbb{R}^d}, \quad x \in \mathbb{R}^d. \tag{4.3.3}$$

Moreover, one has that

$$\sup_{x\in\mathbb{R}^d} \|\text{Hess}\, U_0 g(x)\|_{H.S.} \leq \|C^{-1}\|_{op}\, \|C\|_{op}^{1/2}\, \|g\|_{Lip}. \tag{4.3.4}$$

Remark 4.3.4. 1. If $C = \sigma^2 \mathbf{I}_d$ for some $\sigma > 0$ (that is, if Z is composed of i.i.d. centered Gaussian random variables with common variance equal to σ^2), then

$$\|C^{-1}\|_{op}\, \|C\|_{op}^{1/2} = \|\sigma^{-2}\mathbf{I}_d\|_{op}\, \|\sigma^2\mathbf{I}_d\|_{op}^{1/2} = \sigma^{-1}.$$

2. Unlike formulae (4.1.1) and (4.1.2) (associated with one-dimensional Gaussian approximations) the relation (4.3.2) and the Stein equation (4.3.3) involve second-order differential operators. A discussion of this fact is detailed e.g. in [25, Theorem 4].

Proof of Lemma 4.3.3. We start by proving Point (ii). First observe that, without loss of generality, we can suppose that $Z = (Z_1, ..., Z_d) = (X(h_1), ...X(h_d))$, where X is an isonormal Gaussian process over some Hilbert space $\mathfrak{H}$, the kernels h_i belong to $\mathfrak{H}$ ($i = 1, ..., d$), and $\langle h_i, h_j\rangle_{\mathfrak{H}} = E(X(h_i)X(h_j)) = E(Z_i Z_j) = C(i,j)$. By using the change of variable $2u = -\log t$, one can rewrite $U_0 g(x)$ as follows

$$U_0 g(x) = \int_0^\infty \{E[g(e^{-u}x + \sqrt{1-e^{-2u}}Z)] - E[g(Z)]\} du.$$

Now define $\widetilde{g}(Z) = g(Z) - E[g(Z)]$, and observe that $\widetilde{g}(Z)$ is by assumption a centered element of $\mathrm{L}^2(X)$. For $q \geq 0$, denote by $J_q(\widetilde{g}(Z))$ the projection of $\widetilde{g}(Z)$ on the qth Wiener chaos, so that $J_0(\widetilde{g}(Z)) = 0$. According to Mehler's formula (4.2.6),

$$E[g(e^{-u}x + \sqrt{1-e^{-2u}}Z)]|_{x=Z} - E[g(Z)] = E[\widetilde{g}(e^{-u}x + \sqrt{1-e^{-2u}}Z)]|_{x=Z} = T_u\widetilde{g}(Z),$$

where x denotes a generic element of $\mathbb{R}^d$. In view of (4.2.5), it follows that

$$U_0 g(Z) = \int_0^\infty T_u \widetilde{g}(Z) du = \int_0^\infty \sum_{q\geq 1} e^{-qu} J_q(\widetilde{g}(Z)) du = \sum_{q\geq 1} \frac{1}{q} J_q(\widetilde{g}(Z)) = L^{-1}\widetilde{g}(Z).$$

It is easily seen that the function $U_0 g$ is an element of $\mathscr{C}^2(\mathbb{R}^d)$. By exploiting the differential representation (4.3.1) in the case $Y = Z$, one deduces that

$$\langle C, \text{Hess} U_0 g(Z)\rangle_{H.S.} - \langle Z, \nabla U_0 g(Z)\rangle_{\mathbb{R}^d} = L U_0 g(Z) = L L^{-1}\widetilde{g}(Z) = g(Z) - E[g(Z)].$$

Since the matrix C is positive definite, we infer that the support of the law of Z coincides with $\mathbb{R}^d$, and therefore (e.g. by a continuity argument) we obtain that

$$\langle C, \text{Hess} U_0 g(x)\rangle_{H.S.} - \langle x, \nabla U_0 g(x)\rangle_{\mathbb{R}^d} = g(x) - E[g(Z)],$$

for every $x \in \mathbb{R}^d$. This yields that the function $U_0 g$ solves the Stein's equation (4.3.3). To prove the estimate (4.3.4), we first recall that there exists a unique non-singular *symmetric* matrix A such that $A^2 = C$, and that one has that $A^{-1}Z \sim \mathscr{N}_d(0, \mathbf{I}_d)$. Now write $U_0 g(x) = h(A^{-1}x)$, where

$$h(x) = \int_0^1 \frac{1}{2t} E[g_A(\sqrt{t}x + \sqrt{1-t}A^{-1}Z) - g_A(A^{-1}Z)]dt,$$

and $g_A(x) = g(Ax)$. Note that, since $A^{-1}Z \sim \mathscr{N}_d(0, \mathbf{I}_d)$, the function h solves the Stein's equation

$$\Delta h(x) - \langle x, \nabla h(x) \rangle_{\mathbb{R}^d} = g_A(x) - E[g_A(Y)],$$

where $Y \sim \mathscr{N}_d(0, \mathbf{I}_d)$. We can now use the same arguments as in the proof of Lemma 3 in [25] (by inspection of this last reference, one sees that here is the point where we use the subexponential growth assumptions) to deduce that

$$\sup_{x \in \mathbb{R}^d} \|\text{Hess}\, h(x)\|_{H.S.} \leq \|g_A\|_{Lip} \leq \|A\|_{op} \|g\|_{Lip}. \tag{4.3.5}$$

On the other hand, by noting $h_{A^{-1}}(x) = h(A^{-1}x)$, one obtains by standard computations (recall that A is symmetric) that

$$\text{Hess}\, U_0 g(x) = \text{Hess}\, h_{A^{-1}}(x) = A^{-1} \text{Hess}\, h(A^{-1}x) A^{-1},$$

yielding

$$\begin{aligned}
\sup_{x \in \mathbb{R}^d} \|\text{Hess}\, U_0 g(x)\|_{H.S.} &= \sup_{x \in \mathbb{R}^d} \|A^{-1} \text{Hess}\, h(A^{-1}x) A^{-1}\|_{H.S.} \\
&= \sup_{x \in \mathbb{R}^d} \|A^{-1} \text{Hess}\, h(x) A^{-1}\|_{H.S.} \\
&\leq \|A^{-1}\|_{op}^2 \sup_{x \in \mathbb{R}^d} \|\text{Hess}\, h(x)\|_{H.S.} && (4.3.6) \\
&\leq \|A^{-1}\|_{op}^2 \, \|A\|_{op} \, \|g\|_{Lip} && (4.3.7) \\
&\leq \|C^{-1}\|_{op} \, \|C\|_{op}^{1/2} \, \|g\|_{Lip}. && (4.3.8)
\end{aligned}$$

The chain of inequalities appearing in formulae (4.3.6)–(4.3.8) are mainly a consequence of the usual properties of the Hilbert-Schmidt and operator norms. Indeed, to prove inequality (4.3.6) we used the relations

$$\begin{aligned}
\|A^{-1} \text{Hess}\, h(x) A^{-1}\|_{H.S.} &\leq \|A^{-1}\|_{op} \, \|\text{Hess}\, h(x) A^{-1}\|_{H.S.} \\
&\leq \|A^{-1}\|_{op} \, \|\text{Hess}\, h(x)\|_{H.S.} \, \|A^{-1}\|_{op} \; ;
\end{aligned}$$

relation (4.3.7) is a consequence of (4.3.5); finally, to show the inequality (4.3.8), one uses the fact that

$$\|A^{-1}\|_{op} \leq \sqrt{\|A^{-1}A^{-1}\|_{op}} = \sqrt{\|C^{-1}\|_{op}}$$

$$\|A\|_{op} \;\leq\; \sqrt{\|AA\|_{op}} = \sqrt{\|C\|_{op}}\,.$$

We are now left with the proof of Point (i) in the statement. The fact that a vector $Y \sim \mathscr{N}_d(0, C)$ necessarily verifies (4.3.2) can be proved by standard integration by parts. On the other hand, suppose that Y verifies (4.3.2). Then, according to Point (ii), for every $g \in \mathscr{C}^2(\mathbb{R}^d)$ such that g is globally Lipschitz and g and its partial derivatives have subexponential growth,

$$E(g(Y)) - E(g(Z)) = E(\langle C, \mathrm{Hess} f(Y)\rangle_{H.S.} - \langle Y, \nabla f(Y)\rangle_{\mathbb{R}^d}) = 0,$$

where $Z \sim \mathscr{N}_d(0, C)$. Since the collection of all such functions g generates the Borel σ-field on $\mathbb{R}^d$, this implies that $Y \stackrel{\text{Law}}{=} Z$, thus yielding the desired conclusion. □

The following statement is the main result of this paper. Its proof makes a crucial use of the integration by parts formula (4.2.4) discussed in Section 4.2.

