Communications on Stochastic Analysis
Vol. 7, No. 3 (2013) 349-372

Serials Publications
www.serialspublications.com

STEIN'S METHOD FOR BROWNIAN APPROXIMATIONS

L. COUTIN* AND L. DECREUSEFOND*

Abstract. Motivated by a theorem of Barbour, we revisit some of the classical limit theorems in probability from the viewpoint of the Stein method. We setup the framework to bound Wasserstein distances between some distributions on infinite dimensional spaces. We show that the convergence rate for the Poisson approximation of the Brownian motion is as expected proportional to $\lambda^{-1/2}$ where λ is the intensity of the Poisson process. We also exhibit the speed of convergence for the Donsker Theorem and for the linear interpolation of the Brownian motion.

1. Introduction

Among the classics in probability theory, one can cite the approximation in distribution of a Brownian motion by a normalized compensated Poisson process of intensity going to infinity or the celebrated Donsker theorem which says that a symmetric random walk conveniently normalized approaches a Brownian motion in distribution. Though the topology of the convergence in distribution is known to derive from a distance on the space of probability measures, to the best of our knowledge, we are aware of only one result precising the speed of convergence in one of these two theorems. In [1], Barbour estimated the distance between the distribution of a normalized compensated Poisson process of intensity λ and the distribution of a Brownian motion. The common space on which these two processes are compared is taken as the space of rcll functions, denoted by $\mathfrak{D}([0,\,1],\,\mathbf{R})$ equipped with the distance:

$$d_0(\omega,\,\eta)=\inf_{\Phi\in\text{Hom}([0,1])}(\|\omega\circ\Phi-\eta\|_\infty+\|\Phi-\text{Id}\,\|_\infty),$$

where $\text{Hom}([0,1])$ is the set of increasing homeomorphisms of $[0,1]$. It is proved in [1] that the speed of convergence is not $\lambda^{-1/2}$ as expected but that there exists a non negligible corrective term. This additional term exists because the sample-paths of the two processes do not really belong to the same space: Continuous functions are a rather special class of rcll functions and sample-paths of Poisson process even normalized are never continuous whatever the value of the intensity. Thus there seems to be an unavoidable gap between the two kind of trajectories in the considered approximation. Actually, the additional term is related to the

Received 2012-9-5; Communicated by the editors.
2010 *Mathematics Subject Classification.* Primary 60F15; Secondary 60H07, 60G15, 60G55, 60F05.
Key words and phrases. Donsker theorem, Malliavin calculus, Stein's method, Wasserstein distance.
* Both authors were partially supported by ANR-10-BLAN-0121.

modulus of continuity of the Brownian motion, i.e. in some sense, it measures the cost to approximate a continuous function by a purely discontinuous one.

We circumvent this problem by considering Poisson and Brownian sample-paths as elements of the same space. In fact, the Poisson sample-paths, like the trajectories of the other processes we are considering in this paper, belong to a much smaller space than $\mathfrak{D}([0,\,1],\,\mathbf{R})$: They all are piecewise differentiable, i.e. of the form $\sum_{n\in\mathbf{N}} r_n(t-t_n)\mathbf{1}_{[t_n,\,t_{n+1})}$ where $(t_n,\,n\geq 1)$ is an increasing sequence of real and r_n are differentiable functions. An indicator function is not continuous but it has more property than being rcll. In particular, it belongs to $\mathcal{I}_{\beta,\,p}$ for any $p\geq 1$ and any $\beta<1/p$ (see below for the definition of $\mathcal{I}_{\beta,\,p}$). On the other hand, Brownian trajectories are $(1/2-\epsilon)$-Hlder continuous so that they belong to $\mathcal{I}_{\beta,\,p}$ any $p\geq 1$ and any $\beta<1/2$. Therefore, the natural candidates to support both the distribution of piecewise differentiable processes and that of the Brownian motion are the spaces $\mathcal{I}_{\beta,\,2}$ for any $\beta<1/2$. The original problem is then reduced to the computation of the distance between between a given distribution and a Gaussian law on some Hilbert space.

The Stein method is known for a long time to give the speed of convergence of many Gaussian approximations (see for instance [3] and references therein). The usual approach requires some sort of coupling to derive the pertinent estimates. It is only recently that the mixing of Stein approach and Malliavin calculus proved its efficiency (see [14] for a thorough analysis of this line of thought): The search of ad-hoc couplings in the Stein method is here bypassed by using integration by parts formula in the sense of Malliavin calculus. In particular, it has been used for approximations of point processes functionals [5, 6, 15]. But to the best of our knowledge, up to the notable exceptions of [1] and [20], all these investigations consider finite dimensional Gaussian random variables. We here develop the framework for a Stein theory on Hilbert spaces thus circumventing many of the technicalities of [20] which considers Banach valued random variables. Our approach requires two types of Malliavin *gradients* : One used to characterize the target (Gaussian) measure, one built on the probability space of the measure to be compared to the Gaussian measure, used to perform the necessary integration by parts. For the impatient reader, the actual method can be explained informally in dimension 1. Imagine that we want to precise the speed of convergence of the well-known limit in distribution:

$$\frac{1}{\sqrt{\lambda}}(X_\lambda-\lambda)\xrightarrow{\lambda\to\infty}\mathcal{N}(0,\,1),$$

where X_λ is a Poisson random variable of parameter λ. We consider the Wasserstein distance between the distribution of $\tilde{X}_\lambda=\lambda^{-1/2}(X_\lambda-\lambda)$ and $\mathcal{N}(0,\,1)$, which is defined as

$$\operatorname{dist}_W(\tilde{X}_\lambda,\,\mathcal{N}(0,\,1))=\sup_{F\in 1-\mathrm{Lip}}\mathbb{E}\left[F(\tilde{X}_\lambda)\right]-\mathbb{E}\left[F(\mathcal{N}(0,\,1))\right], \tag{1.1}$$

where $1-\mathrm{Lip}$ is the set of one Lipschitz function from $\mathbf{R}$ into itself. The well known Stein Lemma stands that for any $F\in 1-\mathrm{Lip}$, there exists $H_F\in\mathcal{C}_b^2$ such that for all $x\in\mathbf{R}$,

$$F(x)-\mathbb{E}\left[F(\mathcal{N}(0,\,1))\right]=x\,H_F(x)-H_F'(x).$$

Moreover,

$$\|H'_F\|_\infty \le 1,\ \|H''_F\|_\infty \le 2.$$

Hence, instead of the right-hand-side of (1.1), we are lead to estimate

$$\sup_{\|H'\|_\infty\le 1,\ \|H''\|_\infty\le 2} \mathbb{E}\left[\tilde{X}_\lambda H(\tilde{X}_\lambda) - H'(\tilde{X}_\lambda)\right]. \tag{1.2}$$

This is where the Malliavin-Stein approach differs from the classical line of thought. In order to transform the last expression, instead of constructing a coupling, we resort to the integration by parts formula for functionals of Poisson random variable. The next formula can be checked by hand or viewed as a consequence of (5.1):

$$\mathbb{E}\left[\tilde{X}_\lambda G(\tilde{X}_\lambda)\right] = \sqrt{\lambda}\ \mathbb{E}\left[G(\tilde{X}_\lambda + 1/\sqrt{\lambda}) - G(\tilde{X}_\lambda)\right].$$

Hence (1.2) is transformed into

$$\sup_{\|H'\|_\infty\le 1,\ \|H''\|_\infty\le 2} \mathbb{E}\left[\sqrt{\lambda}(H(\tilde{X}_\lambda + 1/\sqrt{\lambda}) - H(\tilde{X}_\lambda)) - H'(\tilde{X}_\lambda)\right]. \tag{1.3}$$

According to the Taylor formula

$$H(\tilde{X}_\lambda + 1/\sqrt{\lambda}) - H(\tilde{X}_\lambda) = \frac{1}{\sqrt{\lambda}}H'(\tilde{X}_\lambda) + \frac{1}{2\lambda}H''(\tilde{X}_\lambda + \theta/\sqrt{\lambda}),$$

where $\theta \in (0,1)$. If we plug this expansion into (1.3), the term containing H' is miraculously vanishing and we are left with only the second order term. This leads to the estimate (compare to Theorem 5.1):

$$\operatorname{dist}_W\left(\tilde{X}_\lambda,\ \mathcal{N}(0,\,1)\right) \le \frac{1}{\sqrt{\lambda}}\cdot$$

The remainder of this paper consists in generalizing these computations to the infinite dimensional setting. We show that our method is applicable in three different settings: Whenever the alea on which the approximate process is built upon is either the Poisson space, the Rademacher space or the Wiener space.

This paper is organized as follows. After some preliminaries, we construct the Wiener measure on the Besov-Liouville spaces and $l^2(\mathbf{N})$, using the Itô-Nisio Theorem. Section 4 is devoted to the development of the abstract version of the Stein method for Hilbert valued random variables. In Section 5 to Section 7, we exemplify this general scheme of reasoning successively for the Poisson approximation of the Brownian motion, for the linear interpolation of the Brownian motion and for the Donsker theorem. In Section 8, we show that by a transfer principle, similar results can be obtained for other Gaussian processes like the fractional Brownian motion, extending some earlier results [7].

2. Preliminaries

2.1. Tensor products of Hilbert spaces. For X and Y two Hilbert spaces, $\mathfrak{B}(X,Y)$ is the set of multilinear complex-valued forms over $X \times Y$. For $x \in X$ and $y \in Y$, the bilinear form $x \otimes y$ is defined by

$$x \otimes y(f,g) = \langle x,\ f\rangle_X\ \langle y,\ g\rangle_Y,$$

for any $f \in X$ and $g \in Y$. We denote by $\mathfrak{B}_f(X,Y)$, the linear span of such simple bilinear forms. It is equipped with the norm

$$\|\sum_{i=1}^N \alpha_i x_i \otimes y_i\|^2_{\mathfrak{B}_f(X,Y)} = \sum_{i=1}^N |\alpha_i|^2 \|x_i\|_X^2 \|y_i\|_Y^2.$$

The tensor product $X \otimes Y$ is the completion of $\mathfrak{B}_f(X,Y)$ with respect to this norm. A continuous linear map A from X to Y can be viewed as an element of $X \otimes Y$ by the identification :

$$\tilde{A}(f,g) = \langle Af,\, g\rangle_Y \text{ for } f \in X \text{ and } g \in Y.$$

