



Laboratoire
Paul Painlevé

DÉPARTEMENT DE
MATHÉMATIQUES



Université
de Lille

ÉCOLE DOCTORALE MADIS-631 : MATHÉMATIQUES, SCIENCES DU NUMÉRIQUE ET DE
LEURS INTERACTIONS

THÈSE

pour obtenir le grade de docteur délivré par

l'Université de Lille

dans la spécialité : **Mathématiques**

présentée et soutenue publiquement par

Julie Gamain

le 08 décembre 2023

Contributions à l'étude des matrices aléatoires et à l'inférence statistique des EDPS par le calcul de Stein-Malliavin

Directeur de thèse : **Ciprian Tudor**

Jury

M. Ivan Nourdin,

M. Yvik Swan,

M. Jean-Christophe Breton,

Mme. Céline Duval,

M. Ciprian Tudor,

Professeur, Université du Luxembourg

Professeur, Université Libre de Bruxelles

Professeur, Université de Rennes

Professeure, Université de Lille

Professeur, Université de Lille

Rapporteur

Rapporteur

Président du jury

Examinatrice

Directeur

À mon père.

Résumé

Le cadre sur lequel s'érige cette thèse est assez vaste puisqu'il regroupe un large champ d'études, allant des équations stochastiques aux dérivées partielles à la théorie des matrices et tenseurs aléatoires, dont le point commun est l'analyse des p -variations de processus. L'outil majeur qui a permis son élaboration est *le calcul de Malliavin combiné à la méthode de Stein*.

Cette thèse se dessine en trois parties distinctes dont la première évoquera l'inférence statistique des équations différentielles stochastiques. En premier lieu, elle mettra en avant l'estimation des paramètres de dérive et de diffusion pour les équations des ondes et de la chaleur fractionnaire dirigées par un bruit blanc en temps et en espace, qui a la possibilité d'être non-linéaire. Plus précisément, ces travaux porteront sur une exploration minutieuse du comportement des variations quadratiques de la solution associées à ces équations.

En effet, d'une part, nous établissons un théorème central limite obtenu par l'application du célèbre théorème du quatrième moment sur les variations quadratiques spatiales et temporelles de la solution de l'équation des ondes dirigée par un bruit en temps et en espace.

En outre, nous nous sommes évertués à l'estimation du paramètre de dérive de l'équation de la chaleur fractionnaire dirigée par un bruit gaussien non linéaire, blanc en temps et en espace. Pour ce faire, nous avons tiré profit de l'existence d'une connexion forte entre le mouvement brownien fractionnaire et la solution spatiale de l'équation de la chaleur fractionnaire dirigée par un bruit gaussien additif. Au moyen d'un critère d'approximation basé sur les accroissements de la solution linéaire, nous procédons à l'investigation autour de la convergence des variations quadratiques spatiales de la solution dans le cas non linéaire.

Par ailleurs, les deux autres chapitres constituant la thèse donneront accès à une réflexion autour de l'étude asymptotique des matrices et tenseurs de Wishart par l'utilisation de la méthode de Stein-Malliavin.

Déjà largement étudiés au cours des dernières années, les auteurs ont examiné ces objets lorsque les composantes de la matrice initiale associée à la matrice de Wishart étaient gaussiennes, log-concaves ou de loi complètement générale et avec la considération supplémentaire d'une corrélation nulle, partielle ou totale. Notre contribution consistera à apporter un résultat lorsque les coefficients de la matrice initiale demeurent gaussiens mais en relâchant l'hypothèse d'équidistribution. Afin d'y parvenir nous choisissons des coefficients particuliers, correspondant aux accroissements temporels de la solution de l'équation stochastique des ondes. Nous traitons un cas similaire pour l'équation stochastique de la chaleur qui donnera la possibilité d'obtenir une convergence presque sûre vers une matrice diagonale.

Notre souhait a ensuite été de généraliser les résultats concernant les matrices de Wishart à une certaine classe de tenseurs aléatoires, dont la dénomination serait les tenseurs de Wishart. Leur intérêt grandissant dû aux développements des domaines applicatifs nous a conduit à cet axe de recherche pertinent. En particulier, nous nous sommes concentrés sur la situation dans laquelle les composantes de notre vecteur initial associé au tenseur de Wishart sont des variables aléatoires appartenant au second chaos de Wiener. Plus précisément, ce sont des accroissements de processus de Rosenblatt et de surcroît la loi des coefficients n'est pas gaussienne. Il s'est avéré que les termes de l'hyperdiagonale du tenseur, vu comme un vecteur, dominaient et correspondaient aux p -variations du processus de Rosenblatt. Grâce à une étude fine de ces variations, nous parvenons à approcher la loi de notre tenseur de Wishart par celle d'un vecteur aléatoire dont la majorité des composantes sont des variables aléatoires de Rosenblatt.

Abstract

The thesis framework is quite broad, gathering a wide range of studies, from stochastic partial differential equations to the theory of random matrices and random tensors, whose common theme is the analysis of p -variations of stochastic processes. The main tool employed is Malliavin's calculus combined with Stein's method.

This thesis is split into three distinct parts, the first one focusing on the statistical inference of stochastic partial differential equations.

Firstly, this chapter highlights the drift or diffusion parameters estimation for fractional heat and wave equations driven by a space time white noise. The random noise may be nonlinear. More precisely, those works rely on the study of asymptotic behavior of quadratic variations of the solution associated to these equations. Actually, on the one hand, we establish a central limit theorem obtained by applying the well-known fourth moment theorem to the spatial and temporal quadratic variations of the solution to the wave equation driven by a space time white noise.

On the other hand, we also estimate the drift parameter of the fractional stochastic heat equation driven by a nonlinear Gaussian space time white noise. To achieve this, we use the strong connection between the fractional Brownian motion and the solution of fractional heat equation driven by additive Gaussian noise. Through an approximation criterion based on the increments of the linear solution, we investigate the convergence of the spatial quadratic variations of the solution in the nonlinear case.

Furthermore, the other two chapters constituting the thesis provide insights into the asymptotic study of random matrices and random tensors using the Stein-Malliavin method.

This problem has already been widely studied in the last years, when the components of the starting matrix associated to Wishart matrix are Gaussian, log-concave, or with a general distribution, with zero, partial, or total correlation. Our contribution consists in obtaining results when the entries of the initial matrix are still Gaussian but relaxing the equi-distribution assumption. In order to do it, we choose certain particular entries, namely the temporal increments of the solution to the stochastic wave equation. We make a similar analysis for the stochastic heat equation, applying the same methodology.

Then, we intended to generalize the findings concerning the Wishart matrices to a class of random tensors, the so-called Wishart tensors. Their growing significance for practical aspects of the real-life led many researchers to their study. In particular, we focused on the situation where the components of our initial random vector associated to Wishart tensor are random variables belonging to the second Wiener chaos. In fact, they are increments of Rosenblatt process, and therefore, the distribution of the entries is not Gaussian. It turned out that the terms on the hyperdiagonal of the tensor, viewed as a vector, are dominant and they are related to the p -variations of the Rosenblatt process. Through a detailed analysis of these variations, we are able to approximate the distribution of our Wishart tensor with that of a random vector, whose components are Rosenblatt random variables or null.

Remerciements

Nombreuses sont les personnes qui ont été d'un appui indéniable durant la réalisation de cette thèse, auxquelles je souhaiterais exprimer ma profonde gratitude et mes innombrables remerciements.

Tout d'abord, je tenais à remercier mon directeur de thèse, Ciprian Tudor. Sans lui, rien n'aurait été possible, il m'a laissé l'opportunité de pouvoir développer et enrichir mes compétences grâce à ses connaissances et son expertise qui ont été essentielles pour l'aboutissement de ce travail. Ses conseils éclairés, son soutien constant et la proposition de passionnants sujets de recherche auront fait de ce doctorat, une très belle expérience. Je n'aurais pas pu espérer meilleur directeur de thèse que vous et je voulais vous adresser de nouveau mes immenses remerciements.

Désormais, j'aimerais exprimer ma reconnaissance envers Ivan Nourdin et Yvik Swan qui ont accepté la lourde tâche de rapporteur. Je les remercie d'avoir pris soin de relever toutes les erreurs et coquilles qu'ils ont pu apercevoir et je suis véritablement honorée de pouvoir les compter parmi les membres de mon jury. Leurs travaux ont été une source d'inspiration au cours de ce doctorat. Pour l'anecdote, le *livre bleu* intitulé *Normal approximation with Malliavin calculus* d'Ivan Nourdin et Giovanni Peccati est devenu *mon livre sacré* au cours de ces trois dernières années. Un grand merci pour votre travail qui permet aux novices de pouvoir se pencher sur le calcul de Malliavin et la méthode de Stein.

Je tenais également à saluer la présence de Jean-Christophe Breton et Céline Duval, membres de mon jury.

J'adresse mes remerciements envers les membres du laboratoire Paul Painlevé, plus précisément envers Antoine Ayache, Mylène Maïda, Philippe Heinrich, mes anciens enseignants. Je pense aussi à Sabine, toujours disponible et efficace et à Frédérique.

Le bureau 12 aura été un lieu d'échanges marqué par les rencontres avec de belles personnes dont les discussions ont toujours été très enrichissantes. Je songe à François, Ivan, Antonin mais aussi Florent, Jérôme et Théo.

Je pense aussi à Obayda Assaad qui a accompagné mes premiers pas en tant que doctorante durant la crise sanitaire puis à Charles-Philippe Diez. Je me sens extrêmement chanceuse d'avoir été encadrée par des personnes aussi talentueuses, à l'écoute et pourvues de gentillesse.

Pour clore ce doctorat, l'arrivée de Jérémy Zurcher fut la cerise sur le *Gâteaux*, merci à toi JéZu pour tes conseils avisés, ta bienveillance et tes précieuses relectures.

Ayant fait tout mon cursus au sein de l'université de Lille, j'ai vu partir beaucoup de mes amis après la licence et le master mais les souvenirs restent encreés. Encadrer les TD dans cette faculté où j'ai effectué toutes mes études a fait resurgir de vieux souvenirs, des moments passés avec eux. J'aimerais saluer toutes ces personnes que j'ai eu la chance de rencontrer dans le mythique bâtiment M1 car ils ont rendu ces années inoubliables. Je pense à Enguerrand, Paul, Henri, Marine, Margot, François, Jordan, Johan, Nico, Sandro, Marion et bien d'autres qui j'espère ne me tiendront pas rigueur d'avoir oublié de les citer dans l'instant où j'écris ces lignes.

Je porte une attention particulière à Max, notre jeune marié et à Mayday qui arrivent également à la fin de leur doctorat. Je leur souhaite beaucoup de courage et de succès. Tenez bon, la ligne d'arrivée est encore loin mais bien plus proche que celle de la SaintéLyon.

Pour finir, je souhaiterais adresser ma profonde reconnaissance à mon père Hervé qui m'a toujours accompagnée et encadrée au cours de mes études, dans mon sport et à ma mère Maryline. Je pense aussi et surtout à mes adorables frère et soeur, Lucyle et Evann, à ma famille.

Table des matières

Résumé	5
Abstract	7
Remerciements	9
Organisation	17
I Un outil analytique en probabilité : de la théorie de Malliavin à l’alliance avec la méthode de Stein	19
1 Calcul de Malliavin et processus stochastiques	21
1.1 Introduction au calcul de Malliavin	21
1.2 Processus stochastiques	27
1.2.1 Définitions	27
1.2.2 Mouvement brownien fractionnaire	28
1.2.3 Mouvement brownien bi-fractionnaire	29
1.2.4 Processus d’Hermite	31
2 La fructueuse rencontre entre Stein et Malliavin	35
2.1 Distances	36
2.2 Méthode de Stein unidimensionnelle sur l’espace gaussien	38
2.3 De la méthode de Stein au calcul de Malliavin : le théorème du quatrième moment	41
2.4 Méthode de Stein multidimensionnelle et TCL pour les matrices aléatoires	44
2.4.1 Méthode de Stein multidimensionnelle sur l’espace gaussien	44
2.4.2 Approximation gaussienne dans le cas multivarié pour des vecteurs aléatoires et matrices aléatoires	49
II Axes de recherche : EDPS, matrices et tenseurs aléatoires	53
3 Équations stochastiques aux dérivées partielles dirigées par un bruit blanc en temps et en espace	55
3.1 Introduction aux équations aux dérivées partielles stochastiques	55
3.2 Propriétés liées aux SPDEs de la chaleur et des ondes	59
3.2.1 Équation stochastique de la chaleur	60
3.2.2 Équation stochastique de la chaleur fractionnaire	61
3.2.3 Équation stochastique des ondes	62

3.3	Résultats	63
3.3.1	Résumé du chapitre 6	63
3.3.2	Résumé du chapitre 7	67
4	Matrices aléatoires	73
4.1	Introduction aux matrices aléatoires	73
4.1.1	Approche spectrale pour l'étude des matrices aléatoires	74
4.1.2	Analyse non spectrale des matrices de Wishart	77
4.2	Étude asymptotique des matrices aléatoires par la méthode de Stein Malliavin	82
4.2.1	Coefficients gaussiens	82
4.2.2	Coefficients non gaussiennes	90
4.3	Résultats	95
4.3.1	Résumé du chapitre 8	95
4.3.2	Résumé du chapitre 9	100
5	Étude asymptotique des tenseurs aléatoires	107
5.1	Introduction aux tenseurs aléatoires	107
5.2	Analyse non spectrale des tenseurs de Wishart	110
5.2.1	Étude des tenseurs de Wishart par la méthode de Stein-Malliavin	110
5.2.2	Étude des tenseurs de Wishart basée sur une nouvelle utilisation de la méthode de Stein	114
5.3	Résultats : Résumé du chapitre 10	116
 III Inférence statistique pour les équations différentielles stochastiques : estimation de paramètres		 125
6	Quadratic variation and drift parameter estimation for the stochastic wave equation with space-time white noise	127
6.1	Introduction	127
6.2	Preliminaries	129
6.3	Asymptotic behavior of the temporal quadratic variation and estimation of the drift parameter	130
6.3.1	Convergence in $L^2(\Omega)$ and almost sure	131
6.3.2	Central Limit Theorem	132
6.3.3	Optimality of the rate of convergence	136
6.4	Estimation of the drift parameter	137
6.5	Spatial quadratic variation	141
6.5.1	Limit behavior of the spatial quadratic variation	141
6.5.2	On parameter estimation via quadratic variation	143
6.6	Appendix : Multiple stochastic integrals and the Malliavin derivative	145
7	Exact variation and drift parameter estimation for the nonlinear fractional stochastic heat equation	149
7.1	Introduction	149
7.2	Preliminaries : On the stochastic fractional heat equation and its solution	151
7.2.1	The Dalang-Walsh integral	151
7.2.2	Properties of the Green kernel	152
7.2.3	Some technical lemmas on the Green kernel	153
7.3	On the increments of the fractional Brownian motion	153

7.4	The spatial quadratic variation of the solution in the nonlinear case	157
7.4.1	The linear case	157
7.4.2	The nonlinear case	158
7.5	Drift parameter estimation	165
7.6	Appendix : Proofs of the technical lemmas	168
IV Etude du comportement asymptotique des matrices et tenseurs de Wishart par la méthode de Stein-Malliavin		169
8	Random matrices and the stochastic wave equation	171
8.1	Introduction	171
8.2	The stochastic wave equation and its quadratic variation	173
8.2.1	The spatial quadratic variation	174
8.2.2	Temporal quadratic variation	180
8.3	The Wishart matrix	183
8.3.1	The case where entries are temporal increments of the solution	185
8.3.2	The case where entries are spatial increments of the solution	189
8.4	Appendix : Wiener chaos and Malliavin derivative	190
9	High-dimensional regime for Wishart matrices based on the increments of the solution to the stochastic heat equation	193
9.1	Introduction	193
9.2	The stochastic heat equation and its quadratic variation	195
9.2.1	Basic properties of the solution	195
9.2.2	Limit behavior of the spatial quadratic variation	196
9.2.3	Renormalized spatial quadratic variation and Central Limit Theorem	200
9.3	The Wishart matrix	204
9.4	Appendix	208
9.4.1	Basics of the Malliavin calculus	208
9.4.2	Normal approximation and Malliavin calculus	209
10	Limit behavior in high-dimensional regime for Wishart tensors with Rosenblatt entries	211
10.1	Introduction	211
10.2	Preliminaries	213
10.2.1	Wiener chaos and multiple stochastic integrals	213
10.2.2	A general product formula	214
10.2.3	The Rosenblatt process	215
10.3	The starting matrix and the associated Wishart tensor	215
10.4	The p -variation of the Rosenblatt process	216
10.4.1	The limit of the term in the second Wiener chaos	219
10.4.2	The terms in the chaos of order $p > 2$ are negligible	224
10.5	Asymptotic behavior of the Wishart tensor	226
10.5.1	The limit of the hyper-diagonal terms	227
10.5.2	The limit behavior of the other terms	227
10.5.3	Conclusion	229
Bibliographie		230

Organisation

La première partie de ce manuscrit est dédiée à l'introduction de la théorie initiée par Paul Malliavin. Elle nous servira à avancer les principaux outils nécessaires dans nos thèmes de recherche, notamment les processus stochastiques et l'aboutissement de la rencontre entre la méthode de Stein et la théorie fraîchement citée.

La seconde partie reprendra toutes les notions abordées dans nos articles, allant des équations stochastiques aux dérivées partielles aux matrices et tenseurs aléatoires. Nous fournirons alors un résumé de ceux-ci dans chacune des parties respectives, à savoir ;

1. *Quadratic variation and drift parameter estimation for the stochastic wave equation with space-time white noise*, en collaboration avec Obayda Assaad et Ciprian Tudor [8],
2. *Exact variation and drift parameter estimation for the nonlinear fractional stochastic heat equation*, en collaboration avec Ciprian Tudor [55],
3. *Random matrices and the stochastic wave equation*, en collaboration avec Ciprian Tudor [53],
4. *High-dimensional regime for Wishart matrices based on the increments of the solution to the stochastic heat equation*, en collaboration avec David A.C. Mollinedo et Ciprian Tudor [52],
5. *Limit behavior in high-dimensional regime for Wishart tensors with Rosenblatt entries*, en collaboration avec Ciprian Tudor [54].

Les parties suivantes renvoient aux axes de recherche que nous venons d'énumérer.

Nous vous souhaitons une agréable lecture !

Partie I

**Un outil analytique en probabilité :
de la théorie de Malliavin à l'alliance
avec la méthode de Stein**

Calcul de Malliavin et processus stochastiques

1.1 Introduction au calcul de Malliavin

Le calcul de Malliavin, autrement dénommé calcul stochastique des variations, est un calcul différentiel infini-dimensionnel sur l'espace de Wiener ; l'espace canonique du mouvement brownien. Il a été développé dans les années 1970 grâce à Paul Malliavin [86] et a permis l'avancement d'une preuve probabiliste du théorème d'Hörmander sur les opérateurs différentiels hypoelliptiques et l'élaboration d'une puissante méthode pour montrer l'existence et la régularité de la densité de vecteurs aléatoires.

Son développement a considérablement été motivé par la conception d'outils probabilistes afin de répondre aux besoins des applications ; la portée des applications en finance, plus précisément liées aux modèles basés sur le mouvement brownien. Nous renvoyons au livre de E. Alos et D. Garcia Lorite [1] ou au travail de F. Cosmao, F. Dupuy et A. Guillon [36] pour plus de précisions concernant ce domaine applicatif.

Le calcul de Malliavin a également trouvé sa place dans l'étude de la régularité des lois de probabilité de solutions d'EDS/EDPS, la méthode de Stein aboutissant à la méthode de Stein-Malliavin dont on peut trouver une description dans [93] et plus récemment l'inférence statistique. Le travail fondateur de Paul Malliavin inspira de nombreux travaux autour de l'hypoellipticité d'opérateurs différentiels elliptiques dégénérés d'ordre 2 et l'étude des équations différentielles stochastiques.

Dans la suite, nous n'exposerons pas de manière détaillée la théorie de Malliavin. Nous nous limitons aux outils nécessaires à la compréhension de ce manuscrit. Toutefois, nous proposons de renvoyer les curieux lecteurs vers les livres de David Nualart [100], Ivan Nourdin et Giovanni Peccati [93], Laurent Decreusefond [42], pour plus de détails concernant le calcul de Malliavin. Pour une élégante introduction à cette théorie, nous nous tournons aussi vers la thèse d'Hélène Halconruy [56] qui fournit une pertinente approche au calcul de Malliavin et expose différents points de vue pour le présenter (point de vue variationnel ou approche chaotique).

Comme nous l'avons déjà évoqué, le calcul de Malliavin était, historiquement, un calcul différen-

tiel en dimension infinie sur l'espace de Wiener, $\mathcal{C}([0, 1], \mathbb{R}^d)$. Par analogie au calcul différentiel classique sur les espaces de Banach, nous pouvons percevoir les trajectoires du mouvement brownien comme les vecteurs. Aux fonctions de la théorie usuelle se substituent les fonctionnelles qui opèrent sur l'espace des trajectoires. De plus, le gradient, ou bien l'opérateur différentiation qui est la dérivée de Malliavin traduit la dépendance d'une variable aléatoire par rapport aux accroissements du mouvement brownien. L'opérateur adjoint de l'opérateur de dérivation au sens de Malliavin se réfère à l'opérateur de divergence.

La définition d'un opérateur de dérivation n'a pas été tâche facile. En effet, une manière assez naturelle de le définir était en tant que dérivée de Fréchet (une fonction différentiable au sens de Fréchet est continue).

Cette manière d'imaginer la dérivée n'a pu aboutir puisque, par exemple, les solutions d'équations différentielles stochastiques avec des coefficients réguliers ne sont pas continues pour la norme de l'espace de Wiener. En ce sens, voir la dérivée de Malliavin sur l'espace de Wiener comme une dérivée de Fréchet n'est pas approprié.

La bonne manière d'interpréter la dérivée de Malliavin comme une dérivée directionnelle s'est faite dans la direction de l'espace de Cameron-Martin (voir [42]).

Nous allons dessiner les contours de la théorie dans un cadre plus général, celui défini sur un champ gaussien infini-dimensionnel. Nous allons définir la dérivée de Malliavin sur la classe de variables aléatoires cylindriques et l'étendre à un sous espace des variables aléatoires de carrés intégrables en utilisant un critère de densité.

Le point de départ est de considérer un espace de Hilbert réel séparable $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$. Sur cet espace de Hilbert, nous définissons le concept central de la théorie, qui est le processus isonormal.

Définition 1. *Le processus $(X(h), h \in \mathcal{H})$, indexé par des éléments de \mathcal{H} est un processus gaussien isonormal s'il forme une famille gaussienne centrée définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, et s'il est tel que pour tout $h, g \in \mathcal{H}$,*

$$\mathbb{E}[X(h)X(g)] = \langle h, g \rangle_{\mathcal{H}}.$$

En particulier, un processus gaussien isonormal est un isomorphisme entre un espace gaussien fermé dont les variables aléatoires sont centrées et appartenant à $L^2(\Omega)$ et un espace de Hilbert séparable \mathcal{H} .

L'existence d'un tel processus gaussien isonormal sur l'espace de Hilbert \mathcal{H} est assurée par le théorème de Kolmogorov.

De plus, notons qu'en se donnant un espace de Hilbert séparable \mathcal{H} , nous pouvons montrer qu'il existe un unique processus gaussien isonormal sur \mathcal{H} . Par exemple, si nous considérons l'espace de Hilbert $\mathcal{H} = L^2((0, \infty), \mathcal{B}(0, \infty), \lambda)$, où λ est la mesure de Lebesgue, nous pouvons associer à \mathcal{H} , un processus gaussien isonormal W tel que $B_t := W(\mathbf{1}_{[0, t]})$, où $(B_t)_{t \geq 0}$ est le mouvement Brownien standard.

Dans la suite nos éléments sont définis sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, où \mathcal{F} est la σ -algèbre engendrée par $X = (X(h))_{h \in \mathcal{H}}$.

Soit l'espace \mathbf{S} , l'ensemble des variables aléatoires dites cylindriques qui est composé des variables aléatoires F de la forme

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

où h_1, \dots, h_n sont des éléments de \mathcal{H} et $f: \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction appartenant à l'espace de Schwartz $\mathcal{S}(\mathbb{R}^n)$. Notons que l'espace \mathbf{S} est dense dans $L^p(\Omega)$ pour tout entier $p \geq 1$.

La dérivée de Malliavin d'une variable aléatoire F appartenant à l'espace \mathbf{S} est un élément de $L^2(\Omega, \mathcal{H})$ défini par

$$DF = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X(h_1), \dots, X(h_n)) h_i.$$

Par exemple, si on se donne un mouvement Brownien standard, nous pouvons l'écrire comme un processus gaussien isonormal $B_t = X(\mathbf{1}_{[0,t]})$ dont la dérivée est $D_s B_t = \mathbf{1}_{[0,t]}(s)$. L'opérateur de dérivation de Malliavin associée à une variable aléatoire, un processus.

Par itération, nous pouvons déterminer la q -ième dérivée comme un élément de $L^2(\Omega, \mathcal{H}^{\otimes q})$, (où $\mathcal{H}^{\otimes q}$ désigne le q -ième produit tensoriel symétrique).

Définition 2. Soit $F \in \mathbf{S}$ et soit $q \geq 1$ un entier. La q -ième dérivée de Malliavin (par rapport à X) de la variable aléatoire cylindrique F est un élément de $L^2(\Omega, \mathcal{H}^{\otimes q})$ définie par

$$D^q F := \sum_{i_1, \dots, i_q=1}^n \frac{\partial^q f}{\partial x_{i_1} \dots \partial x_{i_q}}(X(h_1), \dots, X(h_n)) h_{i_1} \otimes \dots \otimes h_{i_q}.$$

Remarque 1. La dérivée de Malliavin ne dépend pas de l'écriture de $F \in \mathbf{S}$.

L'opérateur de dérivation au sens de Malliavin a été défini sur l'espace des variables aléatoires cylindriques \mathbf{S} . Cependant, il est possible de l'étendre à des espaces plus "gros". Comme \mathbf{S} est dense dans $L^p(\Omega)$ pour tout entier $p \geq 1$, nous voudrions étendre l'opérateur D^q à $L^p(\Omega)$. Le problème étant que cet opérateur n'est pas continu, et les méthodes usuelles d'extension ne peuvent pas être appliquées. Il faut alors se contenter d'une extension moins riche.

Pour $p \in [1, \infty)$ et un entier $q \geq 1$, l'opérateur de dérivation $D^q : \mathbf{S} \subset L^p(\Omega) \rightarrow L^p(\Omega, \mathcal{H}^{\otimes q})$ est un opérateur fermable. Nous pouvons en trouver une preuve dans le livre de Nourdin et Peccati [93] (Proposition 2.3.4) ou dans l'excellent mémoire de M2 de Jérémy Zurcher [128].

Nous définissons l'espace $\mathbb{D}^{q,p}$, qui est l'adhérence de \mathbf{S} pour la norme,

$$\|F\|_{\mathbb{D}^{q,p}} = (\mathbb{E}[|F|^p] + \mathbb{E}[\|DF\|_{\mathcal{H}}^p] + \dots + \mathbb{E}[\|D^q F\|_{\mathcal{H}^{\otimes q}}^p])^{1/p}.$$

De cette manière, pour tout $q \geq 1$, la q -ième dérivée de Malliavin D^q peut être étendue à l'espace $\mathbb{D}^{q,p}$, qui est appelé le domaine de l'opérateur fermé D^q sur $L^p(\Omega)$ et se nomme espace de Gross-Stroock-Sobolev.

Une propriété importante que vérifie la dérivée de Malliavin est la règle de la chaîne.

Proposition 1. Soit $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction lipschitzienne et soit $F = (F_1, \dots, F_n)$ un vecteur aléatoire dont les composantes appartiennent à $\mathbb{D}^{1,p}$ pour $p > 1$. Si la loi de F est absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^n , alors $\varphi(F) \in \mathbb{D}^{1,p}$ et

$$D\varphi(F) = \sum_{i=1}^n \frac{\partial \varphi}{\partial x_i}(F) DF_i.$$

Soit $q \geq 1$ un entier fixé. L'opérateur adjoint d'ordre q de l'opérateur de dérivation $D^q : \mathbb{D}^{q,2} \rightarrow L^2(\Omega, \mathcal{H}^{\otimes q})$ est noté δ^q . Nous notons $\text{Dom} \delta^q$, le domaine de δ^q tel que

$$\text{Dom} \delta^q = \left\{ u \in L^2(\Omega, \mathcal{H}^{\otimes q}); \exists c > 0, \forall F \in \mathbb{D}^{q,2}, |\mathbb{E}[\langle D^q F, u \rangle_{\mathcal{H}^{\otimes q}}]| \leq c \sqrt{\mathbb{E}[F^2]} \right\}.$$

L'existence de l'opérateur de divergence est une conséquence du théorème de Riesz.

Il existe un lien intrinsèque entre la dérivée de Malliavin et l'opérateur de divergence donné par la formule d'intégration par parties ou formule de dualité.

Définition 3. Si $u \in \text{Dom} \delta^q$, alors $\delta^q(u)$ est l'unique élément de $L^2(\Omega)$ caractérisé par la formule de dualité ;

$$\mathbb{E}[F \delta^q(u)] = \mathbb{E}[\langle D^q F, u \rangle_{\mathcal{H}^{\otimes q}}], \text{ pour tout } F \in \mathbf{S}.$$

Proposition 2. Soient $F \in \mathbb{D}^{1,2}$ et $u \in \text{Dom} \delta$ tels que $\mathbb{E}[F^2 \|u\|_{\mathcal{H}}^2]$, $\mathbb{E}[F^2 \delta(u)^2]$ et $\mathbb{E}[\langle DF, u \rangle_{\mathcal{H}}^2]$ sont finies. Alors $Fu \in \text{Dom} \delta$ et

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

Remarque 2. L'opérateur de divergence permet d'introduire une notion d'intégration pour les processus non-adaptés et généralise l'intégrale d'Itô.

Nous cherchons désormais à définir les chaos de Wiener de manière à apporter une description de l'intégrale stochastique multiple. Pour cela, nous rappelons la notion de polynôme d'Hermite à l'aide de la formule de Rodrigues.

Définition 4. Le p -ième polynôme d'Hermite est donné par :

$$\forall p \geq 1, H_p(x) := (-1)^p e^{\frac{x^2}{2}} \frac{d^p}{dx^p} \left[e^{-\frac{x^2}{2}} \right].$$

Par exemple $H_1(x) = x$, $H_2(x) = x^2 - 1$, $H_3(x) = x^3 - 3x$.

Définition 5 (Chaos de Wiener). Soit $(X(h), h \in \mathcal{H})$ un processus gaussien isonormal défini sur l'espace de Hilbert \mathcal{H} . Pour tout $q \geq 0$ nous notons \mathfrak{H}_q la fermeture du sous espace vectoriel de $L^2(\Omega, \mathcal{F}, \mathbb{P})$ engendré par les variables aléatoires $\{H_q(X(h)), h \in \mathcal{H}, \|h\|_{\mathcal{H}} = 1\}$.

Remarque 3. 1. Nous avons $\mathfrak{H}_0 = \mathbb{R}$ et $\mathfrak{H}_1 = \{X(h), h \in \mathcal{H}\}$.

2. En particulier, lorsque $q = 1$ on retrouve toutes les variables aléatoires gaussiennes. En revanche les chaos d'ordre supérieur $q \geq 2$ contiennent des variables aléatoires non-gaussiennes (voir le théorème 11).

Cela nous permet de définir l'intégrale multiple à l'aide des polynômes d'Hermite.

Définition 6. Soit $f \in \mathcal{H}$ telle que $\|f\|_{\mathcal{H}} = 1$. Alors pour tout entier $q \geq 1$ on définit l'application $H_q(X(h)) = I_q(h^{\otimes q})$ qui s'étend en application linéaire isométrique du produit tensoriel symétrique $\mathcal{H}^{\odot q}$ (muni de la norme $\|\cdot\|_{\mathcal{H}^{\odot q}} = \frac{1}{\sqrt{q!}} \|\cdot\|_{\mathcal{H}^{\otimes q}}$) vers le q -ième chaos de Wiener.

On obtient donc pour tout $f \in \mathcal{H}^{\odot p}$, $g \in \mathcal{H}^{\odot p'}$ et $p, p' \geq 1$,

$$\mathbb{E}(I_p(f) I_{p'}(g)) = \delta_{p,p'} \times p! \langle f, g \rangle_{\mathcal{H}^{\otimes p}}$$

où $\delta_{p,p'}$ est le symbole de Kronecker.

En particulier, il est possible de montrer que les polynômes d'Hermite vérifient la relation suivante.

Si $Y, Z \sim \mathcal{N}(0, 1)$ sont des variables aléatoires conjointement gaussiennes, alors pour tout $n, m \geq 0$,

$$\mathbb{E}[H_m(Y) H_n(Z)] = \begin{cases} n! (\mathbb{E}[YZ])^n & \text{si } m = n, \\ 0 & \text{sinon.} \end{cases}$$

Il en découle que si $n \neq m$, les chaos \mathfrak{H}_n et \mathfrak{H}_m sont orthogonaux pour le produit scalaire défini sur $L^2(\Omega)$ et que $\bigoplus_{n=0}^{\infty} \mathfrak{H}_n$ est en somme directe dans $L^2(\Omega)$. En ce sens, on peut montrer que $L^2(\Omega)$

peut s'écrire comme $\bigoplus_{n=0}^{\infty} \mathfrak{H}_n$.

Cela signifie que toute variable aléatoire de carré intégrable peut se décomposer comme

$$F = \mathbb{E}[F] + \sum_{n=1}^{\infty} F_n,$$

où $F_n \in \mathfrak{H}_n$ et la série converge dans $L^2(\Omega)$. Ainsi, il en résulte la proposition suivante en utilisant la définition 6.

Proposition 3. *Toute variable aléatoire $F \in L^2(\Omega)$ peut s'écrire comme,*

$$F = \mathbb{E}[F] + \sum_{p=1}^{\infty} I_p(f_p)$$

pour une unique collection de noyau $f_p \in \mathcal{H}^{\odot p}$.

Nous allons maintenant énoncer deux formules très importantes dans la théorie de Malliavin. Le premier résultat concerne le produit d'intégrales multiples d'ordres différents. Il permet notamment de donner une expression de linéarisation.

Théorème 1. *Soient $p, q \geq 1$. Si $f \in \mathcal{H}^{\odot p}$ et $g \in \mathcal{H}^{\odot q}$, alors*

$$I_p(f)I_q(g) = \sum_{r=0}^{\min(p,q)} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}(f \tilde{\otimes}_r g). \quad (1.1)$$

Dans le cas où $\mathcal{H} = L^2(T)$, la contraction d'ordre $r = 1, \dots, p$, notée $f \otimes_r g$, est une fonction de $L^2(T^{p+q-2r})$. Pour tout $f \in L^2(T^p)$ et $g \in L^2(T^q)$, nous avons

$$(f \otimes_r g)(t_1, \dots, t_{p+q-2r}) = \int_{T^r} f(u_1, \dots, u_r, t_1, \dots, t_{p-r}) g(u_1, \dots, u_r, t_{p-r+1}, \dots, t_{p+q-2r}) du_1 \dots du_r. \quad (1.2)$$

Remarque 4. 1. $f \otimes_r g$ n'est pas nécessairement une fonction symétrique (même si f, g sont symétriques) et nous notons $f \tilde{\otimes}_r g$ sa symétrisée. Pour obtenir $f \tilde{\otimes}_r g$, il suffit d'appliquer la définition de la symétrisée des fonctions f et g .

2. Si f et g sont symétriques et que nous prenons $f \in L^2_S(T^p)$ et $g \in L^2_S(T^q)$ alors $f \tilde{\otimes}_r g = f \otimes_r g$. Notons que $L^2_S(T)$ désigne l'espace des fonctions symétriques appartenant à $L^2(T)$.
3. Par convention, la contraction d'ordre 0 est égale au produit tensoriel, i.e $f \otimes_0 g = f \otimes g$ et si $p = q = r$ alors $f \otimes_p g = \langle f, g \rangle_{L^2(T^p)}$.

Le second résultat évoque le caractère d'hypercontractivité des chaos. La nécessité d'une telle propriété est primordiale puisqu'elle permet d'aboutir au fait que dans un chaos, toutes les normes $L^q(\Omega)$ sont équivalentes.

Théorème 2 (Hypercontractivité). *Pour tout $p, q \geq 1$, il existe une constante $0 < k(p, q) < \infty$ dépendant seulement de p et de q , telle que*

$$(\mathbb{E}[|Y|^q])^{1/q} \leq k(p, q) \mathbb{E}[Y^2]^{\frac{1}{2}}, \quad (1.3)$$

pour toute variable aléatoire $Y = I_p(f)$.

La preuve de ce résultat découle de l'hypercontractivité du semi-groupe d'Ornstein-Uhlenbeck que nous pouvons définir de la manière suivante.

Définition 7. Soit $f_n \in L^2_{\mathbb{S}}(\mathbb{R}_+^n)$. On définit le semi-groupe d'Ornstein–Uhlenbeck $(P_t)_{t \geq 0}$, pour tout $F = \sum_{n=0}^{\infty} I_n(f_n) \in L^2(\Omega)$ par :

$$P_t(F) = \sum_{n=0}^{\infty} e^{-nt} I_n(f_n).$$

Nous pouvons aussi définir cet opérateur par la formule de Mehler.

Théorème 3. Soient X et X' deux processus gaussiens isonormaux indépendants définis sur l'espace de Hilbert \mathcal{H} . Pour tout $F \in L^2(\Omega)$ mesurable par rapport à X , il existe $\phi : \mathbb{R}^{\mathcal{H}} \rightarrow \mathbb{R}$, tel que $\phi(X) = F$ et

$$P_t(F) = \mathbb{E}' \left[\phi(e^{-t}X + \sqrt{1 - e^{-2t}}X') \right], \quad (1.4)$$

où \mathbb{E}' est l'espérance mathématique sous la loi de X' .

À cet opérateur s'associe son générateur infinitésimal.

Définition 8. Le générateur du semi-groupe d'Ornstein–Uhlenbeck est noté L et défini par,

$$LF = - \sum_{n=0}^{\infty} n I_n(f_n),$$

pour F dans le domaine $\text{Dom}(L)$,

$$\text{Dom}(L) = \left\{ F = \sum_{n=0}^{\infty} I_n(f_n), \sum_{n=0}^{\infty} n^2 n! \|f_n\|_{L^2(T^n)}^2 < \infty \right\}$$

En particulier on a une propriété qui lie le générateur infinitésimal aux opérateurs de dérivation et divergence de Malliavin.

Proposition 4. Soit $F \in L^2(\Omega)$, alors $F \in \text{Dom}(L)$ si et seulement si $F \in \mathbb{D}^{1,2}$ et $DF \in \text{Dom}\delta$ et dans ce cas,

$$\delta(DF) = -LF.$$

Nous donnons à titre d'exemple le cas où $F = I_n(f)$, pour tout $f \in \mathcal{H}^{\otimes n}$. Dans cette situation, nous avons d'une part $DF = nI_{n-1}(f)$ et $\delta DF = nI_n(f)$. D'autre part, nous obtenons que $LF = -nI_n(f) = -\delta DF$.

Pour en finir avec ces concepts introductifs et représentatifs de la théorie de Malliavin, nous proposons de donner une définition du pseudo-inverse de l'opérateur L .

Définition 9. L'opérateur pseudo inverse de L , noté L^{-1} , est défini pour tout $F \in L^2(\Omega)$, par :

$$L^{-1}F = - \sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n).$$

Pour tout $F \in L^2(\Omega)$, $L^{-1}F \in \text{Dom}(L)$ et $LL^{-1}F = F - \mathbb{E}[F]$.

1.2 Processus stochastiques

1.2.1 Définitions

On se place sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. On rappelle qu'un processus stochastique $X = (X_t)_{t \in T}$, où T est un ensemble non vide, est une collection de variables aléatoire définies sur $(\Omega, \mathcal{F}, \mathbb{P})$. C'est-à-dire, pour tout $t \in T$, l'application $X_t : \Omega \rightarrow \mathbb{R}$ est \mathcal{F} -mesurable.

Nous allons énoncer des propriétés importantes qui peuvent être caractéristiques de certains processus. La première que nous évoquons est la stationnarité. Si un processus est à accroissements stationnaires, cela signifie que sa loi est invariante par translation. En particulier nous donnons, aussi, la définition d'un processus faiblement et strictement stationnaire.

Définition 10 (Stationnarité). *Soit $T = \mathbb{R}$ ou $T = \mathbb{R}_+$.*

1. *Un processus stochastique $(X_t)_{t \in T}$ est dit **strictement stationnaire** si pour tout $t_1, \dots, t_n \in T$ et pour tout $k \in T$, les vecteurs aléatoires, $(X_{t_1+k}, \dots, X_{t_n+k})$ et $(X_{t_1}, \dots, X_{t_n})$ ont même loi au sens des distributions finies dimensionnelles.*
2. *Un processus de carré intégrable $(X_t)_{t \in T}$ est **faiblement stationnaire** si $\mathbb{E}[X_{t+h}] = \mathbb{E}[X_t]$ pour tout $t, h \in T$ et*

$$\text{Cov}(X_t, X_s) = \text{Cov}(X_{t-s}, X_0) \text{ pour tout } s, t \in T,$$

de sorte que $\text{Cov}(X_t, X_s)$ soit une fonction de $t - s$ pour tout $s \leq t$.

3. *Un processus $X = (X_t)_{t \in T}$ est à **accroissements stationnaires** si pour tout $h \in T$, le processus stochastique*

$$(X_{t+h} - X_t)_{t \in T}$$

a la même loi que le processus X , au sens des distributions finies dimensionnelles.

Remarque 5. *Si $X = (X_t)_{t \in T}$ est un processus gaussien centré, les notions de faiblement stationnaire et strictement stationnaire coïncident.*

Définition 11. *Soit $(X_t)_{t \geq 0}$ un processus de carré intégrable à accroissements stationnaires. On définit sa fonction d'auto-covariance $r_n := \mathbb{E}[(X_1 - X_0)(X_{n+1} - X_n)]$.*

1. *Si*

$$\sum_{n \in \mathbb{N}} |r(n)| < \infty,$$

le processus est à mémoire courte.

2. *Si*

$$\sum_{n \in \mathbb{N}} |r(n)| = \infty,$$

le processus est à mémoire longue.

Une seconde propriété pouvant caractériser les processus stochastiques est l'auto-similarité. Un processus auto-similaire est invariant en loi par changement d'échelle à un facteur multiplicatif près.

Définition 12 (Auto-similarité). *Un processus stochastique $(X_t)_{t \geq 0}$ est dit auto-similaire d'ordre H s'il existe $H > 0$ tel que pour tout $c > 0$ les processus $(X_{ct})_{t \geq 0}$ et $(c^H X_t)_{t \geq 0}$ ont même loi au sens des distributions finies dimensionnelles. H s'appelle l'indice d'auto-similarité.*

Dans la partie suivante, nous proposons de donner des exemples de processus stochastiques qui vérifient les propriétés énoncées et qui nous serviront au cours de cette thèse.

1.2.2 Mouvement brownien fractionnaire

En 1940, Andreï Kolmogorov [77] introduisit le mouvement brownien fractionnaire dans ses travaux sur les espaces de Hilbert. Initialement peu populaire, c'est sa propriété de mémoire longue qui fera sa notoriété. En effet, son histoire prend surtout racine lorsque l'hydrologue Harold Edwin Hurst [62] met en évidence des corrélations dans les données du niveau des crues annuelles du Nil. Il s'agit donc de trouver un processus autre que le mouvement brownien (puisque l'une de ses principales propriétés est d'être un processus à accroissements indépendants) dont les accroissements sont corrélés afin de permettre la modélisation de ce phénomène. Nous renvoyons au livre d'Ivan Nourdin [92] pour une mise en contexte mathématique.

En 1968, Mandelbrot remarque lui aussi une dépendance à long terme dans ses travaux en finance et avec Van Ness [88], ils populariseront le processus stochastique autrefois appelé *Wiener Helix*, sous le nom de mouvement brownien fractionnaire.

Nous introduirons les principales propriétés de ce processus, qui nous seront utiles au cours de cette thèse. Nous vous proposons de consulter les livres de David Nualart [100], Ivan Nourdin [92] et Ciprian Tudor [116], [117] pour retrouver les preuves des propriétés énoncées.

Définition 13 (Mouvement brownien Fractionnaire). *Soit $T \subset \mathbb{R}$. Le mouvement brownien fractionnaire $(B_t^H)_{t \in T}$ dit fBm (pour fractional Brownian motion) est un processus gaussien centré auto-similaire d'ordre $H \in (0, 1]$ à accroissements stationnaires de fonction de covariance,*

$$\mathbb{E}[B_t^H B_s^H] = \frac{1}{2} (|t|^{2H} + |s|^{2H} - |t-s|^{2H}),$$

et tel que $\mathbb{E}[(B_1^H)^2] = 1$.

Nous pouvons montrer que sa fonction de covariance est symétrique définie positive. L'existence du fBm est alors assurée par le théorème de Kolmogorov.

Remarque 6. 1. En prenant $T = \mathbb{R}$ et $H = \frac{1}{2}$ dans la définition 13, nous obtenons le mouvement brownien bilatéral $(W_t)_{t \in \mathbb{R}}$ dit two-sided Brownian motion en anglais défini par

$$W_t = \begin{cases} W_t^1 & \text{si } t \geq 0, \\ W_{-t}^2 & \text{si } t < 0, \end{cases}$$

où $(W_t^1)_{t \geq 0}$ et $(W_t^2)_{t \geq 0}$ sont deux mouvements browniens standards indépendants.

2. Si $T = [0, \infty)$, le processus gaussien $(B_t^H)_{t \geq 0}$ est appelé mouvement brownien fractionnaire standard de paramètre de Hurst $H \in (0, 1]$. En particulier, pour $H = \frac{1}{2}$, nous retrouvons le mouvement brownien classique.

Le fBm de paramètre de Hurst $H \in (0, 1)$ peut aussi être exprimé comme une intégrale de Wiener. Une représentation possible est celle en moyenne mobile. Soient $(B_t^H)_{t \geq 0}$ un mouvement brownien fractionnaire standard et $(W_t)_{t \geq 0}$ un processus de Wiener, alors cette représentation nous donne que :

$$B_t^H = \frac{1}{\Gamma(\frac{1}{2} + H)} \int_{-\infty}^t \left((t-s)_+^{H-\frac{1}{2}} - (-s)_+^{H-\frac{1}{2}} \right) dW_s,$$

où Γ représente la fonction Gamma d'Euler. Ce n'est pas la seule représentation stochastique existante. En effet, nous pouvons le représenter comme un processus de Volterra par exemple. Même si le fBm standard peut-être vu comme une extension du mouvement brownien standard, il ne conserve pas le caractère markovien et le fait d'être une semi-martingale. Nous proposons d'énumérer ci-dessous quelques propriétés du mouvement brownien fractionnaire standard.

1. Le fBm standard $(B_t^H)_{t \geq 0}$ est stable par renversement du temps. Cela signifie que les processus $\left(t^{2H} B_{\frac{1}{t}}^H\right)_{t \geq 0}$ et $(B_t^H)_{t \geq 0}$ ont la même loi au sens des distributions finies dimensionnelles.
2. Nous avons que $\mathbb{E}|B_s^H - B_t^H|^2 = |s - t|^{2H}$, et en appliquant le théorème de continuité de Kolmogorov-Centsov, nous pouvons montrer que le fBm d'indice $H \in (0, 1)$ admet une version dont les trajectoires sont h lderiennes d'ordre $\delta \in (0, H)$.
3. Le fBm n'est pas un processus de Markov pour $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$.
4. Pour $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$, le fBm n'est pas non plus une semi-martingale.
5. Soit r la fonction d'auto-covariance telle que $r(n) := \mathbb{E}[B_1^H (B_{n+1}^H - B_n^H)]$.
 - (a) Si le param tre de Hurst $H \in (0, \frac{1}{2}]$ alors

$$\sum_{n \in \mathbb{Z}} |r(n)| < \infty,$$

et dans cette situation, le processus est   memoire courte.

- (b) En revanche pour $H \in (\frac{1}{2}, 1)$, le processus est   memoire longue car

$$\sum_{n \in \mathbb{Z}} |r(n)| = \infty.$$

L'association d'un processus isonormal au fBm standard permet la construction d'une int grale de Wiener par rapport   celui-ci. Pour une description compl te des r sultats nous renvoyons aux travaux de L. Decreusefond et A.S.  st nel [43].

Remarque 7. *Le mouvement brownien fractionnaire est  troitement li  aux solutions d' quations diff rentielles stochastiques et notamment   la solution spatiale $(u(t, x))_{x \in \mathbb{R}}$ de l' quation de la chaleur avec laplacien fractionnaire d'ordre $\alpha \in (1, 2]$, dirig e par un bruit blanc en temps et en espace W et donn e par,*

$$\frac{\partial}{\partial t} u(t, x) = -(-\Delta)^{\frac{\alpha}{2}} u(t, x) + \dot{W}(t, x), \quad t \geq 0, x \in \mathbb{R}. \quad (1.5)$$

En effet, pour tout $t \in (0, T]$, le processus gaussien centr  $(u(t, x))_{x \in \mathbb{R}}$ peut se d composer comme ;

$$u(t, x) = m_\alpha V(t, x) + Y(t, x), \quad \forall x \in \mathbb{R},$$

o  $(Y(t, x))_{x \in \mathbb{R}}$ est un processus gaussien centr    trajectoires suffisamment r guli res, m_α une constante d pendant de α que nous expliciterons plus tard et $V(t, x)$ un processus stochastique dont la loi co cide avec celle d'un mouvement brownien fractionnaire d'indice $H = \frac{\alpha-1}{2}$.

1.2.3 Mouvement brownien bi-fractionnaire

Le succ s du mouvement brownien fractionnaire provient de son aspect attractif pour la mod lisation dans beaucoup de domaines applicatifs tels que la finance ou la t l communication. Ceci est notamment d  aux diff rentes propri t s que nous avons  voqu es,   savoir, l'auto-similarit  et ses accroissements stationnaires. Cependant, en 2003, C. Houdr  et J. Villa [63] ont fait remarquer que pour de grands accroissements, le fBm ne semble plus  tre appropri  pour la mod lisation de certaines donn es. Ils ont alors cherch  un processus qui gardait en partie les propri t s d'auto-similarit , de stationnarit  des petits accroissements et la gaussianit  mais en l' largissant pour le besoin des applications. Le processus qu'ils ont alors introduit est le mouvement brownien bi-fractionnaire d pendant de deux indices H et K . Nous pouvons trouver une exposition des r sultats li s   ce processus dans les livres de Ciprian Tudor [116], [117].

Définition 14. Soient $H \in (0, 1)$, $K \in (0, 1]$ deux paramètres et $T \subset \mathbb{R}$. Le mouvement brownien bi-fractionnaire de paramètres de Hurst H et K est défini comme un processus gaussien centré $(B_t^{H,K})_{t \in T}$ dont la covariance est donnée pour tout $s, t \in T$ par,

$$\mathbb{E}[B_t^{H,K} B_s^{H,K}] = R_{H,K}(t, s) = \frac{1}{2^K} ((t^{2H} + s^{2H})^K - |t - s|^{2HK}).$$

Le mouvement brownien bi-fractionnaire est une extension du mouvement brownien fractionnaire, les deux processus coïncident pour $K = 1$.

La fonction de covariance étant bien définie positive, l'existence de ce processus découle du théorème de Kolmogorov.

Il partage beaucoup de propriétés avec le mouvement brownien fractionnaire dont nous dressons une liste non-exhaustive.

1. Le processus $(B_t^{H,K})_{t \in \mathbb{R}}$ est un processus auto-similaire d'ordre HK .
2. Il n'est pas à accroissements stationnaires sauf dans le cas où il coïncide avec le fBm, c'est-à-dire lorsque $K = 1$, pour $B^{H,1}$ avec $H \in (0, 1)$.
3. Si $2HK \neq 1$, alors $B^{H,K}$ n'est pas une semi-martingale.
4. L'inégalité (1.6) montre que le processus $B^{H,K}$ est une quasi-hélice au sens de J-P. Kahane [71].

De plus, nous pouvons dire que le mouvement brownien bi-fractionnaire $B^{H,K}$ préserve l'Hölder continuité du fBm car c'est un processus à trajectoires hölderiennes d'ordre δ pour tout $0 < \delta < H$. Cela provient de la proposition suivante.

Proposition 5. Soient $H \in (0, 1)$, $K \in (0, 1]$ et $(B_t^{H,K})_{t \geq 0}$ un mouvement brownien bi-fractionnaire. Pour tout $s, t \in \mathbb{R}$, nous avons,

$$2^{-K}|t - s|^{2HK} \leq \mathbb{E}[(B_t^{H,K} - B_s^{H,K})^2] \leq 2^{1-K}|t - s|^{2HK}. \quad (1.6)$$

Il n'existe pas de représentation en intégrale de Wiener comme possède le mouvement brownien fractionnaire. Cependant, P. Lei et D. Nualart [80] ont révélé une relation entre ces deux processus.

Théorème 4. Soit $(B_t^{H,K})_{t \geq 0}$ un mouvement Brownien bi-fractionnaire. Considérons $(W_\theta)_{\theta \geq 0}$ un processus de Wiener indépendant de $B^{H,K}$. Pour tout $t \geq 0$, nous définissons un processus gaussien centré $(X_t^K)_{t \geq 0}$ tel que

$$X_t^K = \int_0^\infty (1 - e^{-\theta t}) \theta^{-\frac{1+K}{2}} dW_\theta.$$

Posons $X_t^{H,K} := X_{t^{2H}}^K$, alors les processus

$$\left(C_1 X_t^{H,K} + B_t^{H,K}\right)_{t \geq 0} \quad \text{et} \quad \left(C_2 B_t^{H,K}\right)_{t \geq 0}$$

ont la même loi, où $C_1 = \sqrt{\frac{2^{-K}K}{\Gamma(1-K)}}$ et $C_2 = 2^{\frac{1-K}{2}}$.

Les trajectoires de $(X_t^K)_{t \geq 0}$ sont régulières, ce processus admet une modification dont les trajectoires sont absolument continues et infiniment différentiable sur $(0, \infty)$.

Remarque 8. 1. Le mouvement brownien bi-fractionnaire est également étroitement lié à la solution temporelle $(u(t,x))_{t \geq 0}$ de l'équation de la chaleur avec laplacien fractionnaire dirigée par un bruit blanc en temps et en espace, donnée par l'équation (1.5). L'expression de sa covariance permet de constater que le processus $(u(t,x))_{t \geq 0}$ a la même loi que le processus gaussien $\left(c_{2,\alpha} B_t^{\frac{1}{2}, 1-\frac{1}{\alpha}}\right)_{t \geq 0}$ avec $c_{2,\alpha}$ une constante dépendant de α .

$\left(B_t^{\frac{1}{2}, 1-\frac{1}{\alpha}}\right)_{t \geq 0}$ est en fait le mouvement brownien bi-fractionnaire pour $H = \frac{1}{2}$ et $K = 1 - \frac{1}{\alpha}$.

Comme une relation entre le fBm et le mouvement brownien bi-fractionnaire est donnée par le théorème 4, nous pouvons dire que

$$(u(t,x) + Y_t)_{t \geq 0} \stackrel{(d)}{\equiv} \left(2^{\frac{1}{2\alpha}} c_{2,\alpha} B_t^{\frac{1}{2} - \frac{1}{2\alpha}}\right)_{t \geq 0}$$

où $(Y_t)_{t \geq 0}$ est un processus gaussien à trajectoires absolument continues et $\left(B_t^{\frac{1}{2} - \frac{1}{2\alpha}}\right)_{t \geq 0}$

un mBf de paramètre de Hurst $H = \frac{1}{2} - \frac{1}{2\alpha}$.

2. Nous pouvons remarquer le même type de relation entre le mouvement brownien fractionnaire et la solution temporelle de l'équation de la chaleur dirigée par un bruit blanc en temps et en espace.

Notons qu'un calcul stochastique par rapport au mouvement brownien bi-fractionnaire a été développé dans l'article [107] de F. Russo et C. Tudor.

1.2.4 Processus d'Hermité

Désormais, nous proposons de donner une autre généralisation du mouvement brownien fractionnaire, en s'affranchissant du caractère gaussien mais en gardant ses propriétés fondamentales. Nous voulons donc préserver l'auto-similarité du processus et ses accroissements stationnaires, dont la dernière caractéristique citée n'était plus valable pour le mouvement brownien bi-fractionnaire.

En 1979, l'intérêt grandissant pour l'étude des processus auto-similaires a poussé Dobrushin, Major [44] et Taqqu [112] à s'intéresser à ces objets aléatoires aux accroissements stationnaires et possédant une dépendance à long terme. Cette dépendance se traduit par le fait que la corrélation entre les accroissements de nos processus vérifierait celle du mouvement brownien fractionnaire $(B_t^H)_{t \geq 0}$ et en particulier, quand $s \rightarrow \infty$

$$\mathbb{E}[(B_{t+1}^H - B_t^H)(B_{t+s+1}^H - B_{s+t}^H)] \sim H(2H - 1)s^{2H-2}.$$

Cette décroissance lente vers 0 de la corrélation exprime alors la dépendance à long terme. Les auteurs ont cherché à déterminer la limite d'une somme de variables aléatoires qui vérifiaient ces propriétés, sous le caractère de dépendance à long terme. Les considérations mathématiques sont spécifiées dans l'énoncé suivant.

Soient g une fonction mesurable de rang d'Hermité k (le rang du premier coefficient non-nul dans la décomposition de g dans la base des polynômes d'Hermité) et $(\xi_n)_{n \geq 0}$ une suite de variables aléatoires gaussiennes, centrées et réduites telles que

$$r(n) := \mathbb{E}[\xi_0 \xi_n] = n^{\frac{2H-2}{k}} L(n),$$

où $H \in (\frac{1}{2}, 1)$ et $k \geq 1$. On suppose que $\mathbb{E}[g(\xi_0)] = 0$ et $\text{Var}[g(\xi_0)] < \infty$.

Une fonction mesurable $L : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ est dite à variations lentes à l'infini si

$$\forall m \geq 0, \quad \frac{L(nm)}{L(n)} \xrightarrow{n \rightarrow \infty} 1.$$

Alors on a que,

$$\frac{1}{n^H} \sum_{j=1}^{[nt]} g(\xi_j)$$

converge quand $n \rightarrow \infty$ au sens des distributions finies dimensionnelles vers le processus

$$Z_H^k(t) = c(H, k) \int_{\mathbb{R}^k} \int_0^t \left(\prod_{j=1}^k (s - y_j)_+^{-\left(\frac{1}{2} + \frac{1-H}{k}\right)} \right) ds dB(y_1) \dots dB(y_k),$$

où $x_+ = \max(0, x)$ et l'intégrale définie ci-dessus est une intégrale multiple de Wiener-Itô d'ordre k par rapport au mouvement brownien bilatéral $(B(y))_{y \in \mathbb{R}}$. $c(H, k)$ est une constante de renormalisation qui assure que $\mathbb{E}[(Z_H^k(t))^2] = 1$, telle que

$$c(H, k)^2 = \left(\frac{\beta\left(\frac{1}{2} - \frac{1-H}{k}, \frac{2H-2}{k}\right)^k}{k! H (2H-1)} \right)^{-1}.$$

Le processus $(Z_H^k(t))_{t \geq 0}$ est appelé le processus d'Hermite d'ordre k et de paramètre H . Il apparaît comme la limite, au sens des distributions finies dimensionnelles, d'une somme de variables aléatoires. Ce que nous venons d'énoncer n'est rien d'autre qu'un théorème non-central limite. Nous remarquons que le processus d'Hermite peut s'écrire sous la forme d'une intégrale stochastique multiple d'ordre k ,

$$Z_H^k(t) = I_k(f_t),$$

avec

$$f_t(u_1, \dots, u_k) = \int_0^t \left(\prod_{j=1}^k (s - y_j)_+^{-\left(\frac{1}{2} + \frac{1-H}{k}\right)} \right) ds.$$

Son noyau f_t étant dans $L^2(\mathbb{R}^k)$ pour tout $t \geq 0$, le processus d'Hermite est bien défini.

Comme nous l'avons évoqué, il vérifie une bonne majorité des propriétés associées au fBm (voir [116]).

1. Le processus d'Hermite a la même fonction de covariance que son précurseur le mouvement brownien fractionnaire pour $H \in (1/2, 1)$. En effet, pour tout $k \geq 1$, et $s, t \geq 0$, elle est donnée par

$$\mathbb{E}[Z_H^k(t)Z_H^k(s)] = R_H(t, s) = \frac{1}{2} (t^{2H} + s^{2H} - |t-s|^{2H}).$$

2. Le processus d'Hermite est un processus auto-similaire d'ordre $H \in (\frac{1}{2}, 1)$.
3. Il est à accroissements stationnaires, pour tout $s > 0$ les processus $(Z_H^k(t+s) - Z_H^k(t))_{t \geq 0}$ et $(Z_H^k(t))_{t \geq 0}$ coïncident au sens des distributions finies-dimensionnelles.
4. C'est un processus à trajectoires hölderiennes d'ordre δ avec $0 < \delta < H$. De plus, il a été récemment prouvé que presque-sûrement ses trajectoires ne sont nulles part dérivables. Ce résultat est dû à A. Ayache dans [12].

5. Il présente une dépendance à long terme.
6. Dans l'article [87], M. Maejima et C. Tudor construisent une intégrale de Wiener par rapport au processus d'Hermite. Cette étude a été nécessaire pour pouvoir établir les solutions des équations de Ornstein–Uhlenbeck avec un bruit de type Hermite.

Pour $k = 1$, le processus Z_H^1 est gaussien et correspond au mouvement brownien fractionnaire de paramètre de Hurst H . En revanche pour $k \geq 2$, le processus Z_H^k n'est plus gaussien, il appartient au k -ème chaos de Wiener. Z_H^2 est connu comme étant le processus de Rosenblatt.

Cas particulier : le processus de Rosenblatt

Le processus de Rosenblatt est un processus non gaussien appartenant au second chaos de Wiener. Il a été largement étudié par M. Taqqu [113] et C. Tudor [120], [30], [116]. C'est un processus d'Hermite $Z_H^k(t)$ et par conséquent il vérifie les mêmes propriétés qui lui sont attribuées. Il est alors défini pour $k = 2$ par

$$Z_H^2(t) = d(H) \int_{\mathbb{R}^2} \int_0^t (s - y_1)_+^{\frac{H}{2}-1} (s - y_2)_+^{\frac{H}{2}-1} ds dB(y_1) dB(y_2),$$

avec

$$d(H) = \frac{\sqrt{2H(2H-1)}}{\beta\left(\frac{H}{2}, H-1\right)}. \quad (1.7)$$

β désigne la fonction bêta de Euler. En particulier, nous pouvons écrire ce processus comme une intégrale stochastique multiple d'ordre 2,

$$Z_H^2(t) = I_2(L_t).$$

Nous appelons L_t le noyau de Rosenblatt. Il est déterminé pour tout $t > 0$ par,

$$L_t(y_1, y_2) := d(H) \int_0^t (s - y_1)_+^{\frac{H}{2}-1} (s - y_2)_+^{\frac{H}{2}-1} ds.$$

La fructueuse rencontre entre Stein et Malliavin

Dans cette partie, nous introduirons un théorème central limite quantitatif basé sur la méthode de Stein-Malliavin pour des variables, vecteurs et matrices aléatoires dont les éléments vivent dans un chaos de Wiener. Cette méthode est la pierre angulaire des travaux de recherches constituant cette thèse.

Les variables aléatoires gaussiennes occupent un rôle majeur dans la théorie des probabilités et des statistiques, puisqu'elles permettent de modéliser le comportement global d'une multitude de petits phénomènes. Par exemple, si nous décidons d'observer la taille des individus d'une population, nous obtiendrions une répartition qui ressemblerait fortement à celle de la loi normale, la célèbre courbe de Gauss, qui a la forme d'une cloche. En particulier, le théorème central limite que nous abrégerons sous la notation, non moins connue, TCL est un théorème fondamental en théorie des probabilités. Découvert dans un premier temps par De Moivre au XVIII-ème siècle pour des variables aléatoires indépendantes suivant une loi de Bernoulli - connu sous le nom du théorème de Moivre-Laplace - c'est au début du XIX-ème siècle que Laplace généralisera et donnera la formulation du TCL que nous connaissons à ce jour. Ce théorème permet d'établir la convergence en loi de la somme d'une suite de variables aléatoires indépendantes et équidistribuées vers la loi normale. Dans la suite, nos variables aléatoires sont toutes définies sur un espace probabilisé approprié $(\Omega, \mathcal{F}, \mathbb{P})$. Nous rappelons ci-dessous un énoncé du TCL.

Théorème 5 (Théorème Central Limite). *Soit $(X_i)_{1 \leq i \leq n}$ une suite de variables aléatoires indépendantes, identiquement distribuées (i.i.d) et de carré intégrable (non constantes). On suppose que $\text{Var}[X_1] > 0$. Notons Z une variable aléatoire centrée telle que $Z \sim \mathcal{N}(0, 1)$ et \bar{X}_n , la moyenne empirique de la suite $(X_i)_{1 \leq i \leq n}$,*

$$\bar{X}_n := \frac{1}{n} \sum_{i=1}^n X_i,$$

alors

$$\sqrt{n} \frac{\bar{X}_n - \mathbb{E}(X_1)}{\sqrt{\text{Var}[X_1]}} \xrightarrow[n \rightarrow \infty]{\text{Loi}} Z.$$

Des généralisations du TCL ont été apportées par Lindeberg et Lyapounov pour des variables

aléatoires indépendantes mais sans hypothèse d'équidistribution. Malheureusement, le TCL ne permet pas de déterminer la vitesse de convergence de la loi de notre suite de variables aléatoires vers la loi gaussienne. Ce n'est qu'en 1941-1942 que Berry et Esséen ont découvert indépendamment une approximation de l'erreur commise entre la loi de nos variables aléatoires et la loi normale. En effet, le théorème de Berry-Esséen permet de donner une borne effective de la vitesse de convergence d'une suite de variables aléatoires i.i.d vers une loi gaussienne standard.

Théorème 6 (Berry-Esséen). *Soit $(X_i)_{1 \leq i \leq n}$ une suite de variables aléatoires i.i.d centrées, telle que $\mathbb{E}[|X_1|^3] < \infty$. On note $Y_n := \sqrt{n} \frac{\bar{X}_n}{\sigma}$, où $\sigma^2 := \mathbb{E}[X_1^2]$ (avec $\sigma > 0$) et $Z \sim \mathcal{N}(0, 1)$. Alors, il existe une constante universelle $C > 0$ telle que pour tout $n \geq 1$,*

$$d_{\text{Kol}}(Y_n, Z) \leq C \frac{\mathbb{E}[|X_1|^3]}{\sigma^3 \sqrt{n}}.$$

d_{Kol} correspond à la distance de Kolmogorov donnée dans la section 2.1.

La question qui peut se poser est alors celle de la qualité d'une telle approximation. En 1972, Stein utilise une autre approche pour majorer la distance entre la loi d'une variable aléatoire X et la gaussienne centrée réduite $\mathcal{N}(0, 1)$. En particulier, cette méthode permet de retrouver les résultats de Berry-Esséen. Dans la suite, nous présenterons la méthode de Stein et nous l'utiliserons en la combinant au calcul de Malliavin afin d'établir un TCL quantitatif pour des variables aléatoires vivant dans un chaos de Wiener.

2.1 Distances

Pour préparer l'introduction à la méthode de Stein, nous définissons ci-dessous une notion de distance entre des lois de probabilités qui sera nécessaire dans la suite de notre travail.

La définition suivante concerne la distance entre les lois de deux variables aléatoires.

Définition 15. *Soit F et G deux variables aléatoires réelles. Soit \mathcal{A} une classe de fonctions telle que si $h \in \mathcal{A}$, $h(F)$ et $h(G) \in L^1(\Omega)$ alors*

$$\forall h \in \mathcal{A}, \quad \mathbb{E}[h(F)] = \mathbb{E}[h(G)] \Rightarrow F \stackrel{\text{loi}}{=} G.$$

La distance entre les lois de F et G telles que $h(F)$ et $h(G) \in L^1(\Omega)$ est définie par

$$d(F, G) = \sup_{h \in \mathcal{A}} |\mathbb{E}h(F) - \mathbb{E}h(G)|, \tag{2.1}$$

1. Lorsque \mathcal{A} est l'ensemble des fonctions lipschitziennes $h : \mathbb{R} \rightarrow \mathbb{R}$ telles que $\|h\|_{\text{Lip}} \leq 1$, où

$$\|h\|_{\text{Lip}} = \sup_{x \neq y, x, y \in \mathbb{R}} \frac{|h(x) - h(y)|}{|x - y|}$$

alors (2.1) donne la distance de Wasserstein (aussi appelée distance de 1-Wasserstein).

2. Quand \mathcal{A} est l'ensemble des fonctions indicatrices $\{\mathbf{1}_{(-\infty, z]}, z \in \mathbb{R}\}$, on obtient la distance de Kolmogorov.
3. Pour $\mathcal{A} = \{\mathbf{1}_B, B \in \mathcal{B}(\mathbb{R})\}$, on a la distance en variation totale.

Nous généralisons la définition précédente, pour des vecteurs aléatoires.

Définition 16. Soit \mathcal{G} une classe de fonctions mesurables. Soient X et Y deux vecteurs aléatoires de \mathbb{R}^d tels que tout $g \in \mathcal{G}$, $g(X), g(Y) \in L^1(\Omega)$. La distance entre les lois de X et Y induite par \mathcal{G} est donnée par :

$$d(X, Y) = \sup_{g \in \mathcal{G}} |\mathbb{E}[g(X)] - \mathbb{E}[g(Y)]|,$$

où la classe de fonctions \mathcal{G} vérifie,

$$\forall g \in \mathcal{G}, \mathbb{E}[g(X)] = \mathbb{E}[g(Y)] \Rightarrow X \stackrel{\text{loi}}{=} Y.$$

En particulier, il existe différentes distances, selon la classe de fonctions \mathcal{G} , à savoir :

1. La distance de 1-Wasserstein (ou Kantorovich-Wasserstein) - que nous appellerons par abus de langage, distance de Wasserstein - entre la loi de X et Y est notée $d_W(X, Y)$. Elle est obtenue en prenant \mathcal{G} , l'ensemble des fonctions $g : \mathbb{R}^d \rightarrow \mathbb{R}$, telles que $\|g\|_{\text{Lip}} \leq 1$, où

$$\|g\|_{\text{Lip}} = \sup_{\substack{x_d \neq y_d \\ x_d, y_d \in \mathbb{R}^d}} \frac{|g(x_d) - g(y_d)|}{\|x_d - y_d\|_{\mathbb{R}^d}},$$

où $\|\cdot\|_{\mathbb{R}^d}$ est la norme euclidienne standard sur \mathbb{R}^d .

2. La distance en variation totale entre la loi de X et la loi de Y . Elle est notée $d_{TV}(X, Y)$ et est obtenue en prenant \mathcal{G} , la collection de fonctions $g : \mathbb{R}^d \rightarrow \mathbb{R}$, du type $g(x_1, \dots, x_d) = \mathbf{1}_B(x_1, \dots, x_d)$, où $B \in \mathcal{B}(\mathbb{R}^d)$. Ainsi

$$d_{TV}(X, Y) = \sup_{B \in \mathcal{B}(\mathbb{R}^d)} |\mathbb{P}[X \in B] - \mathbb{P}[Y \in B]|.$$

3. La distance de Kolmogorov entre la loi de X et Y , notée $d_{\text{Kol}}(X, Y)$ est obtenue en prenant la classe de fonctions $g : \mathbb{R}^d \rightarrow \mathbb{R}$, du type $g(x_1, \dots, x_d) = \mathbf{1}_{(-\infty, z_1]}(x_1) \cdots \mathbf{1}_{(-\infty, z_d]}(x_d)$, où $z_1, \dots, z_d \in \mathbb{R}$. Cette distance est donnée par la formule suivante :

$$d_{\text{Kol}}(X, Y) = \sup_{z_1, \dots, z_d \in \mathbb{R}^d} |\mathbb{P}[X \in (-\infty, z_1] \times \dots \times (-\infty, z_d)] - \mathbb{P}[Y \in (-\infty, z_1] \times \dots \times (-\infty, z_d)]|.$$

Nous serons également amenés à évoquer une autre distance, la distance d_2 .

Définition 17. Si X et Y sont deux vecteurs aléatoires de \mathbb{R}^d , la distance d_2 entre leur loi de probabilité est donnée par :

$$d_2(X, Y) = \sup_{\|h''\|_{\infty} \leq 1} |\mathbb{E}h(X) - \mathbb{E}h(Y)|,$$

où

$$\|h''\|_{\infty} = \sup_{x \in \mathbb{R}^d} \sup_{1 \leq i, j \leq d} \left| \frac{\partial^2 h}{\partial x_i \partial x_j}(x) \right|.$$

Comme nous évoquerons largement au cours de cette thèse la distance entre les lois de deux matrices aléatoires, nous proposons de rappeler la définition de la norme d'opérateur et la norme de Hilbert-Schmidt.

Définition 18. 1. Le produit scalaire de Hilbert-Schmidt et la norme Hilbert-Schmidt définis sur la classe des matrices carrées de taille d sont notés respectivement $\langle \cdot, \cdot \rangle_{\text{HS}}$ et $\|\cdot\|_{\text{HS}}$. Il est défini comme suit : pour toutes matrices A et B , $\langle A, B \rangle_{\text{HS}} := \text{Tr}(AB^T)$ et $\|A\|_{\text{HS}} = \sqrt{\langle A, A \rangle_{\text{HS}}}$, où $\text{Tr}(\cdot)$ est l'opérateur trace usuel.

2. L'opérateur norme d'une matrice carrée A de taille d est donnée $\|A\|_{\text{op}} := \sup_{\|x\|_{\mathbb{R}^d}=1} \|Ax\|_{\mathbb{R}^d}$.

Définition 19. Soient \mathcal{X}, \mathcal{Y} , deux matrices aléatoires de $\mathcal{M}_n(\mathbb{R})$, avec $n \geq 1$. On note $d_W(\mathcal{X}, \mathcal{Y})$, la distance de Wasserstein entre les lois de \mathcal{X} et \mathcal{Y} .

$$d_W(\mathcal{X}, \mathcal{Y}) = \sup_{\|g\|_{\text{Lip}} \leq 1} |\mathbb{E}g(\mathcal{X}) - \mathbb{E}g(\mathcal{Y})|,$$

où la norme lipschitzienne $\|g\|_{\text{Lip}}$ de $g : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}$ est définie par

$$\|g\|_{\text{Lip}} = \sup_{A \neq B, A, B \in \mathcal{M}_n(\mathbb{R})} \frac{|g(A) - g(B)|}{\|A - B\|_{\text{HS}}}.$$

Dans la partie suivante, nous proposons de présenter une méthode permettant de quantifier l'erreur commise dans l'approximation de la loi d'une variable aléatoire par une loi gaussienne standard, qui ne sera pas basée sur l'étude des fonctions caractéristiques, mais sur une caractérisation de la loi normale centrée réduite, à l'aide de l'espérance et d'une fonction absolument continue. Nous présenterons ensuite la méthode de Stein combinée au calcul de Malliavin afin de déterminer des bornes effectives entre nos variables aléatoires et la loi normale.

2.2 Méthode de Stein unidimensionnelle sur l'espace gaussien

L'objectif de cette partie est de rappeler les principales idées de la méthode de Stein. Pour un exposé plus général, nous vous renvoyons au livre de Chen, Goldstein et Shao [25], au travail fondateur de Stein [110] ou au chapitre 3 du livre de Nourdin et Peccati [93].

Nous rappelons que nos variables aléatoires sont toutes définies sur un espace probabilisé approprié $(\Omega, \mathcal{F}, \mathbb{P})$.

La méthode de Stein permet d'estimer la proximité entre la loi d'une variable aléatoire et la loi gaussienne standard $Z \sim \mathcal{N}(0, 1)$. L'estimation de cette proximité entre deux lois de variables aléatoires repose sur la notion de distance pour laquelle nous sommes capables de quantifier l'écart entre ces deux lois. La distance entre deux variables aléatoires X et Y notée $d(X, Y)$ correspond à l'une des distances suivantes d_{Kol} , d_{TV} , d_W respectivement la distance de Kolmogorov, la distance en variation totale et la distance de Wasserstein associée à la classe de fonctions \mathcal{G} respective.

Le point de départ de la méthode de Stein est de caractériser la loi normale centrée réduite par un opérateur différentiel du premier ordre. Avant de donner une définition de cet opérateur, remarquons le fait suivant, qui caractérise les variables aléatoires gaussiennes.

Lemme 1 (Caractérisation de la loi gaussienne standard). Soient $Z \sim \mathcal{N}(0, 1)$ et $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction absolument continue telle que $\mathbb{E}|f'(Z)| < \infty$. Alors $\mathbb{E}|Zf(Z)| < \infty$ et l'égalité suivante est satisfaite,

$$\mathbb{E}[Zf(Z)] = \int_{\mathbb{R}} xf(x) \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = \int_{\mathbb{R}} f'(x) \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = \mathbb{E}[f'(Z)].$$

Dans le cas où f est dérivable bornée, il s'agit d'une simple intégration par parties. Ce résultat crucial nous montre donc que si $Z \sim \mathcal{N}(0, 1)$ alors pour une certaine fonction f suffisamment régulière, $\mathbb{E}[f'(Z)] = \mathbb{E}[Zf(Z)]$. L'implication réciproque est également vraie puisque la loi gaussienne est entièrement déterminée par ses moments. Soit N une variable aléatoire et prenons pour tout $n \geq 0$, $f(x) = x^n$, alors $\mathbb{E}[N^{n+1}] = n\mathbb{E}[N^{n-1}]$, caractérisant la relation de récurrence entre les moments de la gaussienne. Cette équivalence est formulée par le lemme de Stein que nous énonçons ci-dessous. Ainsi toute variable aléatoire qui vérifie (2.2) est une variable aléatoire gaussienne et réciproquement. Nous pouvons dire que l'équation (2.2) caractérise entièrement les variables aléatoires gaussiennes et nous remarquons que cette méthode ne fait pas appel à l'étude des fonctions caractéristiques. Par conséquent, nous notons \mathcal{L} l'opérateur différentiel du premier ordre défini pour tout $x \in \mathbb{R}$ par $\mathcal{L}f(x) = f'(x) - xf(x)$.

Lemme 2 (Lemme de Stein). *Soit N une variable aléatoire. $N \stackrel{\text{Loi}}{=} Z \sim \mathcal{N}(0, 1)$ si et seulement si pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ absolument continue telle que $\mathbb{E}[f'(Z)] < \infty$,*

$$\mathbb{E}[\mathcal{L}f(N)] = 0. \quad (2.2)$$

L'idée de Stein repose sur le fait que si nous disposons d'une variable aléatoire N dont on ne connaît pas la loi :

- *Pourrions-nous dire, en observant que la quantité $\mathbb{E}[\mathcal{L}f(N)]$ est proche de 0 que la loi de la variable aléatoire N est très proche d'être gaussienne ?*
- *Et dans ce cas, à quel point est-elle proche d'être gaussienne ?*

Finalement, nous allons voir que la méthode de Stein permet de donner une borne quantitative de l'erreur commise dans l'approximation de la loi d'une variable aléatoire par la loi normale en subsituant la distance entre les distributions de celles-ci par la caractérisation (2.2).

Définition 20 (Equation de Stein). *Soit $Z \sim \mathcal{N}(0, 1)$ et $h : \mathbb{R} \rightarrow \mathbb{R}$ une fonction borélienne telle que $h(Z) \in L^1(\Omega)$. L'équation de Stein associée à h est donnée par :*

$$\mathcal{L}f(x) = h(x) - \mathbb{E}[h(Z)], \quad (2.3)$$

où f vérifie les conditions suivantes selon h et telle qu'il existe une version de sa dérivée f' qui satisfasse l'équation de Stein (2.3).

1. *Si $h : \mathbb{R} \rightarrow [0, 1]$ est une fonction borélienne, alors (2.3) admet une solution f absolument continue telle que $\|f\|_\infty \leq \sqrt{\frac{\pi}{2}}$ et $\|f'\|_\infty \leq 2$.*
2. *Si $h(x) = \mathbf{1}_{(-\infty, z]}(x)$, $z \in \mathbb{R}$, alors (2.3) admet une solution f qui est bornée par $\frac{\sqrt{2\pi}}{4}$, absolument continue, différentiable sauf pour un nombre fini de points et telle que $\|f'\|_\infty \leq 1$.*
3. *Si $h : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction K -Lipschitzienne, où $K > 0$, alors (2.3) admet pour solution f de classe \mathcal{C}^1 et telle que $\|f'\|_\infty \leq \sqrt{\frac{2}{\pi}}K$.*

En fin de compte, si nous remplaçons x par une variable aléatoire N et nous prenons l'espérance de part et d'autre de (2.3), puis le supremum sur la classe de fonction associée à la distance que nous souhaitons considérer, l'équation de Stein de la définition 20 nous fournit une égalité entre notre opérateur différentiel de premier ordre et la distance entre les lois de deux variables aléatoires. Autrement dit :

Regarder la distance entre la loi d'une variable aléatoire N et la loi normale revient à s'intéresser au supremum de la quantité $\mathbb{E}[\mathcal{L}f(N)]$ sur une classe de fonctions f solution de (2.3).

Nous rappelons que d correspond à l'une des distances suivantes d_{Kol} , d_{TV} , d_{W} respectivement la distance de Kolmogorov, la distance en variation totale et la distance de Wasserstein. Notons \mathcal{F}_d la classe de Stein associée à la distance d , c'est-à-dire la classe des fonctions solutions de l'équation de Stein associée à h . Par exemple, la classe de Stein associée à une loi normale est définie par :

- Pour la distance en variation totale, que l'on notera \mathcal{F}_{TV} , la classe de Stein correspond à $\mathcal{F}_{\text{TV}} = \{f : \|f\|_{\infty} \leq \sqrt{\frac{\pi}{2}}, \|f'\|_{\infty} \leq 2\}$.
- La classe de Stein adaptée pour distance de Kolmogorov est

$$\mathcal{F}_{\text{Kol}} = \left\{ f : \|f\|_{\infty} \leq \frac{\sqrt{2\pi}}{4}, \|f'\|_{\infty} \leq 1 \right\}.$$

- Pour la distance de Wasserstein, la classe de Stein est donnée par

$$\mathcal{F}_{\text{W}} = \left\{ f : \mathbb{R} \rightarrow \mathbb{R} \in \mathcal{C}^1 : \|f'\|_{\infty} \leq \sqrt{\frac{2}{\pi}} \right\}.$$

Dans ce cas, et grâce à l'équation (2.3), on a réduit le problème initial, qui était d'approximer la loi d'une variable aléatoire N par la loi normale, par une expression plus simple, à savoir, l'espérance d'une certaine fonctionnelle de notre variable aléatoire N . En effet,

$$\begin{aligned} d(N, Z) &= \sup_{h \in \mathcal{G}} |\mathbb{E}[h(N)] - \mathbb{E}[h(Z)]| \\ &\leq \sup_{f \in \mathcal{F}_d} |\mathbb{E}[\mathcal{L}f(N)]| \\ &= \sup_{f \in \mathcal{F}_d} |\mathbb{E}[f'(N)] - \mathbb{E}[Nf(N)]|, \end{aligned} \tag{2.4}$$

où \mathcal{G} est une classe de fonctions mesurables.

Remarque 9. 1. Nous avons appliqué la méthode de Stein sur l'espace gaussien, cependant, il existe une méthode de Stein pour d'autres lois, en particulier sur l'espace de Poisson appelée méthode de Chen-Stein (1975) que nous pouvons retrouver dans le livre [25], mais aussi pour l'approximation d'une variable aléatoire par une loi Gamma dans [84].

2. Pour établir ces méthodes de Stein, la technique est similaire à celle présentée précédemment ;

(a) Il faut convertir le problème d'approximation de la loi d'une variable aléatoire par la loi cible, à l'aide d'une caractérisation. En général, on caractérise la loi cible par un opérateur différentiel. Soit μ une mesure de probabilité sur \mathbb{R} et prenons X une variable aléatoire réelle. Dans ce cas, X suit la loi μ si et seulement si pour tout $f \in \mathcal{F}$,

$$\mathbb{E}[\mathcal{A}_{\mu}f(X)] = 0.$$

On dit que \mathcal{A}_{μ} est un opérateur caractéristique de μ agissant sur une certaine classe de fonctions \mathcal{F} .

— Pour la loi Gamma $\Gamma(r, \lambda)$, l'opérateur est défini par

$$\mathcal{A}_{\mu}f(x) := xf'(x) + (r - \lambda x)f(x),$$

pour $f : \mathbb{R} \rightarrow \mathbb{R}$.

— Pour la loi de Poisson $\mathcal{P}(\lambda)$, cet opérateur n'est pas différentiel. Il est donné par

$$\mathcal{A}_\mu f(x) := xf(x) - \lambda f(x+1),$$

pour $f : \mathbb{N} \rightarrow \mathbb{N}$.

- (b) Le problème d'approximation se réduit alors à une équation différentielle qu'il faut résoudre pour trouver la classe d'appartenance des fonctions solutions de l'équation Stein.
- (c) Finalement, le problème initial, se réduit au contrôle de l'espérance d'une certaine fonctionnelle de la variable aléatoire que l'on cherche à estimer.

Nous avons vu dans cette partie que nous pouvions quantifier l'erreur commise dans l'approximation de la loi d'une variable aléatoire et la loi normale grâce à la borne (2.4). En particulier, l'expression de cette borne se calcule assez bien dans le cas où les variables aléatoires vivent dans un chaos de Wiener, cela permet d'établir un théorème central limite quantitatif pour les intégrales stochastiques multiples. Dans la suite, nous combinons la méthode de Stein au calcul de Malliavin afin d'énoncer le célèbre théorème du quatrième moment.

2.3 De la méthode de Stein au calcul de Malliavin : le théorème du quatrième moment

Nous considérons ici, un espace de Hilbert \mathcal{H} auquel on associe un processus isonormal $\{X(h), h \in \mathcal{H}\}$ défini sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et tel que pour tout $h, g \in \mathcal{H}$, $\mathbb{E}[X(h)X(g)] = \langle h, g \rangle_{\mathcal{H}}$. Nous voulons établir une présentation du théorème du quatrième moment qui constitue un théorème central limite, sans hypothèse d'indépendance et pour une suite de variables aléatoires appartenant à un chaos de Wiener d'ordre $q \geq 2$. Plus précisément, il nous permet de contrôler la convergence en loi de notre suite de variables aléatoires $F_n := I_q(f_n)$, vers une loi gaussienne à travers des bornes effectives issues de la méthode de Stein et de donner un critère de convergence vers une telle loi. L'objectif de cette partie n'étant pas de faire un exposé détaillé mais simplement d'apporter les principes techniques et conceptuels. Une description complète des résultats se trouve dans le livre fondateur de Nourdin et Peccati [93].

Intéressons-nous tout d'abord aux variables aléatoires de carré intégrable et dérivables au sens de Malliavin. En effet, il est pertinent de se pencher sur le cas de ces variables aléatoires appartenant à l'espace $\mathbb{D}^{1,2}$ (espace introduit dans la section 1.1), puisque nous disposons, pour celles-ci, d'une intégration par parties qui permet de calculer la borne de l'équation (2.4) à l'aide des opérateurs de Malliavin.

Lemme 3 (Nourdin et Peccati (2007)). *Soient F et $G \in \mathbb{D}^{1,2}$ et $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 dont la dérivée est bornée. Alors nous avons,*

$$\mathbb{E}[Fg(G)] = \mathbb{E}[F]\mathbb{E}[g(G)] + \mathbb{E}[g'(G)\langle DG, -DL^{-1}F \rangle_{\mathcal{H}}]. \quad (2.5)$$

Nous renvoyons au théorème 2.9.1 du livre [93] pour une preuve du résultat. En utilisant la borne (2.4) obtenue par la méthode de Stein dans la partie précédente, et en prenant $G = F = N$ dans le lemme 3, où $N \in \mathbb{D}^{1,2}$ est une variable aléatoire centrée et $Z \sim \mathcal{N}(0, 1)$, nous obtenons

les inégalités suivantes :

$$\begin{aligned}
 d(N, Z) &\leq \sup_{f \in \mathcal{F}_d} [\mathbb{E}[\mathcal{L}f(N)]] \\
 &= \sup_{f \in \mathcal{F}_d} [\mathbb{E}[f'(N)] - \mathbb{E}[Nf(N)]] \\
 &= \sup_{f \in \mathcal{F}_d} [\mathbb{E}[f'(N)] - \mathbb{E}[f'(N)\langle DN, -DL^{-1}N \rangle_{\mathcal{H}}]] \\
 &= \sup_{f \in \mathcal{F}_d} [\mathbb{E}[f'(N)](1 - \mathbb{E}[\langle DN, -DL^{-1}N \rangle_{\mathcal{H}}])] \\
 &\leq \sup_{f \in \mathcal{F}_d} \|f'\|_{\infty} \mathbb{E}[1 - \mathbb{E}[\langle DN, -DL^{-1}N \rangle_{\mathcal{H}}]]. \tag{2.6}
 \end{aligned}$$

Le théorème que nous énonçons ci-dessous résulte simplement de l'utilisation de la nouvelle borne (2.6).

Théorème 7. *Soit $Z \sim \mathcal{N}(0, 1)$ et $N \in \mathbb{D}^{1,2}$ une variable aléatoire centrée telle que $\mathbb{E}[N^2] = 1$. On suppose que la loi de N est absolument continue par rapport à la mesure de Lebesgue. Alors on a*

$$d(N, Z) \leq K_d \mathbb{E}[|1 - \langle DN, -DL^{-1}N \rangle_{\mathcal{H}}|]$$

où K_d est une constante dépendant de la distance considérée.

Dans ce que nous avons fait jusqu'à présent, nous avons toujours considéré des variables aléatoires centrées, de variance égale à 1 et établi des convergences vers des variables aléatoires gaussiennes standards. Un résultat pertinent serait aussi de l'avoir directement pour des variables aléatoires centrées dont la variance est plus générale, $\sigma^2 > 0$. En fait, dans la pratique, il est rare d'avoir une variance égale à un donc nous citerons une "variante" du théorème 7 pour plus de généralité. Le théorème suivant découle du fait élémentaire que si $Y \sim \mathcal{N}(0, \sigma^2)$ alors $\frac{Y}{\sigma} \sim \mathcal{N}(0, 1)$, avec $\sigma > 0$. Il suffit ensuite de remplacer Z par $\frac{Y}{\sigma}$ dans l'équation de Stein (2.3) puis utiliser la même méthode de bornitude de la distance pour trouver (2.7).

Théorème 8. *Soit $Z \sim \mathcal{N}(0, \sigma^2)$ et $N \in \mathbb{D}^{1,2}$ une variable aléatoire centrée telle que $\mathbb{E}[N^2] = \sigma^2 > 0$. On suppose que la loi de N est absolument continue par rapport à la mesure de Lebesgue. Alors on a*

$$d(N, Z) \leq K_d \mathbb{E}[|\sigma^2 - \langle DN, -DL^{-1}N \rangle_{\mathcal{H}}|] \tag{2.7}$$

où K_d est une constante dépendant de la distance considérée et de la variance σ^2 .

En particulier, il se trouve que la borne donnée dans le théorème 8 ci-dessus peut être calculée explicitement lorsque N est une intégrale stochastique multiple $N = I_q(f)$. Cela est dû au fait qu'une intégrale multiple se comporte comme un vecteur propre pour l'inverse de l'opérateur d'Ornstein–Uhlenbeck L^{-1} associé à la valeur propre $-\frac{1}{q}$. On a donc $L^{-1}N = -\frac{1}{q}N$. Ainsi nous avons que

$$\langle DN, -DL^{-1}N \rangle_{\mathcal{H}} = \frac{1}{q} \|DN\|_{\mathcal{H}}^2.$$

Nous sommes désormais en mesure d'énoncer le théorème quantitatif du quatrième moment qui repose sur le résultat établi par le théorème 8, dans le cas où l'on considère une suite de variables aléatoires appartenant à un chaos d'ordre $q \geq 2$.

Théorème 9 (Théorème quantitatif du quatrième moment). *Soit un entier $q \geq 1$ fixé. Soit $(F_n)_{n \geq 1} = (I_q(f_n))_{n \geq 1}$, avec $f_n \in \mathcal{H}^{\odot q}$ une suite d'intégrales multiples appartenant à un chaos*

de Wiener telle que

$$\mathbb{E}[F_n^2] \xrightarrow[n \rightarrow \infty]{} \sigma^2.$$

Alors $F_n \xrightarrow[n \rightarrow \infty]{\text{Loi}} Z \sim \mathcal{N}(0, \sigma^2)$ si et seulement si

$$\|DF_n\|_{\mathcal{H}}^2 \xrightarrow[n \rightarrow \infty]{L^2(\Omega)} q\sigma^2.$$

De plus, nous avons la borne effective suivante

$$d(F_n, Z) \leq C \left[\sqrt{\text{Var}(\|DF_n\|_{\mathcal{H}}^2)} + |\mathbb{E}[\|DF_n\|_{\mathcal{H}}^2] - q\sigma^2| \right]. \quad (2.8)$$

C est une constante strictement positive.

C'est un résultat assez récent puisqu'il date du début des années 2000. Nous donnons ensuite un énoncé de sa forme plus habituelle.

C'est en 2005, que Nualart et Peccati dans [99], donnent les premiers critères pour qu'une suite d'intégrales stochastiques multiples d'ordre $q \geq 2$ convergent en loi vers une gaussienne. Cela correspond au trois premiers points du théorème suivant qui est le fameux théorème du quatrième moment. En 2007, Nualart et Ortiz-Latorre dans [101], apportent une assertion supplémentaire et équivalente aux autres, qui correspond au deux derniers points ci-dessous.