Theorem 4.3.5. *Fix $d \geq 2$ and let $C = \{C(i,j) : i,j = 1,...,d\}$ be a $d \times d$ positive definite matrix. Suppose that $Z \sim \mathscr{N}_d(0, C)$ and that $F := (F_1, \ldots, F_d)$ is a $\mathbb{R}^d$-valued random vector such that $E[F_i] = 0$ and $F_i \in \mathbb{D}^{1,2}$ for every $i = 1, \ldots, d$. Then,*

$$d_{\mathrm{W}}(F, Z) \;\leq\; \|C^{-1}\|_{op}\, \|C\|_{op}^{1/2} \sqrt{E\|C - \Phi(DF)\|_{H.S}^2} \tag{4.3.9}$$

$$= \|C^{-1}\|_{op}\, \|C\|_{op}^{1/2} \sqrt{\sum_{i,j=1}^{d} E[(C(i,j) - \langle DF_i, -L^{-1}DF_j\rangle_{\mathfrak{H}})^2]}, \tag{4.3.10}$$

where we write $\Phi(DF)$ to indicate the matrix

$$\Phi(DF) = \{\langle DF_i, -L^{-1}DF_j\rangle_{\mathfrak{H}} : 1 \leq i,j \leq d\}.$$

Proof. We start by proving that, for every $g \in \mathscr{C}^2(\mathbb{R}^d)$ such that g and its partial derivatives have subexponential growth,

$$|E[g(F)] - E[g(Z)]| \leq \|C^{-1}\|_{op}\, \|C\|_{op}^{1/2}\, \|g\|_{Lip} \sqrt{E\|C - \Phi(DF)\|_{H.S}^2}.$$

To prove such a claim, observe that, according to Point (ii) in Lemma 4.3.3, $E[g(F)] - E[g(Z)] = E[\langle C, \mathrm{Hess}U_0 g(F)\rangle_{H.S.} - \langle F, \nabla U_0 g(F)\rangle_{\mathbb{R}^d}]$. Moreover,

$$\begin{aligned}
&\left|E[\langle C, \mathrm{Hess}U_0 g(F)\rangle_{H.S.} - \langle F, \nabla U_0 g(F)\rangle_{\mathbb{R}^d}]\right| \\
&= \left|E\left[\sum_{i,j=1}^{d} C(i,j)\partial_{ij}^2 U_0 g(F) - \sum_{i=1}^{d} F_i \partial_i U_0 g(F)\right]\right| \\
&= \left|\sum_{i,j=1}^{d} E\left[C(i,j)\partial_{ij}^2 U_0(g)(F)\right] - \sum_{i=1}^{d} E\left[\left(LL^{-1}F_i\right)\partial_i U_0(g)F\right]\right| \quad (\text{since } E(F_i) = 0)
\end{aligned}$$

$$
\begin{aligned}
&= \left| \sum_{i,j=1}^{d} E\left[C(i,j)\partial^2_{ij}U_0 g(F)\right] + \sum_{i=1}^{d} E\left[\delta(DL^{-1}F)_i \partial_i U_0 g(F)\right] \right| \quad \text{(since } \delta D = -L\text{)} \\
&= \left| \sum_{i,j=1}^{d} E\left[C(i,j)\partial^2_{ij}U_0 g(F)\right] - \sum_{i=1}^{d} E\left[\langle D(\partial_i U_0 g(F)), -DL^{-1}F_i\rangle_{\mathfrak{H}}\right] \right| \quad \text{(by (4.2.4))} \\
&= \left| \sum_{i,j=1}^{d} E\left[C(i,j)\partial^2_{ij}U_0 g(F)\right] - \sum_{i,j=1}^{d} E\left[\partial^2_{ji}U_0 g(F)\langle DF_j, -DL^{-1}F_i\rangle_{\mathfrak{H}}\right] \right| \quad \text{(by (4.2.3))} \\
&= \left| \sum_{i,j=1}^{d} E\left[\partial^2_{ij}U_0 g(F)\big(C(i,j) - \langle DF_i, -DL^{-1}F_j\rangle_{\mathfrak{H}}\big)\right] \right| \\
&= \left| E\langle \operatorname{Hess} U_0 g(F), C - \Phi(DF)\rangle_{H.S.} \right| \\
&\leq \sqrt{E\|\operatorname{Hess} U_0 g(F)\|^2_{H.S}} \sqrt{E\|C - \Phi(DF)\|^2_{H.S}} \quad \text{(by the Cauchy-Schwarz inequality)} \\
&\leq \|C^{-1}\|_{op}\, \|C\|^{1/2}_{op}\, \|g\|_{Lip} \sqrt{E\|C - \Phi(DF)\|^2_{H.S}} \quad \text{(by (4.3.4)).}
\end{aligned}
$$

To prove the Wasserstein estimate (4.3.9), it is sufficient to observe that, for every globally Lipschitz function g such that $\|g\|_{Lip} \leq 1$, there exists a sequence $\{g_k : k \geq 1\}$, such that:

(i) for each k, the function g_k and its partial derivatives have subexponential growth;

(ii) for each k, one has that $\|g_k\|_{Lip} \leq \|g\|_{Lip}$;

(iii) as $k \to \infty$, $\|g_k - g\|_\infty \downarrow 0$.

□

Observe that Theorem 4.3.5 generalizes relation (4.1.4) (that was proved in [86, Theorem 3.1]). We now aim at applying Theorem 4.3.5 to vectors of multiple stochastic integrals.

Corollary 4.3.6. *Fix $d \geq 2$ and $1 \leq q_1 \leq \ldots \leq q_d$. Consider a vector $F := (F_1, \ldots, F_d) = (I_{q_1}(f_1), \ldots, I_{q_d}(f_d))$ with $f_i \in \mathfrak{H}^{\odot q_i}$ for any $i = 1 \ldots, d$. Let $Z \sim \mathscr{N}_d(0, C)$, with C positive definite. Then,*

$$
d_{\mathrm{W}}(F, Z) \leq \|C^{-1}\|_{op}\, \|C\|^{1/2}_{op} \sqrt{\sum_{1\leq i,j\leq d} E\left[\left(C(i,j) - \frac{1}{q_j}\langle DF_i, DF_j\rangle_{\mathfrak{H}}\right)^2\right]}. \tag{4.3.11}
$$

Proof. We have $-L^{-1}F_j = \frac{1}{q_j} F_j$ so that the desired conclusion follows from (4.3.10). □

When one applies Corollary 4.3.6 in concrete situations (see e.g. Section 4.4 below), one can use the following result in order to evaluate the RHS of (4.3.11).

Lemma 4.3.7. *Let $F = I_p(f)$ and $G = I_q(g)$, with $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$ $(p, q \geq 1)$. Let a be a real constant. If $p = q$, one has the estimate:*

$$E\left[\left(a - \frac{1}{p}\langle DF, DG\rangle_{\mathfrak{H}}\right)^2\right] \leq (a - p!\langle f, g\rangle_{\mathfrak{H}^{\otimes p}})^2$$
$$+\frac{p^2}{2}\sum_{r=1}^{p-1}(r-1)!^2\binom{p-1}{r-1}^4(2p-2r)!\big(\|f\otimes_{p-r}f\|^2_{\mathfrak{H}^{\otimes 2r}} + \|g\otimes_{p-r}g\|^2_{\mathfrak{H}^{\otimes 2r}}\big).$$

On the other hand, if $p < q$, one has that

$$E\left[\left(a - \frac{1}{q}\langle DF, DG\rangle_{\mathfrak{H}}\right)^2\right] \leq a^2 + p!^2\binom{q-1}{p-1}^2(q-p)!\|f\|^2_{\mathfrak{H}^{\otimes p}}\|g\otimes_{q-p}g\|_{\mathfrak{H}^{\otimes 2p}}$$
$$+\frac{p^2}{2}\sum_{r=1}^{p-1}(r-1)!^2\binom{p-1}{r-1}^2\binom{q-1}{r-1}^2(p+q-2r)!\big(\|f\otimes_{p-r}f\|^2_{\mathfrak{H}^{\otimes 2r}} + \|g\otimes_{q-r}g\|^2_{\mathfrak{H}^{\otimes 2r}}\big).$$

Remark 4.3.8.