Conversely, for $x \in X$ and $y \in Y$, the operator $x \otimes y$ can be seen either as an element of $X \otimes Y$ or as a continuous map from X into Y via the identification :

$$x \otimes y(f) = \langle x,\, f\rangle_X\, y.$$

We recall that for X an Hilbert space and A a linear continuous map from X into itself, A is said to be trace-class whenever the series $\|A\|_{\mathcal{S}_1} := \sum_{n\ge 1} |(Af_n,\, f_n)_X|$ is convergent for one (hence any) complete orthonormal basis $(f_n,\, n \ge 1)$ of X. When A is trace-class, its trace is defined as $\operatorname{trace}(A) = \sum_{n\ge 1}(Af_n,\, f_n)_X$. It is then straightforward that for $x,\, y \in X$, the operator $x \otimes y$ is trace-class and that $\operatorname{trace}(x \otimes y) = \sum_{n\ge 1}(y,\, f_n)_X(x,\, f_n)_X = \langle x,\, y\rangle_X$ according to the Parseval formula. The trace-class operators is a two sided ideal of the set of bounded compact operators: If A is trace-class and B is bounded, then $A \circ B$ is trace-class and (see [21])

$$|\operatorname{trace}(A \circ B)| \le \|A \circ B\|_{\mathcal{S}_1} \le \|A\|_{\mathcal{S}_1} \|B\|, \tag{2.1}$$

where $\|B\|$ is the operator norm of B. It is easily seen that when A is symmetric and non-negative, $\|A\|_{\mathcal{S}_1}$ is equal to $\operatorname{trace}(A)$. We also need to introduce the notion of partial trace. For any vector space X, $\operatorname{Lin}(X)$ is the set of linear operator from into itself. For X and Y two Hilbert spaces, the partial trace operator along X can be defined as follows: it is the unique linear operator

$$\operatorname{trace}_X \;:\; \operatorname{Lin}(X \otimes Y) \longrightarrow \operatorname{Lin}(Y)$$

such that for any $R \in \operatorname{Lin}(Y)$, for any trace class operator S on X,

$$\operatorname{trace}_X(S \otimes R) = \operatorname{trace}_X(S)\, R.$$

2.2. Besov-Liouville spaces. This part is devoted to the presentation of the so-called Besov-Liouville spaces. A complete exposition can be found in [18]. For $f \in \mathcal{L}^p([0,1];\; dt)$, (denoted by $\mathcal{L}^p$ for short) the left and right fractional integrals of f are defined by :

$$(I^\alpha_{0^+} f)(x) = \frac{1}{\Gamma(\alpha)} \int_0^x f(t)(x-t)^{\alpha-1}\, \mathrm{d}t \;,\; x \ge 0,$$
$$(I^\alpha_{1^-} f)(x) = \frac{1}{\Gamma(\alpha)} \int_x^1 f(t)(t-x)^{\alpha-1}\, \mathrm{d}t \;,\; x \le 1,$$

where $\alpha > 0$ and $I^0_{0^+} = I^0_{1^-} = \mathrm{Id}$. For any $\alpha \geq 0$, $p, q \geq 1$, any $f \in \mathcal{L}^p$ and $g \in \mathcal{L}^q$ where $p^{-1} + q^{-1} \leq \alpha + 1$, we have :

$$\int_0^1 f(s)(I^\alpha_{0^+} g)(s)\,\mathrm{d}s = \int_0^1 (I^\alpha_{1^-} f)(s) g(s)\,\mathrm{d}s. \tag{2.2}$$

For $p \in [1, +\infty]$, the Besov-Liouville space $I^\alpha_{0^+}(\mathcal{L}^p) := \mathcal{I}^+_{\alpha,p}$ is usually equipped with the norm :

$$\|I^\alpha_{0^+} f\|_{\mathcal{I}^+_{\alpha,p}} = \|f\|_{\mathcal{L}^p}. \tag{2.3}$$

Analogously, the Besov-Liouville space $I^\alpha_{1^-}(\mathcal{L}^p) := \mathcal{I}^-_{\alpha,p}$ is usually equipped with the norm :

$$\|I^{-\alpha}_{1^-} f\|_{\mathcal{I}^-_{\alpha,p}} = \|f\|_{\mathcal{L}^p}.$$

We then have the following continuity results (see [9, 18]) :

Theorem 2.1. *i. If $0 < \alpha < 1$, $1 < p < 1/\alpha$, then $I^\alpha_{0^+}$ is a bounded operator from $\mathcal{L}^p$ into $\mathcal{L}^q$ with $q = p(1 - \alpha p)^{-1}$.*

ii. For any $0 < \alpha < 1$ and any $p \geq 1$, $\mathcal{I}^+_{\alpha,p}$ is continuously embedded in $\mathrm{Hol}_0(\alpha - 1/p)$ provided that $\alpha - 1/p > 0$. $\mathrm{Hol}_0(\nu)$ denotes the space of α Hölder-continuous functions, null at time 0, equipped with the usual norm.

iii. For any $0 < \alpha < \beta < 1$, $\mathrm{Hol}_0(\beta)$ is compactly embedded in $\mathcal{I}_{\alpha,\infty}$.

iv. By $I^{-\alpha}_{0^+}$, respectively $I^{-\alpha}_{1^-}$, we mean the inverse map of $I^\alpha_{0^+}$, respectively $I^\alpha_{1^-}$. The relation $I^\alpha_{0^+} I^\beta_{0^+} f = I^{\alpha+\beta}_{0^+} f$, respectively $I^\alpha_{1^-} I^\beta_{1^-} f = I^{\alpha+\beta}_{1^-} f$, holds whenever $\beta > 0$, $\alpha + \beta > 0$ and $f \in \mathcal{L}^1$.

v. For $\alpha p < 1$, the spaces $\mathcal{I}^+_{\alpha,p}$ and $\mathcal{I}^-_{\alpha,p}$ are canonically isomorphic. We will thus use the notation $\mathcal{I}_{\alpha,p}$ to denote any of this spaces.

We now recall the definition and properties of Besov-Liouville spaces of negative orders. The proofs can be found in [4].

Denote by $\mathcal{D}_+$ the space of $\mathcal{C}^\infty$ functions defined on $[0, 1]$ and such that $\phi^{(k)}(0) = 0$, for all $k \in \mathbf{N}$. Analogously, set $\mathcal{D}_-$ the space of $\mathcal{C}^\infty$ functions defined on $[0, 1]$ and such that $\phi^{(k)}(1) = 0$, for all $k \in \mathbf{N}$. They are both equipped with the projective topology induced by the semi-norms

$$p_k(\phi) = \sum_{j \leq k} \|\phi^{(j)}\|_\infty, \quad \forall k \in \mathbf{N}.$$

Let $\mathcal{D}'_+$, resp. $\mathcal{D}'_-$, be their strong topological dual. It is straightforward that $\mathcal{D}_+$ is stable by $I^\beta_{0^+}$ and $\mathcal{D}_-$ is stable $I^\beta_{1^-}$, for any $\beta \in \mathbf{R}^+$. Hence, guided by (2.2), we can define the fractional integral of any distribution (i.e., an element of $\mathcal{D}'_-$ or $\mathcal{D}'_+$):

$$\text{For } T \in \mathcal{D}'_-;\ I^\beta_{0^+} T :\ \phi \in \mathcal{D}_- \mapsto < T, I^\beta_{1^-} \phi >_{\mathcal{D}'_-, \mathcal{D}_-},$$

$$\text{For } T \in \mathcal{D}'_+;\ I^\beta_{1^-} T :\ \phi \in \mathcal{D}_+ \mapsto < T, I^\beta_{0^+} \phi >_{\mathcal{D}'_+, \mathcal{D}_+}.$$

We introduce now our Besov-Liouville spaces of negative order as follows.

Definition 2.2. For $\beta > 0$ and $r > 1$, $\mathcal{I}^+_{-\beta,r}$ (resp. $\mathcal{I}^-_{-\beta,r}$) is the space of distributions $T \in \mathcal{D}'_-$ (resp. $T \in \mathcal{D}'_+$) such that $I^\beta_{0^+} T$ (resp. $I^\beta_{1^-} T$) belongs to $\mathcal{L}^r$. The norm of an element T in this space is the norm of $I^\beta_{0^+} T$ in $\mathcal{L}^r$ (resp. of $I^\beta_{1^-} T$).

Theorem 2.3. *For $\beta > 0$ and $r > 1$, the dual space of $\mathcal{I}^+_{\beta,r}$ (resp. $\mathcal{I}^-_{\beta,r}$) is canonically isometrically isomorphic to $I^{-\beta}_{1^-}(\mathcal{L}^{r^*})$ (resp. $I^{-\beta}_{0^+}(\mathcal{L}^{r^*})$,) where $r^* = r(r-1)^{-1}$. Moreover, for $\beta \geq \alpha \geq 0$ and $r > 1$, $I^{\beta}_{1^-}$ is continuous from $\mathcal{I}^-_{-\alpha,r}$ into $\mathcal{I}^-_{\beta-\alpha,r}$.*

The first part of the next theorem is a deep result which can be found in [19]. We complement it by the computation of the Hilbert-Schmidt norm of the canonical embedding κ_α from $\mathcal{I}^+_{\alpha,2}$ into $\mathcal{L}^2$.

Theorem 2.4. *The canonical embedding κ_α from $\mathcal{I}^-_{\alpha,2}$ into $\mathcal{L}^2$ is Hilbert-Schmidt if and only if $\alpha > 1/2$. Moreover,*

$$c_\alpha := \|\kappa_\alpha\|_{HS} = \|I^\alpha_{0^+}\|_{HS} = \|I^\alpha_{1^-}\|_{HS} = \frac{1}{2\Gamma(\alpha)}\left(\frac{1}{\alpha(\alpha-1/2)}\right)^{1/2}. \tag{2.4}$$

Proof. Let $(e_n,\, n \geq 1)$ be a CONB of $\mathcal{L}^2$ then $(h^\alpha_n = I^\alpha_{1^-}(e_n),\ n \in \mathbf{N})$ is a CONB of $\mathcal{I}^-_{\alpha,2}$ and

$$\begin{aligned}\|\kappa_\alpha\|^2_{\mathrm{HS}} = \sum_{n\geq 1} \|h^\beta_n\|^2_{\mathcal{L}^2} = \sum_n \|I^\alpha_{1^-}(e_n)\|^2_{\mathcal{L}^2} = \|I^\alpha_{1^-}\|^2_{\mathrm{HS}} \\ = \frac{1}{\Gamma(\alpha)^2}\iint_{[0,\,1]^2}(t-s)^{2\alpha-2}\,\mathrm{d}s\,\mathrm{d}t,\end{aligned}$$

and the result follows by straightforward quadrature. The same reasoning shows also that $c_\alpha = \|I^\alpha_{0^+}\|_{HS}$. □

For any $\tau \in [0,\,1]$, let ϵ_τ the Dirac measure at point τ. In view of Theorem [2.1], assertion i, ϵ_τ belongs to $(\mathcal{I}^-_{\alpha,\,2})'$ for any $\alpha > 1/2$. As will be apparent below, we need to estimate the norm ϵ_τ in this space.