Théorème 10 (Théorème du quatrième moment). *Pour $n \geq 2$, soit $F_n = I_q(f_n)$, une suite de variables aléatoires appartenant au q -ème chaos de Wiener, avec $q \geq 2$. Supposons que $\mathbb{E}[F_n^2] \xrightarrow[n \rightarrow \infty]{} \sigma^2 > 0$. Alors les assertions suivantes sont équivalentes :*

1. F_n converge en loi vers $\mathcal{N}(0, \sigma^2)$ quand $n \rightarrow \infty$.
2. $\mathbb{E}[F_n^4] \xrightarrow[n \rightarrow \infty]{} 3\sigma^4 = \mathbb{E}[Z^4]$.
3. Pour tout $r = 1, \dots, q-1$, $\|f_n \otimes_r f_n\|_{\mathcal{H}^{\otimes(2q-2r)}} \xrightarrow[n \rightarrow \infty]{} 0$.
4. $\text{Var}(\|DF_n\|_{\mathcal{H}}^2) \xrightarrow[n \rightarrow \infty]{} 0$.
5. $\|DF_n\|_{\mathcal{H}}^2$ converge dans $L^2(\Omega)$ vers $q\sigma^2$ quand $n \rightarrow \infty$.

Remarque 10. Les deux dernières conditions du théorème 10 proviennent de l'article [101] de Nualart et Ortiz-Latorre, et permettent d'établir un autre critère de non-gaussianité des chaos. En effet, supposons que F appartienne au q -ème chaos de Wiener et que $F \sim \mathcal{N}(0, 1)$ ou que $\mathbb{E}[\|DF\|_{\mathcal{H}}^4] = q^2$. Ce qui signifie que $\text{Var}(\|DF\|_{\mathcal{H}}^2) = 0$ alors $\|DF\|_{\mathcal{H}}^2$ est constante. Par conséquent F est nulle ou dans le premier chaos.

Théorème 11 (Non-gaussianité des Chaos). *Soient un entier $q \geq 2$ fixé et F une variable aléatoire appartenant au q -ème chaos de Wiener telle que $\mathbb{E}[F^2] = 1$. Alors la loi de F ne peut pas être gaussienne et $\mathbb{E}[\|DF\|_{\mathcal{H}}^4] \neq q^2$.*

Nous avons vu jusqu'à présent, la méthode de Stein-Malliavin dans le cas unidimensionnel. Nous voudrions désormais développer cette méthode dans le cas multidimensionnel afin d'être capables d'énoncer un TCL pour des vecteurs aléatoires puis des matrices aléatoires. Par conséquent, nous présenterons dans la partie suivante la méthode de Stein multidimensionnelle.

2.4 Méthode de Stein multidimensionnelle et TCL pour les matrices aléatoires

Dans cette nouvelle partie, nous introduirons une généralisation du théorème du quatrième moment, c'est-à-dire un théorème central limite multivarié pour des vecteurs aléatoires dont les éléments appartiennent à un chaos de Wiener. L'analyse asymptotique des matrices aléatoires par la méthode de Stein-Malliavin est le socle sur lequel repose les résultats des articles de cette thèse, c'est donc naturellement que nous sommes amenés à évoquer ce qui suit. Afin de parvenir à l'énoncé d'un tel TCL pour les vecteurs aléatoires et a fortiori pour les matrices aléatoires, nous évoquerons dans un premier temps la méthode de Stein multidimensionnelle. Cela nous permettra de comprendre les heuristiques liées à la méthode de Stein-Malliavin multivariée.

2.4.1 Méthode de Stein multidimensionnelle sur l'espace gaussien

Le principe de la méthode de Stein multidimensionnelle est le même que dans le cas unidimensionnel, non plus pour des variables aléatoires mais des vecteurs aléatoires. En effet, l'objectif est de trouver une majoration de la distance entre la loi d'un vecteur aléatoire et celle d'un vecteur gaussien. Pour cela, il est nécessaire de disposer d'une caractérisation pour la loi gaussienne multivariée, de la même manière que dans le cas univarié. La distance utilisée sera la distance de 1-Wasserstein pour les vecteurs aléatoires à valeurs dans \mathbb{R}^d .

En particulier, nous donnons dans le lemme 4 une généralisation du lemme 1. La caractérisation de loi gaussienne standard multivariée est toujours basée sur une formule d'intégration par parties pour une certaine classe de fonctions.

Lemme 4 (Caractérisation de la loi gaussienne multivariée). *Soit $X \sim \mathcal{N}_d(0, I_d)$ un vecteur gaussien centré. On considère $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction différentiable bornée telle que*

$$\mathbb{E}|\partial_{x_k} f(X_1, \dots, X_d)| < \infty,$$

alors on a pour $k = 1, \dots, d$,

$$\int_{\mathbb{R}^d} x_k f(x_1, \dots, x_d) d\mathbb{P}_{X_1, \dots, X_d}(x_1, \dots, x_d) = \int_{\mathbb{R}^d} \partial_{x_k} f(x_1, \dots, x_d) d\mathbb{P}_{X_1, \dots, X_d}(x_1, \dots, x_d), \quad (2.9)$$

où $\partial_{x_k} f$ correspond à la dérivée partielle de f par rapport à la k -ème variable.

Preuve : En utilisant une intégration par parties et puisque $\mathbb{E}|\partial_{x_k} f(X_1, \dots, X_d)| < \infty$, on dispose des égalités suivantes.

$$\begin{aligned}
 & \int_{\mathbb{R}^d} x_k f(x_1, \dots, x_d) d\mathbb{P}_{X_1, \dots, X_d}(x_1, \dots, x_d) \\
 = & \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} x_k f(x_1, \dots, x_d) \exp\left(-\frac{1}{2} \sum_{l=1}^d x_l^2\right) dx_1 \dots dx_d \\
 = & \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} f(x_1, \dots, x_d) \exp\left(-\frac{1}{2} \sum_{l=1, l \neq k}^d x_l^2\right) \partial_{x_k} \left[-\exp\left(-\frac{1}{2} x_k^2\right)\right] dx_1 \dots dx_d \\
 = & \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^{d-1}} dx_1 \dots dx_{k-1} dx_{k+1} \dots dx_d \exp\left(-\frac{1}{2} \sum_{l=1, l \neq k}^d x_l^2\right) \\
 & \int_{\mathbb{R}} dx_k f(x_1, \dots, x_d) \partial_{x_k} \left[-\exp\left(-\frac{1}{2} x_k^2\right)\right] \\
 = & \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^{d-1}} dx_1 \dots dx_{k-1} dx_{k+1} \dots dx_d \exp\left(-\frac{1}{2} \sum_{l=1, l \neq k}^d x_l^2\right) \\
 & \int_{\mathbb{R}} dx_k \exp\left(-\frac{1}{2} x_k^2\right) \partial_{x_k} f(x_1, \dots, x_d) \\
 = & \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2} \sum_{l=1}^d x_l^2\right) \partial_{x_k} f(x_1, \dots, x_d) dx_1 \dots dx_d \\
 = & \int_{\mathbb{R}^d} \partial_{x_k} f(x_1, \dots, x_d) d\mathbb{P}_{X_1, \dots, X_d}(x_1, \dots, x_d).
 \end{aligned}$$

□

Ainsi en réécrivant (2.9), on peut dire que si $X := (X_1, \dots, X_d)$ est un vecteur gaussien standard de \mathbb{R}^d alors pour f donnée par le lemme 4,

$$\forall k, \mathbb{E}[X_k f(X_1, \dots, X_d)] = \mathbb{E}[\partial_{x_k} f(X_1, \dots, X_d)].$$

En réalité ce résultat nous fournit une caractérisation pour chacune des composantes du vecteur gaussien mais nous voudrions une caractérisation plus générale, qui soit vraie pour une combinaison linéaire des éléments du vecteur aléatoire. Nous noterons par \mathcal{C}_b^2 , l'ensemble des fonctions de classe \mathcal{C}^2 dont les dérivées premières et secondes sont bornées.

Proposition 6. Soient $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction continue bornée de classe \mathcal{C}_b^2 et X un vecteur gaussien standard de $\mathbb{R}^d \sim \mathcal{N}_d(\mathbf{0}, I_d)$ alors,

$$\mathbb{E}[\langle X, \nabla f(X) \rangle_{\mathbb{R}^d}] = \mathbb{E}[\Delta f(X)]. \quad (2.10)$$

Preuve : Nous pouvons dire en utilisant le lemme 4 que

$$\begin{aligned}
 \mathbb{E}[\langle X, \nabla f(X) \rangle_{\mathbb{R}^d}] &= \sum_{i=1}^d \mathbb{E}\left[X_i \frac{\partial f}{\partial x_i}(X_1, \dots, X_d)\right] \\
 &= \sum_{i=1}^d \mathbb{E}\left[\frac{\partial^2 f}{\partial x_i^2}(X_1, \dots, X_d)\right] \\
 &= \mathbb{E}[\Delta f(X)].
 \end{aligned}$$

□

Nous allons voir qu'un vecteur aléatoire centré de \mathbb{R}^d vérifiant uniquement la relation (2.10) est gaussien. Pour cette raison, nous allons repartir de la définition de l'égalité en loi pour deux vecteurs aléatoires de \mathbb{R}^d .

Définition 21. Soient X et Y deux vecteurs aléatoires de \mathbb{R}^d . On dit que X a même loi que Y si pour toute fonction $h : \mathbb{R}^d \rightarrow \mathbb{R}$ continue et bornée,

$$\mathbb{E}[h(X)] = \mathbb{E}[h(Y)]. \quad (2.11)$$

Afin de pouvoir donner une caractérisation de la loi gaussienne multivariée comme nous l'avons fait dans le cas unidimensionnel en énonçant le lemme de Stein, nous allons réécrire l'équation (2.11) d'une autre manière. Pour cela, soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction continue bornée de classe \mathcal{C}_b^2 , nous noterons $(P_t)_{t \geq 0}$ le semi-groupe d'Ornstein–Uhlenbeck associé à la fonction $f \in L^q(\gamma)$ pour $q \geq 1$, où γ est la mesure gaussienne standard, et défini selon la formule suivante pour Z un vecteur gaussien standard de \mathbb{R}^d par

$$P_t f(x) = \mathbb{E} \left[f \left(\exp(-t)x + \sqrt{1 - \exp(-2t)}Z \right) \right].$$

Nous voulons montrer que si X est un vecteur aléatoire de \mathbb{R}^d dont on ne connaît pas la loi et vérifiant l'égalité (2.10) alors il est gaussien. Nous noterons $P_\infty f := \lim_{t \rightarrow \infty} P_t f$. Ainsi nous pouvons dire que :

$$\begin{aligned} & \mathbb{E}[f(X)] - \mathbb{E}[f(Z)] \\ &= P_0 f(X) - P_\infty f(X) \\ &= - \int_0^\infty \frac{\partial}{\partial t} P_t f(X) dt \\ &= \int_0^\infty e^{-t} \mathbb{E}[\langle \nabla f(e^{-t}X + \sqrt{1 - e^{-2t}}Z), X \rangle_{\mathbb{R}^d}] \\ & \quad - \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} \mathbb{E}[\langle \nabla f(e^{-t}X + \sqrt{1 - e^{-2t}}Z), Z \rangle_{\mathbb{R}^d}] dt \\ &= \int_0^\infty e^{-t} \mathbb{E}[\langle \nabla f(e^{-t}X + \sqrt{1 - e^{-2t}}Z), X \rangle_{\mathbb{R}^d}] \\ & \quad - e^{-2t} \mathbb{E}[\Delta f(e^{-t}X + \sqrt{1 - e^{-2t}}Z)] dt \end{aligned} \quad (2.12)$$

L'égalité (2.12) découle du fait que Z est un vecteur aléatoire gaussien standard auquel on applique le résultat (2.9) du lemme 4. En fait, comme on a supposé que X vérifiait (2.10), l'égalité (2.12) devient,

$$\int_0^\infty e^{-2t} \mathbb{E}[\Delta f(e^{-t}X + \sqrt{1 - e^{-2t}}Z)] - e^{-2t} \mathbb{E}[\Delta f(e^{-t}X + \sqrt{1 - e^{-2t}}Z)] dt = 0.$$

Ainsi,

$$\mathbb{E}[f(X)] = \mathbb{E}[f(Z)],$$

et par conséquent X a même loi que Z , c'est un vecteur gaussien standard de \mathbb{R}^d . Nous désignons l'opérateur L tel que,

$$Lf(X) := \Delta f(X) - \langle X, \nabla f(X) \rangle_{\mathbb{R}^d}.$$

Remarque 11. *L'opérateur L peut être vu comme le générateur infinitésimal du semi-groupe d'Ornstein–Uhlenbeck, déterminé par la définition 8.*

Nous pouvons dire que X et $Z \sim \mathcal{N}(0, I_d)$ ont même loi si et seulement si $\mathbb{E}[Lf(X)] = 0$.

C'est ce que nous énonçons dans le lemme 5, qui généralise le lemme 2 de Stein pour des vecteurs aléatoires gaussiens standards.

Lemme 5 (Lemme de Stein multidimensionnel). *Soit $X := (X_1, \dots, X_d)$ un vecteur aléatoire de \mathbb{R}^d . X suit une loi gaussienne standard multivariée, et on note $X \stackrel{\text{loi}}{=} Z \sim \mathcal{N}_d(0, I_d)$ si et seulement si pour toute fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$ de classe \mathcal{C}_b^2 ,*

$$\mathbb{E}[Lf(X)] = 0.$$

Nous pouvons alors nous demander :

Si $X := (X_1, \dots, X_d)$ est un vecteur aléatoire centré à valeurs dans \mathbb{R}^d dont les éléments ont une variance égale à 1 et nous observons que la quantité $\mathbb{E}[\langle X, \nabla f(X) \rangle_{\mathbb{R}^d} - \Delta f(X)]$ est très proche de 0 alors pouvons-nous dire que X est proche d'être un vecteur gaussien standard ?

Nous chercherons notamment à quantifier l'erreur commise dans l'approximation de la loi d'un vecteur aléatoire par la loi gaussienne multivariée.

Définition 22 (Equation de Stein multidimensionnelle). *Soit $Z \sim \mathcal{N}_d(0, I_d)$ un vecteur gaussien standard de \mathbb{R}^d . Soit $h : \mathbb{R}^d \rightarrow \mathbb{R}$ tel que $\mathbb{E}|h(Z)| < \infty$. L'équation de Stein associée à h est donnée par l'équation différentielle partielle suivante :*

$$Lf(x) = h(x) - \mathbb{E}[h(Z)]. \quad (2.13)$$

Une fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$ de classe \mathcal{C}_b^2 satisfaisant l'équation (2.13) pour tout $x \in \mathbb{R}^d$ est solution de (2.13).

Le résultat reste vrai dans le cas où Z est un vecteur gaussien de \mathbb{R}^d dont la matrice de covariance est C . Le lemme suivant, que nous pouvons retrouver dans l'article fondateur de la méthode de Stein-Malliavin multidimensionnel [96] nous fournit une équation de Stein dont les composantes du vecteur aléatoire sont corrélées entre elles. Il généralise le lemme 2 de l'article [24] de Chatterjee et Meckes.

Lemme 6 (I.Nourdin, G.Peccati et A.Réveillac (2008)). *1. Soit un entier $d \geq 2$ et soit $C = \{C_{ij} : i, j = 1, \dots, d\}$, une matrice réelle, symétrique, définie positive de taille d . Soit $N := (N_1, \dots, N_d)$ un vecteur aléatoire à valeurs dans \mathbb{R}^d . Alors $N \sim \mathcal{N}_d(0, C)$ si et seulement si, pour toute fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 telle que $\mathbb{E}|\langle C, \text{Hess}f(N) \rangle_{\text{HS}}| + \mathbb{E}|\langle N, \nabla f(N) \rangle_{\mathbb{R}^d}| < \infty$, nous avons*

$$\mathbb{E}[\langle N, \nabla f(N) \rangle_{\mathbb{R}^d} - \langle C, \text{Hess}f(N) \rangle_{\text{HS}}] = 0.$$

2. Notons Z un vecteur gaussien centré de matrice de covariance C , $Z \sim \mathcal{N}_d(0, C)$. Soit $g : \mathbb{R}^d \rightarrow \mathbb{R}$, une fonction de classe $\mathcal{C}_b^2(\mathbb{R}^d)$. Alors, la fonction $U_0(g)$ définie par,

$$U_0g(x) := \int_0^1 \frac{1}{2t} \mathbb{E}[g(\sqrt{t}x + \sqrt{1-t}Z) - g(Z)] dt$$

est solution de l'équation différentielle partielle suivante, (où f est la fonction inconnue)

$$g(x) - \mathbb{E}[g(Z)] = \langle x, \nabla f(x) \rangle_{\mathbb{R}^d} - \langle C, \text{Hess}f(x) \rangle_{\text{HS}}, \quad x \in \mathbb{R}^d. \quad (2.14)$$

De plus, nous avons la borne suivante

$$\sup_{x \in \mathbb{R}^d} \|\text{Hess}U_0g(x)\|_{\text{HS}} \leq \|C^{-1}\|_{\text{op}} \|C\|_{\text{op}}^{\frac{1}{2}} \|g\|_{\text{Lip}}. \quad (2.15)$$

Remarque 12. Le fait que la borne (2.15) dépende de la norme lipschitzienne de g est important puisque cela engendre une condition sur la distance à considérer pour établir une borne de type Berry-Esséen pour des vecteurs aléatoires dont les composantes sont dérivables au sens de Malliavin. Nous allons voir que la classe des fonctions considérées est la classe des fonctions dont la norme lipschitzienne est inférieure à 1.

Preuve : La preuve du lemme 6 se trouvant dans l'article [96] nous montrons uniquement que

$$\sup_{x \in \mathbb{R}^d} \|\text{Hess}U_0g(x)\|_{\text{HS}} \leq \|C^{-1}\|_{\text{op}} \|C\|_{\text{op}}^{\frac{1}{2}} \|g\|_{\text{Lip}}.$$

Soit Y un vecteur gaussien standard de \mathbb{R}^d , i.e : $Y \sim \mathcal{N}_d(0, I_d)$. En utilisant la formule d'intégration par parties (2.9), on a :

$$\begin{aligned} \frac{\partial^2 U_0g}{\partial x_i \partial x_j}(x) &= \int_0^1 \frac{1}{2} \mathbb{E} \left[\frac{\partial^2 g}{\partial x_i \partial x_j} \left(\sqrt{t}x + \sqrt{1-t}Y \right) \right] dt \\ &= \int_0^1 \frac{1}{2\sqrt{1-t}} \mathbb{E} \left[Y_i \frac{\partial g}{\partial x_j} \left(\sqrt{t}x + \sqrt{1-t}Y \right) \right] dt. \end{aligned}$$

Ainsi,

$$\text{Hess}U_0g(x) = \int_0^1 \frac{1}{2\sqrt{1-t}} \mathbb{E} \left[Y \left(\nabla g \left(\sqrt{t}x + \sqrt{1-t}Y \right) \right)^T \right] dt.$$

Soit A une matrice carrée de taille d fixée. Dans ce cas,

$$\begin{aligned} |\langle \text{Hess}U_0g(x), A \rangle_{\text{HS}}| &= \left| \int_0^1 \frac{1}{2\sqrt{1-t}} \mathbb{E} \left[\langle A^T Y, \nabla g \left(\sqrt{t}x + \sqrt{1-t}Y \right) \rangle \right] dt \right| \\ &\leq \|g\|_{\text{Lip}} \mathbb{E} |A^T Y| \leq \|g\|_{\text{Lip}} \|A\|_{\text{HS}}. \end{aligned}$$

On a montré que $\|\text{Hess}U_0g(x)\|_{\text{HS}} \leq \|g\|_{\text{Lip}}$. Pour obtenir (2.15), il suffit d'utiliser le fait que si $Z \sim \mathcal{N}_d(0, C)$ alors $C^{-\frac{1}{2}}Z \sim \mathcal{N}(0, I_d)$. \square

A partir de ce résultat, nous sommes en mesure de quantifier l'erreur d'approximation commise en voulant approcher la loi d'un vecteur aléatoire par la loi d'un vecteur aléatoire gaussien de matrice de covariance C . Naturellement, la notion de distance apparait pour mesurer cet écart. Soit N un vecteur aléatoire de \mathbb{R}^d . Afin d'approximer sa loi par une loi gaussienne multivariée, nous prenons l'espérance puis le supremum à droite et à gauche de (2.14) sur une certaine classe de fonctions, ainsi,

$$d(N, Z) = \sup_{g \in \mathcal{G}} |\mathbb{E}[g(N)] - \mathbb{E}[g(Z)]| \leq \sup_{f \in \mathcal{F}_d} |\mathbb{E}[\langle N, \nabla f(N) \rangle_{\mathbb{R}^d} - \langle C, \text{Hess}f(N) \rangle_{\text{HS}}]|, \quad (2.16)$$

où \mathcal{F}_d est la classe de Stein associée à la distance d . Dans la suite, nous utiliserons cette majoration dans le but d'établir un théorème de type Berry-Esséen pour des vecteurs aléatoires dont les composantes sont dérivables au sens de Malliavin. C'est ainsi qu'entre en scène la théorie de Stein-Malliavin multivariée.

2.4.2 Approximation gaussienne dans le cas multivarié pour des vecteurs aléatoires et matrices aléatoires

Nous voulons désormais établir une version quantitative et multidimensionnelle du théorème du quatrième moment. Par conséquent, nous devons quantifier l'écart dans l'approximation de la loi de notre vecteur aléatoire par celle d'un vecteur gaussien d -dimensionnel. Il est alors nécessaire d'utiliser la borne (2.16), fraîchement établie. La distance que nous utiliserons ici, sera la distance de Wasserstein, car comme nous l'avons vu dans la preuve du lemme 6, la norme lipschitzienne de g intervient dans la majoration (2.15). En effet, cette distance est intéressante puisqu'elle est associée à l'ensemble des fonctions dont la norme lipschitzienne est inférieure à 1.

En vertu du lemme 3, nous disposons d'une formule d'IPP pour des variables aléatoires de carrés intégrables et dérivables au sens de Malliavin, i.e : appartenant à l'espace $\mathbb{D}^{1,2}$. En appliquant le résultat à la quantité $\mathbb{E}[\langle N, \nabla f(N) \rangle_{\mathbb{R}^d} - \langle C, \text{Hess}f(N) \rangle_{\text{HS}}]$, nous parvenons à obtenir l'inégalité (2.17). En particulier nous avons, en appliquant successivement le fait que $\text{LL}^{-1}X = X - \mathbb{E}[X]$, la dualité et la règle de la chaîne,

$$\begin{aligned}
 & \mathbb{E}[\langle N, \nabla f(N) \rangle_{\mathbb{R}^d} - \langle C, \text{Hess}f(N) \rangle_{\text{HS}}] \\
 = & \mathbb{E} \left[\sum_{i=1}^d N_i \frac{\partial f}{\partial x_i}(N_1, \dots, N_d) - \sum_{i,j=1}^d C_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(N_1, \dots, N_d) \right] \\
 = & \sum_{i=1}^d \mathbb{E} \left[\text{LL}^{-1} N_i \frac{\partial f}{\partial x_i}(N_1, \dots, N_d) \right] - \sum_{i,j=1}^d \mathbb{E} \left[C_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(N_1, \dots, N_d) \right] \\
 = & \sum_{i=1}^d \mathbb{E} \left[-\delta(DL^{-1}N_i) \frac{\partial f}{\partial x_i}(N_1, \dots, N_d) \right] - \sum_{i,j=1}^d \mathbb{E} \left[C_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(N_1, \dots, N_d) \right] \\
 = & \sum_{i=1}^d \mathbb{E} \left[\left\langle -DL^{-1}N_i, D \left(\frac{\partial f}{\partial x_i}(N_1, \dots, N_d) \right) \right\rangle_{\mathcal{H}} \right] - \sum_{i,j=1}^d \mathbb{E} \left[C_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(N_1, \dots, N_d) \right] \\
 = & \sum_{i,j=1}^d \mathbb{E} \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(N_1, \dots, N_d) \langle -DL^{-1}N_i, DN_j \rangle_{\mathcal{H}} - C_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(N_1, \dots, N_d) \right] \\
 = & \sum_{i,j=1}^d \mathbb{E} \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(N_1, \dots, N_d) (\langle -DL^{-1}N_i, DN_j \rangle_{\mathcal{H}} - C_{ij}) \right].
 \end{aligned}$$

Finalement, à l'aide de l'inégalité de Cauchy Schwarz et du fait que le supremum de la hessienne de la solution est borné dans (2.15), nous obtenons un théorème central limite quantitatif pour des vecteurs aléatoires dont les composantes sont de carrés intégrables et dérivables au sens de Malliavin.

Théorème 12 (I.Nourdin, G.Peccati et A.Réveillac (2008)). *Soient $d \geq 2$ et $C = \{C_{ij} : i, j = 1, \dots, d\}$, une matrice carrée symétrique de taille d définie positive. Supposons que $Z \sim \mathcal{N}_d(0, C)$ et notons $X = (X_1, \dots, X_d)$ un vecteur aléatoire de \mathbb{R}^d tel que pour tout $i = 1, \dots, d$, $\mathbb{E}[X_i] = 0$ et $X_i \in \mathbb{D}^{1,2}$. Alors,*

$$d_W(X, Z) \leq \|C^{-1}\|_{op} \|C\|_{\delta p}^{\frac{1}{2}} \sqrt{\sum_{i,j=1}^d \mathbb{E} \left[(C_{ij} - \langle DX_i, -DL^{-1}X_j \rangle_{\mathcal{H}})^2 \right]}. \quad (2.17)$$

Remarque 13. 1. $\|\cdot\|_{\text{op}}$ correspond à la norme d'opérateur.

2. La distance de Wasserstein est nécessaire pour contrôler la dérivée seconde de la solution de l'équation de Stein associée à la fonction test g .

Nous donnons ci-dessous une généralisation du théorème du quatrième moment, énoncé par Peccati et Tudor en 2005 dans [102] pour des vecteurs aléatoires dont les composantes sont des intégrales multiples stochastiques. Ce résultat est ensuite amélioré en 2007 par Nualart et Ortiz-Latorre dans [101] pour un vecteur gaussien standard. Nous précisons cette amélioration dans la prochaine remarque.

Théorème 13 (Peccati et Tudor (2005)). Soient $d \geq 2$ et $C = \{C_{ij} : i, j = 1, \dots, d\}$ une matrice carrée de taille d définie positive. Soient $1 \leq q_1 \leq \dots \leq q_d$. Pour tout $n \geq 1$ et $i = 1, \dots, d$, prenons $f_i^{(n)}$ appartenant à $\mathcal{H}^{\odot q_i}$. Supposons que

$$F^{(n)} = (F_1^{(n)}, \dots, F_d^{(n)}) := (I_{q_1}(f_1^{(n)}), \dots, I_{q_d}(f_d^{(n)})), \quad n \geq 1,$$

tel que

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[F_i^{(n)} F_j^{(n)} \right] = C_{ij}, \quad 1 \leq i, j \leq d.$$

Alors, les assertions suivantes sont équivalentes :

1. $F^{(n)}$ converge en loi vers un vecteur gaussien centré de matrice de covariance C quand $n \rightarrow \infty$.
2. Pour tout $1 \leq i \leq d$, $\mathbb{E} \left[\left(F_i^{(n)} \right)^4 \right] \xrightarrow{n \rightarrow \infty} 3C_{ii}^2$.
3. Pour tout $1 \leq i \leq d$, $F_i^{(n)}$ converge en loi vers une variable aléatoire gaussienne centrée de variance C_{ii} quand $n \rightarrow \infty$.
4. Pour tout $1 \leq i \leq d$ et $1 \leq r \leq q_i - 1$, $\|f_i^{(n)} \otimes_r f_i^{(n)}\|_{\mathcal{H}^{\otimes 2(q_i-r)}} \xrightarrow{n \rightarrow \infty} 0$

Remarque 14. En 2007, Nualart et Ortiz-Latorre dans [101] donnent une condition supplémentaire équivalente pour $C = I_d$. Cette condition équivalente est donnée par :

$$- \text{ Pour tout } 1 \leq i \leq d, \|DF_i^{(n)}\|_{\mathcal{H}}^2 \xrightarrow[n \rightarrow \infty]{L^2} q_i.$$

Ce théorème est la contrepartie du célèbre théorème du quatrième moment pour des vecteurs aléatoires de \mathbb{R}^d . En appliquant le théorème 12 à un vecteur aléatoire dont les composantes sont des intégrales multiples d'ordre arbitraire, nous obtenons le corollaire 1. Nourdin et Zheng dans [98] s'appuieront sur ce corollaire pour quantifier la distance entre la loi d'une matrice de Wishart et la matrice du GOE que nous appellerons par abus de langage une matrice gaussienne.

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

$$d_W(F, Z) \leq \|C^{-1}\|_{\text{op}} \|C\|_{\text{op}}^{\frac{1}{2}} \sqrt{\sum_{1 \leq i, j \leq d} \mathbb{E} \left[\left(C(i, j) - \frac{1}{q_j} \langle DF_i, -DF_j \rangle_{\mathcal{H}} \right)^2 \right]}. \quad (2.18)$$

Pour parvenir à un théorème central limite pour les matrices aléatoires, nous utiliserons les résultats établis précédemment, pour les vecteurs aléatoires.

Les matrices que nous sommes amenés à étudier au cours de cette thèse sont assez spécifiques,

ce sont des matrices de Wishart. Elles correspondent à des matrices de variance-covariance empirique. Elles sont donc symétriques et carrées. Bien que les théorèmes que nous venons de voir ne puissent pas être appliqués directement en l'état, en 2018, Nourdin et Zheng dans [98] décident d'utiliser la notion de demi-vecteur, appelée *half vector* en anglais, qui permet de passer d'une matrice symétrique à un vecteur. L'idée est ensuite d'effectuer une majoration assez intuitive de la distance entre la loi de deux matrices par celle entre deux vecteurs. Dans la suite, nous définissons la notion de demi-vecteur associé à une matrice symétrique.

Définition 23. Si $\mathcal{A} = (A_{ij})_{1 \leq i, j \leq n}$ est une matrice aléatoire carrée et symétrique de taille n , notons $\mathcal{A}^{\text{half}}$, le demi-vecteur associé à la matrice \mathcal{A} . Il correspond au vecteur aléatoire de taille $\frac{n(n+1)}{2}$, formé par la diagonale supérieure de la matrice, à savoir :

$$\mathcal{A}^{\text{half}} = (A_{11}, A_{12}, \dots, A_{1n}, A_{22}, A_{23}, \dots, A_{2n}, \dots, A_{nn}).$$

Grâce au demi-vecteur, l'étude asymptotique des matrices de Wishart revient à s'intéresser aux lois des vecteurs aléatoires associés à ces matrices en cherchant la distance avec une loi gaussienne multivariée. Ainsi, la méthode de Stein-Malliavin pour les matrices aléatoires reposent sur le théorème 12 et le lemme suivant.

Lemme 7. Soient \mathcal{X} et \mathcal{Y} deux matrices aléatoires symétriques de $\mathcal{M}_n(\mathbb{R})$ et $\mathcal{X}^{\text{half}}$, $\mathcal{Y}^{\text{half}}$, leur demi-vecteur associé. Alors

$$d_W(\mathcal{X}, \mathcal{Y}) \leq \sqrt{2} d_W(\mathcal{X}^{\text{half}}, \mathcal{Y}^{\text{half}}).$$

Par conséquent, pour évaluer la distance matricielle, il suffit d'analyser la diagonale supérieure de la matrice (dont la diagonale est incluse). La constante $\sqrt{2}$ apparaît assez naturellement. En effet, nous pouvons dire que si nous avons une matrice de Wishart de taille n alors son demi-vecteur associé sera de taille $\frac{n(n+1)}{2}$.

Il faut ensuite calculer les quantités $\mathbb{E} \left[(\mathbb{E}[Z_{ij}Z_{ab}] - \langle DW_{ij}, -DL^{-1}W_{ab} \rangle_{\mathcal{H}})^2 \right]$ du théorème 12 afin de quantifier l'erreur commise dans l'approximation de notre demi-vecteur par un vecteur gaussien. Les $(W_{ij})_{1 \leq i, j \leq n}$ correspondent aux coefficients de la matrice de Wishart et $(Z_{ij})_{1 \leq i, j \leq n}$ correspondent à une matrice du GOE. Nous détaillons dans le chapitre 4 l'étude des matrices de Wishart par la méthode de Stein-Malliavin.

Partie II

**Axes de recherche : EDPS, matrices
et tenseurs aléatoires**

Équations stochastiques aux dérivées partielles dirigées par un bruit blanc en temps et en espace

Ce chapitre sera dédié à l'analyse des équations stochastiques aux dérivées partielles dites EDPS ou SPDE en anglais. Dans la section 3.1, nous chercherons à motiver l'expression de la solution évolutive de ces équations. La section 3.2 s'intéressera plus particulièrement aux relations entre les solutions des EDPS et les processus gaussiens introduits auparavant. Nous nous pencherons ensuite sur le sujet de l'inférence statistique et nous proposerons des estimations de paramètres basées sur les variations quadratiques des solutions. En particulier, ces parties correspondent aux résumés des articles [55] et [8].

3.1 Introduction aux équations aux dérivées partielles stochastiques

Les équations aux dérivées partielles stochastiques (EDPS) généralisent les équations aux dérivées partielles classiques en ajoutant des forces stochastiques et en permettant de modéliser plus justement des phénomènes physiques que nous pouvons rencontrer dans la réalité.

Le problème que J. Walsh a initialement posé dans son cours [125] était la considération d'un système physique pouvant être représenté par une équation aux dérivées partielles et dont la perturbation aléatoire serait un bruit blanc en temps et en espace, appelé *space-time white noise* en anglais.

Définition 24 (Bruit blanc en temps et en espace). *Soit $\mathcal{B}_b(\mathbb{R}^d)$ l'ensemble des sous-ensembles boréliens bornés de \mathbb{R}^d et soit $T = \mathbb{R}_+ \times \mathcal{B}_b(\mathbb{R}^d)$. Le bruit blanc en temps et en espace est un champ aléatoire gaussien $(W(x), x \in T)$ de fonction de covariance,*

$$\mathbb{E}[W(x)W(y)] = \min(t,s)\lambda(A \cap B), \tag{3.1}$$

pour $x = (t, A)$ et $y = (s, B)$, pour $t, s \geq 0$ et $A, B \in \mathcal{B}_b(\mathbb{R}^d)$. λ correspond à la mesure de Lebesgue sur \mathbb{R}^d .

La dénomination de ce bruit provient du fait qu'il se comporte comme un processus de Wiener par rapport au temps et par rapport à sa variable spatiale. En effet, pour $A \in \mathcal{B}_b(\mathbb{R}^d)$, le processus $(W(t,A))_{t \geq 0}$ a la même loi que $(\sqrt{\lambda(A)}B_t)_{t \geq 0}$, où $(B_t)_{t \geq 0}$ est un mouvement brownien standard. De plus, pour $t \geq 0$ fixé, le processus $(W(t,A), A \in \mathcal{B}_b(\mathbb{R}^d))$ a le même comportement que $(\sqrt{t}W(x))_{x \in \mathbb{R}^d}$, où $(W(x))_{x \in \mathbb{R}^d}$ est un mouvement brownien anisotrope à d paramètres.

Remarque 15. *Nous pouvons associer un processus isonormal au bruit blanc en temps et en espace.*

Dans cette thèse, le cadre dans lequel nous nous plaçons est celui dessiné par les équations différentielles stochastiques dirigées par un bruit blanc en temps et en espace. De cette manière, nous désirons nous focaliser sur la définition d'une solution pour les EDPS de la forme suivante ;

$$Lu(t,x) = \sigma(u(t,x))\dot{W}(t,x), \quad t \geq 0, x \in \mathbb{R}^d, \quad (3.2)$$

où L est un opérateur différentiel linéaire, σ est une fonction lipschitzienne et W est notre bruit stochastique, de covariance donnée par (3.1). Plus précisément, nous considérons dans cette partie que l'opérateur sera l'un des trois suivants,

- l'opérateur de la chaleur $L = \frac{\partial}{\partial t} - \frac{1}{2}\Delta$,
- l'opérateur la chaleur fractionnaire $L = \frac{\partial}{\partial t} + (-\Delta)^{\frac{\alpha}{2}}$,
- ou bien l'opérateur des ondes $L = \frac{\partial^2}{\partial t^2} - \Delta$,

où Δ est le laplacien standard défini sur \mathbb{R}^d et $(-\Delta)^{\frac{\alpha}{2}}$ correspond au laplacien fractionnaire. Sans plus attendre, nous rappelons la définition du laplacien fractionnaire, donnée par sa transformée de Fourier.

Définition 25. *Soit $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Nous définissons l'opérateur laplacien fractionnaire, pour $\alpha \in (1, 2]$ par :*

$$\mathcal{F}\left((-\Delta)^{\frac{\alpha}{2}}\varphi\right)(\xi) = |\xi|^\alpha \mathcal{F}(\varphi)(\xi), \quad \text{pour } \xi \in \mathbb{R}^d.$$

D'autres opérateurs différentiels linéaires existent et permettent de définir d'autres équations physiques. Nous vous renvoyons au cours de M. Sanz-Solé [108] pour plus de détails.

Nous cherchons maintenant à donner les solutions fondamentales de l'équation homogène associée

$$Lu(t,x) = 0 \quad t \geq 0, x \in \mathbb{R}^d. \quad (3.3)$$

Lorsque L est l'opérateur de la chaleur ou des ondes, les solutions fondamentales sont bien connues et peuvent se trouver à l'aide de la transformée de Fourier. Notons $\|\cdot\|$ la norme euclidienne sur \mathbb{R}^d . En utilisant la transformée de Fourier inverse, nous pouvons les définir de la manière suivante.

1. Si L est l'opérateur de la chaleur alors la solution fondamentale associée est

$$G(t,x) = \begin{cases} (2\pi t)^{-\frac{d}{2}} \exp\left(-\frac{\|x\|^2}{2t}\right) & \text{si } t > 0, x \in \mathbb{R}^d, \\ 0 & \text{si } t \leq 0, x \in \mathbb{R}^d. \end{cases} \quad (3.4)$$

2. Si L est l'opérateur de la chaleur fractionnaire,

$$G_\alpha(t,x) = \int_{\mathbb{R}^d} e^{-i\langle \xi, x \rangle} e^{-t\|\xi\|^\alpha} d\xi, \quad t > 0, x \in \mathbb{R}^d. \quad (3.5)$$

Notifions que pour $\alpha = 2$, le noyau de la chaleur fractionnaire et celui de la chaleur classique sont "presque égaux". En effet, cela provient de la constante $\frac{1}{2}$ apparaissant dans le noyau de Green de l'équation de la chaleur, et

$$G_2(t, x) = G(2t, x).$$

Ce noyau vérifie plusieurs propriétés qui ont été utilisées dans l'article [55].

3. Si L est l'opérateur des ondes, son noyau de Green associé est

$$G_1(t, x) = \begin{cases} \frac{1}{2} \mathbf{1}_{\{|x| < t\}} & \text{si } d = 1, \\ \frac{1}{2\pi} \frac{1}{\sqrt{t^2 - |x|^2}} \mathbf{1}_{\{|x| < t\}} & \text{si } d = 2, \\ \frac{1}{4\pi} \sigma_t & \text{si } d = 3. \end{cases} \quad (3.6)$$

σ_t est la mesure de surface sur la sphère $\{x \in \mathbb{R}^3; \|x\| = t\}$. Pour $d \geq 4$, $G(t, \cdot)$ est une distribution à support compact sur \mathbb{R}^d . Une manière plus simple de définir G_1 en toute dimension repose sur sa transformée de Fourier, pour tout $d \geq 1$, (avec $\xi \neq 0$)

$$\forall \xi \in \mathbb{R}^d, t > 0 \quad \mathcal{F}G_1(t, \cdot)(\xi) = \frac{\sin(t\|\xi\|)}{\|\xi\|}.$$

$\|\cdot\|$ indique la norme euclidienne sur \mathbb{R}^d .

Notre problème initial était de vouloir donner une solution à l'équation (3.2). Pour cela, nous proposons tout d'abord de considérer l'équation linéaire pour $\sigma \equiv 1$, soit

$$Lu(t, x) = \dot{W}(t, x), \quad t \geq 0, x \in \mathbb{R}^d. \quad (3.7)$$

En remplaçant notre bruit W dans (3.7) par une force déterministe $f \in \mathcal{C}^{1,2}([0, \infty) \times \mathbb{R}^d)$,

$$Lu(t, x) = f(t, x), \quad t \geq 0, x \in \mathbb{R}^d, \quad (3.8)$$

la théorie classique des équations aux dérivées partielles permet de fournir une solution qui est basée sur le principe de Duhamel. Si Γ est l'un des noyaux de Green proposés ci-dessus, en résolvant l'équation non homogène (3.8) alors par la théorie classique, la solution est donnée pour tout $t \geq 0$ et $x \in \mathbb{R}^d$ par

$$u(t, x) = \int_0^t \int_{\mathbb{R}^d} \Gamma(t-s, x-y) f(s, y) dy ds.$$

Ensuite, en substituant f par le bruit W dans (3.8), nous aurions $W(s, y) ds dy$ qui devrait être formellement interprété comme $W(ds, dy)$. Nous obtenons une solution évolutive, appelée *mild solution* de l'équation (3.7), telle que pour tout $t \geq 0$ et $x \in \mathbb{R}^d$,

$$u(t, x) = \int_0^t \int_{\mathbb{R}^d} \Gamma(t-s, x-y) W(ds, dy). \quad (3.9)$$

La question qui se pose alors est la suivante ;

L'intégrale de Wiener par rapport au bruit blanc en temps et en espace est-elle bien définie ?

Il se trouve que l'intégrale de Wiener (3.9) par rapport au bruit blanc en temps et en espace est bien définie si la fonction,

$$(s, y) \mapsto \Gamma(t - s, x - y) \mathbf{1}_{[0, t](s)}$$

appartient à l'espace $L^2(\mathbb{R}_+ \times \mathbb{R})$. Un calcul rapide permet de montrer que la solution évolutive de l'équation linéaire (3.7) existe si et seulement si $d = 1$, pour Γ correspondant au noyau de la chaleur (3.4), de la chaleur fractionnaire (3.5) ou celui des ondes (3.6).

Pour $d \geq 2$, nous ne pouvons plus définir de cette manière la solution évolutive de (3.7) et c'est ainsi que la théorie de Walsh-Dalang entre en scène. Dans un premier temps, J. Walsh va s'appuyer sur la définition de l'intégrale d'Itô, en intégrant non plus par rapport à une martingale mais à une mesure de martingale.

Définition 26 (Mesure de martingale). *Soit un espace de probabilité $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$. On définit la notion de mesure de martingale $(M_t(A), t \in \mathbb{R}^+, A \in \mathcal{B}_b(\mathbb{R}^d))$ sur $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ avec les hypothèses suivantes.*

- $t \mapsto M_t(A)$ est une martingale.
- $A \mapsto M_t(A)$ est une $L^2(\Omega, \mathbb{F}, P)$ -mesure.
- Soit Q une mesure signée. La covariance s'écrit

$$\mathbb{E}(M_t(A)M_t(B)) = Q([0, t] \times A \times B).$$

L'intégrale stochastique est d'abord définie pour les processus simples. Afin de pouvoir étendre cette définition, Walsh est amené à introduire la notion de martingale dite *worthy*, qui va permettre de considérer des processus plus généraux. Nous renvoyons à l'introduction de J. Walsh [125] pour une présentation approfondie ou à l'introduction de la thèse de O. Assaad [7]. Finalement la théorie de Walsh va permettre de donner une solution de l'équation de la chaleur en toute dimension. En revanche, dans le cas où les intégrandes sont des distributions, cette théorie atteint ses limites. La solution de l'équation des ondes dirigées par un bruit blanc en temps et en espace étant une mesure pour $d = 3$ et une distribution pour $d > 3$, elle est seulement bien définie pour $d \leq 2$.

Pour remédier à cela, R. Dalang [39] présente une manière de définir la solution de l'équation des ondes en toute dimension. Il passe alors par la considération d'un bruit gaussien blanc en temps et coloré en espace, dit spatialement homogène qui est un processus gaussien centré $\{M(A), A \in \mathcal{B}_b(\mathbb{R}_+ \times \mathbb{R})\}$ de covariance

$$\mathbb{E}[M(A)M(B)] = \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}_A(t, x) \mathbf{1}_B(t, x) f(x - y) dx dy,$$

où A et B appartiennent à $\mathcal{B}_b(\mathbb{R}_+ \times \mathbb{R})$ et f est une fonction positive définie sur \mathbb{R}^d au sens des distributions. Le théorème de Bochner-Schwartz permet alors d'affirmer que f est la transformée de Fourier d'une mesure tempérée μ sur \mathbb{R}^d . Avec ce type de bruit, la condition pratique de Dalang provenant du calcul de (3.11) nous donne une solution en toute dimension pour l'équation des ondes et de la chaleur.

Proposition 7. *Soit L l'opérateur des ondes ou de la chaleur pour tout $d \geq 1$. Une condition pour l'existence d'une solution est*

$$\int_{\mathbb{R}^d} \frac{1}{1 + |\xi|^2} \mu(d\xi) < +\infty.$$

Remarque 16. *Le bruit blanc en temps et en espace est obtenu pour $f = \delta_0$ et μ correspond à la mesure de Lebesgue.*

Dans le cas des équations stochastiques non linéaire, la théorie de Walsh-Dalang a aussi permis de définir des solutions. Pour cela, considérons une nouvelle fois l'équation (3.2). Heuristiquement, une solution de cette équation peut toujours être donnée par la théorie classique, en choisissant une force déterministe f à la place de notre bruit puis en la substituant par le bruit gaussien spatialement homogène W . Ainsi la solution évolutive de (3.2) est donnée pour tout $t \geq 0$ et $x \in \mathbb{R}^d$ par,

$$u(t, x) = \int_0^t \int_{\mathbb{R}^d} \Gamma(t-s, x-y) \sigma(u(s, y)) W(ds, dy). \quad (3.10)$$

Notons qu'une condition suffisante pour que cette intégrale soit bien définie est que u soit mesurable et \mathcal{F}_t -adaptée et que

$$\sup_{t \in [0, T]} \sup_{x \in \mathbb{R}^d} \mathbb{E}|u(t, x)|^2 < \infty, \quad \text{pour tout } T > 0,$$

où $\mathcal{F}_t := \sigma(\{W_s(A); 0 \leq s \leq t, A \in \mathcal{B}_b(\mathbb{R}^d)\}) \vee \mathcal{N}$, avec \mathcal{N} est la tribu engendrée par les ensembles de mesure nulle.

Théorème 14. *Considérons les hypothèses suivantes,*

1. *La solution $\Gamma(\cdot, \star)$ de l'équation homogène (3.3) est une fonction déterministe en \cdot à valeur dans l'espace dual de l'espace de Schwartz $\mathcal{S}'(\mathbb{R}^d)$ telle que*

$$\int_0^T \int_{\mathbb{R}^d} \mu(d\xi) |\mathcal{F}\Gamma(t, \cdot)(\xi)|^2 < \infty \quad (3.11)$$

avec $T > 0$ et Γ est une mesure positive de la forme $\Gamma(t, dx)dt$ telle que

$$\sup_{t \in [0, T]} \Gamma(t, \mathbb{R}^d) < \infty.$$

2. *$\sigma(\cdot)$ est une fonction lipschitzienne.*

Alors (3.2) admet une unique solution $u(t, x)$ bien définie et donnée par (3.10). De plus, cette solution est L^2 -continue et pour tout $T > 0$ et $p \geq 1$,

$$\sup_{t \in [0, T]} \sup_{x \in \mathbb{R}^d} \mathbb{E}|u(t, x)|^p < \infty.$$

Remarque 17. *La théorie de Walsh-Dalang ne permet pas de considérer des bruits colorés en temps et colorés ou non en espace. En revanche, nous vous renvoyons au chapitre 5 de l'introduction [13] de R. Balan, et au cours de M. Sanz-Solé [108].*

3.2 Propriétés liées aux SPDEs de la chaleur et des ondes

La loi des solutions des équations différentielles stochastiques est souvent liée à celle des processus gaussiens, à savoir, le mouvement brownien standard, fractionnaire, et même bi-fractionnaire. Nous vous proposons de vous tourner vers le livre de C. Tudor [117] pour une preuve des résultats constituant cette section.

3.2.1 Équation stochastique de la chaleur

L'équation de la chaleur est l'une des équations fondamentales de la physique et fut introduite au XIX-ième siècle par le scientifique J. Fourier. Elle permet de décrire la variation de la température à l'intérieur d'un matériau au fil du temps et en fonction de l'emplacement spatial.

Cette équation joue un rôle crucial dans de nombreux domaines de la science et de l'ingénierie, notamment en thermodynamique, en géophysique, en météorologie, et en génie thermique.

Elle peut être utilisée pour modéliser les phénomènes de conduction thermique, de transfert de chaleur dans des systèmes où plusieurs matériaux sont en contact. En effet, lorsqu'un matériau est en contact avec un objet chaud ou froid, la chaleur se propage à travers celui-ci par conduction thermique. Dans ce cas, l'équation de la chaleur permet de prédire comment la température varie à l'intérieur du matériau au fil du temps en réponse à ces différences de température.

Une perturbation extérieure peut venir affecter la température d'un matériau et l'ajout d'un bruit semble pertinent. En ce sens, l'équation stochastique de la chaleur permet de mieux refléter la réalité.

Nous allons maintenant nous intéresser à la loi de la solution de l'équation stochastique linéaire de la chaleur dirigée par un bruit blanc en temps et en espace, dont la fonction de covariance est donnée par (3.1). Nous nous restreignons au cas où $d = 1$, puisque seul ce cas est considéré dans ce manuscrit.

L'intégrale de Wiener par rapport au bruit blanc en temps et en espace étant bien définie pour $d = 1$, nous savons qu'une solution évolutive de cette équation est donnée par (3.9) où $\Gamma = G$ correspond au noyau de Green (3.4) de l'équation de la chaleur homogène. Nous rappelons que l'équation considérée est la suivante ;

$$\begin{cases} \frac{\partial}{\partial t} u(t, x) = \frac{1}{2} \Delta u(t, x) + \dot{W}(t, x), & t \geq 0, x \in \mathbb{R}, \\ u(0, x) = 0, & \text{pour tout } x \in \mathbb{R}. \end{cases}$$

Soit $x \in \mathbb{R}$ fixé et considérons le processus $(u(t, x))_{t \geq 0}$. Sa covariance est donnée pour tout $s, t \geq 0$ par,

$$\mathbb{E}[u(t, x)u(s, x)] = \frac{1}{\sqrt{2\pi}} \left(\sqrt{t+s} - \sqrt{|t-s|} \right).$$

Cette covariance correspond à celle du mouvement brownien bi-fractionnaire à constante près, pour les paramètres de Hurst $H = K = \frac{1}{2}$, et nous avons le résultat suivant.

Proposition 8. *Soit $x \in \mathbb{R}$ fixé. Le processus stochastique $(u(t, x))_{t \geq 0}$ coïncide en loi, modulo une constante avec le mouvement brownien bi-fractionnaire $\left(B_t^{\frac{1}{2}, \frac{1}{2}} \right)_{t \geq 0}$ de paramètres de Hurst $H = K = \frac{1}{2}$. Plus précisément,*

$$(u(t, x))_{t \geq 0} \stackrel{(d)}{=} \left(\pi^{-\frac{1}{4}} B_t^{\frac{1}{2}, \frac{1}{2}} \right)_{t \geq 0}.$$

Le processus gaussien $(u(t, x))_{t \geq 0}$ vérifie donc les mêmes propriétés que le mouvement bi-fractionnaire.

- C'est un processus auto-similaire d'indice d'auto-similarité $\frac{1}{4}$.
- Le lien entre le mouvement brownien bi-fractionnaire et le fBm avancé par le théorème 4, nous conduit à

$$\left(\pi^{\frac{1}{4}} u(t, x) + X_t \right)_{t \geq 0} \stackrel{(d)}{=} \left(2^{\frac{1}{4}} B_t^{\frac{1}{4}} \right)_{t \geq 0}$$

- où $(X_t)_{t \geq 0}$ est un processus gaussien à trajectoires absolument continues \mathcal{C}^∞ et indépendantes de la solution u et $B_t^{\frac{1}{4}}$ un fBm de paramètre de Hurst $H = \frac{1}{4}$.
- D'après l'encadrement (1.6) effectué pour le mouvement bi-fractionnaire, il existe deux constantes $0 < C_1 < C_2$ telles que pour tout $s, t \geq 0$, $x \in \mathbb{R}$,

$$C_1 |t - s|^{\frac{1}{2}} \leq \mathbb{E}|u(t, x) - u(s, x)|^2 \leq C_2 |t - s|^{\frac{1}{2}}.$$

Les constantes C_1, C_2 ne dépendent pas de $x \in \mathbb{R}$. Et en particulier, pour tout $x \in \mathbb{R}$, les trajectoires $t \mapsto u(t, x)$ sont δ -höldériennes pour $\delta \in (0, \frac{1}{4})$.

En espace, la solution $(u(t, x))_{x \in \mathbb{R}}$ ne dépend pas directement du fBm ou du mouvement brownien bi-fractionnaire. Cependant il est possible de donner une décomposition de la solution spatiale par rapport au processus de Wiener bilatéral perturbé par un processus gaussien à trajectoires lisses. La solution en espace est hölderienne d'ordre $\delta \in (0, \frac{1}{2})$ et nous avons la proposition suivante.

Proposition 9. *Pour tout $x, y \in \mathbb{R}$, $s, t \geq 0$, nous avons $p \geq 2$ tel que,*

$$\mathbb{E}|u(t, x) - u(s, y)|^p \leq C_p \left[|t - s|^{\frac{p}{4}} + |x - y|^{\frac{p}{2}} \right].$$

3.2.2 Équation stochastique de la chaleur fractionnaire

L'équation de la chaleur permet de modéliser la propagation de la chaleur dans des matériaux isotropes, ainsi elle n'est pas toujours très représentative de la réalité si le matériau ne satisfait pas des conditions optimales et est plutôt anisotrope. Par conséquent, cela peut avoir des impacts sur l'évolution du flux thermique. Afin de pallier ce problème et prendre en compte ces phénomènes, les physiciens et mathématiciens se sont tournés vers le laplacien fractionnaire, au lieu du laplacien classique. En effet, le laplacien classique provient de la loi de Fourier qui est phénoménologique. Ils ont alors considéré l'équation suivante.

$$\frac{\partial}{\partial t} u(t, x) = -(-\Delta)^{\frac{\alpha}{2}} u(t, x) + \dot{W}(t, x), \quad t \in [0, T], x \in \mathbb{R}.$$

Comme pour l'équation de la chaleur stochastique, nous proposons de rappeler sa relation avec le mouvement brownien bi-fractionnaire et fractionnaire. Considérons le processus $(u(t, x))_{t \geq 0}$ avec $x \in \mathbb{R}$ fixé, solution de l'équation (3.19) telle que,

$$u(t, x) = \int_0^t \int_{\mathbb{R}} G_\alpha(t - s, x - y) ds dy \quad x \in \mathbb{R}, t \geq 0,$$

où G_α est le noyau de Green (3.5). Nous donnons l'expression de sa covariance temporelle, pour tout $s, t \geq 0$ et $x \in \mathbb{R}$ fixé,

$$\mathbb{E}[u(t, x)u(s, x)] = c_{1, \alpha} \left[(t + s)^{1 - \frac{1}{\alpha}} - |t - s|^{1 - \frac{1}{\alpha}} \right],$$

où

$$c_{1, \alpha} = \frac{1}{2\pi(\alpha - 1)} \Gamma\left(\frac{1}{\alpha}\right).$$

Par conséquent, nous pouvons remarquer que le processus $(u(t, x))_{t \geq 0}$ et le mouvement brownien bi-fractionnaire de paramètres de Hurst $H = \frac{1}{2}$ et $K = 1 - \frac{1}{\alpha}$ ont même loi,

$$(u(t, x))_{t \geq 0} \stackrel{(d)}{=} \left(c_{2, \alpha} \mathcal{B}_t^{1-\frac{1}{\alpha}} \right)_{t \geq 0},$$

où $c_{2, \alpha}^2 = c_{1, \alpha} 2^{1-\frac{1}{\alpha}}$.

En particulier, le point 1 de la remarque 8 évoque cette situation et rappelle aussi le lien entre la solution $(u(t, x))_{t \geq 0}$ et le fBm obtenu par le théorème 4. L'encadrement (1.6) permet également de montrer que les trajectoires $t \mapsto u(t, x)$ sont δ -höldériennes pour $\delta \in (0, \frac{1}{2}(1 - \alpha))$.

En ce qui concerne la solution spatiale de l'équation de la chaleur fractionnaire dirigée par un bruit blanc en temps et en espace, nous vous renvoyons vers la remarque 7 qui rappelle que la loi de $(u(t, x))_{x \in \mathbb{R}}$ coïncide avec celle du fBm.

L'encadrement énoncé dans la proposition suivante, nous permettra de fournir un argument sur la régularité des trajectoires.

Proposition 10. *Soit $(u(t, x), t \geq 0, x \in \mathbb{R})$. Alors pour tout $t > 0$, $x, y \in \mathbb{R}$ avec $|x - y|$ assez petit, il existe deux constantes $0 < C_1 < C_2$ telles que,*

$$C_1 |x - y|^{\alpha-1} \leq \mathbb{E} |u(t, x) - u(t, y)|^2 \leq C_2 |x - y|^{\alpha-1}.$$

En particulier, pour $t > 0$ fixé, l'application $x \mapsto u(t, x)$ est δ -höldérienne pour $\delta \in (0, \frac{\alpha-1}{2})$. De manière générale les trajectoires $(t, x) \mapsto u(t, x)$ sont continues presque partout sur $(0, \infty) \times \mathbb{R}$. Et pour tout $x, y \in \mathbb{R}$ et $s, t \geq 0$, nous avons,

$$\forall p \geq 2, \quad \mathbb{E} |u(t, x) - u(s, y)|^p \leq C_p \left[|t - s|^{\frac{p(\alpha-1)}{2\alpha}} + |x - y|^{\frac{p(\alpha-1)}{2}} \right].$$

3.2.3 Équation stochastique des ondes

Une autre équation fondamentale de la physique est l'équation des ondes, aussi appelée équation de d'Alembert. Elle permet la description de la propagation des ondes à travers l'espace et le temps. En particulier, elle peut modéliser la manière dont les ondes se déplacent et interagissent avec leur environnement, que ce soit sous forme de vagues sonores, de rayonnement électromagnétique, ou même de vibrations mécaniques.

L'ajout d'un bruit permet de refléter plus justement la réalité. Par exemple si nous pensons à une onde sonore, les interférences sont à prendre en compte et c'est ce que permet l'ajout d'un bruit.

Nous allons nous intéresser à l'équation des ondes dirigée par un bruit blanc en temps et en espace. Nous rappelons qu'elle est donnée sous la forme suivante.

$$\begin{cases} \frac{\partial^2 u}{\partial t^2}(t, x) = \Delta u(t, x) + \dot{W}(t, x), & t > 0, x \in \mathbb{R}^d \\ u(0, x) = 0, & x \in \mathbb{R}^d \\ \frac{\partial u}{\partial t}(0, x) = 0, & x \in \mathbb{R}^d. \end{cases} \quad (3.12)$$

Nous nous penchons plus particulièrement sur le cas où $d = 1$. Dans cette situation, la solution de l'équation (3.12) s'écrit,

$$u(t, x) = \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y) W(ds, dy), \quad t > 0, x \in \mathbb{R},$$

où G_1 est le noyau de Green des ondes donné par (3.6) pour $d = 1$.

Il s'avère que le processus $(u(t, x))_{t \geq 0}$ a même loi au sens des distributions finies dimensionnelles que le processus $(\frac{1}{2}W_{t^2})_{t \geq 0}$, où $(W_t)_{t \geq 0}$ est un processus de Wiener standard. En effet, le calcul de sa fonction de covariance pour $x \in \mathbb{R}$ fixé, nous donne que pour tout $t, s \geq 0$,

$$\mathbb{E}[u(t, x)u(s, x)] = \frac{1}{4}(t \wedge s)^2.$$

Il vérifie alors les propriétés suivantes :

- Le processus $(u(t, x))_{t \geq 0}$ est auto-similaire d'ordre 1.
- Pour tout $t, s \geq 0$,

$$\mathbb{E}|u(t, x) - u(s, x)|^2 = \frac{1}{4}(\max(t, s)^2 - \min(t, s)^2).$$

- Soit $T > 0$. Pour tout $s, t \in [0, T]$, nous avons,

$$\mathbb{E}|u(t, x) - u(s, x)|^2 \leq C_2|t - s|$$

avec $C_2 = C_2(T) > 0$. Il en découle que $t \mapsto u(t, x)$ est presque sûrement δ -höldérienne pour $\delta \in (0, \frac{1}{2})$.

- $(u(t, x))_{t \geq 0}$ est un processus à accroissements indépendants.

En ce qui concerne son comportement en espace, nous pouvons observer un lien entre la loi de $(u(t, x))_{x \in \mathbb{R}}$ et celle d'un mouvement brownien bilatéral. Sa fonction de covariance est notamment donnée pour tout $t \geq 0$ et $x, y \in \mathbb{R}$,

$$\mathbb{E}(u(t, x)u(t, y)) = \frac{1}{4} \left(\frac{|y - x|}{2} - t \right) \mathbf{1}_{\{|y - x| < 2t\}}.$$

Il vérifie les propriétés suivantes ;

- $(u(t, x))_{x \in \mathbb{R}}$ est stationnaire.
- Si $|x - y| \geq 2t$, alors $u(t, x)$ et $u(t, y)$ sont des variables aléatoires indépendantes.
- Les trajectoires $x \mapsto u(t, x)$ sont $\frac{1}{2}$ -höldérienne. En particulier, si nous considérons $(u(t, x), t \geq 0, x \in \mathbb{R})$, pour $t \in [0, T]$ fixé, il existe deux constantes $0 \leq C_{2,t} \leq C_1$ telles que,

$$C_{2,t}|x - y| \leq \mathbb{E}(|u(t, x) - u(t, y)|^2) \leq C_1|x - y|$$

pour tout $x, y \in [-M, M]$ distincts et $M > 0$.

3.3 Résultats

3.3.1 Résumé du chapitre 6

Dans l'article [8], nous nous sommes intéressés au comportement asymptotique des variations quadratiques en temps puis en espace de la solution de l'équation stochastique des ondes dirigée par un bruit blanc en temps et en espace. Cette analyse nous a permis d'établir un estimateur fortement consistant du paramètre de dérive basé sur les variations quadratiques de la solution temporelle et également une estimation du paramètre de diffusion.

Nous considérons l'équation des ondes dirigées par un bruit blanc en temps et en espace W , dont nous rappelons l'expression,

$$\begin{cases} \frac{\partial^2 u}{\partial t^2}(t, x) = \theta \Delta u(t, x) + \sigma \dot{W}(t, x), & t > 0, x \in \mathbb{R} \\ u(0, x) = 0, & x \in \mathbb{R} \\ \frac{\partial u}{\partial t}(0, x) = 0, & x \in \mathbb{R}, \end{cases} \quad (3.13)$$

où le paramètre de dérive θ et le paramètre de diffusion σ sont strictement positifs. Le noyau de Green G_1 est la solution fondamentale de l'équation homogène des ondes $\left(\frac{\partial^2}{\partial t^2} u - \Delta u = 0\right)$ et est donné pour $t > 0$ et $x \in \mathbb{R}$ par,

$$G_1(t, x) = \frac{1}{2} \mathbf{1}_{\{|x| < t\}}.$$

La solution de l'équation de (3.13) est définie pour tout $t > 0$ et $x \in \mathbb{R}$ par

$$u(t, x) = \sigma \int_0^t \int_{\mathbb{R}} G_1(\sqrt{\theta}t - s, x - y) W(ds, dy). \quad (3.14)$$

Dans ce résumé, nous nous focaliserons sur le résultat principal de l'article qui est l'estimation du paramètre de dérive à partir des variations quadratiques de la solution de l'équation stochastique des ondes (les résultats concernant le paramètre de diffusion sont obtenus de manière similaire). Nous noterons u_1 la solution (3.14) de l'équation (3.13) lorsque $\theta = \sigma = 1$. Nous appellerons $T_{N,x}$ les variations quadratiques temporelles associées au processus gaussien centré $(u_1(t, x))_{t \geq 0}$ que nous définissons pour $x \in \mathbb{R}$ fixé comme,

$$T_{N,x}(u_1) := \sum_{i=0}^{N-1} (u_1(t_{i+1}, x) - u_1(t_i, x))^2,$$

où $t_i = A_1 + \frac{i}{N}(A_2 - A_1)$ correspond à une subdivision de l'intervalle $[A_1, A_2]$.

Comportement asymptotique des variations quadratiques temporelles $T_{N,x}$

La solution u_1 peut s'écrire comme une intégrale stochastique d'ordre un, notée I_1 , et pour $i = 0, \dots, N-1$, nous pouvons écrire que,

$$u_1(t, x_{i+1}) - u_1(t, x_i) = I_1(g_{x,i}),$$

avec $s > 0, y \in \mathbb{R}$,

$$g_{x,i}(s, y) = G_1(t_{i+1} - s, x - y) \mathbf{1}_{(0, t_{i+1})}(s) - G_1(t_i - s, x - y) \mathbf{1}_{(0, t_i)}(s). \quad (3.15)$$

De cette manière, en appliquant la formule produit (1.1) aux variations quadratiques temporelles, nous obtenons que

$$T_{N,x}(u_1) = I_2 \left(\sum_{i=0}^{N-1} g_{x,i}^{\otimes 2} \right) + \mathbb{E}[T_{N,x}(u_1)],$$

où I_2 est une intégrale multiple stochastique d'ordre deux. Une étude préalable concernant la covariance de la solution spatiale nous a permis de montrer que le processus $(u_1(t, x))_{t \geq 0}$ avait la

même loi, au sens des distributions finies dimensionnelles que le processus $(\frac{1}{2}W_{t^2})_{t \geq 0}$, où $(W_t)_{t \geq 0}$ est un processus de Wiener. En particulier, pour $x \in \mathbb{R}$ le processus $(u_1(t, x))_{t \geq 0}$ est auto-similaire d'indice égale à 1 et il est à accroissements indépendants. La variance de ses accroissements étant égale pour tout $0 \leq s \leq t$ à,

$$\mathbb{E}(u_1(t, x) - u_1(s, x))^2 = \frac{1}{4}(t^2 - s^2),$$

nous pouvons facilement en déduire que

$$\mathbb{E}[T_{N,x}(u_1)] = \sum_{i=0}^{N-1} \mathbb{E}(u_1(t_{i+1}, x) - u_1(t_i, x))^2 = \sum_{i=0}^{N-1} \frac{1}{4}(t_{i+1}^2 - t_i^2) = \frac{1}{4}(A_2^2 - A_1^2).$$

Dès à présent, nous voudrions montrer que les variations quadratiques temporelles convergent dans $L^2(\Omega)$ et presque sûrement vers $\frac{1}{4}(A_2^2 - A_1^2)$ quand $N \rightarrow \infty$. Le fait que pour tout $i = 0, 1, \dots, N-1$, les accroissements $u_1(t_{i+1}, x) - u_1(t_i, x)$ suivent une loi gaussienne $\mathcal{N}(0, \frac{1}{4}(t_{i+1}^2 - t_i^2))$, nous donne que,

$$\begin{aligned} \mathbb{E} \left| T_{N,x}(u_1) - \frac{1}{4}(A_2^2 - A_1^2) \right|^2 &= \text{Var}(T_{N,x}(u_1)) = \text{Var}(Z^2) \frac{1}{16} \sum_{i=0}^{N-1} (t_{i+1}^2 - t_i^2)^2 \\ &\leq C \sum_{i=0}^{N-1} (t_{i+1} - t_i)^2 \leq C \frac{1}{N} \xrightarrow{N \rightarrow \infty} 0, \end{aligned}$$

où $Z \sim \mathcal{N}(0, 1)$. Nous obtenons alors la proposition suivante.

Proposition 11. *Pour tout $x \in \mathbb{R}$, la suite $(T_{N,x}(u_1), N \geq 1)$ converge, quand $N \rightarrow \infty$, vers $\frac{1}{4}(A_2^2 - A_1^2)$ dans $L^2(\Omega)$ et presque sûrement.*

L'utilisation de la propriété d'hypercontractivité des chaos (1.3) et le lemme de Borel-Cantelli engendrent la convergence presque sûre. Cette propriété d'hypercontractivité implique aussi la convergence dans $L^p(\Omega)$, pour tout $p \geq 2$.

Désormais, nous voudrions appliquer le théorème 9, du quatrième moment, aux variations quadratiques temporelles. Nous devons alors effectuer une renormalisation de $(T_{N,x}(u_1))_{N \geq 1}$, de sorte qu'elles puissent s'écrire comme une suite d'intégrales stochastiques d'ordre $q = 2$, de variance finie. Nous posons,

$$\begin{aligned} \tilde{T}_{N,x}(u_1) &= \sqrt{N} \left[\sum_{i=0}^{N-1} (u_1(t_{i+1}, x) - u_1(t_i, x))^2 - \frac{1}{4}(A_2^2 - A_1^2) \right] \\ &= I_2(f_N) \quad \text{avec } f_N = \sqrt{N} \sum_{i=0}^{N-1} g_{x,i}^{\otimes 2}, \end{aligned} \tag{3.16}$$

où les $g_{x,i}$ sont données par (3.15). Ainsi, nous obtenons que $\tilde{T}_{N,x}$ est centré et de variance telle que,

$$\mathbf{E}[\tilde{T}_{N,x}^2] \xrightarrow{N \rightarrow \infty} \frac{(A_2^3 - A_1^3)}{6}(A_2 - A_1) := K(A_1, A_2) < \infty. \tag{3.17}$$

Nous montrons alors que $\|D\tilde{T}_{N,x}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}$ converge dans $L^2(\Omega)$ quand $N \rightarrow \infty$ vers $2K(A_1, A_2)$ et par conséquent, nous obtenons la convergence en loi de la suite $(\tilde{T}_{N,x}(u_1))_{N \geq 1}$ vers une loi gaussienne centrée de variance $K(A_1, A_2)$.

Théorème 15. La suite $(\tilde{T}_{N,x}(u_1))_{N \geq 1}$, définie par (3.16) converge en loi quand $N \rightarrow \infty$, vers $\mathcal{N}(0, K(A_1, A_2))$, avec $K(A_1, A_2)$ donnée par (3.17), et pour N assez grand,

$$d_W \left(\tilde{T}_{N,x}(u), \mathcal{N}(0, K(A_1, A_2)) \right) \leq C \frac{1}{\sqrt{N}}.$$

Remarque 18. Nous nous intéressons également à l'optimalité du taux de convergence des variations quadratiques temporelles vers la loi gaussienne. L'application du théorème 1.9 de H. Biermé, A. Bonami, I. Nourdin et G. Peccati [16] permet de prouver que le taux optimal reste de l'ordre de $N^{-\frac{1}{2}}$.

Estimation du paramètre de dérive

Notons désormais u_θ la solution (3.14) de l'équation (3.13) lorsque le paramètre de diffusion $\sigma = 1$. L'objectif est donc de construire un estimateur fortement consistant de θ , basé sur les variations quadratiques temporelles

$$T_{N,x}(u_\theta) := \sum_{i=0}^{N-1} (u_\theta(t_{i+1}, x) - u_\theta(t_i, x))^2.$$

Pour obtenir cet estimateur, nous voulons établir la convergence presque sûre de ces variations quadratiques en nous appuyant sur la proposition 11. Nous considérons la transformation suivante, pour tout $(t, x) \in [0, \infty) \times \mathbb{R}$,

$$v_\theta(t, x) = u_\theta \left(\frac{t}{\sqrt{\theta}}, x \right).$$

v_θ est alors solution de

$$\frac{\partial}{\partial t} v_\theta(t, x) = \Delta v_\theta(t, x) + \theta^{-\frac{1}{4}} \tilde{W}(t, x)$$

avec \tilde{W} un bruit blanc en temps et en espace. Nous pouvons alors remarquer que $(\theta^{\frac{1}{4}} v_\theta(t, x), t \geq 0, x \in \mathbb{R})$ et $(u_1(t, x), t \geq 0, x \in \mathbb{R})$ ont même loi au sens des distributions finies dimensionnelles. Nous avons,

$$\begin{aligned} V_{N,x}(u_\theta) &= \sum_{i=0}^{N-1} \left(v_\theta(\sqrt{\theta} t_{i+1}, x) - v_\theta(\sqrt{\theta} t_i, x) \right)^2 \\ &= \theta^{-\frac{1}{2}} \sum_{i=0}^{N-1} \left(\theta^{\frac{1}{4}} v_\theta(\sqrt{\theta} t_{i+1}, x) - \theta^{\frac{1}{4}} v_\theta(\sqrt{\theta} t_i, x) \right)^2 \\ &\stackrel{(d)}{=} \theta^{-\frac{1}{2}} \sum_{i=0}^{N-1} \left(\tilde{u}_1(\sqrt{\theta} t_{i+1}, x) - \tilde{u}_1(\sqrt{\theta} t_i, x) \right)^2, \end{aligned}$$

où

$$\tilde{u}_1(t, x) = \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y) \tilde{W}(ds, dy).$$

Comme nous savons que $(\sqrt{\theta} t_i, i = 0, \dots, N)$ est une partition de l'intervalle $[\sqrt{\theta} A_1, \sqrt{\theta} A_2]$, nous obtenons grâce à la proposition 11,

$$\sum_{i=0}^{N-1} \left(\tilde{u}_1(\sqrt{\theta} t_{i+1}, x) - \tilde{u}_1(\sqrt{\theta} t_i, x) \right)^2 \xrightarrow{N \rightarrow \infty} \frac{1}{4} \left((\sqrt{\theta} A_2)^2 - (\sqrt{\theta} A_1)^2 \right) = \frac{1}{4} \theta (A_2^2 - A_1^2).$$

Proposition 12. Soit $t_i = A_1 + \frac{i}{N}(A_2 - A_1)$ une partition de l'intervalle $[A_1, A_2]$ avec $i = 0, 1, \dots, N$. Soit u_θ définie par (3.14) pour $\sigma = 1$. Alors pour $x \in \mathbb{R}$ fixé,

$$V_{N,x}(u_\theta) \xrightarrow{N \rightarrow \infty} \frac{1}{4}(A_2^2 - A_1^2)\theta^{\frac{1}{2}} \text{ presque sûrement et dans } L^2(\Omega).$$

Par conséquent, nous obtenons une estimation du paramètre de dérive :

$$\hat{\theta}_N := \left(\frac{4V_{N,x}(u_\theta)}{A_2^2 - A_1^2} \right)^2, \quad (3.18)$$

qui est un estimateur fortement consistant de θ .

Remarque 19. À l'aide du théorème 15 et de la méthode delta, nous fournissons un TCL pour l'estimateur, en particulier, nous montrons que $\sqrt{N}(\hat{\theta}_N - \theta)$ converge vers une variable aléatoire gaussienne centrée.

Une estimation du paramètre de diffusion σ basée sur les variations quadratiques temporelles est aussi possible. Pour cela, il suffit de considérer la solution (3.14) de l'équation (3.13) lorsque $\theta = 1$, que nous notons $(u_\sigma(t, x), t > 0, x \in \mathbb{R})$, et de remarquer que pour tout $t \geq 0$ et $x \in \mathbb{R}$, $u_\sigma(t, x) = \sigma u_1(t, x)$. En se basant sur la proposition 11, nous parvenons à obtenir une convergence presque sûre pour les variations quadratiques temporelles de la solution u_σ . Finalement, il en découle que

$$\hat{\sigma}_N^2 := \left(\frac{1}{4}(A_2^2 - A_1^2) \right)^{-1} \sum_{i=0}^{N-1} (u_\sigma(t_{i+1}, x) - u_\sigma(t_i, x))^2$$

est un estimateur fortement consistant de σ^2 .

Remarque 20. Lorsque nous effectuons un travail similaire pour analyser le comportement asymptotique des variations quadratiques spatiales de la solution $(u_\theta(t, x), x \in \mathbb{R})$, nous remarquons que la limite ne dépend pas de θ . L'estimation du paramètre de dérive basée sur les variations quadratiques spatiales n'étant alors pas possible, nous avons trouvé une estimation du paramètre de diffusion à partir de celles-ci, de la même manière que dans le cas temporel.

3.3.2 Résumé du chapitre 7

Dans cette section, nous vous proposons un résumé de l'article [55]. Ce papier est orienté sur l'estimation du paramètre de dérive de l'équation de la chaleur fractionnaire, dirigée par un bruit gaussien non linéaire, blanc en temps et en espace. L'estimation de ce paramètre sera basée sur les variations quadratiques spatiales de la solution. De cette manière, nous sommes conduit à les analyser minutieusement et à établir leur convergence dans $L^1(\Omega)$. Pour effectuer cette étude, nous nous appuyerons sur une relation établie entre le mouvement brownien fractionnaire et la solution spatiale de l'équation considérée.

Il existe différents résultats dans la littérature traitant de la construction d'un estimateur pour le paramètre de dérive. Par exemple, dans [104] les auteurs J. Pospisil et R. Tribe fournissent une estimation de ce paramètre, dans le cas où l'équation de la chaleur est dirigée par un bruit non linéaire, blanc en temps et en espace. En 2019, Z. Khalil-Mahdi et C. Tudor considèrent l'équation de la chaleur fractionnaire dirigée par un bruit gaussien additif [76]. Quant au cas non linéaire de l'équation de la chaleur fractionnaire, les auteurs R. Dalang et C. Chong [26] ont récemment étudié les p -variations associées à la solution temporelle.

Dans un premier temps, nous rappelons que l'équation concernée est la suivante :

$$\begin{cases} \frac{\partial}{\partial t} u_\theta(t, x) = -\theta(-\Delta)^{\frac{\alpha}{2}} u_\theta(t, x) + \sigma(u_\theta(t, x)) \dot{W}(t, x), & t \in [0, T], x \in \mathbb{R}, \\ u_\theta(0, x) = 0, & \text{pour tout } x \in \mathbb{R}, \end{cases} \quad (3.19)$$

où $\theta > 0$ est le fameux paramètre de dérive que nous allons estimer. La perturbation aléatoire W dans l'équation (3.19) est un bruit blanc en temps et en espace tandis que $(-\Delta)^{\frac{\alpha}{2}}$ correspond au laplacien fractionnaire d'ordre $\alpha \in (1, 2]$. Il est déterminé par la définition 25.

La solution de l'équation (3.19) est donnée pour $t \in [0, T]$ et $x \in \mathbb{R}$ par

$$u_\theta(t, x) = \int_0^t \int_{\mathbb{R}} G_\alpha(\theta(t-s), x-y) \sigma(u_\theta(s, y)) W(ds, dy), \quad (3.20)$$

où l'intégrale stochastique par rapport à $W(ds, dy)$ est une intégrale de Walsh-Dalang et G_α est le noyau de Green, correspondant à la solution fondamentale de l'équation de la chaleur fractionnaire. Ses propriétés nous seront très utiles pour établir les résultats.

L'objectif est d'analyser les variations quadratiques de la solution (3.20) afin d'établir une convergence dans $L^1(\Omega)$ et dont la limite contiendra le paramètre θ . Pour cela, nous proposons d'étudier les variations quadratique de u_θ lorsque $\sigma \equiv 1$ et $\theta = 1$. Nous utiliserons ensuite un argument d'approximation entre le cas linéaire et non linéaire, pour parvenir à nos fins.

Nous notons u_0 , la solution de l'équation de la chaleur fractionnaire dirigée par un bruit blanc en temps et en espace, lorsque $\theta = 1$ et $\sigma \equiv 1$ dans (3.20). Par abus de langage, nous l'appellerons simplement, solution linéaire. Nous désignons par $V_{N,t}$, les variations quadratiques spatiales, en particulier pour $t > 0$ fixé,

$$V_{N,t}(u_0) = N^{\alpha-2} \sum_{i=0}^{N-1} (u_0(t, x_{i+1}) - u_0(t, x_i))^2, \quad (3.21)$$

où

$$x_i = \frac{i}{N} \text{ pour } i = 0, 1, \dots, N. \quad (3.22)$$

Nous analysons le comportement asymptotique de $(V_{N,t})_{N \geq 0}$.

Étape 1 : Décomposition de la solution linéaire

Un article de M. Foodun, D. Khoshnevisan et P. Mahboubi [51] utilise le fait que $(u_0(t, x))_{x \in \mathbb{R}}$ est un mouvement brownien fractionnaire perturbé. Comme nous l'avons déjà évoqué dans la remarque 7, nous pouvons décomposer la solution comme un fBm et un processus gaussien centré dont les trajectoires sont \mathcal{C}^∞ .

— Pour tout $t \in (0, T]$, le processus gaussien centré $(u_0(t, x))_{x \in \mathbb{R}}$ s'écrit comme

$$u_0(t, x) \stackrel{(d)}{=} m_\alpha U(t, x) + Y(t, x), \quad x \in \mathbb{R}. \quad (3.23)$$

- $(U(t, x))_{x \in \mathbb{R}}$ est le mouvement brownien fractionnaire de paramètre de Hurst $H = \frac{\alpha-1}{2}$.
- m_α est une constante dépendant de $\alpha \in (1, 2]$ telle que,

$$m_\alpha := \left(2\Gamma(\alpha) \left| \cos\left(\frac{\alpha\pi}{2}\right) \right| \right)^{-\frac{1}{2}}.$$

— $(Y(t, x))_{x \in \mathbb{R}}$ est un processus gaussien satisfaisant

$$\mathbb{E}|Y(t, x) - Y(t, y)|^2 \leq C_T |x - y|^2 \quad (3.24)$$

pour tout $t \in [0, T]$ et $x, y \in \mathbb{R}$.

Le point de départ est alors de regarder la convergence des variations quadratiques de la solution $(u_0(t, x))_{x \in \mathbb{R}}$ en utilisant celle du fBM.

Etape 2 : Etude des variations quadratiques de la solution linéaire

Nous disposons du résultat suivant.

Proposition 13. *Soit $(B_t^H)_{t \geq 0}$ un mouvement brownien fractionnaire de paramètre de Hurst $H \in (0, \frac{3}{4})$. Pour $N \geq 1$, notons*

$$V_N(B^H) = N^{2H-1} \sum_{i=0}^{N-1} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2.$$

Alors la suite $(V_N(B^H), N \geq 1)$ converge vers 1 quand $N \rightarrow \infty$ dans $L^2(\Omega)$.

Remarque 21. *Le fait de prendre $H \in (0, \frac{3}{4})$ est important puisque notre résultat découle du fait que la fonction de covariance du fBm appartient à $l^2(\mathbb{N})$ si et seulement si $H < \frac{3}{4}$.*

Nous savons d'après (3.23) que $(u_0(t, x))_{x \in \mathbb{R}}$ se décompose comme un fBm et un processus gaussien aux trajectoires régulières. En particulier, nous pouvons réécrire $(V_{N,t})_{N \geq 0}$ de la manière suivante

$$\begin{aligned} V_{N,t}(u_0) &\stackrel{(d)}{=} m_\alpha^2 V_N(U(t, x)) + 2m_\alpha N^{2H-1} \sum_{i=0}^{N-1} (U(t, x_{i+1}) - U(t, x_i))(Y(t, x_{i+1}) - Y(t, x_i)) \\ &\quad + N^{2H-1} \sum_{i=0}^{N-1} (Y(t, x_{i+1}) - Y(t, x_i))^2. \end{aligned}$$

Par la proposition 13, nous avons directement que $m_\alpha^2 V_N(U(t, x))$ converge vers m_α^2 dans $L^2(\Omega)$. Les autres termes convergent vers 0 dans $L^1(\Omega)$ quand $N \rightarrow \infty$. En effet, en utilisant (3.24),

$$\begin{aligned} &\mathbb{E} \left| N^{2H-1} \sum_{i=0}^{N-1} (U(t, x_{i+1}) - U(t, x_i))(Y(t, x_{i+1}) - Y(t, x_i)) \right| \leq CN^{2H-1} \sum_{i=0}^{N-1} \frac{1}{N^{2+H}} \\ &\leq C \frac{1}{N^{2-H}} \xrightarrow{N \rightarrow \infty} 0 \end{aligned}$$

et

$$\mathbb{E} \left| N^{2H-1} \sum_{i=0}^{N-1} (Y(t, x_{i+1}) - Y(t, x_i))^2 \right| \leq CN^{2H-1} \sum_{i=0}^{N-1} \frac{1}{N^2} \leq C \frac{1}{N^{2-2H}} \xrightarrow{N \rightarrow \infty} 0,$$

puisque $H \in (0, \frac{3}{4})$. Il en découle la proposition 14.

Proposition 14. *Considérons la suite $(V_{N,t}(u_0), N \geq 1)$ donnée par (3.21) avec $t \in [0, T]$ fixé. Alors*

$$V_{N,t}(u_0) \xrightarrow{N \rightarrow \infty} m_\alpha^2 \text{ in } L^1(\Omega),$$

où m_α est la constante apparaissant dans la décomposition (3.23).

Etape 3 : Etude des variations quadratiques de la solution non linéaire

Désormais, nous désignons par $u(t, x)$ la solution (3.20) où $\theta = 1$, et nous l'appelons solution non linéaire, par abus de langage. Nous cherchons maintenant à étudier la convergence des variations quadratiques de $(u(t, x))_{x \in \mathbb{R}}$. Pour cela, l'idée est d'approximer les accroissements $u(t, x_{i+1}) - u(t, x_i)$ de la solution non linéaire par $\sigma(u(t(\delta), x_i))(u_0(t, x_{i+1}) - u_0(t, x_i))$, où $\delta = \frac{1}{N}$ et $t(\delta)$ est un point proche de t , tel que $t - t(\delta) = \delta^\gamma$ avec γ à spécifier (cf. Proposition 15). Nous voulons montrer que la suite $(V_{N,t}(u))_{N \geq 0}$ (donnée par (3.21) en remplaçant u_0 par u) converge quand $N \rightarrow \infty$ dans $L^1(\Omega)$ vers

$$m_\alpha^2 \int_0^1 \sigma^2(u(t, x)) dx.$$

Nous effectuons alors une décomposition sous la forme suivante ;

$$\begin{aligned} V_{N,t}(u) - m_\alpha^2 \int_0^1 \sigma^2(u(t, x)) dx &= N^{\alpha-2} \sum_{i=0}^{N-1} \left[(u(t, x_{i+1}) - u(t, x_i))^2 - \sigma^2(u(t(\delta), x_i)) (\tilde{\Delta}u_0)^2(x_i, \delta) \right] \\ &\quad + \sum_{i=0}^{N-1} \sigma^2(u(t(\delta), x_i)) \left(N^{\alpha-2} (\tilde{\Delta}u_0)^2(x_i, \delta) - m_\alpha^2 \frac{1}{N} \right) \\ &\quad + m_\alpha^2 \left[\frac{1}{N} \sum_{i=0}^{N-1} \sigma^2(u(t(\delta), x_i)) - \int_0^1 \sigma^2(u(t, x)) dx \right] \\ &:= B_{1,N} + B_{2,N} + B_{3,N}. \end{aligned}$$

La notation $(\tilde{\Delta}u_0)(x, \delta)$ correspond à une modification des accroissements spatiaux de la solution u_0 ,

$$\begin{aligned} (\tilde{\Delta}u_0)(x, \delta) &= \int_{t(\delta)}^t \int_{\mathbb{R}} (G_\alpha(t-s, x + \delta - y) - G_\alpha(t-s, x - y)) W(ds, dy) \\ &\quad + \int_0^{t(\delta)} \int_{\mathbb{R}} (G_\alpha(t-s, x + \delta - y) - G_\alpha(t-s, x - y)) \tilde{W}(ds, dy), \end{aligned}$$

avec \tilde{W} , une copie indépendante du bruit blanc en temps et en espace W . Pour établir la convergence dans $L^1(\Omega)$, il faut montrer que les termes $\mathbb{E}|B_{1,N}|$, $\mathbb{E}|B_{2,N}|$ et $\mathbb{E}|B_{3,N}|$ tendent vers 0 quand $N \rightarrow \infty$. Nous remarquons que $B_{3,N}$ va converger presque sûrement vers 0 quand $N \rightarrow \infty$ par l'approximation de la somme de Riemann. Nous montrons aussi que ce terme converge dans $L^1(\Omega)$.

Pour traiter le cas $\mathbb{E}|B_{1,N}|$ nous utilisons l'inégalité de Cauchy–Schwarz et la proposition suivante.

Proposition 15. *Pour $t > 0, x \in \mathbb{R}$, et $\delta > 0$ assez petit, nous avons*

$$\mathbb{E} \left| (u(t, x + \delta) - u(t, x)) - \sigma(u(t(\delta), x)) (\tilde{\Delta}u_0)(x, \delta) \right|^2 \leq C_T \delta^{\frac{4(\alpha-1)}{\alpha+1}}. \quad (3.25)$$

En fait, la valeur de $t(\delta)$ est choisie de sorte que la borne (3.25) soit optimale. Pour ce faire,

nous devons écrire, $(u(t, x + \delta) - u(t, x)) - \sigma(u(t(\delta), x))(\tilde{\Delta}u_0)(x, \delta)$ comme

$$\begin{aligned}
 & (u(t, x + \delta) - u(t, x)) - \sigma(u(t(\delta), x))(\tilde{\Delta}u_0)(x, \delta) \\
 = & \int_{t(\delta)}^t \int_{\mathbb{R}} (G_\alpha(t-a, x + \delta - y) - G_\alpha(t-a, x - y)) (\sigma(u(a, y)) - \sigma(u(t(\delta), x))) W(da, dy) \\
 & + \int_0^{t(\delta)} \int_{\mathbb{R}} (G_\alpha(t-a, x + \delta - y) - G_\alpha(t-a, x - y)) \sigma(u(a, y)) W(da, dy) \\
 & - \int_0^{t(\delta)} \int_{\mathbb{R}} (G_\alpha(t-a, x + \delta - y) - G_\alpha(t-a, x - y)) \sigma(u(t(\delta), x)) \tilde{W}(da, dy) \\
 := & A_1(x, \delta) + A_2(x, \delta) + A_3(x, \delta).
 \end{aligned}$$

Grâce aux propriétés du noyau G_α , nous obtenons que,

$$\mathbb{E}[A_1(x, \delta)^2] \leq C\delta^{\gamma(2-\frac{2}{\alpha})},$$

$$\mathbb{E}[A_2(x, \delta)^2] \leq C_T \delta^2 \delta^{-\gamma\frac{3-\alpha}{\alpha}},$$

et

$$\mathbb{E}[A_3(x, \delta)^2] \leq C_T \delta^2 \delta^{-\gamma\frac{3-\alpha}{\alpha}}.$$

On cherche la valeur de γ pour laquelle la borne (3.25) soit optimale. Il s'agit simplement de résoudre l'équation $\gamma(2-\frac{2}{\alpha}) = 2 - \gamma\frac{3-\alpha}{\alpha}$. De cette manière $\gamma = \frac{2\alpha}{\alpha+1}$. Et dans ce cas, nous avons

$$t(\delta) = t - \left(\frac{1}{N}\right)^\gamma \text{ avec } \gamma = \frac{2\alpha}{\alpha+1}.$$

Finalement, l'indépendance entre $(\tilde{\Delta}u_0)(x, \delta)$ et la σ -algèbre engendrée par $\{u_s, 0 \leq s \leq t(\delta)\}$ induit que,

$$\mathbb{E}|B_{1,N}| \leq CN^{\frac{(\alpha-1)(\alpha-3)}{2(\alpha+1)}} \xrightarrow{N \rightarrow \infty} 0.$$

En ce qui concerne le terme $B_{2,N}$, nous devons souligner que la convergence dans $L^1(\Omega)$ est obtenue par le biais de l'analyse des variations quadratiques du fBm et en particulier du fait que

$$\mathbb{E} \left[N^{2H-1} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2 - \frac{1}{N} \right]^2 = \frac{2}{N^2},$$

avec $(B_t^H)_{t \geq 0}$ un mouvement brownien fractionnaire de paramètre de Hurst $H \in (0, 1)$.

Par conséquent, il en découle le théorème suivant.

Théorème 16. *Soit $t \in (0, T]$. Considérons $(u(t, x))_{x \in \mathbb{R}}$ et notons $V_{N,t}(u)$ ses variations quadratiques. Alors pour tout $t > 0$,*

$$V_{N,t}(u) \xrightarrow{N \rightarrow \infty} m_\alpha^2 \int_0^1 \sigma^2(u(t, x)) dx \text{ dans } L^1(\Omega),$$

et pour N assez grand,

$$\mathbb{E} \left| V_{N,t}(u) - m_\alpha^2 \int_0^1 \sigma^2(u(t, x)) dx \right| \leq C \left(\frac{1}{N} \right)^{\frac{(\alpha-1)(3-\alpha)}{2(\alpha+1)}},$$

où la constante $C > 0$ ne dépend pas de t .

Etape 4 : Estimation du paramètre

Nous considérons désormais l'équation (3.19). Grâce au théorème 16 et par une manipulation de la solution u_θ donnée par (3.20), nous obtenons la convergence des variations quadratiques associées à la solution $(u_\theta(t, x))_{x \in \mathbb{R}}$ dans $L^1(\Omega)$.