1. Recall that $E\big(I_p(f)I_q(g)\big) = \begin{cases} p!\langle f, g\rangle_{\mathfrak{H}^{\otimes p}} & \text{if } p = q, \\ 0 & \text{otherwise.}\end{cases}$
2. In order to estimate the right-hand side of (4.3.11), we see that it is sufficient to asses the quantity $\|f_i \otimes_r f_i\|_{\mathfrak{H}^{\otimes 2(q_i - r)}}$ for any $i \in \{1, \dots, d\}$ and $r \in \{1, \dots, q_i - 1\}$ on the one hand, and $\langle f_i, f_j\rangle_{\mathfrak{H}^{\otimes q_i}}$ for any $1 \leq i, j \leq d$ such that $q_i = q_j$ on the other hand.

Proof of Lemma 4.3.7 (see also [95, Lemma 2]). Without loss of generality, we can assume that $\mathfrak{H} = L^2(A, \mathscr{A}, \mu)$, where $(A, \mathscr{A})$ is a measurable space, and μ is a σ-finite and non-atomic measure. Thus, we can write

$$\begin{aligned}
\langle DF, DG\rangle_{\mathfrak{H}} &= p\,q\,\langle I_{p-1}(f), I_{q-1}(g)\rangle_{\mathfrak{H}} = p\,q\int_A I_{p-1}\big(f(\cdot,t)\big)I_{q-1}\big(g(\cdot,t)\big)\mu(dt) \\
&= p\,q\int_A\sum_{r=0}^{p\wedge q-1} r!\binom{p-1}{r}\binom{q-1}{r}I_{p+q-2-2r}\big(f(\cdot,t)\widetilde{\otimes}_r g(\cdot,t)\big)\mu(dt) \\
&= p\,q\sum_{r=0}^{p\wedge q-1} r!\binom{p-1}{r}\binom{q-1}{r}I_{p+q-2-2r}(f\widetilde{\otimes}_{r+1}g) \\
&= p\,q\sum_{r=1}^{p\wedge q}(r-1)!\binom{p-1}{r-1}\binom{q-1}{r-1}I_{p+q-2r}(f\widetilde{\otimes}_r g).
\end{aligned}$$

It follows that

$$E\left[\left(a - \frac{1}{q}\langle DF, DG\rangle_{\mathfrak{H}}\right)^2\right] \tag{4.3.12}$$

$$
= \begin{cases} a^2 + p^2 \sum_{r=1}^{p} (r-1)!^2 \binom{p-1}{r-1}^2 \binom{q-1}{r-1}^2 (p+q-2r)! \|f \widetilde{\otimes}_r g\|^2_{\mathfrak{H}^{\otimes(p+q-2r)}} \text{ if } p < q, \\ \\ (a - p!\langle f, g\rangle_{\mathfrak{H}^{\otimes p}})^2 + p^2 \sum_{r=1}^{p-1} (r-1)!^2 \binom{p-1}{r-1}^4 (2p-2r)! \|f \widetilde{\otimes}_r g\|^2_{\mathfrak{H}^{\otimes(2p-2r)}} \text{ if } p = q. \end{cases}
$$

If $r < p \leq q$ then

$$
\begin{aligned}
\|f \widetilde{\otimes}_r g\|^2_{\mathfrak{H}^{\otimes(p+q-2r)}} &\leq \|f \otimes_r g\|^2_{\mathfrak{H}^{\otimes(p+q-2r)}} = \langle f \otimes_{p-r} f, g \otimes_{q-r} g\rangle_{\mathfrak{H}^{\otimes 2r}} \\
&\leq \|f \otimes_{p-r} f\|_{\mathfrak{H}^{\otimes 2r}} \|g \otimes_{q-r} g\|_{\mathfrak{H}^{\otimes 2r}} \\
&\leq \frac{1}{2} \left(\|f \otimes_{p-r} f\|^2_{\mathfrak{H}^{\otimes 2r}} + \|g \otimes_{q-r} g\|^2_{\mathfrak{H}^{\otimes 2r}} \right).
\end{aligned}
$$

If $r = p < q$, then

$$
\|f \widetilde{\otimes}_p g\|^2_{\mathfrak{H}^{\otimes(q-p)}} \leq \|f \otimes_p g\|^2_{\mathfrak{H}^{\otimes(q-p)}} \leq \|f\|^2_{\mathfrak{H}^{\otimes p}} \|g \otimes_{q-p} g\|_{\mathfrak{H}^{\otimes 2p}}.
$$

If $r = p = q$, then $f \widetilde{\otimes}_p g = \langle f, g\rangle_{\mathfrak{H}^{\otimes p}}$. By plugging these last expressions into (4.3.12), we deduce immediately the desired conclusion. □

Let us now recall the following result, which is a collection of some of the findings contained in the papers by Peccati and Tudor [100] and Nualart and Ortiz-Latorre [95].

Theorem 4.3.9 (See [95, 100]). *Fix $d \geq 2$ and let $C = \{C(i,j) : i,j = 1, ..., d\}$ be a $d \times d$ positive definite matrix. Fix integers $1 \leq q_1 \leq \ldots \leq q_d$. For any $n \geq 1$ and $i = 1, \ldots, d$, let $f_i^{(n)}$ belong to $\mathfrak{H}^{\odot q_i}$. Assume that*

$$
F^{(n)} = (F_1^{(n)}, \ldots, F_d^{(n)}) := (I_{q_1}(f_1^{(n)}), \ldots, I_{q_d}(f_d^{(n)})) \quad n \geq 1,
$$

is such that

$$
\lim_{n \to \infty} E[F_i^{(n)} F_j^{(n)}] = C(i,j), \quad 1 \leq i,j \leq d. \tag{4.3.13}
$$

Then, as $n \to \infty$, the following four assertions are equivalent:

(i) For every $1 \leq i \leq d$, $F_i^{(n)}$ converges in distribution to a centered Gaussian random variable with variance $C(i,i)$.

(ii) For every $1 \leq i \leq d$, $E\left[(F_i^{(n)})^4\right] \to 3C(i,i)^2$.

(iii) For every $1 \leq i \leq d$ and every $1 \leq r \leq q_i - 1$, $\|f_i^{(n)} \otimes_r f_i^{(n)}\|_{\mathfrak{H}^{\otimes 2(q_i - r)}} \to 0$.

(iv) The vector $F^{(n)}$ converges in distribution to a d-dimensional Gaussian vector $\mathscr{N}_d(0, C)$.

Moreover, if $C(i,j) = \delta_{ij}$, where δ_{ij} is the Kronecker symbol, then either one of conditions (i)–(iv) above is equivalent to the following:

(v) For every $1 \leq i \leq d$, $\|DF_i^{(n)}\|^2_{\mathfrak{H}} \xrightarrow{L^2} q_i$.

We conclude this section by pointing out the remarkable fact that, for vectors of multiple Wiener-Itô integrals of arbitrary length, *the Wasserstein distance metrizes the weak convergence towards a Gaussian vector with positive definite covariance*. Note that the next statement also contains a generalization of the multidimensional results proved in [95] to the case of an arbitrary covariance.