Lemma 2.5. *For any $\alpha > 1/2$, for any $\tau \in [0,\,1]$, the image of ϵ_τ by j_α, the canonical isometry between $(\mathcal{I}^-_{\alpha,\,2})'$ and $\mathcal{I}^-_{\alpha,\,2}$, is the function*

$$j_\alpha(\epsilon_\tau)\,:\, s \longmapsto I^\alpha_{1^-}\Big((.-\tau)^{\alpha-1}_+\Big)(s)$$

and

$$\|j_\alpha(\epsilon_\tau)\|^2_{\mathcal{I}^-_{\alpha,\,2}} = \sum_{k\in\mathbf{N}} |h^\alpha_k(\tau)|^2 = \frac{(1-\tau)^{2\alpha-1}}{(2\alpha-1)\Gamma(\alpha)^2}. \tag{2.5}$$

Proof. By definition of the dual product, for any $h = I^\alpha_{1^-}(\dot h)$ where $\dot h \in \mathcal{L}^2$,

$$\begin{aligned}\langle \epsilon_\tau,\, h\rangle_{(\mathcal{I}^-_{\alpha,\,2})',\mathcal{I}^-_{\alpha,\,2}} = h(\tau) = \frac{1}{\Gamma(\alpha)}\int_\tau^1 (s-\tau)^{\alpha-1}\dot h(s)\,\mathrm{d}s \\ = \Big(\frac{1}{\Gamma(\alpha)}(.-\tau)^{\alpha-1}_+,\ \dot h\Big)_{\mathcal{L}^2} = \langle I^\alpha_{1^-}\Big((.-\tau)^{\alpha-1}_+\Big),\ h\rangle_{\mathcal{I}^-_{\alpha,\,2},\mathcal{I}^-_{\alpha,\,2}},\end{aligned}$$

hence the first assertion. Moreover, according to Parseval identity in $\mathcal{L}^2$, we have

$$\sum_{k\in\mathbf{N}} |h_k^\alpha(\tau)|^2 = \sum_{k\in\mathbf{N}} \langle \epsilon_\tau,\, h_k^\alpha\rangle^2_{(\mathcal{I}^-_{\alpha,\,2})',\,\mathcal{I}^-_{\alpha,\,2}} = \sum_{k\in\mathbf{N}} \langle j_\alpha(\epsilon_\tau),\, h_k^\alpha\rangle^2_{\mathcal{I}^-_{\alpha,\,2},\,\mathcal{I}^-_{\alpha,\,2}}$$
$$= \frac{1}{\Gamma(\alpha)^2}\sum_{k\in\mathbf{N}} \Big((.-\tau)_+^{\alpha-1},\ e_k\Big)^2_{\mathcal{L}^2} = \frac{1}{\Gamma(\alpha)^2}\|(.-\tau)_+^{\alpha-1}\|^2_{\mathcal{L}^2} = \|j_\alpha(\epsilon_\tau)\|^2_{\mathcal{I}^-_{\alpha,\,2}}.$$

Then (2.5) follows by quadrature. □

3. Gaussian Structures on Hilbert Spaces

In order to compare quantitatively the distribution of a piecewise differentiable process with that of a Brownian motion, we need to consider a functional space to which the sample-paths of both processes belong to. Ordinary Brownian motion is known to have sample-paths Hlder continuous of any order smaller than 1/2. Thus Theorem [2.1] ensures that its sample-paths belongs to $\mathcal{I}_{\beta,\infty} \subset \mathcal{I}_{\beta,2}$ for any $\beta < 1/2$. Moreover, a simple calculation shows that for any $\alpha \in (0,1)$,

$$\mathbf{1}_{[a,+\infty)} = \Gamma(\alpha)\, I^\alpha_{0^+}((.-a)_+^{-\alpha}).$$

Hence $\mathbf{1}_{[a,+\infty)}$ belongs to $\mathcal{I}_{1/2-\epsilon,\,2}$ for any $\epsilon > 0$. This implies that random step functions belong to $\mathcal{I}_{\beta,2}$ for any $\beta < 1/2$. The space of choice may thus be any space $\mathcal{I}_{\beta,\,2}$ for any $\beta < 1/2$. The closer to $1/2$ β is, the most significant the distance is but the the greater the error bound is.

3.1. Gaussian structure on Besov-Liouville spaces. We start from the Itô-Nisio theorem to construct the Wiener measure on $\mathcal{I}_{\beta,\,2}$. Let $(X_n,\, n \geq 1)$ be a sequence of independent centered Gaussian random variables of unit variance defined on a common probability space $(\Omega,\, \mathcal{A},\, \mathrm{P})$. Let $(e_n,\, n \geq 1)$ be a complete orthonormal basis of $\mathcal{L}^2([0,\,1])$. Then

$$B(t) := \sum_{n\geq 1} X_n I^1_{0^+}(e_n)(t)$$

converges almost-surely for any $t \in [0,\,1]$. From [10], we already know that the convergence holds uniformly with respect to t and thus that B is continuous. To prove that the convergence holds in $L^2(\Omega;\, \mathcal{I}_{\beta,\,2})$, it suffices to show that

$$\sum_{n\geq 1} \|I^1_{0^+} e_n\|^2_{\mathcal{I}_{\beta,\,2}} = \sum_{n\geq 1} \|I^{1-\beta}_{0^+} e_n\|^2_{\mathcal{L}^2} = \|I^{1-\beta}_{0^+}\|^2_{\mathrm{HS}} < \infty. \tag{3.1}$$

From Theorem [2.4], we know that $I^{1-\beta}$ is an Hilbert-Schmidt operator from $\mathcal{L}^2$ into itself if and only if $1-\beta > 1/2$, i.e. $\beta < 1/2$. Thus, for $\beta < 1/2$, the distribution of B defines the Wiener measure on $\mathcal{I}_{\beta,\,2}$. We denote this measure by μ_β. Note that (3.1) implies that the embedding from $\mathcal{I}_{1-\beta,\,2}$ into $\mathcal{L}^2$ is also Hilbert-Schmidt and that its Hilbert-Schmidt norm is $\|I^{1-\beta}_{0^+}\|_{\mathrm{HS}}$. By the very definition of the scalar product on $\mathcal{I}_{\beta,\,2}$, for $\eta \in \mathcal{I}_{\beta,\,2}$, we have

$$\begin{aligned}\mathbb{E}_{\mu_\beta}\left[\exp(i\langle \eta,\, \omega\rangle_{\mathcal{I}_{\beta,\,2}})\right] &= \mathbb{E}_{\mathrm{P}}\left[\exp(i\sum_{n\ge 1}\int_0^1 (I^{1-\beta}_{1^-}\circ I^{-\beta}_{0^+})\eta(s)\ e_n(s)\,\mathrm{d}s\ X_n)\right]\\ &= \exp(-\frac{1}{2}\sum_{n\ge 1}\left(\int_0^1 (I^{1-\beta}_{1^-}\circ I^{-\beta}_{0^+})\eta(s)\ e_n(s)\,\mathrm{d}s\right)^2)\\ &= \exp(-\frac{1}{2}\|(I^{1-\beta}_{1^-}\circ I^{-\beta}_{0^+})\eta\|^2_{\mathcal{L}^2})\\ &= \exp(-\frac{1}{2}\int_0^1 (I^{1-\beta}_{0^+}\circ I^{1-\beta}_{1^-})\dot\eta(s)\ \dot\eta(s)\,\mathrm{d}s),\end{aligned}$$

where $\dot\eta$ is the unique element of $\mathcal{L}^2$ such that $\eta = I^{\beta}_{0^+}\dot\eta$. Thus, μ_β is a Gaussian measure on $\mathcal{I}_{\beta,\,2}$ of covariance operator given by

$$V_\beta = I^{\beta}_{0^+}\circ I^{1-\beta}_{0^+}\circ I^{1-\beta}_{1^-}\circ I^{-\beta}_{0^+}.$$

This means that

$$\mathbb{E}_{\mu_\beta}\left[\exp(i\langle \eta,\, \omega\rangle_{\mathcal{I}_{\beta,\,2}})\right] = \exp(-\frac{1}{2}\langle V_\beta\eta,\, \eta\rangle_{\mathcal{I}_{\beta,\,2}}).$$

We could thus in principle make all the computations in $\mathcal{I}_{\beta,\,2}$. It turns out that we were not able to be explicit in the computations of some traces of some involved operators the expressions of which turned to be rather straightforward in $l^2(\mathbf{N})$ (where $\mathbf{N}$ is the set of positive integers). This is why we transfer all the structure to $l^2(\mathbf{N})$. This is done at no loss of generality nor precision since there exists a bijective isometry between $\mathcal{I}_{\beta,\,2}$ and $l^2(\mathbf{N})$.

3.2. Gaussian structure on $l^2(\mathbf{N})$. Actually, the canonical isometry is given by the Fourier expansion of the β-th *derivative* of an element of $\mathcal{I}_{\beta,\,2}$. As is, that would not be explicit enough for the computations to come to be tractable. We take benefit from the dual aspect of a time indexed point process. On the one hand, as mentioned above, the sample-path of a point process is of the form

$$t \mapsto \sum_{n\ge 1}\mathbf{1}_{[t_n,\,1]}(t),$$

where $(t_n,\, n\ge 1)$ is a strictly increasing sequence of reals, all but a finite number greater than 1, and thus belongs to $\mathcal{I}_{\beta,\,2}$ for any $\beta < 1/2$ as shown above. On the other hand, it can be seen as a locally finite point measure defined by

$$f \in \mathcal{L}^2([0,1]) \mapsto \sum_{n\ge 1} f(t_n).$$

Said otherwise, we have the following identities. For $(h,\omega)\in \mathcal{I}^-_{1-\beta,2}\times\mathcal{I}^+_{\beta,2}$

$$\begin{aligned}\int_0^1 h(s)d\omega_s &:= \sum_{n\ge 1} h(t_n)\mathbf{1}_{[0,\,1]}(t_n)\\ &=< h, I^{-1}_{0^+}(\omega) >_{\mathcal{I}^-_{1-\beta,2},\ I^+_{\beta-1,2}} = (I^{\beta-1}_{1^-}(h), I^{-\beta}_{0^+}(\omega))_{\mathcal{L}^2}.\end{aligned} \tag{3.2}$$

Recall that $(e_n,\, n \in \mathbf{N})$ is a complete orthonormal basis of $\mathcal{L}^2$ and set $h_n^{1-\beta} = I_{1^-}^{1-\beta}(e_n)$. Then $(h_n^{1-\beta},\, n \in \mathbf{N})$ is a complete orthonormal basis of $\mathcal{I}^-_{1-\beta,2}$. Consider the map $\mathfrak{J}_\beta$ defined by:

$$\begin{aligned}\mathfrak{J}_\beta \,:\, \mathcal{I}^+_{\beta,2} &\longrightarrow l^2(\mathbf{N})\\ \omega &\longmapsto \sum_{n\in\mathbf{N}} \int_0^1 h_n^{1-\beta}(s)\,\mathrm{d}\omega(s)\; x_n,\end{aligned}$$

where $(x_n,\, n \in \mathbf{N})$ is the canonical orthonormal basis of $l^2(\mathbf{N})$.

Theorem 3.1. *The map $\mathfrak{J}_\beta$ is a bijective isometry from $\mathcal{I}_{\beta,2}$ into $l^2(\mathbf{N})$. Its inverse is given by:*

$$\begin{aligned}\mathfrak{J}_\beta^{-1} \,:\, l^2(\mathbf{N}) &\longrightarrow \mathcal{I}^+_{\beta,\,2}\\ \sum_{n\in\mathbf{N}} \alpha_n\, x_n &\longmapsto \sum_{n\in\mathbf{N}} \alpha_n I_{0^+}^{\beta}(e_n).\end{aligned}$$

Proof. In view of 3.2, we have

$$\begin{aligned}\|\mathfrak{J}_\beta\omega\|^2_{l^2(\mathbf{N})} &= \sum_{n\in\mathbf{N}} \left(\int_0^1 h_n^{1-\beta}(s)\,\mathrm{d}\omega(s)\right)^2\\ &= \sum_{n\in\mathbf{N}} (e_n, I_{0^+}^{-\beta}\omega)^2_{\mathcal{L}^2}\\ &= \|I_{0^+}^{-\beta}\omega\|^2_{\mathcal{L}^2},\end{aligned}$$

according to Parseval equality. Thus by the definition of the norm on $\mathcal{I}_{\beta,2}$, $\mathfrak{J}_\beta$ is an isometry. Since

$$\int_0^1 h_n^{1-\beta}(s)\,\mathrm{d}\omega(s) = (e_n,\, I_{0^+}^{-\beta}(\omega))_{\mathcal{L}^2},$$

the inverse of $\mathfrak{J}_\beta$ is clearly given by

$$\begin{aligned}\mathfrak{J}_\beta^{-1} \,:\, l^2(\mathbf{N}) &\longrightarrow \mathcal{I}^+_{\beta,2}\\ \sum_{n\geq 1} \alpha_n\, x_n &\longmapsto \sum_{n\geq 0} \alpha_n I_{0^+}^{\beta}(e_n).\end{aligned}$$

The proof is thus complete. □

We thus have the commutative diagram.

$$\begin{array}{ccc}\mathcal{I}_{\beta,\,2} & \xrightarrow{\ \mathfrak{J}_\beta\ } & l^2(\mathbf{N})\\ {\scriptstyle V_\beta}\big\downarrow & & \big\downarrow{\scriptstyle S_\beta:=\mathfrak{J}_\beta\circ V_\beta\circ\mathfrak{J}_\beta^{-1}}\\ \mathcal{I}_{\beta,\,2} & \xrightarrow{\ \mathfrak{J}_\beta\ } & l^2(\mathbf{N})\end{array}$$

According to the properties of Gaussian measure (see [11]), we have the following result.