Proposition 16. *Soit u_θ donnée par (3.20) et notons $V_{N,t}(u_\theta)$, les variations quadratiques données par (3.21), associées à u_θ . Alors, nous obtenons que*

$$V_{N,t}(u_\theta) \xrightarrow[N \rightarrow \infty]{} \theta^{-1} m_\alpha^2 \int_0^1 \sigma^2(u_\theta(t, x)) dx \text{ dans } L^1(\Omega),$$

et pour N assez grand,

$$\mathbf{E} \left| V_{N,t}(u_\theta) - \theta^{-1} m_\alpha^2 \int_0^1 \sigma^2(u_\theta(t, x)) dx \right| \leq C \left(\frac{1}{N} \right)^{\frac{(\alpha-1)(3-\alpha)}{2(\alpha+1)}}.$$

Puisque la limite dépend de θ , nous proposons de construire l'estimateur $\hat{\theta}_N$ tel que,

$$\hat{\theta}_N = \frac{m_\alpha^2}{N} \cdot \frac{\sum_{i=0}^{N-1} \sigma^2(u_\theta(t, x_i))}{V_{N,t}(u_\theta)}, \quad N \geq 1, \quad t \in [0, T].$$

Par la proposition 16, nous avons que

$$\frac{m_\alpha^2 \int_0^1 \sigma^2(u_\theta(t, x)) dx}{V_{N,t}(u_\theta)} \xrightarrow[N \rightarrow \infty]{L^1(\Omega)} \theta,$$

et nous décomposons l'estimateur de sorte que

$$\begin{aligned} \hat{\theta}_N &= \frac{m_\alpha^2 \left(\frac{1}{N} \sum_{i=0}^{N-1} \sigma^2(u_\theta(t, x_i)) - \int_0^1 \sigma^2(u_\theta(t, x)) dx \right)}{V_{N,t}(u_\theta)} \\ &\quad + \frac{m_\alpha^2 \int_0^1 \sigma^2(u_\theta(t, x)) dx}{V_{N,t}(u_\theta)}. \end{aligned}$$

La première somme converge presque sûrement vers 0 quand $N \rightarrow \infty$ par approximation d'une somme de Riemann. Ainsi $\hat{\theta}_N$ converge en probabilité vers θ . C'est un estimateur consistant du paramètre de dérive.

Remarque 22. *Dans notre travail, nous avons aussi cherché à construire un estimateur basé sur les variations quadratiques spatiales en moyenne temporelle. Nous obtenons le même type de résultat que dans la proposition 16, ce qui nous permet d'obtenir un autre estimateur consistant de θ lorsque nous disposons des observations à différents instants t .*

Matrices aléatoires

La majeure partie de cette thèse est consacrée à l'analyse du comportement des matrices aléatoires de grande taille. Nous proposons de fournir une présentation assez brève des résultats fondamentaux se trouvant dans la littérature. Nous examinerons, ensuite plus en détails les résultats liés à la méthode de Stein-Malliavin.

4.1 Introduction aux matrices aléatoires

De nos jours, l'étude des matrices aléatoires constitue un important sujet de recherche dû à l'avènement de l'informatique et des masses de données de grande dimension (le big-data). Elle s'est principalement développée pour ses nombreux champs d'application, en finance, physique théorique, télécommunication, statistique, mais aussi pour des problèmes mathématiques, en théorie des graphes par exemple.

L'analyse des matrices aléatoires prend racine, bien plus tôt, dans les travaux [127] du statisticien John Wishart en 1928. Il initie le sujet par l'étude des matrices de covariance empirique dont la taille est fixe, et dont les éléments sont gaussiens.

Plus tardivement, dans les années cinquante, cette théorie connaît un nouvel essor grâce aux recherches en physique nucléaire [126] du physicien Eugène Wigner. Ce dernier s'est intéressé à la spectroscopie des atomes lourds. En particulier, les niveaux d'énergie atomique sont les valeurs propres d'un opérateur hermitien sur un espace de Hilbert, à savoir, l'hamiltonien du système. Ces niveaux d'énergie ne pouvant pas être étudiés directement lorsque le système atomique est complexe, Wigner a eu l'idée de modéliser cet hamiltonien par une matrice hermitienne (GUE). Cela lui a permis de comprendre les propriétés statistiques de cet opérateur et d'introduire un théorème qui portera son nom, que nous présenterons dans la partie suivante.

Une nouvelle impulsion à la théorie des matrices aléatoires de grande taille est donnée en 1967 par Marchenko et Pastur dans [89], en fixant la condition que lorsque le nombre de lignes et de colonnes de la matrice sont du même ordre, la mesure spectrale empirique des matrices de covariance de Wishart suit une loi de Marchenko-Pastur.

Dans un premier temps, nous nous intéresserons aux éléments que nous venons d'évoquer puis, nous verrons comment la théorie a évolué ces dernières années, pour parvenir à l'étude des

matrices de Wishart par la méthode de Malliavin-Stein dont les résultats sont nombreux. Ce sujet, largement étudié au cours de cette thèse, sera généralisé sur les tenseurs aléatoires.

4.1.1 Approche spectrale pour l'étude des matrices aléatoires

Nous proposons d'énoncer le théorème de Wigner. Il concerne les matrices symétriques de dimension $n \times n$, dont tous les coefficients sur et en dehors de la diagonale sont indépendants et de même loi. Notons que le théorème s'étend aux coefficients complexes et dans cette situation, on supposera que la matrice est hermitienne pour obtenir la symétrie.

Une telle matrice est appelée matrice de Wigner. Nous fournissons une définition de celle-ci dans le cas réel.

Définition 27 (Matrice de Wigner). *Soit $\mathbf{G}_n = (G_{ij})_{1 \leq i, j \leq n}$ une matrice symétrique de taille n . \mathbf{G}_n est une matrice de Wigner réelle si les coefficients sont des variables aléatoires réelles, centrées, indépendantes et telles que :*

- Les variables aléatoires $(G_{ij})_{1 \leq i < j \leq n}$ sont i.i.d suivant une loi de probabilité ν définie sur \mathbb{R} .
- Les termes diagonaux $(G_{ii})_{1 \leq i \leq n}$ sont des variables aléatoires indépendantes et de même loi μ définie sur \mathbb{R} .

Une classe particulière de matrices de Wigner sera utilisée au cours de cette thèse ; ce sont les matrices possédant des propriétés d'invariance en loi par l'action du groupe orthogonal $\mathcal{O}_n(\mathbb{R})$. En physique, Wigner utilisait ces matrices symétriques lorsque l'hamiltonien du système atomique était stable par renversement du temps et par rotation. Cet ensemble est appelé Ensemble Gaussien Orthogonal (GOE).

Définition 28 (GOE). *Une matrice de Wigner (réelle) $\mathbf{G}_n = (G_{ij})_{1 \leq i, j \leq n}$ est une matrice appartenant à l'ensemble gaussien orthogonal $GOE(n, \sigma^2)$ si les coefficients sont des variables aléatoires indépendantes gaussiennes centrées telles que :*

$$\mathbf{G}_n = \begin{cases} G_{ii} \sim \mathcal{N}(0, 2\sigma^2) & \text{pour } i = 1, \dots, n, \\ G_{ij} \sim \mathcal{N}(0, \sigma^2) & \text{pour } 1 \leq i < j \leq n. \end{cases}$$

Remarque 23. *Il existe d'autres ensembles connus de matrices de Wigner, à savoir :*

- Si l'hamiltonien est seulement invariant par rotation, la matrice employée sera hermitienne dont la loi est invariante par l'action du groupe unitaire $\mathbf{U}(n)$, que l'on appellera GUE pour Ensemble Gaussien Unitaire.
- Si le système atomique est stable par renversement du temps mais pas par rotation, la matrice considérée sera une matrice quaternionique réelle dont la loi est invariante par l'action du groupe unitaire symplectique \mathbf{USp}_n . Cet ensemble est appelé Ensemble Gaussien Symplectique et noté par le sigle GSE.

Comme nous l'avons vu en introduction, Wigner a focalisé son analyse sur les valeurs propres de ces matrices. Il a ainsi étudié la distribution de celles-ci pour comprendre l'hamiltonien du système atomique. Afin de décrire mathématiquement la répartition des valeurs propres, nous introduisons la mesure spectrale empirique d'une matrice.

Définition 29 (Mesure spectrale empirique). *Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique de taille n , dont les valeurs propres sont notées $\lambda_1 \leq \dots \leq \lambda_n$. On définit la mesure spectrale de A par :*

$$\mu_A := \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i},$$

où δ est la mesure de Dirac. Pour $\Delta \subset \mathbb{R}$ on a,

$$\mu_A(\Delta) = \frac{1}{n} \text{Card}\{1 \leq i \leq n, \lambda_i \in \Delta\}.$$

Les principaux problèmes en théorie des matrices aléatoires ont donc été de s'interroger sur la convergence de cette mesure pour une famille de matrices aléatoires et d'analyser le comportement des valeurs propres extrêmes.

Wigner va alors démontrer un comportement limite non aléatoire de la mesure spectrale empirique de \mathbf{G}_n . Nous noterons, μ_{sc} la loi du demi-cercle, aussi appelée loi de Wigner dont la densité par rapport à la mesure de Lebesgue est donnée par :

$$d\mu_{sc}(x) = \frac{\sqrt{4-x^2}}{2\pi} \mathbf{1}_{[-2,2]}(x) dx.$$

Théorème 17 (Théorème de Wigner). *Considérons $\mathbf{G}_n := (G_{ij})_{1 \leq i \leq j \leq n}$ une matrice de Wigner donnée par la définition 27. On suppose que les termes non diagonaux ont une variance égale à 1, i.e : pour tout $1 \leq i < j \leq n$, $\mathbb{E}[G_{ij}^2] = 1$. Alors presque sûrement, la mesure spectrale $\mu_{\frac{\mathbf{G}_n}{\sqrt{n}}}$ converge étroitement vers la loi du demi cercle μ_{sc} . C'est-à-dire que, presque sûrement, pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ continue et bornée,*

$$\lim_{n \rightarrow \infty} \int f(x) d\mu_{\frac{\mathbf{G}_n}{\sqrt{n}}}(x) = \int f(x) d\mu_{sc}(x).$$

Le théorème de Wigner fournit une description macroscopique du spectre de la matrice \mathbf{G}_n . Naturellement, les chercheurs ont voulu obtenir des informations plus fines concernant les valeurs propres et en particulier le comportement asymptotique des valeurs propres extrémales. Les premiers à obtenir la convergence de la plus grande valeur propre ont été Füredi et Komlos [50] en utilisant la méthode des grandes traces. Les auteurs considéraient, par ailleurs, une hypothèse assez restrictive ; tous les moments des coefficients de la matrice de Wigner devaient être finis. Ce sont les chercheurs Baï et Yin [23] qui parviennent à relâcher cette condition, en considérant l'existence d'un moment d'ordre quatre, pour aboutir au résultat énoncé ci-dessous.

Théorème 18. *Considérons $\tilde{\mathbf{G}}_n = \frac{1}{\sqrt{n}} \mathbf{G}_n$, où $\mathbf{G}_n = (G_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ est une matrice de Wigner de taille n . On suppose que les coefficients $(G_{ii})_{1 \leq i \leq n}$ sont i.i.d centrés de variance finie et que les termes $(G_{ij})_{1 \leq i < j \leq n}$ sont i.i.d de loi μ telles que*

$$\int x\mu(dx) = 0, \int |x|^2\mu(dx) = \sigma^2 < \infty, \int |x|^4\mu(dx) < \infty.$$

Alors quand n tend vers l'infini, la plus grande valeur propre de $\tilde{\mathbf{G}}_n$, notée λ_{\max} converge presque sûrement vers 2σ .

L'approche pionnière de Wigner a été le point de départ du développement de la théorie des matrices aléatoires. Plus précisément, nous allons nous intéresser aux propriétés spectrales des matrices de covariance empirique. Ce sujet particulièrement essentiel en statistique multivariée mais aussi en télécommunications a permis d'établir de nombreux tests statistiques.

L'étude spectrale des matrices de Wishart a été initiée par les travaux de Marchenko et Pastur [89]. Sous la condition d'un régime spécial entre les lignes et les colonnes de la matrice initiale associée à la matrice de Wishart, ils ont obtenu la loi limite de la mesure spectrale empirique de ces matrices de covariance empirique. Avant d'énoncer ce théorème, nous définissons les éléments évoqués.

Définition 30 (Matrice de Wishart). Soit $\mathcal{X}_{n,d} := (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ une matrice rectangulaire de taille $n \times d$ dont les coefficients sont des variables aléatoires réelles. On appelle matrice de Wishart associée à $\mathcal{X}_{n,d}$, la matrice notée :

$$\mathcal{W}_{n,d} := \frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T.$$

$\mathcal{X}_{n,d}^T$ correspond à la matrice transposée de $\mathcal{X}_{n,d}$.

Remarque 24. 1. Une matrice de Wishart est la résultante du produit matriciel de \mathcal{X} avec sa transposée ainsi la matrice obtenue est symétrique de taille n dont les coefficients sont déterminés par

$$\mathcal{W}_{n,d} := \begin{cases} W_{ii} = \frac{1}{d} \sum_{k=1}^d X_{ik}^2 & \text{pour } i = 1, \dots, n, \\ W_{ij} = \frac{1}{d} \sum_{k=1}^d X_{ik} X_{jk} & \text{pour } 1 \leq i < j \leq n. \end{cases} \quad (4.1)$$

Ces composantes correspondent, en réalité, à la variance empirique des variables aléatoires pour les termes diagonaux et la covariance empirique des termes situés en dehors de la diagonale. Ainsi ces matrices de Wishart sont des matrices de covariance empirique.

2. $\mathcal{W}_{n,d}$ possède n valeurs propres que l'on notera par ordre croissant $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$.

Pour $c > 0$, notons $\mu_{MP}^{(c)}$ la loi de Marchenko-Pastur définie par :

$$\mu_{MP}^{(c)}(dx) = \max \left\{ 1 - \frac{1}{c}, 0 \right\} \delta_0 + \frac{\sqrt{(x - (1 - \sqrt{c})^2)((1 + \sqrt{c})^2 - x)}}{2\pi c x} \mathbf{1}_{[(1 - \sqrt{c})^2, (1 + \sqrt{c})^2]}(x) dx.$$

Théorème 19 (Marchenko-Pastur). Soit $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ une matrice de taille $n \times d$ dont les coefficients sont des variables aléatoires i.i.d centrées et à valeurs réelles telles que $\mathbb{E}[X_{ij}^2] = 1$. Supposons que $d := d(n)$ soit une suite d'entiers telle que $\frac{n}{d} \rightarrow c > 0$. Considérons la matrice de Wishart $\mathcal{W}_{n,d}$ associée à $\mathcal{X}_{n,d}$, alors presque sûrement la mesure spectrale empirique de $\mathcal{W}_{n,d}$, notée $\mu_{\mathcal{W}_{n,d}}$ converge étroitement vers la loi de Marchenko-Pastur $\mu_{MP}^{(c)}$. C'est-à-dire, presque sûrement, pour toute fonction f continue bornée, on a

$$\lim_{n \rightarrow \infty} \int f(x) d\mu_{\mathcal{W}_{n,d}}(x) = \int f(x) d\mu_{MP}^{(c)}(x).$$

Les matrices de Wishart ont été largement étudiées au cours des dernières décennies et pour cause, elles trouvent de multiples applications. En biologie, par exemple, les chercheurs se sont focalisés sur la corrélation entre les gènes, en se questionnant sur la probabilité qu'un certain gène présent dans un organisme en impacte un autre. Les biologistes sont alors amenés à modéliser la corrélation entre de multiples gènes par des matrices de Wishart.

Dans les réseaux de communication, elles sont utilisées pour modéliser les canaux de communication dans les réseaux sans fil et les réseaux de communication à antennes multiples.

En finance et en économie, les statisticiens se sont intéressés aux interactions entre différents actifs du marché. Les matrices de Wishart sont alors utilisées dans l'analyse des séries temporelles financières pour modéliser les dépendances entre les actifs financiers, estimer les volatilités, et étudier les risques associés aux portefeuilles d'actifs. Pour plus de détails et d'applications en télécommunication, nous renvoyons le lecteur au livre de Romain Couillet et Mérouane Debbah [37].

La méthode spectrale employée pour analyser les matrices de Wishart n'est pas toujours la plus judicieuse. Pour certaines applications, il arrive que le nombre de lignes et de colonnes de la matrice initiale $\mathcal{X}_{n,d}$, associée à $\mathcal{W}_{n,d}$, ne soient pas du même ordre par exemple. Dans cette situation, une approche différente a été proposée, il a alors été question de regarder sous quelles conditions imposées sur les lignes et les colonnes de la matrice initiale $\mathcal{X}_{n,d}$, nous parvenons à approcher notre matrice de Wishart par une matrice gaussienne.

4.1.2 Analyse non spectrale des matrices de Wishart

— LE RÉGIME CLASSIQUE :

Dans ce régime, nous supposons que le nombre de lignes n de la matrice initiale $\mathcal{X}_{n,d}$ est fixé. Le nombre de colonnes peut être aussi grand que l'on veut. Pour simplifier, considérons que les coefficients de $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ sont i.i.d et centrés réduits. Dans ce cas, par la loi forte des Grands Nombres, nous pouvons dire que la matrice de Wishart $\mathcal{W}_{n,d}$ associée à $\mathcal{X}_{n,d}$, définies par (4.1), converge presque sûrement vers \mathcal{I}_n quand d tend vers l'infini, i.e :

$$\mathcal{W}_{n,d} \xrightarrow[d \rightarrow \infty]{p.s} \mathcal{I}_n,$$

où \mathcal{I}_n est la matrice identité de taille n . En particulier, nous avons bien pour $1 \leq i < j$ que,

$$W_{ii} = \frac{1}{d} \sum_{k=1}^d X_{ik}^2 \xrightarrow[d \rightarrow \infty]{p.s} \mathbb{E}[X_{11}^2] = 1,$$

$$W_{ij} = \frac{1}{d} \sum_{k=1}^d X_{ik}X_{jk} \xrightarrow[d \rightarrow \infty]{p.s} \mathbb{E}[X_{i1}X_{j1}] = \mathbb{E}[X_{i1}]\mathbb{E}[X_{j1}] = 0.$$

De plus, par le théorème central limite multivarié et en supposant que le moment d'ordre 4 des termes diagonaux de la matrice initiale $\mathcal{X}_{n,d}$ sont finis ; $\mathbb{E}[X_{11}^4] < \infty$, nous avons :

$$\sqrt{d}(\mathcal{W}_{n,d} - \mathcal{I}_n) \xrightarrow[d \rightarrow \infty]{loi} \mathcal{Z}_n,$$

où \mathcal{Z}_n est une matrice du GOE. Une question se pose alors assez naturellement : **Que se passe-t-il lorsque le nombre de lignes n n'est plus fixé ?**

La réponse à cette question est l'enjeu du *high dimensional regime*.

— LE High Dimensional Regime :

L'émergence des grands jeux de données engendrée par l'ère du Big data a soulevé une nouvelle alternative d'étude des matrices aléatoires. De nos jours, le nombre d'observations et de variables deviennent extrêmement grands et considérer un nombre de lignes n fixé n'est absolument plus adapté. Il faut alors prendre en compte que le nombre de variables n et d'observations d peuvent tendre vers l'infini simultanément, sans être du même ordre. C'est ce que nous appelons le *high dimensional regime*. Dans cette situation, nous ne pouvons plus utiliser les résultats précédents, ni même ceux de l'approche spectrale puisque le théorème de Marchenko-Pastur est valable pour n et d du même ordre.

Dans la suite, nous exposerons quelques résultats de la méthode non-spectrale pour l'étude des matrices de Wishart dans le *high dimensional regime*. La compréhension de

ces matrices est fondamentale pour répondre aux besoins des applications. Les auteurs se sont donc questionnés sur l'approximation des matrices de Wishart par des matrices dont on connaît bien la loi des coefficients ; les matrices gaussiennes, en prenant des conditions sur le nombre de lignes et de colonnes de la matrice.

Etude des matrices aléatoires dans le cadre du *high dimensional regime*

Les premiers travaux concernant ce sujet remontent à Bubeck, Ding, Eldan et Racz [20] dans lequel les auteurs s'intéressent aux potentielles structures géométriques présentes dans les graphes aléatoires.

Il existe un lien étroit entre les matrices aléatoires de Wishart et les graphes aléatoires puisque l'étude des matrices aléatoires de grandes tailles part de la volonté de vouloir extraire des informations des grands graphes provenant des données de réseaux sociaux, réseaux de gènes et de neurones en biologie.

En théorie des graphes, on définit un graphe comme étant un ensemble de points, liés par une certaine structure. L'un des plus connu est le graphe d'Erdős-Rényi, noté $\mathbf{G}(n, p)$ dont les sommets $\{1, \dots, n\}$ sont potentiellement reliés par $\frac{n(n-1)}{2}$ arêtes présentes avec probabilité $p \in [0, 1]$. La probabilité qu'il n'y ait pas d'arête entre deux sommets est $1 - p$.

Remarque 25. $\mathbf{G}(n, p)$ est un graphe non orienté qui ne contient ni boucle, ni arête multiple.

En particulier, les auteurs de [20] se sont focalisés sur la représentation des sommets d'un graphe aléatoire, en se questionnant sur l'existence d'une certaine structure géométrique entre les sommets ou sur le fait que les connections sont plutôt purement aléatoires.

Ils cherchent alors à comparer le modèle d'Erdős-Rényi à un graphe géométrique aléatoire $\mathbf{G}(n, p, d)$ possédant n sommets où p est la probabilité qu'il y ait une arête entre deux sommets.

Le dernier objet cité est un graphe dans lequel chaque sommet est repéré par un point dans un espace métrique. Une arête existe entre deux sommets si et seulement si la distance entre ces deux points est plus petite que le seuil de l'espace sous-jacent considéré.

Bubeck, Ding, Eldan et Racz, considèrent une famille de vecteurs aléatoires gaussiens standards X_1, \dots, X_n uniformément distribués sur la sphère euclidienne $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : \|x\|_2 = 1\}$. Dans $\mathbf{G}(n, p, d)$, deux sommets $i \in [n]$ et $j \in [n]$ sont connectés par une arête si et seulement si $\langle X_i, X_j \rangle \geq t_{p,d}$ où la valeur seuil $t_{p,d} \in [-1, 1]$ telle que $\mathbb{P}(\langle X_i, X_j \rangle \geq t_{p,d}) = p$, pour $i \neq j$. En d'autres termes, sur ce type de graphe, une arête est présente entre deux sommets si la corrélation entre les deux vecteurs de \mathbb{R}^d associés à chacun des sommets est assez forte.

En identifiant la matrice d'adjacence de $\mathbf{G}(n, p)$ à une matrice du GOE, et celle de $\mathbf{G}(n, p, d)$ à une matrice de Wishart, les auteurs montrent que sous le régime $\frac{d}{n^3} \rightarrow \infty$, la distance en variation totale entre la matrice de Wishart dont les coefficients sont gaussiens et indépendants et celle du GOE tend vers 0.

Ainsi, leurs résultats montrent que la géométrie dans un graphe aléatoire est perdue quand la dimension de l'espace sous-jacente est plus grande que le cube du nombre de sommets.

Remarque 26. 1. Les coefficients diagonaux de la matrice d'adjacence sont nuls puisqu'il ne peut pas y avoir de boucle sur un sommet. Par conséquent l'étude [20] n'inclut pas la diagonale de la matrice de Wishart.

2. Dans le même temps, les auteurs Jiang et Li [68] obtiennent un résultat similaire cependant l'approche est différente.

En 2018, Racz et Richey [106] reprennent les résultats précédemment établis et abordent le sujet de la transition de phase. Plus précisément, ils analysent la distance en variation totale entre la matrice de Wishart et celle du GOE lorsque n^3 est du même ordre que d , i.e : $\frac{n^3}{d} \rightarrow c \in (0, \infty)$. Ils considèrent que les coefficients de la matrice initiale $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ associée à $\mathcal{W}_{n,d}$, sont des variables aléatoires gaussiennes i.i.d et ils notent \mathcal{Z}_n une matrice du GOE symétrique de taille n , dont les coefficients sur la diagonale sont centrés, de variance égale à 2 et les éléments en dehors sont des variables aléatoires gaussiennes standards. Ils renormalisent \mathcal{Z}_n de sorte que les coefficients aient le même moment d'ordre 1 ainsi ils obtiennent le théorème suivant.

Théorème 20. *Soient $\mathcal{W}_{n,d}$ la matrice de Wishart associée à $\mathcal{X}_{n,d}$ et $\widetilde{\mathcal{Z}}_n$, la matrice du GOE. Posons $d = d(n)$ tel que $\frac{d}{n^3} \rightarrow c \in (0, \infty)$. Alors,*

$$\lim_{n \rightarrow \infty} d_{TV}(\mathcal{W}_{n,d}, \widetilde{\mathcal{Z}}_n) = \text{Erf} \left(\frac{1}{4\sqrt{3}\sqrt{c}} \right),$$

où la fonction erreur $\text{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

Racz et Richey concluent que la transition de phase de la matrice de Wishart à une GOE est lisse.

Remarque 27. *Lorsque $c \rightarrow 0$ ou $c \rightarrow \infty$, la distance en variation totale tend vers 1 ou 0 et nous retrouvons ainsi les résultats de Bubeck, Ding, Eldan et Racz [20] et Jiang et Li [68].*

Certains auteurs ont voulu étendre les résultats des articles [20] et [68] en élargissant la loi sur les coefficients de la matrice initiale. Bubeck et Ganguly [21] ont montré que si les coefficients de $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ sont i.i.d suivant une loi μ suffisamment lisse, la distance en variation totale entre la matrice de Wishart et une matrice gaussienne est toujours bornée lorsque l'on pose la condition $d \gg n^3$, à un facteur logarithmique près. De manière informelle, cela signifie que l'approximation d'une matrice de Wishart par une matrice de Wigner gaussienne est statistiquement valide si le nombre d'observations est au moins le cube du nombre de variables. En réalité, pour éviter d'avoir des difficultés techniques liées à la méthode utilisée, ils analysent la matrice de Wishart en retirant la diagonale à savoir :

$$\mathcal{W}_{n,d} = \frac{\mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - \text{diag} \left(\mathcal{X}_{n,d} \mathcal{X}_{n,d}^T \right)}{\sqrt{d}}. \quad (4.2)$$

Les coefficients de la matrice initiale suivent une loi log-concave μ . Une mesure μ de densité f est dite log-concave si $f(\cdot) = e^{-\varphi(\cdot)}$, pour une fonction convexe φ .

Pour établir la preuve de ce résultat, ils ne travaillent pas directement avec la distance en variation totale mais avec l'entropie relative. En effet, par l'inégalité de Pinsker, ils majorent la distance en variation totale entre la loi de leur matrice de Wishart et celle de la gaussienne par l'entropie relative. Cela leur permet d'utiliser la règle de la chaîne pour l'entropie relative et la convexité de celle-ci.

Théorème 21 (Bubeck et Ganguly (2016)). *Soit μ une mesure log-concave et $\mathcal{W}_{n,d}$ une matrice de Wishart définie par (4.2), considérée comme un vecteur de \mathbb{R}^{n^2} . Alors on a l'estimation suivante :*

Si $\frac{d}{n^3 \log^2(d)} \rightarrow \infty$, $d_{TV}(\mathcal{W}_{n,d}, \mathcal{Z}_n) \rightarrow 0$, où \mathcal{Z}_n est une matrice de Wigner gaussienne, dont les termes diagonaux sont nuls et les termes non diagonaux sont i.i.d suivant une loi gaussienne standard.

- Remarque 28.**
1. L'utilisation de la règle de la chaîne pour l'entropie relative ne permet pas d'inclure les éléments diagonaux.
 2. Pour la mesure log-concave μ , il n'est pas vraiment possible de trouver une expression simple de la densité de la loi de la matrice de Wishart comme on pouvait le faire dans le cas gaussien. C'est notamment ce qu'avaient fait [20] et [106] pour établir la preuve de leur résultat à l'aide de la distance en variation totale.
 3. Bubeck et Ganguly étendent les résultats de Artstein, Ball, Barthe et Naor [6] en établissant une généralisation multi-dimensionnelle d'un TCL basée sur l'entropie.

À l'issue de cet article, les interrogations sont multiples et les principales avancées seraient d'ajouter une certaine forme de corrélation dans la matrice initiale, d'inclure la diagonale de la matrice de Wishart et même de considérer des tenseurs aléatoires de Wishart.

En 2020, dans un brillant papier, Dan Mikulincer [90] répond à tous ces questionnements. Pour cela, il se base sur une nouvelle application de la méthode de Stein et sur l'utilisation de la discrèpence de Stein à la place des distances usuelles (variation totale, Wasserstein). En effet, l'idée de cette méthode repose sur le fait que la déviation entre le noyau de Stein $\tau : \mathbb{R}^n \rightarrow \mathcal{M}_n(\mathbb{R})$ d'une mesure μ de \mathbb{R}^n et la matrice identité I_d , fournit l'écart entre une mesure μ et la loi gaussienne standard de \mathbb{R}^n . La discrèpence de Stein est alors définie de la manière suivante :

$$S(\mu) := \inf_{\tau} \sqrt{\int \|\tau(x) - I_d\|_{\text{HS}}^2 d\mu(x)}, \quad (4.3)$$

où nous rappelons que l'application mesurable τ est un noyau de Stein pour la mesure μ de \mathbb{R}^n si on a l'égalité suivante pour une fonction test $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$

$$\int \langle x, f(x) \rangle d\mu(x) = \int \langle \tau(x), \text{Jac } f(x) \rangle_{\text{HS}} d\mu(x),$$

avec $\text{Jac } f(x)$, la matrice jacobienne de f .

Dans [90], Mikulincer va s'appuyer sur la structure des vecteurs aléatoires de \mathbb{R}^n au lieu d'analyser directement et globalement une matrice ou un tenseur aléatoire. L'heuristique est d'introduire une application qui "transforme" un vecteur en tenseur d'ordre p . Autrement dit, si X est un vecteur aléatoire de \mathbb{R}^n , il va considérer une application $\psi : \mathbb{R}^n \rightarrow \mathbb{R}^{n^{\otimes p}}$ telle que $\psi(X) = X^{\otimes p} - \mathbb{E}[X^{\otimes p}]$. Cela va alors lui permettre de trouver un nouveau noyau de Stein τ pour $X^{\otimes p} - \mathbb{E}[X^{\otimes p}]$.

La clé pour parvenir à la borne (4.4) est l'utilisation de la propriété de sous-additivité de la discrèpence de Stein. Cette propriété est valide uniquement lorsque $X^{\otimes p}$ est isotrope (la loi μ du vecteur X de \mathbb{R}^n est centrée et la matrice de covariance est l'identité), il est donc amené à utiliser une renormalisation en s'appuyant sur une transformation linéaire $A : (\mathbb{R}^n)^{\otimes p} \rightarrow V$ de sorte que $A(X^{\otimes p} - \mathbb{E}[X^{\otimes p}])$ soit cette fois-ci isotrope. De cette manière, la propriété de sous-additivité de la discrèpence de Stein implique que,

$$S^2(A(X^{\otimes p} - \mathbb{E}[X^{\otimes p}])) \leq 2\|A\|_{\text{op}}^2 \mathbb{E} \left[\|\tau(X^{\otimes p} - \mathbb{E}[X^{\otimes p}])\|_{\text{HS}}^2 \right] + 2 \dim V.$$

Ainsi, en ayant recours au lemme 8 suivant, il parvient à majorer le terme d'espérance du noyau de Stein associé au tenseur par l'espérance du vecteur aléatoire et se ramène donc à une situation plus simple à analyser.

Lemme 8 (Mikulincer (2020)). *Soient X un vecteur aléatoire de \mathbb{R}^n et $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une application différentiable presque partout, et telle que $X \stackrel{\text{loi}}{=} \phi(G)$, où G est un vecteur aléatoire gaussien standard de \mathbb{R}^n . Alors pour tout $p \geq 2$, il existe un noyau de Stein τ de $X^{\otimes p} - \mathbb{E}[X^{\otimes p}]$ tel que ;*

$$\mathbb{E} \left[\|\tau(X^{\otimes p} - \mathbb{E}[X^{\otimes p}])\|_{\text{HS}}^2 \right] \leq p^4 n \sqrt{\mathbb{E} \left[\|X\|_2^{8(p-1)} \right]} \sqrt{\mathbb{E} \left[\|\text{Jac } \phi(G)\|_{\text{op}}^8 \right]}.$$

Grâce à ces résultats, Mikulincer établit une borne effective pour l'erreur d'approximation de la loi de la matrice de Wishart associée à $\mathcal{X}_{n,d}$ (dont les coefficients vérifient plusieurs hypothèses) par une loi gaussienne. En particulier :

Théorème 22 (Mikulincer (2020)). *Soit $X \sim \mu$ un vecteur aléatoire isotrope (i.e : centré de matrice de covariance égale à l'identité) de \mathbb{R}^n et soit $G \sim \mathcal{N}(0, I_d)$, un vecteur aléatoire gaussien standard de \mathbb{R}^n . Supposons que $X \stackrel{\text{loi}}{=} \phi(G)$, pour $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$, localement lipschitzienne et $A : \text{Sym}^p(\mathbb{R}^n) \rightarrow V \subset \text{Sym}^p$ une transformation linéaire telle que $A_*W_{n,d}^p(\mu)$ soit isotrope. Alors pour tout entier naturel $p \geq 2$,*

$$S^2\left(A_*W_{n,d}^p(\mu)\right) \leq 2\|A\|_{\text{op}}^2 p^4 \frac{n}{d} \sqrt{\mathbb{E}\left[\|X\|_2^{8(p-1)}\right]} \sqrt{\mathbb{E}\left[\|\text{Jac } \phi(G)\|_{\text{op}}^8\right]} + \frac{2n^p}{d}. \quad (4.4)$$

Notons que $\text{Sym}^p(\mathbb{R}^n)$ est l'espace tensoriel symétrique dont la base est $\{e_{j_1}e_{j_2}\dots e_{j_p} | 1 \leq j_1 \leq j_2 \leq \dots \leq j_p \leq n\}$, et $W_{n,d}^p(\mu)$ correspond à la loi du tenseur de Wishart ;

$$\frac{1}{\sqrt{d}} \sum_{i=1}^d (X_i^{\odot p} - \mathbb{E}[X_i^{\odot p}]), \quad (4.5)$$

dont les coefficients $X_i \sim \mu$ sont i.i.d et $X_i^{\odot p}$ correspond au p -ième produit tensoriel symétrique de X_i . En employant le théorème 22, il trouve divers résultats :

1. Mikulincer parvient à répondre à une première interrogation de Bubeck et Ganguly [21] puisqu'il montre qu'une structure de dépendance est possible, au contraire de Bubeck et Ganguly qui avaient supposé une totale indépendance des coefficients. Il considère une matrice de Wishart dont les composantes de la matrice initiale sont log-concaves et corrélées. Dans ce premier résultat (Théorème 2 dans [90]), il n'inclut pas la diagonale afin d'avoir plus de contrôle sur les valeurs propres de la matrice de covariance de $W_{n,d}^p(\mu)$. Cependant, il ajoute l'hypothèse que la mesure μ des coefficients soit également "inconditionnelle", de sorte à pouvoir minorer la plus petite des valeurs propres par une constante dépendant de p .
2. Mikulincer répond ensuite à l'intuition de Bubeck et Ganguly [21] (Théorème 4 dans [90]) puisqu'il montre qu'en considérant une matrice de Wishart dont la diagonale est incluse et dont les coefficients de la matrice initiale sont log-concaves et totalement indépendants, l'approximation est toujours vraie. Cependant afin de pouvoir utiliser le théorème 22, il doit s'assurer que la matrice de covariance de $W_{n,d}^2(\mu)$ est uniformément bornée en contrôlant la variance des termes diagonaux, en particulier il montre qu'ils sont minorés par $\frac{1}{100}$ ou égaux à 1.
3. Il généralise le résultat principal de Bubeck et Ganguly, le théorème 21, au cas des tenseurs aléatoires (Théorème 3 [90]) et en employant une classe de mesures plus larges que l'hypothèse de log-concavité.

Remarque 29. *Attention, les résultats ne sont pas directement comparables à ceux de Bubeck et Ganguly puisque Mikulincer utilise la discrédence de Stein tandis que Bubeck et Ganguly utilisent l'entropie relative. Dans le cas où l'information de Fischer de $W_{n,d}^p(\mu)$ est finie, on peut majorer l'entropie relative par la discrédence de Stein, à l'aide de l'inégalité HSI de Ledoux, Nourdin et Peccati [79].*

Les hypothèses concernant la loi des composantes de la matrice initiale étant dominantes dans les études que nous venons de citer, nous proposons d'évoquer un résultat de Fang et Koike

[49], qui s'appuie sur la méthode des paires échangeables. Ce résultat très général ne donne pas de condition sur la loi des coefficients, mis à part une condition d'indépendance totale, et qu'elles admettent un moment d'ordre quatre fini. En revanche, la méthode utilisée ne permet pas d'inclure la diagonale de la matrice de Wishart. Ils obtiennent le taux de convergence optimal de $\sqrt{\frac{n^3}{d}}$ sous la condition que le moment d'ordre 6 des coefficients de la matrice initiale $\mathcal{X}_{n,d}$ soit fini.

Jusqu'à présent nous avons vu que l'inclusion de la diagonale d'une matrice de Wishart soulève de grandes problématiques pour l'étude de l'approximation de la loi de celles-ci par des lois gaussiennes multi-dimensionnelles. Nous proposons, dans la partie consécutive, l'étude asymptotique de ces matrices par la méthode de Stein-Malliavin, qui va permettre, non seulement, d'inclure la diagonale, mais aussi de répondre aux besoins de corrélation partielle ou totale et coefficients non gaussiens de la matrice initiale.

4.2 Étude asymptotique des matrices aléatoires par la méthode de Stein Malliavin

Le principe avancé par Nourdin et Zheng [98], en 2018, constitue une base pour l'étude des fluctuations des matrices de Wishart des articles suivants [18], [97], [19], [45], [46], [53], [52]. Ces articles prennent en compte des hypothèses différentes que dans [98] sur les coefficients de la matrice initiale $\mathcal{X}_{n,d}$ associée à la matrice de Wishart, à savoir la non stationnarité, la corrélation totale, la non gaussianité, ou encore la non équidistribution des coefficients.

Leur idée repose sur le Théorème 3.5 [96], établi par Nourdin, Peccati et Réveillac. Ce résultat permet d'évaluer la distance entre la loi d'un vecteur aléatoire dont les composantes sont des variables aléatoires centrées, de carrés intégrables et dérivables au sens de Malliavin. Dans ce qui suit, nous détaillons la méthode employée et les résultats des différents articles.

4.2.1 Coefficients gaussiens

Limite gaussienne

En 2018, Nourdin et Zheng suggèrent d'étudier le comportement asymptotique des matrices de Wishart en se basant sur la méthode de Stein-Malliavin. Pour cela, ils s'appuient sur l'utilisation du théorème central limite quantitatif 12 pour des vecteurs aléatoires, établi auparavant dans [96]. Ainsi, les auteurs de [98] associent à leur matrice de Wishart qui est symétrique par définition, un demi-vecteur afin de pouvoir utiliser la borne du corollaire 1. En le combinant au lemme 7, il en résulte une majoration de l'erreur commise dans l'approximation d'une matrice de Wishart par une matrice du GOE (Ensemble Gaussien Orthogonal).

Dans leur pertinent article [98], Nourdin et Zheng parviennent à retrouver les résultats de l'article [20] et [68], en utilisant la distance de Wasserstein, à la place de la distance en variation totale. En particulier, ils réussissent même à améliorer ces résultats, puisqu'ils ajoutent de la corrélation sur les coefficients.

Sous certaines conditions concernant la fonction de covariance, que nous détaillerons ensuite, ils trouvent que leur matrice de Wishart renormalisée dont les coefficients de la matrice initiale $\mathcal{X}_{n,d}$ sont gaussiens et partiellement corrélés peut être approximée par une matrice GOE lorsque

$\frac{n^3}{d}$ tend vers 0.

Intéressons-nous désormais aux hypothèses posées par Nourdin et Zheng dans leur papier [98]. Pour cela, considérons \mathcal{H} un espace de Hilbert séparable muni de son produit scalaire $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ et de sa norme $\|\cdot\|_{\mathcal{H}}$, auquel on lui associe un processus isonormal gaussien tel que :

$$X = \{X(h), h \in \mathcal{H}\}.$$

Soit $\{e_{ij} : i, j \geq 1\} \subset \mathcal{H}$ une famille vérifiant :

$$\langle e_{ij}, e_{kl} \rangle_{\mathcal{H}} = s(j-l) \mathbf{1}_{\{i=k\}},$$

où $s : \mathbb{Z} \rightarrow \mathbb{R}$ est une fonction de corrélation, et telle que $s(0) = 1$. Grâce à ce contexte, nous pouvons avancer les hypothèses concernant les éléments de la matrice de Wishart.

- Les composantes de la matrice initiale $\mathcal{X}_{n,d}$ sont des variables aléatoires gaussiennes indépendamment distribuées. En particulier, $X_{ij} = X(e_{ij}) \sim \mathcal{N}(0, 1)$ et $\mathbb{E}[X_{ij}X_{kl}] = s(j-l) \mathbf{1}_{\{i=k\}}$, ce qui implique que les coefficients sont aussi stationnaires. En effet, par définition, lorsqu'on dispose de variables aléatoires gaussiennes stationnaires, la fonction de covariance dépend uniquement de $j-l$.
- La corrélation a donc lieu sur une même ligne de la matrice et est donnée par la fonction $s(j-l)$. Il n'y a pas de corrélation d'une ligne à une autre ligne puisque les auteurs posent la condition grâce à l'indicatrice.

Nous énonçons désormais, un des résultats principaux de l'article de Nourdin et Zheng.

Théorème 23 (Nourdin et Zheng (2018)). *Notons la matrice recentrée et renormalisée :*

$$\widetilde{\mathcal{W}}_{n,d} := \sqrt{d} \left(\frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - \mathcal{I}_n \right).$$

Soit $\mathcal{G}_{n,d}^{(s)} = (G_{ij})_{1 \leq i, j \leq n}$, une matrice aléatoire symétrique de taille n telle que son vecteur aléatoire associé soit le suivant :

$$(G_{11}, \dots, G_{1n}, G_{21}, \dots, G_{2n}, \dots, G_{n1}, \dots, G_{nn})^T,$$

vecteur aléatoire gaussien de \mathbb{R}^{n^2} , qui a la même matrice de covariance que

$$(W_{11}, \dots, W_{1n}, W_{21}, \dots, W_{2n}, \dots, W_{n1}, \dots, W_{nn})^T.$$

Alors, pour tout $n, d \geq 1$, nous avons,

$$d_W(\widetilde{\mathcal{W}}_{n,d}, \mathcal{G}_{n,d}^{(s)}) \leq \sqrt{\frac{192n^3}{d} \left(\sum_{|k| \leq d} |s(k)|^{\frac{4}{3}} \right)^3}.$$

De plus, si $s \in l^2(\mathbb{Z})$, notons $\mathcal{Z}_n^{(s)} = (Z_{ij})_{1 \leq i, j \leq n}$, une matrice aléatoire symétrique telle que $Z_{ii} \sim \mathcal{N}\left(0, 2\|s\|_{l^2(\mathbb{Z})}^2\right)$ et $Z_{ij} \sim \mathcal{N}\left(0, \|s\|_{l^2(\mathbb{Z})}^2\right)$ pour $i < j$, $Z_{ij} = Z_{ji}$ pour $i > j$ et $(Z_{ij})_{1 \leq i, j \leq n}$, sont des variables aléatoires indépendantes,

$$d_W(\mathcal{G}_{n,d}^{(s)}, \mathcal{Z}_n^{(s)}) \leq \frac{2\sqrt{n(n+1)}}{\|s\|_{l^2(\mathbb{Z})}} \left(\sum_{|k| \leq d} s(k)^2 + \frac{1}{d} \sum_{|k| \leq d} |k|s(k)^2 \right).$$

Remarque 30. Si la fonction de corrélation $s \in l^{\frac{4}{3}}(\mathbb{Z})$ alors nous obtenons la même borne que dans le cas où les coefficients sont totalement indépendants. En effet, on obtient que $d_W(\widetilde{\mathcal{W}}_{n,d}, \mathcal{G}_{n,d}^{(s)}) \leq C\sqrt{\frac{n^3}{d}}$.

La première étape de la preuve de ce théorème consiste en l'association d'un demi-vecteur à la matrice de Wishart. Il faut ensuite utiliser le théorème d'approximation d'un vecteur aléatoire par un vecteur gaussien qui repose sur l'estimation de la quantité (2.18), du corollaire 1, à savoir

$$\sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq k \leq l \leq n}} \text{Var} \left(\frac{1}{2} \langle DW_{ij}, DW_{kl} \rangle_{\mathcal{H}} \right).$$

En fait, à l'aide de la formule produit et l'utilisation de l'inégalité de Young, les auteurs établissent pour $1 \leq i \leq j \leq n$ et $1 \leq k \leq l \leq n$:

$$\begin{aligned} \text{Var} \left(\frac{1}{2} \langle DW_{ij}, DW_{kl} \rangle_{\mathcal{H}} \right) &\leq \frac{C}{d^2} \sum_{k,l,u,v=1}^d s(k-l)s(l-u)s(u-v)s(v-k) \\ &\leq \frac{C}{d} \left(\sum_{|k| \leq d} |s(k)|^{\frac{4}{3}} \right)^3. \end{aligned}$$

Notons que si $i \neq j \neq k \neq l$, le produit scalaire sur l'espace de Hilbert \mathcal{H} , $\langle DW_{ij}, DW_{kl} \rangle_{\mathcal{H}}$ est nul et finalement la somme à calculer initialement $\sum_{\substack{1 \leq i \leq j \leq n \\ 1 \leq k \leq l \leq n}}$ peut être remplacée par $\sum_{i,j,k,l \in \mathcal{I}}$ où

$$\mathcal{I} := \{(i, j, k, l) \in \{1, \dots, n\}^4 : i, j, k, l \text{ ne sont pas mutuellement distincts}\}. \quad (4.6)$$

Le cardinal de cet ensemble étant $n^4 - n(n-1)(n-2)(n-3) \leq 6n^3$, on parvient alors au résultat de Nourdin et Zheng. Cet argument combinatoire est très important puisqu'il est employé dans d'autres articles utilisant la théorie de Stein-Malliavin pour l'étude asymptotique des matrices de Wishart et conduit à retrouver le ratio n^3/d pour la convergence d'une matrice de Wishart vers une matrice du GOE. Le lemme 7 permet de conclure sur l'approximation.

Remarque 31. Nourdin et Zheng s'intéresseront également au cas où la fonction de corrélation n'appartient plus à $l^{\frac{4}{3}}(\mathbb{Z})$, en choisissant celle du mouvement brownien fractionnaire de paramètre de Hurst $H \in (0, 1)$. Dans cette situation, lorsque $H \in (0, \frac{3}{4})$, la fonction de corrélation

$$s_H(k) = \frac{1}{2} (|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}), \quad k \in \mathbb{Z},$$

appartient à $l^2(\mathbb{Z})$. Les auteurs de [98] retrouvent alors une majoration de l'erreur commise dans l'approximation de $\widetilde{\mathcal{W}}_{n,d}$ par une matrice du GOE, de l'ordre de $\sqrt{\frac{n^3}{d}}$ lorsque $H \in (0, \frac{1}{2})$. Dans le cas où $H = \frac{1}{2}$, ils retombent exactement dans le cas d'indépendance totale, et constatent le même régime asymptotique que dans les articles [20] et [68]. En revanche lorsque $H \geq \frac{3}{4}$, la convergence de la matrice de Wishart vers une matrice du GOE n'est plus assurée mais Nourdin et Zheng présentent un théorème non central limite, conduisant à la convergence de notre matrice de Wishart vers une matrice de Rosenblatt-Wishart (définition 31).

L'hypothèse de stationnarité des composantes étant assez restrictive, c'est en 2020 que Bourguin et Dang dans l'article [18] chercheront à relaxer ce critère adopté par Nourdin et Zheng

dans [98], sur les coefficients de la matrice initiale $\mathcal{X}_{n,d}$. Ils maintiennent l'hypothèse d'une corrélation partielle mais supposent, néanmoins que les coefficients de la matrice initiale soient auto-similaires. Ils reprennent en fait les hypothèses de l'article de Harnett et Nualart dans [57] dans lequel les auteurs établissent une version du théorème de Breuer-Major pour une classe de processus gaussiens continus auto-similaires et à accroissements non stationnaires. Autrement dit, en prenant certaines conditions sur la fonction de covariance Nualart et Harnett prouvent qu'une fonctionnelle d'accroissements d'un processus converge en loi vers une variable aléatoire gaussienne.

Ainsi, en se basant sur une certaine forme de fonction de covariance, pour laquelle les coefficients ne sont plus stationnaires, Bourguin et Dang élargissent les résultats dans [98]. Pour cela, ils considèrent les hypothèses suivantes sur les coefficients de la matrice initiale $\mathcal{X}_{n,d}$:

- Soit \mathcal{H} un espace de Hilbert tel que pour une collection de $(e_{ik})_{i,k \in \mathbb{N}}$ dans \mathcal{H} , $\mathbb{E}[X_{ik}X_{jl}] = \langle e_{ik}, e_{jl} \rangle_{\mathcal{H}}$. On a $\langle e_{ik}, e_{jl} \rangle_{\mathcal{H}} = 0$ pour $i \neq j$, il n'y a donc pas de corrélation entre deux lignes différentes de la matrice et $\|e_{ik}\|_{\mathcal{H}} = 1$. Cependant, l'indépendance n'est pas totale puisque, les éléments d'une même ligne sont corrélés.
- Soit $\{Y^i : i \in \mathbb{N}\}$ des copies i.i.d d'un processus gaussien Y . Les composantes de la matrice initiale $\mathcal{X}_{n,d}$ sont données pour $1 \leq k \leq d$ par :

$$X_k^i = \frac{Y_{k+1}^i - Y_k^i}{\|Y_{k+1}^i - Y_k^i\|_{L^2(\Omega)}},$$

où $\{Y_t : t \geq 0\}$ est un processus auto-similaire d'indice $\beta \in (0, 1)$ tel que

$$\mathbb{E}[Y_s Y_t] = s^{2\beta} \phi\left(\frac{t}{s}\right).$$

La fonction ϕ caractérise la fonction de covariance de Y . Elle est choisie par Harnett et Nualart dans [57], de sorte que les accroissements soient "proche d'être stationnaires". Ainsi ils considèrent ϕ telle que $\mathbb{E}[(Y_{t+s} - Y_t)^2] \sim s^\alpha$ lorsque $s \rightarrow 0$ avec $\alpha \in (0, 2\beta]$. En particulier, sous les hypothèses (H.1) dans [57],

$$\phi(x) = -\lambda(x-1)^\alpha + \psi(x), \quad (4.7)$$

où $\lambda > 0$ et ψ est une fonction deux fois différentiable sur $[1, \infty)$. Les dérivées premières et seconde de ϕ et ψ vérifient plusieurs hypothèses selon la valeur du paramètre α . Elles sont appelées (H.1) et (H.2), pour plus de détails le lecteur pourra consulter [57].

Sous les conditions que nous venons d'énoncer, Bourguin et Dang obtiennent la convergence de leur matrice de Wishart correctement renormalisée vers une matrice du GOE lorsque $0 < \alpha \leq \frac{3}{2}$. Ils utilisent pour cela, la méthode de Stein-Malliavin pour les matrices aléatoires, établie par Nourdin et Zheng dans [98].

Théorème 24 (Bourguin et Dang (2020)). *1. Soit $0 < \alpha < \frac{3}{2}$, dans ce cas, la renormalisation de la matrice de Wishart associée à la matrice initiale $\mathcal{X}_{n,d}$ est donnée par*

$$\mathcal{W}_{n,d} = \sqrt{d} \left(\frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - \mathcal{I}_n \right).$$

Alors la matrice de Wishart $\mathcal{W}_{n,d}$ est proche de l'ensemble gaussien orthogonal \mathcal{Z}_n où $Z_{ii} \sim \mathcal{N}(0, 2\sigma^2)$ et $Z_{ij} \sim \mathcal{N}(0, \sigma^2)$ pour $i \neq j$ avec σ^2 est définie par (4.8). De plus, nous obtenons la borne quantitative suivante :

$$d_W(\mathcal{W}_{n,d}, \mathcal{Z}_n) \leq C(n^{\frac{3}{2}} r(\alpha, \nu) + nd^{2\alpha-3} + nd^{-1}),$$

où

$$r(\alpha, \nu) = \begin{cases} d^{\frac{2\alpha-3}{2(9-2\alpha)}} & \text{si } \alpha < 1 \text{ et } \alpha + \nu < 2 \\ d^{-\frac{1}{2}} & \text{si } \alpha < 1 \text{ et } \alpha + \nu \geq 2 \\ d^{-\frac{1}{2}} & \text{si } 1 \leq \alpha < \frac{5}{4} \\ d^{-\frac{1}{2}} (\ln d)^{\frac{3}{2}} & \text{si } \alpha = \frac{5}{4} \\ d^{2\alpha-3} & \text{si } \frac{5}{4} < \alpha < \frac{3}{2}. \end{cases}$$

2. Soit $\alpha = \frac{3}{2}$. La renormalisation de la matrice de Wishart associée à la matrice initiale $\mathcal{X}_{n,d}$ est donnée par

$$\widetilde{\mathcal{W}}_{n,d} = \frac{\sqrt{d}}{\ln d} \left(\frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - \mathcal{I}_n \right).$$

Alors la matrice de Wishart $\widetilde{\mathcal{W}}_{n,d}$ est proche de la matrice du GOE, $\widetilde{\mathcal{Z}}_n$ où $Z_{ii} \sim \mathcal{N}(0, 2\rho^2)$ et pour $i \neq j$, $Z_{ij} \sim \mathcal{N}(0, \rho^2)$, avec

$$\rho^2 := \lim_{d \rightarrow \infty} \frac{1}{d \ln d} \sum_{k,l=1}^d \langle e_{ik}, e_{il} \rangle_{\mathcal{H}}^2 = \frac{9}{32}.$$

De plus,

$$d_W(\widetilde{\mathcal{W}}_{n,d}, \widetilde{\mathcal{Z}}_n) \leq C \frac{n^{\frac{3}{2}}}{\ln d}.$$

Bourguin et Dang utilisent le lemme 7 et la borne du théorème 12 afin d'obtenir leurs résultats. Plus précisément, ils sont amenés à majorer la quantité

$$\sum_{k,l,m,p=1}^d \langle e_{ik}, e_{il} \rangle_{\mathcal{H}} \langle e_{im}, e_{ip} \rangle_{\mathcal{H}} \langle e_{ik}, e_{im} \rangle_{\mathcal{H}} \langle e_{il}, e_{ip} \rangle_{\mathcal{H}}.$$

Puisque la fonction de covariance utilisée a une certaine forme ; $\phi(x) = -\lambda(x-1)^\alpha + \psi(x)$, la variance des coefficients de la matrice de Wishart $\mathcal{W}_{n,d}$, notée

$$\sigma_d^2 := \frac{1}{d} \sum_{k,l=1}^d \langle e_{ik}, e_{il} \rangle_{\mathcal{H}}^2$$

vérifie :

$$\sigma^2 := \lim_{d \rightarrow \infty} \sigma_d^2 = \frac{1}{2} \sum_{m \in \mathbb{Z}} \left(\frac{1}{2} (|m+1|^\alpha + |m-1|^\alpha - 2|m|^\alpha) \right)^2. \quad (4.8)$$

En fait ceci découle du fait que sous les hypothèses (H.1) de Harnett et Nualart dans [57], nous avons que la covariance des coefficients de la matrice initiale $\mathcal{X}_{n,d}$ vérifie :

- Pour $0 < 2s \leq \frac{t}{3} \leq r \leq t - 2s$,

$$\mathbb{E}[(Y_t - Y_{t-s})(Y_r - Y_{r-s})] = \lambda(r-s)^{2\beta-\alpha} [(t-r-s)^\alpha + (t-r+s)^\alpha - 2(t-r)^\alpha] + g_3(r, t, s),$$

où $g_3(r, t, s)$ provient du lemme 3.2 de [57].

Ainsi, après quelques majorations et en utilisant (4.8) les auteurs parviennent à obtenir le résultat principal du théorème 24.

Remarque 32. 1. L'approximation des matrices de Wishart dont les coefficients de la matrice initiale sont non-stationnaires est du même ordre que dans le cas de [98] où les composantes sont stationnaires. Dans le cas où $\alpha < 1$ et $\alpha + \nu < 2$, l'approximation est moins bonne, cela est dû au fait que les auteurs de [18] ne peuvent pas utiliser la borne suivante,

$$\begin{aligned} & \sum_{k,l,m,p=1}^d \langle e_{ik}, e_{il} \rangle_{\mathcal{H}} \langle e_{im}, e_{ip} \rangle_{\mathcal{H}} \langle e_{ik}, e_{im} \rangle_{\mathcal{H}} \langle e_{il}, e_{ip} \rangle_{\mathcal{H}} \\ & \leq C \left(\sum_{m=-d+1}^{d-1} \left| \frac{1}{2} (|m+1|^\alpha + |m-1|^\alpha - 2|m|^\alpha) \right|^{\frac{4}{3}} \right)^3 \end{aligned}$$

puisque l'inégalité de Young n'est pas applicable.

2. Lorsque $\alpha > 3/2$ il n'y a plus de TLC mais les auteurs de [18] parviennent à établir un théorème non central limite puisque la matrice de Wishart converge dans ce cas vers une matrice de Rosenblatt (définition 31).

Jusqu'à présent, nous avons étudié les situations dans lesquelles les coefficients de la matrice initiale étaient gaussiens et partiellement corrélés.

Que peut-on dire lorsque les coefficients sont totalement corrélés ?

En 2022, Nourdin et Pu dans l'article [97] cherchent à améliorer les résultats de [98] lorsque les coefficients de la matrice initiale $\mathcal{X}_{n,d}$ sont totalement corrélés. En effet, dans le dernier papier cité, les auteurs avaient établi un TCL pour une matrice de Wishart dont les composantes de la matrice initiale étaient gaussiennes et avec une corrélation totale. Ils avaient, pour cela, utilisé le théorème 6.1.2 du livre de Nourdin et Peccati [93] et employé la distance d_2 . A l'aide d'une majoration de la distance de Wasserstein par la distance d_2 , ils parvenaient à obtenir des bornes comparables aux résultats d'indépendance partielle et totale mais nettement moins bonne. En particulier, le théorème 12 n'avait pas pu être utilisé dans cette situation puisqu'il n'était pas évident que la matrice de covariance soit inversible.

A l'aide de certaines hypothèses sur la fonction de corrélation Nourdin et Pu pallient ce problème, et fournissent la même borne que dans les articles précédemment cités. Dans la suite, nous présentons les hypothèses considérées par Nourdin et Pu dans leur papier [97].

Considérons \mathcal{H} un espace de Hilbert muni de son produit scalaire $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ et de sa norme $\|\cdot\|_{\mathcal{H}}$. Soit $\{e_{ij} : i, j \geq 1\} \subset \mathcal{H}$, une famille telle que pour tout $i, j \geq 1$, $\|e_{ij}\|_{\mathcal{H}} = 1$ et

$$\langle e_{ij}, e_{kl} \rangle_{\mathcal{H}} = r(i-k)s(j-l),$$

où $s, r : \mathbb{Z} \rightarrow \mathbb{R}$ sont deux fonctions de corrélation satisfaisant $s(0) = r(0) = 1$. On associe à l'espace de Hilbert \mathcal{H} , un processus isonormal $X = \{X(h), h \in \mathcal{H}\}$ qui vérifie les conditions suivantes :

1. $X_{ij} = X(e_{ij}) \sim \mathcal{N}(0, 1)$. Les coefficients de la matrice initiale $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ sont donc gaussiens et corrélés. En effet, la famille $\{e_{ij} : i, j \geq 1\} \subset \mathcal{H}$ n'est pas orthogonale, et on a $\mathbb{E}[X_{ij}X_{kl}] = r(i-k)s(j-l)$.
2. Les composantes situées sur une même ligne et sur une même colonne sont toujours stationnaires puisque leur fonction de covariance dépend uniquement de $j-l$ ou $i-k$ respectivement.
3. La matrice de Wishart considérée est renormalisée et notée $\widetilde{\mathcal{W}}_{n,d} = (\widetilde{W}_{ij})_{1 \leq i, j \leq n}$, où,

$$\widetilde{W}_{ij} = \frac{1}{\sqrt{d}} \sum_{k=1}^d (X_{ik}X_{jk} - r(i-j)).$$

En tenant compte de ces hypothèses, les auteurs de [97] établissent le régime asymptotique de la matrice de Wishart $\widetilde{\mathcal{W}}_{n,d}$.

Théorème 25 (Nourdin et Pu (2022)). *Supposons que*

$$\|r\|_{l^1(\mathbb{Z})} < \frac{\sqrt{6}}{2}. \quad (4.9)$$

Soit $\mathcal{G}_{n,d}^{r,s} = (G_{ij})_{1 \leq i,j \leq n}$, une matrice carrée de taille n , symétrique, associée au vecteur aléatoire gaussien

$$(G_{11}, \dots, G_{1n}, G_{21}, \dots, G_{2n}, \dots, G_{n1}, \dots, G_{nn}).$$

On suppose qu'il est centré et qu'il a la même matrice de covariance C que le vecteur aléatoire

$$(\widetilde{W}_{11}, \dots, \widetilde{W}_{1n}, \widetilde{W}_{21}, \dots, \widetilde{W}_{2n}, \dots, \widetilde{W}_{n1}, \dots, \widetilde{W}_{nn}).$$

Alors pour tout $n, d \geq 1$,

$$d_W(\widetilde{\mathcal{W}}_{n,d}, \mathcal{G}_{n,d}^{r,s}) \leq \frac{\|r\|_{l^1(\mathbb{Z})}^{\frac{3}{2}}}{3 - 2\|r\|_{l^1(\mathbb{Z})}^2} \sqrt{\frac{32n^3}{d} \left(\sum_{|k| \leq d} |s(k)|^{\frac{4}{3}} \right)^3}.$$

Remarque 33. 1. La condition (4.9) posée sur r est primordiale pour pouvoir utiliser le corollaire 1. En effet, la matrice de covariance C est inversible seulement si $\|r\|_{l^1(\mathbb{Z})} < \frac{\sqrt{6}}{2}$. Sous cette condition, nous constatons que C est à diagonale strictement dominante, c'est-à-dire que le module de chaque coefficient diagonal est strictement plus grand que la somme des modules des autres coefficients de sa ligne. Par conséquent en utilisant le lemme de Hadamard, comme C est une matrice à diagonale strictement dominante, on en conclut qu'elle est inversible.

2. Sous la condition que $s \in l^{\frac{4}{3}}(\mathbb{Z})$, Nourdin et Pu obtiennent alors la même borne que dans [98] dans le cas de corrélation partielle et que dans le cas d'indépendance totale [20] et [68] à savoir,

$$d_W(\widetilde{\mathcal{W}}_{n,d}, \mathcal{G}_{n,d}^{r,s}) \leq C \sqrt{\frac{n^3}{d}}.$$

Limite non gaussienne

Intéressons-nous désormais à la situation où une matrice de Wishart converge vers une matrice non gaussienne. Nous nous replaçons dans le cadre des hypothèses introduites dans la section précédente de l'article [98]. Ainsi les composantes de la matrice initiale considérée sont gaussiennes, issues d'un processus à accroissements stationnaires, corrélées sur une même ligne dont la fonction de covariance s correspond à celle du MBf. En 2018, Nourdin et Zheng sont les premiers à introduire la notion de matrice de Rosenblatt-Wishart et à construire un théorème non central limite pour une matrice Wishart dont la limite était une matrice de Rosenblatt-Wishart. Ils définissent cette matrice comme suit.

Définition 31 (Matrice de Rosenblatt-Wishart). *Soit un processus gaussien centré*

$$B = \{B_t = (B_t^1, \dots, B_t^n), t \in \mathbb{R}_+\}$$

où les B^i pour $i = 1, \dots, n$, sont des copies indépendantes d'un mouvement brownien fractionnaire de paramètre de Hurst $H \in (0, 1)$ et de fonction de covariance $R_H(t, s) = \frac{1}{2}(s^{2H} + t^{2H} - |t - s|^{2H})$. Soit $\mathcal{S}_{n,d} = (S_{ij})_{1 \leq i, j \leq n}$ une matrice carrée de taille n telle que :

$$S_{ij}(d) = \begin{cases} d \sum_{p=0}^{d-1} \left(B_{\frac{p+1}{d}}^i - B_{\frac{p}{d}}^i \right) \left(B_{\frac{p+1}{d}}^j - B_{\frac{p}{d}}^j \right) & \text{si } i \neq j, \\ d \sum_{p=0}^{d-1} \left(\left(B_{\frac{p+1}{d}}^i - B_{\frac{p}{d}}^i \right)^2 - d^{-2H} \right) & \text{si } i = j. \end{cases}$$

La matrice de Rosenblatt-Wishart $\mathcal{R}_n^{(H)} = (R_{ij})_{1 \leq i, j \leq n}$ est une matrice symétrique de taille $n \times n$ dont les coefficients sont donnés par $S_{ij}(d) \xrightarrow[d \rightarrow \infty]{L^2(\Omega)} R_{ij}$ pour tout $i \leq i, j, \leq n$.

Remarque 34. La limite dans $L^2(\Omega)$ existe, le lecteur pourra consulter la preuve de la proposition 3.1 de l'article [98].

En considérant la fonction de corrélation du mouvement brownien fractionnaire dont le paramètre de Hurst $H \in (3/4, 1)$, Nourdin et Zheng, ne peuvent plus établir un TCL pour leur matrice de Wishart. En effet, dans ce cas la fonction de corrélation $s \notin l^2(\mathbb{Z})$ et par conséquent le théorème 23 n'est plus utilisable pour déterminer une majoration de l'erreur commise dans l'approximation d'une matrice de Wishart par une matrice du GOE. En étudiant le comportement limite de chacune des composantes de la matrice de Wishart, ils constatent que celles-ci convergent vers une distribution de Rosenblatt. Ils en déduisent que leur matrice de Wishart est proche d'une matrice de Rosenblatt-Wishart lorsque $d^{\frac{3-4H}{2}}$ devient plus grand que le nombre de ligne n .

Théorème 26 (Nourdin et Zheng (2018)). Soit $s_H(k) := \frac{1}{2}(|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H})$ avec $H \in (3/4, 1)$, posons,

$$\widehat{\mathcal{W}}_{n,d}^{s_H} = d^{2-2H} \left(\frac{1}{d} \mathcal{X}_{n,d}^{s_H} (\mathcal{X}_{n,d}^{s_H})^T - \mathcal{I}_n \right)$$

et $\mathcal{R}_n^{(H)}$, une matrice de Rosenblatt-Wishart, carrée de taille n . Alors, il existe une constante finie $c_H > 0$ dépendant seulement du paramètre de Hurst H , telle que pour tout $n, d \geq 1$,

$$d_W(\widehat{\mathcal{W}}_{n,d}^{(s_H)}, \mathcal{R}_n^{(H)}) \leq c_H n d^{\frac{3-4H}{2}}.$$

Ce résultat établi dans [98] par Nourdin et Zheng est donc valable lorsque les coefficients de la matrice initiale $\mathcal{X}_{n,d}$ sont les accroissements d'un mouvement brownien fractionnaire dont le paramètre de Hurst $H \in (3/4, 1)$. Pour rappel, le seul processus gaussien auto-similaire à accroissements stationnaires est le fBm.

En relâchant l'hypothèse de stationnarité, comme dans le cas où la limite était gaussienne et en gardant le fait que les coefficients de la matrice initiale $\mathcal{X}_{n,d}$ soient auto-similaires et corrélés selon la fonction de covariance ϕ définie par (4.7), les auteurs Bourguin et Dang dans [18] établissent un théorème non central limite. Nous rappelons que bien que les accroissements ne soient pas stationnaires, ils sont proches de l'être puisque les auteurs considèrent que les éléments de la matrice initiale $\mathcal{X}_{n,d}$ sont les accroissements d'un processus gaussien Y vérifiant $\mathbb{E}[(Y_{t-s} - Y_t)^2] \sim s^\alpha$. Nous avons vu que dans le cas où $0 < \alpha \leq 3/2$ la limite était gaussienne. En revanche, pour $\alpha \in (\frac{3}{2}, 2)$, Bourguin et Dang obtiennent le résultat suivant.

Théorème 27 (Bourguin et Dang (2020)). Soit $\frac{3}{2} < \alpha < 2$. On considère la matrice de Wishart renormalisée :

$$\widehat{\mathcal{W}}_{n,d} = \frac{1}{d^{\alpha-2}} \left(\frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - \mathcal{I}_n \right).$$

Alors $\widehat{\mathcal{W}}_{n,d}$ converge au sens des lois finies dimensionnelles vers une matrice de Rosenblatt-Wishart $\mathcal{R}_n = (R_{ij})_{1 \leq i, j \leq n}$ où $R_{ij} = \lim_{d \rightarrow \infty} \widehat{W}_{ij}$ dans $L^2(\Omega)$. De plus, on obtient la borne suivante

$$d_W(\widehat{\mathcal{W}}_{n,d}, \mathcal{R}_n) \leq Cnd^{\frac{3-2\alpha}{2}}.$$

Remarque 35. 1. Bourguin et Dang obtiennent le même ordre de convergence que Nourdin et Zheng lorsque les coefficients de la matrice initiale sont les accroissements d'un mouvement brownien fractionnaire avec $\frac{3}{4} \leq H < 1$, en prenant $\alpha = 2\beta = 2H$.

2. Nous pouvons alors dire que plus l'indice d'auto-similarité est grand, plus il y a de chance pour que la matrice limite soit non gaussienne.

Dans la partie consécutive, nous discuterons du comportement asymptotique des matrices de Wishart dont les coefficients de $\mathcal{X}_{n,d}$ sont non-gaussiens. D'une part, nous verrons que la matrice limite sera gaussienne lorsque les coefficients seront totalement indépendants à l'aide d'un certain critère. D'autre part, nous constaterons une limite non-gaussienne lorsque les composantes de la matrice initiale seront corrélées.

4.2.2 Coefficients non gaussiennes

Limite gaussienne

Dans [19], Bourguin, Diez et Tudor s'intéressent au cas où les coefficients de la matrice initiale $\mathcal{X}_{n,d}$ associée à la matrice de Wishart sont indépendants et non gaussiens. Plus précisément, ils considèrent une matrice $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ dont les coefficients appartiennent à des chaos de Wiener d'ordre arbitraire associés à un processus gaussien isonormal $W = \{W(h), h \in \mathcal{H}\}$, où \mathcal{H} est un espace de Hilbert séparable. Les hypothèses sont les suivantes :

— Pour tout $1 \leq i \leq n, 1 \leq j \leq d$, les coefficients de la matrice $\mathcal{X}_{n,d}$ sont définis par :

$$X_{ij} = I_{q_i}(f_{ij}),$$

avec $f_{ij} \in \mathcal{H}^{\otimes q_i}$ et $q_i \geq 1$, pour tout $1 \leq i \leq n$. I_q est l'intégrale multiple de Wiener-Itô d'ordre q associée au processus gaussien isonormal W . Par conséquent, les éléments d'une même ligne i vivent dans le même chaos d'ordre q_i et l'ordre du chaos peut changer en passant d'une ligne à l'autre.

— Les coefficients ne sont pas identiquement distribués mais les auteurs supposent qu'ils ont même moment d'ordre deux et quatre, en d'autres termes, ils prennent ;

$$\mathbb{E}[X_{ij}^2] = 1 \text{ et } \mathbb{E}[X_{ij}^4] = m_4 < \infty.$$

— Les coefficients de $\mathcal{X}_{n,d}$ doivent être indépendants, c'est-à-dire qu'il faut que la contraction d'ordre un des noyaux des intégrales multiples de Wiener-Itô, soit égale à 0 pour que les chaos soient indépendants. Ce critère correspond au résultat obtenu par Üstünel et Zakai dans [123] que nous rappelons ci-dessous.

Théorème 28 (Üstünel-Zakai). *Pour tout $n, m \geq 1$, soit $f \in \mathcal{H}^{\odot n}$ et $g \in \mathcal{H}^{\odot m}$. Les intégrales multiples de Wiener $I_n(f)$ et $I_m(g)$ sont indépendantes si et seulement si, $f \otimes_1 g = 0$ presque partout sur $\mathcal{H}^{\otimes m+n-2}$.*

Sous ces hypothèses, Bourguin, Diez et Tudor, montrent que la matrice de Wishart dont les coefficients de $\mathcal{X}_{n,d}$ sont non-gaussiens, converge vers une matrice gaussienne. Afin de "quantifier la vitesse de convergence" de la loi de la matrice de Wishart vers une matrice limite de Wigner gaussienne dont les coefficients sont indépendants, les auteurs utilisent le théorème 12 de l'article [96] de Nourdin, Peccati, Réveillac. Il en résulte que l'approximation de la matrice de Wishart par cette matrice de Wigner est valide sous la condition que $\frac{n^3}{d} \rightarrow 0$, comme dans les cas précédemment traités où les coefficients étaient gaussiens.

Théorème 29 (Bourguin, Diez, Tudor (2021)). *Soit $\mathcal{Z}_n = (Z_{ij})_{1 \leq i, j \leq n}$, une matrice de Wigner gaussienne dont les coefficients sont indépendants (si $m_4=3$ on a une matrice GOE),*

$$\begin{cases} Z_{ii} \sim \mathcal{N}(0, m_4 - 1) & \text{pour } 1 \leq i \leq n, \\ Z_{ij} \sim \mathcal{N}(0, 1) & \text{pour } 1 \leq i < j \leq n, \\ Z_{ij} = Z_{ji} & \text{pour } 1 \leq i < j \leq n. \end{cases}$$

On considère une matrice de Wishart renormalisée $\widetilde{\mathcal{W}}_{n,d}$ associée à la matrice initiale $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$, donnée par $\widetilde{\mathcal{W}}_{n,d} := \sqrt{d} \left(\frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - \mathcal{I}_n \right)$. Alors pour tout $n \geq 1$, $\widetilde{\mathcal{W}}_{n,d}$ converge en loi, composante par composante vers la matrice \mathcal{Z}_n quand $d \rightarrow \infty$. De plus, il existe une constante C positive telle que :

$$d_W(\widetilde{\mathcal{W}}_{n,d}, \mathcal{Z}_n) \leq C \sqrt{\frac{n^3}{d}}.$$

Pour établir la preuve de cette approximation, les auteurs associent à leur matrice de Wishart le vecteur "half" et utilisent le lemme 7 afin de pouvoir utiliser la borne du théorème 12 de Nourdin, Peccati et Réveillac. Il s'agit alors de calculer la quantité suivante :

$$\sum_{i,j,a,b=1}^n \mathbb{E} \left[\left(\langle DW_{ij}, -DL^{-1}W_{ab} \rangle_{\mathcal{H}} - [Z_{ij}Z_{ab}] \right)^2 \right].$$

Les auteurs sont amenés à introduire la notion de variables aléatoires fortement indépendantes pour des chaos de Wiener.

Définition 32. *Deux variables aléatoires X et Y admettant une décomposition en chaos de Wiener,*

$$X = \sum_{n=0}^{\infty} I_n(f_n) \text{ et } Y = \sum_{m=0}^{\infty} I_m(g_m),$$

où pour tout $n, m \geq 0$, $f_n \in \mathcal{H}^{\odot n}$ et $g_m \in \mathcal{H}^{\odot m}$, sont fortement indépendantes si chaque chaos de X est indépendant de chaque composante de Y , i.e. pour tout $n, m \geq 0$, les variables aléatoires $I_n(f_n)$ et $I_m(g_m)$ sont indépendantes.

En fait, cette notion est largement utilisée pour dresser la preuve du théorème principal de [19] puisqu'une multitude de propriétés sur les coefficients de la matrice de Wishart considérée en découlent et permettent de simplifier considérablement les calculs.

En particulier, les coefficients de la matrice initiale $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ étant indépendants, les variables aléatoires X_{ik}^2 et X_{il}^2 sont fortement indépendantes par le lemme 2 de [19], ce qui va

permettre aux auteurs de majorer la quantité $\mathbb{E} \left[\left(\langle D\tilde{W}_{ii}, -DL^{-1}\tilde{W}_{ii} \rangle_{\mathcal{H}} - (m_4 - 1) \right)^2 \right]$ par $\frac{C(i)}{d}$, où $C(i)$ est une constante dépendant de i .

Les coefficients de la matrice de Wishart dans le cas non diagonal étant donnés pour $i \neq j$ par

$$\tilde{W}_{ij} = \frac{1}{\sqrt{d}} \sum_{k=1}^d X_{ik} X_{jk},$$

nous remarquons que $X_{ik} X_{jk} = I_{q_i+q_j}(f_{ik} \otimes f_{jk})$. Cela provient de l'utilisation de la formule produit, du critère d'Üstünel et Zakai et du fait que si $f \otimes_1 g = 0$ presque partout sur $\mathcal{H}^{\otimes m+n-2}$ pour $f \in \mathcal{H}^{\otimes n}$ et $g \in \mathcal{H}^{\otimes m}$ alors $f \otimes_r g = 0$ presque partout sur $\mathcal{H}^{\otimes m+n-2r}$ pour tout $1 \leq r \leq \min(n, m)$. Ainsi on obtient réellement, $X_{ik} X_{jk} = I_{q_i}(f_{ik}) I_{q_j}(f_{jk}) = I_{q_i+q_j}(f_{ik} \otimes f_{jk})$.

Grâce à cet argument, ils parviennent à majorer la quantité $\mathbb{E} \left[\left(\langle D\tilde{W}_{ij}, -DL^{-1}\tilde{W}_{ij} \rangle_{\mathcal{H}} - 1 \right)^2 \right] \leq \frac{C(i,j)}{d}$, où $C(i, j)$ est une constante dépendant de i et j .

Finalement, lorsque i, j, a et b sont mutuellement distincts, le produit scalaire est nul ;

$\langle DW_{ij}, -DL^{-1}W_{ab} \rangle_{\mathcal{H}} = 0$, dû au fait que les variables aléatoires sont indépendantes.

Les cas $i = a$ et $j \neq b$, et $i \neq a$ avec $j = b$ sont majorés de la même manière. La somme initiale se réduit à la somme $\sum_{i,j,a,b \in \mathcal{I}}$ dont le cardinal de l'ensemble \mathcal{I} défini par (4.6) est majoré par $6n^3$. Ainsi ils obtiennent le théorème 29.

Peu de temps après, ces mêmes auteurs ont étendu leurs résultats en choisissant des coefficients encore plus généraux. Dans l'article [46], Diez et Tudor ont considéré que les coefficients de la matrice initiale étaient indépendants et s'exprimaient comme des intégrales de Skorohod, de sorte que les éléments suivent une loi de probabilité très générale. Cette situation, dans laquelle les composantes constituant $\mathcal{X}_{n,d}$ ne sont pas identiquement distribuées, couvre un large nombre de cas puisque toute variable aléatoire centrée de carré intégrable peut être exprimée comme une intégrale de Skorohod. Les hypothèses considérées sont les suivantes :

1. Les coefficients sont centrés, de carré intégrable et peuvent s'écrire comme une somme infinie d'intégrales multiples stochastiques par rapport à un processus gaussien isonormal. Pour tout $1 \leq i \leq n, 1 \leq j \leq d$, on définit les coefficients de la matrice initiale $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ par,

$$X_{ij} = \sum_{p \geq 1} I_p(f_p^{(i,j)}) = \delta(u_{i,j}) \text{ avec } u_{i,j}(t) = \sum_{p \geq 1} I_{p-1} \left(f_p^{(i,j)}(\cdot, t) \right), t \in T,$$

où $T \subset \mathbb{R}$ et $f_p^{(i,j)} \in L_S^p(T^p)$. Les processus $\{u_{ij}(t), t \in T\}$ sont suffisamment réguliers au sens du calcul de Malliavin.

2. Les coefficients de $\mathcal{X}_{n,d}$ sont des variables aléatoires fortement indépendantes, c'est-à-dire que pour tout $p, q \geq 1$ et $(i, j) \neq (k, l)$, les variables aléatoires $I_p(f_p^{(i,j)})$ et $I_q(f_q^{(k,l)})$ sont indépendantes.
3. Les coefficients de la matrice initiale ont même moment d'ordre deux et d'ordre quatre,

$$\mathbb{E}[X_{ij}^2] = 1,$$

et

$$\mathbb{E}[X_{ij}^4] = m_4 < \infty.$$

Sous ces conditions, Diez et Tudor montrent que la matrice de Wishart renormalisée $\mathcal{W}_{n,d}$ associée à $\mathcal{X}_{n,d}$ converge composante par composante vers une matrice de Wigner gaussienne dont les coefficients sont indépendants.

Théorème 30 (Diez et Tudor (2021)). *Soit une matrice de Wigner $\mathcal{Z}_n = (Z_{ij})_{1 \leq i, j \leq n}$ dont les coefficients sont mutuellement indépendants :*

$$\begin{cases} Z_{ii} \sim \mathcal{N}(0, m_4 - 1) & \text{si } i = j, \\ Z_{ij} \sim \mathcal{N}(0, 1) & \text{si } i \neq j, \\ Z_{ij} = Z_{ji} & \text{pour } 1 \leq i < j \leq n. \end{cases}$$

Considérons la matrice de Wishart $\mathcal{W}_{n,d} = \sqrt{d} \left(\frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - \mathcal{I}_n \right)$ associée à la matrice initiale $\mathcal{X}_{n,d} := (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$. Alors pour tout $n \geq 1$, la matrice $\mathcal{W}_{n,d}$ converge en loi, composante par composante, vers la matrice de Wigner \mathcal{Z}_n quand $d \rightarrow \infty$. De plus, pour tout $n, d \geq 1$, il existe une constante $C > 0$ telle que ;

$$d_W(\mathcal{W}_{n,d}, \mathcal{Z}_n) \leq C \frac{n^{\frac{9}{4}}}{d^{\frac{1}{4}}}.$$

Pour établir ce résultat, les auteurs ne peuvent plus utiliser la borne usuelle du théorème 12 de Nourdin, Peccati et Réveillac. En effet, ce résultat ne permet pas d'estimer la distance de Wasserstein entre deux vecteurs aléatoires dont les composantes s'expriment comme des intégrales de Skorohod. Pour pallier ce problème, ils appliquent la proposition 2.3 de Huang, Nualart et Viitasaari [61], pour laquelle une majoration de la distance d_2 d'un vecteur aléatoire dont les composantes sont des intégrales de Skorohod et un vecteur gaussien est fournie. Afin d'obtenir une borne en distance de Wasserstein, comparable aux autres articles, ils utilisent le lemme suivant.

Lemme 9. *Si \mathcal{X} et \mathcal{Y} sont deux matrices aléatoires de $\mathcal{M}_n(\mathbb{R})$, alors*

$$d_W(\mathcal{X}, \mathcal{Y}) \leq 4n^{\frac{1}{4}} \sqrt{d_2(\mathcal{X}^{\text{half}}, \mathcal{Y}^{\text{half}})}.$$

Cette majoration est le prix à payer pour des coefficients très généraux. Ainsi une matrice de Wishart, dont les composantes de la matrice initiale sont des variables aléatoires indépendantes et de carrés intégrables, converge composante par composante vers une matrice aléatoire dont les coefficients sont gaussiens et indépendants, lorsque le nombre de colonne d est d'un plus grand ordre que n^9 . L'estimation de la borne s'appuie sur des propriétés concernant l'intégrale de Skorohod. Le principe de la preuve reste le même que dans l'article [19].

Remarque 36. 1. *L'article [19] est en fait un cas particulier de celui-ci, puisque dans [19] les variables aléatoires s'écrivent comme une somme finie de chaos indépendants.*

2. *L'article [46] ne couvre pas les mêmes résultats que dans le papier [49] de Fang et Koike. Ces derniers mesurent la distance de Wasserstein entre la loi de la matrice de Wishart vue comme un vecteur $(W_{ij}, 1 \leq i < j \leq n)$, sans considérer la diagonale de la matrice. En fait, ils utilisent la méthode des paires échangeables qui ne permet pas d'inclure la diagonale, en l'incluant, ils obtiendraient une borne moins bonne que Diez et Tudor.*

Les résultats précédents ont montré qu'en gardant l'hypothèse d'indépendance totale sur les coefficients de la matrice initiale dont ceux-ci suivent une loi de probabilité très générale, la convergence vers une matrice gaussienne tient toujours. Dans la suite, nous nous focaliserons sur un cas où l'on ajoutera de la corrélation sur des coefficients non gaussiens.

Limite non gaussienne

Dans [45], Diez et Tudor élargissent les résultats des articles [98] et [19] en se focalisant sur le comportement asymptotique d'une matrice de Wishart dont les coefficients de la matrice initiale $\mathcal{X}_{n,d}$ sont composés des accroissements du processus d'Hermite. Pour cela, notons $(Z_t^{(q,H)})_{t \geq 0}$ le processus d'Hermite d'ordre $q \geq 1$ dont l'indice d'auto-similarité est $H \in (1/2, 1)$. Nous rappelons que lorsque $q = 1$ le processus d'Hermite est un processus gaussien, qui n'est autre que le fBm et nous ramène au cas de l'article [98]. Lorsque $q \geq 2$, le processus n'est plus gaussien, en particulier pour $q = 2$, nous retrouvons le processus de Rosenblatt et le cas de l'article [19].

Diez et Tudor dans [45] autorisent une corrélation partielle sur une même ligne de la matrice comme dans les articles cités, cependant, les coefficients ne sont pas identiquement distribués. En outre, les auteurs considèrent les éléments suivants :

- Soit $B = (B^{(1)}, \dots, B^{(n)})$, un mouvement brownien n -dimensionnel. $I_q^{(i)}$ désigne l'intégrale stochastique multiple d'ordre $q \geq 1$ par rapport au processus de Wiener $B^{(i)}$ pour $i = 1, \dots, n$, tel que $B^{(i)}$ soit indépendant de $B^{(j)}$, pour $i \neq j$.
Pour tout $1 \leq i \leq n, 1 \leq j \leq d$, les coefficients de $\mathcal{X}_{n,d}$ sont donnés par :

$$X_{ij} = Z_j^{(q_i, H, i)} - Z_{j-1}^{(q_i, H, i)} = I_{q_i}^{(i)}(L_j - L_{j-1}), \text{ pour } q_i \geq 2$$

avec L_t le noyau du processus d'Hermite.

- Par conséquent, les lignes ne sont pas corrélées entre elles, puisque le processus de Wiener de la ligne i est indépendant de celui de la ligne j pour $i \neq j$.
Seuls les éléments d'une même ligne i de la matrice initiale sont corrélés et leur corrélation est déterminée par la fonction de covariance du processus d'Hermite :

$$\mathbb{E}[X_{ij}X_{ik}] = \rho_H(j-k) := \frac{1}{2} (|j-k+1|^{2H} + |j-k-1|^{2H} - 2|j-k|^{2H}).$$

- Les coefficients de $\mathcal{X}_{n,d}$ sont centrés et de variance égale à 1, en particulier, les éléments d'une même ligne i sont identiquement distribués, grâce à la propriété de stationnarité du processus d'Hermite. En revanche, la loi des variables aléatoires n'est pas forcément la même d'une ligne à une autre, puisque l'ordre q_i de l'intégrale multiple dépend de la ligne i .

L'étude des variations quadratiques du processus d'Hermite établie dans [30] permet de donner le comportement des termes diagonaux de la matrice de Wishart associée à la matrice initiale $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$. En effet, les variations quadratiques du processus d'Hermite renormalisées en fonction de l'ordre q de l'intégrale multiple, convergent dans $L^2(\Omega)$ vers une variable aléatoire de Rosenblatt. Ainsi, la renormalisation de la matrice de Wishart dépendra également de l'ordre du chaos, noté q_0 et correspondant au terme chaotique de plus bas degré.

Ainsi, les auteurs montrent que les coefficients diagonaux d'ordre q_0 convergent dans $L^2(\Omega)$ vers une variable aléatoire de Rosenblatt, tandis que les termes diagonaux appartenant à des chaos plus grands, c'est-à-dire, des termes supérieurs à q_0 convergent dans $L^2(\Omega)$ vers 0. Les termes non diagonaux, quant à eux, convergent vers 0 dans $L^2(\Omega)$.

Il en résulte que seuls les termes diagonaux dont l'ordre du chaos est le plus bas, contribuent à déterminer de manière non triviale, le comportement de la matrice aléatoire limite. Finalement, Diez et Tudor montrent que la matrice de Wishart, correctement renormalisée, converge vers une matrice diagonale dont les coefficients sont 0 ou une variable aléatoire de Rosenblatt si l'ordre du chaos de la matrice initiale était de degré le plus bas par rapport aux autres chaos.

Théorème 31 (Diez et Tudor (2021)). Soit $\mathcal{R}_n^H = (R_{ij})_{1 \leq i, j \leq n}$, une matrice diagonale dont les coefficients sont donnés par :

$$R_{ii}^H = Z_1^{(q_i, \frac{2H-2}{q_i} + 1, i)} \mathbf{1}_{q_i=q_0}.$$

On considère la matrice de Wishart renormalisée $\widetilde{\mathcal{W}}_{n,d}^0$ associée à la matrice initiale $\mathcal{X}_{n,d}$, dont les composantes sont définies par :

$$\widetilde{W}_{ii}^0 = c_{q_0, H} d^{\frac{2-2H}{q_0}-1} \sum_{k=1}^d [(Z_k^{(q_i, H, i)} - Z_{k-1}^{(q_i, H, i)})^2 - 1] \text{ pour } i = j,$$

$$\widetilde{W}_{ij}^0 = c_{q_0, H} d^{\frac{2-2H}{q_0}-1} \sum_{k=1}^d (Z_k^{(q_i, H, i)} - Z_{k-1}^{(q_i, H, i)})(Z_k^{(q_j, H, j)} - Z_{k-1}^{(q_j, H, j)}) \text{ pour } i \neq j,$$

où $c_{q_0, H}$ est une constante de renormalisation dépendant de q_0 et H . Alors pour tout $n \geq 1$, $d \geq 3$ et d assez grand :

$$d_W(\widetilde{\mathcal{W}}_{n,d}^0, \mathcal{R}_n^H) \leq C \begin{cases} nd^{\frac{2-2H}{q_0}-\frac{1}{2}} & \text{si } H \in (\frac{1}{2}, \frac{3}{4}), \\ nd^{-\frac{1}{2q_0}} & \text{si } H = \frac{3}{4}, \\ nd^{\frac{2H-2}{q_0}} & \text{si } H \in (\frac{3}{4}, 1). \end{cases}$$

Et pour $n \geq 1$, $q = 2$ et d assez grand :

$$d_W(\widetilde{\mathcal{W}}_{n,d}^0, \mathcal{R}_n^H) \leq C \begin{cases} nd^{\frac{1}{2}-H} & \text{si } H \in (\frac{1}{2}, \frac{3}{4}), \\ nd^{-\frac{1}{4}} \sqrt{\ln d} & \text{si } H = \frac{3}{4}, \\ nd^{H-1} & \text{si } H \in (\frac{3}{4}, 1). \end{cases}$$

Remarque 37. Lorsque $q_1 = \dots = q_n = 2$, nous retrouvons le résultat de l'article [19], dans cette situation les coefficients sont identiquement distribués.

Ainsi, le fait d'ajouter de la corrélation sur des coefficients non gaussiens conduit à une matrice limite non-gaussienne. Nous verrons dans l'un des articles constituant cette thèse que ce résultat peut être généralisé pour les tenseurs aléatoires dont les coefficients sont des accroissements de processus de Rosenblatt.

4.3 Résultats

4.3.1 Résumé du chapitre 8

Dans l'article [53], nous nous focalisons sur le comportement asymptotique d'une matrice de Wishart dont les coefficients de la matrice initiale sont les accroissements de la solution de l'équation des ondes dirigée par un bruit blanc en temps et en espace. Nous suggérons de considérer les cas :

- Temporel, c'est-à-dire, lorsque la variable d'espace est fixée.
- Spatial, lorsque le temps est fixé à son tour.

Cette étude se révèle intéressante puisque la structure spéciale des coefficients de la matrice initiale dans le cas temporel ne permet pas de conclure directement sur la nature de la limite de la matrice de Wishart.

Bien que les éléments soient gaussiens et totalement indépendants, les coefficients ne sont pas identiquement distribués.

Dans le cas spatial, nous verrons que le comportement asymptotique de la matrice de Wishart peut se déduire du résultat principal de l'article de Nourdin et Zheng [98]. Cependant, le théorème ne s'appliquant pas directement du fait de la structure spéciale des coefficients, nous sommes amenés à considérer une matrice de Wishart avec une renormalisation différente.

Notons que l'équation des ondes constitue un modèle pour la propagation d'une onde électromagnétique, sonore ou mécanique par exemple. Dans un système contenant plusieurs ondes, nous pouvons penser à différentes ondes se propageant dans un milieu ou dans le vide, il peut y avoir des interférences.

Notre modèle permet de mesurer la corrélation dans un tel système composé de plusieurs ondes, perturbées par une force aléatoire.

Cas Temporel. Dans une première partie, nous suggérons de fixer la variable d'espace notée $x \in \mathbb{R}$ et nous considérons le processus gaussien $(u(t, x), t \geq 0)$, solution de l'équation stochastique des ondes dirigée par un bruit blanc en temps et en espace.

Nous voulons étudier le comportement asymptotique d'une matrice de Wishart associée à la matrice initiale $\mathcal{Y}_{n,d}$ dont les coefficients sont gaussiens, totalement indépendants mais ne sont pas identiquement distribués, ni stationnaires. En particulier, nous considérons que les coefficients sont donnés par les accroissements de $(u(t, x), t \geq 0)$.

Par conséquent, les résultats [98], [97], [18] déjà évoqués ne peuvent pas s'appliquer. La forme des coefficients de la matrice de Wishart considérée, ne permet pas l'utilisation du théorème 12 puisque les coefficients correspondent au produit de deux intégrales multiples par rapport à des bruits différents.

Notons que cette étude ne rentre pas non plus dans le cadre du résultat très général de Fang et Koike [49] car les auteurs n'incluent pas la diagonale de la matrice de Wishart.