Proposition 4.3.10. *Fix $d \geq 2$, let C be a positive definite $d \times d$ symmetric matrix, and let $1 \leq q_1 \leq \ldots \leq q_d$. Consider vectors*

$$F^{(n)} := (F_1^{(n)}, \ldots, F_d^{(n)}) = (I_{q_1}(f_1^{(n)}), \ldots, I_{q_d}(f_d^{(n)})), \quad n \geq 1,$$

with $f_i^{(n)} \in \mathfrak{H}^{\odot q_i}$ for every $i = 1 \ldots, d$. Assume moreover that $F^{(n)}$ satisfies condition (4.3.13). Then, as $n \to \infty$, the following three conditions are equivalent:

(a) $d_{\mathrm{W}}(F^{(n)}, Z) \to 0$.

(b) For every $1 \leq i \leq d$, $q_i^{-1}\|DF_i^{(n)}\|_{\mathfrak{H}}^2 \xrightarrow{\mathrm{L}^2} C(i,i)$ and, for every $1 \leq i \neq j \leq d$, $\langle DF_i, -DL^{-1}F_j\rangle_{\mathfrak{H}} = q_j^{-1}\langle DF_i, DF_j\rangle_{\mathfrak{H}} \xrightarrow{\mathrm{L}^2} C(i,j)$.

(c) $F^{(n)}$ converges in distribution to $Z \sim \mathscr{N}_d(0, C)$.

Proof. Since convergence in the Wasserstein distance implies convergence in distribution, the implication (a) → (c) is trivial. The implication (b) → (a) is a consequence of relation (4.3.11). Now assume that (c) is verified, that is, $F^{(n)}$ converges in law to $Z \sim \mathscr{N}_d(0, C)$ as n goes to infinity. By Theorem 4.3.9 we have that, for any $i \in \{1, \ldots, d\}$ and $r \in \{1, \ldots, q_i - 1\}$,

$$\|f_i^{(n)} \otimes_r f_i^{(n)}\|_{\mathfrak{H}^{\otimes 2(q_i - r)}} \underset{n\to\infty}{\longrightarrow} 0.$$

By combining Corollary 4.3.6 with Lemma 4.3.7 (see also Remark 4.3.8(2)), one therefore easily deduces that, since (4.3.13) is in order, condition (b) must necessarily be satisfied. □

4.4 Applications

4.4.1 Convergence of marginal distributions in the functional Breuer-Major CLT

In this section, we use our main results in order to derive an explicit bound for the celebrated *Breuer-Major CLT* for fractional Brownian motion (fBm). We recall that a fBm $B = \{B_t : t \geq 0\}$, with Hurst index $H \in (0,1)$, is a centered Gaussian process, started from zero and with covariance function $E(B_s B_t) = R(s,t)$, where

$$R(s,t) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right); \quad s, t \geq 0.$$

For any choice of the Hurst parameter $H \in (0,1)$, the Gaussian space generated by B can be identified with an isonormal Gaussian process of the type $X = \{X(h) : h \in \mathfrak{H}\}$, where the real and separable Hilbert space $\mathfrak{H}$ is defined as follows: (i) denote by $\mathscr{E}$ the set of all $\mathbb{R}$-valued step functions on $[0,\infty)$, (ii) define $\mathfrak{H}$ as the Hilbert space obtained by closing $\mathscr{E}$ with respect to the scalar product

$$\left\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \right\rangle_{\mathfrak{H}} = R(t,s).$$

In particular, with such a notation, one has that $B_t = X(\mathbf{1}_{[0,t]})$. The reader is referred e.g. to [94] for more details on fBm, including crucial connections with fractional operators. We also define $\rho(\cdot)$ to be the covariance function associated with the stationary process $x \mapsto B_{x+1} - B_x$ ($x \in \mathbb{R}$), that is

$$\rho(x) := \frac{1}{2}\big(|x+1|^{2H} + |x-1|^{2H} - 2|x|^{2H}\big) \underset{|x|\to\infty}{\sim} H|2H-1|\,|x|^{2H-2}.$$

Now fix an integer $q \geq 2$, assume that $H < 1 - \frac{1}{2q}$ and set

$$S_n(t) = \frac{1}{\sigma\sqrt{n}} \sum_{k=0}^{\lfloor nt \rfloor - 1} H_q(B_{k+1} - B_k), \quad t \geq 0,$$

where H_q is the qth Hermite polynomial defined as

$$H_q(x) = (-1)^q \mathrm{e}^{x^2/2} \frac{d^q}{dx^q} \mathrm{e}^{-x^2/2}, \quad x \in \mathbb{R},$$

and where

$$\sigma = \sqrt{q! \sum_{r\in\mathbb{Z}} \rho^2(r)}.$$

According e.g. to the main results in [20] or [46], one has the following CLT:

$$\{S_n(t),\, t \geq 0\} \xrightarrow[n\to\infty]{\text{f.d.d.}} \text{standard Brownian motion,}$$

where 'f.d.d.' indicates convergence in the sense of finite-dimensional distributions. To our knowledge, the following statement contains the first multidimensional bound for the Wasserstein distance ever proved for $\{S_n(t),\, t \geq 0\}$.

Theorem 4.4.1. *For any fixed $d \geq 1$ and $0 = t_0 < t_1 < \ldots < t_d$, there exists a constant c, (depending uniquely on d, H and $(t_0, t_1, \ldots, t_d)$, <u>and not</u> on n) such that, for every $n \geq 1$:*

$$d_{\mathrm{W}}\left(\left(\frac{S_n(t_i) - S_n(t_{i-1})}{\sqrt{t_i - t_{i-1}}}\right)_{1\leq i\leq d}; \mathscr{N}_d(0, \mathbf{I}_d)\right) \leq c \times \begin{cases} n^{-\frac{1}{2}} & \text{if } H \in (0, \frac{1}{2}] \\ n^{H-1} & \text{if } H \in (\frac{1}{2}, \frac{2q-3}{2q-2}] \\ n^{qH-q+\frac{1}{2}} & \text{if } H \in (\frac{2q-3}{2q-2}, \frac{2q-1}{2q}) \end{cases}.$$

Proof. Fix $d \geq 1$ and $t_0 = 0 < t_1 < \ldots < t_d$. In the sequel, c will denote a constant independent of n, which can differ from one line to another.

First, observe that

$$\frac{S_n(t_i) - S_n(t_{i-1})}{\sqrt{t_i - t_{i-1}}} = I_q(f_i^{(n)})$$

with

$$f_i^{(n)} = \frac{1}{\sigma\sqrt{n}\sqrt{t_i - t_{i-1}}} \sum_{k=\lfloor nt_{i-1}\rfloor}^{\lfloor nt_i\rfloor - 1} \mathbf{1}_{[k,k+1]}^{\otimes q}.$$

In [86], proof of Theorem 4.1, it is shown that, for any $i \in \{1, \ldots, d\}$ and $r \in \{1, \ldots, q_i - 1\}$:

$$\|f_i^{(n)} \otimes_r f_i^{(n)}\|_{\mathfrak{H}^{\otimes 2(q_i - r)}} \leq c \times \begin{cases} n^{-\frac{1}{2}} & \text{if } H \in (0, \frac{1}{2}] \\ n^{H-1} & \text{if } H \in (\frac{1}{2}, \frac{2q-3}{2q-2}] \\ n^{qH-q+\frac{1}{2}} & \text{if } H \in (\frac{2q-3}{2q-2}, \frac{2q-1}{2q}) \end{cases}. \tag{4.4.1}$$