Theorem 3.2. *Let μ_β denote the Wiener measure on $\mathcal{I}_{\beta,2}$. Denote $m_\beta = \mathfrak{J}_\beta^* \mu_\beta$, then m_β is the Gaussian measure on $l^2(\mathbf{N})$ such that for any $v \in l^2(\mathbf{N})$,*

$$\int_{l^2(\mathbf{N})} \exp(i\, v.u)\, dm_\beta(u) = \exp(-\frac{1}{2} S_\beta v.v)$$

with the following notation

$$\|x\|^2_{l^2(\mathbf{N})} = \sum_{n=1}^\infty |x_n|^2 \text{ and } x.y = \sum_{n=1}^\infty x_n y_n, \text{ for all } x,\, y \in l^2(\mathbf{N}).$$

For the sake of simplicity, we also denote by a dot the scalar product in $l^2(\mathbf{N})^{\otimes k}$ for any integer k.

In view of Theorem [3.1], it is straightforward that the map S_β admits the representation:

$$S_\beta = \sum_{n \geq 1} \sum_{k \geq 1} \langle h_n^{1-\beta},\, h_k^{1-\beta} \rangle_{\mathcal{L}^2}\; x_n \otimes x_k.$$

By $\mathcal{C}_b^k(l^2(\mathbf{N});\, X)$, we denote the space of k-times Frchet differentiable functions from $l^2(\mathbf{N})$ into an Hilbert space X with bounded derivatives: A function F belongs to $\mathcal{C}_b^k(l^2(\mathbf{N});\, X)$ whenever

$$\|F\|_{\mathcal{C}_b^k(l^2(\mathbf{N});\, X)} := \sup_{j=1,\cdots,k} \sup_{x \in l^2(\mathbf{N})} \|\nabla^{(j)} F(x)\|_{X \otimes l^2(\mathbf{N})^{\otimes j}} < \infty.$$

Definition 3.3. The *Ornstein-Uhlenbeck semi-group* on $(l^2(\mathbf{N}),\, m_\beta)$ is defined for any $F \in L^2(l^2(\mathbf{N}),\, m_\beta;\, X)$ by

$$P_t^\beta F(u) := \int_{l^2(\mathbf{N})} F(e^{-t} u + \sqrt{1 - e^{-2t}}\, v)\, \mathrm{d}m_\beta(v),$$

where the integral is a Bochner integral.

The following properties are well known.

Lemma 3.4. *The semi-group P^β is ergodic in the sense that for any $u \in l^2(\mathbf{N})$,*

$$P_t^\beta F(u) \xrightarrow{t \to \infty} \int f\, dm_\beta.$$

Moreover, if F belongs to $\mathcal{C}_b^k(l^2(\mathbf{N});\, X)$, then $\nabla^{(k)}(P_t^\beta F) = \exp(-kt) P_t^\beta (\nabla^{(k)} F)$ so that we have

$$\int_{l^2(\mathbf{N})} \int_0^\infty \sup_{u \in l^2(\mathbf{N})} \|\nabla^{(k)}(P_t^\beta F)(u)\|_{l^2(\mathbf{N})^{\otimes(k)} \otimes X}\, dt \;\; dm_\beta(u) \leq \frac{1}{k} \|F\|_{\mathcal{C}_b^k(l^2(\mathbf{N});\, X)}.$$

For Hilbert valued functions, we define A^β as follows.

Definition 3.5. Let A^β denote the linear operator defined for $F \in \mathcal{C}_b^2(l^2(\mathbf{N});\, X)$ by:

$$(A^\beta F)(u) = u.(\nabla F)(u) - \operatorname{trace}_{l^2(\mathbf{N})}(S_\beta \circ \nabla^{(2)} F(u)), \text{ for all } u \in l^2(\mathbf{N}).$$

We still denote by A^β the unique extension of A^β to its maximal domain.

Theorem 3.6. *The map A^β is the infinitesimal generator of P^β in the sense that for $F \in \mathcal{C}_b^2(l^2(\mathbf{N});\, X)$: for any $u \in l^2(\mathbf{N})$,*

$$P_t^\beta F(u) = F(u) - \int_0^t A^\beta P_s^\beta F(u)\, ds. \tag{3.3}$$

Proof. By its very definition,

$$A^\beta F(u) = \frac{d}{dt} P_t^\beta F(u)\Big|_{t=0}.$$

If $F \in \mathcal{C}_b^2(l^2(\mathbf{N});\, X)$, it is clear that

$$\begin{aligned}\frac{d}{dt} P_t F(u) = -e^{-t} \int_{l^2(\mathbf{N})} & u.\nabla F(e^{-t}u + \sqrt{1-e^{-2t}}\, v)\, \mathrm{d}m_\beta(v) \\ &+ \frac{e^{-2t}}{\sqrt{1-e^{-2t}}} \int_{l^2(\mathbf{N})} v.\nabla F(e^{-t}u + \sqrt{1-e^{-2t}}\, v)\, \mathrm{d}m_\beta(v).\end{aligned} \tag{3.4}$$

The rest of the proof boils down to show that

$$\begin{aligned}&\int_{l^2(\mathbf{N})} v.\nabla F(e^{-t}u + \sqrt{1-e^{-2t}}\, v)\, \mathrm{d}m_\beta(v) \\ &= \int_{l^2(\mathbf{N})} \mathrm{trace}_{l^2(\mathbf{N})}(S_\beta \circ \nabla^{(2)} F(u))\, \mathrm{d}m_\beta(v).\end{aligned} \tag{3.5}$$

Taking that for granted, the result follows by setting $t = 0$ in (3.4). Now, for ν_Γ the centered Gaussian measure on $\mathbf{R}^n$ of covariance matrix Γ, it is tedious but straightforward to show that

$$\int_{\mathbf{R}^n} \nabla F(y).y\, \mathrm{d}\nu_\Gamma(y) = \int_{\mathbf{R}^n} \mathrm{trace}(\Gamma \circ \nabla^{(2)} F(y))\, \mathrm{d}\nu_\Gamma(y). \tag{3.6}$$

Let $(g_n^\beta,\, n \in \mathbf{N})$ be CONB of $\mathcal{I}_{\beta,\,2}$ which reduces S_β, i.e.

$$S_\beta = \sum_{n \in \mathbf{N}} \lambda_n(S_\beta)\, g_n^\beta \otimes g_n^\beta,$$

where $(\lambda_n(S_\beta),\, n \in \mathbf{N})$ is the set of eigenvalues of S_β. Let π_N the orthogonal projection in $\mathcal{I}_{\beta,2}$, on $\mathrm{span}\{g_n^\beta,\, n \le N\}$, $u_N = \pi_N u$ and $u_n^\perp = u - u_N$. Denote by $\nu_n = \pi_N^* m_\beta$ and $\mu_n^\perp = (\mathrm{Id} - \pi_N)^* m_\beta$. By the properties of Gaussian measures,

$$\begin{aligned}\int_{l^2(\mathbf{N})} v.\nabla F(v)\, \mathrm{d}m_\beta(v) &= \int_{l^2(\mathbf{N})} (v_N + v_N^\perp).\nabla F(v_N + v_N^\perp)\, \mathrm{d}\nu_n(v_N)\, \mathrm{d}\nu_N^\perp(v_N^\perp) \\ &= A_1^N + A_2^N.\end{aligned}$$

Since $F \in \mathcal{C}_b^2(l^2(\mathbf{N});\, X)$,

$$|A_2^N| \le \|\nabla F\|_\infty \left(\int_{l^2(\mathbf{N})} |v_N^\perp|^2\, \mathrm{d}\nu_N^\perp(v_N^\perp) \right)^{1/2}.$$

Since $\nu_N^\perp$ is a Gaussian measure on $\mathcal{I}_{\beta,\,2}$ whose covariance kernel is $\pi_N^\perp S_\beta \pi_N^\perp$, we have

$$\int_{l^2(\mathbf{N})} |v_N^\perp|^2\, \mathrm{d}\nu_N^\perp(v_N^\perp) = \mathrm{trace}(\pi_N^\perp S_\beta \pi_N^\perp).$$

Since $\pi_N^\perp$ tends to the null operator as N goes to infinity, A_2^N tends to 0. Moreover, $\Gamma_N = \pi_N S_\beta \pi_N$ tends in trace norm to S_β, hence for any $u \in l^2(\mathbf{N})$,

$$\operatorname{trace}(\tilde{\Gamma}_N \circ \nabla^{(2)} F(u)) \xrightarrow{N\to\infty} \operatorname{trace}(S_\beta \circ \nabla^{(2)} F(u)),$$

where $\tilde{\Gamma}_N(u_N + u_N^\perp) = \Gamma_N(u_N)$ for any $u = u_N + u_N^\perp$ in $l^2(\mathbf{N})$. According to (3.6),

$$\int_{\mathbf{R}^N} \nabla F(u_N + u_N^\perp).u_N \, \mathrm{d}\nu_N(u_N) = \int_{\mathbf{R}^N} \operatorname{trace}(\tilde{\Gamma}_N \circ \nabla^{(2)} F(u_N + u_N^\perp)) \, \mathrm{d}\nu_N(u_N).$$

Hence

$$A_1^N = \int_{l^2(\mathbf{N})} \operatorname{trace}(\tilde{\Gamma}_N \circ \nabla^{(2)} F(u)) \, \mathrm{d}\nu(u),$$

and by dominated convergence, we get (3.5). □

3.3. Notation. Before going further, we summarize the notation

- $x.y$: canonical scalar product on $l^2(\mathbf{N})^{\otimes(k)}$
- $\langle f, g\rangle_{\mathcal{I}_{\alpha,2}}$: canonical scalar product on $\mathcal{I}_{\alpha,2}$
- ∇F : gradient of a Fréchet differentiable F defined on $l^2(\mathbf{N})$
- μ_β (respectively m_β) : Gaussian measure on $\mathcal{I}_{\beta,2}$ (resp. $l^2(\mathbf{N})$)
- $h_n^\alpha = I_{1^-}^\alpha(e_n)$ where $(e_n,\, n \in \mathbf{N})$ is a CONB of $\mathcal{L}^2$
- $(x_n,\, n \in \mathbf{N})$ the canonical basis of $l^2(\mathbf{N})$
- c_α : Hilbert-Schmidt norm of $I_{1^-}^\alpha$

4. Stein Method

For μ and ν two probability measures on $\mathbf{R}^{\mathbf{N}}$ equipped with its Borel σ-field, we define a distance by

$$\rho_{\mathfrak{T}}(\nu,\, \mu) = \sup_{\|F\|_{\mathfrak{T}} \le 1} \int F \, \mathrm{d}\nu - \int F \, \mathrm{d}\mu.$$

where $\mathfrak{T}$ is a normed space of test functions (the norm of which is denoted by $\|.\|_{\mathfrak{T}}$). If $\mathfrak{T}$ is the set 1-Lipschitz functions on $l^2(\mathbf{N})$, then $\rho_{\mathfrak{T}}$ corresponds to the optimal transportation problem for the cost function $c(x,\, y) = \|x - y\|_{l^2(\mathbf{N})}$, $x,\, y \in l^2(\mathbf{N})$ (see [22]). For technical reasons (as in [16]) mainly due to the infinite dimension, we must restrict the space $\mathfrak{T}$ to smaller subsets. We thus introduce the distances ρ_j for $j \ge 1$ as

$$\rho_j(\nu,\, \mu) = \sup_{\|F\|_{\mathcal{C}_b^j(l^2(\mathbf{N});\, \mathbf{R})} \le 1} \int F \, \mathrm{d}\nu - \int F \, \mathrm{d}\mu.$$

However, these weaker distances still metrize the space of weak convergence of probability measures on $l^2(\mathbf{N})$.