Nous détaillons de manière plus précise le contexte dans lequel nous nous plaçons. Les coefficients de la matrice initiale que nous noterons $\mathcal{Y}_{n,d} := (Y_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ sont donnés par les accroissements de la solution de l'équation des ondes dirigée par un bruit blanc en temps et en espace. Plus précisément, pour $1 \leq i \leq n, 1 \leq j \leq d$

$$Y_{ij} = u^{(i)}(j, x) - u^{(i)}(j-1, x),$$

où

$$u^{(i)}(t, x) = \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y) W^{(i)}(ds, dy),$$

avec les $W^{(i)}$ pour $1 \leq i \leq n$, des bruits blancs en temps et en espace, indépendants et $G_1(t, x) = \frac{1}{2} \mathbf{1}_{\{|x| < t\}}$. Ainsi il n'y a pas de corrélation entre différentes lignes de notre matrice initiale $\mathcal{Y}_{n,d}$ puisque les bruits sont indépendants d'une ligne à l'autre. De plus, l'analyse de la covariance des accroissements, nous donne que pour $x \in \mathbb{R}$ et pour tout $0 \leq a < b \leq s < t$,

$$\mathbb{E} \left[(u^{(i)}(t, x) - u^{(i)}(s, x))(u^{(i)}(b, x) - u^{(i)}(a, x)) \right] = 0.$$

Ce qui implique que les variables aléatoires gaussiennes Y_{ij} et Y_{ik} sont indépendantes pour $j \neq k$. Ainsi, nous nous retrouvons dans un cas d'indépendance totale.

En calculant la variance des coefficients, nous trouvons qu'ils suivent la loi gaussienne suivante, pour $1 \leq i \leq n, 1 \leq j \leq d$,

$$Y_{ij} \sim \mathcal{N}\left(0, \frac{1}{4}(2j-1)\right).$$

En l'occurrence, les coefficients ne sont pas identiquement distribués, ni stationnaires.

Remarque 38. *En bref, le fait de considérer des coefficients correspondant aux accroissements temporels de la solution de l'équation des ondes permet de révéler une structure particulière, indépendance totale, non stationnarité, non-équadistribution.*

Nous définissons les coefficients de notre matrice de Wishart renormalisée par $\overline{\mathcal{W}}_{n,d} = (\overline{W}_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$

$$\overline{\mathcal{W}}_{n,d} = \sqrt{d} \left(\frac{1}{d^2} \mathcal{Y}_{n,d} \mathcal{Y}_{n,d}^T - \frac{1}{4} \mathcal{I}_n \right).$$

Remarque 39. 1. *Le travail de renormalisation de la matrice de Wishart passe par une étude fine des variations quadratiques temporelles de la solution de l'équation stochastique des ondes. En effet, ce type de matrice étant des matrices de covariance empirique, les termes diagonaux correspondent à la variance empirique des éléments et dominent les termes de covariance empirique. L'utilisation de la formule de Cauchy-Schwarz permet de le remarquer.*

2. *Le facteur $-\frac{1}{4} \mathcal{I}_n$ signifie que nous retranchons une matrice diagonale dont les coefficients sont tous égaux à $\frac{1}{4}$ afin que toutes les éléments de la matrice de Wishart soient centrés. Cela provient du fait que l'espérance des variations quadratiques est égale à $\frac{1}{4}$;*

$$\mathbb{E} \left[\frac{1}{d^2} \sum_{k=1}^d (u(k,x) - u(k-1,x))^2 \right] = \frac{1}{4}.$$

3. *Afin d'obtenir un théorème central limite, il faut que la variance des coefficients de la matrice de Wishart soit finie et non nulle. Ainsi, nous devons renormaliser par $d^{-\frac{3}{2}}$.*

Les coefficients de $\overline{\mathcal{W}}_{n,d}$ sont donnés par :

$$\overline{W}_{ii} = \frac{1}{d^{\frac{3}{2}}} \sum_{k=1}^d \left(u^{(i)}(k+1,x) - u^{(i)}(k,x) \right)^2 - \sqrt{d} \frac{1}{4} = \frac{1}{d^{\frac{3}{2}}} \sum_{k=0}^{d-1} I_2^{(i)}(h_{k,x}), \quad (4.10)$$

et

$$\overline{W}_{ij} = \frac{1}{d^{\frac{3}{2}}} \sum_{k=1}^d \left(u^{(i)}(k+1,x) - u^{(i)}(k,x) \right) \left(u^{(j)}(k+1,x) - u^{(j)}(k,x) \right) = \frac{1}{d^{\frac{3}{2}}} \sum_{k=0}^{d-1} I_1^{(i)}(h_{k,x}) I_1^{(j)}(h_{k,x}). \quad (4.11)$$

Plus précisément, les coefficients non diagonaux de la matrice de Wishart $\overline{\mathcal{W}}_{n,d}$ sont des intégrales stochastiques multiples d'ordre 1 par rapport à différents processus gaussiens $W^{(i)}$ et $W^{(j)}$. Ainsi, la formule produit ne peut pas s'appliquer pour obtenir une intégrale stochastique multiple d'ordre 2 et comme les coefficients ne sont pas identiquement distribués ni stationnaires, le théorème 12 de Nourdin, Peccati et Réveillac, déjà largement évoqué ne peut pas être employé. Cependant, nous pouvons exprimer les coefficients de la matrice de Wishart comme une intégrale de Skorohod, telles que pour tout $1 \leq i, j \leq n$;

$$\overline{W}_{ij} = \delta^{(i)}(v^{(j)}) \text{ où } v^{(j)} = \frac{1}{d^{\frac{3}{2}}} \sum_{k=0}^{d-1} I_1^{(j)}(h_{k,x}) h_{k,x}.$$

En effet, puisque les intégrandes sont déterministes, l'intégrale de Skorohod coïncide avec l'intégrale de Wiener-Itô. Ainsi, nous proposons un nouveau critère pour établir la distance entre la loi d'un vecteur aléatoire dont les composantes s'écrivent comme des intégrales de Skorohod et un vecteur gaussien. Nous rappelons que nous notons $D^{(i)}$, la dérivée de Malliavin, $\delta^{(i)}$, l'intégrale de Skorohod et $\mathbf{D}^{k,p,i}$ l'espace de Malliavin-Sobolev associé.

Théorème 32. *Soit $F = (F_1, \dots, F_n)$ un vecteur aléatoire de dimension n tel que pour tout $1 \leq i \leq n$,*

$$F_i = \delta^{(i)}(u^{(i)}),$$

avec $u^{(i)} \in \mathbf{D}^{1,2,i}$ et $F_i \in \mathbf{D}^{1,4,i}$. Notons $\Gamma = (\Gamma_{ij})_{1 \leq i, j \leq n}$ une matrice symétrique définie positive de taille $n \times n$. Alors,

$$d_W(F, \mathcal{N}(0, \Gamma)) \leq K \sqrt{\sum_{i,j=1}^n \mathbb{E}[\Gamma_{ij} - \langle u^{(i)}, D^{(i)} F_j \rangle_{\mathcal{H}}]},$$

où \mathcal{H} est un espace de Hilbert et K une constante.

L'étude de la convergence en loi des variations quadratiques temporelles de la solution de l'équation des ondes par le théorème du quatrième moment correctement renormalisé, nous amène à la loi limite gaussienne $\mathcal{N}(0, \frac{1}{6})$. Ainsi, il en découle que la matrice limite à considérer est une matrice GOE, notée $\overline{\mathcal{Z}}_n = (\overline{Z}_{ij})_{1 \leq i, j \leq n}$ dont les éléments sont définis par ;

$$\begin{cases} \overline{Z}_{ii} \sim \mathcal{N}(0, \frac{1}{6}) & \text{pour } 1 \leq i \leq n, \\ \overline{Z}_{ij} \sim \mathcal{N}(0, \frac{1}{12}) & \text{pour } 1 \leq i < j \leq n, \\ \overline{Z}_{ij} = \overline{Z}_{ji} & \text{pour } 1 \leq j < i \leq n. \end{cases} \quad (4.12)$$

Remarque 40. *La variance des termes non-diagonaux provient du fait que $\mathbb{E}[\overline{W}_{ij}^2] = \frac{1}{12} - \frac{1}{48d^2}$.*

Afin d'obtenir l'estimation de $\overline{\mathcal{W}}_{n,d}$ par la matrice du GOE \overline{Z}_n définie par (4.12), en utilisant le théorème 32, nous devons intuitivement passer d'une distance 1-Wasserstein matricielle à une distance vectorielle. Il faut donc associer à notre matrice un vecteur dit demi-vecteur de la définition 23. En appliquant le lemme 7 de Nourdin et Zheng, majorant la distance de 1-Wasserstein des matrices aléatoires par celle des demi-vecteurs, nous pouvons employer le critère annoncé.

IDÉE DE LA PREUVE : Pour prouver la convergence de la loi du demi-vecteur de taille $\frac{n(n+1)}{2}$ associé à $\overline{\mathcal{W}}_{n,d}$ vers la loi d'un vecteur gaussien, il faut donc évaluer les quantités :

$$\mathbb{E} \left[\left(\langle v^{(j)}, D^{(i)} \overline{W}_{ab} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})} - \mathbb{E}[\overline{Z}_{ij} \overline{Z}_{ab}] \right)^2 \right].$$

1. **Cas où $i = j = a = b$.** Par la formule produit, et le fait que $\langle h_{k,x}, h_{l,x} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})} = 0$ si $k \neq l$ car les accroissements sont indépendants, on obtient,

$$\begin{aligned} \langle v^{(i)}, D^{(i)} \overline{W}_{ii} \rangle &= \frac{2}{d^3} \sum_{k,l=0}^{d-1} I_1^{(i)}(h_{k,x}) I_1^{(i)}(h_{l,x}) \langle h_{k,x}, h_{l,x} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})} \\ &= \frac{2}{d^3} \sum_{k=0}^{d-1} I_2^{(i)}(h_{k,x}^{\otimes 2}) \|h_{k,x}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 + \mathbb{E}[\langle v^{(i)}, D^{(i)} \overline{W}_{ii} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})}] \\ &= \frac{2}{d^3} \sum_{k=0}^{d-1} I_2^{(i)}(h_{k,x}^{\otimes 2}) \|h_{k,x}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 + \frac{1}{6} - \frac{1}{24d^2}. \end{aligned}$$

La dernière ligne provient du calcul de l'espérance de la quantité $\mathbb{E}[\langle v^{(i)}, D^{(i)} \overline{W}_{ii} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})}] = \mathbb{E}[\overline{W}_{ii}^2]$. Ainsi, via la propriété d'orthogonalité des chaos, et l'isométrie,

$$\begin{aligned} & \mathbb{E} \left[\left(\langle v^{(i)}, D^{(i)} \overline{W}_{ii} \rangle - \frac{1}{6} \right)^2 \right] \\ &= \mathbb{E} \left[\frac{2}{d^3} \sum_{k=0}^{d-1} I_2^{(i)}(h_{k,x}^{\otimes 2}) \|h_{k,x}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 \right]^2 + \left(\frac{1}{24d^2} \right)^2 \\ &= \frac{C}{d^6} \sum_{k=0}^{d-1} (2k+1)^4 + \left(\frac{1}{24d^2} \right)^2 \leq C \frac{1}{d}. \end{aligned}$$

Remarquons que la quantité $\|h_{k,x}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 = \mathbb{E}[(u(k+1,x) - u(k,x))^2] = 2k+1$, d'où le résultat.

2. **Cas où $i \neq a, b$ et j quelconque.** Notons

$$\mathcal{M}_0 := \{(i, j, a, b) \in \{1, \dots, n\}^4 : i \neq a \neq b \text{ et } j \text{ quelconque}\}$$

alors

$$\max_{i,j,a,b \in \mathcal{M}_0} \mathbb{E}[\langle u^{(j)}, D^{(i)} \overline{W}_{ab} \rangle^2] = 0.$$

Nous avons naturellement que $\mathbb{E}[\langle v^{(j)}, D^{(i)} \overline{W}_{a,b} \rangle^2] = 0$ lorsque $i \neq a \neq b$. En effet, dans cette situation $\overline{W}_{a,b}$ dépend uniquement des bruits $W^{(a)}$ et $W^{(b)}$ et par conséquent la dérivée de Malliavin par rapport au processus isonormal $W^{(i)}$, notée $D^{(i)}$ vaut 0.

3. **Autres cas.** Ce sont les cas où j est quelconque et $i = a$ et/ou $i = b$, nous obtenons par exemple pour $i = a$

$$\mathbb{E}[\langle u^{(j)}, D^{(i)} \overline{W}_{ib} \rangle^2] \leq C \frac{1}{d}.$$

Les autres termes sont majorés par la même quantité.

Ainsi, la borne en n^3/d vient du fait qu'il y ait $\text{Card}(\mathcal{M}_0) = n^2(n-1)(n-2)$ termes nuls donc il restera $n^4 - n^2(n-1)(n-2) \leq 3n^3$ termes non nuls.

Théorème 33. *Considérons la matrice de Wishart renormalisée $\overline{\mathcal{W}}_{n,d} = (W_{ij})_{1 \leq i,j \leq n}$ dont les coefficients sont (4.10) et (4.11). Alors quand d tend vers l'infini, la matrice $\overline{\mathcal{W}}_{n,d}$ est proche de la loi d'une matrice du GOE, $\overline{\mathcal{Z}}_n$ donnée par (4.12). De plus, pour tout $n, d \geq 1$ assez grand,*

$$d_W(\overline{\mathcal{W}}_{n,d}, \overline{\mathcal{Z}}_n) \leq C \sqrt{\frac{n^3}{d}},$$

où C est une constante.

Remarque 41. 1. *Autrement dit, lorsque le nombre de colonnes est significativement plus grand que le cube du nombre de lignes, l'approximation de notre matrice de Wishart par une matrice du GOE est valide.*

2. *Nous obtenons toujours le même ratio n^3/d alors que la renormalisation de la matrice de Wishart est $d^{3/2}$.*

Cas spatial. Nous considérons le processus gaussien $(u(t, x), x \in \mathbb{R})$ et t fixé. Nous suggérons d'analyser le comportement asymptotique d'une matrice de Wishart dont les coefficients de la matrice initiale $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ sont donnés par les accroissements spatiaux de la solution de l'équation des ondes dirigée par un bruit blanc en temps et en espace. En d'autres termes, les coefficients sont définis comme suit :

$$X_{ij} = u^{(i)}(t, j) - u^{(i)}(t, j-1) = \int_0^t \int_{\mathbb{R}} [G_1(t-s, j-y) - G_1(t-s, j-1-y)] W^{(i)}(ds, dy).$$

où $W^{(i)}, i = 1, \dots, n$ sont des copies indépendantes de W , un bruit blanc temps et espace et G_1 la solution fondamentale de l'équation des ondes homogène (sans bruit).

L'étude de la covariance des coefficients de $\mathcal{X}_{n,d}$ révèle que les accroissements sont stationnaires et tels que

$$\mathbb{E}[X_{ij} X_{i'j'}] = f_t(j-j') \mathbf{1}_{i=i'},$$

où f_t est une fonction continue sur $[0, \infty)$. Nous avons alors une structure de corrélation sur une même ligne dont la fonction de corrélation $f_t \in l^1(\mathbb{Z}) \subset l^{4/3}(\mathbb{Z})$. Par conséquent, les coefficients étant gaussiens identiquement distribués et stationnaires, avec une structure de corrélation partielle, l'étude entre dans le cas des auteurs Nourdin et Zheng. Cependant, les résultats ne peuvent pas s'appliquer directement en l'état puisque les coefficients de la matrice initiale ne vérifient pas exactement les hypothèses considérées dans [98]. Nous sommes amenés à effectuer une étude des variations quadratiques spatiales de la solution de l'équation des ondes, dans le but de trouver la bonne renormalisation.

Dans [98], les auteurs utilisent le théorème 12 pour établir leur résultat dont il est nécessaire que les composantes soient des intégrales multiples stochastiques. Les éléments sont donc des variables aléatoires centrées, ainsi nous devons recentrer les coefficients de la matrice de Wishart. La variance des éléments $\mathbb{E}[X_{ij}^2] = e_t$, une constante dépendant du temps t fixé, nous considérons ;

$$\mathcal{W}_{n,d} = \sqrt{d} \left(\frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - e_t \mathcal{I}_n \right).$$

De plus, pour répondre aux hypothèses de Nourdin et Zheng, nous introduisons un processus isonormal $(B(h), h \in \mathcal{H})$, où \mathcal{H} est un espace de Hilbert séparable, telle que pour tout $h, g \in \mathcal{H}$, $\mathbb{E}[B(h)B(g)] = \langle h, g \rangle_{\mathcal{H}}$.

Nous notons $\mathcal{A}_{n,d} = (A_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ une matrice aléatoire, de sorte que les coefficients soient égaux à $A_{ij} = B(e_{ij})$, avec $(e_{ij}, i, j \geq 1) \subset \mathcal{H}$, telle que

$$\langle e_{ij}, e_{i'j'} \rangle_{\mathcal{H}} = \mathbf{1}_{i=i'} f_t(j-j').$$

Les coefficients de la matrice $\mathcal{X}_{n,d}$ ont donc même loi que les composantes de $\mathcal{A}_{n,d}$ et par conséquent, la matrice de Wishart associée à $\mathcal{A}_{n,d}$ et $\mathcal{X}_{n,d}$ ont même loi. Nous en concluons par le théorème 23 de Nourdin et Zheng que la matrice de Wishart associée à $\mathcal{X}_{n,d}$ peut être approximée par une matrice du GOE lorsque $\frac{n^3}{d} \rightarrow 0$.

4.3.2 Résumé du chapitre 9

Dans l'article [52], nous voulons analyser l'évolution de la température entre deux points consécutifs j et $j-1$ où $j = 1, \dots, d$, d'un système composé de plusieurs tiges notées $i = 1, \dots, n$

(correspondant au nombre de lignes de la matrice initiale $\mathcal{X}_{n,d}$), soumis à une perturbation externe affectant la température de ces tiges. L'objectif est de comprendre le comportement asymptotique des corrélations dans le système.

Pour cela, nous choisissons une matrice initiale comportant des données qui coïncident avec les accroissements de la solution de l'équation de la chaleur dirigée par un bruit blanc en temps et en espace, lorsque le temps est fixé. C'est-à-dire que si nous considérons une tige, d'une certaine épaisseur, qui serait chauffée d'un côté par une première source et refroidie d'autre part par une seconde source, la température ne serait pas la même partout à l'instant initial. Nos données vont nous fournir l'évolution de la température entre deux points consécutifs de la tige, lorsque le temps est fixé. Plus le point sera proche de la source chaude, plus la température sera haute et plus le point choisi, sera proche de la source froide, plus la température sera basse. Nous noterons $u(t,x)$ le point dans la tige, au temps t fixé et à la position x . Ainsi, notre matrice initiale à n lignes et d colonnes, que nous noterons $\mathcal{X}_{n,d}$, contiendra sur une même ligne l'écart de température entre deux points consécutifs de la tige, de longueur d .

En d'autres termes, considérons $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$, une matrice aléatoire dont les coefficients sont donnés pour $1 \leq i \leq n, 1 \leq j \leq d$ par,

$$X_{ij} = u^{(i)}(t, j) - u^{(i)}(t, j-1).$$

$u^{(i)}(t,x)$ correspond à la solution de l'équation de la chaleur par rapport au bruit blanc en temps et en espace $W^{(i)}$, pour $i = 1, \dots, n$, indépendant du bruit $W^{(k)}$, pour $i \neq k$. Ainsi, nous pouvons réécrire les coefficients de la manière suivante pour tout $1 \leq i \leq n, 1 \leq j \leq d$

$$\begin{aligned} X_{i,j} &= \int_0^t \int_{\mathbb{R}} [G(t-s, j-y) - G(t-s, j-1-y)] W^{(i)}(ds, dy) \\ &= I_1^{(i)}(g_{t,j-1}), \end{aligned}$$

où G est la solution fondamentale de l'équation de la chaleur homogène et

$$g_{t,j-1} := (G(t-s, j-y) - G(t-s, j-1-y)) \mathbf{1}_{[0,t]}(s), \quad (s,y) \in \mathbb{R}_+ \times \mathbb{R}.$$

Remarquons que les coefficients de la matrice $\mathcal{X}_{n,d}$ sont gaussiens et corrélés sur une même ligne et leur corrélation est déterminée par celle des accroissements de la solution de l'équation de la chaleur dirigée par un bruit blanc temps et espace. Comme les bruits blancs sont indépendants, $W^{(i)}$ est indépendant de $W^{(k)}$ pour $i \neq k$, les lignes de la matrice $\mathcal{X}_{n,d}$ sont indépendantes entre elle. La corrélation dans la matrice n'est donc pas totale. Considérons maintenant la matrice de Wishart $\mathcal{W}_{n,d} = (W)_{1 \leq i, j \leq n}$ définie par :

$$\mathcal{W}_{n,d} = \frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T, \tag{4.13}$$

dont les coefficients sont donnés par :

$$W_{i,i} = \frac{1}{d} \sum_{k=1}^d X_{i,k}^2, \quad \text{pour } i = 1, \dots, n$$

et

$$W_{i,j} = \frac{1}{d} \sum_{k=1}^d X_{i,k} X_{j,k}, \quad \text{pour } 1 \leq i \neq j \leq n.$$

Nous noterons que $\mathbb{E}[W_{i,i}] = A_t$, une constante dépendant seulement du temps et que l'espérance des termes non-diagonaux est nulle car $X_{i,k}$ est indépendant de $X_{j,k}$ pour $i \neq j$. Les coefficients diagonaux de la matrice de Wishart correspondent aux variations quadratiques spatiales de la solution de l'équation de la chaleur et nous les analysons afin d'obtenir une convergence presque sûre. Les termes non diagonaux sont négligeables puisqu'ils tendent vers 0 dans $L^2(\Omega)$. Nous avons alors le premier résultat suivant.

Lemme 10. *Soit la matrice de $\mathcal{W}_{n,d} = (W_{i,j})_{1 \leq i,j \leq n}$ donnée par (4.13). Alors pour tout $n \geq 1$, la matrice converge composante par composante, quand $d \rightarrow \infty$, vers la matrice $A_t \mathcal{I}_n$ dans $L^2(\Omega)$ et presque sûrement, où \mathcal{I}_n est la matrice identité de taille n .*

Remarque 42. *Pour avoir la convergence presque sûre de notre matrice de Wishart $\mathcal{W}_{n,d}$, il faut que la fonction de corrélation f_t appartienne à $l^2(\mathbb{Z})$ et $\sum_{k \geq 1} k f_t(k)^2 < \infty$.*

Dans l'article de Nourdin et Zheng, ils n'établissent pas le régime classique, en fixant n , comme nous l'avons fait. Il faudrait pour cela qu'ils fixent la condition sur leur fonction de corrélation $s \in l^2(\mathbb{Z})$ et que

$$\sum_{k \geq 1} k s(k)^2 < \infty.$$

Le fait de connaître la structure de la corrélation nous permet d'effectuer cette étude.

La preuve du lemme 10 s'appuie sur l'analyse des variations quadratiques de $(u(t,x), x \in \mathbb{R})$ et $t \geq 0$ fixé.

Remarque 43. *Il existe beaucoup de résultats dans la littérature concernant l'étude des variations quadratiques de l'équation de la chaleur, cependant la partition sur laquelle nous nous focalisons, n'est pas abordée. En général les auteurs se focalisent sur une partition de l'intervalle $[0, 1]$. Dans notre cas nous regardons les variations quadratiques spatiales pour $x \rightarrow u(t,x)$ sur l'intervalle $[1, n]$.*

IDÉE DE LA PREUVE : Pour $t \geq 0$ fixé, nous définissons la suite des variations quadratiques

$$S_{N,t} := \frac{1}{N} \sum_{j=0}^{N-1} (u(t, j+1) - u(t, j))^2,$$

et telle que $\mathbb{E}[S_{N,t}] = A_t$, avec A_t une constante dépendant seulement de t . Cela vient du fait que la variance des accroissements de la solution est égale à A_t .

Il faut alors montrer que les variations quadratiques spatiales de la solution $(u(t,x), x \in \mathbb{R})$ convergent dans $L^2(\Omega)$ et presque sûrement vers A_t . Pour cela, nous remarquons que la suite $(S_{N,t})_{N \geq 0}$ s'écrit comme une intégrale stochastique d'ordre 2 d'après la formule produit et la propriété d'isométrie des intégrales stochastiques multiples. Par la suite, cela nous conduit à l'analyse de la sommabilité de la fonction de covariance. En particulier nous trouvons que :

$$\begin{aligned}
 \mathbb{E}[(S_{N,t} - A_t)^2] &= \frac{1}{N^2} \mathbb{E} \left[\sum_{j=0}^{N-1} I_2(g_{t,j}^{\otimes 2}) \right]^2 = \frac{2}{N^2} \sum_{i,j=0}^{N-1} \langle g_{t,i}, g_{t,j} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 \\
 &= \frac{2}{N^2} \sum_{i=0}^{N-1} \underbrace{[\mathbb{E}(u(t, i+1) - u(t, i))^2]}_{:=A_t} \\
 &\quad + \frac{4}{N^2} \sum_{i,j=0; i>j}^{N-1} \underbrace{[\mathbb{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j))]}_{:=f_{1,t}(i-j) + f_{2,t}(i-j)}^2 \\
 &= \frac{2A_t^2}{N} + \frac{4}{N^2} \sum_{i,j=0; i>j}^{N-1} [f_{1,t}(i-j) + f_{2,t}(i-j)]^2 \\
 &= \frac{2A_t^2}{N} + \frac{4}{N} \sum_{k=1}^{N-1} (f_{1,t}(k) + f_{2,t}(k))^2 - \frac{4}{N^2} \sum_{k=1}^{N-1} k(f_{1,t}(k) + f_{2,t}(k))^2
 \end{aligned}$$

En particulier, nous montrons que les deux dernières sommes sont finies quand $N \rightarrow \infty$ (voir le lemme 2.1 et le lemme 2.2 de l'article [52]). Ainsi nous obtenons la convergence des termes diagonaux dans $L^2(\Omega)$. L'étude préalable de la fonction de covariance des accroissements de la solution a son importance, elle appartient notamment à $l^1(\mathbb{Z})$.

Remarque 44. *Pour passer de la convergence $L^2(\Omega)$ à la convergence presque sûre nous utilisons l'inégalité de Tchebychev combinée à la propriété d'hypercontractivité des intégrales multiples. La conclusion suit en appliquant le théorème de Borel-Cantelli.*

Le deuxième point que nous avons voulu considérer dans cet article est l'erreur commise dans l'approximation de notre matrice de Wishart associée à $\mathcal{X}_{n,d}$ par une matrice du GOE.

Pour déterminer la limite asymptotique des coefficients de cette matrice de Wishart, nous avons dû étudier la convergence en loi des variations quadratiques de $(u(t, x), x \in \mathbb{R})$. En ce sens, et puisque nous voulons établir un TCL pour les variations quadratiques, nous allons nous appuyer sur le théorème 9 dénommé théorème du quatrième moment quantitatif. Pour cela, nous devons renormaliser $(S_{N,t})_{N \geq 0}$, de sorte que les coefficients soient centrés, et que la suite ait une variance finie, strictement positive.

Nous notons $(V_{N,t})_{N \geq 0}$ les variations quadratiques renormalisées. Pour $N \geq 1$ et $t > 0$ fixé,

$$V_{N,t} = \sqrt{N}(S_{N,t} - A_t) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} I_2(g_{t,j}^{\otimes 2}) = I_2 \left(\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} g_{t,j}^{\otimes 2} \right).$$

Par conséquent, $\mathbb{E}[V_{N,t}] = 0$ et la variance est telle que $\mathbb{E}[V_{N,t}^2] = a_t$, où a_t est toujours une constante qui dépend du temps t .

De ce fait, nous sommes amenés à reconsidérer la matrice de Wishart associée à $\mathcal{X}_{n,d,t}$ pour établir un TCL. Il faut donc la renormaliser de la même manière que les variations quadratiques, ainsi $\overline{\mathcal{W}}_{n,d,t} = (\overline{W}_{i,j,t})_{1 \leq i,j \leq n}$ telles que ;

$$\overline{\mathcal{W}}_{n,d,t} = \sqrt{d} \left(\frac{1}{d} \mathcal{X}_{n,d,t} \mathcal{X}_{n,d,t}^T - A_t \mathcal{I}_n \right),$$

dont les coefficients sont désormais donnés pour tout $1 \leq i \leq n$,

$$\overline{W}_{i,i,t} = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} I_2^{(i)}(g_{t,k}^{\otimes 2}), \quad (4.14)$$

et pour tout $1 \leq i, j \leq n$ avec $i \neq j$,

$$\bar{W}_{i,j,t} = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} I_1^{(i)}(g_{t,k}) I_1^{(j)}(g_{t,k}). \quad (4.15)$$

Par abus de langage, nous noterons $\|\cdot\|$ et $\langle \cdot, \cdot \rangle$ à la place de $\|\cdot\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}$ et $\langle \cdot, \cdot \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})}$. Nous voulons montrer que les variations quadratiques convergent en loi vers une variable aléatoire gaussienne. À l'aide du théorème du quatrième moment nous devons montrer que $\mathbb{E}\|DV_{N,t}\|^2 \xrightarrow{N \rightarrow \infty} 2a_t$ pour que la suite converge en loi vers une gaussienne centrée de variance a_t . Nous voulons également obtenir une borne de type Berry-Esséen pour évaluer la vitesse de convergence. Ainsi, il faut calculer les quantités

$$\sqrt{\mathbf{Var}(\|DV_{N,t}\|^2) + |\mathbb{E}(\|DV_{N,t}\|^2 - 2a_t)|}.$$

Dans un premier temps, nous vérifions que

$$\mathbb{E}\|DV_{N,t}\|^2 = 2\mathbb{E}V_{N,t}^2 \xrightarrow{N \rightarrow \infty} 2a_t.$$

Ainsi $V_{N,t}$ converge en loi vers une variable aléatoire gaussienne $\mathcal{N}(0, a_t)$. En ce qui concerne la vitesse de convergence, nous avons que

$$\begin{aligned} \mathbf{Var}(\|DV_{N,t}\|^2) &= \mathbb{E}\left[\left(\|DV_{N,t}\|^2 - \mathbb{E}\|DV_{N,t}\|^2\right)^2\right] \\ &= \frac{16}{N^2} \mathbb{E}\left(\sum_{i,j=0}^{N-1} I_2(g_{t,i} \otimes g_{t,j}) \langle g_{t,i}, g_{t,j} \rangle\right)^2 \\ &= \frac{32}{N^2} \sum_{i,j,k,l=0}^{N-1} \langle g_{t,i}, g_{t,j} \rangle \langle g_{t,k}, g_{t,l} \rangle \langle g_{t,i}, g_{t,k} \rangle \langle g_{t,j}, g_{t,l} \rangle. \end{aligned}$$

Dans le lemme 2.4 de [52], la dernière somme est majorée par $C_t N$, où C_t est une constante qui dépend du temps. Finalement,

$$\mathbb{E}\left[\left(\|DV_{N,t}\|_{\mathcal{H}}^2 - \mathbb{E}\left[\|DV_{N,t}\|_{\mathcal{H}}^2\right]\right)^2\right] \leq \frac{C_t}{N}.$$

Cela nous permet de définir la matrice gaussienne limite $\bar{\mathcal{L}}_n$ dont tous les coefficients sont indépendants, à savoir,

$$\begin{cases} \bar{Z}_{ii} \sim \mathcal{N}(0, a_t) & \text{pour } 1 \leq i \leq n, \\ \bar{Z}_{ij} \sim \mathcal{N}(0, \frac{a_t}{2}) & \text{pour } 1 \leq i < j \leq n, \\ \bar{Z}_{ij} = \bar{Z}_{ji} & \text{pour } 1 \leq j < i \leq n. \end{cases} \quad (4.16)$$

La matrice initiale $\mathcal{X}_{n,d}$ ayant des coefficients gaussiens stationnaires et partiellement corrélés dont la fonction de corrélation des accroissements de la solution de l'équation de la chaleur $f_t \in l^1(\mathbb{Z}) \subset l^{4/3}(\mathbb{Z})$, nous pouvons appliquer le théorème de Nourdin et Zheng [98].

Soit \mathcal{H} un espace de Hilbert réel séparable muni de son produit scalaire $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ et soit $\{e_{i,j}, i, j \geq 1\} \subset \mathcal{H}$ une famille telle que

$$\langle e_{i,j}, e_{i',j'} \rangle = \mathbf{1}_{\{i=i'\}} f_t(j - j').$$

Nous notons $\{B(h), h \in \mathcal{H}\}$ un processus isonormal indexé par \mathcal{H} tel que $\mathbb{E}[B(h_1)B(h_2)] = \langle h_1, h_2 \rangle_{\mathcal{H}}$ pour tout $h_1, h_2 \in \mathcal{H}$.

De cette manière, nous pouvons définir une matrice $\mathcal{C}_{n,d} = (C_{i,j})_{1 \leq i \leq n, 1 \leq j \leq d}$, dont les coefficients sont donnés par $C_{i,j} = B(e_{i,j})$, et nous considérons la matrice de Wishart associée à $\mathcal{C}_{n,d}$

$$\mathcal{W}_{n,d,\mathcal{C}} = \sqrt{d} \left(\frac{1}{d} \mathcal{C}_{n,d} \mathcal{C}_{n,d}^T - A_t \mathcal{I}_n \right).$$

Les coefficients des matrices étant de même loi et ayant la même structure de corrélation, nous avons $\mathcal{X}_{n,d,t} \stackrel{(d)}{=} \mathcal{C}_{n,d}$ and $\overline{\mathcal{W}}_{n,d,t} \stackrel{(d)}{=} \mathcal{W}_{n,d,\mathcal{C}}$, où $\stackrel{(d)}{=}$ est l'égalité au sens des distributions finies dimensionnelles.

Considérons également une matrice gaussienne $\mathcal{G}_n = (G_{i,j})_{1 \leq i, j \leq n}$ de même matrice de covariance que $\overline{\mathcal{W}}_{n,d,t}$ et $\mathcal{W}_{n,d,\mathcal{C}} = (W_{i,j,C})_{1 \leq i, j \leq n}$, i.e. pour tout $1 \leq i, j, k, l \leq n$,

$$\mathbb{E}[G_{i,j}G_{k,l}] = \mathbb{E}[\overline{W}_{i,j,t}\overline{W}_{k,l,t}] = \mathbb{E}[W_{i,j,C}W_{k,l,C}].$$

Alors, nous trouvons que

$$d_W(\overline{\mathcal{W}}_{n,d,t}, \overline{\mathcal{Z}}_{n,t}) = d_W(\mathcal{W}_{n,d,\mathcal{C}}, \overline{\mathcal{Z}}_{n,t}) \leq d_W(\mathcal{W}_{n,d,\mathcal{C}}, \mathcal{G}_n) + d_W(\mathcal{G}_n, \overline{\mathcal{Z}}_{n,t}).$$

Par le résultat de Nourdin et Zheng [98], nous pouvons dire que

$$d_W(\mathcal{W}_{n,d,\mathcal{C}}, \mathcal{G}_n) \leq C_t \sqrt{\frac{n^3}{d}}. \quad (4.17)$$

En utilisant les vecteurs half associés aux matrices \mathcal{G}_n et $\overline{\mathcal{Z}}_{n,t}$ et la majoration du lemme 7, nous obtenons la borne suivante

$$d_W(\mathcal{G}_n, \overline{\mathcal{Z}}_{n,t}) \leq k \sqrt{\frac{n^2}{d}}, \quad (4.18)$$

avec k une constante. Il en résulte le théorème suivant.

Théorème 34. *Soit $\overline{\mathcal{W}}_{n,d,t}$, la matrice de Wishart renormalisée, dont les coefficients sont donnés par (4.14) et (4.15). Alors pour tout $n \geq 1$, la matrice $\overline{\mathcal{W}}_{n,d,t}$ est proche de la loi de la matrice $\overline{\mathcal{Z}}_n$, définie par (4.16), quand $d \rightarrow \infty$. De plus, il existe une constante C_t positive telle que, pour tout $n, d \geq 1$ suffisamment grands,*

$$d_W(\overline{\mathcal{W}}_{n,d,t}, \overline{\mathcal{Z}}_n) \leq C_t \sqrt{\frac{n^3}{d}}.$$

Le comportement asymptotique de notre matrice de Wishart indique que lorsque le nombre de points consécutifs considéré d devient beaucoup plus grand que le cube du nombre de tiges étudié, cela revient à considérer une matrice du GOE dont les coefficients sont totalement indépendants.

Remarque 45. *Nous obtenons un résultat similaire dans le cas où nous considérons que les coefficients de la matrice initiale sont donnés par les accroissements temporels de la solution de l'équation de la chaleur dirigée par un bruit blanc en temps et en espace. Dans ce cas, les accroissements ne sont pas stationnaires mais la solution se comporte comme un mouvement Brownien bi-fractionnaire, i.e :*

$$(u(t, x), t \geq 0) \stackrel{(d)}{\equiv} \left(\pi^{-\frac{1}{4}} B_t^{\frac{1}{2}, \frac{1}{2}}, t \geq 0 \right),$$

où $(B_t^{\frac{1}{2}, \frac{1}{2}}, t \geq 0)$ est le mouvement Brownien bi-fractionnaire de paramètres de Hurst $H = K = \frac{1}{2}$. Le comportement de telles matrices de Wishart a été étudié par Bourguin et Dang dans l'exemple 1 de [18].

Étude asymptotique des tenseurs aléatoires

Nous avons vu que les matrices de Wishart ont été largement étudiées. Cependant, pour le besoin des applications en statistique ou en machine learning, il est nécessaire d'exploiter une extension de celles-ci, qui est basée sur l'analyse des tenseurs aléatoires.

En effet, ces objets jouent un rôle important dans le domaine de l'apprentissage automatique lorsque les jeux de données sont extrêmement conséquents et soumis à de l'incertitude ou à des variations aléatoires. Ils peuvent permettre une initialisation de modèles ; par exemple, pour entraîner des réseaux de neurones, il est courant d'initialiser les poids des neurones de manière aléatoire afin d'éviter que le modèle ne se coince dans un minimum local.

Ils peuvent aussi être utilisés pour générer des données synthétiques pour des expériences comme dans [38] ou pour augmenter des bases de données afin de généraliser des modèles en appliquant des transformations aléatoires aux données d'entraînement.

5.1 Introduction aux tenseurs aléatoires

Nous venons d'énoncer quelques applications des tenseurs aléatoires. Dans cette partie, nous proposons dans un premier temps de détailler plus explicitement leur nécessité et leur utilisation. Par exemple, ces objets peuvent évaluer l'état de la distribution d'un réseau. Dans l'article [38], les auteurs veulent détecter et localiser les éventuelles anomalies dans un réseau en se basant sur l'observation des données. Les anomalies sont détectées en analysant les variations de corrélation des données. Ils s'appuient sur l'étude spectrale des matrices aléatoires de covariance empirique. En particulier, lorsqu'il y a une anomalie sur le réseau, modélisée par un bruit, la distribution spectrale empirique des données ne suit plus une loi de Marchenko-Pastur dû aux valeurs propres extrêmes. En revanche, dans un cas stable, c'est-à-dire sans anomalie, la distribution spectrale empirique d'une matrice de Wishart doit converger vers une loi de Marchenko-Pastur.

Ils sont donc amenés à étudier la statistique linéaire des valeurs propres basée sur une fonction test. L'étude de cette statistique permet d'indiquer le positionnement de l'anomalie dans le réseau. Les auteurs remarquent également que lorsque les jeux de données sont trop petits, la méthode ne se révèle pas assez efficace et fournissent un algorithme permettant d'augmenter les jeux de données en utilisant le produit tensoriel. L'étude des tenseurs se montre finalement plus efficace puisqu'elle permet de détecter plus tôt les anomalies dans le réseau.

L'exploitation de la structure des tenseurs aléatoires peut aussi fournir des estimations de paramètres pour des modèles à variables aléatoires latentes. Pour rappel, les modèles à variables latentes sont des modèles statistiques qui visent à expliquer les relations entre les variables observées, en introduisant des variables latentes. Ces variables dites cachées sont supposées avoir une influence sur les données observées. Ces modèles sont utilisés dans divers domaines, tels que la statistique, en machine learning, la psychométrie et bien d'autres. Dans l'article [4], les auteurs réduisent l'estimation des paramètres à la décomposition orthogonale d'un tenseur symétrique.

Pour la nécessité d'un tel domaine applicatif, la théorie des tenseurs aléatoires a dû se développer progressivement. On s'est interrogé s'il était possible d'élargir l'étude spectrale des matrices aléatoires à celles dont les coefficients sont obtenus par le produit tensoriel de vecteurs aléatoires. Ce problème a notamment été traité en théorie de l'informatique quantique. En particulier, les auteurs Ambainis, Harrow et Hastings [3] se sont penchés sur le sujet.

En se basant sur des techniques similaires à la théorie des matrices aléatoires, la méthode de la trace et d'autres méthodes introduites dans [3], ils obtiennent une propriété limite sur la plus grande valeur propre de leur matrice et la convergence de la mesure spectrale empirique associée. Ils ont donc élargi les résultats connus sur la convergence de la distribution spectrale empirique d'une matrice de Wishart vers une loi de Marchenko-Pastur. Plus précisément, les auteurs choisissent p -vecteurs aléatoires unitaires de $(\mathbb{C}^d)^{\otimes k}$, notés $|\varphi_1\rangle, \dots, |\varphi_p\rangle$.

Ambainis, Hastings et Harrow considèrent dans [3] une matrice de covariance empirique ou matrice contenant des moments empiriques d'une collection de produit de vecteurs aléatoires,

$$\mathcal{M}_{p,d,k} = \sum_{s=1}^p \varphi_s,$$

où φ_s correspond au produit du vecteur $|\varphi_s\rangle$ par son adjoint $|\varphi_s\rangle^* = \langle \varphi_s|$, tel que $\varphi_s := |\varphi_s\rangle\langle \varphi_s|$, et $|\varphi_s\rangle = |\varphi_s^1\rangle \otimes |\varphi_s^2\rangle \otimes \dots \otimes |\varphi_s^k\rangle$, avec $|\varphi_s^1\rangle, \dots, |\varphi_s^k\rangle \in \mathbb{C}^d$. Les vecteurs aléatoires $|\varphi_s^a\rangle$ sont des états individuels uniformément distribués sur la sphère unité \mathbb{C}^d .

En notant $\lambda_1, \dots, \lambda_R$ les valeurs propres non nulles de $\mathcal{M}_{p,d,k}$, avec $R = \text{Rg } \mathcal{M}_{p,d,k}$ (en général, $R = \min(p, d^k)$ et les valeurs propres sont distinctes), ils établissent le résultat suivant.

Théorème 35 (Ambainis, Harrow et Hastings (2010)). *Presque sûrement, la mesure spectrale empirique de $\mathcal{M}_{p,d,k}$ converge étroitement vers une loi de type Marchenko-Pastur de paramètre c tel que $\frac{p}{d^k} \rightarrow c \in (0, \infty)$, quand $d \rightarrow \infty$, et $k > 1$:*

$$\frac{\sqrt{((1 + \sqrt{c})^2 - x)(x - (1 - \sqrt{c})^2)}}{2\pi cx} \mathbf{1}_{(1 - \sqrt{c})^2 \leq x \leq (1 + \sqrt{c})^2}.$$

En d'autres termes, nous considérons X_1, \dots, X_k , k -vecteurs aléatoires i.i.d de \mathbb{R}^d et nous désignons \mathbf{Y} , le vecteur aléatoire de \mathbb{R}^{d^k} , obtenu en effectuant le produit tensoriel des X_1, \dots, X_k ;

$$\mathbf{Y} = X_1 \otimes \dots \otimes X_k.$$

Introduisons maintenant p copies du vecteur aléatoire \mathbf{Y} , notées $\{\mathbf{Y}_s\}_{s=1, \dots, p}$. La matrice $\mathcal{M}_{p,d,k}$ est obtenue en effectuant la somme de tous les produits des vecteurs de \mathbb{R}^{d^k} , \mathbf{Y}_s et sa transposée \mathbf{Y}_s^T , pour $s = 1, \dots, p$. La matrice $\mathcal{M}_{p,d,k} := (M_{ij})_{1 \leq i, j \leq d^k}$ est donc une matrice de covariance empirique de taille $d^k \times d^k$ puisque les coefficients sont donnés par ;

$$\begin{cases} M_{ii} = \sum_{s=1}^p \left(Y_s^{(i)} \right)^2 & \text{si } i = j, \\ M_{ij} = \sum_{s=1}^p Y_s^{(i)} Y_s^{(j)} & \text{si } i \neq j, \end{cases} \quad (5.1)$$

et dont les termes $Y_s^{(i)}$ correspondent à la i -ème composante du vecteur aléatoire \mathbf{Y}_s de \mathbb{R}^{d^k} tels que ;

$$Y_s^{(i)} = \prod_{l=1}^k X_l^{(i_l)} \quad \text{où } i \text{ est un } k\text{-multi-indice } i = \{i_1, \dots, i_k\} \text{ pour } i_1, \dots, i_k \in [d].$$

D'après le résultat de Ambainis, Harrow et Hastings la mesure spectrale empirique de cette matrice vérifie en partie le théorème 19 de Marchenko-Pastur de paramètre c .

Remarque 46. *Ces considérations ont été largement étudiées et nous appelons cette version ; version non symétrique des modèles de tenseurs aléatoires. Pour une généralisation de ces résultats, nous renvoyons à l'article de Lytova [85]. Plus précisément, cet auteur a étudié le comportement asymptotique de la distribution spectrale empirique d'une matrice de Wishart dont les coefficients sont obtenus par le produit tensoriel de vecteurs aléatoires corrélés et a établi un TCL pour la statistique linéaire des valeurs propres liées à un tenseur de Wishart.*

Plus tard, les chercheurs ont porté leur intérêt sur le d -produit tensoriel $x^{\otimes d}$ d'un vecteur aléatoire $x \in \mathbb{R}^n$ dont les composantes sont indépendantes à la place de considérer le produit tensoriel de d vecteurs aléatoires i.i.d x_1, \dots, x_d de \mathbb{R}^n comme dans [3], [114] et [85], c'est-à-dire $x_1 \otimes \dots \otimes x_d$.

L'objectif étant de montrer, une nouvelle fois, la convergence vers la loi de Marchenko-Pastur pour la mesure spectrale empirique d'une matrice de Wishart composée de coefficients donnés par des tenseurs aléatoires vectorisés. Dans ce cas, le modèle utilisé se nomme version symétrique des modèles de tenseurs aléatoires.

Les prérequis sont de considérer un vecteur aléatoire isotrope (centré de matrice de covariance I_d) noté $x = (x_j)_{1 \leq j \leq n}$ de \mathbb{R}^n dont les composantes sont indépendantes. En effectuant le d -produit tensoriel $x^{\otimes d}$, on obtient un vecteur de \mathbb{R}^{n^d} , dont on va retirer tous les éléments diagonaux et symétriques, de sorte que le vecteur obtenu $\mathbf{x} \in \mathbb{R}^{\binom{d}{n}}$. De cette manière la j -ème composante est donnée par,

$$\mathbf{x}_{\mathbf{J}} = \prod_{j \in \mathbf{J}} x_j = x_{j_1} x_{j_2} \dots x_{j_d}, \quad \text{où } \mathbf{J} = \{\{j_1, \dots, j_d\} \in \{1, \dots, n\}^d : j_1 < j_2 < \dots < j_d\}.$$

Un résultat de Bryson, Vershynin et Zhao [22], montre que presque sûrement la distribution spectrale empirique de la matrice de Wishart $\mathscr{W}_{p,m} := \frac{1}{m} \mathscr{Y}_{p,m} \mathscr{Y}_{p,m}^T$ associée à la matrice initiale $\mathscr{Y}_{p,m}$ dont les colonnes sont formées des vecteurs $\mathbf{x} \in \mathbb{R}^p$, où $p = \binom{d}{n}$ et sont indépendantes, converge étroitement vers une loi de Marchenko-Pastur de paramètre $\lambda \in (0, \infty)$. Ce résultat est valide lorsque $d = o(n^{\frac{1}{3}})$ et $\frac{p}{m} \rightarrow \lambda \in (0, \infty)$ pour $p \rightarrow \infty$.

Remarque 47. *Ce modèle est généré par peu de variables aléatoires indépendantes puisque les auteurs ont utilisé un seul vecteur de \mathbb{R}^n pour générer les composantes de la matrice initiale. Les colonnes de la matrice initiale correspondant au produit tensoriel du vecteur $x \in \mathbb{R}^n$ dont les composantes sont indépendantes, ne le sont donc plus pour le vecteur $\mathbf{x} \in \mathbb{R}^{\binom{d}{n}}$, ce qui constitue davantage de difficultés.*

Le champ de recherche concernant les tenseurs aléatoires s'est peu à peu étendu pour se rapprocher des thèmes abordés dans la théorie des matrices aléatoires.

En 2019, Vershynin [124] s'intéresse aux inégalités de concentration. En même temps, les auteurs Jiang et Xie [69] se focalisent sur la distribution de l'entrée maximale hors diagonale, d'un tenseur aléatoire. En effet, cette grande question s'est aussi posée sur l'étude des matrices de Wishart, et en particulier dans le papier [67] de Jiang sorti en 2004. L'auteur avait déjà montré que la plus grande entrée d'une matrice de covariance empirique convergeait en loi vers

une distribution de Gumbel sous des conditions assez fortes, qui avait fait l'objet de recherches intensives pour relâcher les hypothèses imposées sur les moments.

Dans [69], l'idée est de généraliser cette étude sur les matrices de Wishart afin de l'élargir aux tenseurs aléatoires. Plus précisément, ils considèrent n vecteurs aléatoire de \mathbb{R}^p , noté $X_k = (x_{k1}, \dots, x_{kp})^T$, pour $1 \leq k \leq n$. Les vecteurs X_1, \dots, X_n sont des vecteurs aléatoires supposés indépendants et de même loi. Soit $\mathbf{T} \in \mathbb{R}^{p^m}$, un tenseur aléatoire tel que :

$$\mathbf{T} = \sum_{k=1}^n \underbrace{X_k \otimes \dots \otimes X_k}_{m \text{ produits}} = \left(\sum_{k=1}^n x_{ki_1} x_{ki_2} \dots x_{ki_m} \right)_{1 \leq i_1, \dots, i_m \leq p}.$$

En analysant le comportement asymptotique de l'entrée maximale W_n du tenseur aléatoire \mathbf{T} dont les termes diagonaux ont été ôtés telle que,

$$W_n := \max_{1 \leq i_1 < \dots < i_m \leq p} \frac{1}{\sqrt{n}} \left| \sum_{k=1}^n x_{ki_1} x_{ki_2} \dots x_{ki_m} \right|,$$

Jiang et Xie prouvent que $W_n^2 - 2m \log p + \log \log p \xrightarrow{(d)} \theta$, si $p \rightarrow \infty$ et $\log p = o(n^{\frac{\alpha}{2m-\alpha}})$ quand $n \rightarrow \infty$ avec $0 < \alpha < 1$, sous des conditions concernant les moments des coefficients de X_k . θ suit une distribution de Gumbel, $F_\theta(z) = \exp\left(-\frac{1}{m! \sqrt{m\pi}} e^{-\frac{z}{\theta}}\right)$, $z \in \mathbb{R}$. Comme la dimension de p peut être d'un ordre exponentiel de n , ce régime est appelé *ultra-high dimensional case*. Sous des hypothèses différentes, les auteurs montrent encore la convergence de l'entrée maximale du tenseur \mathbf{T} vers une loi de Gumbel, pour le régime *high dimensional case*.

L'axe de recherche sur lequel nous allons plus particulièrement nous centraliser est celui de l'approximation de la loi d'un tenseur aléatoire vu comme un vecteur aléatoire par la loi d'un vecteur gaussien. Dans la suite, nous introduisons ces nouvelles considérations et nous énonçons quelques résultats obtenus lorsque nous posons différentes conditions sur les coefficients ; gaussianité ou composantes non gaussiennes, indépendance totale ou ajout de corrélation.

5.2 Analyse non spectrale des tenseurs de Wishart

Dans ce nouveau chapitre, nous suggérons de présenter dans un premier temps une analyse des tenseurs dits de Wishart, en utilisant la méthode de Malliavin-Stein. Nous évoquerons ensuite des résultats basés sur une application de la méthode de Stein.

5.2.1 Étude des tenseurs de Wishart par la méthode de Stein-Malliavin

Soit $\mathbb{X}_j = (X_{1j}, \dots, X_{nj})^T$, un vecteur colonne qui peut-être vu comme la j -ème colonne de la matrice initiale $\mathcal{X}_{n,d}$. Soit $\{e_1, \dots, e_n\}$ la base canonique de \mathbb{R}^n , nous pouvons réécrire le vecteur \mathbb{X}_j comme

$$\mathbb{X}_j = \sum_{i=1}^n X_{ij} e_i.$$

Le produit tensoriel d'ordre p de $\mathbb{X}_j \in \mathbb{R}^n$ est un vecteur de taille \mathbb{R}^{n^p} et défini par la formule suivante :

$$\mathbb{X}_j^{\otimes p} = \sum_{i_1, \dots, i_p=1}^n \left(\prod_{k=1}^p X_{i_k j} \right) e_{i_1} \otimes \dots \otimes e_{i_p}. \quad (5.2)$$

Par exemple, pour $p = 2$, nous obtenons un vecteur de taille \mathbb{R}^{n^2} , à savoir :

$$\mathbb{X}_j^{\otimes 2} = (X_{1j}^2, X_{1j}X_{2j}, \dots, X_{1j}X_{nj}, X_{2j}X_{1j}, X_{2j}^2, \dots, X_{2j}X_{nj}, \dots, X_{nj}X_{1j}, X_{nj}X_{2j}, \dots, X_{nj}^2)^T.$$

En particulier, en prenant $\frac{1}{\sqrt{d}} \sum_{j=1}^d \mathbb{X}_j^{\otimes 2}$, nous remarquons que ce vecteur colonne correspond à une matrice de Wishart dont les colonnes ont été mises bout à bout. C'est dans ce contexte que va se baser l'étude des tenseurs dits de Wishart. Nous considérons donc :

$$\frac{1}{\sqrt{d}} \sum_{j=1}^d \mathbb{X}_j^{\otimes p} = \sum_{i_1, \dots, i_p=1}^n \frac{1}{\sqrt{d}} \sum_{j=1}^d \left(\prod_{k=1}^p X_{i_k j} \right) e_{i_1} \otimes \dots \otimes e_{i_p},$$

qui peut être aussi vu comme le vecteur aléatoire $\mathbf{Y} \in \mathbb{R}^{n^p}$ donné par :

$$\left(\mathbf{Y}_i = \frac{1}{\sqrt{d}} \sum_{j=1}^d \prod_{k=1}^p X_{i_k j}, 1 \leq i_1, \dots, i_p \leq n \right).$$

Dans la suite, nous étudierons ce vecteur sans les termes diagonaux, c'est-à-dire, nous considérons le tenseur

$$\mathcal{Y}_{n,d} = \left(\mathbf{Y}_i := \frac{1}{\sqrt{d}} \sum_{j=1}^d \prod_{k=1}^p X_{i_k j}, j \in \Delta_p \right).$$

où l'ensemble Δ_p est tel que

$$\Delta_p := \{ \{i_1, \dots, i_p\} \in \{1, \dots, n\}^p : i_1, \dots, i_p \text{ sont mutuellement distincts} \}$$

et dont le cardinal vaut $\text{Card}\Delta_p = n(n-1) \cdots (n-p+1) = p! \binom{n}{p}$.

Plusieurs chercheurs se sont intéressés au comportement asymptotique de ces tenseurs de Wishart en prenant différentes hypothèses sur les composantes. La plus naturelle, qui fut l'une des premières études en ce sens, est de considérer des coefficients totalement indépendants et gaussiens. Dans [98], en 2018, Nourdin et Zheng traitent cette situation, en suggérant que les coefficients sont des variables aléatoires gaussiennes, indépendantes et identiquement distribuées. En particulier, lorsque $\frac{n^{2p-1}}{d} \rightarrow 0$, le tenseur $\mathcal{Y}_{n,d}$ est proche de la loi d'un vecteur gaussien standard de $\mathbb{R}^{p! \binom{n}{p}}$.

Théorème 36 (Nourdin et Zheng (2018)). *Soit $p \geq 2$ et Z_v un vecteur gaussien standard de taille $v = \text{Card}\Delta_p$. Alors nous avons,*

$$d_W(\mathcal{Y}_{n,d}, Z_v) = \mathcal{O} \left(\sqrt{\frac{n^{2p-1}}{d}} \right).$$

Pour établir ce théorème, Nourdin et Zheng utilisent la même méthode que pour les matrices de Wishart. En effet, ils vont analyser les termes diagonaux supérieurs de $\mathcal{Y}_{n,d}$, de la même manière que les auteurs examinaient le demi-vecteur associé à la matrice de Wishart. Ils se focalisent donc sur le vecteur aléatoire à valeurs dans $\mathbb{R}^{\binom{n}{p}}$, défini comme :

$$\mathcal{Y}_{n,d}^\uparrow = (\mathbf{Y}_i, i \in \Delta_p^\uparrow),$$

où $\Delta_p^\uparrow = \{i \in \Delta_p : i_1 < i_2 < \dots < i_p\}$. Ils fournissent ensuite une extension du lemme 7 ; ils montrent en particulier que la distance de Wasserstein entre $\mathcal{Y}_{n,d}$ et le vecteur gaussien standard de \mathbb{R}^V est majorée par la distance de Wasserstein entre $\mathcal{Y}_{n,d}^\uparrow$ et \mathbf{Z}_V^\uparrow , multipliée par $\sqrt{p!}$, dépendant de la taille du tenseur. Cette constante, assez intuitive dans le cas des matrices, l'est toujours pour les tenseurs puisqu'elle correspond aux nombres d'éléments identiques. Pour rappel, les matrices de Wishart étant symétriques, il y avait les mêmes termes en dessous et au dessus de la diagonale d'où la constante $\sqrt{2}$. Pour les tenseurs, il y a $p!$ fois chaque terme en dehors de la diagonale d'où le résultat. En d'autres termes, nous avons que,

$$d_W(\mathcal{Y}_{n,d}, \mathbf{Z}_V) \leq \sqrt{p!} d_W(\mathcal{Y}_{n,d}^\uparrow, \mathbf{Z}_V^\uparrow). \quad (5.3)$$

Ainsi, en exploitant cette borne et en se servant du théorème 12 de Nourdin, Peccati et Réveillac, les auteurs obtiennent leurs résultats par un argument combinatoire.

Quelques années plus tard, en 2021, Nourdin et Pu ajoutent une condition de corrélation totale sur les composantes du tenseur aléatoire $\mathcal{Y}_{n,d}$. Dans leur papier [97], ils prennent les mêmes hypothèses que pour les matrices de Wishart. Ayant déjà été largement introduites dans la partie précédente, nous rappelons qu'ils considèrent un espace de Hilbert \mathcal{H} muni de son produit scalaire $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, de sa norme $\|\cdot\|_{\mathcal{H}}$ et d'une famille $\{e_{ij} : i, j \geq 1\} \subset \mathcal{H}$ telle que pour tout $i, j \geq 1$, $\|e_{ij}\|_{\mathcal{H}} = 1$ et

$$\langle e_{ij}, e_{kl} \rangle_{\mathcal{H}} = r(i-k)s(j-l).$$

$s, r : \mathbb{Z} \rightarrow \mathbb{R}$ sont deux fonctions de corrélation satisfaisant $s(0) = r(0) = 1$. Ils associent à \mathcal{H} , un processus isonormal $X = \{X(h), h \in \mathcal{H}\}$ de sorte que les coefficients soient gaussiens et identiquement distribués $X_{ij} = X(e_{ij}) \sim \mathcal{N}(0, 1)$. Les composantes sont corrélées puisque la famille $\{e_{ij} : i, j \geq 1\} \subset \mathcal{H}$ n'est pas orthogonale, et on a $\mathbb{E}[X_{ij}X_{kl}] = r(i-k)s(j-l)$. Par conséquent, sous ces hypothèses les composantes du tenseurs aléatoires $\mathcal{Y}_{n,d}$ sont totalement corrélées. Nourdin et Pu parviennent à montrer que sous la condition (5.4), la matrice de covariance \tilde{C} du vecteur $\mathcal{Y}_{n,d}^\uparrow$ est à diagonale strictement dominante donc inversible par le théorème de Hadamard. Il en résulte la borne quantitative suivante pour la distance de Wasserstein entre le tenseur $\mathcal{Y}_{n,d}$ et un vecteur gaussien de \mathbb{R}^V .

Théorème 37 (Nourdin et Pu (2021)). *Soit $\tilde{\mathcal{Y}}_{n,d}$ et $\mathbf{Z} = (\mathbf{Z}_i : i \in \Delta_p)$ un vecteur gaussien de \mathbb{R}^V qui a la même matrice de covariance que $\tilde{\mathcal{Y}}_{n,d}$. Supposons que la fonction de covariance r satisfasse :*

$$\left(1 - (\|r\|_{l^1(\mathbb{Z})} - 1) \left(p! \|r\|_{l^1(\mathbb{Z})}^{p-1} + \frac{p! - 1}{2}\right)\right) > 0. \quad (5.4)$$

Alors il existe une constante $C_p > 0$ dépendant uniquement de p et $\|r\|_{l^1(\mathbb{Z})}$, telle que :

$$d_W(\tilde{\mathcal{Y}}_{n,d}, \mathbf{Z}) \leq C_p \sqrt{\frac{d}{|\sum_{k,l=1}^d s(k-l)^p|} \left(\sum_{|k| \leq d} |s(k)|^{\frac{4}{3}}\right)^3 \frac{n^{2p-1}}{d}}.$$

Remarque 48. 1. Si la fonction de corrélation $s \in l^{\frac{4}{3}}(\mathbb{Z})$ et si p est pair ou $s(k) \geq 0$ pour tout $k \in \mathbb{Z}$, on obtient la même borne que pour le cas d'indépendance totale, dans [98].

2. La preuve de ce théorème s'appuie sur la borne 5.3, l'utilisation du bien connu théorème 12, du fait que la matrice de covariance \tilde{C} soit inversible sous la condition (5.4).

Dans les deux cas que nous avons mentionné, les coefficients étaient toujours identiquement distribués. Dans la pratique, cette condition est assez restrictive et il peut sembler pertinent d'étudier des situations où les composantes ne suivent pas la même loi. En particulier, dans l'article [47], Dhoyer et Tudor traitent le cas où les coefficients initiaux sont indépendants et appartiennent à un chaos de Wiener, dont l'ordre de ce chaos change d'une ligne à l'autre. Ce travail peut être vu comme une extension de l'article [19], que nous avons déjà évoqué dans le cadre des matrices aléatoires.

Nous suggérons de rappeler que lorsque les composantes de la matrice initiale associée à la matrice de Wishart étaient posées de cette manière, elle convergait toujours vers une matrice de Wigner gaussienne. Nous pouvons donc nous attendre à ce que le tenseur aléatoire $\mathcal{Y}_{n,d}$ converge vers un certain vecteur gaussien.

\mathcal{H} est un espace de Hilbert séparable auquel on associe un processus isonormal $(W(h), h \in \mathcal{H})$. Ils considèrent une matrice initiale $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ dont les coefficients sont donnés pour tout $1 \leq i \leq n$ et $1 \leq j \leq d$ par ; $X_{ij} = I_{q_i}(f_{i,j})$, avec $q_1, \dots, q_n \geq 1$ et $f_{i,j} \in \mathcal{H}^{\otimes q_i}$. Dhoyer et Tudor supposent que les coefficients soient mutuellement indépendants mais pas identiquement distribués puisque l'ordre du chaos q_i peut changer d'une ligne à l'autre. Néanmoins, ils choisissent de prendre des coefficients dont les moments d'ordre deux sont égaux, en particulier, pour tout $1 \leq i \leq n, 1 \leq j \leq d$, $\mathbb{E}[X_{ij}^2] = 1$.

Remarque 49. *Contrairement à l'article [19], il n'est pas exigé que les moments d'ordre 4 soient finis et égaux puisque dans ce papier les auteurs n'incluent pas les termes diagonaux du tenseur.*

Sous ces conditions Dhoyer et Tudor vont analyser le tenseur d'ordre p associé à cette matrice, noté :

$$\mathcal{Y}_{n,d}^p = (Y_{i_1, \dots, i_p}, i_1, \dots, i_p \in \Delta_p).$$

Notifions que les composantes de ce tenseur aléatoire peuvent s'exprimer comme une intégrale multiple d'ordre $q_{i_1} + \dots + q_{i_p}$, en appliquant la formule produit et le critère d'Üstünel et Zakai (puisque les chaos sont indépendants) ;

$$Y_{i_1, \dots, i_p} = \frac{1}{\sqrt{d}} \sum_{j=1}^d I_{q_{i_1} + \dots + q_{i_p}}(f_{i_1, j} \otimes \dots \otimes f_{i_p, j}).$$

A l'aide de propriétés sur l'indépendance des chaos de Wiener et de l'utilisation du théorème 12, ils parviennent à établir que le tenseur aléatoire $\mathcal{Y}_{n,d}^p$ converge en loi vers un vecteur gaussien, lorsque $\frac{n^{2p-1}}{d} \rightarrow 0$.

Théorème 38 (Dhoyer et Tudor (2023)). *Considérons le vecteur $\mathcal{Y}_{n,d}^p = (Y_{j_1, \dots, j_p}, (j_1, \dots, j_p) \in \Delta_p)$ et $\mathbf{Z}_v = (Z_a, a \in \mathbb{R}^v)$. Alors pour tout $n \geq 1$, $\mathcal{Y}_{n,d}^p$ converge en loi quand $d \rightarrow \infty$ vers le vecteur gaussien v -dimensionnel, \mathbf{Z}_v et pour n assez grand,*

$$d_W(\mathcal{Y}_{n,p}^p, \mathbf{Z}_v) \leq C \sqrt{\frac{n^{2p-1}}{d}}.$$

Dans cette thèse, nous proposons d'étendre ces résultats sur les tenseurs aléatoires en choisissant des coefficients non gaussiens et corrélés sur une même ligne dont la corrélation est celle d'un processus de Rosenblatt. Avant de proposer un résumé de l'article [54], nous évoquons d'autres résultats obtenus sur les tenseurs aléatoires dont les coefficients sont plus généraux, et basés sur une utilisation innovante de la méthode de Stein.

5.2.2 Étude des tenseurs de Wishart basée sur une nouvelle utilisation de la méthode de Stein

Bien que Mikulincer ait proposé dans [90] des résultats sur les matrices aléatoires dont les coefficients étaient log-concaves, il a aussi voulu étendre les résultats de Bubeck et Ganguly [21] en considérant des tenseurs de Wishart à la place des matrices de covariance empirique.

Soient $(X_i)_{1 \leq i \leq d}$ des variables aléatoires i.i.d et μ une mesure isotrope sur \mathbb{R}^n telles que $X_i \sim \mu$. Notons $X_i^{\odot p}$ le produit tensoriel symétrique d'ordre $p \geq 2$, le tenseur de Wishart est donné par (4.5). Nous noterons $\tilde{W}_{n,d}^p(\mu)$ la loi de (4.5) dont les termes diagonaux ont été ôtés et les coefficients suivent la loi μ . Comme il l'avait prouvé pour les matrices de Wishart, Mikulincer montre que la structure de mesure produit de μ n'est pas nécessaire pour obtenir l'approximation de $\tilde{W}_{n,d}^p(\mu)$ par une loi gaussienne, en obtenant la même vitesse de convergence $\frac{n^{2p-1}}{d}$ que Nourdin et Pu [97] et Nourdin et Zheng [98]. En gardant les mêmes hypothèses et la même méthode que pour le cas de matrices de covariance, que nous avons évoqué dans la sous-section consacrée à l'approche non-spectrale des matrices de Wishart, Mikulincer obtient l'extension aux tenseurs aléatoires dont les coefficients admettent une dépendance et sont log-concaves. Il garde la condition que μ soit aussi inconditionnelle afin de parvenir à majorer la plus petite valeur propre de la matrice de covariance qui intervient dans le théorème 22 et lui donne le résultat.

Il propose également d'autres perspectives concernant l'étude des tenseurs aléatoires (théorème 3 dans [90]) en reprenant la condition d'indépendance des coefficients; il impose que μ soit une mesure produit. En particulier, il montre que l'approximation du tenseur de Wishart par une gaussienne est valide lorsque la distribution est à queue lourde, de type sous-exponentielle. En effet, il va le montrer pour μ log-concave mais aussi appartenant à une classe plus large : toutes les marginales de μ doivent satisfaire l'inégalité de Poincaré.

Remarque 50. *Toutes les mesures log-concaves vérifient l'inégalité de Poincaré. Il va donc élargir la condition de log-concavité sur la mesure μ , à la classe des distributions sous-exponentielles qui sont à queues lourdes et satisfont aussi l'inégalité de Poincaré (et même plus).*

Pour montrer ce résultat, Mikulincer utilise son principal théorème 22 qui permet d'estimer la discrèpence de Stein (4.3) à l'aide d'une borne effective se reposant sur la structure initiale composant le tenseur aléatoire, c'est-à-dire un vecteur de \mathbb{R}^n . Dans la suite C est une constante qui varie d'une ligne à l'autre. Nous rappelons également les notations suivantes;

- $X \sim \mu$ est un vecteur aléatoire isotrope (i.e : centré de matrice de covariance égale à l'identité) de \mathbb{R}^n .
- $G \sim \mathcal{N}(0, I_d)$, un vecteur aléatoire gaussien standard de \mathbb{R}^n .
- $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$, localement lipschitzienne, nous avons $X \stackrel{\text{loi}}{=} \phi(G)$.
- $A : \text{Sym}^p(\mathbb{R}^n) \rightarrow V \subset \text{Sym}^p$ est une transformation linéaire de manière à ce que la loi de $\tilde{W}_{n,d}^p(\mu)$ soit isotrope.

Dans la borne du théorème 22, il est amené à calculer $\sqrt{\mathbb{E} \left[\|X\|_2^{8(p-1)} \right]}$.

Cette quantité est immédiatement majorée par n^{2p-2} , à une constante près puisque les lois log-concaves et les lois μ dont toutes les marginales satisfont l'inégalité de Poincaré vérifient

$$\mathbb{E}[\|X\|_2^m] \leq C \mathbb{E}[\|X\|_2^2]^{\frac{m}{2}}.$$

Dans cette partie, tous les coefficients étant indépendants (puisque μ est une mesure produit), $\tilde{W}_{n,d}^p(\mu)$ est directement isotrope et n'a pas besoin d'être renormalisée. Par conséquent $\|A\|_{\text{op}} = 1$ et il ne reste plus qu'à estimer $\sqrt{\mathbb{E}[\|\text{Jac } \phi(G)\|_{\text{op}}^8]}$.

Comme tous les éléments du vecteur initial X sont indépendants, dans cette situation,

$$\|\text{Jac } \phi(x)\|_{\text{op}} \leq C(1 + \|x\|_{\infty}).$$

Finalement, Mikulincer obtient que $\sqrt{\mathbb{E}[\|\text{Jac } \phi(G)\|_{\text{op}}^8]} \leq C \log(n)^2$ pour le cas log concave et $\sqrt{\mathbb{E}[\|\text{Jac } \phi(G)\|_{\text{op}}^8]} \leq C \log(n)^4$ pour la classe plus générale dont les marginales de μ satisfont l'inégalité de Poincaré.

Finalement, malgré la grande généralité des résultats de Mikulincer qui permettent de donner une approximation d'un tenseur aléatoire dont les coefficients sont indépendants et de distribution à queue lourde de type sous-exponentielle, par une loi gaussienne, les auteurs Fang et Koike [49] relâchent complètement l'hypothèse concernant la loi des éléments.

Ils suggèrent que leur résultat sur les matrices de covariance empirique - évoqué dans la section concernant l'approche non-spectrale pour l'étude des matrices de Wishart - est encore valide pour les tenseurs aléatoires.

En utilisant la méthode des paires échangeables, ils proposent l'approximation d'un tenseur de Wishart, vu comme un vecteur dont les coefficients ont une loi très générale, par une loi gaussienne standard multivariée qui est valide lorsque $d \gg n^{2p-1}$. Fang et Koike supposent l'indépendance et l'équidistribution des coefficients. En ce sens, ils améliorent les résultats de [98] et [90]. Plus précisément, les considérations sont les suivantes,

- Soient $(X_{i_a k})_{1 \leq a \leq p, 1 \leq k \leq d}$, une suite de variables aléatoires centrées, réduites, indépendantes et identiquement distribuées, de moment d'ordre quatre fini $\mathbb{E}[X_{11}^4] < \infty$. Notons

$$\mathbf{W} = \{W_{i_1, \dots, i_p} : 1 \leq i_1 < \dots < i_p \leq n\},$$

un tenseur aléatoire dont les coefficients sont donnés, pour $p \geq 2$ par,

$$W_{i_1 \dots i_p} = \frac{1}{\sqrt{d}} \sum_{k=1}^d X_{i_1 k} \dots X_{i_p k}.$$

- Fang et Koike considèrent le tenseur aléatoire \mathbf{W} comme un vecteur de $\mathbb{R}^{\binom{n}{p}}$. Cela signifie que si nous nous plaçons dans le cas $p = 2$, nous obtenons une matrice de Wishart dont les colonnes ont été mises bout à bout et dans laquelle nous avons retiré les termes diagonaux; les termes $(X_{ii}^2)_{1 \leq i \leq n}$.
- Il en résulte alors que \mathbf{W} peut être approché par un vecteur gaussien standard de $\mathbb{R}^{\binom{n}{p}}$ lorsque $d \gg n^{2p-1}$. Par conséquent les auteurs retrouvent la même approximation et la même condition sur n et d que Nourdin et Zheng mais en relâchant l'hypothèse de gaussianité et de distribution à queue lourde de Mikulincer [90].

Dans les résultats précédemment étudiés, nous avons vu que l'approximation de la loi d'un tenseur aléatoire par une loi gaussienne multivariée était possible lorsque les coefficients étaient ;

1. gaussiennes avec une corrélation totale [97] (dont le cas particulier de l'indépendance totale est traité dans [98]),
2. log-concaves et partiellement corrélées [90],
3. et totalement indépendantes, de loi très générale comme dans [47], [90] et [49].

L'hypothèse de non-gaussianité des coefficients est rarement combinée à la condition de corrélation. Dans cette thèse, nous proposons de nous pencher sur un cas dont les coefficients sont partiellement corrélés et non gaussiens en incluant l'hyper-diagonale du tenseur aléatoire. Nous élargissons ainsi les résultats de l'article [19] en trouvant qu'une approximation de la loi de ce tenseur aléatoire par un vecteur dont les composantes sont des variables aléatoires de Rosenblatt est valide.

Les résultats évoqués jusqu'à présent n'incluaient pas l'hyper-diagonale des tenseurs aléatoires dû fait d'une plus grande complexité lors de l'étude de ceux-ci. Mikulincer dans [90] a largement évoqué le problème de l'inclusion des termes diagonaux.

Dans la suite, nous proposons de résumer l'article [54] concernant l'étude d'un tenseur de Wishart dont les composantes de la matrice initiale sont des accroissements de processus de Rosenblatt.

5.3 Résultats : Résumé du chapitre 10

Dans ce chapitre, nous proposons un résumé de l'article [54] qui traite de l'approximation d'un tenseur aléatoire, vu comme un vecteur aléatoire de \mathbb{R}^{n^p} dont les coefficients sont non-gaussiens et partiellement corrélés, par un vecteur non-gaussien dont les composantes sont indépendantes. Nous considérons une matrice initiale $\mathcal{X}_{n,d} = (X_{ij})_{1 \leq i \leq n, 1 \leq j \leq d}$ de taille $n \times d$ dont les coefficients sont des accroissements de processus de Rosenblatt. Nous rappelons que ce sont des processus non gaussiens à accroissements stationnaires et auto-similaires, vivant dans le second chaos de Wiener. Ainsi, les coefficients de la matrice initiale peuvent donc être exprimés comme des intégrales stochastiques multiples d'ordre deux.

Pour cela, nous nous donnons n processus de Wiener (bilatéraux) $(B^{(i)}(y), y \in \mathbb{R})$ pour $i = 1, \dots, n$ tels que ;

$$X_{ij} = Z_j^{H,i} - Z_{j-1}^{H,i} = I_2^{(i)}(L_j^H - L_{j-1}^H), \quad 1 \leq i \leq n, 1 \leq j \leq d,$$

où $Z^{H,1}, \dots, Z^{H,n}$ constituent une famille de processus de Rosenblatt indépendants puisque les $(B^{(i)}(y), y \in \mathbb{R})$ sont indépendants pour $i = 1, \dots, n$ et L^H est le noyau du processus de Rosenblatt (5.5), tel que pour tout $y_1, y_2 \in \mathbb{R}$, $H \in (\frac{1}{2}, 1)$ et $d(H)$ une constante dépendant de H ;

$$L_t^H(y_1, y_2) = d(H) \int_0^t (u - y_1)_+^{\frac{H}{2}-1} (u - y_2)_+^{\frac{H}{2}-1} du. \quad (5.5)$$

La notation $I_2^{(i)}$ désigne une intégrale stochastique multiple d'ordre deux par rapport au processus de Wiener $(B^{(i)}(y), y \in \mathbb{R})$.

Remarque 51. 1. *Sous ces considérations, il est aisé de constater que nos coefficients sont corrélés sur une même ligne i puisque ce sont les accroissements d'un processus de Rosenblatt par rapport au processus de Wiener $(B^{(i)}(y), y \in \mathbb{R})$. Leur corrélation est donnée pour tout $1 \leq i, i' \leq n, 1 \leq j, j' \leq d$;*

$$\mathbb{E}[X_{ij} X_{i'j'}] = \rho_H(j - j') \mathbf{1}_{i=i'},$$

$$\text{où } \rho_H(v) = \frac{1}{2} (|v+1|^{2H} + |v-1|^{2H} - 2|v|^{2H}).$$

2. *Les coefficients sur différentes lignes sont indépendants puisque ce sont des intégrales multiples par rapport à des processus de Wiener indépendants.*

3. Un résultat de Bourguin, Diez et Tudor [19] fournit l'approximation d'une matrice de Wishart associée à $\mathcal{X}_{n,d}$ par une matrice diagonale dont les composantes sont soit des variables aléatoires de Rosenblatt ou bien sont nulles. En ce sens, nous voulons élargir ce résultat aux tenseurs de Wishart.

Par conséquent, nous associons à $\mathcal{X}_{n,d}$ un tenseur aléatoire et nous étudions son comportement asymptotique lorsque le nombre de lignes n et de colonnes d deviennent très grands.

Construction du tenseur de Wishart

Soit $\mathbb{X}_i = (X_{1i}, \dots, X_{ni})^T$, la i -ième colonne de $\mathcal{X}_{n,d}$. Nous nous intéressons au produit tensoriel d'ordre p noté $\mathbb{X}_i^{\otimes p} \in \mathbb{R}^{n^p}$ du vecteur $\mathbb{X}_i \in \mathbb{R}^n$ donné par (5.2). Plus précisément, nous considérons un p -tenseur de Wishart centré :

$$\sum_{i=1}^d \mathbb{X}_i^{\otimes p} - \mathbb{E} \left[\mathbb{X}_i^{\otimes p} \right] = \sum_{j_1, \dots, j_p=1}^n \sum_{i=1}^d \left(\prod_{k=1}^p X_{j_k i} - \mathbb{E} \left[\prod_{k=1}^p X_{j_k i} \right] \right) e_{j_1} \otimes \dots \otimes e_{j_p},$$

où $\{e_{j_1}, \dots, e_{j_p}\}$ est une base de \mathbb{R}^{n^p} . Nous utiliserons plus couramment la représentation vectorielle du tenseur, pour $j_1, \dots, j_p \in \{1, \dots, n\}$

$$Y_{j_1, \dots, j_p} = \sum_{i=1}^d \left(\prod_{k=1}^p X_{j_k i} - \mathbb{E} \left[\prod_{k=1}^p X_{j_k i} \right] \right). \quad (5.6)$$

L'objectif de cet article est d'analyser le comportement asymptotique dans le régime à grandes dimensions (*High dimensional regime*) du vecteur aléatoire \mathbf{Y} de \mathbb{R}^{n^p} ,

$$\mathbf{Y} = (Y_{j_1, \dots, j_p}, j_1, \dots, j_p = 1, \dots, n). \quad (5.7)$$

Nous allons montrer que le tenseur aléatoire correctement renormalisé vu comme un vecteur de \mathbb{R}^{n^p} peut être assimilé à un vecteur non-gaussien \mathbf{R} de \mathbb{R}^{n^p} dont les coefficients sont des variables aléatoires de Rosenblatt sur l'hyper-diagonale et nuls en dehors de celle-ci, lorsque le d devient très grand.

Nous définissons le vecteur limite \mathbf{R} comme

$$\mathbf{R} = (R_{j_1, \dots, j_p}, j_1, \dots, j_p = 1, \dots, n), \quad (5.8)$$

où pour chaque $j_1, \dots, j_p = 1, \dots, n$,

$$R_{j_1, \dots, j_p} = \begin{cases} K(H, p) Z_1^H & \text{si } j_1 = \dots = j_p, \\ 0 & \text{si } j_1, \dots, j_p \text{ ne sont pas tous identiques,} \end{cases}$$

où

$$K(H, p) := d(H)^{p-1} \beta \left(\frac{1}{2}, 1 \right)^{p-1} \sum_{\substack{\alpha \in \mathcal{A}_{p,2} \\ |\alpha|=p-1}} C_\alpha C_{1,\alpha}, \quad (5.9)$$

avec $d(H)$, la constante de renormalisation du processus de Rosenblatt définie par (1.7) et β correspond à la fonction bêta de Euler. Les autres notations sont détaillées ensuite.

Nous évaluons également la distance de 1-Wasserstein vectorielle associée à cette convergence. Pour effectuer l'étude du comportement asymptotique de \mathbf{Y} , nous analysons avec précision l'hyper-diagonale du tenseur aléatoire, c'est-à-dire, lorsque $j_1 = j_2 = \dots = j_p$, qui correspond

aux p -variations du processus de Rosenblatt. L'autre partie sera destinée à l'analyse des termes en dehors de l'hyper-diagonale, i.e : le cas où les j_1, \dots, j_p ne sont pas tous égaux.

Traitement de l'hyper-diagonale : Étude des p -variations du processus de Rosenblatt

La première étape consiste à étudier le comportement asymptotique de l'hyper-diagonale du tenseur, c'est-à-dire, le comportement de $Y_{j, \dots, j}$ pour $j = j_1 = \dots = j_p$ donné par (5.6) ;

$$Y_{j, \dots, j} = \sum_{i=1}^d (X_{ji})^p - \mathbb{E}[(X_{ji})^p]. \quad (5.10)$$

Comme nous le disions, ceci revient à examiner les p -variations d'un processus de Rosenblatt. En effet, en réécrivant $Y_{j, \dots, j}$ et en utilisant la propriété d'auto-similarité du processus de Rosenblatt, nous avons

$$\begin{aligned} Y_{j, \dots, j} &= \sum_{i=1}^d \left(Z_i^{H,j} - Z_{i-1}^{H,j} \right)^p - \mathbb{E} \left(Z_i^{H,j} - Z_{i-1}^{H,j} \right)^p \\ &\stackrel{(d)}{=} d^{Hp} \sum_{i=1}^d \left(Z_{\frac{i}{d}}^{H,j} - Z_{\frac{i-1}{d}}^{H,j} \right)^p - \mathbb{E} \left(Z_{\frac{i}{d}}^{H,j} - Z_{\frac{i-1}{d}}^{H,j} \right)^p \\ &= d^{Hp} \sum_{i=1}^d (I_2(f_{i,d}))^p - \mathbb{E} (I_2(f_{i,d}))^p, \end{aligned}$$

où le noyau est donné par

$$f_{i,d}(y_1, y_2) = d(H) \int_{\frac{i-1}{d}}^{\frac{i}{d}} (u - y_1)_+^{\frac{H}{2}-1} (u - y_2)_+^{\frac{H}{2}-1} du. \quad (5.11)$$

Afin d'étudier les p -variations, notre stratégie consiste à développer $(I_2(f_{i,d}))^p$ en chaos de Wiener et analyser séparément le comportement asymptotique de chacun des chaos ainsi obtenus. Pour calculer le produit effectué p fois d'intégrales stochastiques d'ordre deux $(I_2(f_{i,d}))^p$, nous utilisons une généralisation de la formule produit. Ce résultat a été obtenu par les auteurs Garino, Nourdin, Nualart et Salamat dans [58] qui ont développé une extension de la formule produit pour le n -produit d'intégrales multiples d'ordre q .

EXTENSION DE LA FORMULE PRODUIT

Soient deux entiers $n, q \geq 2$ et h_1, \dots, h_n , n fonctions appartenant à $L^2_{\mathbb{S}}(\mathbb{R}^q)$, l'ensemble des fonctions symétriques de $L^2(\mathbb{R}^q)$. L'intérêt de cette extension de la formule produit est de noter $I_q(h_1) \times \dots \times I_q(h_n)$ comme une somme d'intégrales stochastiques multiples. Nous considérons les notations suivantes ;

- Soit $\mathcal{A}_{n,q}$ un ensemble de multi-indices $\alpha = (\alpha_{ij}, 1 \leq i < j \leq n)$ tel que pour tout $k = 1, \dots, n$,

$$\sum_{1 \leq i < j \leq n} \alpha_{ij} \mathbf{1}_{k \in \{i, j\}} \leq q.$$

Les α_{ij} représentent le nombre de variables dans h_i qui sont contractées avec h_j . La longueur de α est donnée par $|\alpha| = \sum_{1 \leq i < j \leq n} \alpha_{ij}$.

- Le nombre de variables dans h_k qui ne sont pas contractées est noté $\beta_k^0 = q - \sum_{1 \leq i < j \leq n} \alpha_{ij} \mathbf{1}_{k \in \{i, j\}}$,

pour $1 \leq k \leq n$ et remarquons que $\sum_{k=1}^n \beta_k^0 = nq - 2|\alpha|$.

Nous notons aussi que $\beta_k = \sum_{k=1}^n \beta_k^0$ pour $k = 1, \dots, n$ et $\beta_0 = 0$ par convention.

— Pour $1 \leq i, j \leq n$, tel que $i \neq j$, soit

$$u^{i,j} = \alpha_{\min(i,j), \max(i,j)}.$$

Pour tout $\alpha \in \mathcal{A}_{n,q}$ nous pouvons définir la multi-contraction

$$\begin{aligned} & \otimes_{\alpha}(h_1, \dots, h_n)(\xi_1, \dots, \xi_{nq-2|\alpha|}) \\ &= \int_{\mathbb{R}^{|\alpha|}} \prod_{k=1}^n h_k \left(s_1^{k,1}, \dots, s_{u^{k,1}}^{k,1}, \dots, s_1^{k,n}, \dots, s_{u^{k,n}}^{k,n}, \xi_{1+\beta_{k-1}}, \dots, \xi_{\beta_k} \right) \prod_{1 \leq i < j \leq n} ds_1^{i,j} \dots ds_{u^{i,j}}^{i,j}. \end{aligned}$$

Notons que cette multi-contraction n'est pas nécessairement symétrique.

Sous ces considérations, nous pouvons énoncer une généralisation de la formule produit.

Lemme 11 (Garino, Nourdin, Nualart et Salamat (2021)). *Soient $n, q \geq 2$ des entiers et $h_i \in L_S^2(\mathbb{R}^q)$ pour $i = 1, \dots, n$. Nous avons,*

$$\prod_{k=1}^n I_q(h_k) = \sum_{\alpha \in \mathcal{A}_{n,q}} C_{\alpha} I_{nq-2|\alpha|}(\otimes_{\alpha}(h_1, \dots, h_n)), \quad (5.12)$$

où

$$C_{\alpha} = \frac{q^{|\alpha|}}{\prod_{k=1}^n \beta_k^{0!} \prod_{1 \leq i < j \leq n} \alpha_{ij}!}.$$

En utilisant le lemme 11 énoncé ci-dessus, et en prenant $n = p$ et $q = 2$ nous pouvons réécrire $I_2(f_{i,d})^p$ comme

$$I_2(f_{i,d})^p = \sum_{\alpha \in \mathcal{A}_{p,2}} C_{\alpha} I_{2p-2|\alpha|}(\otimes_{\alpha}(f_{i,d}, \dots, f_{i,d})).$$

- Remarque 52.**
1. La formule produit généralisée nous fournit alors une somme d'intégrales stochastiques multiples allant de l'ordre 2 à l'ordre $2p$ plus un terme constant qui est "l'intégrale multiple d'ordre zéro" et correspond à $\mathbb{E}[I_2(f_{i,d})^p]$.
 2. L'intégrale d'ordre 2 n'ayant pas le même comportement asymptotique que les termes d'ordre supérieur, nous étudions ce cas de manière isolée. Les termes d'ordre supérieur sont en fait négligeables par rapport à celui-ci lorsqu'on effectue la renormalisation puisqu'ils ont une limite triviale dans $L^2(\Omega)$.
 3. Il est nécessaire de calculer la multi-contraction $\otimes_{\alpha}(f_{i,d}, \dots, f_{i,d})(\xi_1, \xi_2, \dots, \xi_{2p-2|\alpha|})$, afin de pouvoir manipuler le noyau des intégrales multiples stochastiques et établir la convergence du chaos d'ordre deux vers une variable aléatoire de Rosenblatt.

Pour plus de confort, nous notons les p -variations renormalisées du processus de Rosenblatt auxquelles nous avons appliqué la formule produit généralisée comme suit,

$$\tilde{V}_{1,N} = N^{Hp-H} \sum_{i=0}^{N-1} (I_2(f_{i,N}))^p - \mathbb{E}(I_2(f_{i,N}))^p = N^{-H} \sum_{k=0}^{p-1} \sum_{\substack{\alpha \in \mathcal{A}_{p,2} \\ |\alpha|=k}} C_{\alpha} V_{N,\alpha}^{(2p-2|\alpha|)}, \quad (5.13)$$

où

$$V_{N,\alpha}^{(2p-2|\alpha|)} := N^{Hp} \sum_{i=0}^{N-1} I_{2p-2|\alpha|}(\otimes_{\alpha}(f_{i,N}, \dots, f_{i,N})). \quad (5.14)$$

Dans notre article, nous montrons le résultat suivant, nécessaire à l'analyse du comportement asymptotique de l'hyper-diagonale de notre tenseur. Il permet de donner la convergence dans $L^2(\Omega)$ des p -variations du processus Rosenblatt vers une variable aléatoire de Rosenblatt.

Théorème 39. *Considérons la suite $(\tilde{V}_{1,N}, N \geq 1)$ donnée par (5.13). Alors,*

$$\tilde{V}_{1,N} \xrightarrow[N \rightarrow \infty]{L^2(\Omega)} K(H, p) Z_1^H,$$

où Z^H est un processus de Rosenblatt de paramètre de Hurst $H \in (\frac{1}{2}, 1)$ et $K(H, p)$, une constante donnée par (5.9). De plus pour N assez grand,

$$d_W \left(\tilde{V}_{1,N}, K(H, p) Z_1^H \right) \leq c(H, p) \begin{cases} N^{\frac{1}{2}-H}, & \text{si } H \in (\frac{1}{2}, \frac{3}{4}), \\ \sqrt{\log(N)} N^{-\frac{1}{4}} & \text{si } H = \frac{3}{4}, \\ N^{H-1} & \text{si } H \in (\frac{3}{4}, 1). \end{cases}$$

Idée de la preuve

La convergence vers la variable aléatoire de Rosenblatt provient de la forme particulière du noyau du chaos d'ordre 2. Nous allons donc montrer que,

1. Le chaos d'ordre 2 converge vers une variable aléatoire de Rosenblatt.
2. Les chaos d'ordre supérieur convergent vers 0 dans $L^2(\Omega)$.