Moreover, when $1 \leq i < j \leq d$, we have:

$$\begin{aligned} &\left|\langle f_i^{(n)}, f_j^{(n)}\rangle_{\mathfrak{H}^{\otimes q}}\right| \\ = &\left|\frac{1}{\sigma^2\, n\sqrt{t_i - t_{i-1}}\sqrt{t_j - t_{j-1}}} \sum_{k=\lfloor nt_{i-1}\rfloor}^{\lfloor nt_i\rfloor - 1} \sum_{l=\lfloor nt_{j-1}\rfloor}^{\lfloor nt_j\rfloor - 1} \rho^q(l-k)\right| \\ = &\ \frac{c}{n}\left|\sum_{|r|=\lfloor nt_{j-1}\rfloor - \lfloor nt_i\rfloor + 1}^{\lfloor nt_j\rfloor - \lfloor nt_{i-1}\rfloor - 1} \left[(\lfloor nt_j\rfloor - 1 - r) \wedge (\lfloor nt_i\rfloor - 1) - (\lfloor nt_{j-1}\rfloor - r) \vee (\lfloor nt_{i-1}\rfloor)\right] \rho^q(r)\right| \\ \leq &\ c\, \frac{\lfloor nt_i\rfloor - \lfloor nt_{i-1}\rfloor - 1}{n} \sum_{|r| \geq \lfloor nt_{j-1}\rfloor - \lfloor nt_i\rfloor + 1} \left|\rho(r)\right|^q = O\big(n^{2qH-2q+1}\big), \quad \text{as } n \to \infty, \end{aligned} \tag{4.4.2}$$

the last equality coming from

$$\sum_{|r| \geq N} \left|\rho(r)\right|^q = O\Big(\sum_{|r| \geq N} |r|^{2qH-2q}\Big) = O(N^{2qH-2q+1}), \quad \text{as } N \to \infty.$$

Finally, by combining (4.4.1), (4.4.2), Corollary 4.3.6 and Lemma 4.3.7, we obtain the desired conclusion. ☐

4.4.2 Vector-valued functionals of finite Gaussian sequences

Let $Y = (Y_1, ..., Y_n) \sim \mathscr{N}_n(0, \mathbf{I}_n)$, and let $f : \mathbb{R}^n \to \mathbb{R}$ be an absolutely continuous function such that f and its partial derivatives have subexponential growth at infinity. The following result has been proved by Chatterjee in [24], in the context of limit theorems for linear statistics of eigenvalues of random matrices. We use the notation d_{TV} to indicate the *total variation distance* between laws of real valued random variables.

Proposition 4.4.2 (Lemma 5.3 in [24])**.** *Assume that the random variable $W = f(Y)$ has zero mean and unit variance, and denote by $Z \sim \mathscr{N}(0,1)$ a standard Gaussian random variable. Then,*

$$d_{\mathrm{TV}}(W,Z) \leq 2\mathrm{Var}(T(Y))^{1/2},$$

where the function $T(\cdot)$ is defined as

$$T(y) = \int_0^1 \frac{1}{2\sqrt{t}} \sum_{i=1}^{n} E\big[\frac{\partial f}{\partial y_i}(y)\frac{\partial f}{\partial y_i}(\sqrt{t}\,y + \sqrt{1-t}\;Y)\big]dt.$$

In what follows, we shall use Theorem 4.3.5 in order to deduce a multidimensional generalization of Proposition 4.4.2 (with the Wasserstein distance replacing total variation).

Proposition 4.4.3. *Let $Y \sim \mathscr{N}_n(0,K)$, where $K = \{K(i,l) : i,l = 1,...,n\}$ is a $n \times n$ positive definite matrix. Consider absolutely continuous functions $f_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1,...,d$. Assume that each random variable $f_j(Y)$ has zero mean, and also that each function f_j and its partial derivatives have subexponential growth at infinity. Denote by $Z \sim \mathscr{N}_d(0,C)$ a Gaussian vector with values in $\mathbb{R}^d$ and with positive definite covariance matrix $C = \{C(a,b) : a,b = 1,...,d\}$. Finally, write $W = (W_1,...,W_d) = (f_1(Y),...,f_d(Y))$. Then,*

$$d_{\mathrm{W}}(W,Z) \leq \|C^{-1}\|_{op}\;\|C\|_{op}^{1/2}\sqrt{\sum_{a,b=1}^{d} E[(C(a,b) - T_{ab}(Y))^2]}$$

where the functions $T_{ab}(\cdot)$ are defined as

$$T_{ab}(y) = \int_0^1 \frac{1}{2\sqrt{t}} \sum_{i,j=1}^{n} K(i,j)E\big[\frac{\partial f_a}{\partial y_i}(y)\frac{\partial f_b}{\partial y_j}(\sqrt{t}\,y + \sqrt{1-t}\;Y)\big]dt.$$

Proof. Without loss of generality, we can assume that $Y = (Y_1,...,Y_n) = (X(h_1),...,X(h_n))$, where X is an isonormal Gaussian process over some Hilbert space $\mathfrak{H}$, and $\langle h_i, h_l\rangle_{\mathfrak{H}} = E(X(h_i)X(h_l)) = K(i,l)$. According to Theorem 4.3.5, it is therefore sufficient to show that, for every $a,b = 1,...,d$,

$$T_{ab}(Y) = \langle DW_a, -DL^{-1}W_b\rangle_{\mathfrak{H}}.$$

To prove this last claim, introduce the two $\mathfrak{H}$-valued functions $\Theta_a(y)$ and $\Theta_b(y)$, defined for $y \in \mathbb{R}^d$ as follows:

$$\Theta_a(y) = \sum_{i=1}^{n} \frac{\partial f_a}{\partial y_i}(y)h_i$$

and

$$\Theta_b(y) = \int_0^1 \frac{1}{2\sqrt{t}} \sum_{j=1}^{n} \left\{E\big[\frac{\partial f_b}{\partial y_j}(\sqrt{t}\,y + \sqrt{1-t}\;Y)\big]h_j\right\}dt.$$

By using (4.2.2), it is easily seen that $\Theta_a(Y) = DW_a$. Moreover, by using e.g. formula (3.46) in [86], one deduces that $\Theta_b(Y) = -DL^{-1}W_b$. Since $T_{ab}(Y) = \langle\Theta_a(Y), \Theta_b(Y)\rangle_{\mathfrak{H}}$, the conclusion is immediately obtained. □

By specializing the previous statement to the case $n = d$ and $f_j(y) = y_j$, $j = 1, ..., d$, one obtains the following simple bound on the Wasserstein distance between Gaussian vectors of the same dimension (the proof is straightforward and omitted).

Corollary 4.4.4. *Let $Y \sim \mathscr{N}_d(0, K)$ and $Z \sim \mathscr{N}_d(0, C)$, where K and C are two positive definite covariance matrices. Then,*

$$d_{\mathrm{W}}(Y, Z) \leq Q(C, K) \times \|C - K\|_{H.S.},$$

where

$$Q(C, K) := \min\{\|C^{-1}\|_{op} \, \|C\|_{op}^{1/2}, \|K^{-1}\|_{op} \, \|K\|_{op}^{1/2}\}.$$

Acknowledgments. Part of this paper has been written while I. Nourdin and G. Peccati were visiting the Institute of Mathematical Sciences of the National University of Singapore, in the occasion of the "Workshop on Stein's Method", from March 31 until April 4, 2008. These authors heartily thank A. Barbour, L. Chen, K.-P. Choi and A. Xia for their kind hospitality and support.

Bibliographie

[1] Y. Aït Sahalia and J. Jacod. Testing for jumps in a discretely observed process. *Preprint to appear in Ann. Statist.*, 2006.

[2] Y. Aït Sahalia and J. Jacod. Volatility estimators for discretely sampled lévy processes. *Ann. Statist.*, 35(1) :355–392, 2007.

[3] Y. Aït Sahalia and J. Jacod. Fisher's information for discretely sampled lévy processes. *Econometrica*, 76(4) :727–761, 2008.

[4] D. Aldous and G. Eagleson. On mixing and stability of limit theorems. *Ann. Probab.*, 6(2) :325–331, 1978.

[5] E. Alòs, O. Mazet, and D. Nualart. Stochastic calculus with respect to Gaussian processes. *Ann. Probab.*, 29(2) :766–801, 2001.

[6] R. Averkamp and C. Houdré. Stein estimation for infinitely divisible laws. *ESAIM Probab. Stat.*, 10 :269–276, 2006.

[7] A. Ayache, S. Léger, and M. Pontier. Drap brownien fractionnaire. (french) [the fractional brownian sheet]. *Potential Anal.*, 17(1) :31–43, 2002.

[8] A. D. Barbour. Stein's method for diffusion approximations. *Probab. Theory Related Fields*, 84(3) :297–322, 1990.

[9] O. Barndorff-Nielsen, S. Graversen, J. Jacod, and N. Shephard. Limit theorems for bipower variation in financial econometrics. *Econometric Theory*, 22(4) :677–719, 2006.