Theorem 4.1. *Let $(\nu_n,\, n \ge 1)$ be a sequence of probability measures on $l^2(\mathbf{N})$ such that some $j \ge 1$,*

$$\rho_j(\nu_n,\, \mu) \xrightarrow{n\to\infty} 0.$$

Then $(\nu_n,\, n \ge 1)$ converges weakly to μ in $l^2(\mathbf{N})$:

$$\int F \, d\nu_n \xrightarrow{n\to\infty} \int F \, d\mu,$$

for any F bounded and continuous from $l^2(\mathbf{N})$ into $\mathbf{R}$.

Proof. As Hilbert spaces admit arbitrarily smooth partition of unity [12], for $j \geq 2$, one can mimic the proof of [8, page 396] (see also [2]) which corresponds to ρ_1. □

Say that $\mu = m_\beta$ is our reference measure, that is the measure we want the other measures to be compared to. Stein method relies on the characterization of m_β as the stationary measure of the ergodic semi-group P^β. In view of (3.3), for $j \geq 2$,

$$\rho_j(\nu, m_\beta) = \sup_{\|F\|_{\mathcal{C}_b^j(l^2(\mathbf{N});\, \mathbf{R})} \leq 1} \int_{l^2(\mathbf{N})} \int_0^\infty A^\beta P_t^\beta F(x) \,\mathrm{d}t \,\mathrm{d}\nu(x).$$

Thanks to the integration by parts induced by Malliavin calculus, we can control the right-hand-side integrand and obtain bounds on $\rho_j(\nu, m_\beta)$. To be more illustrative, the Stein method works as follows: construct a process $(t \mapsto \mathfrak{X}(x, t))$ constant in distribution if its initial condition x is distributed according to m_β. Moreover, for any initial distribution, the law of $\mathfrak{X}(x,t)$ tends to m_β as t goes to infinity. Stein method then consists in going back in time, from infinity to 0, controlling along the way the derivative of the changes, yielding a bound on the distance between the two initial measures. Other versions (coupling, size-bias, etc) are just other ways to construct another process $\mathfrak{X}$. In these approaches, for every ν, the couplings are ad-hoc whereas Malliavin calculus gives a certain kind of universality as it depends only on the underlying alea. Malliavin structures are well established for sequences of Bernoulli random variables, Poisson processes, Gaussian processes and several other spaces (see [17]). In what follows, we show an example of the machinery for each of these three examples.

The core of the method can be summarized in the following theorem.

Hypothesis I. For X a Hilbert space, $H \in l^2(\mathbf{N}) \otimes X$ and α a non-negative real, we say that the probability measure ν satisfies $\operatorname{Hyp}(X, H, \alpha)$ whenever for any $G \in \mathcal{C}_b^2(l^2(\mathbf{N});\, l^2(\mathbf{N}))$

$$\Big| \int_{l^2(\mathbf{N})} x.G(x) \,\mathrm{d}\nu(x) - \int_{l^2(\mathbf{N})} \operatorname{trace}(\operatorname{trace}_X(H \otimes H) \circ \nabla G(x)) \,\mathrm{d}\nu(x) \Big| \\ \leq \alpha \, \|\nabla^{(2)} G\|_\infty. \quad (4.1)$$

Theorem 4.2 (Stein method)**.** *Assume that* $\operatorname{Hyp}(X, H, \alpha)$ *holds. Then, if* $\alpha > 0$*,*

$$\rho_3(\nu, m_\beta) \leq \frac{1}{2} \, \| \operatorname{trace}_X(H \otimes H) - S_\beta \|_{\mathcal{S}_1} + \frac{\alpha}{3}. \quad (4.2)$$

If $\alpha = 0$*,*

$$\rho_2(\nu, m_\beta) \leq \frac{1}{2} \| \operatorname{trace}_X(H \otimes H) - S_\beta \|_{\mathcal{S}_1}. \quad (4.3)$$

Remark 4.3. The two terms in the right-hand-side of (4.2) are of totally different nature. The trace term really measures the effect of the approximation scheme whereas the second term comes from a sort of curvature of the space on which is built the approximate process. As will become evident in the examples below, this term is zero when the Malliavin gradient satisfies the chain rule formula and non-zero otherwise.

Proof. For $\alpha > 0$, for $F \in \mathcal{C}_b^3$, according to Lemma [3.4] and Theorem [3.6], we have

$$\mathbb{E}_\nu [F] - \mathbb{E}_{m_\beta} [F] \\ = -\int_{l^2(\mathbf{N})} \int_0^\infty x.\nabla P_t^\beta F(x) - \operatorname{trace}\left(S_\beta \circ \nabla^{(2)} P_t^\beta F(x)\right) \mathrm{d}t \, \mathrm{d}\nu(x). \quad (4.4)$$

Applying $\operatorname{Hyp}(X, H, \alpha)$ to $G = \nabla P_t^\beta F$, we have

$$\begin{aligned} \left|\mathbb{E}_\nu [F] - \mathbb{E}_{m_\beta} [F]\right| \le &\left|\mathbb{E}_\nu \left[\int_0^\infty \operatorname{trace}(\operatorname{trace}_X(H \otimes H) - S_\beta) \circ \nabla^{(2)} P_t^\beta F) \, \mathrm{d}t\right]\right| \\ &+ \alpha \, \mathbb{E}_\nu \left[\int_0^\infty \|\nabla^{(3)} P_t^\beta F\|_\infty \, \mathrm{d}t\right] \\ &\le \frac{1}{2} \|\nabla^{(2)} F\|_\infty \| \operatorname{trace}_X(H \otimes H) - S_\beta\|_{\mathcal{S}_1} + \frac{\alpha}{3} \|\nabla^{(3)} F\|_\infty, \end{aligned}$$

according to Lemma [3.4] and Equation (2.1). If $\alpha = 0$, the very same lines show that the second order differential of F is sufficient to have a bound of $\mathbb{E}_\nu [F] - \mathbb{E}_{m_\beta} [F]$. □

5. Normal Approximation of Poisson Processes

Let $\chi_{[0,\,1]}$ the space of locally finite measures on [0, 1] equipped with the vague topology. We identify a point measure $\omega = \sum_{n\in\mathbf{N}} \delta_{t_n}$ with the one dimensional process

$$N \,:\, t \in [0,\,1] \longmapsto \int_0^t \mathrm{d}\omega(s) = \sum_{n\in\mathbf{N}} \mathbf{1}_{[0,\,t]}(t_n).$$

The measure ν_λ is the only measure on $(\chi_{[0,\,1]}, \mathfrak{B}(\chi_{[0,\,1]}))$ such that the canonical process N is a Poisson process of intensity $\lambda \, \mathrm{d}\tau$. It is well known that for a Poisson process N of intensity λ, the process

$$N_\lambda(t) = \frac{1}{\sqrt{\lambda}} \left(N(t) - \lambda t\right)$$

converges in distribution on $\mathfrak{D}$ to a Brownian motion as λ goes to infinity. For any $\beta < 1/2$, we want to precise the rate of convergence.

5.1. Malliavin calculus for Poisson process. For a real valued functional F on $\chi_{[0,\,1]}$, it is customary to define the discrete gradient as

$$D_\tau F(N) = F(N + \epsilon_\tau) - F(N), \text{ for any } \tau \in [0,\,1],$$

where $N + \epsilon_\tau$ is the point process N with an extra atom at time τ. We denote by $\mathbb{D}_{2,1}^\lambda$ the set of square integrable functionals F such that

$$\|F\|_{2,1,\lambda}^2 := \mathbb{E}_{\nu_\lambda} \left[F^2\right] + \mathbb{E}_{\nu_\lambda} \left[\int_0^1 |D_\tau F|^2 \lambda \, \mathrm{d}\tau\right]$$

is finite. A process $G \in L^2(\nu_\lambda \times d\tau)$ is said to belong to $\operatorname{Dom}\delta^\lambda$ whenever there exists $c > 0$ such that

$$\mathbb{E}_{\nu_\lambda}\left[\int_0^1 D_\tau F\ G_\tau\,\lambda\,\mathrm{d}\tau\right] \le c\,\|F\|_{L^2(\nu_\lambda)},$$

for any $F \in \mathbb{D}^\lambda_{2,1}$. The adjoint of D, denoted by δ^λ is then defined by the following relationship:

$$\mathbb{E}_{\nu_\lambda}\left[F\ \delta^\lambda(G)\right] = \lambda\ \mathbb{E}_{\nu_\lambda}\left[\int_0^1 D_\tau F\ G_\tau\,\mathrm{d}\tau\right]. \tag{5.1}$$

Moreover, it is well known that for G deterministic, δ^λ coincides with the compensated integral with respect to the Poisson process, i.e.

$$\delta^\lambda G = \int_0^1 G_\tau(\,\mathrm{d}\omega(\tau) - \lambda\,\mathrm{d}\tau),$$

and that $D\delta^\lambda G = G$.

5.2. Convergence theorem.

Theorem 5.1. *Let* $H_\lambda = \lambda^{-1/2}\sum_{n\ge1} h_n^{1-\beta}\otimes x_n = \lambda^{-1/2}H_1$. *We denote by* ν^*_λ *the distribution of* $\mathfrak{J}_\beta N_\lambda$ *in* $l^2(\mathbf{N})$. *The measure* ν^*_λ *satisfies* $\operatorname{Hyp}(\mathcal{L}^2([0,1]),\,H_1,\,a)$ *with*

$$a = \frac{(1-\beta)^{3/2}}{5-6\beta}\,\frac{c^3_{1-\beta}}{\sqrt{\lambda}} \le \frac{c^3_{1-\beta}}{2\sqrt{\lambda}}. \tag{5.2}$$

Hence

$$\rho_3(\nu^*_\lambda,\,m_\beta) \le \frac{a}{3\sqrt{\lambda}}.$$

Remark 5.2. From its very definition, it is clear that

$$\mathfrak{J}_\beta N_\lambda = \frac{1}{\sqrt{\lambda}}\sum_{n\ge1}\delta^\lambda(h_n^{1-\beta})\ x_n$$

where $(x_n,\,n\ge1)$ is the canonical orthonormal basis of $l^2(\mathbf{N})$. Note also that

$$D\mathfrak{J}_\beta N_\lambda = \frac{1}{\sqrt{\lambda}}\sum_{n\ge1} h_n^{1-\beta}\otimes x_n.$$

It is because of this particular form of $D\mathfrak{J}_\beta N_\lambda$ as an infinite series of simple bilinear forms on $\mathcal{I}^-_{1-\beta,2}\otimes l^2(\mathbf{N})$ that the computations to come are feasible. To compare, if we view N_λ as an element of $\mathcal{I}_{\beta,\,2}$, then

$$D_\tau N_\lambda(t) = I^\beta_{0^+}((\tau - .)_+^{-\beta})(t).$$

Since there is no *decoupling* in this expression between the τ variable and the t variable, the computations are intractable; hence the need to resort to the Gaussian structure on $l^2(\mathbf{N})$.