ANALYSE DU CHAOS D'ORDRE 2

La projection de $\tilde{V}_{1,N}$ sur le chaos d'ordre 2 est obtenue lorsque $|\alpha| = p - 1$. Nous notons cette projection

$$\tilde{V}_N^{(2)} = \sum_{\substack{\alpha \in \mathcal{A}_{p,2} \\ |\alpha|=p-1}} C_\alpha N^{-H} N^{Hp} \sum_{i=0}^{N-1} I_2(\otimes_\alpha(f_{i,N}, \dots, f_{i,N})).$$

De cette manière la somme $\beta_1^0 + \dots + \beta_p^0 = 2$, et nous remarquons qu'il y a au plus deux indices β_k^0 non nuls pour $k = 1, \dots, p$. Ainsi nous parvenons à écrire la multi-contraction de la manière suivante,

$$\begin{aligned} \otimes_\alpha(f_{i,N}, \dots, f_{i,N})(\xi_1, \xi_2) &= d(H)^p \beta \left(\frac{H}{2}, 1-H \right)^{p-1} \int_{\frac{i}{N}}^{\frac{i+1}{N}} \dots \int_{\frac{i}{N}}^{\frac{i+1}{N}} du_1 \dots du_p \\ &\quad \left(\prod_{1 \leq a < b \leq p} |u_a - u_b|^{(H-1)\alpha_{ab}} \right) (u_r - \xi_1)_+^{\frac{H}{2}-1} (u_s - \xi_2)_+^{\frac{H}{2}-1}, \end{aligned}$$

pour $r, s = 1, \dots, p$, où $f_{i,N}$ est définie par (5.11) et $d(H)$ par (1.7). En considérant que $\tilde{V}_N^{(2)}$ peut aussi s'écrire comme ;

$$\tilde{V}_N^{(2)} = d(H)^p \beta \left(\frac{1}{2}, 1 \right)^{p-1} \sum_{\substack{\alpha \in \mathcal{A}_{p,2} \\ |\alpha|=p-1}} C_\alpha I_2(g_{N,\alpha}),$$

où

$$g_{N,\alpha} := N^{-H} N^{Hp} \sum_{i=0}^{N-1} \int_{\frac{i}{N}}^{\frac{i+1}{N}} \dots \int_{\frac{i}{N}}^{\frac{i+1}{N}} du_1 \dots du_p \left(\prod_{1 \leq a < b \leq p} |u_a - u_b|^{(H-1)\alpha_{ab}} \right) (u_r - \xi_1)_+^{\frac{H}{2}-1} (u_s - \xi_2)_+^{\frac{H}{2}-1},$$

nous allons chercher à montrer que $g_{N,\alpha}$ converge vers le noyau du processus de Rosenblatt. Pour cela, nous allons, dans un premier temps, réécrire $g_{N,\alpha}$ en effectuant un changement de

variables $\tilde{u}_k = (u_k - \frac{i}{N})N$, puis le décomposer en un terme $g_{N,1}$ qui est une somme de Riemann et un terme de reste r_N ;

$$\begin{aligned} g_{N,\alpha} &= \frac{C_{1,\alpha}}{N} \sum_{i=0}^{N-1} \left(\frac{u_r}{N} + \frac{i}{N} - \xi_1 \right)_+^{\frac{H}{2}-1} \left(\frac{u_s}{N} + \frac{i}{N} - \xi_2 \right)_+^{\frac{H}{2}-1} \\ &= \frac{C_{1,\alpha}}{N} \sum_{i=0}^{N-1} \left(\frac{i}{N} - \xi_1 \right)_+^{\frac{H}{2}-1} \left(\frac{i}{N} - \xi_2 \right)_+^{\frac{H}{2}-1} \\ &+ \frac{C_{1,\alpha}}{N} \sum_{i=0}^{N-1} \left[\left(\frac{u_r}{N} + \frac{i}{N} - \xi_1 \right)_+^{\frac{H}{2}-1} \left(\frac{u_s}{N} + \frac{i}{N} - \xi_2 \right)_+^{\frac{H}{2}-1} - \left(\frac{i}{N} - \xi_1 \right)_+^{\frac{H}{2}-1} \left(\frac{i}{N} - \xi_2 \right)_+^{\frac{H}{2}-1} \right] \\ &= g_{N,1}(\xi_1, \xi_2) + r_N(\xi_1, \xi_2) \end{aligned}$$

où

$$C_{1,\alpha} = \int_{[0,1]^p} du_1 \dots du_p \left(\prod_{1 \leq a < b \leq p} |u_a - u_b|^{(H-1)\alpha_{ab}} \right). \quad (5.15)$$

Remarque 53. La constante $C_{1,\alpha}$ est bien définie, cela découle du lemme 13 de [58].

— $g_{N,1}$ est une somme de Riemann qui converge simplement vers son intégrale :

$$C_{1,\alpha} \int_0^1 dx (x - \xi_1)_+^{\frac{H}{2}-1} (x - \xi_2)_+^{\frac{H}{2}-1} = d(H)^{-1} C_{1,\alpha} L_1^H(\xi_1, \xi_2), \text{ pour } \xi_1, \xi_2 \in \mathbb{R}.$$

Comme il a été prouvé dans [48], [120] et [116], $(g_{N,1})_{N \geq 0}$ est une suite de Cauchy dans l'espace de Hilbert $L^2(\mathbb{R}^2)$, cela implique que la suite de variables aléatoires $I_2(g_{N,\alpha})$ est de Cauchy donc convergente dans l'espace $L^2(\Omega)$ complet. Nous en concluons que,

$$I_2(g_{N,\alpha}) \xrightarrow[N \rightarrow \infty]{L^2(\Omega)} d(H)^{-1} C_{1,\alpha} Z_1^H,$$

où Z^H est le processus de Rosenblatt, $d(H)$ est une constante donnée par (1.7) et $C_{1,\alpha}$ par (5.15).

- Pour le terme de reste r_N , nous parvenons à montrer qu'il converge vers 0 dans $L^2(\mathbb{R}^2)$.
- Pour résumer, la projection de $\tilde{V}_{1,N}$ sur le chaos d'ordre 2 que nous avons notée $\tilde{V}_N^{(2)}$ converge dans $L^2(\Omega)$ quand $N \rightarrow \infty$ vers

$$d(H)^p \beta \left(\frac{1}{2}, 1 \right)^{p-1} \sum_{\substack{\alpha \in \mathcal{A}_{p,2} \\ |\alpha|=p-1}} C_\alpha C_{1,\alpha} d(H)^{-1} Z_1^H = K(H, p) Z_1^H.$$

ANALYSE DES CHAOS D'ORDRE SUPÉRIEUR

Les termes d'ordre supérieur situés sur l'hyper-diagonale du tenseur convergent vers 0 dans $L^2(\Omega)$. En particulier, on a obtenu le résultat suivant,

Proposition 17. Soit $0 \leq k \leq p-2$ et notons

$$\tilde{V}_N^{2p-2k} = N^{-H} \sum_{k=0}^{p-2} \sum_{\substack{\alpha \in \mathcal{A}_{p,2} \\ |\alpha|=k}} C_\alpha V_{N,\alpha}^{(2p-2|\alpha|)},$$

dont l'expression de $V_{N,\alpha}^{(2p-2|\alpha|)}$ est donnée par (5.14). Alors,

$$\tilde{V}_N^{(2p-2k)} \xrightarrow[N \rightarrow \infty]{L^2(\Omega)} 0.$$

De plus,

$$\mathbb{E} \left| \tilde{V}_N^{(2p-2k)} \right|^2 \leq c(H, p) \begin{cases} N^{1-2H} & \text{si } H \in \left(\frac{1}{2}, \frac{3}{4}\right), \\ \log(N)N^{-\frac{1}{2}} & \text{si } H = \frac{3}{4}, \\ N^{2H-2} & \text{si } H \in \left(\frac{3}{4}, 1\right). \end{cases}$$

Finalement, ces résultats nous permettent de conclure sur le comportement asymptotique de l'hyper-diagonale $Y_{j,\dots,j}$ du tenseur aléatoire.

Proposition 18. *Pour $j = 1, \dots, n$, soit $\tilde{Y}_{j,\dots,j} := d^{-H}Y_{j,\dots,j}$ donné par (5.10). Alors*

$$\tilde{Y}_{j,\dots,j} \xrightarrow[d \rightarrow \infty]{} K(H, p)Z_1^{H,j}$$

et

$$\mathbb{E} \left| \tilde{Y}_{j,\dots,j} - K(H, p)Z_1^{H,j} \right|^2 \leq c(H, p) \begin{cases} d^{1-2H} & \text{if } H \in \left(\frac{1}{2}, \frac{3}{4}\right), \\ \log(d)d^{-\frac{1}{2}} & \text{if } H = \frac{3}{4}, \\ d^{2H-2} & \text{if } H \in \left(\frac{3}{4}, 1\right). \end{cases}$$

Il ne nous reste plus qu'à évaluer les termes en dehors de l'hyperdiagonale. Plus particulièrement, grâce à l'indépendance des processus de Rosenblatt, nous obtenons la convergence vers 0 dans $L^2(\Omega)$.

Évaluation des termes en dehors de l'hyper-diagonale

Dans cette situation, nous considérons que les indices j_1, \dots, j_p de Y_{j_1, \dots, j_p} donné par (5.6), ne sont pas tous égaux, i.e. $\text{Card}\{j_1, \dots, j_p\} \geq 2$. Nous avons ;

$$\begin{aligned} Y_{j_1, \dots, j_p} &= \sum_{i=0}^{d-1} \left[\left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} \left(Z_{i+1}^{H,2} - Z_i^{H,2} \right)^{p_2} \dots \left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} \right. \\ &\quad \left. - \mathbb{E} \left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} \left(Z_{i+1}^{H,2} - Z_i^{H,2} \right)^{p_2} \dots \left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} \right], \end{aligned}$$

où $p_1 + \dots + p_x = p$, $x \geq 2$ et $Z^{H,1}, \dots, Z^{H,x}$ sont des processus de Rosenblatt indépendants. Nous obtenons ainsi que,

$$\begin{aligned} Y_{j_1, \dots, j_p} &= \sum_{i=0}^{d-1} \left[\left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} - \mathbb{E} \left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} \right] \left[\left(Z_{i+1}^{H,2} - Z_i^{H,2} \right)^{p_2} - \mathbb{E} \left(Z_{i+1}^{H,2} - Z_i^{H,2} \right)^{p_2} \right] \\ &\quad \dots \left[\left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} - \mathbb{E} \left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} \right]. \end{aligned}$$

Pour établir la convergence des termes non diagonaux vers zéro dans $L^2(\Omega)$ quand $d \rightarrow \infty$, nous devons nous focaliser sur $\mathbb{E}[Y_{j_1, \dots, j_p}^2]$. Puisque les processus de Rosenblatt sont indépendants, nous avons,

$$\begin{aligned} &\mathbb{E}(Y_{j_1, \dots, j_p})^2 \\ &= \sum_{i,j=0}^{N-1} \mathbb{E} \left[\left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} - \mathbb{E} \left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} \right] \left[\left(Z_{j+1}^{H,1} - Z_j^{H,1} \right)^{p_1} - \mathbb{E} \left(Z_{j+1}^{H,1} - Z_j^{H,1} \right)^{p_1} \right] \\ &\quad \dots \mathbb{E} \left[\left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} - \mathbb{E} \left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} \right] \left[\left(Z_{j+1}^{H,x} - Z_j^{H,x} \right)^{p_x} - \mathbb{E} \left(Z_{j+1}^{H,x} - Z_j^{H,x} \right)^{p_x} \right] \end{aligned}$$

Finalement, nous appliquons la formule produit généralisée (5.12) pour $i, j = 0, \dots, d-1$, et pour $a = 1, \dots, x$,

$$\mathbb{E} \left[\left(Z_{i+1}^{H,a} - Z_i^{H,a} \right)^{Pa} - \mathbb{E} \left(Z_{i+1}^{H,a} - Z_i^{H,a} \right)^{Pa} \right] \left[\left(Z_{j+1}^{H,a} - Z_j^{H,a} \right)^{Pa} - \mathbb{E} \left(Z_{j+1}^{H,a} - Z_j^{H,a} \right)^{Pa} \right].$$

Nous obtenons que les termes en dehors de l'hyper-diagonale vérifient la borne suivante.

Proposition 19. *Pour tout $j_1, \dots, j_p \in \{1, \dots, n\}$ tels que $\text{Card}\{j_1, \dots, j_p\} \geq 2$, nous avons, pour d assez grand,*

$$\mathbb{E} \left| \tilde{Y}_{j_1, \dots, j_p} \right|^2 \leq c(H, p) \begin{cases} d^{1-2H} & \text{si } H \in \left(\frac{1}{2}, \frac{3}{4} \right) \\ \log(d) d^{-\frac{1}{2}} & \text{si } H = \frac{3}{4} \\ d^{2H-2} & \text{si } H \in \left(\frac{3}{4}, 1 \right). \end{cases}$$

En particulier,

$$\tilde{Y}_{j_1, \dots, j_p} \xrightarrow[d \rightarrow \infty]{L^2(\Omega)} \mathbf{0}.$$

Désormais, la dernière chose qu'il nous reste à évaluer est la distance de 1-Wasserstein vectorielle entre $\tilde{\mathbf{Y}} = (d^{-H} Y_{j_1, \dots, j_p}, j_1, \dots, j_p = 1, \dots, p)$ et sa limite \mathbf{R} donnés par (5.7) et (5.8) respectivement. Pour cela,

1. Nous utilisons l'invariance en loi de la distance de Wasserstein.
2. La distance de 1-Wasserstein est majorée par la distance dans $L^2(\Omega)$ à l'aide de la formule suivante.

De ce fait, nous pouvons dire que

$$d_W(\tilde{\mathbf{Y}}, \mathbf{R}) = d_W(\mathbf{Y}^{(1)}, \mathbf{R}) \leq \sqrt{\sum_{j_1, \dots, j_p=1}^n \mathbb{E} \left[\left(Y_{j_1, \dots, j_p}^{(1)} - R_{j_1, \dots, j_p} \right)^2 \right]}, \quad (5.16)$$

où $\mathbf{Y}^{(1)} = \left(Y_{j_1, \dots, j_p}^{(1)}, j_1, \dots, j_p = 1, \dots, n \right)$ et

$$Y_{j_1, \dots, j_p}^{(1)} = d^{-H} d^{Hp} \sum_{i=1}^d \left[\prod_{k=1}^p \left(Z_{\frac{i+1}{d}}^{H, j_k} - Z_{\frac{i}{d}}^{H, j_k} \right)^p - \mathbb{E} \prod_{k=1}^p \left(Z_{\frac{i+1}{d}}^{H, j_k} - Z_{\frac{i}{d}}^{H, j_k} \right)^p \right].$$

En utilisant la majoration de la distance de Wasserstein (5.16) combinée aux propositions 18 et 19 nous aboutissons au théorème suivant.

Théorème 40. *Considérons le vecteur aléatoire $\tilde{\mathbf{Y}}$ donné par (5.7). Alors $\tilde{\mathbf{Y}}$ converge dans $L^2(\Omega)$ composante par composante quand $d \rightarrow \infty$, vers le vecteur aléatoire \mathbf{R} introduit dans l'équation (5.8). De plus, pour n et d assez grands, nous disposons d'une majoration de l'erreur commise dans l'approximation de $\tilde{\mathbf{Y}}$ par un vecteur de Rosenblatt ;*

$$d_W(\tilde{\mathbf{Y}}, \mathbf{R}) \leq c(H, p) \begin{cases} n^{\frac{p}{2}} d^{\frac{1}{2}-H} & \text{si } H \in \left(\frac{1}{2}, \frac{3}{4} \right) \\ n^{\frac{p}{2}} \sqrt{\log(d)} d^{-\frac{1}{4}} & \text{si } H = \frac{3}{4} \\ n^{\frac{p}{2}} d^{H-1} & \text{si } H \in \left(\frac{3}{4}, 1 \right). \end{cases}$$

Partie III

**Inférence statistique pour les
équations différentielles
stochastiques : estimation de
paramètres**

Quadratic variation and drift parameter estimation for the stochastic wave equation with space-time white noise

This article is published in Stochastics and Dynamics.

Joint work with Obayda J. Assaad and Ciprian A. Tudor.

Abstract. We study the quadratic variations (in time and in space) of the solution to the stochastic wave equation driven by a space-time white noise. We give their limit (almost surely and in $L^2(\Omega)$) and we prove that these variations satisfy, after a proper renormalization, a Central Limit Theorem. We apply the quadratic variation to define and analyze estimators for the drift parameter of the wave equation.

6.1 Introduction

We deal with the following stochastic partial differential equation (SPDE in the sequel)

$$\frac{\partial^2 u}{\partial t^2}(t, x) = \theta \Delta u(t, x) + \sigma \dot{W}(t, x), \quad t > 0, x \in \mathbb{R}, \quad (6.1)$$

with vanishing initial value $u(0, x) = \frac{\partial u}{\partial t}(0, x) = 0$ for every $x \in \mathbb{R}$. In (6.1), W is a space-time white noise, while Δ denotes the standard Laplacian operator on \mathbb{R} . The equation (6.1) represents *the linear stochastic wave equation driven by the space-time white noise*. This stochastic partial differential equation constitutes a model for a vibrating string (of infinite length, under an ideal context, with uniform mass, neglecting the air resistance etc) perturbed by a random force which behaves as a Wiener process in time and in space. The value $u(t, x)$ constitutes a model for the vertical displacement from the x -axis of the string at time t and at position x (in a coordinate system with x on the horizontal line and u on the vertical line). We will discuss the estimation of the drift parameter $\theta > 0$ and of the diffusion parameter $\sigma > 0$ in (6.1). The drift parameter

is usually interpreted as the tension in the string while the diffusion parameter is related to the magnitude of the random perturbation.

The statistical inference for stochastic partial differential equations constitutes an intensive research direction. The reader may consult the recent surveys [31] or [83] for a description of the classical approach to parameter estimation in various SPDEs. One may also consult the website <https://sites.google.com/prod/view/stats4spdes/> for a more complete list of publications on these topics.

We aim at estimating the drift and diffusion parameters $\theta, \sigma > 0$ based on the observation of the solution u at discrete times and at a fixed point in space. That is, we assume that we have at our disposal the observations $(u(t_i, x), i = 0, 1, \dots, N)$ where $t_i, i = 0, 1, \dots, N$ denotes an equidistant partition of an arbitrary interval $[A_1, A_2]$ with $0 \leq A_1 < A_2$ given by $t_i = A_1 + \frac{i}{N}(A_2 - A_1)$ for $i = 0, 1, \dots, N$. Our estimator is constructed by using the quadratic variation in time of the solution u . This is a standard approach and several authors used it in order to estimate the parameters (drift, diffusion or even Hurst parameter) especially for the heat equation and related SPDES. We refer, among others, to, [14], [15], [33] [27], [28], [76], [104], [109], [119], [129]. A maximum likelihood approach has been used in [91]. On the other hand, in contrast to stochastic parabolic equations, the case of hyperbolic SPDEs was paid less attention to. Concerning the bibliography related to the statistical inference for the stochastic wave equation, we refer to [81], [82] for a maximum likelihood estimator, to [66] for a minimum contrast approach, to [10] for a wavelet-type approach and to [73] for the estimation of the Hurst parameter via p -variations when the noise is related to the fractional Brownian motion.

While the diffusion parameter can be estimated directly via the quadratic variations (in time or in space) of the solution to (6.1), the construction of the estimator for the drift parameter is based on some transform of the solution u which allows to replace the parameter θ in front of the random noise (the idea has also been used in e.g. [104] or [76]). This will allow to employ the power variation method in order to construct an estimator for θ in (6.1). We analyze the asymptotic behavior of the temporal quadratic variation sequence of the random field u which solves (6.1) in the mild sense. We prove that this sequence converges almost surely and in $L^2(\Omega)$ to a constant and it also satisfies a Central Limit Theorem (CLT in the sequel). The proof uses the techniques of the Stein-Malliavin calculus and this method also leads to some results concerning the rate of convergence in this CLT. We also tackle the limit behavior of the spatial quadratic variation of the solution. For the stochastic heat equation (see e.g. [104]), it is possible to estimate the drift parameter via the spatial variation by a similar approach. For the wave equation, we notice that the situation is different : we will see that the spatial variation does not depend asymptotically on the parameter θ . On the other hand, we also notice that the variations in space (as well as the variations in time) are useful to estimate the diffusion parameter.

We organize our paper as follows. Section 2 contains some basic facts related to the distributional and trajectorial properties of the solution to the stochastic wave equation with additive space-time white noise. In Section 3, we analyze the limit behavior of the temporal quadratic variation of this solution and we apply these results to the parameter estimation problem in Section 4. In Section 5, we consider the quadratic variation in space of the solution to (6.1) and we discuss how it can be applied to parameter estimation. Section 6 is the Appendix where we recall the basic elements concerning the Wiener chaos and the Malliavin derivative.

6.2 Preliminaries

In this section, we present some basic results concerning the probability distribution of the solution to the stochastic wave equation. We consider the non-parametrized situation, i.e. $\theta = \sigma = 1$ in (6.1). That is, we consider the following stochastic wave equation

$$\begin{cases} \frac{\partial^2 u}{\partial t^2}(t, x) = \Delta u(t, x) + \dot{W}(t, x), & t > 0, x \in \mathbb{R} \\ u(0, x) = 0, & x \in \mathbb{R} \\ \frac{\partial u}{\partial t}(0, x) = 0, & x \in \mathbb{R}. \end{cases} \quad (6.2)$$

Here, Δ is the Laplacian on \mathbb{R} and $W = (W(t, A); t \geq 0, A \in \mathcal{B}_b(\mathbb{R}))$ is a real valued centered Gaussian field, over a given complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ with covariance :

$$\mathbf{E}\left(W(t, A)W(s, B)\right) = (t \wedge s)\lambda(A \cap B), \text{ for every } A, B \in \mathcal{B}_b(\mathbb{R}), \quad (6.3)$$

where λ is the Lebesgue measure and $\mathcal{B}_b(\mathbb{R})$ is the set of the Borel-subsets of \mathbb{R} with finite Lebesgue measure. This is usually called "the space-time white noise".

Let G_1 be the fundamental solution of the homogeneous wave equation $\frac{\partial^2 u}{\partial t^2} - \Delta u = 0$. It is known that G_1 is given by, for $t > 0$ and $x \in \mathbb{R}$.

$$G_1(t, x) = \frac{1}{2} 1_{\{|x| < t\}}. \quad (6.4)$$

The mild solution to (6.2) is a square-integrable process $u = (u(t, x); t \geq 0, x \in \mathbb{R})$ which is defined by

$$u(t, x) = \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y) W(ds, dy). \quad (6.5)$$

The random field given by (6.5) is a centered Gaussian random field, due to the fact that the noise is additive and the initial data is non-random. It is well-known that the process u is well-defined in $L^2(\Omega)$, i.e.

$$\mathbf{E}(u(t, x)^2) < \infty, \text{ for every } t, x.$$

Indeed, the mild solution of (6.2) exists when the Wiener integral (6.5) with respect to the Gaussian process W is well-defined for every t and x , i.e. when the Green kernel (6.4) is such that $(s, y) \mapsto G_1(t-s, x-y) \in L^2(\mathbb{R}_+ \times \mathbb{R})$ for every fixed t, x .

We will give some properties of the mild solution (6.5) which be needed later. Let us start by computing the covariance in time of the solution.

Lemma 1. *Let $(u(t, x), t \geq 0, x \in \mathbb{R})$ be given by (6.5) and assume that $x \in \mathbb{R}$ is fixed. For every $s, t \geq 0$ we have*

$$\mathbf{E}u(t, x)u(s, x) = \frac{1}{4}(t \wedge s)^2. \quad (6.6)$$

Proof : Assume $0 \leq s \leq t$ and $x \in \mathbb{R}$. By the Itô's isometry of the Wiener integral and by using

the expression of the Green kernel G_1 , we have

$$\begin{aligned} \mathbf{E}u(t,x)u(s,x) &= \int_0^{s \wedge t} dr \int_{\mathbb{R}} dy G_1(t-r, x-y) G_1(s-r, x-y) \\ &= \frac{1}{4} \int_0^s dr \int_{\mathbb{R}} dy \mathbf{1}_{\{|x-y| \leq t-r\}} \mathbf{1}_{\{|x-y| \leq s-r\}} \\ &= \frac{1}{4} \int_0^s dr \int_{r-s}^{s-r} dy = \frac{1}{4} \int_0^s 2(s-r) dr = \frac{1}{4} s^2. \end{aligned}$$

□

Remark 1. 1. It follows from Lemma 1 that the Gaussian process $(u(t,x), t \geq 0)$ with $x \in \mathbb{R}$ fixed, is self-similar with self-similarity index equal to 1. Indeed, for every $c > 0$, the processes $(u(ct,x), t \geq 0)$ and $(cu(t,x), t \geq 0)$ have the same finite dimensional distributions.

2. It also follows from (6.6) that $(u(t,x), t \geq 0)$ has the same finite dimensional distributions as $(\frac{1}{2}W_t, t \geq 0)$ where $(W_t, t \geq 0)$ is a standard Wiener process.

Concerning the increment in time of the mild solution (6.5), we have the following.

Lemma 2. Let $(u(t,x), t \geq 0, x \in \mathbb{R})$ be given by (6.5) and let $x \in \mathbb{R}$ be fixed. Then we have

1. For every $0 \leq s \leq t$,

$$\mathbf{E}(u(t,x) - u(s,x))^2 = \frac{1}{4}(t^2 - s^2).$$

2. For every $0 \leq a < b \leq s < t$, we have

$$\mathbf{E}(u(t,x) - u(s,x))(u(b,x) - u(a,x)) = 0.$$

In particular, $u(t,x) - u(s,x)$ and $u(b,x) - u(a,x)$ are independent random variables.

Proof : The identities are immediate consequences of the covariance formula (6.6). The independence of the increments comes from the fact that the process $(u(t,x), t \geq 0)$ is Gaussian. □

6.3 Asymptotic behavior of the temporal quadratic variation and estimation of the drift parameter

In this part, we will study the limit behavior of the quadratic variation in time for the random field $(u(t,x), t \geq 0, x \in \mathbb{R})$ defined by (6.5). For this random field, we can define its *temporal quadratic variation* by fixing the space variable $x \in \mathbb{R}$ and by considering the quadratic variation of the centered Gaussian process $(u(t,x), t \geq 0)$. We can also define its *spatial quadratic variation* by assuming that the time $t \geq 0$ is fixed, see Section 6.5. For the stochastic heat equation, both types of variations can be used in order to estimate the parameters that may appear in the model (see e.g. [104] or [76]).

Concerning the temporal quadratic variation, we will give its limit in $L^2(\Omega)$ and almost sure and, after a suitable renormalization, we will prove that it satisfies a Central Limit Theorem.

6.3.1 Convergence in $L^2(\Omega)$ and almost sure

Let $0 \leq A_1 < A_2$ be two real numbers. Define, for every $x \in \mathbb{R}$,

$$V_{N,x}(u) = \sum_{i=0}^{N-1} (u(t_{i+1},x) - u(t_i,x))^2, \quad (6.7)$$

where

$$t_i = A_1 + \frac{i}{N}(A_2 - A_1), \quad i = 0, 1, \dots, N, \quad (6.8)$$

constitutes a partition of the interval $[A_1, A_2]$. The sequence $(V_{N,x}(u), N \geq 1)$ is called *the temporal quadratic variation of the random field u over the interval $[A_1, A_2]$* .

For every $t \in \mathbb{R}_+$ and $N \geq 1$, the random variable $V_{N,x}(u)$ is the sum of its expectation plus a variable in the second Wiener chaos. Indeed, since for every $i = 0, \dots, N-1$

$$u(t, x_{i+1}) - u(t, x_i) = I_1(g_{x,i}),$$

with, for $s > 0, y \in \mathbb{R}$,

$$g_{x,i}(s,y) = G_1(t_{i+1} - s, x - y)1_{(0,t_{i+1})}(s) - G_1(t_i - s, x - y)1_{(0,t_i)}(s), \quad (6.9)$$

we can write,

$$V_{N,x}(u) = I_2 \left(\sum_{i=0}^{N-1} g_{x,i}^{\otimes 2} \right) + \mathbf{E}V_{N,x}(u), \quad (6.10)$$

where I_2 is the multiple integral of order 2 with respect to the space-time white noise (see Section 6.6 for the presentation of the multiple stochastic integrals). Notice that for $i = 0, \dots, N-1, x \in \mathbb{R}$, by Lemma 2,

$$\begin{aligned} \langle g_{x,i}, g_{x,j} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})} &= \mathbf{E}((u(t_{i+1},x) - u(t_i,x))(u(t_{j+1},x) - u(t_j,x))) \\ &= \begin{cases} \frac{1}{4}(t_{i+1}^2 - t_i^2), & \text{for } i = j \\ 0 & \text{for } i \neq j. \end{cases} \end{aligned} \quad (6.11)$$

In the sequel we denote by C, C_p, \dots strictly positive constant that may change. They depend on the parameter indicated in their notation. By using Lemmas 1 and 2, we obtain the asymptotic behavior of the sequence (6.7).

Proposition 1. *Consider the random field u defined by (6.5). For every $x \in \mathbb{R}$, the sequence $(V_{N,x}(u), N \geq 1)$ converges, as $N \rightarrow \infty$, to $\frac{1}{4}(A_2^2 - A_1^2)$ in $L^2(\Omega)$ and almost surely.*

Proof : Let us start with the $L^2(\Omega)$ -convergence. First notice that

$$\mathbf{E}V_{N,x}(u) = \sum_{i=0}^{N-1} \mathbf{E}(u(t_{i+1},x) - u(t_i,x))^2 = \sum_{i=0}^{N-1} \frac{1}{4}(t_{i+1}^2 - t_i^2) = \frac{1}{4}(A_2^2 - A_1^2).$$

From Lemma 2, for every $i = 0, 1, \dots, N-1$, $u(t_{i+1}, x) - u(t_i, x) \sim \mathcal{N}\left(0, \frac{1}{4}(t_{i+1}^2 - t_i^2)\right)$, and so, if $Z \sim \mathcal{N}(0, 1)$,

$$\begin{aligned} \mathbf{E} \left| V_{N,x}(u) - \frac{1}{4}(A_2^2 - A_1^2) \right|^2 &= \text{Var}(V_{N,x}(u)) = \text{Var} \left(\sum_{i=0}^{N-1} (u(t_{i+1}, x) - u(t_i, x))^2 \right) \\ &= \sum_{i=0}^{N-1} \text{Var} \left((u(t_{i+1}, x) - u(t_i, x))^2 \right) = \text{Var}(Z^2) \frac{1}{16} \sum_{i=0}^{N-1} (t_{i+1}^2 - t_i^2)^2 \\ &\leq \frac{C}{N^2} \sum_{i=0}^{N-1} (t_{i+1} + t_i)^2 \leq C \frac{1}{N} \xrightarrow{N \rightarrow \infty} 0, \end{aligned} \quad (6.12)$$

where we used the independence of the temporal increment of the solution u (Lemma 2, point 2.). This gives the $L^2(\Omega)$ -convergence of $V_{N,x}(u)$ to $\frac{1}{4}(A_2^2 - A_1^2)$ as $N \rightarrow \infty$. To get the almost sure convergence, we use a Borel-Cantelli argument. For every $\gamma > 0$ and $p \geq 1$, we have by hypercontractivity (6.53) (which can be applied due to (6.10)) and (6.12)

$$\begin{aligned} P \left(\left| V_{N,x}(u) - \frac{1}{4}(A_2^2 - A_1^2) \right| \geq N^{-\gamma} \right) &\leq N^{\gamma p} \mathbf{E} \left| V_{N,x}(u) - \frac{1}{4}(A_2^2 - A_1^2) \right|^p \\ &\leq C_p N^{\gamma p} \left(\mathbf{E} \left| V_{N,x}(u) - \frac{1}{4}(A_2^2 - A_1^2) \right|^2 \right)^{\frac{p}{2}} \leq C_p N^{p(\gamma - \frac{1}{2})}, \end{aligned}$$

and by choosing p large enough and $\gamma \in (0, \frac{1}{2})$, we have

$$\sum_{N \geq 1} P \left(\left| V_{N,x}(u) - \frac{1}{4}(A_2^2 - A_1^2) \right| \geq N^{-\gamma} \right) \leq C_p \sum_{N \geq 1} N^{p(\gamma - \frac{1}{2})} < \infty.$$

The conclusion follows by Borel-Cantelli lemma. \square

Remark 2. 1. Proposition 1 and (6.53) imply the convergence in $L^p(\Omega)$, as $N \rightarrow \infty$, of $V_{N,x}(u)$ to its limit $\frac{1}{4}(A_2^2 - A_1^2)$, for every $p \geq 2$.

2. Notice that the stochastic heat equation with additive space-time white noise admits a nontrivial quartic variation in time (see e.g. [111], [104]), while for the stochastic wave equation we have a nontrivial quadratic variation in time (and in space).

6.3.2 Central Limit Theorem

Let u be given by (6.5). For $x \in \mathbb{R}$, we consider the sequence $(U_{N,x}(u), N \geq 1)$ given by

$$\begin{aligned} U_{N,x}(u) &= V_{N,x}(u) - \frac{1}{4}(A_2^2 - A_1^2) \\ &= \sum_{i=0}^{N-1} (u(t_{i+1}, x) - u(t_i, x))^2 - \frac{1}{4}(A_2^2 - A_1^2) \\ &= \sum_{i=0}^{N-1} \left[(u(t_{i+1}, x) - u(t_i, x))^2 - \frac{1}{4}(t_{i+1}^2 - t_i^2) \right]. \end{aligned} \quad (6.13)$$

We will show that the centered quadratic variation sequence $(U_{N,x}(u), N \geq 1)$ satisfies, after a proper renormalization, a CLT. By (6.10), we have for $N \geq 1, x \in \mathbb{R}$,

$$U_{N,x}(u) = I_2 \left(\sum_{i=0}^{N-1} g_{x,i}^{\otimes 2} \right), \quad (6.14)$$

with $g_{x,i}$ given by (6.9). A first result gives the behavior as N tends to infinity of $\mathbf{E}U_{N,x}(u)^2$.

Lemma 3. *Let $(U_{N,x}, N \geq 1)$ be given by (6.13). Then*

$$N\mathbf{E}U_{N,x}^2 \xrightarrow{N \rightarrow \infty} \frac{(A_2^3 - A_1^3)}{6} (A_2 - A_1) := K(A_1, A_2). \quad (6.15)$$

Proof : By the isometry of multiple integrals (6.50) and by applying relation (6.11)

$$\begin{aligned} \mathbf{E}U_{N,x}(u)^2 &= \mathbf{E} \left(\sum_{i=0}^{N-1} I_2 \left(\mathbf{1}_{[t_i, t_{i+1})}^{\otimes 2} \right) \right)^2 = 2 \sum_{i,j=0}^{N-1} \langle g_{x,i}, g_{x,j} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 \\ &= 2 \sum_{i,j=0}^{N-1} [\mathbf{E}(u(t_{i+1}, x) - u(t_i, x))u(t_{j+1}, x) - u(t_j, x))]^2 \\ &= 2 \sum_{i=0}^{N-1} (\mathbf{E}(u(t_{i+1}, x) - u(t_i, x)))^2, \end{aligned}$$

where we used the independence of the increments in time of the process u given in Lemma 2 to deduce that

$$\mathbf{E}(u(t_{i+1}, x) - u(t_i, x))(u(t_{j+1}, x) - u(t_j, x)) = 0 \text{ if } i \neq j.$$

Thus, since

$$t_{i+1}^2 - t_i^2 = \frac{2}{N}A_1(A_2 - A_1) + \frac{2i+1}{N^2}(A_2 - A_1)^2, \quad (6.16)$$

we get

$$\begin{aligned} \mathbf{E}U_{N,x}(u)^2 &= 2 \sum_{i=0}^{N-1} \left(\frac{1}{4}(t_{i+1}^2 - t_i^2) \right)^2 \\ &= \frac{1}{2} \frac{(A_2 - A_1)^2}{N^2} \sum_{i=0}^{N-1} \left(A_1 + \frac{2i+1}{2N}(A_2 - A_1) \right)^2, \end{aligned} \quad (6.17)$$

and thus for N large, $\mathbf{E}U_{N,x}(u)^2$ behaves as

$$\frac{(A_2 - A_1)^2}{2} \left(A_1^2 + A_1(A_2 - A_1) + \frac{1}{3}(A_2 - A_1)^2 \right) \frac{1}{N} = \frac{(A_2^3 - A_1^3)}{6} (A_2 - A_1) \frac{1}{N} = K(A_1, A_2) \frac{1}{N},$$

with $K(A_1, A_2)$ given by (6.15). □

Set

$$\tilde{U}_{N,x}(u) = \sqrt{N} \left[\sum_{i=0}^{N-1} (u(t_{i+1},x) - u(t_i,x))^2 - \frac{1}{4}(A_2^2 - A_1^2) \right]. \quad (6.18)$$

We show that the renormalized quadratic variation sequence $(\tilde{U}_{N,x}, N \geq 1)$ converges in law to a Gaussian limit, more precisely we will prove that

$$\tilde{U}_{N,x}(u) \xrightarrow[N \rightarrow \infty]{(d)} \mathcal{N}(0, K(A_1, A_2)),$$

where $\xrightarrow{(d)}$ denotes the convergence in distribution and $K(A_1, A_2)$ is defined in (6.15). We will use Theorem 4 in the Appendix to obtain the normal convergence of the sequence $\tilde{U}_{N,x}(u)$. Notice that $\tilde{U}_{N,x}(u)$ belongs to the second Wiener chaos for every $x \in \mathbb{R}$ and $N \geq 1$ i.e.

$$\tilde{U}_{N,x}(u) = I_2(f_N) \text{ with } f_N = \sqrt{N} \sum_{i=0}^{N-1} g_{x,i}^{\otimes 2}, \quad (6.19)$$

with $g_{x,i}$ given by (6.9). Hence, Theorem 4 can be applied to this sequence. We prove the following CLT for the sequence (6.18). Recall that d may be any of the following distances : Kolmogorov, total variation or Wasserstein (see the Appendix, formula (6.54) for their definitions).

Theorem 1. *The sequence $(\tilde{U}_{N,x}(u))_{N \geq 1}$ given by (6.18) converges in distribution, as $N \rightarrow \infty$, to $\mathcal{N}(0, K(A_1, A_2))$, with $K(A_1, A_2)$ given by (6.15), and for N large enough,*

$$d\left(\tilde{U}_{N,x}(u), \mathcal{N}(0, K(A_1, A_2))\right) \leq C \frac{1}{\sqrt{N}}.$$

Proof : By Lemma 3, with $K(A_1, A_2)$ defined in (6.15),

$$\mathbf{E} \tilde{U}_{N,x}(u)^2 \xrightarrow[N \rightarrow \infty]{} K(A_1, A_2).$$

In order to apply Theorem 4, we compute the Malliavin derivative of $\tilde{U}_{N,x}(u)$. We have from (6.19), with $g_{x,i}$ from (6.9),

$$D\tilde{U}_{N,x}(u) = 2\sqrt{N} \sum_{i=0}^{N-1} I_1(g_{x,i})g_{x,i},$$

and

$$\begin{aligned} \|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 &= 4N \sum_{i,j=0}^{N-1} I_1(g_{x,i})I_1(g_{x,j}) \langle g_{x,i}, g_{x,j} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})} \\ &= 4N \sum_{i=0}^{N-1} I_1^2(g_{x,i}) \|g_{x,i}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2, \end{aligned}$$

by using the fact that (see (6.11) and Lemma 2)

$$\langle g_{x,i}, g_{x,j} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})} = \mathbf{E}(u(t_{i+1}, x) - u(t_i, x))(u(t_{j+1}, x) - u(t_j, x)) = 0 \text{ for } i \neq j.$$

Therefore,

$$\begin{aligned} \|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 &= 4N \sum_{i=0}^{N-1} \left[I_2(g_{x,i}^{\otimes 2}) + \|g_{x,i}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 \right] \|g_{x,i}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 \\ &= A_N + \mathbf{E}\|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2, \end{aligned}$$

where

$$A_N = 4N \sum_{i=0}^{N-1} I_2(g_{x,i}^{\otimes 2}) \|g_{x,i}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2,$$

and

$$\mathbf{E}\|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 = 4N \sum_{i=0}^{N-1} \|g_{x,i}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^4 = 4N \sum_{i=0}^{N-1} (\mathbf{E}(u(t_{i+1}, x) - u(t_i, x))^2)^2.$$

Now we need to estimate

$$\mathbf{E}\left(\|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 - \mathbf{E}\|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2\right)^2 \text{ and } \mathbf{E}\|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 - K(A_1, A_2).$$

We can write, via (6.50) and (6.11),

$$\begin{aligned} &\mathbf{E}\left(\|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 - \mathbf{E}\|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2\right)^2 = \mathbf{E}A_N^2 \\ &= 16N^2 \sum_{i,j=0}^{N-1} \mathbf{E}I_2(g_{x,i}^{\otimes 2})I_2(g_{x,j}^{\otimes 2}) \|g_{x,i}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 \|g_{x,j}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 \\ &= 32N^2 \sum_{i,j=0}^{N-1} \langle g_{x,i}, g_{x,j} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 \|g_{x,i}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 \|g_{x,j}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 \\ &= 32N^2 \sum_{i=0}^{N-1} \|g_{x,i}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^8 = 32N^2 \sum_{i=0}^{N-1} (\mathbf{E}(u(t_{i+1}, x) - u(t_i, x))^2)^4 \\ &= 32N^2 \sum_{i=0}^{N-1} \left(\frac{1}{4}(t_{i+1}^2 - t_i^2)\right)^4, \end{aligned}$$

and by (6.16), for N large enough,

$$\begin{aligned} &\mathbf{E}\left(\|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 - \mathbf{E}\|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2\right)^2 \\ &= \frac{1}{8}N^2 \sum_{i=0}^{N-1} \left(\frac{2}{N}A_1(A_2 - A_1) + \frac{2i+1}{N^2}(A_2 - A_1)^2\right)^4 \leq C\frac{1}{N}. \end{aligned} \quad (6.20)$$

In order to estimate the second summand in the right-hand side of (6.55), we write,

$$\begin{aligned}
& \mathbf{E} \|D\tilde{U}_{N,x}(u)\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^2 - 2K(A_1, A_2) = 4N \sum_{i=0}^{N-1} \left(\frac{1}{4} (t_{i+1}^2 - t_i^2) \right)^2 - 2K(A_1, A_2) \\
& = \frac{1}{4} N \sum_{i=0}^{N-1} \left(\frac{2}{N} A_1 (A_2 - A_1) + \frac{2i+1}{N^2} (A_2 - A_1)^2 \right)^2 - 2K(A_1, A_2) \\
& \sim C \frac{1}{N},
\end{aligned} \tag{6.21}$$

where the last bound follows from (6.17) (we write $u_n \sim v_n$ to indicate that $\frac{u_n}{v_n} \rightarrow 1$ as $n \rightarrow \infty$). The conclusion is obtained by (6.20), (6.21) and Theorem 4. \square

Remark 3. If $A_1 = 0$ and $A_2 = 1$ then $K(A_1, A_2) = \frac{1}{6}$ and then the sequence (6.18) converges in distribution to $\mathcal{N}(0, \frac{1}{6})$.

6.3.3 Optimality of the rate of convergence

Let us analyze the optimality of the rate of convergence of the sequence $(\tilde{U}_{N,x}, N \geq 1)$ to the standard normal law under the total variation distance (denoted d_{TV} in the sequel). Define, for $N \geq 1$ and $x \in \mathbb{R}$,

$$\bar{U}_{N,x}(u) = \frac{U_{N,x}}{\sqrt{\mathbf{E}U_{N,x}(u)^2}}. \tag{6.22}$$

The estimate (6.56) in the Appendix can be applied to $(\bar{U}_{N,x}(u), N \geq 1)$ since this sequence belongs to the second Wiener chaos and $\mathbf{E}\bar{U}_{N,x}(u)^2 = 1$ for every $N \geq 1$. We have the following result.

Lemma 4. Let $x \in \mathbb{R}$ be fixed and consider the sequence $(\tilde{U}_{N,x}(u), N \geq 1)$ given by (6.18). Then

$$d_{TV}(\bar{U}_{N,x}(u), \mathcal{N}(0, 1)) \propto \frac{1}{\sqrt{N}}.$$

Proof : We need to evaluate $\mathbf{E}\bar{U}_{N,x}(u)^3$. Recall the expression (6.14) of $U_{N,x}(u)$ as a multiple integral of order two, i.e. $U_{N,x} = I_2(g_N)$ with $g_N = \sum_{i=0}^{N-1} g_{x,i}^{\otimes 2}$. We have by (6.50) and (6.51)

$$\mathbf{E}U_{N,x}(u)^3 = 8 \langle g_N, g_N \otimes_1 g_N \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})},$$

with the contraction defined by (6.52). Now, for $N \geq 1$,

$$g_N \otimes_1 g_N = \sum_{i,j=0}^{N-1} (g_{x,i} \otimes g_{x,j}) \langle g_{x,i}, g_{x,j} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})},$$

with $g_{x,i}$ from (6.9). Thus

$$\begin{aligned}
\mathbf{E}U_{N,x}(u)^3 & = 8 \sum_{i,j,k=0}^{N-1} \langle g_{x,i}, g_{x,j} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})} \langle g_{x,j}, g_{x,k} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})} \langle g_{x,k}, g_{x,i} \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})} \\
& = 8 \sum_{i=0}^{N-1} \|g_{x,i}\|_{L^2(\mathbb{R}_+ \times \mathbb{R})}^6 = 8 \sum_{i=0}^{N-1} \left(\frac{1}{4} (t_{i+1}^2 - t_i^2) \right)^3,
\end{aligned}$$

where we used (6.11). By (6.16) and the proof of Lemma 3,

$$\begin{aligned} \mathbf{E}\bar{U}_{N,x}(u)^3 &\sim 8N^{\frac{3}{2}} \sum_{i=0}^{N-1} \left(\frac{1}{4} \left(\frac{2}{N} A_1 (A_2 - A_1) + \frac{2i+1}{N^2} (A_2 - A_1)^2 \right) \right)^3 \\ &\sim C \frac{1}{\sqrt{N}} \end{aligned}$$

for N large enough, with $C > 0$. Since by Theorem 1 (see relation (6.21)),

$$\mathbf{Var} \left(\frac{1}{q} \|DF_N\|_{\mathcal{H}}^2 \right) \sim C \frac{1}{N},$$

the conclusion is obtained via (6.56). □

6.4 Estimation of the drift parameter

Let us now consider the parametrized linear stochastic wave equation

$$\frac{\partial^2 u_\theta}{\partial t^2} = \theta \Delta u_\theta(t, x) + \dot{W}(t, x), \quad t \geq 0, x \in \mathbb{R} \quad (6.23)$$

with vanishing initial condition

$$u_\theta(0, x) = \frac{\partial u_\theta}{\partial t}(0, x) = 0 \text{ for every } x \in \mathbb{R}.$$

We assume $\theta > 0$ and W is the space-time white noise defined by its covariance (6.3). The purpose is to estimate the parameter θ on the basis of the observations $(u(t_i, x), i = 0, \dots, N)$ at a fixed point in space $x \in \mathbb{R}$. We assume that the observation times $t_i, i = 0, 1, \dots, N$ are given by (6.8).

The Green kernel associated to the SPDE (6.23), i.e. the deterministic function which solves $\frac{\partial^2}{\partial t^2} v(t, x) = \theta \Delta v(t, x)$ is $G_1(\sqrt{\theta}t, x)$ with G_1 defined by (6.4). Therefore, the mild solution to the SPDE (6.23) can be written as

$$u_\theta(t, x) = \int_0^t \int_{\mathbb{R}} G_1(\sqrt{\theta}(t-s), x-y) W(ds, dy) \quad (6.24)$$

for every $t \geq 0, x \in \mathbb{R}$. The above Wiener integral with respect to the Gaussian noise W is well-defined.

We consider the following transform of the mild solution (6.24). Denote by

$$v_\theta(t, x) = u_\theta \left(\frac{t}{\sqrt{\theta}}, x \right), \quad (t, x) \in [0, \infty) \times \mathbb{R}. \quad (6.25)$$

We notice the following property for the random field v_θ .

Lemma 5. *The random field $(v_\theta(t, x), t \geq 0, x \in \mathbb{R})$ is a mild solution to the stochastic heat equation*

$$\frac{\partial}{\partial t} v_\theta(t, x) = \Delta v_\theta(t, x) + \theta^{-\frac{1}{4}} \tilde{W}(t, x),$$

with \tilde{W} a space-time white noise, i.e. for every $t \geq 0, x \in \mathbb{R}$,

$$v_\theta(t, x) = \theta^{-\frac{1}{4}} \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y) \tilde{W}(ds, dy). \quad (6.26)$$

Proof : We have, for $t \geq 0, x \in \mathbb{R}$,

$$\begin{aligned} v_\theta(t, x) &= \int_0^{\frac{t}{\sqrt{\theta}}} \int_{\mathbb{R}} G_1(t - \sqrt{\theta}s, x-y) W(ds, dy) \\ &= \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y) W\left(d\left(\frac{s}{\sqrt{\theta}}\right), dy\right) \\ &= \theta^{-\frac{1}{4}} \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y) \tilde{W}(ds, dy), \end{aligned}$$

where $(\tilde{W}(t, x), t \geq 0, x \in \mathbb{R})$ with $\tilde{W}(t, x) = \theta^{\frac{1}{4}} W\left(\frac{t}{\sqrt{\theta}}, x\right)$ is a space-time white noise. \square

Let us deduce the asymptotic behavior of the temporal quadratic variation of the random field u_θ .

Proposition 2. *Let $(t_i, i = 0, 1, \dots, N)$ be the partition of $[A_1, A_2]$ given by (6.8) and let u_θ be defined by (6.24). Denote, for $N \geq 1, x \in \mathbb{R}$,*

$$V_{N,x}(u_\theta) = \sum_{i=0}^{N-1} (u_\theta(t_{i+1}, x) - u_\theta(t_i, x))^2. \quad (6.27)$$

Then

$$V_{N,x}(u_\theta) \xrightarrow{N \rightarrow \infty} \frac{1}{4} (A_2^2 - A_1^2) \theta^{\frac{1}{2}} \text{ almost surely and in } L^2(\Omega).$$

Proof : Notice that, by (6.26),

$$\left(\theta^{\frac{1}{4}} v_\theta(t, x), t \geq 0, x \in \mathbb{R} \right) \stackrel{(d)}{\equiv} (u_1(t, x), t \geq 0, x \in \mathbb{R}), \quad (6.28)$$

where $\stackrel{(d)}{\equiv}$ means the equivalence of finite dimensional distributions (while $\stackrel{(d)}{=}$ stands for the equality in law). Then, by the definition of v_θ (see (6.25)) and (6.28)

$$\begin{aligned} V_{N,x}(u_\theta) &= \sum_{i=0}^{N-1} \left(v_\theta(\sqrt{\theta}t_{i+1}, x) - v_\theta(\sqrt{\theta}t_i, x) \right)^2 \\ &= \theta^{-\frac{1}{2}} \sum_{i=0}^{N-1} \left(\theta^{\frac{1}{4}} v_\theta(\sqrt{\theta}t_{i+1}, x) - \theta^{\frac{1}{4}} v_\theta(\sqrt{\theta}t_i, x) \right)^2 \\ &= \theta^{-\frac{1}{2}} \sum_{i=0}^{N-1} \left(\tilde{u}_1(\sqrt{\theta}t_{i+1}, x) - \tilde{u}_1(\sqrt{\theta}t_i, x) \right)^2, \end{aligned}$$

with \tilde{W} is the space-time white noise from (6.26))

$$\tilde{u}_1(t, x) = \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y) \tilde{W}(ds, dy). \quad (6.29)$$

Since $(\sqrt{\theta}t_i, i = 0, \dots, N)$ represents a partition of the interval $[\sqrt{\theta}A_1, \sqrt{\theta}A_2]$, we have by Proposition 1 that

$$\sum_{i=0}^{N-1} \left(\tilde{u}_1(\sqrt{\theta}t_{i+1}, x) - \tilde{u}_1(\sqrt{\theta}t_i, x) \right)^2 \xrightarrow{N \rightarrow \infty} \frac{1}{4} \left((\sqrt{\theta}A_2)^2 - (\sqrt{\theta}A_1)^2 \right) = \frac{1}{4} \theta (A_2^2 - A_1^2).$$

This implies the conclusion. \square

Remark 4. We can also prove that for $N \geq 1, x \in \mathbb{R}$,

$$\mathbf{E} \left| V_{N,x}(u_\theta) - \frac{1}{4} (A_2^2 - A_1^2) \theta^{\frac{1}{2}} \right|^2 \leq C \frac{1}{N}. \quad (6.30)$$

This is an easy consequence of (6.12).

Via Proposition 2, we define the following estimator for the parameter $\theta > 0$ in (6.23)

$$\hat{\theta}_N := \left(\frac{4V_{N,x}(u_\theta)}{A_2^2 - A_1^2} \right)^2. \quad (6.31)$$

Proposition 3. The estimator $\hat{\theta}_N$ is strongly consistent, that is, $\hat{\theta}_N \xrightarrow{N \rightarrow \infty} \theta$ almost surely.

Proof : It follows directly from Proposition 2. \square

We have the following Central Limit Theorem for the estimator (6.31). Recall that " $\xrightarrow{(d)}$ " stands for the convergence in distribution.

Theorem 2. Let $0 \leq A_1 < A_2$ and let $K(A_1, A_2)$ be given by (6.15). Consider the estimator $\hat{\theta}_N$ given by (6.31). Then

$$\sqrt{N} \left(\hat{\theta}_N^{\frac{1}{2}} - \theta^{\frac{1}{2}} \right) \xrightarrow[N \rightarrow \infty]{(d)} \mathcal{N}(0, H(\theta, A_1, A_2)), \quad (6.32)$$

where $H(\theta, A_1, A_2) = \left(\frac{4}{A_2^2 - A_1^2} \theta^{-\frac{1}{2}} \right)^2 K(\sqrt{\theta}A_1, \sqrt{\theta}A_2)$. Also, for N large enough,

$$d \left(\sqrt{N} \left(\hat{\theta}_N^{\frac{1}{2}} - \theta^{\frac{1}{2}} \right), \mathcal{N}(0, H(\theta, A_1, A_2)) \right) \leq C \frac{1}{\sqrt{N}}. \quad (6.33)$$

Proof : We write, for every $N \geq 1$, with \tilde{u}_1 from (6.29),

$$\begin{aligned} & \hat{\theta}_N^{\frac{1}{2}} - \theta^{\frac{1}{2}} \\ &= \frac{4}{A_2^2 - A_1^2} \sum_{i=0}^{N-1} (u_\theta(t_{i+1}, x) - u_\theta(t_i, x))^2 - \theta^{\frac{1}{2}} \\ &= \frac{4}{A_2^2 - A_1^2} \sum_{i=0}^{N-1} \left(\tilde{u}_1(\sqrt{\theta}t_{i+1}, x) - \tilde{u}_1(\sqrt{\theta}t_i, x) \right)^2 - \theta^{\frac{1}{2}} \\ &= \frac{4}{A_2^2 - A_1^2} \theta^{-\frac{1}{2}} \left[\sum_{i=0}^{N-1} \left(\tilde{u}_1(\sqrt{\theta}t_{i+1}, x) - \tilde{u}_1(\sqrt{\theta}t_i, x) \right)^2 - \frac{1}{4} (A_2^2 - A_1^2) \theta \right]. \end{aligned}$$

Now by Theorem 1, since $(\sqrt{\theta}t_i, i = 0, 1, \dots, N)$ forms a partition of the interval $[\sqrt{\theta}A_1, \sqrt{\theta}A_2]$, we get

$$\sqrt{N} \left[\sum_{i=0}^{N-1} \left(\tilde{u}_1(\sqrt{\theta}t_{i+1}, x) - \tilde{u}_1(\sqrt{\theta}t_i, x) \right)^2 - \frac{1}{4}(A_2^2 - A_1^2)\theta \right] \xrightarrow[N \rightarrow \infty]{(d)} \mathcal{N}(0, K(\sqrt{\theta}A_1, \sqrt{\theta}A_2)),$$

and this gives the conclusion. \square

By applying the delta-method, we can obtain a result similar to the limit theorem (6.33) with $\widehat{\theta}_N$ instead of $\widehat{\theta}_N^{\frac{1}{2}}$, by considering the particular case of the Wasserstein distance, denoted d_W in the sequel. We recall that the Wasserstein distance is defined by (6.54) in the Appendix.

Corollary 1. *Let $\widehat{\theta}_N$ be given by (6.31). Then*

$$\sqrt{N}(\widehat{\theta}_N - \theta) \rightarrow \mathcal{N}(0, L(\theta, A_1, A_2)), \quad (6.34)$$

with $L(\theta, A_1, A_2) = 4\theta H(\theta, A_1, A_2)$ with H given by (6.32). Moreover, for N large enough,

$$d\left(\sqrt{N}(\widehat{\theta}_N - \theta), \mathcal{N}(0, L(\theta, A_1, A_2))\right) \leq C \frac{1}{\sqrt{N}}.$$

Proof : Consider the function $g : \mathbb{R} \rightarrow \mathbb{R}$, $g(x) = x^2$. We can write, for every $N \geq 1$,

$$\sqrt{N}(\widehat{\theta}_N - \theta) = \sqrt{N}\left(g(\widehat{\theta}_N^{\frac{1}{2}}) - g(\theta^{\frac{1}{2}})\right) = g'(a_N)\sqrt{N}\left(\widehat{\theta}_N^{\frac{1}{2}} - \theta^{\frac{1}{2}}\right),$$

with a_N a random point located between $\theta^{\frac{1}{2}}$ and $\widehat{\theta}_N^{\frac{1}{2}}$. Therefore,

$$\sqrt{N}(\widehat{\theta}_N - \theta) = g'(\theta^{\frac{1}{2}})\sqrt{N}\left(\widehat{\theta}_N^{\frac{1}{2}} - \theta^{\frac{1}{2}}\right) + R_N,$$

where $R_N = \sqrt{N}\left(\widehat{\theta}_N^{\frac{1}{2}} - \theta^{\frac{1}{2}}\right)\left(g'(a_N) - g'(\theta^{\frac{1}{2}})\right)$. Consequently, by the triangle inequality and by the definition of the Wasserstein distance, since $\left(g'(\theta^{\frac{1}{2}})\right)^2 H(\theta, A_1, A_2) = L(\theta, A_1, A_2)$,

$$\begin{aligned} & d_W\left(\sqrt{N}(\widehat{\theta}_N - \theta), \mathcal{N}(0, L(\theta, A_1, A_2))\right) \\ & \leq d_W\left(g'(\theta^{\frac{1}{2}})\sqrt{N}\left(\widehat{\theta}_N^{\frac{1}{2}} - \theta^{\frac{1}{2}}\right), g'(\theta^{\frac{1}{2}})\mathcal{N}(0, H(\theta, A_1, A_2))\right) + \mathbf{E}|R_N|. \end{aligned} \quad (6.35)$$

By Theorem 2, the first summand is bounded by $C \frac{1}{\sqrt{N}}$. For the second summand in the right-hand side of (6.35), we have

$$\mathbf{E}|R_N| \leq \left(\mathbf{E}\left[\sqrt{N}\left(\widehat{\theta}_N^{\frac{1}{2}} - \theta^{\frac{1}{2}}\right)\right]^2\right)^{\frac{1}{2}} \left(\mathbf{E}\left(g'(a_N) - g'(\theta^{\frac{1}{2}})\right)^2\right)^{\frac{1}{2}}.$$

Now $\mathbf{E}\left(\sqrt{N}\left(\widehat{\theta}_N^{\frac{1}{2}} - \theta^{\frac{1}{2}}\right)\right)^2$ is bounded by a constant by (6.30) and

$\mathbf{E}\left(g'(a_N) - g'(\theta^{\frac{1}{2}})\right)^2 \leq CE\left(\widehat{\theta}_N^{\frac{1}{2}} - \theta^{\frac{1}{2}}\right)^2 \leq C \frac{1}{N}$ again by (6.30). The conclusion follows. \square

6.5 Spatial quadratic variation

Now we fix $t \geq 0$ and we regard the process $(u(t, x), x \in \mathbb{R})$ defined by (6.5). We study its quadratic variation and we discuss how it can be applied to parameter estimation.

6.5.1 Limit behavior of the spatial quadratic variation

Let us recall (see e.g. [74]) that $(u(t, x), x \in \mathbb{R})$ is a centered Gaussian process with covariance

$$\mathbf{E}u(t, x)u(t, y) = \frac{1}{4} \left(\frac{|x - y|}{2} - t \right)^2 1_{\{|x - y| \leq 2t\}} \text{ for every } t \geq 0, x, y \in \mathbb{R}. \quad (6.36)$$

Fix two real numbers $A_1 < A_2$ and consider the following partition of the interval $[A_1, A_2]$

$$x_i = A_1 + \frac{i}{N}(A_2 - A_1), \quad i = 0, 1, \dots, N. \quad (6.37)$$

Let u be given by (6.5). We define the *spatial quadratic variation* sequence of the random field u by

$$S_{N,t}(u) = \sum_{i=0}^{N-1} [u(t, x_{i+1}) - u(t, x_i)]^2, \quad N \geq 1. \quad (6.38)$$

A CLT for the sequence (6.38) (under a different renormalization than that we consider below) has been obtained in [74]. Although the asymptotic behavior of $S_{N,t}$, as $N \rightarrow \infty$, can be deduced from the proofs in [74], this was not clearly stated in that reference. Let us start by giving the limit (almost surely and in $L^2(\Omega)$) of the sequence (6.38).

As for the temporal quadratic variation, we can express the random variable $S_{N,t}(u)$ as a multiple integral in the second Wiener chaos. Indeed, by writing

$$u(t, x_{i+1}) - u(t, x_i) = I_1(h_{i,t}),$$

with (recall that G_1 was defined by (6.4))

$$h_{i,t}(s, y) = 1_{[0,t]}(s) (G_1(t - s, x_{i+1} - y) - G_1(t - s, x_i - y)), \quad i = 0, 1, \dots, N-1 \text{ and } t \geq 0, \quad (6.39)$$

we get via the product formula (6.51)

$$S_{N,t}(u) = I_2 \left(\sum_{i=0}^{N-1} h_{i,t}^{\otimes 2} \right) + \mathbf{E}S_{N,t}. \quad (6.40)$$

We will analyze the limit behavior of the sequence $(S_{N,t}(u), N \geq 1)$ as $N \rightarrow \infty$. We first recall some useful results concerning the spatial increment of the solution to (6.2). We refer to [122] for their proofs.

Lemma 6. *Let $(u(t, x), t \geq 0, x \in \mathbb{R})$ be the mild solution to (6.2) and fix $t > 0$. If $(x_i, i = 0, \dots, N)$ are given by (6.37), then*

1. *If $t > \frac{1}{2}$, for every $i = 0, 1, \dots, N-1$,*

$$\mathbf{E}(u(t, x_{i+1}) - u(t, x_i))^2 = \frac{t}{2N}(A_2 - A_1) - \frac{1}{8N^2}(A_2 - A_1)^2. \quad (6.41)$$

2. If $t > \frac{1}{2}$, for every $i, j = 0, 1, \dots, N-1$ with $i \neq j$,

$$\mathbf{E} \left[(u(t, x_{i+1}) - u(t, x_i)) (u(t, x_{j+1}) - u(t, x_j)) \right] = -\frac{1}{8N^2}. \quad (6.42)$$

By using the above lemma, we obtain the following limit theorem for the sequence (6.38).

Proposition 4. *Let $(S_{N,t}(u), N \geq 1)$ be given by (6.38) with $t > \frac{1}{2}$ fixed. Then*

$$S_{N,t}(u) \xrightarrow[N \rightarrow \infty]{\text{a.s.}} \frac{t}{2}(A_2 - A_1) \text{ almost surely and in } L^2(\Omega).$$

Proof : By (6.41), we see that

$$\mathbf{E}S_{N,t}(u) = \frac{t}{2}(A_2 - A_1) - \frac{1}{8N}(A_2 - A_1)^2,$$

and consequently

$$\mathbf{E}S_{N,t}(u) - \frac{t}{2}(A_2 - A_1) = -\frac{1}{8N}(A_2 - A_1)^2 \xrightarrow[N \rightarrow \infty]{} 0. \quad (6.43)$$

Now we write

$$S_{N,t} - \frac{t}{2}(A_2 - A_1) = (S_{N,t} - \mathbf{E}S_{N,t}) + \left(\mathbf{E}S_{N,t}(u) - \frac{t}{2}(A_2 - A_1) \right),$$

and then by (6.43) and the orthogonality of Wiener chaoses of different orders (see (6.50)),

$$\begin{aligned} \mathbf{E} \left(S_{N,t} - \frac{t}{2}(A_2 - A_1) \right)^2 &= \mathbf{E} (S_{N,t} - \mathbf{E}S_{N,t})^2 + \mathbf{E} \left(\mathbf{E}S_{N,t}(u) - \frac{t}{2}(A_2 - A_1) \right)^2 \\ &= \mathbf{E} (S_{N,t} - \mathbf{E}S_{N,t})^2 + \frac{1}{64N^2}(A_2 - A_1)^2. \end{aligned}$$

By (6.40) and Lemma 6, with $h_{i,t}$ given by (6.39),

$$\begin{aligned} \mathbf{E} (S_{N,t} - \mathbf{E}S_{N,t})^2 &= 2 \sum_{i,j=0}^{N-1} \langle h_{i,t}, h_{j,t} \rangle^2 \\ &= 2 \sum_{i=0}^{N-1} \|h_{i,t}\|^4 + 2 \sum_{i,j=0, i \neq j}^{N-1} \langle h_{i,t}, h_{j,t} \rangle^2 \\ &= 2 \sum_{i=0}^{N-1} \left(\frac{t}{2N}(A_2 - A_1) - \frac{1}{8N^2}(A_2 - A_1)^2 \right)^2 + 2 \sum_{i,j=0, i \neq j}^{N-1} \left(-\frac{1}{8N^2} \right)^2 \\ &= \frac{t^2}{2N}(A_2 - A_1)^2 + o\left(\frac{1}{N}\right), \end{aligned}$$

and thus,

$$\mathbf{E} \left(S_{N,t} - \frac{t}{2}(A_2 - A_1) \right)^2 = \frac{t^2}{2N}(A_2 - A_1)^2 + o\left(\frac{1}{N}\right) \xrightarrow[N \rightarrow \infty]{} 0. \quad (6.44)$$

The almost sure convergence follows as in the proof of Proposition 1, via Borel-Cantelli and hypercontractivity. Indeed, for $0 < \gamma < \frac{1}{2}$, $p \geq 2$ and for N large enough,

$$\begin{aligned} P \left(\left| S_{N,t} - \frac{t}{2}(A_2 - A_1) \right| \geq N^{-\gamma} \right) &\leq N^{\gamma p} \mathbf{E} \left| S_{N,t} - \frac{t}{2}(A_2 - A_1) \right|^p \\ &\leq C_p N^{\gamma p} \left[\mathbf{E} \left| S_{N,t} - \frac{t}{2}(A_2 - A_1) \right|^2 \right]^{\frac{p}{2}} \leq C_p N^{(\gamma - \frac{1}{2})p}, \end{aligned}$$

where we used (6.44) for the last inequality. Thus, since $\gamma < \frac{1}{2}$,

$$\sum_{N \geq 1} P \left(\left| S_{N,t} - \frac{t}{2}(A_2 - A_1) \right| \geq N^{-\gamma} \right) \leq C_p \sum_{N \geq 1} N^{(\gamma - \frac{1}{2})p} < \infty.$$

□

Let us set for $N \geq 1$ and $t \geq 0$,

$$\tilde{T}_{N,t}(u) = \sqrt{N} \left(S_{N,t}(u) - \frac{t}{2}(A_2 - A_1) \right). \quad (6.45)$$

The following CLT has been obtained in [74] (under a slightly different renormalization).

Theorem 3. *Let $(\tilde{T}_{N,t}(u), N \geq 1)$ be given by (6.45) and let $t > \frac{1}{2}$. Then*

$$\tilde{T}_{N,t}(u) \xrightarrow[N \rightarrow \infty]{(d)} \mathcal{N} \left(0, \frac{t^2}{2}(A_2 - A_1)^2 \right),$$

and for N large enough,

$$d \left(\tilde{T}_{N,t}(u), \mathcal{N} \left(0, \frac{t^2}{2}(A_2 - A_1)^2 \right) \right) \leq C \frac{1}{\sqrt{N}}.$$

Remark 5. *It is also possible to prove as in Lemma 4 that the above estimate is sharp under the total variation distance, that is,*

$$d_{TV} \left(\tilde{T}_{N,t}(u), \mathcal{N} \left(0, \frac{t^2}{2}(A_2 - A_1)^2 \right) \right) \propto C \frac{1}{\sqrt{N}},$$

where the symbol \propto is that used in (6.56).

6.5.2 On parameter estimation via quadratic variation

Let us now consider the solution to the parametrized wave equation $(u_\theta(t, x), t \geq 0, x \in \mathbb{R})$ given by (6.24). By Lemma 5, its covariance can be computed as follows, for $t > 0, x, y \in \mathbb{R}$,

$$\begin{aligned} \mathbf{E}u_\theta(t, x)u_\theta(t, y) &= \mathbf{E}v_\theta(\sqrt{\theta}t, x)v_\theta(\sqrt{\theta}t, y) \\ &= \theta^{-\frac{1}{2}} \mathbf{E}u_1(\sqrt{\theta}t, x)u_1(\sqrt{\theta}t, y) \end{aligned}$$

so by (6.36),

$$\mathbf{E}u_\theta(t, x)u_\theta(t, y) = \frac{1}{4\sqrt{\theta}} \left(\frac{|x-y|}{2} - t\sqrt{\theta} \right)^2 1_{\{|x-y| \leq 2t\sqrt{\theta}\}} \text{ for every } t \geq 0, x, y \in \mathbb{R}. \quad (6.46)$$

Thus, if $2t\sqrt{\theta} > 1$, we find

$$\mathbf{E}(u_\theta(t, x_{i+1}) - u_\theta(t, x_i))^2 = \frac{t}{2N}(A_2 - A_1) - \frac{1}{8\sqrt{\theta}N^2}(A_2 - A_1)^2.$$

Therefore,

$$\mathbf{E}S_{N,t}(u_\theta) = \frac{t}{2}(A_2 - A_1) - \frac{1}{8\sqrt{\theta}N}(A_2 - A_1)^2,$$

and we can prove by following the lines of Proposition 4 that

$$S_{N,t}(u_\theta) \xrightarrow[N \rightarrow \infty]{} \frac{t}{2}(A_2 - A_1) \text{ almost surely and in } L^2(\Omega). \quad (6.47)$$

From (6.47), we notice that the spatial quadratic variation $S_{N,t}(u_\theta)$ does not depend asymptotically on θ and it cannot be used to estimate the drift parameter in (6.23). On the other hand, it can still be used to estimate the diffusion parameter for the wave equation. Indeed, let us consider the SPDE

$$\frac{\partial^2 u_\sigma}{\partial t^2} = \Delta u_\sigma(t, x) + \sigma \dot{W}(t, x), \quad t \geq 0, x \in \mathbb{R} \quad (6.48)$$

with vanishing initial condition

$$u_\sigma(0, x) = \frac{\partial u_\sigma}{\partial t}(0, x) = 0 \text{ for every } x \in \mathbb{R}.$$

We assume again $\sigma > 0$ and W is the space-time white noise whose covariance is given by (6.3). The mild solution to (6.48) is given, for $t \geq 0, x \in \mathbb{R}$, by

$$u_\sigma(t, x) = \sigma \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y) W(ds, dy) = \sigma u(t, x),$$

with G_1 from (6.4) and u from (6.5). It is well-defined since we consider the spatial dimension $d = 1$. Moreover,

$$S_{N,t}(u_\sigma) = \sum_{i=0}^{N-1} [u_\sigma(t, x_{i+1}) - u_\sigma(t, x_i)]^2 = \sigma^2 S_{N,t}(u),$$

with $S_{N,t}(u)$ given by (6.38). Therefore, by Proposition 4,

$$S_{N,t}(u_\sigma) \xrightarrow[N \rightarrow \infty]{} \sigma^2 \frac{t}{2}(A_2 - A_1) \text{ almost surely and in } L^2(\Omega).$$

Consequently, as $N \rightarrow \infty$,

$$\widehat{\sigma}_N^2 := \frac{2S_{N,t}(u_\sigma)}{t(A_2 - A_1)} \rightarrow \sigma^2 \text{ almost surely,}$$

so $\widehat{\sigma}_N$ constitutes a consistent estimator for the diffusion parameter σ in (6.48). A CLT for this estimator can be deduced via Theorem 3.

Let us end this section by two comments.

Remark 6. — *Obviously, the temporal quadratic variation can be also used in order to estimate the parameter σ in (6.48). It follows from Proposition 1 that, for every $x \in \mathbb{R}$, the sequence*

$$\sum_{i=0}^{N-1} (u_\sigma(t_{i+1}, x) - u_\sigma(t_i, x))^2 \xrightarrow[N \rightarrow \infty]{} \sigma^2 \frac{1}{4}(A_2^2 - A_1^2) \text{ almost surely and in } L^2(\Omega),$$

and hence $(\frac{1}{4}(A_2^2 - A_1^2))^{-1} \sum_{i=0}^{N-1} (u_\sigma(t_{i+1}, x) - u_\sigma(t_i, x))^2$ acts as a consistent estimator for σ^2 . The asymptotic normality of the estimator can be also obtained by using the results in Section 3. It would be also interesting to explore the possibility to explore the identification of both parameters θ and σ in (6.1).

- An other interesting problem is the identification of the parameters of the stochastic wave equations when the solution is observed on a space-time grid. The difficulty is related to the sharp estimations of the space-time correlation structure of the increments of the solution.

6.6 Appendix : Multiple stochastic integrals and the Malliavin derivative

The basic tools from the analysis on Wiener space are presented in this section. We will focus on some elementary facts about multiple stochastic integrals. We refer to [100] or [93] for a complete review on the topic.

Consider \mathcal{H} a real separable infinite-dimensional Hilbert space with its associated inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, and $(B(\varphi), \varphi \in \mathcal{H})$ an isonormal Gaussian process on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, which is a centered Gaussian family of random variables such that $\mathbf{E}(B(\varphi)B(\psi)) = \langle \varphi, \psi \rangle_{\mathcal{H}}$ for every $\varphi, \psi \in \mathcal{H}$. Denote by I_q the q th multiple stochastic integral with respect to B , which is an isometry between the Hilbert space $\mathcal{H}^{\odot q}$ (symmetric tensor product) equipped with the scaled norm $\frac{1}{\sqrt{q!}} \|\cdot\|_{\mathcal{H}^{\otimes q}}$ and the Wiener chaos of order q , which is defined as the closed linear span of the random variables $H_q(B(\varphi))$ where $\varphi \in \mathcal{H}$, $\|\varphi\|_{\mathcal{H}} = 1$ and H_q is the Hermite polynomial of degree $q \geq 1$ defined by :

$$H_q(x) = (-1)^q \exp\left(\frac{x^2}{2}\right) \frac{d^q}{dx^q} \left(\exp\left(-\frac{x^2}{2}\right) \right), \quad x \in \mathbb{R}. \quad (6.49)$$

The isometry of multiple integrals can be written as follows : for $p, q \geq 1$, $f \in \mathcal{H}^{\otimes p}$ and $g \in \mathcal{H}^{\otimes q}$

$$\mathbf{E}\left(I_p(f)I_q(g)\right) = \begin{cases} q! \langle \tilde{f}, \tilde{g} \rangle_{\mathcal{H}^{\otimes q}} & \text{if } p = q, \\ 0 & \text{otherwise,} \end{cases} \quad (6.50)$$

where \tilde{f} stands for the symmetrization of f . We have the following product formula : if $f \in \mathcal{H}^{\otimes p}$ and $g \in \mathcal{H}^{\otimes q}$, then

$$I_p(f)I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}(f \otimes_r g), \quad (6.51)$$

where for $r = 0, \dots, p \wedge q$, if $\mathcal{H} = L^2(T)$, the contraction $f \otimes_r g$ is the function in $L^2(T^{p+q-2r})$ given by

$$(f \otimes_r g)(t_1, \dots, t_{p+q-2r}) = \int_{T^r} f(u_1, \dots, u_r, t_1, \dots, t_{p-r}) g(u_1, \dots, u_r, t_{p-r+1}, \dots, t_{p+q-2r}) du_1 \dots du_r. \quad (6.52)$$

An useful property of finite sums of multiple integrals is the hypercontractivity. Namely, if $F = \sum_{k=0}^n I_k(f_k)$ with $f_k \in \mathcal{H}^{\otimes k}$ then

$$\mathbf{E}|F|^p \leq C_p (\mathbf{E}F^2)^{\frac{p}{2}}. \quad (6.53)$$

for every $p \geq 2$. In this work, the role of the Hilbert space \mathcal{H} will be played by $L^2((\mathbb{R}_+ \times \mathbb{R}^d)^q)$.

We denote by D the Malliavin derivative operator that acts on cylindrical random variables of the form $F = g(B(\varphi_1), \dots, B(\varphi_n))$, where $n \geq 1$, $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is a smooth function with compact support and $\varphi_i \in \mathcal{H}$ in the following way

$$DF = \sum_{i=1}^n \frac{\partial g}{\partial x_i}(B(\varphi_1), \dots, B(\varphi_n)) \varphi_i.$$

The operator D is closable and it can be extended to the closure of the set of cylindrical random variables (denotes $\mathbb{D}^{1,2}$) with respect to the norm

$$\|F\|_{1,2}^2 := \mathbf{E}|F|^2 + \mathbf{E}\|DF\|_{\mathcal{H}}^2.$$

If $F = I_p(f)$ with $f \in \mathcal{H}^{\odot p}$ and $p \geq 1$, then

$$DF = pI_{p-1}(f(\cdot, *)),$$

where ”*” stands for $p - 1$ variables.

In our work, we have $\mathcal{H} = L^2(\mathbb{R}_+ \times \mathbb{R})$ while the role of the isonormal process $(B(\varphi), \varphi \in \mathcal{H})$ is played by $(\int_{\mathbb{R}_+} \int_{\mathbb{R}} \varphi(s, y) W(ds, dy), \varphi \in \mathcal{H} = L^2(\mathbb{R}_+ \times \mathbb{R}))$, where W is the space-time white noise.

In our proofs, we used the following theorem which provides a description of the normal approximation of a sequence of multiple stochastic integrals $(G_N)_{N \geq 1}$ under several well-known distances d (Kolmogorov, Total Variation, Wasserstein) (c.f [93] Appendix C). Let us first recall the definition of these distances. Usually, the distance between the laws two of real -valued random variables F and G is defined as

$$d_W(F, G) = \sup_{h \in \mathcal{A}} |\mathbf{E}h(F) - \mathbf{E}h(G)|, \quad (6.54)$$

where \mathcal{A} is a class of of functions. When \mathcal{A} is the set of Lipschitz continuous function $h : \mathbb{R} \rightarrow \mathbb{R}$ such that $\|h\|_{Lip} \leq 1$, where

$$\|h\|_{Lip} = \sup_{x, y \in \mathbb{R}^d, x \neq y} \frac{|h(x) - h(y)|}{|x - y|}$$

then (6.54) gives the Wasserstein distance. When \mathcal{A} is the set of indicator functions $\{1_{-\infty, z], z \in \mathbb{R}\}$ the (6.54) gives the Kolmogoriv distance while for $\mathcal{A} = \{1_B, B \in \mathcal{B}(\mathbb{R})\}$ we have the total variation distance.

Let us state the following result that characterizes the normal convergence of a sequence of multiple stochastic integrals and that also gives an explicit bound for the distance between the law of G_N and the normal law. We refer, among others, to [101], [94] and [93] and references therein for more details.

Theorem 4. Fix $q \geq 1$. Assume that $(G_N)_{N \geq 1} = (I_q(g_N))_{N \geq 1}$ with $g_N \in \mathcal{H}^{\odot q}$, a sequence of random variables belonging to the q th Wiener chaos such that :

$$\mathbf{E}(G_N^2) \xrightarrow{N \rightarrow \infty} \sigma^2.$$

Then, G_N converges in law to $Z \sim \mathcal{N}(0, \sigma^2)$ if and only if

$$\|DG_N\|_{\mathcal{H}}^2 \xrightarrow{N \rightarrow \infty} q\sigma^2.$$

Furthermore,

$$d(G_N, \mathcal{N}(0, \sigma^2)) \leq C \left(\sqrt{\mathbf{Var}(\|DG_N\|_{\mathcal{H}}^2)} + \sqrt{\mathbf{E}(\|DG_N\|_{\mathcal{H}}^2) - q\sigma^2} \right). \quad (6.55)$$

We also recall a result concerning the optimality of the rate of convergence in Theorem 4. Consider $(F_N, N \geq 1)$ a sequence in the q th Wiener chaos. The main result in [95] says that if $\mathbf{E}F_N^2 = 1$ for every $N \geq 1$, then

$$d_{TV}(F_N, \mathcal{N}(0, 1)) \propto M(N), \quad (6.56)$$

where $u_n \propto v_n$ means that $0 < \limsup_{n \rightarrow \infty} \frac{u_n}{v_n} < \infty$ and

$$M(N) := \max \left(\mathbf{E}(F_N^3), \mathbf{Var} \left(\frac{1}{q} \|DF_N\|_{\mathcal{H}}^2 \right) \right).$$

Exact variation and drift parameter estimation for the nonlinear fractional stochastic heat equation

This article is accepted in Japanese Journal of Statistics and Data Science.

Joint work with Ciprian A. Tudor.

Abstract. This work concerns the statistical inference for stochastic partial differential equations. We consider the fractional stochastic heat equation driven by a nonlinear Gaussian space-time white noise and we analyze its mild solution. We actually study the limit behavior of the spatial quadratic variation of its mild solution both in the linear and nonlinear noise cases by obtaining the exact limit of this quadratic variation. We apply these results to parameter estimation. More precisely, we construct an estimator for the drift parameter of the fractional stochastic heat equation with nonlinear noise, which is defined in terms of the quadratic variation and it is based on the observation of the solution at a fixed time and at discrete points in space. The proofs are based on the relation between the solution to the linear fractional stochastic heat equation and the fractional Brownian motion and on a sharp analysis of the Green kernel associated to the fractional Laplacian operator.

7.1 Introduction

We consider the following nonlinear stochastic heat equation

$$\frac{\partial}{\partial t} u_{\theta}(t, x) = -\theta(-\Delta)^{\frac{\alpha}{2}} u_{\theta}(t, x) + \sigma(u_{\theta}(t, x)) \dot{W}(t, x), \quad t \in [0, T], x \in \mathbb{R} \quad (7.1)$$

with $\theta > 0$ and with vanishing initial condition $u_{\theta}(0, x) = 0$ for every $x \in \mathbb{R}$. The random perturbation W in (7.1) is a space-time white noise while the symbol $(-\Delta)^{\frac{\alpha}{2}}$ stands for the fractional

Laplacian operator of order $\alpha \in (1, 2)$. The solution to (7.1) is understood in the mild sense, i.e. for $t \in [0, T]$ and $x \in \mathbb{R}$,

$$u_\theta(t, x) = \int_0^t \int_{\mathbb{R}} G_\alpha(\theta(t-s), x-y) \sigma(u_\theta(s, y)) W(ds, dy) \quad (7.2)$$

where the stochastic integral $W(ds, dy)$ is a Dalang-Walsh integral and G_α is the Green kernel corresponding to the fractional Laplacian. These notions will be described later in Section 7.2.

The statistical inference for stochastic partial differential equations (SPDEs in the sequel) represents a very dynamical research direction. The website <https://sites.google.com/prod/view/stats4spdes/> provides a long list of research articles on the estimation of the parameters in various types of SPDEs. The survey [31] describes several methods to identify the parameters that may appear in stochastic heat, wave or other equations.

Our purpose is to estimate the drift parameter $\theta > 0$ based on the discrete observations of the solution to the SPDE (7.1). We will assume that the solution is observed at fixed time and at discrete points in space. Our estimators will be constructed via the quadratic variations with respect to the space variable of the solution u_θ . The p -variation method is a well-known technical to identify the parameters of SPDEs and it has been used in many works, see, among others, in [2], [11], [33], [32], [34], [35], [27], [28], [14], [15], [60], [72], [76], [104], [59] and [129]. While many of these references concern the case of SPDEs with additive Gaussian noise, there are relatively few works which treat the case of stochastic diffusion coefficient σ in (7.1). Among these works, we refer to [27] for the study of the p -variation of the nonlinear heat equation with standard Laplacian (i.e. $\alpha = 2$ in (7.1)) driven by a Gaussian noise with spatial correlation and to [104] for the case of the standard Laplacian and of the space-time white noise. See also [33] for a statistical analysis of some nonlinear SPDEs. The analysis of the p -variation in time of the solution to the nonlinear fractional stochastic heat equation (on a bounded spatial domain) appears in the recent paper [37].

We aim at extending the results in the above references by analyzing the spatial quadratic variations of the solution to fractional stochastic heat equation with multiplicative space-time white noise. Our approach is based on the link between the solution to the linear heat equation (i. e. $\sigma \equiv 1$ in (7.1)) and the fractional Brownian motion (fBm in the sequel). Indeed, as we explain later, this solution coincides (viewed as a stochastic process with respect to its space variable) with a perturbed fractional Brownian motion (i.e. the sum of a fBm and of a Gaussian process with regular sample paths). Then we apply these results related to fBm in order to construct consistent estimators for the parameter θ in (7.1). Our estimators are expressed in terms of the discrete observations of the solution. The idea to use the link between the solution to the stochastic heat equation with additive noise and fBm in order to get the behavior of the power variation of the solution and to estimate various parameters of the equation is not new. For instance, it has been used in [33], [35], [75] or [76]. See also the recent monograph [117] for a survey of the relationships between fBm and the solutions to stochastic heat equation.

We organized our paper as follows. In Section 2 we recall several basic facts concerning the stochastic heat equation and its solution and the Green kernel associated to the fractional Laplacian. We also included some technical (new) results related to this Green kernel. Section 3 contains some results on the limit behavior of the quadratic variation of the perturbed fBm while Section 4 is the core of this work. Here we included the main result concerning the asymptotic behavior of the spatial quadratic variations of the solution the SPDE (7.1). Section 6 is devoted to the statistical inference, here we construct a consistent estimator for the drift parameter θ in (7.1). The Appendix (Section 7) contains the proofs of our technical lemmas.

7.2 Preliminaries : On the stochastic fractional heat equation and its solution

Let us recall some basic facts concerning the solution to (7.1). The random noise in this SPDE is defined as a centered Gaussian field $(W(t,A), t \in [0, T], A \in \mathcal{B}_b(\mathbb{R}))$ with covariance

$$\mathbf{E}W(t,A)W(s,B) = (t \wedge s)\lambda(A \cap B), \quad s, t \in [0, T], A, B \in \mathcal{B}_b(\mathbb{R}), \quad (7.3)$$

where λ denotes the Lebesgue measure on \mathbb{R} and $\mathcal{B}_b(\mathbb{R})$ is the set of bounded Borel subsets of \mathbb{R} . We will assume that the coefficient $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ is Lipschitz continuous, i.e. for $x, y \in \mathbb{R}$,

$$|\sigma(x) - \sigma(y)| \leq K|x - y|. \quad (7.4)$$

Let us first consider $\theta = 1$ in (7.1). That is, we deal with the SPDE

$$\frac{\partial}{\partial t}u(t,x) = -(-\Delta)^{\frac{\alpha}{2}}u(t,x) + \sigma(u(t,x))\dot{W}(t,x), \quad t \in [0, T], x \in \mathbb{R}. \quad (7.5)$$

The solution to (7.5) is considered in the mild sense, i.e.

$$u(t,x) = \int_0^t \int_{\mathbb{R}} G_{\alpha}(t-s, x-y)\sigma(u(s,y))W(ds, dy), \quad (7.6)$$

where G_{α} stands for the Green kernel associated to the fractional heat equation and the stochastic integral $W(ds, dy)$ is the so-called Dalang-Walsh integral. We recall below some basic facts about these objects.

7.2.1 The Dalang-Walsh integral

For $t \geq 0$, we denote by \mathcal{F}_t the sigma-algebra generated by $(W(s,A), 0 \leq s \leq t, A \in \mathcal{B}_b(\mathbb{R}))$. For a jointly measurable random field $(X(s,y), s \geq 0, y \in \mathbb{R})$ adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$ generated by W and satisfying

$$\mathbf{E} \int_0^{\infty} \int_{\mathbb{R}} X(s,y)^2 dy ds < \infty,$$

we can define the Dalang-Walsh stochastic integral

$$\int_0^{\infty} \int_{\mathbb{R}} X(s,y)W(ds, dy).$$

We refer to [39] or [125] for the detailed construction of this integral. We recall that the Dalang-Walsh stochastic integral satisfies the Itô-type isometry

$$\mathbf{E} \left(\int_0^{\infty} \int_{\mathbb{R}} X(s,y)W(ds, dy) \right)^2 = \mathbf{E} \int_0^{\infty} \int_{\mathbb{R}} X(s,y)^2 dy ds.$$

7.2.2 Properties of the Green kernel

For the definition and other properties of the fractional Laplacian, one may consult, among others, the references [64], [65]. Here, we will mainly work with the Green kernel associated to this operator. The best way to define the Green kernel G_α is via its Fourier transform (we use $\mathcal{F}f$ to denote the Fourier transform of the function f), namely

$$\mathcal{F}G_\alpha(t, \cdot)(\xi) := \int_{\mathbb{R}} e^{ix\xi} G_\alpha(t, x) dx = e^{-t|\xi|^\alpha} \quad (7.7)$$

for $t \in [0, T]$ and $\xi \in \mathbb{R}$, i.e.

$$G_\alpha(t, x) = (2\pi)^{-1} \int_{\mathbb{R}} e^{-ix\xi} e^{-t|\xi|^\alpha} d\xi.$$

Let us list some useful properties of the Green kernel (7.7).

- For every $t > 0$, $G_\alpha(t, \cdot)$ is the density of the stable Lévy process. In particular, this implies

$$\int_{\mathbb{R}} G_\alpha(t, x) dx = 1.$$

- For every $t > 0$, the kernel $G_\alpha(t, x)$ is positive, real-valued and symmetric with respect to the space variable $x \in \mathbb{R}$.
- We have the scaling property

$$G_\alpha(t, x) = t^{-\frac{1}{\alpha}} G_\alpha(1, t^{-\frac{1}{\alpha}} x)$$

for $t > 0, x \in \mathbb{R}$.

- There exists two constants $0 < K'_\alpha < K_\alpha$ such that for every $t \in [0, T], x \in \mathbb{R}$,

$$K'_\alpha \frac{t^{-\frac{1}{\alpha}}}{\left(1 + |t^{-\frac{1}{\alpha}} x|\right)^{1+\alpha}} \leq G_\alpha(t, x) \leq K_\alpha \frac{t^{-\frac{1}{\alpha}}}{\left(1 + |t^{-\frac{1}{\alpha}} x|\right)^{1+\alpha}}. \quad (7.8)$$

The existence and uniqueness of the mild solution (7.6) under assumption (7.4) have been showed in [41] (see also [9]). We will need the following facts concerning the moments and the pathwise regularity of the mild solution (7.6), see again [9], [41].

1. For every $p \geq 1$, we have

$$\sup_{t \in [0, T], x \in \mathbb{R}} \mathbf{E} |u(t, x)|^p \leq C_T. \quad (7.9)$$

2. For every $s, t \in [0, T], x, y \in \mathbb{R}$ and for every $p \geq 2$,

$$\mathbf{E} |u(t, x) - u(s, y)|^p \leq C_T \left[|t - s|^{(1-\frac{1}{\alpha})\frac{p}{2}} + |x - y|^{(\alpha-1)\frac{p}{2}} \right]. \quad (7.10)$$

Let us also recall the Parseval-Plancherel identity

$$\langle f, g \rangle_{L^2(\mathbb{R})} = (2\pi)^{-1} \int_{\mathbb{R}} (\mathcal{F}f)(\xi) \overline{(\mathcal{F}g)(\xi)} d\xi, \quad (7.11)$$

for any $f, g \in L^2(\mathbb{R})$.

7.2.3 Some technical lemmas on the Green kernel

The properties of the Green kernel G_α given by (7.7) will play an important role for the obtention of the main results. In this section, we state two technical results concerning the fundamental solution. Their proofs are postponed to the Appendix.

Lemma 7. *Let G_α be given by (7.7). Then for every $s, t \in [0, T]$ with $s \leq t$ and for every $\delta > 0$, we have*

$$I(s, t, \delta) := \int_0^s da \int_{\mathbb{R}} dy (G_\alpha(t-a, y+\delta) - G_\alpha(t-a, y))^2 \leq C_T \delta^2 (t-s)^{1-\frac{3}{\alpha}}.$$

Lemma 8. *For $0 \leq s \leq t \leq T$, it holds*

$$J(s, t) := \int_s^t da \int_{\mathbb{R}} dy G_\alpha^2(t-a, y) \leq C(t-s)^{1-\frac{1}{\alpha}}.$$

7.3 On the increments of the fractional Brownian motion

There is a strong connection between the solution to the linear stochastic heat equation and the fractional Brownian motion. We recall that the fBm with Hurst parameter $H \in (0, 1)$ is defined as a centered Gaussian process with covariance

$$\mathbf{E} B_t^H B_s^H = \frac{1}{2} (t^{2H} + s^{2H} - |t-s|^{2H}), \quad s, t \geq 0.$$

The behavior of the quadratic variation of the fBm and the correlation structure of its increments will play an important role for the analysis of the quadratic variation of the stochastic fractional heat equation with additive noise ($\sigma \equiv 1$ in (7.5)), and then, by an approximation argument, for the quadratic variation of the nonlinear fractional heat equation. We will start with some lemmas concerning the increments and the variation of the fBm and of the perturbed fBm.

Lemma 9. *Let $(B_t^H)_{t \geq 0}$ be a fBm with Hurst index $H \in (0, 1)$. Then, for every $i = 0, \dots, N-1$,*

$$\mathbf{E} \left[N^{2H-1} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2 - \frac{1}{N} \right]^2 = \frac{2}{N^2}$$

and for every $i, j = 0, \dots, N-1$ with $i \neq j$,

$$\mathbf{E} \left[N^{2H-1} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2 - \frac{1}{N} \right] \left[N^{2H-1} \left(B_{\frac{j+1}{N}}^H - B_{\frac{j}{N}}^H \right)^2 - \frac{1}{N} \right] = \frac{1}{N^2} \rho_H^2(i-j)$$

with

$$\rho_H(k) = \frac{1}{2} (|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}). \quad (7.12)$$

Proof : If $\stackrel{(d)}{=}$ denotes the equality in distribution, we can write, for every $i = 0, \dots, N-1$,

$$N^{2H-1} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2 \stackrel{(d)}{=} \frac{1}{N} Z^2$$

with $Z \sim \mathcal{N}(0, 1)$. Then

$$\mathbf{E} \left[N^{2H-1} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2 - \frac{1}{N} \right]^2 = \frac{1}{N^2} \mathbf{E} (Z^2 - 1)^2 = \frac{2}{N^2}.$$

Take now $i, j = 0, \dots, N-1$ with $i \neq j$. Then, by the scaling property of the fBm,

$$\begin{aligned} & \mathbf{E} \left[N^{2H-1} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2 - \frac{1}{N} \right] \left[N^{2H-1} \left(B_{\frac{j+1}{N}}^H - B_{\frac{j}{N}}^H \right)^2 - \frac{1}{N} \right] \\ &= \frac{1}{N^2} \mathbf{E} [(B_{i+1}^H - B_i^H)^2 - 1] [(B_{j+1}^H - B_j^H)^2 - 1] \\ &= \frac{2}{N^2} (\mathbf{E}(B_{i+1}^H - B_i^H)(B_{j+1}^H - B_j^H))^2 \\ &= \frac{2}{N^2} \rho_H^2 (i - j) \end{aligned}$$

with ρ_H given by (7.12). Above (and in the sequel) we use the fact that if X, Y are centered Gaussian random variables, then $\text{Cov}(X^2, Y^2) = 2(\text{Cov}(X, Y))^2$. \square

Lemma 10. *Let $(Y_t)_{t \geq 0}$ be a centered Gaussian process such that for every $s, t \geq 0$, one has*

$$\mathbf{E} |Y_t - Y_s|^2 \leq \frac{1}{N^{2a}} \quad (7.13)$$

with $a > H$. Let $C_0 > 0$. Assume that $(X_t)_{t \geq 0}$ is a Gaussian process such that

$$X_t = C_0 B_t^H + Y_t, \quad t \geq 0. \quad (7.14)$$

Then for every $i = 0, \dots, N-1$,

$$\mathbf{E} \left[N^{2H-1} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 - \frac{C_0^2}{N} \right]^2 \leq C \frac{1}{N^2} \quad (7.15)$$

and for every $i, j = 0, \dots, N-1$ with $i \neq j$,

$$\mathbf{E} \left[N^{2H-1} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 - \frac{C_0^2}{N} \right] \left[N^{2H-1} \left(X_{\frac{j+1}{N}} - X_{\frac{j}{N}} \right)^2 - \frac{C_0^2}{N} \right] \leq C \frac{1}{N^2} (\rho_H^2 (i - j) + 1) \quad (7.16)$$

with ρ_H from (7.12). Above and in the sequel, we denote by C a generic strictly positive constant that may change from one line to another (or even on the same line).