[10] O. Barndorff-Nielsen and N. Shephard. Realized power variation and stochastic volatility models. *Bernoulli*, 9(4) :243–265, 2003.

[11] O. Barndorff-Nielsen and N. Shephard. Power and bipower variation with stochastic volatility and jumps. *J. Financial Econometrics*, 2 :1–48, 2004.

[12] O. Barndorff-Nielsen, N. Shephard, and M. Winkel. Limit theorems for multipower variation in the presence of jumps. *Stochastic Process. Appl.*, 116(5) :796–806, 2006.

[13] R. Bass and R. Pyke. The space $\mathcal{D}(\mathcal{A})$ and weak convergence for set-indexed processes. *Ann. Probab.*, 13(3) :860–884, 1985.

[14] J. Berger and R. Wolpert. Estimating the mean function of a Gaussian process and the Stein effect. *J. Multivariate Anal.*, 13(3) :401–424, 1983.

[15] J. Bertoin. *Lévy processes.* Number 121 in Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1996.

[16] H. Biermé and F. Richard. A statistical methodology for testing the anisotropy of brownian textures with an application to full-filled digital mammography. *Preprint*, 2007.

[17] H. Biermé and F. Richard. Estimation of anistropic gaussian fields through radon transform. *ESAIM Probab. Stat. (electronic)*, 12 :30–50, 2008.

[18] P. Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics : Probability and Statistics. John Wiley And Sons, Inc., New York, second edition, 1999.

[19] A. Bonami and A. Estrade. Anisotropic analysis of some gaussian models. *J. Fourier Anal. Appl.*, 9(3) :215–236, 2003.

[20] P. Breuer and P. Major. Central limit theorems for nonlinear functionals of gaussian fields. *J. Multivariate Anal.*, 13(3) :425–441, 1983.

[21] K. Burdzy and J. Swanson. A change of variable formula with itô correction term. *Preprint.*, 2008.

[22] R. Cairoli and J. Walsh. Stochastic integrals in the plane. *Acta Math.*, 134 :111–183, 1975.

[23] E. Carlen and E. Pardoux. Differential calculus and integration by parts on Poisson space. *In S. Albeverio, Ph. Blanchard, and D. Testard, editors, Stochastics, Algebra and Analysis in Classical and Quantum Dynamics (Marseille, 1988), Math. Appl., Kluwer Acad. Publ., Dordrecht*, 59 :pages 63–73, 1990.

[24] S. Chatterjee. Fluctuation of eigenvalues and second order Poincaré inequalities. *Preprint. To appear in Probab. Theory Related Fields*, 2007.

[25] S. Chatterjee and E. Meckes. Multivariate normal approximation using exchangeable pairs. *Preprint. To appear in ALEA*, 2008.

[26] L. Chen and Q.-M. Shao. Stein's method for normal approximation. *In An introduction to Stein's method. Volume 4 of Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., Singapore Univ. Press, Singapore*, pages 1–59, 2005.

[27] L.H.Y. Chen and A. Xia. Stein's method, Palm theory and Poisson process approximation. *Ann. Probab.*, 32(3B) :2545–2569, 2004.

[28] P. Cheridito and D. Nualart. Stochastic integral of divergence type with respect to fractional brownian motion with hurst parameter h in $(0, 1/2)$. *Ann. Inst. H. Poincaré Probab. Statist.*, 41 :1049–1081, 2005.

[29] Z. Ciesielski and A. Kamont. Lévy's fractional brownian random field and function spaces. *Acta Sci. Math.*, 60 :99–118, 1995.

[30] S. Cohen, X. Guyon, O. Perrin, and M. Pontier. Identification of an isometric transformation of the standard brownian sheet. *J. Statist. Plann. Inference*, 136(4) :1317–1330, 2006.

[31] L. Coutin, D. Nualart, and C.A. Tudor. Tanaka formula for the fractional brownian motion. *Stochastic Process. Appl.*, 94(2), 2001.

[32] L. Coutin and M. Pontier. Approximation of the fractional brownian sheet via ornstein-uhlenbeck sheet. *ESAIM Probab. Stat.*, 11, 2007.

[33] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 1610 of *Encyclopedia of Mathematics and its Applications.* University Press, Cambridge, 1992.

[34] R.L. Dobrushin and P. Major. Non-central limit theorems for nonlinear functionals of gaussian fields. *Z. Wahrsch. verw. Gebiete*, 50 :27–52, 1979.

[35] M. Dozzi. *Stochastic processes with a multidimensional parameter*, volume 194 of *Pitman Research Notes in Mathematics Series.* Longman Scientific & Technical, Harlow ; copublished with John Wiley & Sons, New York, 1989.

[36] R.M. Dudley. *Real Analysis and Probability.* Cambridge University Press, Cambridge, second edition, 2003.

[37] D. Dürr, S. Goldstein, and J.L. Lebowitz. Asymptotics of particle trajectories in infinite one-dimensional systems with collisions. *Comm. Pure Appl. Math.*, 38 :573–597, 1985.

[38] E.B. Dynkin. *Markov processes*, volume I, II of *Die Grundlehren der Mathematischen Wissenschaften.* Academic Press Inc., New York, 1965.

[39] R.J. Elliott and A.H. Tsoi. Integration by parts for Poisson processes. *J. Multivariate Anal.*, 44(2) :179–190, 1993.

[40] M. Émery. On the azéma martingales. *In Séminaire de Probabilités, XXIII, Lecture Notes in Math., Springer, Berlin*, 1372, 1990.

[41] D. Fourdrinier. *Statistique inférentielle.* Sciences Sup. Dunod, 2002.

[42] D. Fourdrinier, W.E. Strawderman, and M.T. Wells. On the construction of Bayes minimax estimators. *Ann. Stat.*, 26(2) :660–671, 1998.

[43] V. Genon-Catalot and J. Jacod. On the estimation of the diffusion coefficient for multi-dimensional diffusion process. *Ann. Inst. H. Poincaré Probab. Statist.*, 29(1) :119–151, 1993.

[44] V. Genon-Catalot and J. Jacod. Estimation of the diffusion coefficient for diffusion processes : random sampling. *Scandinavian Journal of Statistics*, 21 :193–221, 1994.

[45] V. Genon-Catalot, C. Laredo, and M. Nussbaum. Asymptotic equivalence of poisson intensity and diffusion drift estimation. *Ann. Statist.*, 30(3) :731–753, 2002.

[46] L. Giraitis and D. Surgailis. Clt and other limit theorems for functionals of gaussian processes. *Z. Wahrsch. verw. Gebiete*, 70 :191–212, 1985.

[47] V. Goodman. Harmonic functions on Hilbert space. *J. Funct. Anal.*, 10 :451–470, 1972.

[48] F. Götze. On the rate of convergence in the multivariate clt. *Ann. Probab.*, 19(2) :724–739, 1991.

[49] M. Gradinaru and I. Nourdin. Convergence of weighted power variations of fractional brownian motion. *Preprint.*, 2007.

[50] M. Gradinaru and I. Nourdin. Stochastic volatility : approximation and goodness-of-fit test. *Probab. Math. Statist.*, 28(1) :1–19, 2007.

[51] M. Gradinaru and I. Nourdin. Weighted power variations of fractional brownian motion and application to approximating schemes. *Preprint*, 2007.

[52] M. Gradinaru, I. Nourdin, F. Russo, and P. Vallois. m-order integrals and itô's formula for non-semimartingale processes ; the case of a fractional brownian motion with any hurst index. *Ann. Inst. H. Poincaré Probab. Statist.*, 41 :781–806, 2005.

[53] M. Gradinaru, F. Russo, and P. Vallois. Generalized covariations, local time and stratonovich itô's formula for fractional brownian motion with hurst index $h \geq \frac{1}{4}$. *Ann. Probab.*, 31 :1772–1820, 2001.

[54] L. Gross. Potential theory on Hilbert space. *J. Funct. Anal.*, 1 :123–181, 1967.

[55] D. Guo, S. Shamai, and S. Verdu. Mutual information and minimum mean-square error in gaussian channels. *IEEE Trans. Inform. Theory*, 51 :1261–1282, 2005.