Proof. Let $F \in \mathcal{C}_b^2(l^2(\mathbf{N});\mathbf{R})$ and $x \in l^2(\mathbf{N})$. Denoting by $G(y) = F(y)x$ for $y \in l^2(\mathbf{N})$, we have

$$\begin{aligned}\mathbb{E}\left[\mathfrak{J}_\beta N_\lambda . G(\mathfrak{J}_\beta N_\lambda)\right] &= \frac{1}{\sqrt{\lambda}} \sum_{n\geq 1} \mathbb{E}\left[\delta^\lambda(h_n^{1-\beta})F(\mathfrak{J}_\beta N_\lambda)\right]\ x_n.x \\ &= \frac{1}{\sqrt{\lambda}} \sum_{n\geq 1} \mathbb{E}\left[\int_0^1 h_n^{1-\beta}(\tau) D_\tau F(\mathfrak{J}_\beta N_\lambda)\lambda \,\mathrm{d}\tau\right]\ x_n.x \\ &= \sqrt{\lambda}\ \mathbb{E}\left[\int_0^1 D_\tau F(\mathfrak{J}_\beta N_\lambda).H_1(\tau)\,\mathrm{d}\tau\right].\end{aligned}$$

According to the Taylor formula,

$$\begin{aligned}&D_\tau F(\mathfrak{J}_\beta N_\lambda) = F(\mathfrak{J}_\beta N_\lambda + H_\lambda(\tau)) - F(\mathfrak{J}_\beta N_\lambda) \\ &= \frac{1}{\sqrt{\lambda}}\nabla F(\mathfrak{J}_\beta N_\lambda).H_1(\tau) + \frac{1}{\lambda}\int_0^1 (1-r)\ \nabla^{(2)} F(\mathfrak{J}_\beta N_\lambda + rH_\lambda(\tau)).H_1(\tau)^{\otimes(2)}\,\mathrm{d}r.\end{aligned}$$

Thus we get

$$\begin{aligned}\mathbb{E}\left[\mathfrak{J}_\beta N_\lambda . G(\mathfrak{J}_\beta N_\lambda)\right] &= \mathbb{E}\left[\int_0^1 \nabla G(\mathfrak{J}_\beta N_\lambda).H_1(\tau)^{\otimes(2)}\,\mathrm{d}\tau\right] \\ &+ \lambda^{-1/2}\mathbb{E}\left[\int_0^1\int_0^1 (1-r)\ \nabla^{(2)} G(\mathfrak{J}_\beta N_\lambda).H_1(\tau)^{\otimes(3)}\,\mathrm{d}\tau\,\mathrm{d}r\right]. \quad (5.3)\end{aligned}$$

By linearity and density, (5.3) holds for any $G \in C_b^2(l^2(\mathbf{N}); l^2(\mathbf{N}))$. Note that for any $A = \sum_{n,\,jk\in\mathbf{N}} a_{n,\,k}\, x_n \otimes x_k \in l^2(\mathbf{N}) \otimes l^2(\mathbf{N})$

$$\int_0^1 A.H_1(\tau)^{\otimes(2)}\,\mathrm{d}\tau = \sum_{n,\,k=1}^{\infty} a_{i,j}\left(h_n^{1-\beta},\, h_k^{1-\beta}\right)_{\mathcal{L}^2} = \mathrm{trace}_{l^2(\mathbf{N})}(S_\beta \circ A).$$

Hence

$$\int_0^1 \nabla G(\mathfrak{J}_\beta N_\lambda).H_1(\tau)^{\otimes(2)}\,\mathrm{d}\tau = \mathrm{trace}_{l^2(\mathbf{N})}(S_\beta \circ \nabla G(\mathfrak{J}_\beta N_\lambda)).$$

Since $\nabla^2 G$ is bounded, we have

$$\begin{aligned}&\left|\mathbb{E}\left[\int_0^1\int_0^1 (1-r)\ \nabla^{(2)} G(\mathfrak{J}_\beta N_\lambda + rH_\lambda), H_1(\tau)^{\otimes 3}\,\mathrm{d}r\,\mathrm{d}\tau\right]\right| \\ &\qquad\qquad \leq \frac{1}{2}\|\nabla^{(2)} G\|_\infty \int_0^1 \|H_1(\tau)\|_{l^2(\mathbf{N})}^3\,\mathrm{d}\tau.\end{aligned}$$

Moreover, according to (2.5),

$$\int_0^1 \|H_1(\tau)\|_{l^2(\mathbf{N})}^3\,\mathrm{d}\tau = \int_0^1 \left(\sum_{n\geq 1} h_n^{1-\beta}(\tau)^2\right)^{3/2} d\tau = \frac{(1-\beta)^{3/2}}{5/2-3\beta}c_{1-\beta}^3$$

Hence it follows that

$$\left|\mathbb{E}\left[\mathfrak{J}_\beta N_\lambda . G(\mathfrak{J}_\beta N_\lambda)\right] - \mathbb{E}\left[\mathrm{trace}(S_\beta \circ \nabla G(\mathfrak{J}_\beta N_\lambda))\right]\right| \leq \frac{(1-\beta)^{3/2}}{5/2-3\beta}\frac{c_{1-\beta}^3}{2\sqrt{\lambda}}\|\nabla^{(2)} G\|_\infty,$$

which is Equation (4.1) with $X = \mathcal{L}^2$ and $\alpha = a$ given by (5.2). □

Remark 5.3. It is remarkable that by homogeneity, the partial trace of $H_1 \otimes H_1$ is equal to S_β. The only remaining term in Theorem [4.2] comes from the fact that the discrete gradient does not satisfy the chain rule.

One could also remark that the choice of the space in which we embed the Poisson and Brownian sample-paths (i.e. the choice of the value of β) modifies only the constant but not the order of convergence, which remains proportional to $\lambda^{-1/2}$.

6. Linear Interpolation of the Brownian Motion

For $m \geq 1$, the linear interpolation $B^\dagger_m$ of a Brownian motion $B^\dagger$ is defined by

$$B^\dagger_m(0) = 0 \text{ and } \mathrm{d}B^\dagger_m(t) = m \sum_{j=0}^{m-1} (B^\dagger(j+1/m) - B^\dagger(j/m)) \mathbf{1}_{[j/m,\,(j+1)/m)}(t)\,\mathrm{d}t.$$

Thus $\mathfrak{J}_\beta B^\dagger_m$ is given by

$$\mathfrak{J}_\beta B^\dagger_m = m \sum_{j=0}^{m-1} (B^\dagger(j+1/m) - B^\dagger(j/m)) \sum_{n \in \mathbf{N}} \int_{j/m}^{(j+1)/m} h_n^{1-\beta}(t)\,\mathrm{d}t\; x_n.$$

Consider the $\mathcal{L}^2$-orthonormal functions

$$e_j^m(s) = \sqrt{m}\,\mathbf{1}_{[j/m,\,(j+1)/m)}(s),\ j = 0, \cdots, m-1,\ s \in [0,\,1]$$

and $F^\dagger_m = \operatorname{span}(e_j^m,\ j = 0, \cdots, m-1)$. We denote by $p_{F^\dagger_m}$ the orthogonal projection over $F^\dagger_m$. Since $B^\dagger_m$ is constructed as a function of a standard Brownian motion, we work on the canonical Wiener space $(\mathcal{C}^0([0,\,1];\mathbf{R}),\ \mathcal{I}_{1,\,2},\ m^\dagger)$. The gradient we consider, $D^\dagger$, is the derivative of the usual gradient on the Wiener space and the integration by parts formula reads as:

$$\mathbb{E}_{m^\dagger}\left[F \int_0^1 u(s)\,\mathrm{d}B^\dagger(s)\right] = \mathbb{E}_{m^\dagger}\left[\int_0^1 D^\dagger_s F\ u(s)\,\mathrm{d}s\right] \tag{6.1}$$

for any $u \in \mathcal{L}^2$. Let

$$H^\dagger_m = \sum_{n \in \mathbf{N}} p_{F^\dagger_m} h_n^{1-\beta} \otimes x_n \in \mathcal{L}^2 \otimes l^2(\mathbf{N}).$$

It means that

$$H^\dagger_m(k,\,s) = m \sum_{n \in \mathbf{N}} \sum_{j=0}^{m-1} \left(\int_{j/m}^{(j+1)/m} h_k^{1-\beta}(t)\,\mathrm{d}t\right)\ \mathbf{1}_{[j/m,\,(j+1)/m)}(s)\ x_n.$$

Since the e_j^m's are orthogonal in $\mathcal{L}^2$, we can compute the partial trace as follows.

$$\begin{aligned}&\operatorname{trace}_{\mathcal{L}^2}(H^\dagger_m \otimes H^\dagger_m)\\ &= m \sum_{n \in \mathbf{N}} \sum_{k \in \mathbf{N}} \sum_{j=0}^{m-1} \left(\int_{j/m}^{(j+1)/m} h_k^{1-\beta}(t)\,\mathrm{d}t\right) \left(\int_{j/m}^{(j+1)/m} h_n^{1-\beta}(t)\,\mathrm{d}t\right)\ x_n \otimes x_k.\end{aligned} \tag{6.2}$$

Theorem 6.1. *Let $\nu_m^\dagger$ be the law of $\mathfrak{J}_\beta B_m^\dagger$ on $l^2(\mathbf{N})$. The measure $\nu_m^\dagger$ satisfies* $\operatorname{Hyp}(\mathcal{L}^2,\, H_m^\dagger,\, 0)$*. Hence*

$$\rho_2(\nu_m^\dagger,\, m_\beta) \le \frac{m^{2\beta-1}}{2(1-2\beta)\,\Gamma(1-\beta)^2}.$$

Proof. For G sufficiently regular, according to the definition of B^m and to (6.1), we have

$$\begin{aligned}
&\mathbb{E}\left[\mathfrak{J}_\beta B_m^\dagger . G(\mathfrak{J}_\beta B_m^\dagger)\right] \\
&= \mathbb{E}\left[\sum_{n\in\mathbf{N}} m \sum_{i=0}^{m-1} (B(i+1/m) - B(i/m)) \int_{i/m}^{(i+1)/m} h_n^{1-\beta}(t)\,\mathrm{d}t\; G_n(\mathfrak{J}_\beta B_m^\dagger)\right] \\
&= m \sum_{n\in\mathbf{N}} \sum_{i=0}^{m-1} \int_{i/m}^{(i+1)/m} h_n^{1-\beta}(t)\,\mathrm{d}t\; \mathbb{E}\left[\int_{i/m}^{(i+1)/m} D_s^\dagger G_n(\mathfrak{J}_\beta B_m^\dagger)\,\mathrm{d}s\right] \qquad (6.3)\\
&= \int_0^1 H_m^\dagger(t)\,\mathrm{d}t . \mathbb{E}_{m^\dagger}\left[\int_{i/m}^{(i+1)/m} D_s^\dagger G(\mathfrak{J}_\beta B_m^\dagger)\,\mathrm{d}s\right].
\end{aligned}$$