Proof : We can bound the left-hand side of (7.15) as follows

$$\begin{aligned} & \mathbf{E} \left[N^{2H-1} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 - \frac{C_0^2}{N} \right]^2 \\ &= \mathbf{E} \left[C_0^2 \left(N^{2H-1} (B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H)^2 - \frac{1}{N} \right) + 2N^{2H-1} C_0 (B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H) (Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}}) \right. \\ & \quad \left. + N^{2H-1} (Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}})^2 \right]^2 \\ &\leq C \left[\frac{1}{N^2} + N^{4H-2} \mathbf{E} (B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H)^2 (Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}})^2 + N^{4H-2} \mathbf{E} (Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}})^4 \right], \end{aligned}$$

where we used Lemma 9. By (7.13) and the fact that B^H and Y are Gaussian processes,

$$\begin{aligned} & \mathbf{E}(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H)^2 (Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}})^2 \\ & \leq \left(\mathbf{E}(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H)^4 \right)^{\frac{1}{2}} \left(\mathbf{E}(Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}})^4 \right)^{\frac{1}{2}} \leq C \frac{1}{N^{2H+2a}} \end{aligned}$$

and

$$\mathbf{E}(Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}})^4 \leq C \frac{1}{N^{4a}}.$$

Thus, since $a > H$,

$$\begin{aligned} & \mathbf{E} \left[N^{2H-1} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 - \frac{C_0^2}{N} \right]^2 \\ & \leq C \left[\frac{1}{N^2} + \frac{1}{N^{2+2a-2H}} + \frac{1}{N^{2+4a-4H}} \right] \leq C \frac{1}{N^2}. \end{aligned}$$

Let us prove the claim (7.16). Notice that, for every $N \geq 1$ and for every $i = 0, 1, \dots, N-1$,

$$\begin{aligned} & N^{2H-1} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 - \frac{C_0^2}{N} \\ & = N^{2H-1} \left(\left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 - \mathbf{E} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 \right) + N^{2H-1} \mathbf{E} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 - \frac{C_0^2}{N} \end{aligned} \quad (7.17)$$

and, via Cauchy-Schwarz, inequality and (7.13),

$$\begin{aligned} & \left| N^{2H-1} \mathbf{E} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 - \frac{C_0^2}{N} \right| \\ & \leq 2N^{2H-1} \mathbf{E} \left| B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right| \left| Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}} \right| + N^{2H-1} \mathbf{E} \left| Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}} \right|^2 \\ & \leq CN^{2H-1} \left(\frac{1}{N^{H+a}} + \frac{1}{N^{2a}} \right) \leq C \frac{1}{N^{1+a-H}}. \end{aligned} \quad (7.18)$$

On the other hand, since X is a centered Gaussian process, for every $i \neq j$,

$$\begin{aligned} & \mathbf{E} \left[N^{2H-1} \left(\left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 - \mathbf{E} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 \right) N^{2H-1} \left(\left(X_{\frac{j+1}{N}} - X_{\frac{j}{N}} \right)^2 - \mathbf{E} \left(X_{\frac{j+1}{N}} - X_{\frac{j}{N}} \right)^2 \right) \right] \\ & = 2N^{4H-2} \left(\mathbf{E} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right) \left(X_{\frac{j+1}{N}} - X_{\frac{j}{N}} \right) \right)^2 \end{aligned}$$

and by (7.13) and (7.14),

$$\left| \mathbf{E} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right) \left(X_{\frac{j+1}{N}} - X_{\frac{j}{N}} \right) \right| \leq C \left(\frac{1}{N^{2H}} |\rho_H(i-j)| + \frac{1}{N^{a+H}} + \frac{1}{N^{2a}} \right) \leq C \frac{1}{N^{2H}} (|\rho_H(i-j)| + 1).$$

Consequently,

$$\begin{aligned} & \mathbf{E} \left[N^{2H-1} \left(\left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 - \mathbf{E} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2 \right) N^{2H-1} \left(\left(X_{\frac{j+1}{N}} - X_{\frac{j}{N}} \right)^2 - \mathbf{E} \left(X_{\frac{j+1}{N}} - X_{\frac{j}{N}} \right)^2 \right) \right] \\ & \leq C \frac{1}{N^2} (\rho_H^2(i-j) + 1). \end{aligned} \quad (7.19)$$

By (7.17), (7.18) and (7.19), we obtain the conclusion. \square

Next, we state and prove two results on the quadratic variation of the fBm and of the perturbed fBm.

Proposition 5. *Let $(B_t^H)_{t \geq 0}$ be a fractional Brownian motion with Hurst index $H \in (0, \frac{3}{4})$. For $N \geq 1$, denote*

$$V_N(B^H) = N^{2H-1} \sum_{i=0}^{N-1} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2.$$

Then the sequence $(V_N(B^H), N \geq 1)$ converges to 1 as $N \rightarrow \infty$ in $L^2(\Omega)$.

Proof : It is immediate that $\mathbf{E}V_N = 1$ for every $N \geq 1$ and

$$\begin{aligned} & \mathbf{E}(V_N(B^H) - 1)^2 \\ &= \mathbf{E} \left[N^{2H-1} \sum_{i=0}^{N-1} \left(\left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2 - \mathbf{E} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2 \right) \right]^2 \\ &= N^{4H-2} \sum_{i,j=0}^{N-1} \mathbf{E} \left(\left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2 - \mathbf{E} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2 \right) \left(\left(B_{\frac{j+1}{N}}^H - B_{\frac{j}{N}}^H \right)^2 - \mathbf{E} \left(B_{\frac{j+1}{N}}^H - B_{\frac{j}{N}}^H \right)^2 \right) \\ &= 2N^{4H-2} \sum_{i,j=0}^{N-1} \left(\mathbf{E} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right) \left(B_{\frac{j+1}{N}}^H - B_{\frac{j}{N}}^H \right) \right)^2 \\ &= 2N^{4H-2} \sum_{i=0}^{N-1} \left(\mathbf{E} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right)^2 \right)^2 \\ &\quad + 4N^{4H-2} \sum_{i,j=0; i>j}^{N-1} \left(\mathbf{E} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right) \left(B_{\frac{j+1}{N}}^H - B_{\frac{j}{N}}^H \right) \right)^2 \\ &= \frac{2}{N} + 4 \frac{1}{N^2} \sum_{i,j=0; i>j}^{N-1} \rho_H(i-j)^2 \\ &= \frac{2}{N} + 4 \frac{1}{N^2} \sum_{k=1}^{N-1} (N-k) \rho_H(k)^2 \end{aligned}$$

with ρ_H from (7.12). Since $\rho_H(k)$ behaves as $H(2H-1)k^{2H-2}$ for k large enough, we get

$$\mathbf{E}(V_N(B^H) - 1)^2 \leq C \frac{1}{N}$$

for N large enough, by using $H < \frac{3}{4}$. \square

Proposition 6. *Let $(B_t^H)_{t \geq 0}$ be a fBm with $H \in (0, \frac{3}{4})$ and let $(X_t)_{t \geq 0}$ be a perturbed fBm defined by (7.14). Let*

$$V_N(X) = N^{2H-1} \sum_{i=0}^{N-1} \left(X_{\frac{i+1}{N}} - X_{\frac{i}{N}} \right)^2.$$

Then the sequence $(V_N(X), N \geq 1)$ converges to C_0^2 in $L^1(\Omega)$, with C_0 from (7.14).

Proof : We can write, for every $N \geq 1$,

$$\begin{aligned} V_N(X) &= C_0^2 V_N(B^H) + 2C_0 N^{2H-1} \sum_{i=0}^{N-1} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right) \left(Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}} \right) \\ &\quad + N^{2H-1} \sum_{i=0}^{N-1} \left(Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}} \right)^2. \end{aligned}$$

By Proposition 5, $C_0^2 V_N(B^H)$ converges to C_0^2 in $L^2(\Omega)$. We show below that the other two summands go to zero in $L^1(\Omega)$ as $N \rightarrow \infty$. Indeed, by (7.13),

$$\begin{aligned} \mathbf{E} \left| N^{2H-1} \sum_{i=0}^{N-1} \left(B_{\frac{i+1}{N}}^H - B_{\frac{i}{N}}^H \right) \left(Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}} \right) \right| &\leq CN^{2H-1} \sum_{i=0}^{N-1} \frac{1}{N^{a+H}} \\ &\leq C \frac{1}{N^{a-H}} \xrightarrow{N \rightarrow \infty} 0 \end{aligned}$$

and

$$\mathbf{E} \left| N^{2H-1} \sum_{i=0}^{N-1} \left(Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}} \right)^2 \right| \leq CN^{2H-1} \sum_{i=0}^{N-1} \frac{1}{N^{2a}} \leq C \frac{1}{N^{2a-2H}} \xrightarrow{N \rightarrow \infty} 0,$$

by using the assumption $a > H$. □

7.4 The spatial quadratic variation of the solution in the nonlinear case

Let $t > 0$ be fixed. We set, for every $N \geq 1$,

$$V_{N,t} = N^{\alpha-2} \sum_{i=0}^{N-1} (u(t, x_{i+1}) - u(t, x_i))^2 \quad (7.20)$$

with

$$x_i = \frac{i}{N} \text{ for } i = 0, 1, \dots, N. \quad (7.21)$$

We analyze the asymptotic behavior, as $N \rightarrow \infty$, of the sequence $(V_{N,t}, N \geq 1)$.

The first part of this section will contain a review of some properties of the solution to the linear fractional stochastic heat equation and we will deduce the behavior of its quadratic variation. The second part will be devoted to the nonlinear noise case.

7.4.1 The linear case

We start by introducing the centered Gaussian random field $(u_0(t, x), t \in [0, T], x \in \mathbb{R})$ given by

$$u_0(t, x) = \int_0^t \int_{\mathbb{R}} G_\alpha(t-s, x-y) W(ds, dy). \quad (7.22)$$

That is, u_0 is the mild solution to (7.5) with $\sigma \equiv 1$.

We recall below an interesting fact about the distribution of the random field u_0 . For their proof, we refer to [51], [76], [121].

— For every $t > 0$, the centered Gaussian process $(u_0(t, x), x \in \mathbb{R})$ can be decomposed as

$$u_0(t, x) \stackrel{(d)}{=} m_\alpha U(t, x) + Y(t, x), \quad x \in \mathbb{R} \quad (7.23)$$

where for every $t \in (0, T]$, $(U(t, x), x \in \mathbb{R})$ is a fractional Brownian motion with Hurst index $H = \frac{\alpha-1}{2}$, $m_\alpha = (2\Gamma(\alpha) |\cos(\frac{\alpha\pi}{2})|)^{-\frac{1}{2}}$ and $(Y(t, x), x \in \mathbb{R})$ is a Gaussian process satisfying

$$\mathbf{E}|Y(t, x) - Y(t, y)|^2 \leq C_T |x - y|^2$$

for every $t \in [0, T]$ and $x, y \in \mathbb{R}$.

For $t \in [0, T]$ and $N \geq 1$, define

$$V_{N,t}(u_0) = N^{\alpha-2} \sum_{i=0}^{N-1} (u_0(t, x_{i+1}) - u_0(t, x_i))^2 \quad (7.24)$$

with $x_i, i = 0, \dots, N-1$ given by (7.21). The behavior of the above sequence can be deduced directly from Proposition 6.

Proposition 7. *Consider the sequence $(V_{N,t}(u_0), N \geq 1)$ given by (7.24) with $t \in [0, T]$ fixed. Then*

$$V_{N,t}(u_0) \xrightarrow[N \rightarrow \infty]{} m_\alpha^2 \text{ in } L^1(\Omega),$$

where m_α is the constant appearing in (7.23).

Proof : From (7.23), for every $t \in [0, T]$, the Gaussian process $(u_0(t, x), x \in \mathbb{R})$ is a perturbed fBm. It suffices to apply Proposition 6 with $a = 1$ in (7.13). \square

7.4.2 The nonlinear case

Let us now analyze the asymptotic behavior, as $N \rightarrow \infty$, of the sequence $V_{N,t}$ given by (7.20). The idea is to approximate the increment $u(t, x_{i+1}) - u(t, x_i)$ of the solution to the nonlinear heat equation by $\sigma(u(t(\delta), x_i))(u_0(t, x_{i+1}) - u_0(t, x_i))$ where $x_i, i = 0, \dots, N$ are given by (7.21), $\delta = \frac{1}{N}$, u_0 is the solution to the linear heat equation given by (7.22) and $t(\delta)$ is a point close to t , suitably chosen.

Let us first introduce some notation. We denote by $(\Delta u_0)(x, \delta)$ the spatial increment of u_0 between x and $x + \delta$ with $x \in \mathbb{R}$ and $\delta > 0$, i.e.

$$\begin{aligned} (\Delta u_0)(x, \delta) &= u_0(t, x + \delta) - u_0(t, x) \\ &= \int_0^t \int_{\mathbb{R}} (G_\alpha(t-s, x + \delta - y) - G_\alpha(t-s, x - y)) W(ds, dy). \end{aligned}$$

We will also consider the "modified spatial increment" of u_0 given by

$$\begin{aligned} (\tilde{\Delta} u_0)(x, \delta) &= \int_{t(\delta)}^t \int_{\mathbb{R}} (G_\alpha(t-s, x + \delta - y) - G_\alpha(t-s, x - y)) W(ds, dy) \\ &\quad + \int_0^{t(\delta)} \int_{\mathbb{R}} (G_\alpha(t-s, x + \delta - y) - G_\alpha(t-s, x - y)) \tilde{W}(ds, dy) \end{aligned} \quad (7.25)$$

where \tilde{W} is an independent copy of the space-time white noise W and

$$t(\delta) = t - \left(\frac{1}{N}\right)^\gamma \text{ with } \gamma = \frac{2\alpha}{\alpha+1}. \quad (7.26)$$

Let us notice the following facts :

- For every x, δ , the random variable $(\tilde{\Delta}u_0)(x, \delta)$ is a centered Gaussian random variable with the same distribution as $(\Delta u_0)(x, \delta)$.
- For every x, δ , the random variable $(\tilde{\Delta}u_0)(x, \delta)$ is independent of the sigma-algebra generated by $\{u_s, 0 \leq s \leq t(\delta)\}$. This comes from the fact that W has independent increments in time and \tilde{W} is independent by W .

Next, we prove an approximation result for the spatial increment of the random field (7.6).

Proposition 8. *For $t > 0, x \in \mathbb{R}$, and for $\delta > 0$ small enough, we have*

$$\mathbf{E} \left| (u(t, x + \delta) - u(t, x)) - \sigma(u(t(\delta), x))(\tilde{\Delta}u_0)(x, \delta) \right|^2 \leq C_T \delta^{\frac{4(\alpha-1)}{\alpha+1}}$$

Proof : Notice that, since $(\tilde{\Delta}u_0)(x, \delta)$ is independent of the sigma-algebra generated by $\{u_s, 0 \leq s \leq t(\delta)\}$, the quantity $\sigma(u(t(\delta), x))(\tilde{\Delta}u_0)(x, \delta)$ can be written as, by (7.25),

$$\begin{aligned} & \sigma(u(t(\delta), x))(\tilde{\Delta}u_0)(x, \delta) \\ &= \int_{t(\delta)}^t \int_{\mathbb{R}} (G_\alpha(t-s, x+\delta-y) - G_\alpha(t-s, x-y)) \sigma(u(t(\delta), x)) W(ds, dy) \\ & \quad + \int_0^{t(\delta)} \int_{\mathbb{R}} (G_\alpha(t-s, x+\delta-y) - G_\alpha(t-s, x-y)) \sigma(u(t(\delta), x)) \tilde{W}(ds, dy), \end{aligned}$$

We can express the difference $(u(t, x + \delta) - u(t, x)) - \sigma(u(t(\delta), x))(\tilde{\Delta}u_0)(x, \delta)$ as follows

$$\begin{aligned} & (u(t, x + \delta) - u(t, x)) - \sigma(u(t(\delta), x))(\tilde{\Delta}u_0)(x, \delta) \\ &= \int_{t(\delta)}^t \int_{\mathbb{R}} (G_\alpha(t-a, x+\delta-y) - G_\alpha(t-a, x-y)) (\sigma(u(a, y)) - \sigma(u(t(\delta), x))) W(da, dy) \\ & \quad + \int_0^{t(\delta)} \int_{\mathbb{R}} (G_\alpha(t-a, x+\delta-y) - G_\alpha(t-a, x-y)) \sigma(u(a, y)) W(da, dy) \\ & \quad - \int_0^{t(\delta)} \int_{\mathbb{R}} (G_\alpha(t-a, x+\delta-y) - G_\alpha(t-a, x-y)) \sigma(u(t(\delta), x)) \tilde{W}(da, dy) \\ &:= A_1(x, \delta) + A_2(x, \delta) + A_3(x, \delta). \end{aligned}$$

Consequently,

$$\mathbf{E} \left| (u(t, x + \delta) - u(t, x)) - \sigma(u(t(\delta), x))(\tilde{\Delta}u_0)(x, \delta) \right|^2 \leq 2(\mathbf{E}A_1(x, \delta)^2 + \mathbf{E}A_2(x, \delta)^2 + \mathbf{E}A_3(x, \delta)^2). \quad (7.27)$$

We estimate separately the quantities $\mathbf{E}A_1(x, \delta)^2$, $\mathbf{E}A_2(x, \delta)^2$ and $\mathbf{E}A_3(x, \delta)^2$. For the first one, we can write, since σ satisfies the Lipschitz condition (7.4),

$$\begin{aligned}
& \mathbf{E}A_1(x, \delta)^2 \\
&= \int_{t(\delta)}^t da \int_{\mathbb{R}} dy (G_\alpha(t-a, x+\delta-y) - G_\alpha(t-a, x-y))^2 \mathbf{E}(\sigma(u(a, y)) - \sigma(u(t(\delta), x)))^2 \\
&\leq C \int_{t(\delta)}^t da \int_{\mathbb{R}} dy (G_\alpha(t-a, x+\delta-y) - G_\alpha(t-a, x-y))^2 \mathbf{E}|u(a, y) - u(t(\delta), x)|^2
\end{aligned}$$

and by using the bound (7.10) and $|a - t(\delta)| \leq |t - t(\delta)| \leq \delta^\gamma$

$$\begin{aligned}
& \mathbf{E}A_1(x, \delta)^2 \\
&\leq C \int_{t(\delta)}^t da \int_{\mathbb{R}} dy (G_\alpha(t-a, x+\delta-y) - G_\alpha(t-a, x-y))^2 \left[|x-y|^{\alpha-1} + \delta^{\gamma(1-\frac{1}{\alpha})} \right] \\
&\leq C \int_{t(\delta)}^t da \int_{\mathbb{R}} dy G_\alpha(t-a, x+\delta-y)^2 \left[|x-y|^{\alpha-1} + \delta^{\gamma(1-\frac{1}{\alpha})} \right] \\
&\quad + C \int_{t(\delta)}^t da \int_{\mathbb{R}} dy G_\alpha(t-a, x-y)^2 \left[|x-y|^{\alpha-1} + \delta^{\gamma(1-\frac{1}{\alpha})} \right] \\
&\leq C \int_{t(\delta)}^t da \int_{\mathbb{R}} dy G_\alpha(t-a, x+\delta-y)^2 \left[|x-y+\delta|^{\alpha-1} + |x-y|^{\alpha-1} + \delta^{\gamma(1-\frac{1}{\alpha})} \right] \\
&\leq C \int_{t(\delta)}^t da \int_{\mathbb{R}} dy G_\alpha(t-a, x+\delta-y)^2 \left[|x-y|^{\alpha-1} + \delta^{\alpha-1} + \delta^{\gamma(1-\frac{1}{\alpha})} \right] \\
&\leq C \int_{t(\delta)}^t da \int_{\mathbb{R}} dy G_\alpha(t-a, x+\delta-y)^2 \left[|x-y|^{\alpha-1} + \delta^{\gamma(1-\frac{1}{\alpha})} \right]
\end{aligned}$$

since, by (7.26), $\gamma < \alpha$ and δ is small enough. Recall that the constant $C > 0$ may change from line to line. So,

$$\begin{aligned}
& \mathbf{E}A_1(x, \delta)^2 \\
&\leq C \delta^{\gamma(1-\frac{1}{\alpha})} \int_{t(\delta)}^t da \int_{\mathbb{R}} dy G_\alpha(t-a, x+\delta-y)^2 \\
&\quad + C \int_{t(\delta)}^t da \int_{\mathbb{R}} dy G_\alpha(t-a, x+\delta-y)^2 |x-y|^{\alpha-1} \\
&\leq C \delta^{2\gamma(1-\frac{1}{\alpha})} + C \int_{t(\delta)}^t da \int_{\mathbb{R}} dy G_\alpha(t-a, x+\delta-y)^2 |x-y|^{\alpha-1} \tag{7.28}
\end{aligned}$$

where we used Lemma 8. Now, via the inequality (7.8),

$$\begin{aligned}
& \int_{\mathbb{R}} dy G_\alpha(t-a, x-y)^2 |x-y|^{\alpha-1} = \int_{\mathbb{R}} dy G_\alpha(t-a, y)^2 |y|^{\alpha-1} \\
&\leq C(t-a)^{-\frac{2}{\alpha}} \int_{\mathbb{R}} dy \left(\frac{1}{(1+|(t-a)^{-\frac{1}{\alpha}}y|)^{1+\alpha}} \right)^2 |y|^{\alpha-1} \\
&\leq (t-a)^{-\frac{2}{\alpha} + \frac{1}{\alpha} + \frac{\alpha-1}{\alpha}} \int_{\mathbb{R}} \left(\frac{1}{(1+|y|)^{1+\alpha}} \right)^2 |y|^{\alpha-1} = C(t-a)^{1-\frac{2}{\alpha}}. \tag{7.29}
\end{aligned}$$

By plugging (7.29) into (7.28), we obtain

$$\mathbf{E}A_1(x, \delta)^2 \leq C \delta^{2\gamma(1-\frac{1}{\alpha})} + C \int_{t(\delta)}^t da (t-a)^{1-\frac{2}{\alpha}} \leq C \delta^{\gamma(2-\frac{2}{\alpha})}. \tag{7.30}$$

Let us now estimate $\mathbf{EA}_2(x, \delta)^2$. By using successively the Itô isometry, the bound (7.9) and Lemma 7, we obtain

$$\begin{aligned}
 & \mathbf{EA}_2(x, \delta)^2 \\
 &= \int_0^{t(\delta)} da \int_{\mathbb{R}} dy (G_\alpha(t-a, x+\delta-y) - G_\alpha(t-a, x-y))^2 \mathbf{E} \sigma u(a, y)^2 \\
 &\leq C_T \int_0^{t(\delta)} da \int_{\mathbb{R}} dy (G_\alpha(t-a, x+\delta-y) - G_\alpha(t-a, x-y))^2 \\
 &\leq C_T \delta^2 \delta^{-\gamma \frac{3-\alpha}{\alpha}} = C_T \delta^{\frac{4(\alpha-1)}{\alpha+1}}. \tag{7.31}
 \end{aligned}$$

With the same proof, we get a similar bound for the summand $\mathbf{EA}_3(x, \delta)$, i.e. for $x \in \mathbb{R}$ and $\delta > 0$ small enough,

$$\mathbf{EA}_3(x, \delta) \leq C_T \delta^{\frac{4(\alpha-1)}{\alpha+1}}. \tag{7.32}$$

By combining (7.27) with (7.30), (7.31), (7.32), we obtain the desired conclusion. \square

Let us now state and prove the main result of this section.

Theorem 5. *Consider the random field $(u(t, x), t \in [0, T], x \in \mathbb{R})$ given by (7.6) and let $V_{N,t}$ be given by (7.20). Then for every $t > 0$,*

$$V_{N,t}(u) \xrightarrow[N \rightarrow \infty]{} m_\alpha^2 \int_0^1 \sigma^2(u(t, x)) dx \text{ in } L^1(\Omega)$$

and for N large enough,

$$\mathbf{E} \left| V_{N,t}(u) - m_\alpha^2 \int_0^1 \sigma^2(u(t, x)) dx \right| \leq C \left(\frac{1}{N} \right)^{\frac{(\alpha-1)(3-\alpha)}{2(\alpha+1)}}, \tag{7.33}$$

where the constant $C > 0$ does not depend on t .

Proof : We use the following decomposition, for every $t \in (0, T]$,

$$\begin{aligned}
 & V_{N,t}(u) - m_\alpha^2 \int_0^1 \sigma^2(u(t, x)) dx \\
 &= N^{\alpha-2} \sum_{i=0}^{N-1} \left[(u(t, x_{i+1}) - u(t, x_i))^2 - \sigma^2(u(t(\delta), x_i)) (\tilde{\Delta} u_0)^2(x_i, \delta) \right] \\
 &\quad + \sum_{i=0}^{N-1} \sigma^2(u(t(\delta), x_i)) \left(N^{\alpha-2} (\tilde{\Delta} u_0)^2(x_i, \delta) - m_\alpha^2 \frac{1}{N} \right) \\
 &\quad + m_\alpha^2 \left[\frac{1}{N} \sum_{i=0}^{N-1} \sigma^2(u(t(\delta), x_i)) - \int_0^1 \sigma^2(u(t, x)) dx \right] \\
 &:= B_{1,N} + B_{2,N} + B_{3,N}.
 \end{aligned}$$

where $\delta = \frac{1}{N}$, $t(\delta)$ is given by (7.26) and $\tilde{\Delta} u_0$ is the modified increment (7.25). Let us deal separately with the three summands above. We have, by using Cauchy-Schwarz's inequality and Proposition 8,

$$\begin{aligned}
& \mathbf{E}|B_{1,N}| \\
& \leq N^{\alpha-2} \sum_{i=0}^{N-1} \left[\mathbf{E} \left| (u(t, x_{i+1}) - u(t, x_i)) - \sigma(u(t(\delta), x_i)) (\tilde{\Delta}u_0)(x_i, \delta) \right|^2 \right]^{\frac{1}{2}} \\
& \quad \times \left(\mathbf{E} \left| u(t, x_{i+1}) - u(t, x_i) + \sigma(u(t(\delta), x_i)) \tilde{\Delta}u_0(x_i, \delta) \right|^2 \right)^{\frac{1}{2}} \\
& \leq C \delta^{2\frac{(\alpha-1)}{\alpha+1}} N^{\alpha-2} \sum_{i=0}^{N-1} \left(\mathbf{E} \left| (u(t, x_{i+1}) - u(t, x_i))^2 + \sigma^2(u(t(\delta), x_i)) \tilde{\Delta}u_0^2(x_i, \delta) \right| \right)^{\frac{1}{2}}
\end{aligned}$$

Now by using the fact that $(\tilde{\Delta}u_0)(x, \delta)$ is independent of the sigma-algebra generated by $\{u_s, 0 \leq s \leq t(\delta)\}$, and by (7.10) and (7.4), for every $t > 0$ and $i = 0, \dots, N$,

$$\begin{aligned}
& \mathbf{E} \left| (u(t, x_{i+1}) - u(t, x_i))^2 + \sigma^2(u(t(\delta), x_i)) (\tilde{\Delta}u_0)^2(x_i, \delta) \right| \\
& \leq \mathbf{E} |u(t, x_{i+1}) - u(t, x_i)|^2 + \mathbf{E} |\sigma(u(t(\delta), x_i))|^2 \mathbf{E} \left| (\tilde{\Delta}u_0)(x_i, \delta) \right|^2 \\
& \leq CN^{-(\alpha-1)} + C \mathbf{E} \left| (\tilde{\Delta}u_0)(x_i, \delta) \right|^2 \leq CN^{-(\alpha-1)}. \tag{7.34}
\end{aligned}$$

Consequently,

$$\mathbf{E}|B_{1,N}| \leq CN^{\alpha-2} \delta^{\frac{2(\alpha-1)}{\alpha+1}} \sum_{i=0}^{N-1} N^{-\frac{\alpha-1}{2}} = CN^{\frac{(\alpha-1)(\alpha-3)}{2(\alpha+1)}} \xrightarrow{N \rightarrow \infty} 0 \tag{7.35}$$

because $\alpha \in (1, 2)$. For $B_{2,N}$, we will estimate its norm in $L^2(\Omega)$. It holds, since $(\tilde{\Delta}u_0)(x_i, \delta)$ is independent of $\sigma(u(t(\delta), x_i))$ and via (7.4) and (7.9),

$$\begin{aligned}
& \mathbf{E}|B_{2,N}|^2 \\
& = \sum_{i,j=0}^N \mathbf{E} \sigma^2(u(t(\delta), x_i)) \sigma^2(u(t(\delta), x_j)) \\
& \quad \times \mathbf{E} \left(N^{\alpha-2} (\tilde{\Delta}u_0)^2(x_i, \delta) - m_\alpha^2 \frac{1}{N} \right) \left(N^{\alpha-2} (\tilde{\Delta}u_0)^2(x_j, \delta) - m_\alpha^2 \frac{1}{N} \right) \\
& \leq C \sum_{i,j=0}^N \left| \mathbf{E} \left(N^{\alpha-2} (\tilde{\Delta}u_0)^2(x_i, \delta) - m_\alpha^2 \frac{1}{N} \right) \left(N^{\alpha-2} (\tilde{\Delta}u_0)^2(x_j, \delta) - m_\alpha^2 \frac{1}{N} \right) \right| \\
& = C \sum_{i=0}^{N-1} \mathbf{E} \left(N^{\alpha-2} (\tilde{\Delta}u_0)^2(x_i, \delta) - m_\alpha^2 \frac{1}{N} \right)^2 \\
& \quad + C \sum_{i,j=0; i>j}^N \left| \mathbf{E} \left(N^{\alpha-2} (\tilde{\Delta}u_0)^2(x_i, \delta) - m_\alpha^2 \frac{1}{N} \right) \left(N^{\alpha-2} (\tilde{\Delta}u_0)^2(x_j, \delta) - m_\alpha^2 \frac{1}{N} \right) \right|
\end{aligned}$$

Recall that for every $t \in [0, T]$, $(u_0(t, x), x \in \mathbb{R})$ is a perturbed fBm, due to (7.23). We have, by the inequality (7.15) in Lemma 10,

$$\sum_{i=0}^{N-1} \mathbf{E} \left(N^{\alpha-2} (\tilde{\Delta}u_0)^2(x_i, \delta) - m_\alpha^2 \frac{1}{N} \right)^2 = \sum_{i=0}^{N-1} \mathbf{E} \left(N^{\alpha-2} (\Delta u_0)^2(x_i, \delta) - m_\alpha^2 \frac{1}{N} \right)^2 \leq C \frac{1}{N^2}$$

and by (7.16) in Lemma 10, with ρ_H from (7.12),

$$\begin{aligned} & \sum_{i,j=0;i>j}^N \left| \mathbf{E} \left(N^{\alpha-2} (\tilde{\Delta}u_0)^2(x_i, \delta) - m_\alpha^2 \frac{1}{N} \right) \left(N^{\alpha-2} (\tilde{\Delta}u_0)^2(x_j, \delta) - m_\alpha^2 \frac{1}{N} \right) \right| \\ & \leq C \frac{1}{N^2} \sum_{i,j=0;i>j}^N \left(\rho_{\frac{\alpha-1}{2}}^2(|i-j|) + 1 \right) \leq C \left(\frac{1}{N^2} \sum_{k=1}^N (N-k) \rho_{\frac{\alpha-1}{2}}^2(k) + \frac{1}{N} \right) \leq C \frac{1}{N}. \end{aligned}$$

We conclude that

$$\mathbf{E}|B_{2,N}| \leq C \frac{1}{\sqrt{N}}, \quad (7.36)$$

and this converges to zero as $N \rightarrow \infty$.

Concerning the summand denoted by $B_{3,N}$, it clearly converges almost surely to zero as $N \rightarrow \infty$ by a Riemann sum approximation. Let us regard its convergence in $L^1(\Omega)$. We can express $B_{3,N}$ as

$$\begin{aligned} B_{3,N} &= m_\alpha^2 \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} [\sigma^2(u(t(\delta), x_i)) - \sigma^2(u(t, x))] dx \\ &= m_\alpha^2 \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} [\sigma^2(u(t(\delta), x_i)) - \sigma^2(u(t(\delta), x))] dx \\ &\quad + m_\alpha^2 \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} [\sigma^2(u(t(\delta), x)) - \sigma^2(u(t, x))] dx := B_{3,N}^{(1)} + B_{3,N}^{(2)}. \end{aligned}$$

Regarding $B_{3,N}^{(1)}$, we have, via the Lipschitz assumption (7.4) and the bound (7.9) on the moments of the solution,

$$\begin{aligned} \mathbf{E}|B_{3,N}^{(1)}| &\leq m_\alpha^2 \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} \mathbf{E} |\sigma^2(u(t(\delta), x_i)) - \sigma^2(u(t(\delta), x))| dx \\ &\leq m_\alpha^2 \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} \left(\mathbf{E} |\sigma(u(t(\delta), x_i)) - \sigma(u(t(\delta), x))|^2 \right)^{\frac{1}{2}} \left(\mathbf{E} |\sigma(u(t(\delta), x_i)) + \sigma(u(t(\delta), x))|^2 \right)^{\frac{1}{2}} dx \\ &\leq C \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} \left(\mathbf{E} |u(t(\delta), x_i) - u(t(\delta), x)|^2 \right)^{\frac{1}{2}} dx. \end{aligned}$$

We now use the inequality (7.10) to obtain

$$\mathbf{E}|B_{3,N}^{(1)}| \leq C \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} |x_i - x|^{\frac{\alpha-1}{2}} dx \leq \left(\frac{1}{N} \right)^{\frac{\alpha-1}{2}}. \quad (7.37)$$

By the same argument, we have

$$\begin{aligned} \mathbf{E}|B_{3,N}^{(2)}| &\leq C \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} \left(\mathbf{E} |u(t(\delta), x) - u(t, x)|^2 \right)^{\frac{1}{2}} dx \\ &\leq C \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} |t(\delta) - t|^{\frac{\alpha-1}{2\alpha}} dx = C \left(\frac{1}{N} \right)^{\frac{\gamma(\alpha-1)}{2\alpha}} = C \left(\frac{1}{N} \right)^{\frac{\alpha-1}{\alpha+1}}, \end{aligned} \quad (7.38)$$

where we used (7.10) for the spatial temporal variable. Via (7.37) and (7.38), we find,

$$\mathbf{E}|B_{3,N}| \leq \mathbf{E}|B_{3,N}^{(1)}| + \mathbf{E}|B_{3,N}^{(2)}| \leq C \left(\frac{1}{N} \right)^{\frac{\alpha-1}{\alpha+1}}, \quad (7.39)$$

since $\alpha > 1$. The conclusion follows by putting together the estimates (7.35), (7.36) and (7.39). \square

If u denotes the solution to (7.5), let us introduce its averaged spatial quadratic variation $(V_{N,M}(u), N, M \geq 1)$, by the following formula

$$V_{N,M}(u) = \frac{N^{\alpha-2}}{M} \sum_{j=0}^M \sum_{i=0}^{N-1} (u(t_j, x_{i+1}) - u(t_j, x_i))^2 = \frac{1}{M} \sum_{j=0}^M V_{N,t_j}(u) \quad (7.40)$$

where $x_i, i = 0, 1, \dots, N$ are given by (7.21), $V_{N,t}$ by (7.20) and

$$t_j = \frac{j}{M}, \quad j = 0, 1, \dots, M. \quad (7.41)$$

That is, $V_{N,M}(u)$ constitutes the average of the spatial quadratic variations at different times. A similar idea has been used in the paper [60]. From Theorem 5, we can deduce the asymptotic behavior of $V_{N,M}(u)$ as $N, M \rightarrow \infty$.

Theorem 6. *Consider the sequence $(V_{N,M}(u), N, M \geq 1)$ given by (7.40). Then, as $M, N \rightarrow \infty$,*

$$V_{N,M}(u) \rightarrow m_\alpha^2 \int_0^1 \int_0^1 \sigma^2(u(t, x)) dt dx \text{ in } L^1(\Omega) \quad (7.42)$$

and for N sufficiently large,

$$\mathbf{E} \left| V_{N,M}(u) - m_\alpha^2 \int_0^1 \int_0^1 \sigma^2(u(t, x)) dt dx \right| \leq C \left[\left(\frac{1}{N} \right)^{\frac{(\alpha-1)(3-\alpha)}{2(\alpha+1)}} + \left(\frac{1}{M} \right)^{\frac{\alpha-1}{2\alpha}} \right].$$

Proof : We write, for $N, M \geq 1$,

$$\begin{aligned} & V_{N,M}(u) - m_\alpha^2 \int_0^1 \int_0^1 \sigma^2(u(t, x)) dt dx \\ = & V_{N,M}(u) - \frac{1}{M} \sum_{j=1}^M m_\alpha^2 \int_0^1 \sigma^2(u(t_j, x)) dx + \frac{1}{M} \sum_{j=1}^M m_\alpha^2 \int_0^1 \sigma^2(u(t_j, x)) dx \\ & - m_\alpha^2 \int_0^1 \int_0^1 \sigma^2(u(t, x)) dt dx = D_{1,M,N} + D_{2,M}, \end{aligned}$$

with

$$D_{1,M,N} = V_{N,M}(u) - \frac{1}{M} \sum_{j=1}^M m_\alpha^2 \int_0^1 \sigma^2(u(t_j, x)) dx$$

and

$$D_{2,M} = \frac{1}{M} \sum_{j=1}^M m_\alpha^2 \int_0^1 \sigma^2(u(t_j, x)) dx - m_\alpha^2 \int_0^1 \int_0^1 \sigma^2(u(t, x)) dt dx.$$

By relation (7.33) in Theorem 5, for any $M \geq 1$ and N large enough,

$$\begin{aligned} \mathbf{E}|D_{1,M,N}| &= \frac{1}{M} \mathbf{E} \left| \sum_{j=1}^M \left[V_{N,t_j}(u) - m_\alpha^2 \int_0^1 \sigma^2(u(t_j, x)) dx \right] \right| \\ &\leq \frac{1}{M} \sum_{j=1}^M \mathbf{E} \left| V_{N,t_j}(u) - m_\alpha^2 \int_0^1 \sigma^2(u(t_j, x)) dx \right| \leq C \left(\frac{1}{N} \right)^{\frac{(\alpha-1)(3-\alpha)}{2(\alpha+1)}}. \end{aligned} \quad (7.43)$$

On the other hand,

$$\begin{aligned} \mathbf{E}|D_{2,M}| &= m_\alpha^2 \mathbf{E} \left| \sum_{j=1}^M \int_{t_j}^{t_{j+1}} \left[\int_0^1 \sigma^2(u(t_j, x)) - \sigma^2(u(t, x)) dx \right] dt \right| \\ &\leq C \sum_{j=1}^M \int_{t_j}^{t_{j+1}} \int_0^1 (\mathbf{E}|u(t_j, x) - u(t, x)|^2)^{\frac{1}{2}} dx dt \\ &\leq C \sum_{j=1}^M \int_{t_j}^{t_{j+1}} |t_j - t|^{\frac{\alpha-1}{2\alpha}} dt \leq C \left(\frac{1}{M} \right)^{\frac{\alpha-1}{2\alpha}}. \end{aligned} \quad (7.44)$$

The conclusion is obtained by combing (7.43) and (7.44). \square

Remark 7. Let us also comment on the limit behavior in distribution of the quadratic variations (7.20) or (7.40). This behavior has been studied by several authors when the random noise in (7.5) is additive, i.e. $\sigma(x) = \sigma$ for every $x \in \mathbb{R}$. In this case, it holds that for every $t \in (0, T]$,

$$\sqrt{N} (V_{N,t} - m_\alpha^2 \sigma^2) \xrightarrow[N \rightarrow \infty]{} \mathcal{N}(0, a(\alpha)^2),$$

where $\mathcal{N}(0, a(\alpha)^2)$ denotes the centered Gaussian law with standard deviation equals to $a(\alpha)$, a strictly positive constant depending on α . This follows from e.g. Theorem 6.3 in [35] or Lemma 1 in [76]. See also [33] for the case $\alpha = 2$ or [60], [59] for the case of a bounded spatial domain. In the nonlinear case, as far as we know, the limit distribution of the sequence $V_{N,t}$ as $N \rightarrow \infty$ (or, more generally, of $V_{N,M}(u)$ as $N, M \rightarrow \infty$) remains an open problem (even for the case $\alpha = 2$). On the other hand, when the noise has a spatial correlation given by the Riesz kernel and $\alpha = 2$, a Central Limit Theorem for the power variations of the solution to (7.5) has been obtained in [28].

7.5 Drift parameter estimation

In this part, we consider the parametrized stochastic fractional heat equation given by (7.1). We recall that W is a space-time white noise characterized by its covariance (7.3), σ satisfies the Lipschitz assumption (7.4). We also assume $\theta > 0$ and $\alpha \in (1, 2)$.

We construct an estimator for the parameter θ in (7.1) based on the quadratic variation of the solution u_θ . Via the results in Section 7.4, we deduce the consistency of our estimator. First, recall that the mild solution to (7.1) is given by (7.2). In a first step, we introduce the following transform of u_θ

$$v_\theta(t, x) = u_\theta\left(\frac{t}{\theta}, x\right), \text{ so } u_\theta(t, x) = v_\theta(\theta t, x), \quad t \in [0, T], x \in \mathbb{R}. \quad (7.45)$$

Concerning the random field v_θ , we can write

$$\begin{aligned} v_\theta(t, x) &= \int_0^t \int_{\mathbb{R}} G_\alpha(t - \theta s, x - y) \sigma(u_\theta(s, y)) W(ds, dy) \\ &= \int_0^t \int_{\mathbb{R}} G_\alpha(t - s, x - y) \sigma(u_\theta\left(\frac{s}{\theta}, y\right)) W(d\left(\frac{s}{\theta}\right), dy) \\ &= \theta^{-\frac{1}{2}} \int_0^t \int_{\mathbb{R}} G_\alpha(t - s, x - y) \sigma(v_\theta(s, y)) \tilde{W}(ds, dy) \end{aligned}$$

where we denoted \tilde{W} the random field

$$\tilde{W}(t, A) = \theta^{\frac{1}{2}} W\left(\frac{t}{\theta}, A\right), \quad t \in [0, T], A \in \mathcal{B}_b(\mathbb{R}).$$

By the $\frac{1}{2}$ -selfsimilarity of the space-time white noise in time, it follows that \tilde{W} is also a space-time white noise. So, the random field v_θ satisfies the SPDE

$$v_\theta(t, x) = \int_0^t \int_{\mathbb{R}} G_\alpha(t - s, x - y) \tilde{\sigma}(v_\theta(s, y)) \tilde{W}(ds, dy) \quad (7.46)$$

with $\tilde{\sigma}(x) = \theta^{-\frac{1}{2}} \sigma(x)$ for $x \in \mathbb{R}$. Obviously, $\tilde{\sigma}$ is a Lipschitz function.

We can easily deduce the behavior of the quadratic variation of the mild solution (7.2).

Proposition 9. *Let u_θ be given by (7.2) and set, for $N \geq 1$ and $t \in (0, T]$,*

$$V_{N,t}(u_\theta) = N^{\alpha-2} \sum_{i=0}^{N-1} (u_\theta(t, x_{i+1}) - u_\theta(t, x_i))^2$$

with $x_i, i = 0, \dots, N$ given by (7.21). Then

$$V_{N,t}(u_\theta) \xrightarrow{N \rightarrow \infty} \theta^{-1} m_\alpha^2 \int_0^1 \sigma^2(u_\theta(t, x)) dx \text{ in } L^1(\Omega)$$

and for N large enough,

$$\mathbf{E} \left| V_{N,t}(u_\theta) - \theta^{-1} m_\alpha^2 \int_0^1 \sigma^2(u_\theta(t, x)) dx \right| \leq C \left(\frac{1}{N} \right)^{\frac{(\alpha-1)(3-\alpha)}{2(\alpha+1)}}. \quad (7.47)$$

Proof : From Theorem 5 and relation (7.46), for every $t \in [0, T]$,

$$\begin{aligned} & N^{\alpha-2} \sum_{i=0}^{N-1} (v_\theta(t, x_{i+1}) - v_\theta(t, x_i))^2 \\ & \xrightarrow{N \rightarrow \infty} m_\alpha^2 \int_0^1 (\tilde{\sigma}(v_\theta(t, x)))^2 dx = m_\alpha^2 \int_0^1 \left(\theta^{-\frac{1}{2}} \sigma(v_\theta(t, x)) \right)^2 dx = \frac{m_\alpha^2}{\theta} \int_0^1 \sigma^2(v_\theta(t, x)) dx. \end{aligned}$$

Consequently, due to (7.45),

$$\begin{aligned} V_{N,t}(u_\theta) &= N^{\alpha-2} \sum_{i=0}^{N-1} (v_\theta(\theta t, x_{i+1}) - v_\theta(\theta t, x_i))^2 \\ & \xrightarrow{N \rightarrow \infty} \frac{m_\alpha^2}{\theta} \int_0^1 \sigma^2(v_\theta(\theta t, x)) dx = \frac{m_\alpha^2}{\theta} \int_0^1 \sigma^2(u_\theta(t, x)) dx. \end{aligned}$$

The bound (7.47) is a direct consequence of (7.33). \square

The above result can be easily generalized to the context of the averaged spatial quadratic variation of the random field u_θ .

Corollary 2. For $N, M \geq 1$, let us define

$$V_{N,M}(u_\theta) = \frac{N^{\alpha-2}}{M} \sum_{j=0}^M \sum_{i=0}^{M-1} (u_\theta(t_j, x_{i+1}) - u_\theta(t_j, x_i))^2,$$

where u_θ is defined by (7.2) and $(t_j, j = 0, \dots, M)$, $(x_i, i = 0, \dots, N)$ are the partition given by (7.41) and (7.21), respectively. Then

$$V_{N,M}(u_\theta) \xrightarrow{N, M \rightarrow \infty} \theta^{-1} m_\alpha^2 \int_0^1 \int_0^1 \sigma^2(u_\theta(t, x)) dt dx \text{ in } L^1(\Omega).$$

Moreover, for N large enough,

$$\mathbf{E} \left| V_{N,M}(u_\theta) - \theta^{-1} m_\alpha^2 \int_0^1 \int_0^1 \sigma^2(u_\theta(t, x)) dt dx \right| \leq C \left[\left(\frac{1}{N} \right)^{\frac{(\alpha-1)(3-\alpha)}{2(\alpha+1)}} + \left(\frac{1}{M} \right)^{\frac{\alpha-1}{2\alpha}} \right].$$

Proof : The proof is an immediate consequence of Theorem 6 and of the proof of Proposition 9. \square

Define the estimator

$$\widehat{\theta}_{N,M} = \frac{m_\alpha^2 \frac{1}{MN} \sum_{j=0}^M \sum_{i=0}^{N-1} \sigma^2(u_\theta(t_j, x_i))}{V_{N,t}(u_\theta)}, \quad N \geq 1, t \in [0, T]. \quad (7.48)$$

Notice that $\widehat{\theta}_{N,M}$ can be calculated directly from the data $(u_\theta(t_j, x_i), i = 0, \dots, N, j = 0, 1, \dots, M)$. As a consequence of Corollary 2, we can show that the estimator (7.49) is a consistent estimator the drift parameter θ .

Proposition 10. Let $(\widehat{\theta}_{N,M}, N, M \geq 1)$ be defined by (7.48). Then

$$\widehat{\theta}_{N,M} \xrightarrow{N, M \rightarrow \infty} \theta \text{ in probability.}$$

Proof : By Corollary 2, we have

$$\frac{m_\alpha^2 \int_0^1 \int_0^1 \sigma^2(u_\theta(t, x)) dx dt}{V_{N,M}(u_\theta)} \xrightarrow{N, M \rightarrow \infty} \theta.$$

It remains to observe that

$$\begin{aligned} \widehat{\theta}_{N,M} &= \frac{m_\alpha^2 \left(\frac{1}{MN} \sum_{j=0}^M \sum_{i=0}^{N-1} \sigma^2(u_\theta(t_j, x_i)) - \int_0^1 \int_0^1 \sigma^2(u_\theta(t, x)) dx dt \right)}{V_{N,M}(u_\theta)} \\ &\quad + \frac{m_\alpha^2 \int_0^1 \int_0^1 \sigma^2(u_\theta(t, x)) dx dt}{V_{N,M}(u_\theta)} \end{aligned}$$

and the first summand above converges almost surely to zero as $N \rightarrow \infty$ by a Riemann sum argument. \square

Remark 8. If we dispose on the observation of the solution (7.2) at a single time $t \in (0, T]$, we can use the estimator

$$\widehat{\theta}_N = \frac{m_\alpha^2 \frac{1}{N} \sum_{i=0}^{N-1} \sigma^2(u_\theta(t, x_i))}{V_{N,t}(u_\theta)}, \quad N \geq 1, t \in [0, T]. \quad (7.49)$$

Via Proposition 9, we can similarly show that the estimator (7.49) is a consistent estimator the drift parameter θ .

7.6 Appendix : Proofs of the technical lemmas

Proof of Lemma 7 : We have, by Parseval's identity (7.11),

$$\begin{aligned} I(s, t, \delta) &= (2\pi)^{-1} \int_0^s da \int_{\mathbb{R}} d\xi |\mathcal{F}G_\alpha(t-a, x+\delta-\cdot)(\xi) - \mathcal{F}G_\alpha(t-a, x-\cdot)(\xi)|^2 \\ &= (2\pi)^{-1} \int_0^s da \int_{\mathbb{R}} d\xi \left| e^{i\xi(x+\delta)} - e^{i\xi x} \right|^2 e^{-2(t-a)|\xi|^\alpha} \end{aligned}$$

since, from (7.7),

$$\mathcal{F}G_\alpha(t-a, x-\cdot)(\xi) = e^{i\xi x} e^{-(t-a)|\xi|^\alpha}.$$

Since $\left| e^{i\xi(x+\delta)} - e^{i\xi x} \right|^2 = 2 - 2\cos(\xi\delta)$, we get, by calculating the integral da ,

$$\begin{aligned} I(s, t, \delta) &= 2(2\pi)^{-1} \int_0^s da \int_{\mathbb{R}} d\xi (1 - \cos(\delta\xi)) e^{-2(t-a)|\xi|^\alpha} \\ &= 2(2\pi)^{-1} \int_{\mathbb{R}} d\xi \frac{(1 - \cos(\delta\xi))}{|\xi|^\alpha} \left(e^{-2(t-s)|\xi|^\alpha} - e^{-2t|\xi|^\alpha} \right). \end{aligned}$$

By using the bound $|1 - \cos(x)| \leq x^2$ and the change of variables $(t-s)^{\frac{1}{\alpha}} = \tilde{\xi}$, we can write

$$\begin{aligned} I(s, t, \delta) &\leq 2(2\pi)^{-1} \int_{\mathbb{R}} d\xi \frac{(1 - \cos(\delta\xi))}{|\xi|^\alpha} e^{-2(t-s)|\xi|^\alpha} \\ &\leq 2(2\pi)^{-1} \delta^2 \int_{\mathbb{R}} d\xi |\xi|^{2-\alpha} e^{-2(t-s)|\xi|^\alpha} = 4(2\pi)^{-1} \delta^2 \int_0^\infty d\xi \xi^{2-\alpha} e^{-2(t-s)\xi^\alpha} \\ &= 4(2\pi)^{-1} \delta^2 (t-s)^{-\frac{1}{\alpha} - \frac{2-\alpha}{\alpha}} \int_0^\infty d\xi \xi^{2-\alpha} e^{-\xi^\alpha} \leq C \delta^2 (t-s)^{-\frac{3-\alpha}{\alpha}}. \end{aligned}$$

□

Proof of Lemma 8 : Due to Parseval's relation (7.11),

$$J(s, t) = (2\pi)^{-1} \int_s^t da \int_{\mathbb{R}} d\xi e^{-2(t-a)|\xi|^\alpha} = 2(2\pi)^{-1} \int_s^t da \int_0^\infty d\xi e^{-(t-a)\xi^\alpha}.$$

We perform the change of variables $(t-s)^{\frac{1}{\alpha}} \xi = \tilde{\xi}$ to get

$$\begin{aligned} J(s, t) &= C \int_s^t da (t-s)^{-\frac{1}{\alpha}} \int_{\mathbb{R}} e^{-\tilde{\xi}^\alpha} d\tilde{\xi} \\ &= C (t-s)^{1-\frac{1}{\alpha}}. \end{aligned}$$

□

Partie IV

**Etude du comportement
asymptotique des matrices et
tenseurs de Wishart par la méthode
de Stein-Malliavin**

Random matrices and the stochastic wave equation

This article is submitted.

Joint work with Ciprian A. Tudor.

Abstract. We consider an initial random matrix whose entries are the temporal or spatial increments of the solution to the stochastic wave equation driven by an additive space-time white noise. We associate to it a Wishart matrix, which contains the empirical covariances of the elements of the initial matrix. We study the almost sure limit of this Wishart matrix and also its fluctuations, i.e. the limit in distribution of the renormalized Wishart matrix in the so-called high-dimensional regime. Via the techniques of the Stein-Malliavin calculus, we prove that the renormalized Wishart matrix satisfies a Central Limit Theorem (more precisely, it converges in law to a Gaussian Wigner matrix) and we analyze the rate of convergence under the Wasserstein distance for this limit theorem. To this end, we explore in details the correlation of the temporal and spatial increments of the solution to the stochastic wave equations and we derive some refinements of the Stein-Malliavin techniques.

8.1 Introduction

The class of Wishart matrices represents a special class of random matrices with numerous applications in statistics, multivariate analysis or information theory. Given an initial random matrix $\mathcal{X}_{n,d}$ with n rows and d columns, its associated random matrix is usually defined as $\mathcal{W}_{n,d} = \mathcal{X}_{n,d}\mathcal{X}_{n,d}^T$, where \mathcal{X}^T denotes the transpose of the matrix \mathcal{X} . This class of random matrices has been introduced in [127] and a rather complete description of their properties and their applications can be found in e.g. [17], [70], [105]. The asymptotic behavior of the Wishart matrix when both dimensions n and d tend to infinity has been the object of several works in the last decades. The limit behavior of the associated Wishart matrix clearly depends on the probability distribution of the entries of the starting matrix $\mathcal{X}_{n,d}$. Many situations have been analyzed in the literature. The case when the initial matrix $\mathcal{X}_{n,d}$ has independent or partially correlated

Gaussian entries has been treated in e.g. [20], [21], [18], [68], [106], [97], [98], the case of entries with log-concave distribution has been considered in [21] or [90] (the second reference assuming only the column independence of the entries), while other cases of non-Gaussian entries can be found in e.g. [19], [45], [46] or [49]. In many situations, the limit in distribution of the Wishart matrix is a Gaussian Wigner matrix, but a non-Gaussian limit matrix may appear when the correlation of the entries of the initial matrix is strong enough (examples can be found in [19], [45], [97], [98]).

Our goal is to analyze a new situation when the data collected in our initial matrix is related to solutions to stochastic partial differential equations. More precisely, we study the limit behavior of the Wishart matrix constructed from an initial matrix $\mathcal{X}_{n,d} = (X_{i,j})_{1 \leq i \leq n, 1 \leq j \leq d}$ whose entries $X_{i,j}$ are given by the increments (in time or in space) of the solution to the stochastic wave equation driven by the space-time white noise. The wave equation constitutes a model for the propagation of waves (for example, electromagnetic, radio or wireless waves). Our model measures the long-term behavior of the correlation of a system of multiple waves (or strings) perturbed by a random force. The solution $u(t,x)$ is interpreted at the position at time t of the point x on the wave (which is assumed initially in a horizontal position and it is displaced by an external force). When the spatial increments of the solution are included in the initial matrix $\mathcal{X}_{n,d}$, this means that the i th ($1 \leq i \leq d$) row of this starting matrix contains the information about the displacement on the i th wave between two consecutive spatial points, at fixed time. The Wishart matrix associated to $\mathcal{X}_{n,d}$ includes in its structure the empirical covariance of the displacements of the d waves and its asymptotic behavior indicates how the system evolve when the number of spatial points and the number of rows become bigger and bigger. More exactly, in the case of the spatial increments, we will assume that the entries $X_{i,j}, 1 \leq i \leq n, 1 \leq j \leq d$ are given by

$$X_{i,j} = u^{(i)}(t, j+1) - u^{(i)}(t, j)$$

where $u^{(i)}$ satisfies

$$\frac{\partial^2 u^{(i)}}{\partial t^2} = \Delta u^{(i)} + \dot{W}^{(i)} \tag{8.1}$$

with vanishing initial condition, by assuming that $W^{(i)}, i = 1, \dots, n$ are independent space-time white noises. That means that the elements of the matrix $\mathcal{X}_{n,d}$ situated on different rows are independent while those on the same row are correlated and the correlation is given by the spatial increments of the solution to the stochastic wave equation. Then we will consider the situation when the entries of the starting matrix are the temporal increments over the set of positive integers of the solution to (8.1). In this situation the analysis is somehow simpler, since as we will see in the next section, these temporal increments are independent and thus all the entries of the initial matrix are independent (but not identically distributed).

We will start with the study of the correlation structure, both in time and in space, for the solution to the wave equation and we will deduce the limit behavior of the spatial and temporal quadratic variation, over the set of positive integers, of the solution. This is helpful because the diagonal elements of the Wishart matrix $\mathcal{W}_{n,d}$ are precisely given by these quadratic variations. Then we will use the techniques of the Stein-Malliavin calculus in order to quantify the distance between the probability distribution of the vector associated to the (renormalized) Wishart matrix and its limit in distribution (a Gaussian vector associated to a Wigner matrix with Gaussian entries, see (8.43) for the spatial case and (8.37) for the temporal case). Actually, in order to apply the tools of the Malliavin calculus, we state and prove a criterion that gives a bound for the Wasserstein distance, between a random vector whose components are Skorohod integrals with respect to different Gaussian processes and a centered Gaussian vector with a

given covariance matrix. This approximation is considered in the so-called "high-dimensional regime", that is, when both dimensions n and d are large enough. In both situations, we find a classical bound, that is, the Wasserstein distance between the renormalized Wishart matrix and the Gaussian Wigner matrix is of order less than $\sqrt{\frac{n^3}{d}}$. As mentioned above, the same result has been obtained, among others in [97], [98] by assuming that the starting matrix is a Gaussian random matrix with (partial or total) correlation. More precisely, in these references it is assumed that the elements on the same row of the initial matrix constitutes a stationary sequence of standard normal random variables whose correlation function verifies a sommability condition (actually, the main hypothesis is that this correlation function belongs to the space $\ell^{\frac{4}{3}}(\mathbb{Z})$). In our case, when the entries are the temporal increments of the solution to the wave equation, the stationarity assumption in [97], [98] is not satisfied. When the spatial increments of the solution are considered, the correlation of the elements of our initial matrix satisfy the main assumption in the above references. We will then use the results in [97] and [98] to get our main result. Nevertheless, these results cannot be directly applied since we normalize the Wishart matrix in a different way, by subtracting a diagonal matrix (which is not the case in [97], [98]).

We organized our work as follows. In Section 2 we recall some facts concerning the stochastic wave equation with space-time white noise and its mild solution and we will analyze the limit behavior of the spatial and temporal quadratic variation of the solution. In particular, we make a detailed study of the correlation of the increments of this solution, which can be also helpful for other purposes. In Section 3 we use the tools of the Stein-Malliavin calculus in order to show that the Wishart matrix constructed from a starting matrix whose entries are the increments of the solution to the stochastic wave equation converges in distribution to a Gaussian Wigner matrix. We also estimate the Wasserstein distance between the Wishart matrix and its limiting matrix. Section 4 is the Appendix which contains the elements from Malliavin calculus needed in our work.

8.2 The stochastic wave equation and its quadratic variation

In this part we first recall some known facts concerning the solution to the stochastic wave equation with space-time white noise. Then we will analyze the correlation structure of the increments of the solution and we will deduce the asymptotic behavior of the quadratic variations defined via the increments over the set of positive integers. We discuss separately the spatial and temporal cases.

We will consider the following stochastic partial differential equation

$$\begin{cases} \frac{\partial^2 u}{\partial t^2}(t, x) = \Delta u(t, x) + \dot{W}(t, x), & t > 0, x \in \mathbb{R} \\ u(0, x) = 0, & x \in \mathbb{R} \\ \frac{\partial u}{\partial t}(0, x) = 0, & x \in \mathbb{R}. \end{cases} \quad (8.2)$$

We denoted by Δ the Laplacian operator on \mathbb{R} and $W = (W(t, A); t \geq 0, A \in \mathcal{B}_b(\mathbb{R}))$ is the space-time white noise, it is a real valued centered Gaussian field, over a given complete filtered

probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ with covariance function

$$\mathbf{E}\left(W(t, A)W(s, B)\right) = (t \wedge s)\lambda(A \cap B), \text{ for every } A, B \in \mathcal{B}_b(\mathbb{R}) \quad (8.3)$$

where λ is the Lebesgue measure and $\mathcal{B}_b(\mathbb{R})$ is the set of the Borel subsets of \mathbb{R} with finite Lebesgue measure.

Let G_1 be the fundamental solution of the homogeneous wave equation $\frac{\partial^2}{\partial t^2}u - \Delta u = 0$. It is known that for $t > 0$ and $x \in \mathbb{R}$, the kernel G_1 is given by

$$G_1(t, x) = \frac{1}{2} 1_{\{|x| < t\}}. \quad (8.4)$$

The mild solution to (8.2) is a square-integrable process $u = (u(t, x); t \geq 0, x \in \mathbb{R})$ which is defined by

$$u(t, x) = \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y)W(ds, dy), \quad t \geq 0, x \in \mathbb{R}. \quad (8.5)$$

The above Wiener integral is well-defined, i.e $\mathbf{E}u(t, x)^2 < \infty$ for every $t \geq 0, x \in \mathbb{R}$.

8.2.1 The spatial quadratic variation

In this paragraph we deal with the Gaussian stochastic process $(u(t, x), x \in \mathbb{R})$ with $t > 0$ fixed. We recall (see e.g. [74]) that the spatial covariance of the mild solution (8.5) is given by

$$\mathbf{E}u(t, x)u(t, y) = \frac{1}{4} \left(\frac{|x-y|}{2} - t \right)^2 1_{\{|x-y| < 2t\}}, \quad t \geq 0, x, y \in \mathbb{R}. \quad (8.6)$$

In particular, for $t > 0$ fixed, the Gaussian process $(u(t, x), x \in \mathbb{R})$ is stationary and for every $t > 0, x \in \mathbb{R}$,

$$\mathbf{E}u(t, x)^2 = \frac{1}{4}t^2.$$

Let us start with the following lemma concerning the spatial correlation structure of the mild solution (8.5).

Lemma 11. *Consider the random field $(u(t, x), t \geq 0, x \in \mathbb{R})$ given by (8.5). Let $t > 0$ be fixed. Then*

1. For every integer $j \geq 0$,

$$\mathbf{E}(u(t, j+1) - u(t, j))^2 = \begin{cases} \frac{t}{2} - \frac{1}{8}, & \text{if } 2t > 1 \\ \frac{1}{2}t^2, & \text{if } 2t \leq 1. \end{cases}$$

2. For any integers $i, j \geq 0$ with $i \neq j$, we have

$$\begin{aligned} & \mathbf{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j)) \\ &= \begin{cases} 0, & \text{if } 2t \leq i-j-1 \\ -\frac{1}{4} \left(\frac{i-j-1}{2} - t \right)^2, & \text{if } i-j-1 < 2t \leq i-j \\ \frac{1}{4} \left(\frac{i-j+1}{2} - t \right)^2 - \frac{1}{8}, & \text{if } i-j < 2t \leq i-j+1 \\ -\frac{1}{8} & \text{if } 2t > i-j+1. \end{cases} \end{aligned}$$

Proof : Point 1. is an easy consequence of the covariance formula (8.6). To show 2., we notice that by (8.6), for $i > j$,

$$\begin{aligned} & \mathbf{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j)) \\ = & 2 \times \frac{1}{4} \left(\frac{i-j}{2} - t \right)^2 \mathbf{1}_{\{i-j < 2t\}} - \frac{1}{4} \left(\frac{i-j+1}{2} - t \right)^2 \mathbf{1}_{\{i-j+1 < 2t\}} \\ & - \frac{1}{4} \left(\frac{i-j-1}{2} - t \right)^2 \mathbf{1}_{\{i-j-1 < 2t\}}. \end{aligned} \quad (8.7)$$

We separate our calculation upon several cases.

Case 1 : $2t \leq i-j-1$. In this case all the three summands in the right-hand side of (8.7) vanish and thus

$$\mathbf{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j)) = 0.$$

Case 2 : $i-j-1 < 2t \leq i-j$. In this situation only the last summand in (8.7) is not zero and we find

$$\mathbf{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j)) = -\frac{1}{4} \left(\frac{i-j-1}{2} - t \right)^2.$$

Case 3 : $i-j < 2t \leq i-j+1$. We will have

$$\begin{aligned} & \mathbf{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j)) \\ = & 2 \times \frac{1}{4} \left(\frac{i-j}{2} - t \right)^2 - \frac{1}{4} \left[\left(\frac{i-j}{2} - t \right)^2 - \left(\frac{i-j}{2} - t \right) + \frac{1}{4} \right] \\ = & \frac{1}{4} \left(\frac{i-j}{2} - t \right)^2 + \frac{1}{4} \left(\frac{i-j}{2} - t \right) - \frac{1}{16} \\ = & \frac{1}{4} \left(\frac{i-j+1}{2} - t \right)^2 - \frac{1}{8}. \end{aligned}$$

Case 4 : $2t > i-j+1$. Now, all the summand in the right-hand side of (8.7) contribute, i.e.

$$\begin{aligned} & \mathbf{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j)) \\ = & 2 \times \frac{1}{4} \left(\frac{i-j}{2} - t \right)^2 - \frac{1}{4} \left(\frac{i-j+1}{2} - t \right)^2 \\ & - \frac{1}{4} \left(\frac{i-j-1}{2} - t \right)^2 = -\frac{1}{8}. \end{aligned}$$

□

Let us use the following notation, for $v > 0$

$$\begin{aligned}
 f_{1,t}(v) &= -\frac{1}{4} \left(\frac{v-1}{2} - t \right)^2 \mathbf{1}_{\{v-1 < 2t \leq v\}} \\
 f_{2,t}(v) &= \frac{1}{4} \left[\left(\frac{v}{2} - t \right)^2 + \left(\frac{v}{2} - t \right) - \frac{1}{4} \right] \mathbf{1}_{\{v < 2t \leq v+1\}} \\
 f_{3,t}(v) &= -\frac{1}{8} \mathbf{1}_{\{2t > v+1\}}.
 \end{aligned} \tag{8.8}$$

Then we can write for any integers $i, j \geq 0$ with $i > j$, due to Lemma 11,

$$\mathbf{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j)) = f_{1,t}(i-j) + f_{2,t}(i-j) + f_{3,t}(i-j). \tag{8.9}$$

Notice that the function

$$f_t(v) = f_{1,t}(v) + f_{2,t}(v) + f_{3,t}(v) \tag{8.10}$$

is continuous on $[0, \infty)$.

Let us define, for $t > 0$ fixed and for every $N \geq 1$

$$S_{N,t} = \frac{1}{N} \sum_{j=0}^{N-1} (u(t, j+1) - u(t, j))^2. \tag{8.11}$$

The sequence $(S_{N,t}, N \geq 1)$ will be called the *spatial quadratic variation sequence* of u . By Lemma 11, for every $N \geq 1$ and $t > 0$,

$$\mathbf{E}S_{N,t} = e_t$$

where we denoted

$$e_t = \left(\frac{t}{2} - \frac{1}{8} \right) \mathbf{1}_{\{2t > 1\}} + \frac{1}{2} t^2 \mathbf{1}_{\{2t \leq 1\}}. \tag{8.12}$$

It is useful to give the chaos expression of the quadratic variation. For every $N \geq 1$, $t > 0$, the random variable $S_{N,t}$ can be written as a multiple integral of order two with respect to the space-time white noise W . Actually, for $i = 1, \dots, N$,

$$u(t, i+1) - u(t, i) = I_1(g_{t,i})$$

with

$$g_{t,i}(s, y) = (G_1(t-s, i+1-y) - G_1(t-s, i-y)) \mathbf{1}_{[0,t]}(s), \quad (s, y) \in \mathbb{R}_+ \times \mathbb{R} \tag{8.13}$$

where the fundamental solution G_1 is given by (8.4) and I_q denotes the multiple integral of order $q \geq 1$ with respect to the Gaussian process W . As a consequence of the product formula (8.46),

$$(u(t, i+1) - u(t, i))^2 = I_2(g_{t,i}^{\otimes 2}) + \mathbf{E}(u(t, i+1) - u(t, i))^2 = I_2(g_{t,i}^{\otimes 2}) + e_t$$

with e_t given by (8.12). Therefore,

$$S_{N,t} = \frac{1}{N} \sum_{j=0}^{N-1} \left[I_2(g_{t,i}^{\otimes 2}) + e_t \right] = \frac{1}{N} \sum_{j=0}^{N-1} I_2(g_{t,i}^{\otimes 2}) + e_t. \tag{8.14}$$

Let us analyze the asymptotic behavior of the sequence $S_{N,t}$ as $N \rightarrow \infty$. Throughout this work, we denote by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ the scalar product and the norm in the Hilbert space $L^2(\mathbb{R}_+ \times \mathbb{R})$ (when we used scalar products and norms in other spaces, this will be explicitly showed by our notation).

Proposition 11. *Let $t > 0$ be fixed. Then the sequence $(S_{N,t}, N \geq 1)$ given by (8.11) converges to e_t as $N \rightarrow \infty$ in $L^2(\Omega)$ and almost surely.*

Proof : We show first the convergence in $L^2(\Omega)$. By (8.14), we have, with $g_{t,i}$ given by (8.13),

$$S_{N,t} - e_t = \frac{1}{N} \sum_{j=0}^{N-1} I_2(g_{t,i}^{\otimes 2})$$

and thus, by the isometry formula (8.45),

$$\begin{aligned} \mathbf{E}(S_{N,t} - e_t)^2 &= \frac{1}{N^2} \mathbf{E} \left(\sum_{j=0}^{N-1} I_2(g_{t,i}^{\otimes 2}) \right)^2 = \frac{2}{N^2} \sum_{i,j=0}^{N-1} \langle g_{t,i}, g_{t,j} \rangle^2 \\ &= \frac{2}{N^2} \sum_{i,j=0}^{N-1} (\mathbf{E}(u(t,i+1) - u(t,i))(u(t,j+1) - u(t,j)))^2 \\ &= \frac{2}{N^2} \sum_{j=0}^{N-1} (\mathbf{E}(u(t,j+1) - u(t,j))^2)^2 \\ &\quad + \frac{2}{N^2} \sum_{i,j=0; i \neq j}^{N-1} (\mathbf{E}(u(t,i+1) - u(t,i))(u(t,j+1) - u(t,j)))^2 \\ &= e_t^2 \frac{2}{N} + \frac{4}{N^2} \sum_{i,j=0; i > j}^{N-1} [f_{1,t}(i-j) + f_{2,t}(i-j) + f_{3,t}(i-j)]^2 \end{aligned}$$

where we used the notation (8.8). Now

$$\begin{aligned} &\frac{4}{N^2} \sum_{i,j=0; i > j}^{N-1} (f_{1,t}(i-j) + f_{2,t}(i-j) + f_{3,t}(i-j))^2 \\ &= \frac{4}{N^2} \sum_{k=1}^{N-1} (N-k) (f_{1,t}(k) + f_{2,t}(k) + f_{3,t}(k))^2 \\ &= \frac{4}{N} \sum_{k=1}^N (f_{1,t}(k) + f_{2,t}(k) + f_{3,t}(k))^2 - \frac{4}{N^2} \sum_{k=1}^N k (f_{1,t}(k) + f_{2,t}(k) + f_{3,t}(k))^2. \end{aligned}$$

Consequently,

$$\mathbf{E}(S_{N,t} - e_t)^2 = K_{1,t} \frac{1}{N} + K_{2,t} \frac{1}{N^2} \quad (8.15)$$

where

$$K_{1,t} = 2e_t^2 + 4 \sum_{k=1}^N (f_{1,t}(k) + f_{2,t}(k) + f_{3,t}(k))^2 \quad \text{and} \quad K_{2,t} = 4 \sum_{k=1}^N k (f_{1,t}(k) + f_{2,t}(k) + f_{3,t}(k))^2. \quad (8.16)$$

Actually, the quantities $K_{1,t}, K_{2,t}$ are constants which do not depend on N . Indeed, the sums $\sum_{k=1}^N f_i(k)$ and $\sum_{k=1}^N k f_i(k)$ for $i = 1, 2, 3$ contains a finite number of non-zero terms. This is due to the definition of the functions $f_{i,t}, i = 1, 2, 3$, see (8.8). Therefore $S_{N,t}$ converges to e_t in $L^2(\Omega)$ as $N \rightarrow \infty$.

The almost sure convergence follows by using Borel-Cantelli lemma. Indeed, for $0 < \gamma < \frac{1}{2}$ and for $p \geq 1$ large enough, we have

$$\begin{aligned} & \sum_{N \geq 1} P(|S_{N,t} - e_t| \geq N^{-\gamma}) \leq \sum_{N \geq 1} N^{\gamma p} \mathbf{E}|S_{N,t} - e_t|^p \\ & \leq \sum_{N \geq 1} N^{\gamma p} (\mathbf{E}|S_{N,t} - e_t|^2)^{\frac{p}{2}} \leq C_p \sum_{N \geq 1} N^{p(\gamma - \frac{1}{2})} < \infty \end{aligned}$$

where we used the estimate (8.15) and (8.16) to get $C_p := K_{1,t}^{\frac{p}{2}}$ and the hyper-contractivity property (8.48) of multiple stochastic integrals. □

Now, we define for $N \geq 1$ and $t > 0$, the renormalized quadratic variation in time of the mild solution (8.5)

$$V_{N,t} = \sqrt{N}(S_{N,t} - e_t). \quad (8.17)$$

By Proposition 11,

$$\mathbf{E}V_{N,t}^2 \xrightarrow{N \rightarrow \infty} K_{1,t}$$

with $K_{1,t}$ from (8.16). By a standard procedure based on Malliavin calculus, we can show that the above sequence $(V_{N,t}, N \geq 1)$ satisfies a Central Limit Theorem and we can estimate the distance between its probability distribution and the Gaussian limit.

By $d(\cdot, \cdot)$ we denote one of the following metrics : Kolmogorov, total variation, Wasserstein or Mourier-Fortet (see [93], Appendix C for their definition and properties). We recall the following well-known criterion in Stein-Malliavin calculus which characterizes the normal convergence of a sequence of multiple stochastic integrals. Below $(W(h), h \in \mathcal{H})$ is an isonormal process and $I_q, q \geq 1$ stands for the multiple integral of order q with respect to this isonormal process (see the Appendix).

Theorem 7. Fix $q \geq 1$. Assume that $(G_N)_{N \geq 1} = (I_q(g_N))_{N \geq 1}$ with $g_N \in \mathcal{H}^{\odot q}$, a sequence of random variables belonging to the q th Wiener chaos such that :

$$\mathbf{E}(G_N^2) \xrightarrow{N \rightarrow \infty} \sigma^2.$$

Then, G_N converges in law to $Z \sim \mathcal{N}(0, \sigma^2)$ if and only if

$$\|DG_N\|_{\mathcal{H}}^2 \xrightarrow{N \rightarrow \infty} q\sigma^2.$$

Moreover,

$$d(G_N, \mathcal{N}(0, \sigma^2)) \leq C \left(\sqrt{\mathbf{Var}(\|DG_N\|_{\mathcal{H}}^2)} + \left| \mathbf{E}(\|DG_N\|_{\mathcal{H}}^2) - q\sigma^2 \right| \right). \quad (8.18)$$

For every $N \geq 1$ and $t \geq 0$, the random variable $V_{N,t}$ belongs to the second Wiener chaos. More precisely, it can be written as, due to (8.14),

$$V_{N,t} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} I_2(g_{t,i}^{\otimes 2})$$

with $g_{t,i}$ given by (8.13). Therefore, we can apply Theorem 7 to establish a bound on the speed of convergence of $V_{N,t}$ to its limit. Before this, let us state and prove a technical lemma which will be needed several times in the sequel.

Lemma 12. *Let $t > 0$ be fixed and let $g_{t,i}, i = 0, \dots, N-1$ be given by (8.13). Then for every $N \geq 1$,*

$$\sum_{i,j,k,l=1}^N \langle g_{t,i}, g_{t,j} \rangle \langle g_{t,j}, g_{t,k} \rangle \langle g_{t,k}, g_{t,l} \rangle \langle g_{t,l}, g_{t,i} \rangle \leq C_t N$$

where $C_t > 0$ is a constant depending on t but independent on N .

Proof : We have, by (8.9), for $i \neq j$,

$$\langle g_{t,i}, g_{t,j} \rangle = f_t(|i-j|) 1_{\{|i-j| < 2t+1\}}$$

with

$$f_t(v) = f_{1,t}(v) + f_{2,t}(v) + f_{3,t}(v) \text{ for any integer } v > 0.$$

Notice that, by (8.8),

$$|f_t(|i-j|)| 1_{\{|i-j| < 2t+1\}} \leq C_t 1_{\{|i-j| < 2t+1\}}$$

for every $i, j = 0, \dots, N-1, i \neq j$ with $C_t > 0$. Also, for $i = 0, \dots, N-1$, by using Lemma 1 point 1.

$$\langle g_{t,i}, g_{t,i} \rangle = e_t.$$

Therefore we can write, for every $i, j = 0, 1, \dots, N-1$,

$$|\langle g_{t,i}, g_{t,j} \rangle| \leq C_t 1_{\{|i-j| < 2t+1\}}.$$

By using the change of indices $\tilde{i} = i - j, \tilde{j} = j - k, \tilde{k} = k - l$, we have

$$\begin{aligned} & \sum_{i,j,k,l=1}^N \langle g_{t,i}, g_{t,j} \rangle \langle g_{t,j}, g_{t,k} \rangle \langle g_{t,k}, g_{t,l} \rangle \langle g_{t,l}, g_{t,i} \rangle \\ & \leq C_t \sum_{i,j,k,l=-N}^N 1_{\{|i| \leq 2t+1\}} 1_{\{|j| \leq 2t+1\}} 1_{\{|k| < 2t+1\}} \leq C_t N \sum_{i,j,k=0}^{N-1} 1_{\{|i| \leq 2t+1\}} 1_{\{|j| \leq 2t+1\}} 1_{\{|k| < 2t+1\}} \leq C_t N. \end{aligned}$$

We used the same argument from the proof of Proposition 11 :

the sum $\sum_{i,j,k=0}^{N-1} 1_{\{|i| \leq 2t+1\}} 1_{\{|j| \leq 2t+1\}} 1_{\{|k| < 2t+1\}}$ is bounded by a constant depending on t but not on N , since it contains a finite number of terms (actually at most $(2[t] + 2)^3$ terms) and each term is less than 1. \square

The next result constitutes a Berry-Esséen theorem for the renormalized quadratic variation.

Theorem 8. *Let $t > 0$ be fixed and consider the sequence $(V_{N,t}, N \geq 1)$ given by (8.17). Then as $N \rightarrow \infty$, the sequence $V_{N,t}$ converges in distribution to the Gaussian law $\mathcal{N}(0, K_{1,t})$. Moreover for N large enough,*

$$d(V_{N,t}, \mathcal{N}(0, K_{1,t})) \leq C \frac{1}{\sqrt{N}}.$$

Proof : By (8.18) in Theorem 7,

$$d(V_{N,t}, \mathcal{N}(0, K_{1,t})) \leq C \left[\sqrt{\mathbf{E}(\|DV_{N,t}\|^2 - \mathbf{E}\|DV_{N,t}\|^2)^2} + \sqrt{\mathbf{E}\|DV_{N,t}\|^2 - 2K_{1,t}} \right]. \quad (8.19)$$

Now, we have,

$$DV_{N,t} = \frac{2}{\sqrt{N}} \sum_{i=0}^{N-1} I_1(g_{t,i}) g_{t,i}$$

and

$$\begin{aligned} \|DV_{N,t}\|^2 &= \frac{4}{N} \sum_{i,j=0}^{N-1} I_1(g_{t,i}) I_1(g_{t,j}) \langle g_{t,i}, g_{t,j} \rangle \\ &= \frac{4}{N} \sum_{i,j=0}^{N-1} I_2(g_{t,i} \otimes g_{t,j}) \langle g_{t,i}, g_{t,j} \rangle + \mathbf{E}\|DV_{N,t}\|^2. \end{aligned}$$

Consequently, since $\mathbf{E}\|DV_{N,t}\|^2 = 2\mathbf{E}V_{N,t}^2$ (this is true because $V_{N,t}$ is an element of the second Wiener chaos),

$$\begin{aligned} \mathbf{E}(\|DV_{N,t}\|^2 - \mathbf{E}\|DV_{N,t}\|^2)^2 &= \frac{16}{N^2} \mathbf{E} \left(\sum_{i,j=0}^{N-1} I_2(g_{t,i} \otimes g_{t,j}) \langle g_{t,i}, g_{t,j} \rangle \right)^2 \\ &= C \frac{1}{N^2} \sum_{i,j,k,l=0}^{N-1} \langle g_{t,i}, g_{t,j} \rangle \langle g_{t,j}, g_{t,k} \rangle \langle g_{t,k}, g_{t,l} \rangle \langle g_{t,l}, g_{t,i} \rangle \\ &\leq C_t \frac{1}{N} \end{aligned} \quad (8.20)$$

where the last bound follows from Lemma 12. By (8.15) (with $K_{1,t}, K_{2,t}$ defined by (8.16)), we have :

$$\mathbf{E}V_{N,t}^2 = K_{1,t} + K_{2,t} \frac{1}{N}.$$

Consequently,

$$\mathbf{E}\|DV_{N,t}\|^2 - 2K_{1,t} = 2(\mathbf{E}V_{N,t}^2 - K_{1,t}) = 2K_{2,t} \frac{1}{N}. \quad (8.21)$$

By plugging the estimates (8.20) and (8.21) into (8.19), we obtain the conclusion. \square

8.2.2 Temporal quadratic variation

Now, we will analyze the Gaussian process $(u(t, x), t \geq 0)$ where u is given by (8.5) and the spatial variable $x \in \mathbb{R}$ is fixed. Notice that this process is a centered Gaussian process and we recall the following result from [8] (see Lemma 2 in this reference).

Lemma 13. *Let $(u(t, x), t \geq 0, x \in \mathbb{R})$ be given by (8.5) and let $x \in \mathbb{R}$ be fixed. Then we have*

1. For every $0 \leq s \leq t$,

$$\mathbf{E}(u(t,x) - u(s,x))^2 = \frac{1}{4}(t^2 - s^2)$$

and

$$\mathbf{E}u(t,x)u(s,x) = \frac{1}{4}(t \wedge s)^2. \quad (8.22)$$

2. For every $0 \leq a < b \leq s < t$, we have

$$\mathbf{E}(u(t,x) - u(s,x))(u(b,x) - u(a,x)) = 0.$$

In particular, $u(t,x) - u(s,x)$ and $u(b,x) - u(a,x)$ are independent random variables.

Based on the covariance formula (8.22) and the result stated in Lemma 13, we will get the asymptotic behavior of the temporal quadratic variation of the process u . Let us set, for $N \geq 1$ and $x \in \mathbb{R}$,

$$T_{N,x} = \frac{1}{N^2} \sum_{j=0}^{N-1} (u(j+1,x) - u(j,x))^2 \quad (8.23)$$

where u is the mild solution to the stochastic wave equation. We have,

$$\begin{aligned} \mathbf{E}T_{N,x} &= \frac{1}{N^2} \sum_{j=0}^{N-1} \mathbf{E}(u(j+1,x) - u(j,x))^2 = \frac{1}{N^2} \sum_{j=0}^{N-1} \frac{1}{4}((j+1)^2 - j^2) \\ &= \frac{1}{4N^2} \sum_{j=0}^{N-1} (2j+1) = \frac{1}{4N^2} \left(2 \frac{N(N-1)}{2} + N \right) = \frac{1}{4} \end{aligned} \quad (8.24)$$

for every $N \geq 1$ and $x \in \mathbb{R}$. We will actually prove below that the sequence $(T_{N,x}, N \geq 1)$ converges to $\frac{1}{4}$ almost surely and in mean square.

Proposition 12. *The sequence $(T_{N,x}, N \geq 1)$ converges to $\frac{1}{4}$ almost surely and in $L^2(\Omega)$.*

Proof : We have

$$\mathbf{E} \left(T_{N,x} - \frac{1}{4} \right)^2 = \mathbf{E}T_{N,x}^2 - \frac{1}{2}\mathbf{E}T_{N,x} + \frac{1}{16} = \mathbf{E}T_{N,x}^2 - \frac{1}{16}. \quad (8.25)$$

Let us compute $\mathbf{E}T_{N,x}^2$. We have

$$\begin{aligned} \mathbf{E}T_{N,x}^2 &= \frac{1}{N^4} \sum_{j,k=0}^{N-1} \mathbf{E}(u(j+1,x) - u(j,x))^2 (u(k+1,x) - u(k,x))^2 \\ &= \frac{1}{N^4} \sum_{j=0}^{N-1} \mathbf{E}(u(j+1,x) - u(j,x))^4 \\ &\quad + \frac{1}{N^4} \sum_{j,k=0; j \neq k}^{N-1} \mathbf{E}(u(j+1,x) - u(j,x))^2 \mathbf{E}(u(k+1,x) - u(k,x))^2 \\ &=: t_{1,N} + t_{2,N}. \end{aligned}$$

Since $u(j+1, x) - u(j, x) \sim N(0, \frac{1}{4}(2j+1))$, we get

$$\begin{aligned} t_{1,N} &= \frac{1}{N^4} \frac{3}{16} \sum_{j=0}^{N-1} (2j+1)^2 \\ &= \frac{1}{N^4} \frac{3}{16} \sum_{j=0}^{N-1} (4j^2 + 4j + 1) = \frac{1}{N^4} \frac{3}{16} \left(\frac{4}{3}N^3 - \frac{1}{3}N \right) \\ &= \frac{1}{4N} - \frac{1}{16N^3}. \end{aligned}$$

Also, concerning the second summand denoted by $t_{2,N}$, we have

$$\begin{aligned} t_{2,N} &= \frac{1}{16} \frac{1}{N^4} \sum_{j,k=0; j \neq k}^{N-1} (2j+1)(2k+1) \\ &= \frac{1}{16} - \frac{1}{3} t_{1,N}. \end{aligned}$$

Consequently

$$\mathbf{E}T_{N,x}^2 = \frac{1}{16} + \frac{2}{3} t_{1,N} = \frac{1}{16} + \frac{1}{6N} - \frac{1}{24N^3}.$$

Thus, by (8.25),

$$\mathbf{E} \left(T_{N,x} - \frac{1}{4} \right)^2 = \frac{1}{6N} - \frac{1}{24N^3} \xrightarrow{N \rightarrow \infty} 0 \quad (8.26)$$

and this gives the convergence in $L^2(\Omega)$ of the sequence $T_{N,x}$ to $\frac{1}{4}$. The almost sure convergence is an easy consequence of the standard law of large numbers, since the increments of the solution in time are independent. \square

Now, we set for every $N \geq 1$ and $x \in \mathbb{R}$,

$$U_{N,x} = \sqrt{N} \left(T_{N,x} - \frac{1}{4} \right). \quad (8.27)$$

By relation (8.26) in the proof of Proposition 12,

$$\mathbf{E}U_{N,x}^2 = \frac{1}{6} - \frac{1}{24N^2} \xrightarrow{N \rightarrow \infty} \frac{1}{6}. \quad (8.28)$$

The random variable $U_{N,x}$ can be also expressed as a multiple integral of order two with respect to the space-time white noise. By writing, for every $j \geq 0$,

$$u(j+1, x) - u(j, x) = I_1(h_{j,x})$$

with

$$h_{j,x}(s, y) = 1_{(0, j+1)}(s) G_1(j+1-s, x-y) - 1_{(0, j)}(s) G_1(j-s, x-y) \quad (8.29)$$

we obtain

$$U_{N,x} = I_2(f_{N,x}) \text{ with } f_{N,x} = \frac{1}{N^{\frac{3}{2}}} \sum_{j=0}^{N-1} h_{j,x}^{\otimes 2}.$$

We will use Theorem 7 to show that the sequence (8.27) satisfies a Central Limit Theorem and to evaluate the distance between its law and the limit distribution. Recall that below d stands for one of the following distances : Kolmogorov, total variation, Wasserstein or Fortet-Mourier.

Theorem 9. Consider the sequence $(U_{N,x}, N \geq 1)$ given by (8.27). Then $U_{N,x}$ converges in distribution, as $N \rightarrow \infty$, to the normal law $\mathcal{N}\left(0, \frac{1}{6}\right)$ and for N large enough,

$$d\left(U_{N,x}, \mathcal{N}\left(0, \frac{1}{6}\right)\right) \leq C \frac{1}{\sqrt{N}}.$$

Proof : We have, with $h_{j,x}, j = 0, \dots, N-1$ given by (8.29),

$$DU_{N,x} = \frac{2}{N^{\frac{3}{2}}} \sum_{j=0}^{N-1} I_1(h_{j,x}) h_{j,x}$$

and, since $\langle h_{j,x}, h_{k,x} \rangle = \mathbf{E}((u(j+1,x) - u(j,x))(u(k+1,x) - u(k,x))) = 0$ for $j \neq k$,

$$\begin{aligned} \|DU_{N,x}\|^2 &= \frac{4}{N^3} \sum_{l,k=0}^{N-1} I_1(h_{l,x}) I_1(h_{k,x}) \langle h_{l,x}, h_{k,x} \rangle \\ &= \frac{4}{N^3} \sum_{j=0}^{N-1} I_1(h_{j,x})^2 \|h_{j,x}\|^2 \\ &= \frac{4}{N^3} \sum_{j=0}^{N-1} I_2(h_{j,x}^{\otimes 2}) \|h_{j,x}\|^2 + \mathbf{E} \|DU_{N,x}\|^2 \end{aligned}$$

with (see relation (8.28))

$$\mathbf{E} \|DU_{N,x}\|^2 = 2\mathbf{E} U_{N,x}^2 = \frac{1}{3} - \frac{1}{12N^2}.$$

Therefore, with $C > 0$ a generic constant,

$$\begin{aligned} \text{Var}(\|DU_{N,x}\|^2) &= \frac{16}{N^6} \mathbf{E} \left(\sum_{j=0}^{N-1} I_2(h_{j,x}^{\otimes 2}) \|h_{j,x}\|^2 \right)^2 \\ &= C \frac{1}{N^6} \sum_{j,k=0}^{N-1} \langle h_{j,x}, h_{k,x} \rangle^2 \|h_{j,x}\|^2 \|h_{k,x}\|^2 \\ &= C \frac{1}{N^6} \sum_{j=0}^{N-1} \|h_{j,x}\|^8 = C \frac{1}{N^6} \sum_{j=0}^{N-1} (2j+1)^4 \\ &\leq C \frac{1}{N}. \end{aligned}$$

Thus, by the inequality (8.18),

$$d\left(U_{N,x}, \mathcal{N}\left(0, \frac{1}{6}\right)\right) \leq C \left(\sqrt{\text{Var}(\|DU_{N,x}\|^2)} + \left| \mathbf{E} \|DU_{N,x}\|^2 - \frac{1}{3} \right| \right) \leq C \frac{1}{\sqrt{N}}.$$

□

8.3 The Wishart matrix

In this section we will analyze the asymptotic behavior of the Wishart matrix constructed from an initial random matrix $\mathcal{X}_{n,d}$ with n rows and d columns, whose entries are the spatial or

temporal increments of the mild solution to the stochastic wave equation. We will also evaluate the Wasserstein distance when both sizes n and d are large (i.e. in the high-dimensional regime) between the law of the renormalized Wishart matrix and the Gaussian Wigner matrix which corresponds to its limit in law.

Let us first recall some facts about the Wasserstein distance for random matrices and random vectors.