[56] T.E. Harris. Diffusions with collisions between particles. *J. Appl. Probab.*, 2 :323–338, 1965.

[57] C. Houdré and J. Villa. An example of infinite dimensional quasi-helix stochastic models. *Contemp. Math.*, 336 :195–201, 2003.

[58] E.P. Hsu. Characterization of brownian motion on manifolds through integration by parts. *In Stein's method and applications. Volume 5 of Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., Singapore Univ. Press, Singapore*, pages 185–208, 2005.

[59] I.A. Ibragimov and Y.A. Rozanov. *Gaussian random processes*, volume 9 of *Applications of Mathematics.* Springer-Verlag, New York, 1978.

[60] P. Imkeller. A class of two-parameter stochastic integrators. *Stochastics Stochastics Rep.*, 27(3) :167–188, 1989.

[61] J. Istas and G. Lang. Quadratic variations and estimation of the local hölder index of a gaussian process. *Ann. Inst. H. Poincaré Probab. Stat.*, 33(4) :407–436, 1997.

[62] G. Ivanoff and E. Merzbach. *Set-indexed martingales*, volume 85 of *Monographs on Statistics and Applied Probability.* Chapman & Hall/CRC, 2000.

[63] J. Jacod. Statistics and high frequency data. *Lecture given at the SEMSTAT 2007 held in La Manga (Spain)*, 2007.

[64] J. Jacod. Asymptotic properties of realized power variations and related functionals of semimartingales. *Stoch. Processes and their Appl.*, 118(4) :517–559, 2008.

[65] J. Jacod and J. Mémin. Weak and strong solutions of stochastic differential equations : existence and stability. *Stochastic integrals Lecture Notes in Math., 851, Springer, Berlin-New York, (Proc. Sympos., Univ. Durham, Durham)*, 1981.

[66] J. Jacod and P. Protter. *Probability essentials.* Universitext. Springer-Verlag, Berlin, second edition, 2003.

[67] J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.

[68] W. James and C. Stein. Estimation with quadratic loss. *Proc. 4th Berkeley Sympos. Math. Statist. and Prob. (California Press, Berkeley, Calif. Univ.)*, 1 :361–379, 1961.

[69] Y.M. Kabanov. On extended stochastic integrals. *Theory of Probability and its Applications*, XX(4) :710–722, 1975.

[70] A. Kamont. On the fractional anisotropic wiener field. *Anna . Probab. Math. Statist.*, 16(1) :85–98, 1996.

[71] Y. Kutoyants. Intensity parameter estimation of an inhomogeneous Poisson process. *Problems Control Inform. Theory/Problemy Upravlen. Teor. Inform.*, 8(2) :137–149, 1979.

[72] Y. Kutoyants. *Statistical inference for spatial Poisson processes*, volume 134 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1998.

[73] S. Léger. *Stochastic analysis of multifractal signal and parameters estimations.* PhD thesis, University of Orléans France, 2000.

[74] P. Lei and D. Nualart. A decomposition of the bifractional brownian motioin and some applications. *Preprint*, 2008.

[75] R.S. Liptser and A.N. Shiryaev. *Statistics of random processes II*, volume 6 of *Applications of Mathematics*. Springer-Verlag, Berlin, 2001.

[76] T. Lindstrøm. Fractional brownian fields as integrals of white noise. *Bull. London Math. Soc.*, 25(1) :83–88, 199".

[77] E. Merzbach and D. Nualart. Generalized holomorphic processes and differentiability. *J. Theoret. Probab.*, 2(4) :419–432, 1989.

[78] A. Neuenkirch and I. Nourdin. Exact rate of convergence of some approximation schemes associated to sde's driven by a fractional brownian motion. *J. Theor. Probab.*, 20(4) :871–899, 2007.

[79] G. Neuhaus. On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.*, 42(4) :1285–1295, 1971.

[80] J. Neveu. *Bases mathématiques du calcul des probabilités.* Masson, 1964.

[81] I. Nourdin. Asymptotic behavior of certain weighted quadratic and cubic variations of fractional brownian motion. *Preprint. To appear in Ann. Probab.*, 2007.

[82] I. Nourdin. A simple theory for the study of sde's driven by a fractional brownian motion, in dimension one. *Preprint. To appear in Séminaire de Probabilités* ***XLI***, 2007.

[83] I. Nourdin and D. Nualart. Central limit theorems for multiple skorohod integrals. *Preprint.*, 2008.

[84] I. Nourdin, D. Nualart, and C.A. Tudor. Central and non-central limit theorems for weighted power variations of fractional brownian motion. *Preprint.*, 2007.

[85] I. Nourdin and G. Peccati. Non-central convergence of multiple integrals. *Preprint.*, 2007.

[86] I. Nourdin and G. Peccati. Stein's method on wiener chaos. *Preprint. Under revision for : Probab. Theory Related Fields*, 2007.

[87] I. Nourdin and G. Peccati. Stein's method and exact berry-esséen asymptotics for functionals of gaussian fields. *Preprint.*, 2008.

[88] I. Nourdin and G. Peccati. Weighted power variations of iterated brownian motion. *Electron. J. Probab.*, 13 :1229–1256, 2008.

[89] I. Nourdin, G. Peccati, and A. Réveillac. Multivariate normal approximation using Stein's method and Malliavin calculus. *To appear in Ann. Inst. H. Poincaré Probab. Statist.*, 2008.

[90] I. Nourdin and A. Réveillac. Asymptotic behavior of weighted quadratic variations of fractional brownian motion : the critical case H = 1/4. *Preprint. Submitted.*, 2008.

[91] D. Nualart. On the quadratic variation of two-parameter continuous martingales. *Ann. Probab.*, 12(2) :445–457, 1984.

[92] D. Nualart. Une formule d'itô pour les martingales continues à deux indices et quelques applications. *Ann. Inst. H. Poincaré Probab. Stat.*, 20(3) :251–275, 1984.

[93] D. Nualart. Stochastic calculus with respect to the fractional brownian motion and applications. *Contemp. Math.*, 336 :3–39, 2003.

[94] D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[95] D. Nualart and S. Ortiz-Latorre. Central limit theorem for multiple stochastic integrals and Malliavin calculus. *Stochastic Process. Appl.*, 118(4) :614–628, 2007.

[96] D. Nualart and G. Peccati. Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.*, 33(1) :177–193, 2005.

[97] D. Nualart and J. Vives. A duality formula on the Poisson space and some applications. *In R. Dalang, M. Dozzi, and F. Russo, editors, Seminar on Stochastic Analysis, Random Fields and Applications (Ascona, 1993), Progress in Probability, Birkhäuser, Basel*, 36 :205–213, 1995.

[98] D. Nualart and M. Zakai. Generalized stochastic integrals and the Malliavin calculus. *Probab. Theory Relat. Fields*, 73(2) :255–280, 1986.

[99] G. Peccati. Gaussian approximations of multiple integrals. *Electron. Comm. Prob.*, 12 :350–364, 2007.

[100] G. Peccati and C.A. Tudor. Gaussian limits for vector-valued multiple stochastic integrals. *In Séminaire de Probabilités XXXVIII, Lecture Notes in Math., Springer, Berlin*, pages 247–262, 2005.

[101] N. Privault. Chaotic and variational calculus in discrete and continuous time for the Poisson process. *Stochastics and Stochastics Reports*, 51 :83–109, 1994.

[102] N. Privault. Calcul des variations stochastique pour les martingales. *C. R. Acad. Sci. Paris Sér. I Math.*, 321 :923–928, 1995.

[103] N. Privault. A calculus on Fock space and its probabilistic interpretations. *Bull. Sci. Math.*, 123(2) :97–114, 1999.

[104] N. Privault. *An introduction to stochastic analysis in discrete and continuous settings.* Lecture Notes, 2007.

[105] N. Privault and A. Réveillac. Stein estimation for the drift of Gaussian processes using the Malliavin calculus. *Preprint, to appear in the Annals of Statistics*, 2006.

[106] N. Privault and A. Réveillac. Stein estimation on the Wiener space. *Preprint*, 2006.