Since $D^\dagger$ obeys the chain rule formula,

$$\begin{aligned}
D_s^\dagger G_n(\mathfrak{J}_\beta B_m^\dagger) &= \sum_{k\in\mathbf{N}} \nabla_k G_n(\mathfrak{J}_\beta B_m^\dagger) D_s^\dagger(\mathfrak{J}_\beta B_m^\dagger) \\
&= \sum_{k\in\mathbf{N}} \nabla_k G_n(\mathfrak{J}_\beta B_m^\dagger)\left(m \sum_{l=0}^{m-1} \mathbf{1}_{[l/m,\,(l+1)/m)}(s) \int_{l/m}^{(l+1)/m} h_k(s)\,\mathrm{d}s\right) \\
&= \nabla G_n(\mathfrak{J}_\beta B_m^\dagger) . H_m^\dagger(s). \qquad (6.4)
\end{aligned}$$

Combining (6.3) and (6.4), we get

$$\begin{aligned}
&\mathbb{E}\left[\mathfrak{J}_\beta B_m^\dagger . G(\mathfrak{J}_\beta B_m^\dagger)\right] \\
&\quad = m\mathbb{E}\left[\sum_{k\in\mathbf{N}} \sum_{n\in\mathbf{N}} \sum_{i=0}^{m-1} \nabla_k G_n(\mathfrak{J}_\beta B_m^\dagger) \int_{i/m}^{(i+1)/m} h_n^{1-\beta}(t)\,\mathrm{d}t \int_{i/m}^{(i+1)/m} h_k(s)\,\mathrm{d}s\right] \\
&\quad = \mathbb{E}\left[\operatorname{trace}(\operatorname{trace}_{\mathcal{L}^2}(H_m^\dagger \otimes H_m^\dagger) \circ \nabla G(\mathfrak{J}_\beta B_m^\dagger))\right].
\end{aligned}$$

It follows that $\nu_m^\dagger$ satisfies $\operatorname{Hyp}(\mathcal{L}^2,\, H_m^\dagger,\, 0)$. To conclude, it remains to estimate $\|\operatorname{trace}_{\mathcal{L}^2}(H_m^\dagger \otimes H_m^\dagger) - S_\beta\|_{\mathcal{S}_1}$. According to Pythagorean Theorem, we have

$$\begin{aligned}
&S_\beta - \operatorname{trace}_{\mathcal{L}^2}(H_m^\dagger \otimes H_m^\dagger) \\
&\qquad = \sum_{n\in\mathbf{N}} \sum_{k\in\mathbf{N}} \left((p_{F_m^\dagger} h_n^{1-\beta},\, p_{F_m^\dagger} h_k^{1-\beta})_{\mathcal{L}^2} - (h_n^{1-\beta},\, h_k^{1-\beta})_{\mathcal{L}^2}\right) x_n \otimes x_k \\
&\qquad = \sum_{n\in\mathbf{N}} \sum_{k\in\mathbf{N}} ((\operatorname{Id} - p_{F_m^\dagger}) h_n^{1-\beta},\, (\operatorname{Id} - p_{F_m^\dagger}) h_k^{1-\beta})_{\mathcal{L}^2}\; x_n \otimes x_k.
\end{aligned}$$

Hence $S_\beta - \operatorname{trace}_{\mathcal{L}^2}(H_m^\dagger \otimes H_m^\dagger)$ is a symmetric non-negative operator. Therefore,

$$\begin{aligned}
&\| \operatorname{trace}_{\mathcal{L}^2}(H_m^\dagger \otimes H_m^\dagger) - S_\beta \|_{\mathcal{S}_1} \\
&\qquad \le \sum_{n\in\mathbf{N}} \| (\operatorname{Id} - p_{F_m^\dagger}) h_n^{1-\beta} \|_{\mathcal{L}^2}^2 \\
&\qquad = \sum_{n\in\mathbf{N}} \int_0^1 \left(h_n^{1-\beta}(s) - \sum_{j=0}^{m-1} \int_{j/m}^{(j+1)/m} h_n^{1-\beta}(t)\,\mathrm{d}t\; e_j^m(s) \right)^2 \mathrm{d}s \\
&\qquad = \sum_{n\in\mathbf{N}} \sum_{j=0}^{m-1} \int_{j/m}^{(j+1)/m} \left(h_n^{1-\beta}(s) - m \int_{j/m}^{(j+1)/m} h_n^{1-\beta}(t)\,\mathrm{d}t \right)^2 \mathrm{d}s \\
&\qquad = m \sum_{n\in\mathbf{N}} \sum_{j=0}^{m-1} \int_{j/m}^{(j+1)/m} \left(\int_{j/m}^{(j+1)/m} (h_n^{1-\beta}(s) - h_n^{1-\beta}(t)) m \,\mathrm{d}t \right)^2 \mathrm{d}s \\
&\qquad \le m^2 \sum_{n\in\mathbf{N}} \sum_{j=0}^{m-1} \int_{j/m}^{(j+1)/m} \int_{j/m}^{(j+1)/m} (h_n^{1-\beta}(s) - h_n^{1-\beta}(t))^2 \,\mathrm{d}s\,\mathrm{d}t,
\end{aligned}$$

where the last inequality follows from Jensen inequality. Since $h_n^{1-\beta} = I_{1^-}^{1-\beta}(e_n)$ with $(e_n,\, n \in \mathbf{N})$ being a CONB of $\mathcal{L}^2$, according to Parseval identity,

$$\begin{aligned}
\sum_{n\in\mathbf{N}} (h_n^{1-\beta}(s) - h_n^{1-\beta}(t))^2 &= \frac{1}{\Gamma(1-\beta)^2} \sum_{n\in\mathbf{N}} \Big((.-s)_+^{-\beta} - (.-t)_+^{-\beta},\, e_n \Big)_{\mathcal{L}^2}^2 \\
&= \frac{1}{\Gamma(1-\beta)^2} \int_0^1 \Big((\tau-s)_+^{-\beta} - (\tau-t)_+^{-\beta} \Big)^2 \mathrm{d}\tau.
\end{aligned}$$

Expanding the square and using the monotonicity of the power function, we get

$$\begin{aligned}
\sum_{n\in\mathbf{N}} (h_n^{1-\beta}(s) - h_n^{1-\beta}(t))^2 &\le \frac{1}{(1-2\beta)\Gamma(1-\beta)^2} (|1-s|^{1-2\beta} - |1-t|^{1-2\beta}) \\
&\le \frac{(1-2\beta)^{-1}}{\Gamma(1-\beta)^2} |t-s|^{1-2\beta}.
\end{aligned}$$

It follows that

$$\| \operatorname{trace}_{\mathcal{L}^2}(H_m^\dagger \otimes H_m^\dagger) - S_\beta \|_{\mathcal{S}_1} \le \frac{m^{2\beta-1}}{(1-2\beta)\Gamma(1-\beta)^2}.$$

The proof is thus complete. □

7. Donsker Theorem

The same approach can be applied to have precise asymptotic for the Donsker theorem. Let $X = (X_n,\, n \in \mathbf{N})$ be a sequence of independent and identically distributed Rademacher random variables, i.e. $\mathrm{P}(X_n = \pm 1) = 1/2$ for any n. For any k in $\mathbf{N}$, we set

$$\begin{aligned}
X_k^+ &= (X_1,\, \cdots,\, X_{k-1},\, 1,\, X_{k+1}\, \cdots), \\
X_k^- &= (X_1,\, \cdots,\, X_{k-1},\, -1,\, X_{k+1}\, \cdots).
\end{aligned}$$

The discrete gradient on this probability space is given by

$$D^\sharp_k F(X) = \frac{1}{2}(F(X_k^+) - F(X_k^-)).$$

Then the integration by parts formula reads as

$$\mathbb{E}\left[\sum_{k\in\mathbf{N}} u_k D^\sharp_k F(X)\right] = \mathbb{E}\left[F(X)\sum_{k\in\mathbf{N}} u_k X_k\right] \tag{7.1}$$

for any $u = (u_k,\, k \in \mathbf{N})$ which belongs to $l^2(\mathbf{N})$. The approximating process of the Donsker Theorem is defined by:

$$B^\sharp_m(t) = \frac{1}{\sqrt{m}}\left(\sum_{j=1}^{[mt]} X_j + (mt - [mt])X_{[mt+1]}\right).$$

Hence

$$\mathrm{d}B^\sharp_m(t) = \frac{1}{\sqrt{m}}\sum_{j=1}^{m} X_j\ \mathbf{1}_{[(j-1)/m,\, j/m)}(t)\,\mathrm{d}t.$$

Thus we get

$$\mathfrak{I}_\beta B^\sharp_m = \sum_{n\in\mathbf{N}}\sum_{j=1}^{m}\frac{1}{\sqrt{m}}\,X_j\int_{(j-1)/m}^{j/m} h_n^{1-\beta}(s)\,\mathrm{d}s\ x_j\otimes x_n.$$

Theorem 7.1. *We denote by $\nu^\sharp_m$ the distribution of $\mathfrak{I}_\beta B^\sharp_m$ on $l^2(\mathbf{N})$. The measure $\nu^\sharp_m$ satisfies* $\mathrm{Hyp}(l^2(\mathbf{N}),\, H^\sharp_m,\, 1)$ *where*

$$H^\sharp_m = \sum_{n\in\mathbf{N}}\sum_{j=1}^{m}\frac{1}{\sqrt{m}}\int_{(j-1)/m}^{j/m} h_n^{1-\beta}(s)\,ds\ x_j\otimes x_n.$$

Furthermore, for any $\epsilon > 0$, there exists m_0 such that for $m \geq m_0$,

$$\rho_3(\nu^\sharp_m,\, m_\beta) \leq (1+\epsilon)\ \frac{m^{2\beta-1}}{2(1-2\beta)\,\Gamma(1-\beta)^2}.$$

Proof. According to the integration by parts formula (7.1), we have

$$\begin{aligned}
\mathbb{E}\left[\mathfrak{I}_\beta B^\sharp_m . G(\mathfrak{I}_\beta B^\sharp_m)\right] &= \frac{1}{\sqrt{m}}\,\mathbb{E}\left[\sum_{n\geq 1}\mathfrak{I}_\beta B^\sharp_m(n)\, G_n(\mathfrak{I}_\beta B^\sharp_m)\right]\\
&= \frac{1}{\sqrt{m}}\,\mathbb{E}\left[\sum_{n\geq 1}\sum_{k=1}^{m} h_n^{1-\beta}(k/m) X_n\, G_n(\mathfrak{I}_\beta B^\sharp_m)\right]\\
&= \frac{1}{\sqrt{m}}\,\mathbb{E}\left[\sum_{k=1}^{m}\sum_{n\geq 1} h_n^{1-\beta}(k/m) D^\sharp_k G_n(\mathfrak{I}_\beta B^\sharp_m)\right]\\
&= \mathbb{E}\left[D^\sharp G(\mathfrak{I}_\beta B^\sharp_m).\ H^\sharp_m\right],
\end{aligned} \tag{7.2}$$

According to the Taylor formula,

$$D_j^\sharp G(\mathfrak{J}_\beta B_m^\sharp) = \frac{1}{2}\left(G(\mathfrak{J}_\beta B_m^\sharp + (1-X_j)H_m^\sharp(j)) - G(\mathfrak{J}_\beta B_m^\sharp - (1+X_j)H_m^\sharp(j))\right)$$
$$= \langle \nabla G(\mathfrak{J}_\beta B_m^\sharp),\, H_m^\sharp(j)\rangle_{l^2(\mathbf{N})}(1-X_j+1+X_j)/2$$
$$+ \frac{(1-X_j)^2}{2}\int_0^1 (1-r)\nabla^2 G(\mathfrak{J}_\beta B_m^\sharp + r(1-X_j)H_m^\sharp(j)).\, H_m^\sharp(j)^{\otimes(2)}\,\mathrm{d}r$$
$$+ \frac{(1+X_j)^2}{2}\int_0^1 (1-r)\nabla^2 G(\mathfrak{J}_\beta B_m^\sharp + r(1+X_j)H_m^\sharp(j)).\, H_m^\sharp(j)^{\otimes(2)}\,\mathrm{d}r.$$