Let \mathcal{X}, \mathcal{Y} be two random matrices with values in $\mathcal{M}_n(\mathbb{R})$, $n \geq 1$. We will denote by d_W the Wasserstein distance between the probability distributions of \mathcal{X} and \mathcal{Y} . That is,

$$d_W(\mathcal{X}, \mathcal{Y}) = \sup_{\|g\|_{\text{Lip}} \leq 1} |\mathbf{E}(g(\mathcal{X})) - \mathbf{E}(g(\mathcal{Y}))|,$$

where the Lipschitz norm $\|\cdot\|_{\text{Lip}}$ of $g : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}$ is defined by

$$\|g\|_{\text{Lip}} = \sup_{A \neq B, A, B \in \mathcal{M}_n(\mathbb{R})} \frac{|g(A) - g(B)|}{\|A - B\|_{\text{HS}}},$$

with $\|\cdot\|_{\text{HS}}$ denoting the Hilbert-Schmidt norm on $\mathcal{M}_n(\mathbb{R})$.

We recall that if X, Y are two n -dimensional random vectors, then the Wasserstein distance between them is defined to be

$$d_W(X, Y) = \sup_{\|g\|_{\text{Lip}} \leq 1} |\mathbf{E}(g(X)) - \mathbf{E}(g(Y))|, \quad (8.30)$$

where the Lipschitz norm $\|\cdot\|_{\text{Lip}}$ of $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is defined by

$$\|g\|_{\text{Lip}} = \sup_{x \neq y \in \mathbb{R}^n} \frac{|g(x) - g(y)|}{\|x - y\|_{\mathbb{R}^n}},$$

with $\|\cdot\|_{\mathbb{R}^n}$ denoting the Euclidean norm on \mathbb{R}^n .

If $\mathcal{X} = (X_{ij})_{1 \leq i, j \leq n}$ is an $n \times n$ symmetric random matrix, we associate to it its "half-vector" defined to be the $n(n+1)/2$ -dimensional random vector

$$X^{\text{half}} = (X_{1,1}, X_{1,2}, \dots, X_{1,n}, X_{2,2}, X_{2,3}, \dots, X_{2,n}, \dots, X_{n,n}). \quad (8.31)$$

The following result has been proven in [98]. It gives the link between the distance of two symmetric random matrices and their associated half vectors : if \mathcal{X}, \mathcal{Y} are two symmetric random matrices with values in $\mathcal{M}_n(\mathbb{R})$ then

$$d_W(\mathcal{X}, \mathcal{Y}) \leq \sqrt{2} d_W(X^{\text{half}}, Y^{\text{half}}), \quad (8.32)$$

where $X^{\text{half}}, Y^{\text{half}}$ are the associated half-vectors defined in (8.31). Due to the bound (8.32), the analysis of the distance between two random matrices can be reduced to the study of the distance between their corresponding random vectors.

8.3.1 The case where entries are temporal increments of the solution

The first situation that we treat is when the entries of the starting matrix are the increments in time of the solution to the stochastic wave equation. That is, we consider the random matrix $\mathcal{Y}_{n,d} = (Y_{i,j})_{1 \leq i \leq n, 1 \leq j \leq d}$ with entries given by, for $1 \leq i \leq n, 1 \leq j \leq d$

$$Y_{i,j} = u^{(i)}(j,x) - u^{(i)}(j-1,x) \quad (8.33)$$

where

$$u^{(i)}(t,x) = \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y) W^{(i)}(ds, dy)$$

with $W^{(i)}, 1 \leq i \leq n$ independent space-time white noises, see (8.3). So, for every $1 \leq i \leq n, 1 \leq j \leq d$,

$$\begin{aligned} Y_{i,j} &= \int_0^\infty \int_{\mathbb{R}} [G_1(j-s, x-y) \mathbf{1}_{(0,j)}(s) - G_1(j-1-s, x-y) \mathbf{1}_{(0,j-1)}(s)] W^{(i)}(ds, dy) \\ &= I_1^{(i)}(h_{j-1,x}) \end{aligned}$$

where $h_{j,x}$ are given by (8.29) and $I_p^{(i)}$ denotes the multiple integral of order p with respect to the Gaussian process $W^{(i)}$.

The elements of the matrix $\mathcal{Y}_{n,d}$ are Gaussian and mutually independent (see Lemma 13) and this fact makes the analysis easier. On the other hand, the entries are not identically distributed and the known results in the literature (e.g. [97], [98]) cannot be applied directly. In fact,

$$Y_{i,j} \sim \mathcal{N}\left(0, \frac{1}{4}(2j-1)\right) \quad (8.34)$$

for all $1 \leq i \leq n$ and $1 \leq j \leq d$.

We define the renormalized Wishart matrix $\overline{\mathcal{W}}_{n,d} = (\overline{W}_{i,j})_{1 \leq i, j \leq n}$ by

$$\overline{\mathcal{W}}_{n,d} = \sqrt{d} \left(\frac{1}{d^2} \mathcal{Y}_{n,d} \mathcal{Y}_{n,d}^T - \frac{1}{4} I_n \right)$$

whose elements can be expressed as, due to (8.24),

$$\overline{W}_{i,i} = \frac{1}{d^{\frac{3}{2}}} \sum_{k=1}^d \left(u^{(i)}(k,x) - u^{(i)}(k-1,x) \right)^2 - \sqrt{d} \frac{1}{4} = \frac{1}{d^{\frac{3}{2}}} \sum_{k=0}^{d-1} I_2^{(i)}(h_{k,x}^{\otimes 2}) \quad (8.35)$$

and

$$\begin{aligned} \overline{W}_{i,j} &= \frac{1}{d^{\frac{3}{2}}} \sum_{k=0}^{d-1} \left(u^{(i)}(k+1,x) - u^{(i)}(k,x) \right) \left(u^{(j)}(k+1,x) - u^{(j)}(k,x) \right) \\ &= \frac{1}{d^{\frac{3}{2}}} \sum_{k=0}^{d-1} I_1^{(i)}(h_{k,x}) I_1^{(j)}(h_{k,x}). \end{aligned} \quad (8.36)$$

We also define the Wigner random matrix $\mathcal{Z}_{1,n} = (Z_{1,i,j})_{1 \leq i, j \leq n}$ with the following entries

$$\begin{cases} Z_{1,i,i} \sim \mathcal{N}(0, \frac{1}{6}) \text{ for } 1 \leq i \leq n \\ Z_{1,i,j} \sim \mathcal{N}(0, \frac{1}{12}) \text{ for } 1 \leq i < j \leq n \\ Z_{1,i,j} = Z_{1,j,i} \text{ for } 1 \leq j < i \leq n \end{cases} \quad (8.37)$$

with $(Z_{1,i,j}, 1 \leq i < j \leq n)$ independent.

The components of the random matrix $\bar{W}_{n,d}$ are Skorohod integrals (they can also be viewed as multiple stochastic integrals of order one) but with respect to different Gaussian processes. Therefore, the usual Stein-Malliavin inequality to evaluate the distance between an arbitrary random vector and a Gaussian random vector, such as Theorem 6.1.1 in [93] or Theorem 3.5 in [96], cannot be applied directly. We will state and prove a new criterion which is needed in our situation. Let $(W^{(i)}(h), h \in H)$ be n independent copies of an isonormal process $(W(h), h \in H)$. Let $D^{(i)}, \delta^{(i)}, \mathbb{D}^{k,p,i}$ denote their associated Malliavin derivative, Skorohod integral and Malliavin-Sobolev spaces.

Theorem 10. *Assume $F = (F_1, \dots, F_n)$ is a n -dimensional random vector such that for every $1 \leq i \leq n$,*

$$F_i = \delta^{(i)}(u^{(i)})$$

with $u^{(i)} \in \mathbb{D}^{1,2,i}$ and $F_i \in \mathbb{D}^{1,4,i}$. Let $C = (C_{i,j})_{1 \leq i,j \leq n}$ be a $n \times n$ symmetric positive definite matrix. Then

$$d_W(F, \mathcal{N}(0, C)) \leq C \sqrt{\sum_{i,j=1}^n \mathbf{E}(C_{i,j} - \langle u^{(i)}, D^{(i)} F_j \rangle_H)^2}.$$

Proof : By Theorem 3.5 in [96], we have

$$d_W(F, \mathcal{N}(0, C)) \leq \sup_{g \in \mathcal{C}} |\mathbf{E}\langle C, \text{Hess}U_0g(F) \rangle_{HS} - \mathbf{E}\langle F, \nabla U_0g(F) \rangle_{\mathbb{R}^n}| \quad (8.38)$$

where Hess stands for the Hessian matrix, \mathcal{C} denotes the class of functions $g : \mathbb{R}^n \rightarrow \mathbb{R}$ of class \mathcal{C}^2 with bounded first and second derivative and

$$U_0g(x) = \int_0^1 \mathbf{E} \left[g(\sqrt{t}x + \sqrt{1-t}Z) - g(Z) \right] dt$$

with $Z \sim \mathcal{N}(0, C)$. For $g \in \mathcal{C}$, we can write

$$\begin{aligned} & \mathbf{E}\langle C, \text{Hess}U_0g(F) \rangle_{HS} - \mathbf{E}\langle F, \nabla U_0g(F) \rangle_{\mathbb{R}^n} \\ &= \mathbf{E} \sum_{i,j=1}^n C_{i,j} \frac{\partial^2 U_0g}{\partial x_i \partial x_j}(F) - \mathbf{E} \sum_{i=1}^n F_i \frac{\partial U_0g}{\partial x_i}(F). \end{aligned} \quad (8.39)$$

We can write via the duality formula (8.51), for every $i = 1, \dots, n$

$$\begin{aligned} \mathbf{E} F_i \frac{\partial U_0g}{\partial x_i}(F) &= \mathbf{E} \delta^{(i)}(u^{(i)}) \frac{\partial U_0g}{\partial x_i}(F) \\ &= \mathbf{E} \langle u^{(i)}, D^{(i)} \frac{\partial U_0g}{\partial x_i}(F) \rangle_H = \sum_{i,j=1}^n \frac{\partial^2 U_0g(F)}{\partial x_i \partial x_j}(F) \langle u^{(i)}, D^{(i)} F_j \rangle_H. \end{aligned}$$

Therefore (8.39) becomes

$$\begin{aligned} \mathbf{E}\langle C, \text{Hess}U_0g(F) \rangle_{HS} - \mathbf{E}\langle F, \nabla U_0g(F) \rangle_{\mathbb{R}^n} &= \sum_{i,j=1}^n \frac{\partial^2 U_0g}{\partial x_i \partial x_j}(F) (C_{i,j} - \langle u^{(i)}, D^{(i)} F_j \rangle_H) \\ &= \mathbf{E}\langle \text{Hess}U_0g(F), C - M \rangle_{HS} \end{aligned}$$

where we denoted $M = (M_{i,j})_{1 \leq i,j \leq n}$ with $M_{i,j} = \langle u^{(i)}, D^{(i)} F_j \rangle_H$ for every $i, j = 1, \dots, n$. Consequently, by Cauchy-Schwarz's inequality,

$$\begin{aligned} |\mathbf{E}\langle C, \text{Hess}U_0g(F) \rangle_{HS} - \mathbf{E}\langle F, \nabla U_0g(F) \rangle_{\mathbb{R}^n}| &\leq \sqrt{\mathbf{E}\|\text{Hess}U_0g(F)\|_{HS}^2} \sqrt{\mathbf{E}\|C - M\|_{HS}^2} \\ &\leq \|C^{-1}\|_{op} \|g\|_{Lip} \|C\|_{op}^{\frac{1}{2}} \sqrt{\mathbf{E}\|C - M\|_{HS}^2} \end{aligned}$$

where the last inequality has been proven in Lemma 3.3 in [96]. The conclusion follows from the above estimate and (8.38). \square

As a consequence of Theorem 9 in Section 8.2.2, we can show that for every $n \geq 1$, the matrix $\overline{\mathcal{W}}_{n,d}$ converges componentwise to the above Wigner matrix. We can express the entries of the matrix $\overline{\mathcal{W}}_{n,d}$ as Skorohod integrals. Indeed, for every $1 \leq i, j \leq n$,

$$\overline{W}_{i,j} = \delta^{(i)}(v^{(j)}) \quad \text{with } v^{(j)} = \frac{1}{d^{\frac{3}{2}}} \sum_{k=0}^{d-1} I_1^{(j)}(h_{k,x}) h_{k,x}. \quad (8.40)$$

Therefore, we can use Theorem 10 to measure the distance between the half vector to $\overline{\mathcal{W}}_{n,d}$ and the half vector associated to the limit matrix. We start by evaluating this distance.

Proposition 13. *Consider the random matrix $\overline{\mathcal{W}}_{n,d}$ with entries given by (8.35), (8.36) and let $\mathcal{Z}_{1,n}$ be the Wigner matrix (8.37). Let $\overline{W}_{n,d}^{half}$ and $Z_{1,n}^{half}$ be their associated half vectors, see (8.31). Then for n, d large enough,*

$$d_W(\overline{W}_{n,d}^{half}, Z_{1,n}^{half}) \leq C \sqrt{\frac{n^3}{d}}.$$

Proof : We use Theorem 10 to write

$$d_W(\overline{W}_{n,d}^{half}, Z_n^{half}) \leq C \sqrt{\sum_{i,j,a,b=1}^n \mathbf{E}(\langle v^{(j)}, D^{(i)} \overline{W}_{a,b} \rangle - \mathbf{E} Z_{1,i,j} Z_{1,a,b})^2} \quad (8.41)$$

with $v^{(j)}$ from (8.40). Now, the quantity $\mathbf{E}(\langle v^{(j)}, D^{(i)} \overline{W}_{a,b} \rangle - \mathbf{E} Z_{1,i,j} Z_{1,a,b})^2$ by following several cases. The calculation is made easier due to the independence of the temporal increments of the solution which gives $\langle h_{k,x}, h_{l,x} \rangle = 0$ if $k \neq l$. Let us show below how these quantities are estimated.

Case $i = j = a = b$. Calculation of

$$\mathbf{E} \left(\langle v^{(i)}, D^{(i)} \overline{W}_{i,i} \rangle - \frac{1}{6} \right)^2.$$

We have, by (8.40) and (8.35),

$$\begin{aligned} \langle v^{(i)}, D^{(i)} \overline{W}_{i,i} \rangle &= \frac{2}{d^3} \sum_{k,l=0}^{d-1} I_1^{(i)}(h_{k,x}) I_1^{(i)}(h_{l,x}) \langle h_{k,x}, h_{l,x} \rangle \\ &= \frac{2}{d^3} \sum_{k=0}^{d-1} I_1^{(i)}(h_{k,x})^2 \|h_{k,x}\|^2 = \frac{2}{d^3} \sum_{k=0}^{d-1} I_2^{(i)}(h_{k,x}^{\otimes 2}) \|h_{k,x}\|^2 + \mathbf{E} \langle v^{(i)}, D^{(i)} \overline{W}_{i,i} \rangle \\ &= \frac{2}{d^3} \sum_{k=0}^{d-1} I_2^{(i)}(h_{k,x}^{\otimes 2}) \|h_{k,x}\|^2 + \mathbf{E} \overline{W}_{i,i}^2 \\ &= \frac{2}{d^3} \sum_{k=0}^{d-1} I_2^{(i)}(h_{k,x}^{\otimes 2}) \|h_{k,x}\|^2 + \frac{1}{6} - \frac{1}{24d^2} \end{aligned}$$

where we used $\mathbf{E}\overline{W}_{i,i}^2 = \frac{1}{6} - \frac{1}{24d^2}$ which follows from (8.28). Thus

$$\begin{aligned} & \mathbf{E} \left(\langle v^{(i)}, D^{(i)} \overline{W}_{i,i} \rangle - \frac{1}{6} \right)^2 \\ &= \mathbf{E} \left(\frac{2}{d^3} \sum_{k=0}^{d-1} I_2^{(i)}(h_{k,x}^{\otimes 2}) \|h_{k,x}\|^2 \right)^2 + \left(\frac{1}{24d^2} \right)^2 \\ &= C \frac{1}{d^6} \sum_{k=0}^{d-1} \|h_{k,x}\|^8 + \left(\frac{1}{24d^2} \right)^2 \\ &= C \frac{1}{d^6} \sum_{k=0}^{d-1} (2j+1)^4 + \left(\frac{1}{24d^2} \right)^2 \leq C \frac{1}{d}. \end{aligned}$$

Case $i = a \neq j = b$. Next, by (8.40) and (8.36), since $\mathbf{E}\overline{W}_{i,j}^2 = \frac{1}{12} - \frac{1}{48d^2}$ (this is also a consequence of the estimate (8.28))

$$\begin{aligned} \langle v^{(j)}, D^{(i)} \overline{W}_{i,j} \rangle &= \frac{1}{d^3} \sum_{k,l=0}^{d-1} I_1^{(j)}(h_{k,x}) I_1^{(j)}(h_{l,x}) \langle h_{k,x}, h_{l,x} \rangle \\ &= \frac{1}{d^3} \sum_{k=0}^{d-1} I_2^{(j)}(h_{k,x}^{\otimes 2}) \|h_{k,x}\|^2 + \frac{1}{12} - \frac{1}{48d^2} \end{aligned}$$

which gives

$$\mathbf{E} \left(\langle v^{(j)}, D^{(i)} \overline{W}_{i,j} \rangle - \frac{1}{12} \right)^2 \leq C \frac{1}{d}.$$

The remaining cases can be treated here in a similar manner without any difficulty. We find, if $M = \{(i, j, a, b) \in \{1, \dots, n\}^4, ((i = a) \text{ and } (j \neq b)) \text{ or, } ((i \neq a) \text{ and } (j = b)) \text{ or, } ((i = b) \text{ and } (j \neq a)) \text{ or, } ((j = a) \text{ and } (i \neq b))\}$

$$\max_{i,j,a,b \in M} \mathbf{E} \left(\langle v^{(j)}, D^{(i)} \overline{W}_{i,b} \rangle \right)^2 \leq C \frac{1}{d}.$$

Indeed, if for instance $i = a$ and $j \neq b$, then

$$\mathbf{E} \left[\langle v^{(j)}, D^{(i)} \overline{W}_{i,b} \rangle \right] = \mathbf{E} \left[\frac{1}{d^3} \sum_{k=1}^d \|h_{k,x}\|^2 I_1^{(j)}(h_{k,x}) I_1^{(b)}(h_{k,x}) \right] \leq C \frac{1}{d},$$

exactly as above.

Finally, in the case $(i \neq a, b \text{ and } j \neq a, b)$, we find

$$\mathbf{E} \left(\langle v^{(j)}, D^{(i)} \overline{W}_{a,b} \rangle \right)^2 = 0.$$

By combining the above estimates with (8.41), we obtain the desired conclusion. \square

Now we state the main result of this section.

Theorem 11. Consider the renormalized Wishart matrix $\overline{\mathcal{W}}_{n,d} = (\overline{W}_{i,j})_{1 \leq i,j \leq n}$ with entries given by (8.35), (8.36). Then as $d \rightarrow \infty$, the matrix $\overline{\mathcal{W}}_{n,d}$ converges component-wise to the Wigner matrix $\mathcal{Z}_{1,n}$ given by (8.37). Moreover, for $d, n \geq 1$ large enough

$$d_W(\overline{\mathcal{W}}_{n,d}, \mathcal{Z}_{1,n}) \leq C \sqrt{\frac{n^3}{d}}.$$

Proof : The proof is immediate from Proposition 13 and the bound (8.32). \square

Remark 9. We obtain again the classical bound $\sqrt{\frac{n^3}{d}}$, although the renormalization of the Wishart matrix is different (i.e. the order of renormalization is $d^{-\frac{3}{2}}$).

8.3.2 The case where entries are spatial increments of the solution

We start with the $n \times d$ random matrix $\mathcal{X}_{n,d} = (X_{i,j})_{1 \leq i \leq n, 1 \leq j \leq d}$ with entries given by

$$X_{i,j} = u^{(i)}(t, j) - u^{(i)}(t, j-1) = \int_0^t \int_{\mathbb{R}} [G_1(t-s, j-y) - G_1(t-s, j-1-y)] W^{(i)}(ds, dy).$$

Above $W^{(i)}, i = 1, \dots, n$ are independent copies of the space-time white noise W defined by its covariance (8.3) and G_1 is the fundamental solution to the wave equation given by (8.4).

Therefore, for each $i = 1, \dots, n, j = 1, \dots, d$, the entry $X_{i,j}$ is the j th spatial increment of the mild solution to the wave equation driven by $W^{(i)}$, i.e.

$$u^{(i)}(t, x) = \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y) W^{(i)}(ds, dy).$$

That means that the elements on the same row are correlated and their correlation is the correlation of the spatial increments of the mild solution to the wave equation, which has been studied in Section 8.2. The entries on different rows are independent. We also have, for $t > 0$ and for every $1 \leq i \leq n, 1 \leq j \leq d$,

$$\mathbf{E}X_{i,j}^2 = \mathbf{E}(u(t, j) - u(t, j-1))^2 = e_t$$

where u is defined by (8.5) and e_t is the constant (8.12).

The Wishart matrix associated to $\mathcal{X}_{n,d}$ is defined as

$$\mathcal{W}_{n,d} = \frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - e_t \mathcal{I}_n$$

where \mathcal{I}_n is the $n \times n$ identity matrix. This matrix depends on $t > 0$, i.e. $\mathcal{W}_{n,d} = \mathcal{W}_{n,d,t}$ but we omit the index t in the notation, for simplicity. Via Proposition 11, one can show that $\mathcal{W}_{n,d}$ tends to zero almost surely as $d \rightarrow \infty$, if n is fixed.

Denote by

$$\widetilde{\mathcal{W}}_{n,d} = \sqrt{d} \mathcal{W}_{n,d} = \sqrt{d} (W_{i,j})_{1 \leq i,j \leq n} \quad (8.42)$$

and consider the Wigner matrix $\mathcal{Z}_n = (Z_{i,j})_{1 \leq i,j \leq n}$ with the following entries

$$\begin{cases} Z_{i,i} \sim \mathcal{N}(0, K_{1,t}) \text{ for } 1 \leq i \leq n \\ Z_{i,j} \sim \mathcal{N}\left(0, \frac{K_{1,t}}{2}\right) \text{ for } 1 \leq i < j \leq n \\ Z_{i,j} = Z_{j,i} \text{ for } 1 \leq j < i \leq n \end{cases} \quad (8.43)$$

with $(Z_{i,j}, 1 \leq i < j \leq n)$ independent.

We have the following result concerning the limit behavior in distribution of the Wishart matrix $\widetilde{\mathcal{W}}_{n,d}$.

Theorem 12. *Consider the renormalized Wishart matrix $\widetilde{\mathcal{W}}_{n,d} = (\widetilde{W}_{i,j})_{1 \leq i, j \leq n}$ by (8.42). Then as $d \rightarrow \infty$, the matrix $\widetilde{\mathcal{W}}_{n,d}$ converges componentwise to the Wigner matrix \mathcal{Z}_n given by (8.43). Moreover, for $d, n \geq 1$ large enough,*

$$d_W(\widetilde{\mathcal{W}}_{n,d}, \mathcal{Z}_n) \leq C \sqrt{\frac{n^3}{d}}.$$

Proof : We can apply directly the results in [98] (Theorem 1.2 in this reference). Let $(K, \langle \cdot, \cdot \rangle_K)$ be a real separable Hilbert space and let $(e_{i,j}, i, j \geq 1) \subset K$ be a family such that

$$\langle e_{i,j}, e_{i',j'} \rangle_K = 1_{(i=i')}(j-j')$$

with f_i given by (8.10). Let $(B(h), h \in K)$ be an isonormal process, i.e. a centered Gaussian family with $\mathbf{E}B(h)B(g) = \langle h, g \rangle_K$ for every $h, g \in K$. Consider the random matrix $\mathcal{A}_{n,d} = (A_{i,j})_{1 \leq i \leq n, 1 \leq j \leq d}$ with $A_{i,j} = B(e_{i,j})$ and its associated renormalized random matrix

$$\mathcal{W}_{n,d,A} = \sqrt{d} \left(\frac{1}{d} \mathcal{A}_{n,d} \mathcal{A}_{n,d}^T - e_t \mathcal{I}_n \right).$$

Then clearly

$$\mathcal{X}_{n,d} \stackrel{(d)}{=} \mathcal{A}_{n,d} \text{ and } \mathcal{W}_{n,d} \stackrel{(d)}{=} \mathcal{W}_{n,d,A}.$$

Using the proof of Proposition 1 (see relation (8.16) and the comment after this formula), it is immediate to see that the function f_t given by (8.10) belongs to the space $\ell^{\frac{4}{3}}(\mathbb{Z})$. Then, by Theorem 1.2 in [98], we have

$$d_W(\widetilde{\mathcal{W}}_{n,d,A}, \mathcal{Z}_n) \leq C \sqrt{\frac{n^3}{d}}$$

and this implies the conclusion. \square

8.4 Appendix : Wiener chaos and Malliavin derivative

Let $T = \mathbb{R}_+ \times \mathbb{R}$ and denote by $L^2_S(T^p)$ the set of real-valued symmetric square integrable functions on T^p . Let $(B_t)_{t \in T}$ be a Wiener process. Denote by $B(\varphi) := \int_T \varphi_s dB_s$ the Wiener integral of $\varphi \in H := L^2(T, \mathcal{B}(T), \lambda)$ with respect to the Brownian motion B . We denote by λ the Lebesgue measure and $\mathcal{B}(T)$ stands for the Borel subsets of T . The family $(B(\varphi), \varphi \in H)$ forms an isonormal process, i.e. a Gaussian family of centered random variables such that

$$\mathbf{E}B(\varphi_1)B(\varphi_2) = \langle \varphi_1, \varphi_2 \rangle_H = \int_T \varphi_1(s) \varphi_2(s) ds$$

for any $\varphi_1, \varphi_2 \in H$.

Denote I_n the multiple stochastic integral with respect to B (see [100]). This I_n is actually an isometry between the Hilbert space $H^{\odot n}$ (symmetric tensor product) equipped with the scaled norm $\sqrt{n!} \|\cdot\|_{H^{\otimes n}}$ and the Wiener chaos of order n which is defined as the closed linear span of the random variables $H_n(B(\varphi))$ where $\varphi \in H, \|\varphi\|_H = 1$ and H_n is the Hermite polynomial of degree $n \geq 1$

$$H_n(x) = (-1)^n \exp\left(\frac{x^2}{2}\right) \frac{d^n}{dx^n} \left(\exp\left(-\frac{x^2}{2}\right)\right), \quad x \in \mathbb{R}. \quad (8.44)$$

The isometry of multiple integrals can be written as : if \tilde{f} denotes the symmetrization of the function f , for m, n positive integers,

$$\begin{aligned} \mathbf{E}(I_n(f)I_m(g)) &= n! \langle \tilde{f}, \tilde{g} \rangle_{H^{\otimes n}} \quad \text{if } m = n, \\ \mathbf{E}(I_n(f)I_m(g)) &= 0 \quad \text{if } m \neq n. \end{aligned} \quad (8.45)$$

We will use the product formula for multiple stochastic integrals, if $f \in L^2_S(T^m)$ and $g \in L^2_S(T^n)$, then

$$I_m(f)I_n(g) = \sum_{r=0}^{m \wedge n} r! \binom{m}{r} \binom{n}{r} I_{m+n-2r}(f \otimes_r g) \quad (8.46)$$

where for $r = 0, \dots, m \wedge n$, the contraction $f \otimes_r g$ is the function in $L^2(T^{m+n-2r})$ given by

$$(f \otimes_r g)(t_1, \dots, t_{m+n-2r}) = \int_{T^r} f(u_1, \dots, u_r, t_1, \dots, t_{m-r}) g(u_1, \dots, u_r, t_{m-r+1}, \dots, t_{m+n-2r}) du_1 \dots du_r. \quad (8.47)$$

Notice that $f \otimes_r g$ is not necessarily a symmetric function (even if f, g are symmetric) and we will denote by $f \tilde{\otimes}_r g$ its symmetrization.

An useful property of finite sums of multiple integrals is the hypercontractivity. Namely, if $F = \sum_{k=0}^n I_k(f_k)$ with $f_k \in \mathcal{H}^{\otimes k}$ then

$$\mathbf{E}|F|^p \leq C_p (\mathbf{E}F^2)^{\frac{p}{2}}. \quad (8.48)$$

for every $p \geq 2$.

Let \mathcal{S} be the class of smooth functionals of the form

$$F = f(B_{t_1}, \dots, B_{t_n}), \quad t_1, \dots, t_n \in T, \quad (8.49)$$

with $f \in C^\infty(\mathbb{R}^n)$ with at most polynomial growth (for f and its derivatives). For the random variable (8.49) we define its Malliavin derivative with respect to B by

$$D_t F = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(B_{t_1}, \dots, B_{t_n}) 1_{[0, t_i]}(t), \quad t \in T.$$

The operator D is an unbounded closable operator and it can be extended to the closure of \mathcal{S} with respect to the Malliavin -Sobolev norm

$$\|F\|_{k,p}^p = \mathbf{E}|F|^p + \sum_{j=1}^k \mathbf{E}\|D^{(j)}F\|_{L^2(T^j)}^p, \quad F \in \mathcal{S}, p \geq 2, k \geq 1, \quad (8.50)$$

where $D^{(j)}$ stands for the j th iterated Malliavin derivative. This closure will be denoted by $\mathbb{D}^{k,p}$. The Skorohod integral, denoted by δ , is the adjoint operator of D . Its domain is

$$\text{Dom}(\delta) = \left\{ u \in L^2(T \times \Omega), \mathbf{E} \left| \int_T u_s D_s F ds \right| \leq C \|F\|_2 \right\}$$

and we have the duality relationship

$$\mathbf{E} F \delta(u) = \mathbf{E} \int_T D_s F u_s ds, \quad F \in \mathcal{S}, u \in \text{Dom}(\delta). \quad (8.51)$$

We set $\mathbb{L}^{k,p} = L^p(T; \mathbb{D}^{k,p})$, $k \geq 1, p \geq 2$. This set is a subset of $\text{Dom}(\delta)$ and it is endowed with the norm

$$\|u\|_{k,p}^p = \int_T \left[\mathbf{E} |u_t|^p + \sum_{j=1}^k \mathbf{E} \|D_{s_1, \dots, s_j} u_t\|_{L^2(T^j)}^p \right] dt.$$

The Malliavin derivative D acts on the Wiener chaos as an annihilation operator : if $F = I_n(f)$ with $f \in L^2(T^n)$ symmetric, then $D_t F = n I_{n-1}(f(\cdot, t))$ where " \cdot " stands for $n-1$ variables in T . The pseudo-inverse of the Ornstein–Uhlenbeck operator, denoted by L^{-1} , acts as follows : if $F = I_n(f)$ with $n \geq 1$, then $-L^{-1}F = \frac{1}{n}F$.

High-dimensional regime for Wishart matrices based on the increments of the solution to the stochastic heat equation

This article is accepted on Brazilian Journal of Probability and Statistics.

Joint work with David A.C. Mollinedo and Ciprian A. Tudor.

Abstract. We consider a $n \times d$ random matrix $\mathcal{X}_{n,d}$ whose entries are the spatial increments of the solution to the stochastic heat equation with space-time white noise. We analyze the limit behavior of the associated Wishart matrix, by showing that it converges almost surely to a diagonal matrix (with equal diagonal terms) and the renormalized Wishart matrix satisfies a Central Limit Theorem. Our techniques are based on the analysis on Wiener chaos, Malliavin calculus and Stein's method.

9.1 Introduction

If $\mathcal{X}_{n,d}$ is a $n \times d$ matrix with random entries, then its corresponding Wishart matrix $\mathcal{W}_{n,d}$ is usually defined as $\mathcal{W}_{n,d} = \frac{1}{d} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T$, where \mathcal{X}^T denotes the transpose of the matrix \mathcal{X} . The class of Wishart matrices represents a particular class of sample covariance matrices. The Wishart matrix contains the empirical variances and covariances of the data included in the initial matrix and it constitutes a good estimator for the long-term correlation of the data. They have many applications in multivariate statistical inference. A more complete description of their properties and of their applications can be found in [17], [70], [105] or in the seminal work of [127], where they have been introduced. The analysis of their behavior for large data sets is of interest.

While the first studies on Wishart matrices concerned the behavior in one dimensional regime (i.e. n is fixed and d tends to infinity), in the recent years, several studies focused on the behavior of the Wishart matrix in the high-dimensional regime, i.e. when both dimensions n and d are

large enough. This is due to the fact that the number of observed variables have increased with the development of the information technology. The limit behavior of the Wishart matrix depends on the structure of the random variables contained in the starting matrix $\mathcal{X}_{n,d}$. A large number of papers treated the situation when the initial entries are Gaussian random variables, independent or with (partial) correlation, see among others, in [20], [21], [53], [68], [97], [98], [106], [18]. Other works considered the case of non-Gaussian initial entries, see e.g. [19], [45], [46], [49] or [90].

The contribution of the present paper is to use the Stein-Malliavin calculus in order to analyze the asymptotic behavior of the Wishart matrix associated to an initial random matrix whose entries are the spatial increments of the solution to the stochastic heat equation driven by a space-time white noise. More precisely, we let $(W^{(i)}, 1 \leq i \leq n)$ to be n independent space-time white noises and assume that the line i of the initial matrix $\mathcal{X}_{n,d} = (X_{i,j})_{1 \leq i \leq n, 1 \leq j \leq d}$ contains the spatial increments $(u^{(i)}(t, j+1) - u^{(i)}(t, j), j = 1, \dots, d)$, where the time $t > 0$ is fixed and $u^{(i)}$ represents the solution to the stochastic heat equation driven by the space-time white noise $W^{(i)}$. That means that all the entries of the starting matrix are Gaussian, the entries on different rows are independent and the entries of the same row are correlated since the increments of the solution to the heat equation are correlated. Recall that, if $u(t, x)$ is the solution to the heat equation, then $u(t, x)$ represents the temperature (for example, along a rod connected to a heat source), at time t and at the point x . Thus, we can say that our initial data, on line i (with $1 \leq i \leq n$), contains the variation of the temperature along the rod i , at a fixed time, between two consecutive points in the space. We then study the asymptotic behavior of the associated Wishart matrix, which includes the empirical covariances of the data contained in the initial matrix $\mathcal{X}_{n,d}$. The behavior of the Wishart matrix indicates the asymptotic correlation of the temperature along a system with multiple rods. We notice that the references [97] and [98] also treated a related situation, i.e. the limit behavior of a Wishart matrix associated to an initial random matrix with correlated Gaussian entries. Actually, it is assumed in these references that the entries of the initial matrix on the same row constitutes a stationary sequence of identically distributed Gaussian random variables with a summability assumption on the correlation function of this sequence. In the situation that we consider, we show that our Wishart matrix fits into the context of [98], [97], so we will apply directly some findings in these references. Nevertheless, in contrast to [97], [98], we are able to obtain the convergence toward a diagonal Gaussian matrix. This is due to the concrete correlation structure of our initial entries. Indeed, every non-diagonal element of the Wishart matrix has zero mean and it can be written as an empirical sum whose terms have a relatively weak correlation. It then follows that these non-diagonal terms converges almost surely to zero when $d \rightarrow \infty$.

As we will see, our proofs rely on a sharp analysis of the correlation structure of the spatial increments of the solution to the stochastic heat equation and on the techniques of the Stein-Malliavin calculus. In particular, we derive the asymptotic behavior of the quadratic variation (in time or in space) defined on the whole positive axis, of this solution, and this may have its own interest. Via the techniques of the Malliavin calculus, we evaluate the Wasserstein distance between the renormalized Wishart matrix and its limit matrix (a Gaussian matrix, the so-called GOE matrix). We obtain a classical bound, i.e. we show that this Wasserstein distance's order is lesser than $\sqrt{\frac{n^3}{d}}$. A similar analysis for the wave equation has been done in the reference [53]. We also notice that the case of the temporal increments of the solution as initial data is included in the work [18], since, viewed as a process with respect to the time variable, the solution to the heat equation is a bifractional Brownian motion.

We structured our work as follows. In Section 2, we recall some basic properties of the solution to the stochastic heat equation driven by a space-time white noise and we provide an analysis of the correlation of the spatial increments of this solution. We also deduce the asymptotic behavior of its corresponding quadratic variation. In Section 3, we study the Wishart matrix associated to an initial random matrix constructed from the increments in space of the solution to the stochastic heat equation. In particular, we evaluate and estimate the Wasserstein distance between the renormalized Wishart matrix and its limit distribution (the so-called GOE matrix). Finally, in Section 4, we introduce the basic elements and the most important tools on Wiener chaos and Malliavin calculus necessary to develop this work.

9.2 The stochastic heat equation and its quadratic variation

This part is devoted to the study of the solution to the stochastic heat equation with space-time white noise. We do a sharp analysis of the correlation of its spatial increments over the set of positive integers. We also deduce the limit behavior of the corresponding quadratic variation.

9.2.1 Basic properties of the solution

In this work, we will focus on the stochastic heat equation with space-time white noise

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + \dot{W}, & t > 0, x \in \mathbb{R} \\ u(0, x) = 0, & x \in \mathbb{R}, \end{cases} \quad (9.1)$$

where Δ denote the Laplacian operator on \mathbb{R} and $W = (W(t, A), t \geq 0, A \in \mathcal{B}_b(\mathbb{R}))$ is the space-time white noise, that is, W is a real-valued centered Gaussian field with covariance

$$\mathbb{E}[W(t, A)W(s, B)] = (t \wedge s)\lambda(A \cap B), \quad s, t \geq 0, A, B \in \mathcal{B}_b(\mathbb{R}), \quad (9.2)$$

where λ stands for the Lebesgue measure on \mathbb{R} and $\mathcal{B}_b(\mathbb{R})$ for the set of bounded Borel subsets of \mathbb{R} .

The starting point of the study of the stochastic partial differential equation (SPDE in the sequel) of the form (9.1) is the pioneering work by Walsh ([125]), where the author introduced the so-called random field approach in order to solve SPDEs with (possibly) nonlinear Gaussian noise. Since, this research direction has been expanding rapidly, and nowadays there exists a vast literature on SPDEs in general and on the stochastic heat equation in particular. We refer, among others, to the seminal paper [39] and to the monographs and surveys [29], [40], [78] or [117].

The solution of the heat equation (9.1), in the mild sense, is given by

$$u(t, x) = \int_0^t \int_{\mathbb{R}} G(t-s, x-y) W(ds, dy), \quad (9.3)$$

for every $x \in \mathbb{R}$ and $t \geq 0$, where the function G is the fundamental solution of the heat equation :

$$G(t, x) = \begin{cases} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{|x|^2}{2t}\right) & \text{if } t > 0, x \in \mathbb{R} \\ 0 & \text{if } t \leq 0, x \in \mathbb{R} \end{cases} \quad (9.4)$$

Let us recall the main properties of the solution (9.3). It is well-known that the Wiener integral (9.3) is well-defined as a square integrable random variable since the spatial dimension is equal 1, see Proposition 5.1 in [117].

We also know that if u is the solution of the stochastic heat equation, then its covariance is given by (see [117], Proposition 5.2) :

$$\mathbb{E}[u(t,x) u(s,y)] = \frac{1}{\sqrt{2\pi}} \int_0^{t \wedge s} \frac{1}{\sqrt{t+s-2u}} \exp\left(-\frac{(x-y)^2}{2(t+s-2u)}\right) du \quad (9.5)$$

for every $s, t \geq 0$ and for every $x, y \in \mathbb{R}$. This pretty complex expression will reduce to a simpler one if one fixes the time or the space variable. For other properties of the solution to the SPDE (9.1), we refer to Section 5 in [117]. In particular, let us notice that the solution is stationary with respect to its spatial variable (see Corollary 5.1 in [117]). For the regularity of the sample paths of this solution and its relationship with the fractional Brownian motion, see Proposition 5.13 and Theorem 5.2 in [117].

9.2.2 Limit behavior of the spatial quadratic variation

Intuitively, the situation modelled by heat equation can be seen by considering two rods, whose temperature is not the same. Denote by T_1 the high temperature of the first rod and by T_2 the low temperature along the second rod. If we put them end-to-end, then, T_1 will be lower, and T_2 will increase so that the temperature along these two rods will change. The stochastic heat equation models the temperature variation along a rod, under a random perturbation. As mentioned in the introduction, $u(t,x)$ allows us to give the temperature at a point x and at time t . We notice that in general the temperature at time t between the points x and $x+k$ is not the same because, at the beginning, when we put two rods together, the heat is not equally distributed. Whereas, over time, the temperature will tend to be constant.

In this part, we are going to study the properties of the spatial increments of the solution to the heat equation driven by a space time white noise at a fixed time t . This preliminary part will help us to analyze the renormalized Wishart matrix in the next section.

First of all, we compute the correlation of the spatial increments of the random field (9.3). Notice that, that when $t = s$, the equation (9.5) reduces to

$$\mathbb{E}[u(t,x)u(t,y)] = \frac{\sqrt{t}}{\sqrt{\pi}} e^{-\frac{|y-x|^2}{4t}} + \sqrt{2}|x-y|\operatorname{erf}\left(\frac{|x-y|}{2\sqrt{t}}\right) - \frac{1}{2}|x-y|, \quad (9.6)$$

where erf denotes the error function

$$\operatorname{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2} dz. \quad (9.7)$$

Lemma 14. *Let $t > 0$ be fixed and consider the random field $(u(t,x), t \geq 0, x \in \mathbb{R})$ given by (9.3). Then, we have*

1. For every $j \geq 0$,

$$\mathbb{E}(u(t, j+1) - u(t, j))^2 = A_t,$$

where A_t is a constant depending only on t defined by

$$A_t = \frac{2\sqrt{t}}{\sqrt{\pi}} \left(1 - e^{-\frac{1}{4t}}\right) - 2\sqrt{2}\operatorname{erf}\left(\frac{1}{2\sqrt{t}}\right) + 1. \quad (9.8)$$

2. For any integers $i, j \geq 0$, with $i > j$,

$$\mathbb{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j)) = f_{1,t}(i-j) + f_{2,t}(i-j),$$

where, for $k \in \mathbb{Z}$,

$$f_{1,t}(k) = \frac{\sqrt{t}}{\sqrt{\pi}} \left(2e^{-\frac{k^2}{4t}} - e^{-\frac{(k+1)^2}{4t}} - e^{-\frac{(k-1)^2}{4t}} \right), \quad (9.9)$$

and

$$f_{2,t}(k) = \sqrt{2} \left(2k \operatorname{erf}\left(\frac{k}{2\sqrt{t}}\right) - (k+1) \operatorname{erf}\left(\frac{k+1}{2\sqrt{t}}\right) - (k-1) \operatorname{erf}\left(\frac{k-1}{2\sqrt{t}}\right) \right). \quad (9.10)$$

Notice that $f_{i,t}(k) = f_{i,t}(-k)$ for every $k \in \mathbb{Z}$ and $i = 1, 2$.

Proof : It is an immediate consequence of (9.6). \square

We will need the following technical result, concerning the summability of the quantities $f_{1,t}(k), f_{2,t}(k)$ from the previous lemma.

Lemma 15. Let $f_{1,t}$ and $f_{2,t}$ be given by (9.9) and (9.10), respectively. Then,

$$K_{0,t} := \sum_{k \geq 1} (f_{1,t}(k) + f_{2,t}(k))^2 < \infty \quad \text{and} \quad K_{1,t} := \sum_{k \geq 1} k(f_{1,t}(k) + f_{2,t}(k))^2 < \infty. \quad (9.11)$$

Proof : First, notice that

$$|f_{1,t}(k)| \leq C_t e^{-\frac{(k-1)^2}{4t}}, \quad \text{for all } k \geq 1, \quad (9.12)$$

where C_t is a positive constant depending only on t . Now, by definition of the error function erf we have

$$\begin{aligned} |f_{2,t}(k)| &= \frac{1}{\sqrt{\pi}} \left| 2k \int_0^{\frac{k}{2\sqrt{t}}} e^{-z^2} dz - (k+1) \int_0^{\frac{k+1}{2\sqrt{t}}} e^{-z^2} dz - (k-1) \int_0^{\frac{k-1}{2\sqrt{t}}} e^{-z^2} dz \right| \\ &= \frac{1}{\sqrt{\pi}} \left| k \int_{\frac{k-1}{2\sqrt{t}}}^{\frac{k}{2\sqrt{t}}} e^{-z^2} dz - k \int_{\frac{k}{2\sqrt{t}}}^{\frac{k+1}{2\sqrt{t}}} e^{-z^2} dz - \int_{\frac{k-1}{2\sqrt{t}}}^{\frac{k+1}{2\sqrt{t}}} e^{-z^2} dz \right| \\ &\leq \frac{1}{\sqrt{\pi}} \left(k \int_{\frac{k-1}{2\sqrt{t}}}^{\frac{k+1}{2\sqrt{t}}} e^{-z^2} dz + \int_{\frac{k-1}{2\sqrt{t}}}^{\frac{k+1}{2\sqrt{t}}} e^{-z^2} dz \right) \\ &\leq \left(\frac{k+1}{\sqrt{\pi t}} \right) e^{-\frac{(k-1)^2}{4t}}. \end{aligned} \quad (9.13)$$

The conclusion is obtained via (9.12) and (9.13). \square

For $t > 0$ fixed, we define the sequence $(S_{N,t}, N \geq 1)$ given by

$$S_{N,t} = \frac{1}{N} \sum_{j=0}^{N-1} (u(t, j+1) - u(t, j))^2,$$

where u is the solution to the stochastic heat equation (9.1). Now, by Lemma 14, point 1, we have

$$\mathbb{E}[S_{N,t}] = \frac{1}{N} \sum_{j=0}^{N-1} \mathbb{E}(u(t, j+1) - u(t, j))^2 = A_t,$$

with A_t from (9.8). We can also write the random variable $S_{N,t}$ as an element of the Wiener chaos of order two with respect to the Gaussian process W , by applying the product formula (9.50). Actually, for $i = 1, \dots, N$,

$$u(t, i+1) - u(t, i) = I_1(g_{t,i})$$

with

$$g_{t,i}(s, y) = (G(t-s, i+1-y) - G(t-s, i-y)) 1_{[0,t]}(s), \quad (s, y) \in \mathbb{R}_+ \times \mathbb{R} \quad (9.14)$$

where G is the fundamental solution of the heat equation given by (9.4) and I_q denotes the multiple integral of order $q \geq 1$ with respect to the Gaussian process W . See the Appendix for the basic facts on multiple stochastic integrals. As a consequence of the product formula (6.51), we have

$$(u(t, i+1) - u(t, i))^2 = I_2(g_{t,i}^{\otimes 2}) + \mathbb{E}(u(t, i+1) - u(t, i))^2 = I_2(g_{t,i}^{\otimes 2}) + A_t.$$

Therefore,

$$S_{N,t} = \frac{1}{N} \sum_{j=0}^{N-1} [I_2(g_{t,i}^{\otimes 2}) + A_t] = \frac{1}{N} \sum_{j=0}^{N-1} I_2(g_{t,i}^{\otimes 2}) + A_t. \quad (9.15)$$

We give below the limit, almost surely and in $L^2(\Omega)$, of the sequence $S_{N,t}$, as $N \rightarrow \infty$.

Lemma 16. *For $t > 0$ fixed, as $N \rightarrow \infty$, the sequence $(S_{N,t}, N \geq 1)$ converges to A_t (given by (9.8)) in $L^2(\Omega)$ and almost surely. Moreover,*

$$N\mathbb{E}(S_{N,t} - A_t)^2 \longrightarrow 2A_t^2 + 4K_{0,t}, \quad \text{as } N \rightarrow \infty,$$

where A_t and $K_{0,t}$ are constants defined in the Lemmas 14 and 15, respectively.

Slightly abusing notation, we denote by $\langle \cdot, \cdot \rangle$, instead of $\langle \cdot, \cdot \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R})}$, the scalar product on the space $L^2(\mathbb{R}_+ \times \mathbb{R})$.

Proof : We will do the proof in several steps.

Step 1 : Expression of $\mathbb{E}(S_{N,t} - A_t)^2$. In fact,

$$\begin{aligned} \mathbb{E}(S_{N,t} - A_t)^2 &= \frac{1}{N^2} \mathbb{E} \left[\sum_{j=0}^{N-1} I_2(g_{t,i}^{\otimes 2}) \right]^2 = \frac{2}{N^2} \sum_{i,j=0}^{N-1} \langle g_{t,i}, g_{t,j} \rangle^2 \\ &= \frac{2}{N^2} \sum_{i,j=0}^{N-1} [\mathbb{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j))]^2 \\ &= \frac{2}{N^2} \sum_{i=0}^{N-1} [\mathbb{E}(u(t, i+1) - u(t, i))^2]^2 \\ &\quad + \frac{4}{N^2} \sum_{i,j=0; i>j}^{N-1} [\mathbb{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j))]^2 \\ &= s_{1,N} + s_{2,N}. \end{aligned} \quad (9.16)$$

Step 2 : Expression of $s_{1,N}$ and $s_{2,N}$. From Lemma 14, we get

$$s_{1,N} = \frac{2}{N^2} \sum_{i=0}^{N-1} [\mathbb{E}(u(t, i+1) - u(t, i))^2]^2 = \frac{2}{N^2} \sum_{i=0}^{N-1} A_t^2 = \frac{2A_t^2}{N}. \quad (9.17)$$

Once again, by Lemma 14, it follows that

$$\begin{aligned}
 s_{2,N} &= \frac{4}{N^2} \sum_{i,j=0;i>j}^{N-1} [f_{1,t}(i-j) + f_{2,t}(i-j)]^2 \\
 &= \frac{4}{N^2} \sum_{k=1}^{N-1} (N-k)(f_{1,t}(k) + f_{2,t}(k))^2 \\
 &= \frac{4}{N} \sum_{k=1}^{N-1} (f_{1,t}(k) + f_{2,t}(k))^2 - \frac{4}{N^2} \sum_{k=1}^{N-1} k(f_{1,t}(k) + f_{2,t}(k))^2 \\
 &= \frac{4}{N} s_{2,1,N} - \frac{4}{N^2} s_{2,2,N},
 \end{aligned}$$

where, $s_{2,1,N}$ and $s_{2,2,N}$ satisfy, respectively, that

$$\begin{aligned}
 s_{2,1,N} &:= \sum_{k=1}^{N-1} (f_{1,t}(k) + f_{2,t}(k))^2 \xrightarrow{N \rightarrow \infty} K_{0,t} \quad \text{and} \\
 s_{2,2,N} &:= \sum_{k=1}^{N-1} k(f_{1,t}(k) + f_{2,t}(k))^2 \xrightarrow{N \rightarrow \infty} K_{1,t},
 \end{aligned}$$

with $K_{0,t}$ and $K_{1,t}$ constants defined in the Lemma 15.

Step 3 : Convergence in $L^2(\Omega)$. By the steps 1 and 2, we have

$$\mathbb{E}(S_{N,t} - A_t)^2 = \frac{2A_t^2}{N} + \frac{4}{N} s_{2,1,N} - \frac{4}{N^2} s_{2,2,N} \xrightarrow{N \rightarrow \infty} 0. \quad (9.18)$$

Therefore,

$$N\mathbb{E}(S_{N,t} - A_t)^2 \xrightarrow{N \rightarrow \infty} 2A_t^2 + 4K_{0,t}.$$

For a later use, we also notice that, for N large enough,

$$\begin{aligned}
 N\mathbb{E}(S_{N,t} - A_t)^2 - (2A_t^2 + 4K_{0,t}) &= 4(s_{2,1,N} - K_{0,t}) - \frac{4}{N} s_{2,2,N} \\
 &\leq C_t \frac{1}{N}
 \end{aligned} \quad (9.19)$$

since, via (9.12) and (9.13), $|s_{2,1,N} - K_{0,t}| \leq C e^{-\frac{N^2}{4r}}$ for N large enough.

Step 4 : Almost sure convergence. The almost sure convergence follows by applying Borel-Cantelli lemma. Indeed, for $0 < \gamma < \frac{1}{2}$ and for $p \geq 1$ large enough, we have by using Tchebychev inequality and the hyper-contractivity property of multiple stochastic integrals (see (6.53)) :

$$\begin{aligned}
 \sum_{N \geq 1} P(|S_{N,t} - A_t| \geq N^{-\gamma}) &\leq \sum_{N \geq 1} N^{\gamma p} \mathbb{E}|S_{N,t} - A_t|^p \\
 &\leq \sum_{N \geq 1} N^{\gamma p} (\mathbb{E}|S_{N,t} - A_t|^2)^{\frac{p}{2}} \leq C_p \sum_{N \geq 1} N^{p(\gamma - \frac{1}{2})} < \infty.
 \end{aligned}$$

We have used the fact that $\mathbb{E}(S_{N,t} - A_t)^2 \leq \frac{C_t}{N}$ for N large enough, by (9.18). \square

9.2.3 Renormalized spatial quadratic variation and Central Limit Theorem

For $N \geq 1$ and $t > 0$ fixed, we define the sequence

$$V_{N,t} = \sqrt{N}(S_{N,t} - A_t), \quad (9.20)$$

where A_t is defined by Lemma 14. Now, from (9.15), we see that

$$V_{N,t} = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} I_2(g_{t,j}^{\otimes 2}) \quad (9.21)$$

Thus, $\mathbb{E}[V_{N,t}] = 0$ and considering the result in Lemma 16, we get

$$\mathbb{E}[V_{N,t}^2] = 2A_t^2 + 4s_{2,1,N} - \frac{4}{N}s_{2,2,N} \xrightarrow{N \rightarrow \infty} 2A_t^2 + 4K_{0,t} := a_t \quad (9.22)$$

and, by (9.19),

$$|\mathbb{E}[V_{N,t}^2] - a_t| \leq C_t \frac{1}{N}. \quad (9.23)$$

Above and in the sequel, C_t stands for a strictly positive constant that may change from line to line.

We will study the limiting distribution of the sequence (9.20). Actually, we will use Theorem 15 to show that this sequence satisfies a Central Limit Theorem. A similar result, along an equidistant partition of the interval $[0, 1]$, has been obtained in [5]. We will need the below auxiliary result.

Lemma 17. *Let $t > 0$ be fixed and $g_{t,i}, i = 0, \dots, N-1$ defined by (9.14). Then, for N large enough,*

$$\sum_{i,j,k,l=0}^{N-1} \langle g_{t,i}, g_{t,j} \rangle \langle g_{t,k}, g_{t,l} \rangle \langle g_{t,i}, g_{t,k} \rangle \langle g_{t,j}, g_{t,l} \rangle \leq C_t N,$$

where C_t is a positive constant depending on t .

Proof : By Lemma 14, for every $i, j = 0, \dots, N-1$, we have

$$\langle g_{t,i}, g_{t,i} \rangle = \mathbb{E}(u(t, i+1) - u(t, i))^2 = A_t, \text{ for all } i \geq 0, \quad (9.24)$$

and, for all $i > j$,

$$\langle g_{t,i}, g_{t,j} \rangle = \mathbb{E}(u(t, i+1) - u(t, i))(u(t, j+1) - u(t, j)) = f_{1,t}(i-j) + f_{2,t}(i-j) \quad (9.25)$$

with $f_{1,t}, f_{2,t}$ given by (9.9) and (9.10), respectively. Thus, by considering the decomposition

$$\sum_{i,j,k,l} = \sum_{i=j=k=l} + \sum_{\substack{\text{three equal} \\ \text{indices}}} + \sum_{\substack{\text{two equal} \\ \text{indices}}} + \sum_{i,j,k,l \text{ distinct}},$$

where,

$$\sum_{\substack{\text{three equal} \\ \text{indices}}} = \sum_{i=j=k \neq l} + \sum_{i=j=l \neq k} + \sum_{i=k=l \neq j} + \sum_{j=k=l \neq i} \quad (9.26)$$

and

$$\sum_{\substack{\text{two equal} \\ \text{indices}}} = \sum_{i=j \neq k \neq l} + \sum_{i=k \neq j \neq l} + \sum_{i=l \neq j \neq k} + \sum_{j=k \neq i \neq l} + \sum_{j=l \neq i \neq k} + \sum_{k=l \neq i \neq j}, \quad (9.27)$$

we can write

$$\sum_{i,j,k,l=0}^{N-1} \langle g_{t,i}, g_{t,j} \rangle \langle g_{t,k}, g_{t,l} \rangle \langle g_{t,i}, g_{t,k} \rangle \langle g_{t,j}, g_{t,l} \rangle = S_{4,N,t} + S_{3,N,t} + S_{2,N,t} + S_{1,N,t}, \quad (9.28)$$

where, from (9.24), the first term on the right-hand side gives

$$S_{4,N,t} = \sum_{i=0}^{N-1} \langle g_{t,i}, g_{t,i} \rangle^4 = \sum_{i=0}^{N-1} [\mathbb{E}(u(t, i+1) - u(t, i))^2]^4 = A_t^4 N. \quad (9.29)$$

By (9.26), the second term in (9.28) can be expressed as

$$S_{3,N,t} = S_{3,N,t}^1 + S_{3,N,t}^2 + S_{3,N,t}^3 + S_{3,N,t}^4,$$

and using a simple change of indices it is not difficult to show that

$$S_{3,N,t}^1 = S_{3,N,t}^2 = S_{3,N,t}^3 = S_{3,N,t}^4$$

Thus, by (9.24) and (9.25), and considering the change of indices $k = i - l$, we get the following estimate for N large,

$$\begin{aligned} S_{3,N,t} &= 4S_{3,N,t}^1 = 4 \sum_{i,l=0;i \neq l}^{N-1} \langle g_{t,i}, g_{t,i} \rangle^2 \langle g_{t,i}, g_{t,l} \rangle^2 = 4A_t^2 \sum_{i,l=0;i \neq l}^{N-1} (f_{1,t}(i-l) + f_{2,t}(i-l))^2 \\ &= 8A_t^2 \sum_{i,l=0;i > l}^{N-1} (f_{1,t}(i-l) + f_{2,t}(i-l))^2 = 8A_t^2 \sum_{l=0}^{N-2} \sum_{i=l+1}^{N-1} (f_{1,t}(i-l) + f_{2,t}(i-l))^2 \\ &= 8A_t^2 \sum_{k=1}^{N-1} (N-k)(f_{1,t}(k) + f_{2,t}(k))^2 \\ &\leq C_t N \sum_{k=1}^{N-1} (f_{1,t}(k) + f_{2,t}(k))^2 \leq C_t N, \end{aligned} \quad (9.30)$$

with the last bound being obtained via Lemma 15. Now, from (9.27), the third term in (9.28) can be decomposed as

$$S_{2,N,t} = S_{2,N,t}^1 + S_{2,N,t}^2 + S_{2,N,t}^3 + S_{2,N,t}^4 + S_{2,N,t}^5 + S_{2,N,t}^6,$$

and through a change of indices we have

$$S_{2,N,t}^1 = S_{2,N,t}^2 = S_{2,N,t}^5 = S_{2,N,t}^6 \quad \text{and} \quad S_{2,N,t}^3 = S_{2,N,t}^4.$$

Consequently,

$$S_{2,N,t} = 4S_{2,N,t}^1 + 2S_{2,N,t}^3, \quad (9.31)$$

where

$$\begin{aligned} S_{2,N,t}^1 &= 6A_t \sum_{i,k,l=0; i>k>l}^{N-1} \langle g_{t,k}, g_{t,l} \rangle \langle g_{t,i}, g_{t,k} \rangle \langle g_{t,i}, g_{t,l} \rangle \\ &= 6A_t \sum_{i,k,l=0; i>k>l}^{N-1} (f_{1,t}(k-l) + f_{2,t}(k-l))(f_{1,t}(i-k) + f_{2,t}(i-k)) \\ &\quad \times (f_{1,t}(i-l) + f_{2,t}(i-l)), \end{aligned}$$

and according to the change of indices $k-l = a, i-k = b$, so $i-l = a+b$, it results

$$\begin{aligned} S_{2,N,t}^1 &\leq 6A_t \sum_{a,b \geq 1} \sum_{l=0}^{N-1} |f_{1,t}(a) + f_{2,t}(a)| |f_{1,t}(b) + f_{2,t}(b)| |f_{1,t}(a+b) + f_{2,t}(a+b)| \\ &\leq C_t N \sum_{a,b \geq 1} |f_{1,t}(a) + f_{2,t}(a)| |f_{1,t}(b) + f_{2,t}(b)| |f_{1,t}(a+b) + f_{2,t}(a+b)| \\ &\leq C_t N. \end{aligned} \tag{9.32}$$

Note that we used the estimates (9.12) and (9.13) to justify that the sum over $a, b \geq 1$ is finite. We also see that

$$\begin{aligned} S_{2,N,t}^3 &= 6 \sum_{i,j,k=0; i>j>k}^{N-1} \langle g_{t,i}, g_{t,j} \rangle^2 \langle g_{t,i}, g_{t,k} \rangle^2 \\ &= 6 \sum_{i,j,k=0; i>j>k}^{N-1} (f_{1,t}(i-j) + f_{2,t}(i-j))^2 (f_{1,t}(i-k) + f_{2,t}(i-k))^2 \\ &\leq 6 \left(\sum_{i,j=0; i>j}^{N-1} (f_{1,t}(i-j) + f_{2,t}(i-j))^2 \right) \left(\sum_{a \geq 1} (f_{1,t}(a) + f_{2,t}(a))^2 \right) \\ &\leq C_t N, \end{aligned} \tag{9.33}$$

where the last bound is due to the Lemma 15 and estimate (9.30). Thereby, replacing (9.32) and (9.33) in (9.31), we have

$$S_{2,N,t} \leq C_t N. \tag{9.34}$$

Finally, let us analyze the fourth term $S_{1,N,t}$. In fact, by considering the change of indices $a = i-j, b = k-l, c = i-k$ and then $j-l = c-a+b$, we obtain

$$\begin{aligned} S_{1,N,t} &\leq C \sum_{i,j,k,l=0; i>j>k>l}^{N-1} (f_{1,t}(i-j) + f_{2,t}(i-j))(f_{1,t}(k-l) + f_{2,t}(k-l)) \\ &\quad \times (f_{1,t}(i-k) + f_{2,t}(i-k))(f_{1,t}(j-l) + f_{2,t}(j-l)) \\ &\leq CN \sum_{a,b,c \geq 1} |f_{1,t}(a) + f_{2,t}(a)| |f_{1,t}(b) + f_{2,t}(b)| \\ &\quad \times |f_{1,t}(c) + f_{2,t}(c)| |f_{1,t}(c-a+b) + f_{2,t}(c-a+b)| \\ &\leq C_t N, \end{aligned} \tag{9.35}$$

where $C > 0$ is a constant and with the last estimate obtained again via (9.12) and (9.13).

Therefore, by the estimates (9.29), (9.30), (9.34) and (9.35) we get the conclusion. \square

We show below that the sequence (9.20) satisfies a Central Limit Theorem. Recall that the notation d stands for any of the following distances : Kolmogorov, total variation, Wasserstein or Fortet-Mourier. Their definitions can be found in the Appendix.

Theorem 13. *Let $t > 0$ be fixed and consider the sequence $(V_{N,t}, N \geq 1)$ defined by (9.20). Then as $N \rightarrow \infty$, the sequence $V_{N,t}$ converges in distribution to the Gaussian law $\mathcal{N}(0, a_t)$, with a_t from (9.22). Moreover, for N large enough,*

$$d(V_{N,t}, \mathcal{N}(0, a_t)) \leq C \frac{1}{\sqrt{N}}.$$

Proof : We have, for $t > 0$ and $N \geq 1$, via (9.21), We have,

$$DV_{N,t} = \frac{2}{\sqrt{N}} \sum_{i=0}^{N-1} I_1(g_{t,i}) g_{t,i}$$

and

$$\begin{aligned} \|DV_{N,t}\|^2 &= \frac{4}{N} \sum_{i,j=0}^{N-1} I_1(g_{t,i}) I_1(g_{t,j}) \langle g_{t,i}, g_{t,j} \rangle \\ &= \frac{4}{N} \sum_{i,j=0}^{N-1} I_2(g_{t,i} \otimes g_{t,j}) \langle g_{t,i}, g_{t,j} \rangle + \mathbb{E} \|DV_{N,t}\|^2. \end{aligned}$$

Moreover,

$$\mathbb{E} \|DV_{N,t}\|^2 = 2\mathbb{E} V_{N,t}^2 \xrightarrow{N \rightarrow \infty} 2a_t.$$

In order to apply Theorem 15, we need to evaluate the following quantity

$$\sqrt{\mathbf{Var}(\|DV_{N,t}\|^2)} + |\mathbb{E}(\|DV_{N,t}\|^2 - 2a_t)| =: a_N + b_N.$$

Concerning b_N , we have by (9.23),

$$b_N = |\mathbb{E}(\|DV_{N,t}\|^2 - 2a_t)| = |2(\mathbb{E} V_{N,t}^2 - a_t)| \leq C_t \frac{1}{N}. \quad (9.36)$$

Now, we focus on a_N :

$$\begin{aligned} a_N^2 &= \mathbf{Var}(\|DV_{N,t}\|^2) = \mathbb{E} \left[(\|DV_{N,t}\|^2 - \mathbb{E} \|DV_{N,t}\|^2)^2 \right] \\ &= \frac{16}{N^2} \mathbb{E} \left(\sum_{i,j=0}^{N-1} I_2(g_{t,i} \otimes g_{t,j}) \langle g_{t,i}, g_{t,j} \rangle \right)^2 \\ &= \frac{32}{N^2} \sum_{i,j,k,l=0}^{N-1} \langle g_{t,i}, g_{t,j} \rangle \langle g_{t,k}, g_{t,l} \rangle \langle g_{t,i}, g_{t,k} \rangle \langle g_{t,j}, g_{t,l} \rangle. \end{aligned}$$

By applying Lemma 17, we obtain :

$$\mathbb{E} \left[(\|DV_{N,t}\|_{\mathcal{H}}^2 - \mathbb{E} [\|DV_{N,t}\|_{\mathcal{H}}^2])^2 \right] \leq \frac{C_t}{N}. \quad (9.37)$$

The conclusion is obtained by using (9.36) and (9.37), and by applying Theorem 15. \square

9.3 The Wishart matrix

Based on the results in Section 9.2, we study the Wishart matrix associated to an initial random matrix whose entries are the increments in space of the solution to the stochastic heat equation (9.3). For instance, when the spatial increments are considered, then the data included in the initial matrix models the evolution of the temperature between two consecutive points in the space, at a fixed time.

Let $W^{(i)}, 1 \leq i \leq n$ be n independent space-time white noises, characterized by the covariance function (9.2). Let us consider the random fields $u^{(i)}, 1 \leq i \leq n$ defined by

$$u^{(i)}(t, x) = \int_0^t \int_{\mathbb{R}} G(t-s, x-y) W^{(i)}(ds, dy), \quad t \geq 0, x \in \mathbb{R} \quad (9.38)$$

with the Green kernel G given by (9.4). That is, $u^{(i)}$ constitutes the mild solution to the stochastic heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2} \Delta u + \dot{W}^{(i)}$$

with vanishing initial condition.

Now, for $t > 0$ fixed, we consider the random matrix $\mathcal{X}_{n,d,t} = (X_{i,j,t})_{1 \leq i \leq n, 1 \leq j \leq d}$ whose entries are given by

$$X_{i,j,t} = u^{(i)}(t, j) - u^{(i)}(t, j-1), \quad 1 \leq i \leq n, 1 \leq j \leq d.$$

We can also write, for every $1 \leq i \leq n, 1 \leq j \leq d$,

$$\begin{aligned} X_{i,j,t} &= \int_0^t \int_{\mathbb{R}} [G(t-s, j-y) - G(t-s, j-1-y)] W^{(i)}(ds, dy) \\ &= I_1^{(i)}(g_{t,j-1}), \end{aligned}$$

where $g_{t,j}$ are defined by (9.14). In fact, the random matrix $\mathcal{X}_{n,d,t}$ gather on a same row i , at a fixed time t , all the variation of temperature between two successively spatial points j and $j+1$. All rows change because the random noise is different from one row to another. We can also notice that all the entries of $\mathcal{X}_{n,d,t}$ are centered Gaussian random variables and their variance is constant, i.e. (see Lemma 14)

$$\mathbb{E}(X_{i,j,t}^2) = A_t, \quad \text{for every } 1 \leq i \leq n, 1 \leq j \leq d$$

with A_t given by (9.8).

From Lemma 16, it follows that, as $d \rightarrow \infty$ and with n fixed, the usual Wishart matrix $\frac{1}{d} \mathcal{X}_{n,d,t} \mathcal{X}_{n,d,t}^T$ converges almost surely to $A_t \mathcal{I}_n$, where \mathcal{I}_n is the $n \times n$ identity matrix.

Lemma 18. *Let $t > 0$ be fixed and consider the Wishart matrix $\mathcal{W}_{n,d,t} = (W_{i,j,t})_{1 \leq i, j \leq n}$ defined by $\mathcal{W}_{n,d,t} = \frac{1}{d} \mathcal{X}_{n,d,t} \mathcal{X}_{n,d,t}^T$. Then, as $d \rightarrow \infty$, $\mathcal{W}_{n,d,t}$ converges componentwise to $A_t \mathcal{I}_n$ in $L^2(\Omega)$ and almost surely.*

Proof : Notice that the entries of the matrix $\mathcal{W}_{n,d,t}$ can be expressed as

$$W_{i,i,t} = \frac{1}{d} \sum_{k=1}^d X_{i,k,t}^2 = \frac{1}{d} \sum_{k=1}^d \left(u^{(i)}(t, k) - u^{(i)}(t, k-1) \right)^2$$

and for $i \neq j$,

$$W_{i,j,t} = \frac{1}{d} \sum_{k=1}^d X_{i,k,t} X_{j,k,t} = \frac{1}{d} \sum_{k=1}^d \left(u^{(i)}(t,k) - u^{(i)}(t,k-1) \right) \left(u^{(j)}(t,k) - u^{(j)}(t,k-1) \right).$$

Lemma 16 implies that, for every $1 \leq i \leq n$, $W_{i,i,t}$ converges to A_t as $d \rightarrow \infty$, almost surely and in $L^2(\Omega)$. Also, for $i \neq j$,

$$\mathbb{E}[W_{i,j,t}^2] = \frac{1}{d^2} \sum_{k,l=1}^d \mathbb{E}[X_{i,k,t} X_{i,l,t}] \mathbb{E}[X_{j,k,t} X_{j,l,t}] = \frac{1}{d^2} \sum_{k,l=1}^d f_t(k-l)^2,$$

where

$$f_t(k) = \begin{cases} f_{1,t}(k) + f_{2,t}(k), & \text{if } k \neq 0 \\ A_t & , \text{if } k = 0. \end{cases} \quad (9.39)$$

with $f_{1,t}$, $f_{2,t}$ and A_t defined by (9.9), (9.10) and (9.8), respectively. Thus, by Lemma 15, we get

$$\mathbb{E}[W_{i,j,t}^2] \leq C_t \frac{1}{d} \xrightarrow{d \rightarrow \infty} 0.$$

The almost sure convergence of the non-diagonal terms follows by a standard Borel-Cantelli argument as in the proof of Lemma 16. \square

Next, we consider the renormalized Wishart matrix $\overline{\mathcal{W}}_{n,d,t} = (\overline{W}_{i,j,t})_{1 \leq i,j \leq n}$ associated to $\mathcal{X}_{n,d,t}$ defined as :

$$\overline{\mathcal{W}}_{n,d,t} = \sqrt{d} \left(\frac{1}{d} \mathcal{X}_{n,d,t} \mathcal{X}_{n,d,t}^T - A_t \mathcal{I}_n \right). \quad (9.40)$$

The entries of the renormalized Wishart matrix are given by the following : for every $1 \leq i \leq n$,

$$\begin{aligned} \overline{W}_{i,i,t} &= \sqrt{d} \left(\frac{1}{d} \sum_{k=1}^d X_{i,k,t}^2 - A_t \right) \\ &= \sqrt{d} \left(\frac{1}{d} \sum_{k=1}^d \left(u^{(i)}(t,k) - u^{(i)}(t,k-1) \right)^2 - A_t \right) \\ &= \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} I_2^{(i)}(g_{t,k}^{\otimes 2}), \end{aligned}$$

and for every $1 \leq i, j \leq n$ with $i \neq j$,

$$\begin{aligned} \overline{W}_{i,j,t} &= \frac{1}{\sqrt{d}} \sum_{k=1}^d X_{i,k,t} X_{j,k,t} \\ &= \frac{1}{\sqrt{d}} \sum_{k=1}^d \left(u^{(i)}(t,k) - u^{(i)}(t,k-1) \right) \left(u^{(j)}(t,k) - u^{(j)}(t,k-1) \right) \\ &= \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} I_1^{(i)}(g_{t,k}) I_1^{(j)}(g_{t,k}) \end{aligned}$$

with $g_{t,k}$ from (9.14) and where $I_q^{(i)}$ means the multiple integral of order q with respect to the Gaussian process $W^{(i)}$. We notice the following facts :

- For every $1 \leq i, j \leq n$, the entries $\bar{W}_{i,j,t}$ are centered random variables.
- For every $1 \leq i, j \leq n$, with $i \neq j$, we have

$$\mathbb{E}\bar{W}_{i,i,t}^2 \xrightarrow{d \rightarrow \infty} a_t \text{ and } \mathbb{E}\bar{W}_{i,j,t}^2 = \frac{1}{2}\mathbb{E}\bar{W}_{i,i,t}^2 \xrightarrow{d \rightarrow \infty} \frac{1}{2}a_t,$$

with a_t defined by (9.22). Moreover, for d large enough,

$$\mathbb{E}|\bar{W}_{i,i,t} - a_t|^2 \leq c_t \frac{1}{d} \text{ and } \mathbb{E}\left|\bar{W}_{i,j,t} - \frac{1}{2}a_t\right|^2 \leq c_t \frac{1}{d}. \quad (9.41)$$

The first bound above follows directly from (9.19) and the second one can be obtained in a similar way.

- As a consequence of Theorem 13, one has that for every $1 \leq i \leq n$,

$$\bar{W}_{i,i,t} \xrightarrow{d \rightarrow \infty} \mathcal{N}(0, a_t). \quad (9.42)$$

One can also show directly that for every $1 \leq i, j \leq n$ with $i \neq j$,

$$\bar{W}_{i,j,t} \xrightarrow{d \rightarrow \infty} \mathcal{N}\left(0, \frac{1}{2}a_t\right), \quad (9.43)$$

but this convergence will also follow as a consequence of the results that follow below. Therefore, the renormalized Wishart matrix $\bar{\mathcal{W}}_{n,d,t}$ converges, component by component, as $d \rightarrow \infty$, to a random matrix whose entries are centered Gaussian random variables with variance a_t on the diagonal and $\frac{1}{2}a_t$ outside the diagonal. Such a matrix is called a GOE (Gaussian Orthogonal Ensemble) matrix. More precisely, the GOE matrix $\bar{\mathcal{Z}}_{n,t} = (\bar{Z}_{i,j,t})_{1 \leq i, j \leq n}$ is given by

$$\begin{cases} \bar{Z}_{i,i,t} \sim \mathcal{N}(0, a_t), & \text{for } 1 \leq i \leq n \\ \bar{Z}_{i,j,t} \sim \mathcal{N}(0, \frac{1}{2}a_t), & \text{for } 1 \leq i < j \leq n. \\ \bar{Z}_{j,i,t} = \bar{Z}_{i,j,t} & \text{if } 1 \leq j < i \leq n \end{cases} \quad (9.44)$$

and $(\bar{Z}_{i,j,t}, 1 \leq i \leq j \leq n)$ are independent random variables.

We will evaluate the Wasserstein distance, for n, d large enough (i.e. in the high-dimensional regime) between the renormalized Wishart matrix $\bar{\mathcal{W}}_{n,d,t}$ defined by (9.40) and the GOE matrix (9.44). We refer to Section 9.4.2 in the Appendix for the definition of the Wasserstein distance for random matrices and random vectors.

To follow the main result of this section. Recall that d_W denotes the Wasserstein distance.

Theorem 14. *For $t > 0$ fixed, consider the renormalized Wishart matrix $\bar{\mathcal{W}}_{n,d,t}$, given by (9.40). Then, for every $n \geq 1$, $\bar{\mathcal{W}}_{n,d,t}$ converges in distribution componentwise to the GOE matrix $\bar{\mathcal{Z}}_{n,t}$ given by (9.44), as d goes to infinity. Furthermore, there exists a positive constant C_t such that, for all $n, d \geq 1$ large enough,*

$$d_W(\bar{\mathcal{W}}_{n,d,t}, \bar{\mathcal{Z}}_{n,t}) \leq C_t \sqrt{\frac{n^3}{d}}. \quad (9.45)$$

Proof : We are going to apply the Theorem 1.2 from [98]. Then, initially, we are going to consider the framework to apply this theorem. Thus, let \mathcal{H} be a real separable Hilbert space with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and let $\{e_{i,j}, i, j \geq 1\} \subset \mathcal{H}$ be a family such that

$$\langle e_{i,j}, e_{i',j'} \rangle = \mathbf{1}_{\{i=i'\}} f_t(j - j')$$

where f_t is given by (9.39). Notice that from estimates (9.12) and (9.13) it is easy to show that $f_t \in l^{4/3}(\mathbb{Z})$ (the space of summable sequences), see Lemma 15. Let us also consider $\{B(h), h \in \mathcal{H}\}$ be an isonormal process, that is, a centered Gaussian process indexed by \mathcal{H} such that $\mathbb{E}[B(h_1)B(h_2)] = \langle h_1, h_2 \rangle_{\mathcal{H}}$ for all $h_1, h_2 \in \mathcal{H}$. In this way, we define the random matrix $\mathcal{C}_{n,d} = (C_{i,j})_{1 \leq i \leq n, 1 \leq j \leq d}$, whose entries are given by $C_{i,j} = B(e_{i,j})$, and consider its associated renormalized random matrix

$$\mathcal{W}_{n,d,\mathcal{C}} = \sqrt{d} \left(\frac{1}{d} \mathcal{C}_{n,d} \mathcal{C}_{n,d}^T - A_t \mathcal{I}_n \right).$$

From Lemma 14 and (9.40) we observe that

$$\mathcal{X}_{n,d,t} \stackrel{(d)}{=} \mathcal{C}_{n,d} \quad \text{and} \quad \overline{\mathcal{W}}_{n,d,t} \stackrel{(d)}{=} \mathcal{W}_{n,d,\mathcal{C}}.$$

Let us also introduce a Gaussian matrix $\mathcal{G}_n = (G_{i,j})_{1 \leq i, j \leq n}$ with the same covariance as the matrices $\overline{\mathcal{W}}_{n,d,t}$ and $\mathcal{W}_{n,d,\mathcal{C}}$, i.e. for every $1 \leq i, j, k, l \leq n$,

$$\mathbb{E}[G_{i,j}G_{k,l}] = \mathbb{E}[\overline{W}_{i,j,t}\overline{W}_{k,l,t}] = \mathbb{E}[W_{i,j,C}W_{k,l,C}],$$

where $\mathcal{W}_{n,d,C} = (W_{i,j,C})_{1 \leq i, j \leq n}$. Then, we have

$$\begin{aligned} d_W(\overline{\mathcal{W}}_{n,d,t}, \overline{\mathcal{Z}}_{n,t}) &= d_W(\mathcal{W}_{n,d,C}, \overline{\mathcal{Z}}_{n,t}) \\ &\leq d_W(\mathcal{W}_{n,d,C}, \mathcal{G}_n) + d_W(\mathcal{G}_n, \overline{\mathcal{Z}}_{n,t}). \end{aligned}$$

Therefore, by the framework above, we apply Theorem 1.2 from [98] and obtain

$$d_W(\overline{\mathcal{W}}_{n,d,t}, \mathcal{G}_n) \leq C_t \sqrt{\frac{n^3}{d}}. \quad (9.46)$$

Now, by (9.58) and the definition of the Wasserstein distance,

$$\begin{aligned} d_W(\mathcal{G}_n, \overline{\mathcal{Z}}_{n,t}) &\leq \sqrt{2} d_W(\mathcal{G}_n^{half}, \mathcal{Z}_{n,t}^{half}) \\ &\leq \mathbb{E} \left(\sum_{i,j=1}^n |G_{i,j} - \overline{Z}_{i,j,t}|^2 \right)^{\frac{1}{2}} \leq \left(\sum_{i,j=1}^n \mathbb{E} [|G_{i,j} - \overline{Z}_{i,j,t}|^2] \right)^{\frac{1}{2}}. \end{aligned}$$

By (9.41), for every $1 \leq i, j \leq n$,

$$\mathbb{E} [|G_{i,j} - \overline{Z}_{i,j,t}|^2] \leq c \frac{1}{d}$$

and therefore

$$d_W(\mathcal{G}_n, \overline{\mathcal{Z}}_{n,t}) \leq c \sqrt{\frac{n^2}{d}}. \quad (9.47)$$

By (9.46) and (9.47), we conclude the proof. \square

A similar estimate as in (9.45) has been obtained in [53] in the case of the stochastic wave equation with space-time white noise. A related open problem is the behavior of the Wishart matrix when the entries of the starting matrix are the increments of the solution to stochastic heat equation driven by a Gaussian noise with correlation in time and/or in space. Also, the question of the functional convergence of the matrix-valued process $(\mathcal{W}_{n,d,t}, t \in [0, T])$ may be of interest (a similar problem has been considered in [18]).

9.4 Appendix

Here we recall the basic elements on Wiener chaos and Malliavin calculus as well as some standard results on the normal approximation on Wiener space.

9.4.1 Basics of the Malliavin calculus

Let $T = \mathbb{R}_+ \times \mathbb{R}$ and consider the Hilbert space $H := L^2(T, \mathcal{B}(T), \lambda)$ where we denoted by λ the Lebesgue measure and $\mathcal{B}(T)$ stands for the Borel subsets of T . Let $L_S^2(T^p)$ be the set of real-valued symmetric square integrable functions on T^p . Let $(B_t)_{t \in T}$ be a Wiener process. Denote by $B(\varphi) := \int_T \varphi_s dB_s$ the Wiener integral of $\varphi \in H$ with respect to the Brownian motion B . The family $(B(\varphi), \varphi \in H)$ forms an isonormal process, i.e. a Gaussian family of centered random variables such that

$$\mathbb{E}B(\varphi_1)B(\varphi_2) = \langle \varphi_1, \varphi_2 \rangle_H = \int_T \varphi_1(s)\varphi_2(s)ds$$

for any $\varphi_1, \varphi_2 \in H$.

Denote I_n the multiple stochastic integral of order n with respect to B (see [100]). This I_n is actually an isometry between the Hilbert space $H^{\otimes n}$ (symmetric tensor product) equipped with the scaled norm $\sqrt{n!} \|\cdot\|_{H^{\otimes n}}$ and the Wiener chaos of order n which is defined as the closed linear span of the random variables $H_n(B(\varphi))$ where $\varphi \in H, \|\varphi\|_H = 1$ and H_n is the Hermite polynomial of degree $n \geq 1$

$$H_n(x) = (-1)^n \exp\left(\frac{x^2}{2}\right) \frac{d^n}{dx^n} \left(\exp\left(-\frac{x^2}{2}\right) \right), \quad x \in \mathbb{R}. \quad (9.48)$$

The isometry of multiple integrals can be written as : if \tilde{f} denotes the symmetrization of the function f , for m, n positive integers,

$$\begin{aligned} \mathbb{E}(I_n(f)I_m(g)) &= n! \langle \tilde{f}, \tilde{g} \rangle_{H^{\otimes n}} \quad \text{if } m = n, \\ \mathbb{E}(I_n(f)I_m(g)) &= 0 \quad \text{if } m \neq n. \end{aligned} \quad (9.49)$$

We will use the product formula for multiple stochastic integrals, if $f \in L_S^2(T^m)$ and $g \in L_S^2(T^n)$, then

$$I_m(f)I_n(g) = \sum_{r=0}^{m \wedge n} r! \binom{m}{r} \binom{n}{r} I_{m+n-2r}(f \otimes_r g) \quad (9.50)$$

where for $r = 0, \dots, m \wedge n$, the contraction $f \otimes_r g$ is the function in $L^2(T^{m+n-2r})$ given by

$$\begin{aligned} (f \otimes_r g)(t_1, \dots, t_{m+n-2r}) &= \int_{T^r} f(u_1, \dots, u_r, t_1, \dots, t_{m-r}) \\ &\quad g(u_1, \dots, u_r, t_{m-r+1}, \dots, t_{m+n-2r}) du_1 \dots du_r. \end{aligned} \quad (9.51)$$

Notice that $f \otimes_r g$ is not necessarily a symmetric function (even if f, g are symmetric) and we will denote by $f \tilde{\otimes}_r g$ its symmetrization.

An useful property of finite sums of multiple integrals is the hypercontractivity. Namely, if $F = \sum_{k=0}^n I_k(f_k)$ with $f_k \in \mathcal{H}^{\otimes k}$ then

$$\mathbb{E}|F|^p \leq C_p (\mathbb{E}F^2)^{\frac{p}{2}}. \quad (9.52)$$

for every $p \geq 2$, where the constant $C_p > 0$ depends only on p .

Let \mathcal{S} be the class of smooth functionals of the form

$$F = f(B_{t_1}, \dots, B_{t_n}), \quad t_1, \dots, t_n \in T, \quad (9.53)$$

with $f \in C^\infty(\mathbb{R}^n)$ with at most polynomial growth (for f and its derivatives). For the random variable (9.53) we define its Malliavin derivative with respect to B by

$$D_t F = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(B_{t_1}, \dots, B_{t_n}) 1_{[0, t_i]}(t), \quad t \in T.$$

The operator D is an unbounded closable operator and it can be extended to the closure of \mathcal{S} with respect to the Malliavin-Sobolev norm

$$\|F\|_{k,p}^p = \mathbb{E}|F|^p + \sum_{j=1}^k \mathbb{E}\|D^{(j)}F\|_{L^2(T^j)}^p, \quad F \in \mathcal{S}, p \geq 2, k \geq 1. \quad (9.54)$$

where $D^{(j)}$ stands for the j th iterated Malliavin derivative. This closure will be denoted by $\mathbb{D}^{k,p}$.

The Malliavin derivative D acts on the Wiener chaos as an annihilation operator : if $F = I_n(f)$ with $f \in L^2(T^n)$ symmetric, then $D_t F = nI_{n-1}(f(\cdot, t))$ where “ \cdot ” stands for $n-1$ variables in T .

9.4.2 Normal approximation and Malliavin calculus

We use in this paper a basic result concerning the normal approximation on Wiener space. Let us first recall the definition of some distances between random variables. Usually, the distance between the laws two of real-valued random variables F and G is defined as

$$d(F, G) = \sup_{h \in \mathcal{A}} |\mathbb{E}h(F) - \mathbb{E}h(G)|, \quad (9.55)$$

where \mathcal{A} is a class of functions. When \mathcal{A} is the set of Lipschitz continuous function $h : \mathbb{R} \rightarrow \mathbb{R}$ such that $\|h\|_{Lip} \leq 1$, where

$$\|h\|_{Lip} = \sup_{x, y \in \mathbb{R}^d, x \neq y} \frac{|h(x) - h(y)|}{|x - y|}$$

then (9.55) gives the Wasserstein distance (or the so-called 1-Wasserstein distance). When \mathcal{A} is the set of indicator functions $\{1_{(-\infty, z]}, z \in \mathbb{R}\}$ the (9.55) gives the Kolmogorov distance while for $\mathcal{A} = \{1_B, B \in \mathcal{B}(\mathbb{R})\}$ we have the total variation distance. The reader may consult Appendix C in [93] for a more complete exposition.

The first one is a classical result in Stein-Malliavin calculus (see e.g. [93]).

Theorem 15. Fix $q \geq 1$. Assume that $(G_N)_{N \geq 1} = (I_q(g_N))_{N \geq 1}$ with $g_N \in \mathcal{H}^{\odot q}$, a sequence of random variables belonging to the q th Wiener chaos such that :

$$\mathbb{E}(G_N^2) \xrightarrow{N \rightarrow \infty} \sigma^2.$$

Then, G_N converges in law to $Z \sim \mathcal{N}(0, \sigma^2)$ if and only if

$$\|DG_N\|_{\mathcal{H}}^2 \xrightarrow{N \rightarrow \infty} q\sigma^2.$$

Moreover,

$$d(G_N, \mathcal{N}(0, \sigma^2)) \leq C \left(\sqrt{\mathbf{Var}(\|DG_N\|_{\mathcal{H}}^2)} + \left| \mathbb{E}(\|DG_N\|_{\mathcal{H}}^2) - q\sigma^2 \right| \right). \quad (9.56)$$

It is also possible to define the Wasserstein distance for random matrices and random vectors. To follow we recall these definitions as well as an important result with the relation between both distances.

Let \mathcal{X}, \mathcal{Y} be two random matrices with values in $\mathcal{M}_n(\mathbb{R})$, $n \geq 1$. We will denote by d_W the Wasserstein distance between the probability distributions of \mathcal{X} and \mathcal{Y} . That is,

$$d_W(\mathcal{X}, \mathcal{Y}) = \sup_{\|g\|_{Lip} \leq 1} |\mathbb{E}g(\mathcal{X}) - \mathbb{E}g(\mathcal{Y})|,$$

where the Lipschitz norm $\|g\|_{Lip}$ of $g: \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}$ is defined by

$$\|g\|_{Lip} = \sup_{A \neq B, A, B \in \mathcal{M}_n(\mathbb{R})} \frac{|g(A) - g(B)|}{\|A - B\|_{HS}},$$

with $\|\cdot\|_{HS}$ denoting the Hilbert-Schmidt norm on $\mathcal{M}_n(\mathbb{R})$.

We also recall that if X, Y are two n -dimensional random vectors, then the Wasserstein distance between them is defined to be

$$d_W(X, Y) = \sup_{\|g\|_{Lip} \leq 1} |\mathbb{E}(g(X)) - \mathbb{E}(g(Y))|,$$

where the Lipschitz norm $\|\cdot\|_{Lip}$ of $g: \mathbb{R}^n \rightarrow \mathbb{R}$ is defined by

$$\|g\|_{Lip} = \sup_{x \neq y \in \mathbb{R}^n} \frac{|g(x) - g(y)|}{\|x - y\|_{\mathbb{R}^n}},$$

with $\|\cdot\|_{\mathbb{R}^n}$ denoting the Euclidean norm on \mathbb{R}^n .

If $\mathcal{X} = (X_{i,j})_{1 \leq i, j \leq n}$ is an $n \times n$ symmetric random matrix, we associate to it its "half-vector" defined to be the $n(n+1)/2$ -dimensional random vector

$$X^{half} = (X_{1,1}, X_{1,2}, \dots, X_{1,n}, X_{2,2}, X_{2,3}, \dots, X_{2,n}, \dots, X_{n,n}). \quad (9.57)$$

Thereby, we have the below result proven in [98] :

Lemma 19. *If \mathcal{X}, \mathcal{Y} are two symmetric random matrices with values in $\mathcal{M}_n(\mathbb{R})$ then*

$$d_W(\mathcal{X}, \mathcal{Y}) \leq \sqrt{2} d_W(X^{half}, Y^{half}), \quad (9.58)$$

where X^{half}, Y^{half} are the associated half-vectors defined in (8.31).

Due to the bound (9.58), the analysis of the distance between two symmetric random matrices can be shifted to the study of the distance between their corresponding random vectors.

Limit behavior in high-dimensional regime for Wishart tensors with Rosenblatt entries

This article is submitted.

Joint work with Ciprian A. Tudor.

Abstract. We consider an initial $n \times d$ random matrix with non-Gaussian correlated entries on each row and independent entries from one row to another. The correlation on the rows is given by the correlation of the increments of the Rosenblatt process, which is a non-Gaussian self-similar process with stationary increments, living in the second Wiener chaos. To this initial matrix, we associate a Wishart tensor of length $p \geq 2$. We study the limit behavior in distribution of this Wishart tensor in the high-dimensional regime, i.e. when n, d are large enough. We prove that the vector corresponding to the p -Wishart tensor converges to a n^p -dimensional non-Gaussian vector, with Rosenblatt random variables on its hyper-diagonals and zeros outside the hyper-diagonals.

10.1 Introduction

The class of Wishart matrices constitutes a particular class of random matrices. If $\mathcal{X}_{n,d}$ is a $n \times d$ random matrix, then the Wishart matrix associated to $\mathcal{X}_{n,d}$ is given by $\mathcal{W}_{n,d} = \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T$ where \mathcal{X}^T stands for the transpose of the matrix \mathcal{X} . The Wishart matrices are empirical covariance matrices and so they are natural estimators in many practical applications. They have been introduced in [127] and then studied by many authors. We refer to the surveys [17], [70], [105] for the properties and the applications of the random matrices in general and of the Wishart matrices in particular.

The Wishart tensors are natural extensions of the Wishart matrices. If $p \geq 2$, the p -Wishart tensor associated to the initial matrix $\mathcal{X}_{n,d}$ is defined by

$$\sum_{j=1}^d \left(\mathbb{X}_j^{\otimes p} - \mathbb{E} \mathbb{X}_j^{\otimes p} \right), \tag{10.1}$$

where \mathbb{X}_j is the j th column of the matrix $\mathcal{X}_{n,d}$ and $\mathbb{X}_j^{\otimes p}$ denotes the p -tensor product of \mathbb{X}_j . We will mainly use the vectorial form of the Wishart tensor, i.e.

$$\mathbb{Y} = (Y_{j_1, \dots, j_p}, j_1, \dots, j_p = 1, \dots, n),$$

where for every $j_1, \dots, j_p = 1, \dots, n$,

$$Y_{j_1, \dots, j_p} = \sum_{i=1}^d \left(\prod_{k=1}^p X_{j_k, i} - \mathbb{E} \prod_{k=1}^p X_{j_k, i} \right) \quad (10.2)$$

When $p=2$, the Wishart tensor coincides with the centered Wishart matrix $\mathcal{X}_{n,d} \mathcal{X}_{n,d}^T - \mathbb{E} \mathcal{X}_{n,d} \mathcal{X}_{n,d}^T$. Recently, the random tensors found applications in machine learning (see [4] and [38]). A question of interest for many researchers in the last decades was to understand the asymptotic behavior of the vector \mathbb{Y} when the sizes n and d of the initial matrix tend to infinity, i.e. in the so-called high-dimensional regime. For $p=2$, there exists nowadays an important number of works dealing with the asymptotic behavior of the Wishart matrices in high-dimensional regime. Ones of the first studies concerning the normal approximation for the Wishart matrix was initiated in [20] and in [68]. In these references, the authors considered that the entries of the starting matrix are i.i.d. standard Gaussian random variables and they have shown that the limiting matrix of the renormalized Wishart matrix is a Gaussian matrix when $\frac{n^3}{d}$ converges to zero (see also [106] for related results on the phase transition). Then, these results have been extended by taking non-Gaussian entries and by relaxing the full independence hypothesis. In [21], [19], [46], [49], the authors assumed more complex distribution for the entries (such as log-concave distribution, or in a Wiener chaos, or expressed as Skorohod integrals or even with a very general law having the sixth moment finite), and they found that the Wishart matrix also converges to a Gaussian matrix. The convergence of the Wishart matrix to a Gaussian matrix still holds when the entries are (totally or partially) correlated but the correlation have to be weak enough, we refer to [98], [18] and [90] for a partial correlation, and to [97] for an overall correlation. Non-Gaussian limits may occur in the case of a relatively strong correlation of the initial entries (see e.g. [19], [18], [45], [98], for some examples of non-Gaussian behavior).