[107] N. Privault and A. Réveillac. Superefficient estimation on the Wiener space. *C. R. Acad. Sci. Paris Sér. I Math.*, 343 :607–612, 2006.

[108] N. Privault and A. Réveillac. Stein estimation of poisson process intensities. *Preprint. To appear in Stat. Inference Stoch. Process.*, 2007.

[109] P. Protter. *Stochastic integration and differential equations. A new approach*, volume 21 of *Applications of Mathematics.* Springer-Verlag, Berlin, 1990.

[110] B.L.S. Prakasa Rao. *Statistical inference for diffusion type processes*, volume 8 of *Kendall's Library of Statistics.* Edward Arnold, London, 1999.

[111] G. Reinert. Three general approaches to Stein's method. *In Volume 4 of Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., Singapore Univ. Press, Singapore*, pages 183–221, 2005.

[112] G. Reinert and A. Röllin. Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Preprint.*, 2007.

[113] R. Reiss. *A course on point processes.* Springer Series in Statistics. Springer-Verlag, New York, 1993.

[114] A. Rényi. On mixing sequences of sets. *Acta Math. Acad. Sci. Hung.*, 9 :215–228, 1958.

[115] A. Rényi. On stable sequences of events. *Sankhya, Ser A*, 25 :293–302, 1963.

[116] A. Réveillac. Convergence of finite-dimensional laws of the weighted quadratic variations process for some fractional brownian sheets. *Preprint. To appear in Stoch. Anal. Appl.*, 2007.

[117] A. Réveillac. Estimation of quadratic variation for two-parameter diffusions. *Preprint. To appear in Stochastic Process. Appl.*, 2007.

[118] A. Réveillac. Likelihood ratios and bayesian inference for poisson channels. *Preprint. Submitted.*, 2007.

[119] F. Russo and C.A. Tudor. On the bifractional brownian motion. *Stochastic Process. Appl.*, 116(5) :830–856, 2006.

[120] M. Sanz. r-variations for two-parameter continuous martingales and itô's formula. *Stochastic Process. Appl.*, 32(1) :69–92, 1989.

[121] A. Skorohod. Limit theorems for stochastic processes. *Theory Probab. Appl.*, 1 :261–290, 1956.

[122] C. Stein. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc. 3th Berkeley Sympos. Math. Statist. and Prob. (California Press, Berkeley, Calif. Univ.)*, 1 :197–206, 1956.

[123] 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 (Univ. California, Berkeley, Calif., 1970/1971), Vol. II : Probability theory, Univ. California Press, Berkeley, Calif.*, pages 583–602, 1972.

[124] C. Stein. Estimation of the mean of a multivariate normal distribution. *Ann. Stat.*, 9(6) :1135–1151, 1981.

[125] C. Stein. *Approximate computation of expectations*, volume 7 of *Institute of Mathematical Statistics Lecture Notes, Monograph Series.* Institute of Mathematical Statistics, Hayward, CA, 1986.

[126] M. Straf. Weak convergence of stochastic processes with several parameters. *Proc. of 6th Berkeley Symp. on Math. Statistics and Probability*, 2 :187–221, 1971.

[127] J. Swanson. Variation of the solution to a stochastic heat equation. *Ann. Probab.*, 35(6) :2122–2159, 2007.

[128] M. Taqqu. Convergence of integrated processes of arbitrary hermite rank. *Z. Wahrsch. verw. Gebiete*, 50 :53–83, 1979.

[129] C.A. Tudor and F.G. Viens. Itô formula and local time for the fractional brownian sheet. *Electron. J. Probab.*, 8, 2003.

[130] C.A. Tudor and F.G. Viens. Itô formula for the two-parameter fractional brownian motion using the extended divergence operator. *Stochastics*, 78(6) :443–462, 2006.

[131] C.A. Tudor and F.G. Viens. Statistical aspects of the fractional stochastic calculus. *Ann. Statist.*, 35(3) :1183–1212, 2007.

[132] A.S. Üstünel. *An introduction to analysis on Wiener space*, volume 1610 of *Lecture Notes in Mathematics.* Springer Verlag, 1995.

[133] A.S. Üstünel. Estimation for the additive gaussian channel and monge–kantorovitch measure transportation. *Stochastic Process. Appl.*, 117 :1316–1329, 2007.

[134] S. Verdu. Poisson communication theory. *Invited talk, March 25 1999. The International Technion Communication Day in honor of Israel Bar-David.*, 1999.

[135] R. Wolpert and J. Berger. Incorporating prior information in minimax estimation of the mean of a Gaussian process. *Statistical decision theory and related topics, III, West Lafayette, Ind., Academic Press, New York*, 2 :451–464, 1982.

[136] E. Wong and M. Zakai. A characterization of the kernels associated with the multiple integral representation of some functionals of the wiener process. *Systems Control Lett.*, 2(2) :94–98, 1982.

[137] L. Wu. A new modified logarithmic sobolev inequality for poisson point processes and several applications. *Probab. Theory Related Fields*, 118(3) :427–438, 2000.

[138] M. Zakai. On mutual information, likelihood-ratios and estimation error for addaitive gaussian channel. *IEEE Trans. Inform. Theory*, 51(9) :3017–3024, 2005.

Résumé

Dans cette thèse nous appliquons le calcul de Malliavin à l'estimation statistique de paramètres de certains processus stochastiques et à l'obtention de théorèmes de la limite centrale pour les variations quadratiques à poids de processus fractionnaires et/ou à deux paramètres ainsi qu'à l'approximation gaussienne de mesures de probabilités multidimensionnelles. Dans le Chapitre 1 nous construisons des estimateurs de type Stein pour la dérive de processus gaussiens et pour l'intensité de processus de Poisson. Dans le Chapitre 2 nous calculons l'estimateur bayésien du signal d'entrée d'un canal de Poisson et nous étendons notre résultat aux canaux dont le bruit est une martingale normale possédant la propriété de représentation chaotique. Dans le Chapitre 3 nous établissons des théorèmes de la limite centrale pour les variations quadratiques à poids du drap brownien standard (nous permettant de donner un estimateur asymptotiquement normal de la variation quadratique de certains processus de diffusion à deux paramètres) puis pour celles de certains draps browniens fractionnaires. Dans ce même chapitre nous établissons un théorème de la limite centrale pour les variations quadratiques à poids du mouvement brownien fractionnaire d'indice $H = 1/4$ nous permettant de donner le comportement asymptotique des sommes de Riemann à signe alterné associées au mouvement brownien fractionnaire d'indice $H = 1/4$. Enfin dans le Chapitre 4 nous appliquons la méthode de Stein et du calcul de Malliavin afin d'obtenir des bornes explicites pour l'approximation gaussienne multidimensionnelle de fonctionnelles de champs gaussiens. Nous appliquons en particulier nos résultats aux théorème de la limite centrale de Breuer et Major pour des champs associés à un mouvement brownien fractionnaire.

Abstract

In this thesis we apply the Malliavin calculus to statistical estimation of parameters of stochastic processes and to derive limit theorems for the weighted quadratic variations of one or two-parameter fractional processes and to multidimensional normal approximation of probability measures. In Chapter 1 we construct Stein type estimators for the drift of Gaussian processes and for the intensity of Poisson processes. In Chapter 2, we compute the Bayesian estimator of the input of a Poisson channel then extended to normal martingales with chaotic representation property channels. In Chapter 3 we derive central limit theorems for the weighted quadratic variations of the standard Brownian sheet (applied then to the obtaining of an asymptotically normal estimator of the quadratic variation of some two-parameter diffusion processes) and of some fractional Brownian sheets. Then in this chapter we establish a central limit theorem for the weighted quadratic variations of the fractional Brownian motion with Hurst index $H = 1/4$ leading to the study of the asymptotic behavior of the Riemann sums with alternating signs associated to the fractional brownian motion with Hurst index $H = 1/4$. Finally in Chapter 4 we apply Stein's method and the Malliavin calculus in order to obtain explicit bounds in the multidimensional normal approximation of functionals of gaussian fields. In particular we provide an application to a functional version of the Breuer-Major TCL for fields subordinated to a fractional Brownian motion.