Plugging this latter equation into (7.2), it follows that

$$\mathbb{E}\left[\mathfrak{J}_\beta B_m^\sharp.G(\mathfrak{J}_\beta B_m^\sharp)\right] = \mathbb{E}\left[\sum_{j=1}^m \nabla G(\mathfrak{J}_\beta B_m^\sharp).\, H_m^\sharp(j)\otimes H_m^\sharp(j)\right]$$
$$+ \sum_{z=\pm 1}\mathbb{E}\left[\sum_{j=1}^m \frac{(1-zX_j)^2}{2} \times \int_0^1 (1-r)\nabla^{(2)}G(\mathfrak{J}_\beta B_m^\sharp + r(1-zX_j)H_m^\sharp(j)).\, H_m^\sharp(j)^{\otimes 3}\,\mathrm{d}r\right].$$

Since $(1-zX_j)$ is either 0 or 2 for any $j \geq 1$ and any $z \in \pm 1$, we get

$$\left|\mathbb{E}\left[\mathfrak{J}_\beta B_m^\sharp.G(\mathfrak{J}_\beta B_m^\sharp)\right] - \mathbb{E}\left[\mathrm{trace}(\mathrm{trace}_{l^2(\mathbf{N})}(H_m^\sharp\otimes H_m^\sharp)\circ\nabla^2 G(\mathfrak{J}_\beta B_m^\sharp)))\right]\right|$$
$$\leq \|\nabla^2 G\|_\infty \sum_{j=1}^m \|H_m^\sharp(j)\|^3_{l^2(\mathbf{N})},$$

which is (4.1) with $\alpha = \sum_{j=1}^m \|H_m^\sharp(j)\|^3_{l^2(\mathbf{N})}$. It turns out that according to (6.2),

$$\mathrm{trace}_{l^2(\mathbf{N})}(H_m^\sharp\otimes H_m^\sharp)$$
$$= \sum_{n\in\mathbf{N}}\sum_{k\in\mathbf{N}}\sum_{j=1}^m (h_n^{1-\beta},\, e_{j-1}^m)_{\mathcal{L}^2}(h_n^{1-\beta},\, e_{j-1}^m)_{\mathcal{L}^2}\ x_n\otimes x_k$$
$$= \mathrm{trace}_{\mathcal{L}^2}(H_m^\dagger\otimes H_m^\dagger).$$

Hence we can use the result of Theorem [6.1]. It remains to control the following additional term (due to the fact that $D^\sharp$ does not satisfy the chain rule formula):

$$\sum_{j=1}^m \|H_m^\sharp(j)\|^3_{l^2(\mathbf{N})}.$$

By the very definition of $H^\sharp_m$,

$$\begin{aligned}\|H^\sharp_m(j)\|^2_{l^2(\mathbf{N})} &= \frac{1}{m}\sum_{n\in\mathbf{N}}(h_n^{1-\beta},\, e_j^m)^2_{\mathcal{L}^2}\\ &= \frac{1}{m}\sum_{n\in\mathbf{N}}(e_n,\, I_{0^+}^{1-\beta}(e_j^m))^2_{\mathcal{L}^2}\\ &= \frac{1}{m}\|I_{0^+}^{1-\beta}(e_j^m)\|^2_{\mathcal{L}^2}\\ &= \frac{1}{m\Gamma(1-\beta)^2}\int_0^1\left(\int_{(j-1)/m}^{j/m}(\tau-s)^{-\beta}\,\mathrm{d}s\right)^2\mathrm{d}\tau\\ &\le \frac{m^{2\beta-3}}{(1-\beta)\Gamma(1-\beta)^2}.\end{aligned}$$

Thus

$$\sum_{j=1}^m \|H^\sharp_m(j)\|^3_{l^2(\mathbf{N})} \le \frac{m^{3\beta-7/2}}{(1-\beta)\Gamma(1-\beta)^2}.$$

The dominating term is thus the term in $m^{2\beta-1}$ and the result follows. □

8. Transfer Principle

For X and Y two Hilbert spaces and Θ a continuous linear map from X to Y. Let μ and ν two probability measures on X and μ_Y (respectively ν_Y) their image measure with respect to Θ. Since Θ is linear and continuous, for $F\in F\in \mathcal{C}_b^k(Y,\,\mathbf{R})$, $F\circ\Theta$ belongs to $\mathcal{C}_b^k(X,\,\mathbf{R})$, hence, we have

$$\begin{aligned}\sup_{F\in\mathcal{C}_b^k(Y,\,\mathbf{R})}\int F\,\mathrm{d}\mu_Y-\int F\,\mathrm{d}\nu_Y &= \sup_{F\in\mathcal{C}_b^k(Y,\,\mathbf{R})}\int F\circ\Theta\,\mathrm{d}\mu-\int F\circ\Theta\,\mathrm{d}\nu\\ &\le \sup_{F\in\mathcal{C}_b^k(X,\,\mathbf{R})}\int F\,\mathrm{d}\mu-\int F\,\mathrm{d}\nu.\end{aligned}$$

As an application, we can precise the convergence established in [7]. Note that in this paper, the key tool was also a matter of Hilbert-Schmidt property of some operator.

The fractional Brownian motion of Hurst index $H\in[0,\,1]$ may be defined (see [4]) by

$$B^H(t)=\int_0^t K_H(t,s)\,\mathrm{d}B(s),$$

where

$$K_H(t,r):=\frac{(t-r)^{H-\frac12}}{\Gamma(H+\frac12)}F(\frac12-H,H-\frac12,H+\frac12,1-\frac{t}{r})1_{[0,t)}(r).$$

The Gauss hyper-geometric function $F(\alpha,\beta,\gamma,z)$ (see [13]) is the analytic continuation on $\mathbb{C}\times\mathbb{C}\times\mathbb{C}\backslash\{-1,-2,\ldots\}\times\{z\in\mathbb{C}, Arg|1-z|<\pi\}$ of the power series

$$\sum_{k=0}^{+\infty} \frac{(\alpha)_k(\beta)_k}{(\gamma)_k k!} z^k,$$

and

$$(a)_0 = 1 \text{ and } (a)_k = \frac{\Gamma(a+k)}{\Gamma(a)} = a(a+1)\dots(a+k-1).$$

Moreover, according to [18], K_H is a continuous map from $\mathcal{L}^2$ into $\mathcal{I}^+_{H+1/2,\,2}$. Hence the map $\Theta_H = K_H \circ I^{-1}_{0^+}$ can be defined continuously from $\mathcal{I}^+_{\beta,\,2}$ to $\mathcal{I}^+_{H-(1/2-\beta),\,2}$. Since $\Theta_H B = B^H$, we have the following result.

Theorem 8.1. *For any $H \in [0,\,1]$, for any $1/2 > \epsilon > 0$,*

$$\rho_3\left(\mathfrak{J}_{H-\epsilon}(\int_0^{\cdot} K_H(t,\,s)\,dN^\lambda(s)),\ \mathfrak{J}_{H-\epsilon}(B^H)\right) \le \frac{a}{3\sqrt{\lambda}}.$$

Acknowledgment. This work was motivated by discussions during the Stein's program held at Singapore. The second author would like to thank the National University of Singapore for warm hospitality and generous support.

References

1. Barbour, A. D.: *Stein's method for diffusion approximations*, Probab. Theory Related Fields **84** (1990), no. 3, 297–322.
2. Charbonneau, B., Svyrydov, Y. and Tupper, P. F.: *Weak convergence in the Prokhorov metric of methods for stochastic differential equations*, IMA J. Numer. Anal. **30** (2010), no. 2, 579–594.
3. Chen, L. H. Y., Goldstein, L., and Shao, Q.-M.: *Normal Approximation by Stein's Method*, Probability and its Applications (New York), Springer, Heidelberg, 2011.
4. Decreusefond, L.: *Stochastic calculus with respect to Volterra processes*, Annales de l'Institut Henri Poincaré (B) Probability and Statistics **41** (2005), 123–149.
5. Decreusefond, L., Joulin, A., and Savy, N.: *Upper bounds on Rubinstein distances on configuration spaces and applications*, Communications on stochastic analysis **4** (2010), no. 3, 377–399.
6. Decreusefond, L. and Moyal, P.: *Stochastic Modeling and Analysis of Telecom Networks*, ISTE Ltd and John Wiley & Sons Inc, 2012.
7. Decreusefond, L. and Savy, N.: *Filtered Brownian motion as weak limit of filtered Poisson processes*, Bernoulli **11** (2005), no. 2, 283–292.
8. Dudley, R. M.: *Real Analysis and Probability*, The Wadsworth & Brooks/Cole Mathematics Series, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1989. MR 982264 (91g:60001)
9. Feyel, D. and de La Pradelle, A.: *On fractional Brownian processes*, Potential Anal. **10** (1999), no. 3, 273–288.
10. Itô, K. and Nisio, M.: *On the convergence of sums of independent Banach space valued random variables*, Osaka J. Math. **5** (1968), 35–48.
11. Kuo, H.-H.: *Gaussian Measures in Banach Spaces*, Lecture Notes in Mathematics, Vol. 463, Springer-Verlag, Berlin, 1975.
12. Lang, S. *Differential Manifolds*, Addison-Wesley Publishing Co., Inc., Reading, Mass.-London-Don Mills, Ont., 1972.
13. Nikiforov, A. F. and Uvarov, V. B.: *Special Functions of Mathematical Physics*, Birkhäuser, 1988.
14. Nourdin, I. and Peccati, G.: *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality*, Cambridge University Press, 2012.

15. Peccati, G., Solé, J. L., Taqqu, M. S., and Utzet, F.: *Stein's method and normal approximation of poisson functionals*, Annals of Probability **38** (2010), no. 2, 443–478.
16. Peccati, G. and Zheng, C.: *Multi-dimensional Gaussian fluctuations on the Poisson space*, Electron. J. Probab. **15** (2010), no. 48, 1487–1527.
17. Privault, N. *Stochastic Analysis in Discrete and Continuous Settings with Normal Martingales*, Lecture Notes in Mathematics, vol. 1982, Springer-Verlag, Berlin, 2009.
18. Samko, S. G., Kilbas, A. A., and Marichev, O. I.: *Fractional Integrals and Derivatives*, Gordon and Breach Science, 1993.
19. Schwartz, L.: *Séminaire Laurent Schwartz 1969–1970: Applications Radonifiantes*, Centre de Mathématiques,, 1970.
20. Shih, H. H.: *On Stein's method for infinite-dimensional Gaussian approximation in abstract Wiener spaces*, Journal of Functional Analysis **261** (2011), no. 5, 1236–1283.
21. Simon, B.: *Trace Ideals and Their Applications*, second ed., Mathematical Surveys and Monographs, vol. 120, American Mathematical Society, Providence, RI, 2005.
22. Villani, C.: *Optimal Transport, Old and New*, Lectures Notes in Mathematics, Springer Verlag, New York, 2007.

L. Coutin: Institute of Mathematics, Université Toulouse 3, Toulouse, France
E-mail address — : laure.coutin@math.univ-toulouse.fr

L. Decreusefond: Institut Telecom, Telecom ParisTech, CNRS LTCI, Paris, France
E-mail address — : Laurent.Decreusefond@telecom-paristech.fr