In the case of p -Wishart tensors with $p \geq 3$ there are relatively few results concerning their limit behavior in distribution in high-dimensional regime. The case of initial entries (independent or correlated) with Gaussian distribution has been treated in [98] and [97] while the case of column-independent entries with log-concave distribution has been considered in [90]. It was shown that the corresponding Wishart tensor is close in distribution, when d, n are large enough, to a Gaussian tensor and the Wasserstein distance between the p -Wishart tensor and the limiting Gaussian tensor is of order less than $\sqrt{\frac{n^{2p-1}}{d}}$, extending a well-known result in the case $p=2$. Other results concerning various aspects of the Wishart tensors have been given in [3], [69], [85], [124].

In this work, we deal with a class of Wishart tensors of length $p \geq 2$ constructed from a particular starting matrix with non-Gaussian and correlated entries. That is, we assume that the elements of the starting matrix $\mathcal{X}_{n,d}$ on the i th row ($1 \leq i \leq n$) are increments of the Rosenblatt process $Z^{H,i}$, where $Z^{H,1}, \dots, Z^{H,n}$ are independent Rosenblatt processes. We recall that the Rosenblatt process is a non-Gaussian self-similar process with stationary increments, which lives in the second Wiener chaos. It has the same covariance and correlation structure as the fractional Brownian motion (fBm). Therefore, our initial matrix contains correlated entries on each row and these entries are independent from one row to another. Moreover, the correlation on the rows is the same as in the case when the entries are the increments of the fBm (as considered in [97],

[98]) but the non-Gaussian character will lead to a different behavior of the associated tensor. We actually prove that in high-dimensional regime, the tensor constructed from such an initial matrix converges in distribution to a Rosenblatt tensor, which has Rosenblatt random variables on hyper-diagonals (the terms in (10.2) with $j_1 = \dots = j_p$) and zeros outside its hyper-diagonals. Our proofs exploit the properties of the Wiener chaos, a product formula for several multiple stochastic integrals and a sharp analysis of the multi-contractions involving the kernel of the Rosenblatt process. In particular, we obtain results concerning the behavior of the p -variation of the Rosenblatt process, which is new and it has its own interest.

The structure of our work is as follows. Section 2 contains some basics facts concerning the Wiener chaos and the Rosenblatt process. In Section 3 we describe our starting random matrix while in Section 4 we study the limit behavior of the p -variation of the Rosenblatt process. Based on the results obtained in Section 4, we get in Section 5 the asymptotic behavior, in high-dimensional regime, of the p -Wishart tensor with Rosenblatt entries. We show that this tensor is close, when n, d are large enough, to a Rosenblatt-type tensor and we also obtain the rate of convergence under the Wasserstein distance for this limit theorem.

By $C, c(H), c(H, p), c(H, p, \alpha)$ we denote generic strictly positive constants depending on the parameters appearing in their notation.

10.2 Preliminaries

In this preliminary part, we introduce the basics of the analysis on Wiener chaos needed in our work as well as the Rosenblatt process.

10.2.1 Wiener chaos and multiple stochastic integrals

The basic tools from the analysis on Wiener space are presented in this section. We will focus on some elementary facts about multiple stochastic integrals. We refer to [100] or [93] for a complete review on the topic.

Consider a real separable infinite-dimensional Hilbert space \mathcal{H} with its associated inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, and $(B(\varphi), \varphi \in \mathcal{H})$ an isonormal Gaussian process on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, which is a centered Gaussian family of random variables such that $\mathbb{E}(B(\varphi)B(\psi)) = \langle \varphi, \psi \rangle_{\mathcal{H}}$ for every $\varphi, \psi \in \mathcal{H}$. Denote by I_q the q th multiple stochastic integral with respect to B , which is an isometry between the Hilbert space $\mathcal{H}^{\odot q}$ (symmetric tensor product) equipped with the scaled norm $\sqrt{q!} \|\cdot\|_{\mathcal{H}^{\otimes q}}$ and the Wiener chaos of order q , which is defined as the closed linear span of the random variables $H_q(B(\varphi))$ where $\varphi \in \mathcal{H}$, $\|\varphi\|_{\mathcal{H}} = 1$ and H_q is the Hermite polynomial of degree $q \geq 1$ defined by

$$H_q(x) = \frac{(-1)^q}{q!} \exp\left(\frac{x^2}{2}\right) \frac{d^q}{dx^q} \left(\exp\left(-\frac{x^2}{2}\right) \right), \quad x \in \mathbb{R}. \quad (10.3)$$

The isometry property of multiple integrals can be written as follows : for $p, q \geq 1, f \in \mathcal{H}^{\otimes p}$ and $g \in \mathcal{H}^{\otimes q}$

$$\mathbb{E}\left(I_p(f)I_q(g)\right) = \begin{cases} q! \langle \tilde{f}, \tilde{g} \rangle_{\mathcal{H}^{\otimes q}} & \text{if } p = q, \\ 0 & \text{otherwise,} \end{cases} \quad (10.4)$$

where \tilde{f} stands for the symmetrization of f . We have the following product formula : if $f \in \mathcal{H}^{\otimes p}$ and $g \in \mathcal{H}^{\otimes q}$, then

$$I_p(f)I_q(g) = \sum_{r=0}^{p \wedge q} r! C_q^r C_p^r I_{p+q-2r}(f \otimes_r g), \quad (10.5)$$

where for $r = 0, \dots, p \wedge q$, if $\mathcal{H} = L^2(T)$, the contraction $f \otimes_r g$ is the function in $L^2(T^{p+q-2r})$ given by, for $r = 1, \dots, n$,

$$\begin{aligned} & (f \otimes_r g)(t_1, \dots, t_{p+q-2r}) \\ &= \int_{T^r} f(u_1, \dots, u_r, t_1, \dots, t_{p-r}) g(u_1, \dots, u_r, t_{p-r+1}, \dots, t_{p+q-2r}) du_1 \dots du_r \end{aligned} \quad (10.6)$$

and $f \otimes_0 g = f \otimes g$, the tensor product.

In our work, we have $\mathcal{H} = L^2(\mathbb{R})$ while the role of the isonormal process $(\mathbf{B}(\boldsymbol{\varphi}), \boldsymbol{\varphi} \in \mathcal{H})$ is played by the usual Wiener integral on $L^2(\mathbb{R})$ associated with the Wiener process $(\mathbf{B}(y), y \in \mathbb{R})$.

10.2.2 A general product formula

Fix two integers $n, q \geq 2$. Let h_1, \dots, h_n be n functions in $L_S^2(\mathbb{R}^q)$ (the set of symmetric functions in $L^2(\mathbb{R}^q)$). The purpose is to write the product

$$I_q(h_1) \dots I_q(h_n)$$

as a sum of multiple stochastic integrals. Consider the following notation :

— By $\mathcal{A}_{n,q}$ we denote the set of multi-indices $\boldsymbol{\alpha} = (\alpha_{ij}, 1 \leq i < j \leq n)$ such that

$$\sum_{1 \leq i < j \leq n} \alpha_{ij} \mathbf{1}_{k \in \{i,j\}} \leq q \quad \text{for every } k = 1, \dots, n.$$

The integer number α_{ij} constitutes the number of variables of h_i which are contracted with h_j for $1 \leq i < j \leq n$. Set

$$|\boldsymbol{\alpha}| = \sum_{1 \leq i < j \leq n} \alpha_{ij}.$$

— Let, for $k = 1, \dots, n$,

$$\beta_k^0 = q - \sum_{1 \leq i < j \leq n} \alpha_{ij} \mathbf{1}_{k \in \{i,j\}} \quad \text{and} \quad \beta_k = \sum_{j=0}^k \beta_j^0. \quad (10.7)$$

The value β_k^0 represents the number of variables of the kernel h_k which are not contracted with any of the other variables. Notice that

$$\sum_{k=1}^n \beta_k^0 = nq - 2|\boldsymbol{\alpha}|.$$

— For $1 \leq i < j \leq n$, let

$$u^{i,j} = \boldsymbol{\alpha}_{\min(i,j) \max(i,j)}$$

and consider the multi-contraction

$$\begin{aligned} & \otimes_{\boldsymbol{\alpha}}(h_1, \dots, h_n)(\xi_1, \dots, \xi_{nq-2|\boldsymbol{\alpha}|}) \\ &= \int_{\mathbb{R}^{|\boldsymbol{\alpha}|}} \prod_{k=1}^n h_k \left(s_1^{k,1}, \dots, s_{u^{k,1}}^{k,1}, \dots, s_1^{k,n}, \dots, s_{u^{k,n}}^{k,n}, \xi_{1+\beta_{k-1}}, \dots, \xi_{\beta_k} \right) \prod_{1 \leq i < j \leq n} ds_1^{i,j} \dots ds_{u^{i,j}}^{i,j}. \end{aligned} \quad (10.8)$$

Then, the following product formula holds

$$I_q(h_1) \dots I_q(h_n) = \sum_{\alpha \in \mathcal{A}_{n,q}} C_\alpha I_{nq-2|\alpha|}(\otimes_\alpha(h_1, \dots, h_n)), \quad (10.9)$$

with

$$C_\alpha = \frac{q!^n}{\prod_{k=1}^n \beta_k^0! \prod_{1 \leq i < j \leq n} \alpha_{ij}!}.$$

The details of the proof of (10.9) can be found in [58].

10.2.3 The Rosenblatt process

Let $(Z_t^H, t \geq 0)$ be a Rosenblatt process with self-similarity index $H \in (\frac{1}{2}, 1)$. For every $t \geq 0$, the random variable Z_t^H is defined by

$$Z_t^H = d(H) \int_{\mathbb{R}^2} \left(\int_0^t (u-y_1)_+^{\frac{H}{2}-1} (u-y_2)_+^{\frac{H}{2}-1} du \right) dB(y_1) dB(y_2), \quad (10.10)$$

where $(B(y), y \in \mathbb{R})$ is a (two-sided) Brownian motion, $d(H)$ is a strictly positive normalizing constant and for $a \in \mathbb{R}$, $x_+^a = x^a$ if $x > 0$ and $x_+^a = 0$ otherwise. Alternatively, we can write, for every $t \geq 0$,

$$Z_t^H = I_2(L_t^H), \quad (10.11)$$

where I_q denotes the multiple integral of order $q \geq 1$ with respect to B and L^H is the kernel of the Rosenblatt process given by, for every $y_1, y_2 \in \mathbb{R}$,

$$L_t^H(y_1, y_2) = d(H) \int_0^t (u-y_1)_+^{\frac{H}{2}-1} (u-y_2)_+^{\frac{H}{2}-1} du. \quad (10.12)$$

A more detailed presentation of this stochastic process can be found in [103], [116] or [117]. Let us recall that Z^H is a H -self-similar stochastic process with stationary increments and long memory. Its sample paths are Hölder continuous of order δ for every $\delta \in (0, H)$.

The Rosenblatt process shares the same covariance as the fractional Brownian motion, i.e. for every $s, t \geq 0$,

$$\mathbb{E} Z_t^H Z_s^H = \frac{1}{2} (t^{2H} + s^{2H} - |t-s|^{2H}). \quad (10.13)$$

It follows from (10.13) that for every integer $i \geq 0$,

$$\mathbb{E} (Z_{i+1}^H - Z_i^H) (Z_{j+1}^H - Z_j^H) = \rho_H(i-j), \quad (10.14)$$

with

$$\rho_H(v) = \frac{1}{2} (|v+1|^{2H} + |v-1|^{2H} - 2|v|^{2H}). \quad (10.15)$$

10.3 The starting matrix and the associated Wishart tensor

Let us introduce our initial random matrix $\mathcal{X}_{n,d}$. For $i = 1, \dots, n$, we consider n independent Wiener processes $(B^{(i)}(y), y \in \mathbb{R})$ and we set

$$Z_t^{H,i} = I_2^{(i)}(L_t^H) \text{ for every } i = 1, \dots, n \text{ and } t \geq 0,$$

with the kernel L^H given by (10.12). In this way, $Z^{H,1}, \dots, Z^{H,n}$ constitutes a family of n independent Rosenblatt processes. We consider the random matrix $\mathcal{X}_{n,d} = (X_{i,j})_{1 \leq i \leq n, 1 \leq j \leq d}$ whose entries are given by

$$X_{i,j} = Z_j^{H,i} - Z_{j-1}^{H,i} = I_2^{(i)}(L_j^H - L_{j-1}^H), \quad 1 \leq i \leq n, 1 \leq j \leq d. \quad (10.16)$$

Consequently, the matrix $\mathcal{X}_{n,d}$ admits the following structure :

- The entries on different rows are independent since they are multiple integrals with respect to independent Wiener processes.
- On each row $i = 1, \dots, n$, the elements of the matrix $\mathcal{X}_{n,d}$ are correlated and their correlation is given by the auto-correlation of the increments of the Rosenblatt process. Actually, we have by (10.14) that for every $k, l = 1, \dots, d$,

$$\mathbb{E}X_{ik}X_{il} = \rho_H(k-l),$$

with ρ_H given by (10.15). Thus, we can write, for every $i, j = 1, \dots, n$ and for every $k, l = 1, \dots, d$,

$$\mathbb{E}X_{ik}X_{j,l} = \mathbf{1}_{(i=j)}\rho_H(k-l).$$

Let us introduce the (centered) p -Wishart tensor associated to the initial matrix $\mathcal{X}_{n,d}$ defined by (10.16). We use its vectorial representation, i.e. for $j_1, \dots, j_p \in \{1, \dots, n\}$ we set

$$Y_{j_1, \dots, j_p} = \sum_{i=1}^d \left(\prod_{k=1}^p X_{j_k, i} - \mathbb{E} \prod_{k=1}^p X_{j_k, i} \right) \quad (10.17)$$

and we aim at regarding the asymptotic behavior in high-dimensional regime (i.e. when $n, d \rightarrow \infty$) of the random vector

$$\mathbb{Y} = (Y_{j_1, \dots, j_p}, j_1, \dots, j_p = 1, \dots, n). \quad (10.18)$$

Notice that \mathbb{Y} is a n^p -dimensional random vector. Moreover, if $j_1, \dots, j_p = 1, \dots, n$ are all distinct, then

$$\mathbb{E}Y_{j_1, \dots, j_p} = 0$$

so, in this case, we do not need to subtract the expectation in (10.17). Also, by the independence of the elements on different rows and (10.14),

$$\mathbb{E}Y_{j_1, \dots, j_p}^2 = \sum_{i,a=1}^d \mathbb{E}X_{j_1, i}X_{j_1, a} \dots \mathbb{E}X_{j_p, i}X_{j_p, a} = \sum_{i,a=1}^d \rho_H(i-a)^p,$$

We will analyze separately the hyper-diagonal terms of the form $Y_{j, \dots, j}$ and the terms Y_{j_1, \dots, j_p} where j_1, \dots, j_p are not all the same. Actually, we will see that the hyper-diagonal terms $Y_{j, \dots, j}$ are dominant (in the sense of the $L^2(\Omega)$ -norm) for n, d large. Moreover, the behavior of these hyper-diagonal terms is related to the behaviour of the p -variation of the Rosenblatt process. Therefore, we devote the next section to the study of this p -variation.

10.4 The p -variation of the Rosenblatt process

In this part, we consider a Rosenblatt process $(Z_t^H, t \geq 0)$ given by (10.11) and we define, for every $N \geq 1$,

$$V_N = \sum_{i=0}^{N-1} \left[(Z_{i+1}^H - Z_i^H)^p - \mathbb{E}(Z_{i+1}^H - Z_i^H)^p \right]. \quad (10.19)$$

We aim at studying the limit behavior in distribution, as $N \rightarrow \infty$, of the sequence $(V_N, N \geq 1)$. We first notice (we denote by $\stackrel{(d)}{=}$ the equality in distribution),

$$V_N \stackrel{(d)}{=} N^{Hp} \sum_{i=0}^{N-1} \left[\left(Z_{\frac{i+1}{N}}^H - Z_{\frac{i}{N}}^H \right)^p - \mathbb{E} \left(Z_{\frac{i+1}{N}}^H - Z_{\frac{i}{N}}^H \right)^p \right] =: V_{1,N}. \quad (10.20)$$

The relation (10.20) is true due to the scaling property of the Rosenblatt process. Thus, it suffices to get the limit behavior in law, as $N \rightarrow \infty$, of $V_{1,N}$. Our strategy consists in decomposing in Wiener chaos the random variable $V_{1,N}$ and then we analyze separately the behavior of each chaos component. Notice that

$$V_{1,N} = N^{Hp} \sum_{i=0}^{N-1} [I_2(f_{i,N})^p - \mathbb{E}I_2(f_{i,N})^p], \quad (10.21)$$

where we denoted, for $N \geq 1$ and $i = 0, 1, \dots, N-1$,

$$f_{i,N}(y_1, y_2) = L_{\frac{i+1}{N}}^H(y_1, y_2) - L_{\frac{i}{N}}^H(y_1, y_2) = d(H) \int_{\frac{i}{N}}^{\frac{i+1}{N}} (u - y_1)_+^{\frac{H}{2}-1} (u - y_2)_+^{\frac{H}{2}-1} du, \quad (10.22)$$

for all $y_1, y_2 \in \mathbb{R}$, with L^H from (10.12).

To find the chaos decomposition of (10.21), we will use the general product formula (10.9). The notations below are those from Section 10.2.2. This formula gives

$$V_{1,N} = \sum_{k=0}^{p-1} V_N^{(2p-2k)} \quad (10.23)$$

where for every $k = 0, 1, \dots, p-1$, we denoted

$$V_N^{(2p-2k)} = \sum_{\substack{\alpha \in \mathcal{A}_{p,2} \\ |\alpha|=k}} C_\alpha V_{N,\alpha}^{(2p-2|\alpha|)}, \quad (10.24)$$

with

$$V_{N,\alpha}^{(2p-2|\alpha|)} = N^{Hp} \sum_{i=0}^{N-1} I_{2p-2|\alpha|}(\otimes_\alpha(f_{i,N}, \dots, f_{i,N})). \quad (10.25)$$

The index $2p-2|\alpha|$ in the notation $V_{N,\alpha}^{(2p-2|\alpha|)}$ indicates that this random variable belongs to the Wiener chaos of order $2p-2|\alpha|$. Above, $\otimes_\alpha(f_{i,N}, \dots, f_{i,N})$ is the multi-contraction given by (10.8). In a first step, we calculate this multi-contraction of the kernels $f_{i,N}$. Let us also recall the following formula, used several times in the sequel : if $a, b \in \mathbb{R}$ are such that $a+b < -1$, then

$$\int_{\mathbb{R}} (u-y)_+^a (v-y)_+^b du = \beta(-1-a-b, b+1) |u-v|^{a+b+1}. \quad (10.26)$$

Lemma 20. For $N \geq 1$ and $i = 0, 1, \dots, N-1$, let $f_{i,N}$ be given by (10.22). Then for every $\alpha \in \mathcal{A}_{p,2}$, we have

$$\begin{aligned} \otimes_\alpha(f_{i,N}, \dots, f_{i,N})(\xi_1, \xi_2, \dots, \xi_{2p-2|\alpha|}) &= d(H)^p \beta \left(\frac{H}{2}, 1-H \right)^{|\alpha|} \\ &\int_{\frac{i}{N}}^{\frac{i+1}{N}} \dots \int_{\frac{i}{N}}^{\frac{i+1}{N}} dv_1 \dots dv_p \prod_{1 \leq a < b \leq p} (|v_a - v_b|^{H-1})^{\alpha_{ab}} \prod_{k=1}^p \prod_{l=1+\beta_{k-1}}^{\beta_k} (v_k - \xi_l)_+^{\frac{H}{2}-1} \end{aligned}$$

Proof : The proof is a consequence of Proposition 5 in [58]. \square

Notice that the function $\otimes_{\alpha}(f_{i,N}, \dots, f_{i,N})$ is not necessarily symmetric. We denote by $\tilde{\otimes}_{\alpha}(f_{i,N}, \dots, f_{i,N})$ its symmetrization.

For later purposes, we also need to compute the scalar product of the multi-contractions of two kernels $f_{i,N}$ and $f_{j,N}$.

Lemma 21. *For every $\alpha \in \mathcal{A}_{p,2}$ and for every $i, j = 0, 1, \dots, N-1$, with $f_{i,N}$ from (10.22),*

$$\begin{aligned} & \langle \otimes_{\alpha}(f_{i,N}, \dots, f_{i,N}), \otimes_{\alpha}(f_{j,N}, \dots, f_{j,N}) \rangle_{L^2(\mathbb{R}^{2p-2|\alpha|})} \\ &= d(H)^{2p} \beta \left(\frac{H}{2}, 1-H \right)^{2p} N^{-2p} N^{2(1-H)|\alpha|} \int \dots \int_{[0,1]^{2p}} du_1 \dots du_p dv_1 \dots dv_p \\ & \quad \prod_{1 \leq a < b \leq p} [|u_a - u_b|^{H-1} |v_a - v_b|^{H-1}]^{\alpha_{ab}} \prod_{k=1}^p \left| \frac{u_k - v_k}{N} + \frac{i-j}{N} \right|^{(H-1)\beta_k^0}. \end{aligned} \quad (10.27)$$

Proof : By Lemma 20,

$$\begin{aligned} & \langle \otimes_{\alpha}(f_{i,N}, \dots, f_{i,N}), \otimes_{\alpha}(f_{j,N}, \dots, f_{j,N}) \rangle_{L^2(\mathbb{R}^{2p-2|\alpha|})} \\ &= d(H)^{2p} \beta \left(\frac{H}{2}, 1-H \right)^{2|\alpha|} \int_{\mathbb{R}^{2p-2|\alpha|}} d\xi_1 \dots d\xi_{2p-2|\alpha|} \\ & \quad \int_{\frac{i}{N}}^{\frac{i+1}{N}} \dots \int_{\frac{i}{N}}^{\frac{i+1}{N}} du_1 \dots du_p \prod_{1 \leq a < b \leq p} (|u_a - u_b|^{H-1})^{\alpha_{ab}} \prod_{k=1}^p \prod_{l=1+\beta_{k-1}}^{\beta_k} (u_k - \xi_l)_+^{\frac{H}{2}-1} \\ & \quad \int_{\frac{j}{N}}^{\frac{j+1}{N}} \dots \int_{\frac{j}{N}}^{\frac{j+1}{N}} dv_1 \dots dv_p \prod_{1 \leq a < b \leq p} (|v_a - v_b|^{H-1})^{\alpha_{ab}} \prod_{k=1}^p \prod_{l=1+\beta_{k-1}}^{\beta_k} (v_k - \xi_l)_+^{\frac{H}{2}-1}. \end{aligned}$$

Via the key formula (10.26), we get, for every $k = 1, \dots, p$ and for every $l = 1 + \beta_{k-1}, \dots, \beta_k$,

$$\int_{\mathbb{R}} (u_k - \xi_l)_+^{\frac{H}{2}-1} (v_k - \xi_l)_+^{\frac{H}{2}-1} d\xi_l = \beta \left(1-H, \frac{H}{2} \right) |u_k - v_k|^{H-1},$$

we get,

$$\begin{aligned} & \langle \otimes_{\alpha}(f_{i,N}, \dots, f_{i,N}), \otimes_{\alpha}(f_{j,N}, \dots, f_{j,N}) \rangle_{L^2(\mathbb{R}^{2p-2|\alpha|})} \\ &= d(H)^{2p} \beta \left(1-H, \frac{H}{2} \right)^{2p} \int_{\frac{i}{N}}^{\frac{i+1}{N}} \dots \int_{\frac{i}{N}}^{\frac{i+1}{N}} du_1 \dots du_p \int_{\frac{j}{N}}^{\frac{j+1}{N}} \dots \int_{\frac{j}{N}}^{\frac{j+1}{N}} dv_1 \dots dv_p \\ & \quad \prod_{1 \leq a < b \leq p} (|u_a - u_b|^{H-1})^{\alpha_{ab}} (|v_a - v_b|^{H-1})^{\alpha_{ab}} \prod_{k=1}^p |u_k - v_k|^{(H-1)\beta_k^0}. \end{aligned}$$

Next, we perform the changes of variables, for $k = 1, \dots, p$

$$\tilde{u}_k = \left(u_k - \frac{i}{N} \right) N$$

and similarly,

$$\tilde{v}_k = \left(v_k - \frac{j}{N} \right) N.$$

We get

$$\begin{aligned}
 & \langle \otimes_{\alpha}(f_{i,N}, \dots, f_{i,N}), \otimes_{\alpha}(f_{j,N}, \dots, f_{j,N}) \rangle_{L^2(\mathbb{R}^{2p-2|\alpha|})} \\
 &= d(H)^{2p} \beta \left(1 - H, \frac{H}{2}\right)^{2p} N^{-2p} \int_{[0,1]^{2p}} du_1 \cdots du_p dv_1 \cdots dv_p \\
 & \quad \prod_{1 \leq a < b \leq p} N^{-2\alpha_{ab}(H-1)} [|u_a - u_b|^{H-1} |v_a - v_b|^{H-1}]^{\alpha_{ab}} \prod_{k=1}^p \left| \frac{u_k - v_k}{N} + \frac{i-j}{N} \right|^{(H-1)\beta_k^0}. \\
 &= d(H)^{2p} \beta \left(1 - H, \frac{H}{2}\right)^{2p} N^{-2p-2|\alpha|(H-1)} \int \cdots \int_{[0,1]^{2p}} du_1 \cdots du_p dv_1 \cdots dv_p \\
 & \quad \prod_{1 \leq a < b \leq p} [|u_a - u_b|^{H-1} |v_a - v_b|^{H-1}]^{\alpha_{ab}} \prod_{k=1}^p \left| \frac{u_k - v_k}{N} + \frac{i-j}{N} \right|^{(H-1)\beta_k^0}.
 \end{aligned}$$

□

Remark 10. Notice that the formula (10.27) also holds in the case $|\alpha| = 0$. In this case $\alpha_{i,j} = 0$ for every i, j and $\beta_k^0 = 2$ for $k = 1, \dots, p$, and we have the following;

$$\begin{aligned}
 & \langle \otimes_{\alpha}(f_{i,N}, \dots, f_{i,N}), \otimes_{\alpha}(f_{j,N}, \dots, f_{j,N}) \rangle_{L^2(\mathbb{R}^{2p})} \\
 &= d(H)^{2p} \beta \left(\frac{H}{2}, 1 - H\right)^{2p} N^{-2Hp} \left[2|i-j|^{2H} - |i-j+1|^{2H} - |i-j-1|^{2H}\right]^p.
 \end{aligned}$$

10.4.1 The limit of the term in the second Wiener chaos

By (10.25), the projection of $V_{1,N}$ on the second Wiener chaos is given by

$$V_N^{(2)} = \sum_{\substack{\alpha \in \mathcal{A}_{p,2} \\ |\alpha| = p-1}} C_{\alpha} V_{N,\alpha}^{(2)}, \quad (10.28)$$

with, for $\alpha \in \mathcal{A}_{p,2}$ such that, $|\alpha| = p-1$,

$$V_{N,\alpha}^{(2)} = N^{Hp} \sum_{i=0}^{N-1} I_2(\otimes_{\alpha}(f_{i,N}, \dots, f_{i,N})),$$

where $f_{i,N}$ are given by (10.22). Recall that C_{α} is the constant that appears in the product formula (10.9). Since $\beta_1^0 + \dots + \beta_p^0 = 2$ if $|\alpha| = p-1$, we notice that there are at most two non-vanishing indices among $\beta_k^0, k = 1, \dots, p$. Therefore, by Lemma 20, the multi-contraction $\otimes_{\alpha}(f_{i,N}, \dots, f_{i,N})$ of the second chaos term $V_{N,\alpha}^{(2)}$ can be written as :

$$d(H)^p \beta \left(\frac{H}{2}, 1 - H\right)^{p-1} \int_{\frac{i}{N}}^{\frac{i+1}{N}} \cdots \int_{\frac{i}{N}}^{\frac{i+1}{N}} du_1 \cdots du_p \left(\prod_{1 \leq a < b \leq p} |u_a - u_b|^{\alpha_{ab}(H-1)} \right) (u_r - \xi_1)_{+}^{\frac{H}{2}-1} (u_s - \xi_2)_{+}^{\frac{H}{2}-1}$$

with $r, s = 1, \dots, p$ (notice that we may have $r = s$). A key point consist in understanding the behavior of the above sequence as $N \rightarrow \infty$. This will be done in the below result.

Proposition 14. Let $\alpha \in \mathcal{A}_{p,2}$ with $|\alpha| = p - 1$, $r, s \in \{1, \dots, p\}$ and set, for every $N \geq 1$ and $\xi_1, \xi_2 \in \mathbb{R}$,

$$g_{N,\alpha}(\xi_1, \xi_2) =: g_N(\xi_1, \xi_2) = N^{-H} N^{Hp} \sum_{i=0}^{N-1} \int_{\frac{i}{N}}^{\frac{i+1}{N}} \dots \int_{\frac{i}{N}}^{\frac{i+1}{N}} du_1 \dots du_p \left(\prod_{1 \leq a < b \leq p} |u_a - u_b|^{\alpha_{ab}(H-1)} \right) (u_r - \xi_1)_+^{\frac{H}{2}-1} (u_s - \xi_2)_+^{\frac{H}{2}-1}. \quad (10.29)$$

Then the sequence $(I_2(g_N), N \geq 1)$ converges in $L^2(\Omega)$ to $d(H)^{-1} C_{1,\alpha} Z_1^H$ where Z^H is a Rosenblatt process and $d(H)$ is the constant from (10.10) and

$$C_{1,\alpha} = \int_{[0,1]^p} du_1 \dots du_p \left(\prod_{1 \leq a < b \leq p} |u_a - u_b|^{\alpha_{ab}(H-1)} \right). \quad (10.30)$$

Moreover,

$$\mathbb{E} |I_2(g_N) - d(H)^{-1} C_{1,\alpha} Z_1^H|^2 \leq c(H, p, \alpha) \begin{cases} N^{1-2H}, & \text{if } H \in (\frac{1}{2}, \frac{3}{4}) \\ \log(N) N^{-\frac{1}{2}} & \text{if } H = \frac{3}{4} \\ N^{2H-2} & \text{if } H \in (\frac{3}{4}, 1). \end{cases} \quad (10.31)$$

Proof : After the standard change of variables $\tilde{u}_k = (u_k - \frac{i}{N})N$, we obtain

$$\begin{aligned} g_N(\xi_1, \xi_2) &= N^{-H} N^{Hp} N^{-p} N^{(1-H)(p-1)} \int_{[0,1]^p} du_1 \dots du_p \left(\prod_{1 \leq a < b \leq p} |u_a - u_b|^{\alpha_{ab}(H-1)} \right) \\ &\quad \times \sum_{i=0}^{N-1} \left(\frac{u_r}{N} + \frac{i}{N} - \xi_1 \right)_+^{\frac{H}{2}-1} \left(\frac{u_s}{N} + \frac{i}{N} - \xi_2 \right)_+^{\frac{H}{2}-1} \\ &= \frac{1}{N} \int_{[0,1]^p} du_1 \dots du_p \left(\prod_{1 \leq a < b \leq p} |u_a - u_b|^{\alpha_{ab}(H-1)} \right) \\ &\quad \times \sum_{i=0}^{N-1} \left(\frac{u_r}{N} + \frac{i}{N} - \xi_1 \right)_+^{\frac{H}{2}-1} \left(\frac{u_s}{N} + \frac{i}{N} - \xi_2 \right)_+^{\frac{H}{2}-1}. \end{aligned}$$

We decompose g_N as follows

$$g_N(\xi_1, \xi_2) = g_{N,1}(\xi_1, \xi_2) + r_N(\xi_1, \xi_2)$$

where, with $C_{1,\alpha}$ from (10.30),

$$\begin{aligned} g_{N,1}(\xi_1, \xi_2) &= \int_{[0,1]^p} du_1 \dots du_p \left(\prod_{1 \leq a < b \leq p} |u_a - u_b|^{\alpha_{ab}(H-1)} \right) \frac{1}{N} \sum_{i=0}^{N-1} \left(\frac{i}{N} - \xi_1 \right)_+^{\frac{H}{2}-1} \left(\frac{i}{N} - \xi_2 \right)_+^{\frac{H}{2}-1} \\ &= C_{1,\alpha} \frac{1}{N} \sum_{i=0}^{N-1} \left(\frac{i}{N} - \xi_1 \right)_+^{\frac{H}{2}-1} \left(\frac{i}{N} - \xi_2 \right)_+^{\frac{H}{2}-1} \end{aligned} \quad (10.32)$$

and

$$\begin{aligned} &r_N(\xi_1, \xi_2) \quad (10.33) \\ &= \frac{C_{1,\alpha}}{N} \sum_{i=0}^{N-1} \left[\left(\frac{u_r}{N} + \frac{i}{N} - \xi_1 \right)_+^{\frac{H}{2}-1} \left(\frac{u_s}{N} + \frac{i}{N} - \xi_2 \right)_+^{\frac{H}{2}-1} - \left(\frac{i}{N} - \xi_1 \right)_+^{\frac{H}{2}-1} \left(\frac{i}{N} - \xi_2 \right)_+^{\frac{H}{2}-1} \right]. \end{aligned}$$

First, notice that that the sequence $g_{N,1}$ is a Riemann sum and it converges pointwise, for every $\xi_1, \xi_2 \in \mathbb{R}$, to

$$C_{1,\alpha} \int_0^1 dx (x - \xi_1)_+^{\frac{H}{2}-1} (x - \xi_2)_+^{\frac{H}{2}-1} = d(H)^{-1} C_{1,\alpha} L_1^H(\xi_1, \xi_2)$$

where L^H is the kernel of the Rosenblatt process defined by (10.12). It has been showed in [48] or [116] that $g_{N,1}$ is a Cauchy sequence in $L^2(\mathbb{R}^2)$. Therefore, $I_2(g_N)$ converges in $L^2(\Omega)$ as $N \rightarrow \infty$ and its limit coincides with $d(H)^{-1} C_{1,\alpha} Z_1^H$ where Z^H is the Rosenblatt process given by (10.10).

Let us show that r_N given by (10.33) converges to zero in $L^2(\mathbb{R}^2)$, as $N \rightarrow \infty$. From the identity (10.26),

$$\begin{aligned} & \|r_N\|_{L^2(\mathbb{R}^2)}^2 \\ &= c(H, p) \int_{[0,1]^{2p}} du_1 \dots du_p dv_1 \dots dv_p \left(\prod_{1 \leq a < b \leq p} |u_a - u_b|^{\alpha_{ab}(H-1)} |v_a - v_b|^{\alpha_{ab}(H-1)} \right) \\ & \quad \frac{1}{N^2} \sum_{i,j=0}^{N-1} \left[\left| \frac{u_r - v_r}{N} + \frac{i-j}{N} \right|^{H-1} \left| \frac{u_s - v_s}{N} + \frac{i-j}{N} \right|^{H-1} + \left| \frac{i-j}{N} \right|^{2H-2} \right. \\ & \quad \left. - \left| \frac{u_r}{N} + \frac{i-j}{N} \right|^{H-1} \left| \frac{u_s}{N} + \frac{i-j}{N} \right|^{H-1} - \left| \frac{v_r}{N} - \frac{i-j}{N} \right|^{H-1} \left| \frac{v_s}{N} - \frac{i-j}{N} \right|^{H-1} \right] \\ &= r_{N,d} + r_{N,nd}, \end{aligned}$$

where $r_{N,d}$ denotes the diagonal term obtained by taking $i = j$ above while $r_{N,nd}$ is the non-diagonal term corresponding to $i \neq j$. We first notice that

$$\begin{aligned} |r_{N,d}| &= c(H, p) \int_{[0,1]^{2p}} du_1 \dots du_p dv_1 \dots dv_p \prod_{1 \leq a < b \leq p} [|u_a - u_b| |v_a - v_b|]^{(H-1)\alpha_{ab}} \\ & \quad \frac{1}{N^2} \sum_{i=0}^{N-1} \left| \frac{u_r - v_r}{N} \right|^{H-1} \left| \frac{u_s - v_s}{N} \right|^{H-1} - \left| \frac{u_r}{N} \right|^{H-1} \left| \frac{u_s}{N} \right|^{H-1} - \left| \frac{v_r}{N} \right|^{H-1} \left| \frac{v_s}{N} \right|^{H-1} \\ &= N^{1-2H} c(H, p) \int_{[0,1]^{2p}} du_1 \dots du_p dv_1 \dots dv_p \prod_{1 \leq a < b \leq p} [|u_a - u_b| |v_a - v_b|]^{(H-1)\alpha_{a,b}} \\ & \quad \left[|u_r - v_r|^{H-1} |u_s - v_s|^{H-1} - |u_r|^{H-1} |u_s|^{H-1} - |v_r|^{H-1} |v_s|^{H-1} \right] \\ &\leq CN^{1-2H} \xrightarrow{N \rightarrow \infty} 0, \end{aligned} \tag{10.34}$$

where we used $H > \frac{1}{2}$. Concerning the non-diagonal tem, we can write

$$\begin{aligned}
|r_{N,nd}| &\leq c(H,p) \int_{[0,1]^{2p}} du_1 \cdots du_p dv_1 \cdots dv_p \prod_{1 \leq a < b \leq p} [|u_a - u_b| |v_a - v_b|]^{(H-1)\alpha_{ab}} \\
&\quad \frac{1}{N} \sum_{k=1}^{N-1} \left[\left| \frac{u_r - v_r}{N} + \frac{k}{N} \right|^{H-1} \left| \frac{u_s - v_s}{N} + \frac{k}{N} \right|^{H-1} + \left(\frac{k}{N} \right)^{2H-2} \right. \\
&\quad \left. - \left| \frac{u_r}{N} + \frac{k}{N} \right|^{H-1} \left| \frac{u_s}{N} + \frac{k}{N} \right|^{H-1} - \left| \frac{v_r}{N} - \frac{k}{N} \right|^{H-1} \left| \frac{v_s}{N} - \frac{k}{N} \right|^{H-1} \right] \\
&\leq c(H,p) \int_{[0,1]^{2p}} du_1 \cdots du_p dv_1 \cdots dv_p \prod_{1 \leq a < b \leq p} [|u_a - u_b| |v_a - v_b|]^{(H-1)\alpha_{ab}} \\
&\quad \frac{1}{N} \sum_{k=1}^{N-1} \left(\frac{k}{N} \right)^{2H-2} \left[\left(\frac{u_r - v_r}{k} + 1 \right)^{H-1} \left(\frac{u_s - v_s}{k} + 1 \right)^{H-1} + 1 \right. \\
&\quad \left. - \left(\frac{u_r}{k} + 1 \right)^{H-1} \left(\frac{u_s}{k} + 1 \right)^{H-1} - \left(1 - \frac{v_r}{k} \right)^{H-1} \left(1 - \frac{v_s}{k} \right)^{H-1} \right].
\end{aligned}$$

For $x > 0$ close to zero such that $0 \leq u_r x \leq \frac{1}{2}, 0 \leq u_s x \leq \frac{1}{2}$, and $0 \leq v_r x \leq \frac{1}{2}, 0 \leq v_s x \leq \frac{1}{2}$ consider the function

$$\begin{aligned}
g(x) &= ((u_r - v_r)x + 1)^{H-1} ((u_s - v_s)x + 1)^{H-1} + 1 \\
&\quad - (u_r x + 1)^{H-1} (u_s x + 1)^{H-1} - (1 - v_r x)^{H-1} (1 - v_s x)^{H-1}.
\end{aligned}$$

We have $g(0) = 0$ and

$$\begin{aligned}
g'(x) &= (H-1) ((u_r - v_r)((u_r - v_r)x + 1)^{H-2} ((u_s - v_s)x + 1)^{H-1} \\
&\quad + (u_s - v_s)((u_r - v_r)x + 1)^{H-1} ((u_s - v_s)x + 1)^{H-2}) \\
&\quad - (H-1) (u_r (u_r x + 1)^{H-2} (u_s x + 1)^{H-1} + u_s (u_s x + 1)^{H-2} (u_r x + 1)^{H-1}) \\
&\quad + (H-1) (v_r (1 - v_r x)^{H-2} (1 - v_s x)^{H-1} + v_s (1 - v_s x)^{H-2} (1 - v_r x)^{H-1}).
\end{aligned}$$

If $0 \leq u_r x \leq \frac{1}{2}, 0 \leq u_s x \leq \frac{1}{2}, 0 \leq v_r x \leq \frac{1}{2}, 0 \leq v_s x \leq \frac{1}{2}$, it is easy to see that $|g'(x)| \leq c_2(H)$, so

$$\left| g\left(\frac{1}{k}\right) \right| \leq c_2(H) \frac{1}{k} \text{ for } k \geq 2.$$

Hence

$$\begin{aligned}
|r_{N,nd}| &\leq C \left(N^{1-2H} + \frac{1}{N} \sum_{k=2}^{N-1} \left(\frac{k}{N} \right)^{2H-2} \frac{1}{k} \right) \\
&\leq C \left(N^{1-2H} + N^{1-2H} \sum_{k=2}^{N-1} k^{2H-3} \right) \leq CN^{1-2H}. \tag{10.35}
\end{aligned}$$

From (10.34) and (10.35), the sequence r_N defined by (10.33) converges to zero in $L^2(\Omega)$ as $N \rightarrow \infty$. The estimate (10.31) is obtained by following the line of the proof of Proposition 3 in [19]. \square

Remark 11. *The fact the the constant $C_{1,\alpha}$ given by (10.30) is well defined follows as a particular case of Lemma 13 in [58].*

From Proposition 14, we can deduce the convergence of the second chaos component of $V_{1,N}$. Let us consider the following renormalization : for all $\alpha \in \mathcal{A}_{p,2}$ with $|\alpha| \leq p-1$, we set

$$\tilde{V}_{1,N} = \sum_{k=0}^{p-1} \tilde{V}_N^{(2p-2k)} \quad (10.36)$$

where, for $k = 0, \dots, p-1$,

$$\tilde{V}_N^{(2p-2k)} = N^{-H} V_N^{(2p-2k)}, \quad (10.37)$$

and $V_N^{(2p-2k)}$ from (10.24). In particular, the projection of $\tilde{V}_{1,N}$ on the second Wiener chaos reads

$$\tilde{V}_N^{(2)} = \sum_{\substack{\alpha \in \mathcal{A}_{p,2} \\ |\alpha|=p-1}} C_\alpha \tilde{V}_{N,\alpha}^{(2)} \quad (10.38)$$

where $\tilde{V}_{N,\alpha}^{(2)}$ is expressed as

$$\tilde{V}_{N,\alpha}^{(2)} = N^{-H} N^{Hp} \sum_{i=0}^{N-1} I_2(\otimes_\alpha(f_{i,N}, \dots, f_{i,N})). \quad (10.39)$$

From Proposition 14, we now deduce the behavior of the renormalized second chaos term.

Proposition 15. *Consider the sequence $(\tilde{V}_N^{(2)}, N \geq 1)$ given by (10.39). Then*

$$\tilde{V}_N^{(2)} \xrightarrow[N \rightarrow \infty]{} K(H, p) Z_1^H \quad (10.40)$$

in $L^2(\Omega)$, with

$$K(H, p) = d(H)^{p-1} \beta \left(\frac{H}{2}, 1-H \right)^{p-1} \sum_{\substack{\alpha \in \mathcal{A}_{p,2} \\ |\alpha|=p-1}} C_\alpha C_{1,\alpha}. \quad (10.41)$$

Moreover,

$$\mathbb{E} \left| \tilde{V}_N^{(2)} - K(H, p) Z_1^H \right|^2 \leq c(H, p) \begin{cases} N^{1-2H}, & \text{if } H \in \left(\frac{1}{2}, \frac{3}{4} \right) \\ \log(N) N^{-\frac{1}{2}} & \text{if } H = \frac{3}{4} \\ N^{2H-2} & \text{if } H \in \left(\frac{3}{4}, 1 \right). \end{cases}$$

Proof : By (10.24) and Lemma 20,

$$\tilde{V}_N^{(2)} = d(H)^p \beta \left(\frac{H}{2}, 1-H \right)^{p-1} \sum_{\substack{\alpha \in \mathcal{A}_{p,2} \\ |\alpha|=p-1}} C_\alpha I_2(g_{N,\alpha})$$

with $g_{N,\alpha}$ given by (10.29). It suffices to apply Proposition 14. \square

10.4.2 The terms in the chaos of order $p > 2$ are negligible

Let us consider the sequence $\tilde{V}_N^{(2p-2k)}$ given by (10.37) with $0 \leq k \leq p-2$. We will show that these terms are negligible in $L^2(\Omega)$ so the limit of $\tilde{V}_{1,N}$, and thus of the (renormalized) p -variation of the Rosenblatt process will be given by the second chaos component.

Proposition 16. *Let $0 \leq k \leq p-2$ and let $\tilde{V}_N^{(2p-2k)}$ be given by (10.37). Then*

$$\tilde{V}_N^{(2p-2k)} \xrightarrow[N \rightarrow \infty]{} 0 \text{ in } L^2(\Omega)$$

and

$$\mathbb{E} \left| \tilde{V}_N^{(2p-2k)} \right|^2 \leq c(H, p) \begin{cases} N^{1-2H}, & \text{if } H \in \left(\frac{1}{2}, \frac{3}{4}\right) \\ \log(N)N^{-\frac{1}{2}} & \text{if } H = \frac{3}{4} \\ N^{2H-2} & \text{if } H \in \left(\frac{3}{4}, 1\right). \end{cases}$$

Proof : Let us estimate $\mathbb{E} \left(\tilde{V}_N^{2p-2|\alpha|} \right)^2$ for $\alpha \in A_{p,2}$ fixed such that $|\alpha| = k \leq p-2$. We have

$$\begin{aligned} \mathbb{E} \left(\tilde{V}_N^{2p-2k} \right)^2 &= N^{2H(p-1)} (2p-2|\alpha|)! \left\| \sum_{i=0}^{N-1} \tilde{\otimes}_{\alpha}(f_{i,N}, \dots, f_{i,N}) \right\|_{L^2(\mathbb{R}^{2p-|\alpha|})}^2 \\ &\leq N^{2H(p-1)} (2p-2|\alpha|)! \left\| \sum_{i=0}^{N-1} \otimes_{\alpha}(f_{i,N}, \dots, f_{i,N}) \right\|_{L^2(\mathbb{R}^{2p-|\alpha|})}^2 \\ &= N^{2H(p-1)} (2p-2|\alpha|)! \sum_{i,j=0}^{N-1} \langle \otimes_{\alpha}(f_{i,N}, \dots, f_{i,N}), \otimes_{\alpha}(f_{j,N}, \dots, f_{j,N}) \rangle_{L^2(\mathbb{R}^{2p-|\alpha|})} \end{aligned}$$

and then, by Lemma 21,

$$\begin{aligned} &\mathbb{E} \left(\tilde{V}_N^{2p-2k} \right)^2 \\ &\leq (2p-2|\alpha|)! d(H)^{2p} \beta \left(\frac{H}{2}, 1-H \right)^{2p} N^{(H-1)(2p-2|\alpha|)} N^{-2H} \\ &\quad \sum_{i,j=0}^{N-1} \int_{[0,1]^{2p}} du_1 \dots du_p dv_1 \dots dv_p \prod_{1 \leq a < b \leq p} (|u_a - u_b| |v_a - v_b|)^{\alpha_{ab}(H-1)} \\ &\quad \prod_{k=1}^p \left| \frac{u_k - v_k}{N} + \frac{i-j}{N} \right|^{(H-1)\beta_k^0} \\ &= A_{N,d} + A_{N,nd} \end{aligned} \tag{10.42}$$

where $A_{N,d}$ stands for the diagonal part (the terms with $i = j$ above) while $A_{N,nd}$ denotes the non-diagonal part (the terms with $i \neq j$). First, we notice that

$$\begin{aligned}
 A_{N,d} &= (2p-2|\alpha|)!d(H)^{2p}\beta\left(\frac{H}{2}, 1-H\right)^{2p} \int_{[0,1]^{2p}} du_1 \dots du_p dv_1 \dots dv_p \\
 &\quad \prod_{1 \leq a < b \leq p} (|u_a - u_b| |v_a - v_b|)^{\alpha_{ab}(H-1)} \prod_{k=1}^p |u_k - v_k|^{(H-1)\beta_k^0} \\
 &\quad N^{-2H+1} N^{(H-1)(2p-2|\alpha|)} N^{(1-H)(\beta_1^0 + \dots + \beta_p^0)} \\
 &\leq c(p, \alpha, H) N^{1-2H} \xrightarrow{N \rightarrow \infty} 0,
 \end{aligned} \tag{10.43}$$

where we recall that $\beta_1^0 + \dots + \beta_p^0 = 2p - 2|\alpha|$.

Let us deal with the non-diagonal summand $A_{N,nd}$. We use the fact that for $u, v \in [0, 1]$ and $k \geq 2$,

$$k - 1 \leq |u - v + k| \leq k + 1.$$

Therefore,

$$\begin{aligned}
 &A_{N,nd} \\
 &\leq c_2(H, p, \alpha) N^{-2H} N^{(H-1)(2p-2|\alpha|)} \sum_{\substack{i,j=0 \\ i-j \geq 2}}^{N-1} \left(\frac{i-j-1}{N}\right)^{(H-1)(\beta_1^0 + \dots + \beta_p^0)} + c(H, p, \alpha) N^{1-2H} \\
 &\leq c_2(H, p, \alpha) N^{1-2H} \sum_{k=2}^{N-1} (k-1)^{(H-1)(2p-2|\alpha|)} + c(p, H, \alpha) N^{1-2H} \\
 &\leq N^{1-2H} \left(c_2(p, H, \alpha) \sum_{k=1}^{N-2} k^{(H-1)(2p-2|\alpha|)} + c(p, H, \alpha) \right).
 \end{aligned}$$

If $|\alpha| \leq p - 3$ and $H < \frac{3}{4}$, then $\sum_{k \geq 1} k^{(H-1)(2p-2|\alpha|)}$ converges and then

$$A_{N,nd} \leq CN^{1-2H}.$$

If $|\alpha| = p - 2$, then the series $\sum_{k \geq 1} k^{4H-4}$ converges for $H < \frac{3}{4}$ and we get the same estimate as above. If $|\alpha| = p - 2$ and $H > \frac{3}{4}$, then

$$A_{N,nd} \leq C(N^{2H-2} + N^{1-2H}) \leq CN^{2H-2} \xrightarrow{N \rightarrow \infty} 0$$

while if $|\alpha| = p - 2$ and $H = \frac{3}{4}$, then

$$A_{N,nd} \leq CN^{1-2H} (\log(N) + 1) \leq C \log(N) N^{-\frac{1}{2}} \xrightarrow{N \rightarrow \infty} 0.$$

thus, for all $\alpha \in \mathcal{A}_{p,2}$ with $|\alpha| \leq p - 2$, we have

$$A_{N,nd} \leq c(H, p, \alpha) \begin{cases} N^{1-2H}, & \text{if } H \in (\frac{1}{2}, \frac{3}{4}) \\ \log(N) N^{-\frac{1}{2}} & \text{if } H = \frac{3}{4} \\ N^{2H-2} & \text{if } H \in (\frac{3}{4}, 1). \end{cases} \xrightarrow{N \rightarrow \infty} 0. \tag{10.44}$$

We conclude the proof by combining the bounds (10.43), (10.44) and (10.42). \square

We can now state the result concerning the asymptotic behaviour of the centered p -variation of the Rosenblatt process. We also need to define the Wasserstein distance which will be used in the sequel. The Wasserstein distance (denoted d_W in the sequel) between the probability distributions of two random vector $X, Y \in \mathbb{R}^m$ is defined by

$$d_W(X, Y) = \sup_{h \in \mathcal{A}} |\mathbb{E}h(X) - \mathbb{E}h(Y)|$$

where \mathcal{A} is the class of functions $h : \mathbb{R}^m \rightarrow \mathbb{R}$ with $h(X), h(Y) \in L^1(\Omega)$ such that $\|h\|_{Lip} \leq 1$ where $\|\cdot\|$ is the Lipschitz norm, i.e.

$$\|h\|_{Lip} = \sup_{x, y \in \mathbb{R}^m, x \neq y} \frac{|h(x) - h(y)|}{\|x - y\|}$$

with $\|\cdot\|$ the Euclidean norm in \mathbb{R}^m .

Theorem 16. *Consider the sequence $(V_N, N \geq 1)$ given by (10.19) and let us set, for, $N \geq 1$,*

$$\tilde{V}_N = N^{-H} V_N.$$

Then

$$\tilde{V}_N \xrightarrow[N \rightarrow \infty]{} K(H, p) Z_1^H \text{ in distribution}$$

where Z^H is a Rosenblatt process with Hurst index $H \in (\frac{1}{2}, 1)$ and $K(H, p)$ is the constant defined by (10.41). For N large enough,

$$d_W(\tilde{V}_N, K(H, p) Z_1^H) \leq c(H, p) \begin{cases} N^{\frac{1}{2}-H}, & \text{if } H \in (\frac{1}{2}, \frac{3}{4}) \\ \sqrt{\log(N)} N^{-\frac{1}{4}} & \text{if } H = \frac{3}{4} \\ N^{H-1} & \text{if } H \in (\frac{3}{4}, 1). \end{cases}$$

Proof : By (10.20), \tilde{V}_N has the same law as $\tilde{V}_{1,N}$ and by (10.23) and Propositions 15 and 16, the sequence $(\tilde{V}_{1,N}, N \geq 1)$ converges in $L^2(\Omega)$, as N tends to infinity, to $K(H, p, \alpha) Z_1^H$ and satisfies

$$d_W(\tilde{V}_{1,N}, K(H, p) Z_1^H) \leq \left[\mathbb{E} \left| \tilde{V}_{1,N} - K(H, p) Z_1^H \right|^2 \right]^{\frac{1}{2}} \leq c(H, p) \begin{cases} N^{\frac{1}{2}-H}, & \text{if } H \in (\frac{1}{2}, \frac{3}{4}) \\ \sqrt{\log(N)} N^{-\frac{1}{4}} & \text{if } H = \frac{3}{4} \\ N^{H-1} & \text{if } H \in (\frac{3}{4}, 1). \end{cases} \quad (10.45)$$

We conclude since obviously $d_W(\tilde{V}_{1,N}, K(H, p) Z_1^H) = d_W(\tilde{V}_N, K(H, p) Z_1^H)$. □

10.5 Asymptotic behavior of the Wishart tensor

Let us come back to the n^p -dimensional random vector (10.18) with components given by (10.17). We regard its limit behavior as $n, d \rightarrow \infty$. As mentioned in Section 10.3, we will separate the study of the hyper-diagonal terms of the form $Y_{j, \dots, j}$ with $j = 1, \dots, n$ and of the other terms Y_{j_1, \dots, j_p} where $j_1, \dots, j_p \in \{1, \dots, n\}$ are not all the same. Let us consider the following renormalized tensor :

$$\tilde{Y} = \left(\tilde{Y}_{j_1, \dots, j_p}, j_1, \dots, j_p = 1, \dots, n \right) \quad (10.46)$$

where for all $j_1, \dots, j_p \in \{1, \dots, n\}$, we define

$$\tilde{Y}_{j_1, \dots, j_p} = d^{-H} Y_{j_1, \dots, j_p}. \quad (10.47)$$

10.5.1 The limit of the hyper-diagonal terms

Let us take $j_1 = j_2 = \dots = j_p = j$ in (10.17). Then, for $j = 1, \dots, n$, we have

$$Y_{j, \dots, j} = \sum_{k=1}^d \left[\left(Z_k^{H,j} - Z_{k-1}^{H,j} \right)^p - \mathbb{E} \left(Z_k^{H,j} - Z_{k-1}^{H,j} \right)^p \right],$$

where $Z^{H,1}, \dots, Z^{H,n}$ are independent Rosenblatt processes. So

$$\tilde{Y}_{j, \dots, j} = d^{-H} \sum_{k=0}^{d-1} \left[\left(Z_{k+1}^{H,j} - Z_k^{H,j} \right)^p - \mathbb{E} \left(Z_{k+1}^{H,j} - Z_k^{H,j} \right)^p \right].$$

The asymptotic behavior of the hyper-diagonal terms can then be obtained directly from Theorem 16.

Proposition 17. *For $j = 1, \dots, n$, let $\tilde{Y}_{j, \dots, j}$ be given by (10.47) and let $K(H, p)$ be given by (10.41). Then*

$$\tilde{Y}_{j, \dots, j} \xrightarrow{d \rightarrow \infty} K(H, p) Z_1^{H,j}$$

and

$$\mathbb{E} \left| \tilde{Y}_{j, \dots, j} - K(H, p) Z_1^{H,j} \right|^2 \leq c(H, p) \begin{cases} d^{1-2H}, & \text{if } H \in \left(\frac{1}{2}, \frac{3}{4} \right) \\ \log(d) d^{-\frac{1}{2}} & \text{if } H = \frac{3}{4} \\ d^{2H-2} & \text{if } H \in \left(\frac{3}{4}, 1 \right). \end{cases}$$

10.5.2 The limit behavior of the other terms

Let us consider the situation when the indices j_1, \dots, j_p in (10.17) do not coincide all of them, i.e. $\text{card}\{j_1, \dots, j_p\} \geq 2$. We can write

$$Y_{j_1, \dots, j_p} = \sum_{i=0}^{d-1} \left[\left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} \left(Z_{i+1}^{H,2} - Z_i^{H,2} \right)^{p_2} \dots \left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} - \mathbb{E} \left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} \left(Z_{i+1}^{H,2} - Z_i^{H,2} \right)^{p_2} \dots \left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} \right],$$

where $p_1 + \dots + p_x = p, x \geq 2$ and $Z^{H,1}, \dots, Z^{H,x}$ are independent Rosenblatt processes. We can also write

$$Y_{j_1, \dots, j_p} = \sum_{i=0}^{d-1} \left[\left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} - \mathbb{E} \left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} \right] \left[\left(Z_{i+1}^{H,2} - Z_i^{H,2} \right)^{p_2} - \mathbb{E} \left(Z_{i+1}^{H,2} - Z_i^{H,2} \right)^{p_2} \right] \dots \left[\left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} - \mathbb{E} \left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} \right]. \quad (10.48)$$

We will show that all these non-diagonal terms, after renormalization, converge to zero in $L^2(\Omega)$ as $d \rightarrow \infty$.

Proposition 18. *For every $j_1, \dots, j_p \in \{1, \dots, n\}$ with $\text{card}\{j_1, \dots, j_p\} \geq 2$, we have, for d large enough,*

$$\mathbb{E} \left| \tilde{Y}_{j_1, \dots, j_p} \right|^2 \leq c(H, p) \begin{cases} d^{1-2H}, & \text{if } H \in \left(\frac{1}{2}, \frac{3}{4} \right) \\ \log(d) d^{-\frac{1}{2}} & \text{if } H = \frac{3}{4} \\ d^{2H-2} & \text{if } H \in \left(\frac{3}{4}, 1 \right). \end{cases}$$

and in particular $\tilde{Y}_{j_1, \dots, j_p} \xrightarrow{d \rightarrow \infty} 0$ in $L^2(\Omega)$.

Proof : By (10.48),

$$\begin{aligned} & \mathbb{E}(Y_{j_1, \dots, j_p})^2 \tag{10.49} \\ &= \sum_{i,j=0}^{N-1} \mathbb{E} \left[\left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} - \mathbb{E} \left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} \right] \left[\left(Z_{j+1}^{H,1} - Z_j^{H,1} \right)^{p_1} - \mathbb{E} \left(Z_{j+1}^{H,1} - Z_j^{H,1} \right)^{p_1} \right] \\ & \quad \dots \mathbb{E} \left[\left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} - \mathbb{E} \left(Z_{i+1}^{H,x} - Z_i^{H,x} \right)^{p_x} \right] \left[\left(Z_{j+1}^{H,x} - Z_j^{H,x} \right)^{p_x} - \mathbb{E} \left(Z_{j+1}^{H,x} - Z_j^{H,x} \right)^{p_x} \right] \end{aligned}$$

Let us calculate, for $i, j = 0, \dots, d-1$, and for $a = 1, \dots, x$,

$$\mathbb{E} \left[\left(Z_{i+1}^{H,a} - Z_i^{H,a} \right)^{p_a} - \mathbb{E} \left(Z_{i+1}^{H,a} - Z_i^{H,a} \right)^{p_a} \right] \left[\left(Z_{j+1}^{H,a} - Z_j^{H,a} \right)^{p_a} - \mathbb{E} \left(Z_{j+1}^{H,a} - Z_j^{H,a} \right)^{p_a} \right].$$

We assume $a = 1$. By the general product formula (10.9), with the notation from Section 10.2.2,

$$\begin{aligned} & \left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} - \mathbb{E} \left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} \\ &= \sum_{\substack{\alpha^{(1)} \in A_{p_1,2} \\ |\alpha^{(1)}| \leq p_1 - 1}} C_{\alpha^{(1)}} I_{2p_1 - 2|\alpha^{(1)}|} \left(\otimes_{\alpha^{(1)}} (f_i, \dots, f_i) \right). \end{aligned}$$

Above $\alpha^{(1)} \in \mathcal{A}_{p_1,2}$ is the set of indices $\alpha^{(1)} = \left(\alpha_{ij}^{(1)}, 1 \leq i < j \leq p_1 \right)$ with the properties described in Section 10.2.2. We denoted, for $i = 0, \dots, d-1$,

$$f_i = L_{i+1}^H - L_i^H,$$

with L^H from (10.12) and the multi-contraction $\otimes_{\alpha^{(1)}} (f_i, \dots, f_i)$ was defined in (10.8). We will have, by following the computations from Lemmas 20 and 21,

$$\begin{aligned} & \mathbb{E} \left[\left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} - \mathbb{E} \left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} \right] \left[\left(Z_{j+1}^{H,1} - Z_j^{H,1} \right)^{p_1} - \mathbb{E} \left(Z_{j+1}^{H,1} - Z_j^{H,1} \right)^{p_1} \right] \\ &= c(H, p_1, \alpha^{(1)}) \sum_{\substack{\alpha^{(1)} \in A_{p_1,2} \\ |\alpha^{(1)}| \leq p_1 - 1}} C_{\alpha^{(1)}}^2 \left(2p_1 - 2|\alpha^{(1)}| \right)! \int_i^{i+1} \dots \int_i^{i+1} \int_j^{j+1} \dots \int_j^{j+1} du_1 \dots du_{p_1} \\ & \quad dv_1 \dots dv_{p_1} \prod_{1 \leq a < b \leq p_1} [|u_a - u_b|^{H-1} |v_a - v_b|^{H-1}]^{\alpha_{ab}^{(1)}} \prod_{k=1}^{p_1} |u_k - v_k|^{(H-1)\beta_k^{(1),0}}. \end{aligned}$$

with $\beta_k^{(1),0}, k = 0, 1, \dots, p_1$ defined by (10.7) and corresponding to $\alpha^{(1)}$. After the change of variables $\tilde{u}_k = u_k - i, \tilde{v}_k = v_k - i$ for $i = 1, 2, \dots, p_1$, we get

$$\begin{aligned} & \mathbb{E} \left[\left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} - \mathbb{E} \left(Z_{i+1}^{H,1} - Z_i^{H,1} \right)^{p_1} \right] \left[\left(Z_{j+1}^{H,1} - Z_j^{H,1} \right)^{p_1} - \mathbb{E} \left(Z_{j+1}^{H,1} - Z_j^{H,1} \right)^{p_1} \right] \\ &= c(H, p_1, \alpha^{(1)}) \sum_{\substack{\alpha^{(1)} \in A_{p_1,2} \\ |\alpha^{(1)}| \leq p_1 - 1}} C_{\alpha^{(1)}}^2 \left(2p_1 - 2|\alpha^{(1)}| \right)! \int \dots \int_{[0,1]^{2p_1}} du_1 \dots du_{p_1} dv_1 \dots dv_{p_1} \\ & \quad \prod_{1 \leq a < b \leq p_1} [|u_a - u_b|^{H-1} |v_a - v_b|^{H-1}]^{\alpha_{ab}^{(1)}} \prod_{k=1}^{p_1} |u_k - v_k + i - j|^{(H-1)\beta_k^{(1),0}}. \end{aligned}$$

By applying the above formula for every factor in the right-hand side of (10.49) and by using $|u - v + k| \geq k - 1$ for $k \geq 2$ and $u, v \in [0, 1]$, we obtain, after isolating the diagonal term and the term with $i - j = 1$ in (10.49),

$$\begin{aligned} & \mathbb{E}(\tilde{Y}_{j_1, \dots, j_p})^2 \\ & \leq c_1(H, p)d^{1-2H} + c_2(H, p)d^{-2H} \sum_{\substack{i, j=0 \\ i-j \geq 2}}^{d-1} (i-j-1)^{(H-1)(2p_1-2|\alpha_1|+\dots+2p_x-2|\alpha_x|)} \\ & \leq c_1(H, p)d^{1-2H} + c_2(H, p)d^{1-2H} \sum_{k=2}^{d-1} (k-1)^{(H-1)(2p_1-2|\alpha_1|+\dots+2p_x-2|\alpha_x|)}. \end{aligned}$$

Since $p_j - |\alpha_j| \geq 1$ for $j = 1, \dots, x$ and $x \geq 2$, we find

$$\mathbb{E}(\tilde{Y}_{j_1, \dots, j_p})^2 \leq c_1(H, p)d^{1-2H} + c_2(H, p)d^{1-2H} \sum_{k=1}^{d-2} k^{4H-4}.$$

To finish the proof that $\sum_{k \geq 1} k^{4H-4}$ behaves as a constant if $H \in (\frac{1}{2}, \frac{3}{4})$, as $\log(d)$ if $H = \frac{3}{4}$ and as N^{4H-3} if $H \in (\frac{3}{4}, 1)$. \square

10.5.3 Conclusion

Let us conclude the asymptotic behavior of the random Wishart tensor (10.18). We first define the limiting vector. Let

$$\mathbb{R} = (R_{j_1, \dots, j_p}, j_1, \dots, j_p = 1, \dots, n),$$

where for every $j_1, \dots, j_p = 1, \dots, n$,

$$R_{j_1, \dots, j_p} = \begin{cases} K(H, p)Z_1^H & \text{if } j_1 = \dots = j_p \\ \mathbf{0} & \text{if } j_1, \dots, j_p \text{ are not all the same.} \end{cases}$$

Theorem 17. *Consider the vector \mathbb{Y} given by \tilde{Y} . Then \mathbb{Y} converges componentwise, as $d \rightarrow \infty$, to the random vector \mathbb{R} and for n, d large enough,*

$$d_W(\mathbb{Y}, \mathbb{R}) \leq c(H, p) \begin{cases} n^{\frac{p}{2}} d^{\frac{1}{2}-H} & \text{if } H \in (\frac{1}{2}, \frac{3}{4}) \\ n^{\frac{p}{2}} \sqrt{\log(d)} d^{-\frac{1}{4}} & \text{if } H = \frac{3}{4} \\ n^{\frac{p}{2}} d^{H-1} & \text{if } H \in (\frac{3}{4}, 1). \end{cases}$$

Proof : Via the scaling property of the Rosenblatt process,

$$d_W(\mathbb{Y}, \mathbb{R}) = d_W(\mathbb{Y}^{(1)}, \mathbb{R})$$

where

$$\mathbb{Y}^{(1)} = (Y_{j_1, \dots, j_p}^{(1)}, j_1, \dots, j_p = 1, \dots, n)$$

and

$$Y_{j_1, \dots, j_p}^{(1)} = d^{-H} d^{Hp} \sum_{i=1}^d \left[\prod_{k=1}^p \left(Z_{\frac{i+1}{d}}^{H, j_k} - Z_{\frac{i}{d}}^{H, j_k} \right) - \mathbb{E} \prod_{k=1}^p \left(Z_{\frac{i+1}{d}}^{H, j_k} - Z_{\frac{i}{d}}^{H, j_k} \right) \right].$$

Now, by the definition of the Wasserstein distance,

$$d_W(\mathbb{Y}^{(1)}, \mathbb{R}) \leq \sqrt{\sum_{j_1, \dots, j_p=1}^n \mathbb{E} \left(Y_{j_1, \dots, j_p}^{(1)} - R_{j_1, \dots, j_p} \right)^2}. \quad (10.50)$$

By relation (10.45) in Theorem 16, for $j = 1, \dots, n$,

$$\mathbb{E} \left| Y_{j, \dots, j}^{(1)} - K(H, p) Z_1^H \right|^2 \leq c(H, p) \begin{cases} d^{1-2H}, & \text{if } H \in \left(\frac{1}{2}, \frac{3}{4}\right) \\ \log(d) d^{-\frac{1}{2}} & \text{if } H = \frac{3}{4} \\ d^{2H-2} & \text{if } H \in \left(\frac{3}{4}, 1\right). \end{cases}$$

while when j_1, \dots, j_p are not all the same we have by the scaling property of the Rosenblatt process and Proposition 18, the same bound holds for

$$\mathbb{E} \left| \tilde{Y}_{j_1, \dots, j_p}^{(1)} \right|^2 = \mathbb{E} \left| \tilde{Y}_{j_1, \dots, j_p} \right|^2.$$

By combining these estimates and (10.50),

$$d_W(\mathbb{Y}^{(1)}, \mathbb{R}) \leq c(H, p) n^{\frac{p}{2}} \begin{cases} d^{\frac{1}{2}-H}, & \text{if } H \in \left(\frac{1}{2}, \frac{3}{4}\right) \\ \sqrt{\log(d)} d^{-\frac{1}{4}} & \text{if } H = \frac{3}{4} \\ d^{H-1} & \text{if } H \in \left(\frac{3}{4}, 1\right). \end{cases}$$

Bibliographie

- [1] E. Alos and D. Garcia Lorite (2022) : *Malliavin Calculus in Finance*. Chapman and Hall.
- [2] R. Altmeyer, I. Cialenco and G. Pasemann (2020) : *Parameter estimation for semilinear SPDEs from local measurements*. arXiv preprint, [arXiv:2004.14728](https://arxiv.org/abs/2004.14728).
- [3] A. Ambainis, A. W. Harrow and M.B Hastings (2012) : *Random tensor theory : extending random matrix theory to mixtures of random product states*. Communications in Mathematical Physics 310 (1), pages 25-74.
- [4] A. Anandkumar, R.Ge, D. Hsu, S. M. Kakade, and M. Telgarsky (2014) : *Tensor decompositions for learning latent variable models*. Journal of Machine Learning Research 15, pages 2773-2832.
- [5] H. Araya and C.A. Tudor (2021) : *Asymptotic expansion for the quadratic variations of the solution to the heat equation with additive white noise*. Stochastics and Dynamics 21-2.
- [6] S. Artstein, K.M Ball, R. Barthe and A. Nassor (2004) : *On the convergence rate on entropic central limit theorem*. Probability theory and related fields 129 (3), pages 381-390.
- [7] O. Assaad (2021) : *Solutions des équations différentielles stochastiques : analyse asymptotique par la méthode de Malliavin-Stein et estimation statistique*. <https://pepite-depot.univ-lille.fr>
- [8] O. Assaad, J. Gamain and C. Tudor (2022) : *Quadratic variation and drift parameter estimation for the stochastic wave equation with space-time white noise*. Stochastics and Dynamics 22 (7).
- [9] O. Assaad, D. Nualart, C.A. Tudor and L. Viitasaari (2021) : *Quantitative normal approximations for the stochastic fractional heat equation*. Analysis and Computations 10, pages 223–254.
- [10] O. Assaad and C. A. Tudor (2021) : *Wavelet analysis for the solution to the wave equation with fractional noise in time and white noise in space*. ESAIM Probability and Statistics 25, pages 220-257.
- [11] O. Assaad and C. A. Tudor (2021) : *Pathwise analysis and parameter estimation for the stochastic Burgers equation*. Bulletin des Sciences Mathématiques 170.
- [12] A. Ayache (2020) : *Lower bound for local oscillations of Hermite processes*. Stochastic Processes and their applications 130 (8), pages 4593-4607.
- [13] R. Balan (2018) : *A gentle introduction to SPDEs : the random field approach*. <https://arxiv.org/abs/1812.02812>.

-
- [14] M. Bibinger and M. Trabs (2020) : *Volatility estimation for stochastic PDEs using high-frequency observations*. Stochastic Processes and their Applications 130 (5), pages 3005-3052.
- [15] M. Bibinger and M. Trabs (2019) : *On central limit theorems for power variations of the solution to the stochastic heat equation*. Stochastic models, statistics and their applications, pages 69–84.
- [16] H. Biermé, A. Bonami, I. Nourdin and G. Peccati (2012) : *Optimal Berry-Esseen rates on the Wiener space : the barrier of third and fourth cumulants*. ALEA 9 (2), pages 473-500.
- [17] A. N. Bishop, P. Del Moral, and A. Niclas (2018) : *An Introduction to Wishart Matrix Moments*. MAL 11 (2), pages 97-218.
- [18] S. Bourguin and T. Dang (2022) : *High-dimensional regimes for sequences of non-stationary Gaussian correlated Wishart matrices*. Random Matrices : Theory and Applications 1 (1).
- [19] S. Bourguin, Ch-Ph. Diez and C.A. Tudor (2021) : *Limiting behavior of large correlated Wishart matrices with chaotic entries*. Bernoulli 27 (2), pages 1077–1102.
- [20] S. Bubeck, J. Ding, R. Eldan and M.Z. Racz (2015) : *Testing high-dimensional geometry in random graphs*. Random Structures and Algorithms 49 (3), pages 503-532.
- [21] S. Bubeck and S. Ganguly (2016) : *Entropic CLT and phase transition in high dimensional regime Wishart matrix*. International Mathematics Research Notices, pages 1-40.
- [22] J. Bryson, R. Vershynin and H. Zhao (2021) : *Marchenko-Pastur law with relaxed independence conditions*. [arXiv:1912.12724v3](https://arxiv.org/abs/1912.12724v3).
- [23] Z.D. Bai and Y.Q. Yin (1988) : *Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix*. Annals of Probability 16 (4), pages 1729-1741.
- [24] S. Chatterjee and E. Meckes (2008) : *Multivariate normal approximation using exchangeable pairs*. ALEA 4, pages 257-283.
- [25] L.H.Y Chen, L. Goldstein, and Q.M. Shao (2010) : *Normal Approximation by Stein's method*. Probability and its Applications. Springer, Heidelberg.
- [26] C. Chong and R. Dalang (2022) : *Power variations in fractional Sobolev spaces for a class of parabolic stochastic PDEs*. arXiv preprint, [arXiv:2006.15817](https://arxiv.org/abs/2006.15817).
- [27] C. Chong (2020) : *High-frequency analysis of parabolic stochastic PDEs*. The Annals of Statistics 48 (2), pages 1143 - 1167.
- [28] C. Chong (2019) : *High-frequency analysis of parabolic stochastic PDEs with multiplicative noise : Part I*. arXiv preprint [arXiv:1908.04145](https://arxiv.org/abs/1908.04145).
- [29] P-L. Chow (2007) : *Stochastic partial differential equations*. Chapman and Hall. CRC Applied Mathematics and Nonlinear Science Series. Chapman and Hall.
- [30] A. Chronopoulou, C. A. Tudor and F. G. Viens (2001) : *Self-similarity parameter estimation and reproduction property for non-Gaussian Hermite processes*. Communications on Stochastic Analysis 5 (1), pages 161-185.
- [31] I. Cialenco (2018) : *Statistical inference for spdes : an overview*. Statistical Inference for Stochastic Processes 21 (2), pages 309-329.
- [32] I. Cialenco, F. Delgado-Vences, H.-J. Kim (2020) : *Drift estimation for discretely sampled SPDEs*. Stochastics and Partial Differential Equations : Analysis and Computations 8 (2), pages 895-920.
- [33] I. Cialenco and Y. Huang (2020) : *A note on parameter estimation for discretely sampled spdes*. Stochastics and Dynamics 20 (3).
-

-
- [34] I. Cialenco, H.-J. Kim (2022) : *Parameter estimation for discretely sampled stochastic heat equation driven by space-only noise*. Stochastic Processes and their Applications 143 (5), pages 1–30.
- [35] I. Cialenco, H.-J. Kim, and G. Pasemann (2021) : *Statistical analysis of discretely sampled semilinear SPDEs : a power variation approach*. arXiv preprint [arXiv:2103.04211](https://arxiv.org/abs/2103.04211).
- [36] F. Cosmao, F. Dupuy and A. Guillon (2002) : *Le Calcul de Malliavin Appliqué à la Finance*. Groupe de Travail Dirigé par Jean-Frédéric Jouanin, Ashkan Nikeghbali et Thierry Roncalli.
- [37] R. Couillet and M. Debbah (2011) : *Random Matrix Methods for Wireless Communications*. Cambridge University Press.
- [38] X. Shi, R. Qiu, X. He, Z. Ling, H. Yang and L. Chu (2020) : *Early Anomaly Detection and Localization in Distribution Network : A Data-Driven Approach*. arXiv preprint [arXiv:1801.01669v4](https://arxiv.org/abs/1801.01669v4).
- [39] R. Dalang (1999) : *Extending the martingale measure stochastic integral with applications to spatially homogeneous S.P.D.E.'s*. Electronic Journal of Probability 4 (6).
- [40] R. Dalang, D. Khoshnevisan, C. Mueller, D. Nualart, Y. Xiao (2006) : *A minicourse on stochastic partial differential equations*. [Utah-Summer-School.pdf](https://www.math.utah.edu/~dalan/Utah-Summer-School.pdf)
- [41] L. Debbi and M. Dozzi (2005) : *On the solutions of nonlinear stochastic fractional partial differentialequations in one spatial dimension*. Stochastic Processes and their Applications 115, pages 1761-1781.
- [42] L. Decreusefond (2022) : *Selected Topics in Malliavin calculus*. Bocconi and Springer Series 10.
- [43] L. Decreusefond and A.S. Üstünel (1998) : Stochastic analysis of the fractional Brownian motion. Potential Analysis 10, pages 177-214.
- [44] R.L. Dobrushin and P. Major (1979) : Non-central limit theorems for non-linear functionals of Gaussian fields. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 50, pages 27-52.
- [45] Ch-Ph. Diez and C. A. Tudor (2021) : *Non-central limit theorem for large Wishart matrices with Hermite entries*. Journal of Stochastic Analysis 2 (1).
- [46] Ch-Ph. Diez and C. A. Tudor (2021) : *Limit behavior for Wishart matrices with Skorohod integrals*. ALEA Lat. Am. J. Probab. Math. Stat. 18 (2), pages 1625–1641.
- [47] R. Dhoyer and C. A. Tudor (2023) : *Limit behavior in high-dimensional regime for the Wishart tensors in Wiener chaos*. Preprint.
- [48] K. Es-Sebaiy and C. A. Tudor (2011) : *Noncentral limit theorem for cubic variation of a class of self-similar stochastic processes*. Theory of Probability and Its Applications 55 (3), pages 411-431.
- [49] X. Fang and Y. Koike (2020) : *New error bounds in multivariate normal approximations via exchangeable pairs with applications to Wishart matrices and fourth moment theorems*. Preprint.
- [50] Z. Füredi and J. Komlos (1981) : *The eigenvalues of random symmetric matrices*. Combinatorica 1 (3), pages 233-241.
- [51] M. Foodun, D. Khoshnevisan, and P. Mahboubi (2015) : *Analysis of the gradient of the solution to a stochastic heat equation via fractional Brownian motion*. Stochastic Partial Differential Equations : Analysis and Computations 3 (2), pages 133-158.
- [52] J. Gamain, D.A.C Mollinedo and C. Tudor (2022) : *High-dimensional regime for Wishart matrices based on the increments of the solution to the stochastic heat equation*. Brazilian Journal of Probability and Statistics.
-

-
- [53] J. Gamain and C. Tudor (2022) : *Random matrices and the stochastic wave equation*. Submitted.
- [54] J. Gamain and C. Tudor (2023) : *Limit behavior in high-dimensional regime for Wishart tensors with Rosenblatt entries*. Submitted.
- [55] J. Gamain and C. Tudor (2022) : *Exact variation and drift parameter estimation for the nonlinear fractional stochastic heat equation*. Japanese Journal of Statistics and Data Science.
- [56] H. Halconruy (2021) : Calcul de Malliavin et structures de Dirichlet pour des variables aléatoires indépendantes. <https://www.theses.fr>
- [57] D. Harnett and D. Nualart (2018) : *Central limit theorem for functionals of a generalized self-similar Gaussian process*. Stochastic Processes and their Applications 128 (2), pages 404-425.
- [58] V. Garino, I. Nourdin, D. Nualart and M. Salamat (2021) : *Limit theorems for integral functionals of Hermite-driven processes*. Bernoulli 27 (3), pages 1764-1788.
- [59] S. Gaudlitz and M. Reiss (2022) : *Estimation for the reaction term in semi-linear SPDEs under small diffusivity*. arXiv preprint [arXiv:2203.10527](https://arxiv.org/abs/2203.10527).
- [60] F. Hildebrandt and M. Trabs (2021) : *Parameter estimation for SPDEs based on discrete observations in time and space*. Electronic Journal of Statistics 15 (1), pages 2716-2776.
- [61] J. Huang, D. Nualart and L. Viitasaari (2020) : *A central limit theorem for the stochastic heat equation*. Stochastic Processes and their Applications 130 (12), pages 7170-7184.
- [62] H.E. Hurst (1951) : *Long-term storage capacity in reservoirs*. Transactions of the American Society of Civil Engineers 116 (1), pages 400-410.
- [63] C. Houdré and J. Villa (2003) : *An example of infinite dimensional quasi-helix*. Contemporary Mathematics, Amer. Math. Soc. 336, pages 195-201.
- [64] N. Jacob and H.G. Leopold (1993) : *Pseudo differential operators with variable order of differentiation generating Feller semigroups*. Integral Equations Operator Theory 17, pages 544-553.
- [65] N. Jacob, A. Potrykus and J.-L. Wu (2010) : *Solving a non-linear stochastic pseudo-differential equation of Burgers type*. Stochastic Processes and their Applications 120, pages 2447-2467.
- [66] J. Janák (2021) : *Parameter estimation for stochastic wave equation based on observation window*. Acta Applicandae Mathematicae 172 (2).
- [67] T. Jiang (2004) : *The asymptotic distributions of the largest entries of sample correlation matrices*. The Annals of Applied Probability 14 (2), pages 865-880.
- [68] T. Jiang and D. Li (2015) : *Approximation of rectangular beta-laguerre ensembles and large deviations*. Journal of Theoretical Probability 28 (3), pages 804-847.
- [69] T. Jiang and J. Xie (2019) : *Limiting behavior of largest entry of random tensor constructed by high-dimensional data*. Journal of Theoretical Probability 33 (1), pages 1-21.
- [70] I. M. Johnstone (2007) : *High dimensional statistical inference and random matrices*, International Congress of Mathematicians. Zurich : European Mathematical Society Publishing House 1, pages 30-333.
- [71] J.P. Kahane (1981) : *Hélices et quasi-hélices*. Advances in Mathematics B7, pages 417-433.
- [72] Y. Kaino and M. Uchida (2021) : *Parametric estimation for a parabolic linear SPDE model based on sampled data*. Journal of Statistical Planning and Inference 211, pages 190-220.
-

-
- [73] M. Khalil and C. A. Tudor (2018) : *Correlation structure, quadratic variations and parameter estimation for the solution to the wave equation with fractional noise*. Electronic Journal of Statistics 12 (2), pages 3639-3672.
- [74] M. Khalil, C.A. Tudor and M. Zili (2018) : *Spatial variation for the solution to the stochastic linear wave equation driven by additive space-time white noise*. Stochastics and Dynamics 18 (5).
- [75] Z. Khalil-Mahdi and C. A. Tudor (2019) : *On the distribution and q -variation of the solution to the heat equation with Fractional Laplacian*. Probability Theory and Mathematical Statistics 39 (2), pages 315–335.
- [76] Z. Khalil-Mahdi and C. A. Tudor (2019) : *Estimation of the drift parameter for the fractional stochastic heat equation via power variation*. Modern Stochastics : Theory and Applications 6 (4), pages 397–417.
- [77] A.N. Kolmogorov (1940) : *Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum*. Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS (N.S.) 26, pages 115–118.
- [78] N. Krylov (2006) : *On the foundation of the L_p -theory of stochastic partial differential equations*. Stochastic Partial Differential Equations and Applications-VII, Lect. Notes Pure Appl. Math. 245, pages 179–191, Chapman.
- [79] M. Ledoux, I. Nourdin and G. Peccati (2015) : *Stein's method, logarithmic Sobolev and transport inequalities*. Geometric and Functional Analysis 25 (1), pages 256-306.
- [80] P. Lei and D. Nualart (2008) : *A decomposition of the bifractional Brownian motion and some applications*. Statistics and Probability Letters 79 (5), pages 619-624.
- [81] W. Liu, S. V. Lototski (2010) : *Estimating speed and damping in the stochastic wave equation*. Stochastic partial differential equations and applications, pages 191-206, Quad. Mat., 25, Dept. Math., Seconda Univ. Napoli, Caserta.
- [82] W. Liu, S. V. Lototski (2009) : *Parameter estimation in diagonalizable stochastic hyperbolic equations*. arXiv preprint, [arXiv:0906.4353](https://arxiv.org/abs/0906.4353) .
- [83] S. V. Lototski (2009) : *Statistical inference for stochastic parabolic equations : a spectral approach*. Publicacions Matemàtiques 53 (1), pages 3-45.
- [84] H.M. Luk (1994) : *Stein's method for the Gamma distribution and related statistical applications*. ProQuest LLC, Ann Arbor, MI, 1994. Thesis (Ph.D.) University of Southern California.
- [85] A. Lytova (2018) : *Central Limit theorem for linear eigenvalue statistics for a tensor product version of sample covariance matrices*. Journal of Theoretical Probability 31 (2), pages 1024-1057.
- [86] P. Malliavin (1978) : *Stochastic Calculus of variations and hypoelliptic operators*. Proceedings of the International Symposium on Stochastic Differential Equations, pages 195-263.
- [87] M. Maejima and C. A. Tudor (2007) : *Wiener integrals with respect to the Hermite process and a Non-Central Limit Theorem*. Stochastic Analysis and Applications 25 (5), pages 1043 - 1056.
- [88] B.B. Mandelbrot and J.W. Van Ness (1968) : *Fractional Brownian motions, fractional noises and applications*. SIAM Review 10, pages 422–437.
- [89] V.A Marchenko and L.A Pastur (1968) : *Distribution of eigenvalues in certain sets of random matrices*. Mathematics of the USSR-Sbornik 1 (4), pages 507-536.
-

-
- [90] D. Mikulincer (2020) : *A CLT in Stein's distance for generalized Wishart matrices and higher order tensors*. International Mathematics Research Notices 2022 (10), pages 7839–7872.
- [91] B. Markussen (2003) : *Likelihood inference for a discretely observed stochastic partial differential equation*. Bernoulli 9 (5), pages 745–762.
- [92] I. Nourdin (2012) : *Selected Aspects of Fractional Brownian Motion*. Bocconi and Springer Series.
- [93] I. Nourdin and G. Peccati (2012) : *Normal Approximations with Malliavin Calculus From Stein's Method to Universality*. Cambridge University Press.
- [94] I. Nourdin, G. Peccati (2009) : *Stein's method on Wiener chaos*. Probability Theory and Related Fields 145 (1-2), pages 75–118.
- [95] I. Nourdin and G. Peccati (2015) : *The optimal fourth moment theorem*. Proceedings of the American Mathematical Society 143 (7), pages 3123–3133.
- [96] I. Nourdin, G. Peccati and A. Réveillac (2008) : *Multivariate normal approximation using Stein's method and Malliavin calculus*. Annales de l'Institut Henri Poincaré Probabilités et Statistiques 46 (1).
- [97] I. Nourdin and Fei Pu (2022) : *Gaussian fluctuation for Gaussian Wishart matrices of overall correlation*. Statistics & Probability Letters 181 (2).
- [98] I. Nourdin and G. Zheng (2021) : *Asymptotic behavior of large Gaussian correlated Wishart matrices*. Journal of Theoretical Probability 35 (3), pages 1–30.
- [99] D. Nualart and G. Peccati (2005) : *Central limit theorems for sequences of multiple stochastic integrals*. The Annals of Probability 33 (1), pages 177–193.
- [100] D. Nualart (2006) : *Malliavin Calculus and Related Topics*. Second Edition. Springer.
- [101] D. Nualart and S. Ortiz-Latorre (2007) : *Central limit theorem for multiple stochastic integrals and Malliavin calculus*. Stochastic Processes and their Applications 118 (4), pages 614–628.
- [102] G. Peccati and C.A. Tudor (2005) : *Gaussian limits for vector-valued multiple stochastic integrals*. Séminaire de Probabilités XXXVIII, pages 219–245.
- [103] V. Pipiras and M. Taqqu (2017) : *Long-range dependence and self-similarity*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.
- [104] J. Pospisil and R. Tribe (2007) : *Parameter estimates and exact variations for stochastic heat equation heat equations driven by space-time white noise*. Stochastic Analysis and Applications 25 (3), pages 593–611.
- [105] C. E. Rasmussen and C. K. I. Williams (2006) : *Gaussian processes for machine learning*. Adaptive Computation and Machine Learning MIT Press, Cambridge MA.
- [106] M.Z Racz and J. Richey (2018) : *A smooth transition form Wishart to GOE*. Journal of Theoretical Probability 32 (2), pages 898–906.
- [107] F. Russo and C.A. Tudor (2005) : *On the bifractionnal Brownian motion*. Stochastic Processes and their Applications 116 (6), pages 830–856.
- [108] M. Sanz-Solé (2016) : *An Introductory Course on Stochastic Partial Differential Equations*. EPFL, February–May 2016.
- [109] R. Shevchenko, M. Slaoui and C. A. Tudor (2020) : *Generalized k -variations and Hurst parameter estimation for the fractional wave equation via Malliavin calculus*. Journal of Statistical Planning and Inference 207 (3), pages 155–180.
-

-
- [110] Ch. Stein (1972) : *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables*. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability II, pages 583-602.
- [111] J. Swanson (2007) : *Variations of the solution to a stochastic heat equation*. The Annals of Probability 35 (6), pages 2122-2159.
- [112] M. Taqqu (1979) : *Convergence of integrated processes of arbitrary Hermite rank*. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 50, pages 53-83.
- [113] M. Taqqu (1975) : *The Rosenblatt Process*. Springer Science Business Media LLC 2011.
- [114] D. Tieplova (2016) : *Distribution of eigenvalues of sample covariance matrices with tensor product samples*. Zurnal matematicheskoi fiziki analiza geometrii 13 (1), pages 1024-1057.
- [115] C. A. Tudor (2008) : *Analysis of the Rosenblatt process*. ESAIM : Probability and Statistics 12, pages 230-257.
- [116] C. A. Tudor (2013) : *Analysis of variations for self-similar processes. A stochastic calculus approach*. Probability and its Applications (New York). Springer, Cham.
- [117] C.A. Tudor (2022) : *Stochastic Partial Differential Equations with Additive Gaussian Noise*. World Scientific.
- [118] F. Trèves (1975) : *Basic Linear Partial Differential Equations*. Academic Press.
- [119] S. Torres, C.A. Tudor, F.G. Viens (2014) : *Quadratic variations for the fractional-colored stochastic heat equation*. Electronic Journal of Probability 19 (76), pages 1-51.
- [120] C. A. Tudor and F. Viens (2009) : *Variations and estimators for selfsimilarity parameters via Malliavin calculus*. The Annals of Probability 37 (6), pages 2093–213.
- [121] C. A. Tudor and Y. Xiao (2017) : *Sample paths of the solution to the fractional-colored stochastic heat equation*. Stochastics and Dynamics 17 (1).
- [122] C.A. Tudor and N. Yoshida (2018) : *High order asymptotic expansion for Wiener functionals*. Stochastic Processes and their Applications 164 (October 2023), pages 443-492
- [123] A.S. Ustunel and M. Zakai (1989) : *On Independence and Conditioning On Wiener Space*. The Annals of Probability 17 (4), pages 1441-1453.
- [124] R. Vershynin (2019) : *Concentration inequalities for random tensors*. Bernoulli 26 (4), pages 3139-3162
- [125] J.B. Walsh (1986) : *An introduction to stochastic partial differential equations*. École d'Été de Probabilités de Saint-Flour XIV—1984, pages 265-439.
- [126] E. Wigner (1955) : *Characteristic vectors of bordered matrices with infinite dimensions*. Annals of Mathematics 62 (3), pages 548-564.
- [127] J. Wishart (1928) : *The generalized product moment distribution in samples from a normal multivariate population*. Biometrika 20A (1-2), pages 32-52.
- [128] Jérémy Zurcher (2022) : *Convergence in law in Wiener chaos*. <https://drive.google.com/file>.
- [129] M. Zili and E. Zougar (2019) : *Exact variations for stochastic heat equations with piecewise constant coefficients and applications to parameter estimation*. Theory of Probability and Mathematical Statistics 100, pages 75–101.
-