Stein's method for rough paths

Laure Coutin, Laurent Decreusefond

STEIN'S METHOD FOR ROUGH PATHS

L. COUTIN AND L. DECREUSEFOND

Abstract. The original Donsker theorem says that a standard random walk converges in distribution to a Brownian motion in the space of continuous functions. It has recently been extended to enriched random walks and enriched Brownian motion. We use the Stein-Dirichlet method to precise the rate of this convergence in the topology of fractional Sobolev spaces.

1. Introduction

The Donsker theorem says that a random walk

$$X^m(t) = \frac{1}{\sqrt{m}} \sum_{k=1}^{[mt]} X_k$$

where the X_k's are independent, identically distributed random variables with mean 0 and variance 1, converges in a functional space to the Brownian motion B. In the original version (see [8]), the convergence was proved to hold in the space of continuous functions. The first evolution was the paper of Lamperti [14], which proved the convergence in Hölder spaces. Namely, he stated that if the increments of the X_k's are p-integrable then the random walk X^m converges to B in $\mathrm{Hol}(1/2 - 1/p)$. The higher the integrability, the stronger the topology. There are numerous other extensions which can be made to the Donsker theorem. In the 90s, Barbour [2] estimated the rate of convergence in the space $\mathcal{C}$ of continuous functions on $[0, 1]$ equipped with a stronger topology than the usual sup-norm topology. He proved that, provided that $\mathbf{E}\left[|X_k|^3\right]$ is finite, then

$$\sup_{\|F\|_M \le 1} \mathbf{E}\left[F(X^m)\right] - \mathbf{E}\left[F(B)\right] \le c\, \frac{\log m}{\sqrt{m}},$$

where M is roughly speaking, the set of thrice Fréchet differentiable functions on $\mathcal{C}$ with bounded derivatives and $\|F\|_M$ is a function of the supremum of D^iF, $i = 0, \cdots, 3$ over $\mathcal{C}$. The strategy is to compare X^m with B^m, the affine interpolation, of mesh $1/m$, of the Brownian motion and then to compare B^m with B. The latter comparison is a sample-path comparison since the two processes live in the same probability space. This is the part which yields the $\log m$ factor. The former comparison is done via the Stein's method in finite dimension. The rate of this convergence is as usual (see [3, 15]) for Gaussian limits, of the order of $m^{-1/2}$.

In [5], we quantified the rate of convergence of X^m towards B in Besov-Liouville spaces (see (3), which are one scale of fractional Sobolev spaces.

2000 *Mathematics Subject Classification.* 60F17.
Key words and phrases. Donsker theorem, rough paths, Stein method.

The spaces we considered were not included in Hölder spaces but the method could be adapted to obtain convergence rate in Hölder spaces. However, it would not fit to the present context where we are considering enriched-paths (see definition below).

Note that in both [2] and [5], an higher integrability of the X_k's would not improve the convergence rates but would give more flexibility on the choice of the topology in which the convergence holds: The higher the integrability, the higher the Hölder exponent may be chosen.

In [12, Theorem 13.3.3], Friz and Victoir essentially showed that a Lamperti's like result holds for the convergence of the enriched random walk in the sense of rough path to the enriched Brownian motion.

The motivation of this paper is to quantify the rate of this convergence in rough-paths sense. The first difficulty is that the limiting process is no longer a Gaussian process: The Lévy area of a Brownian motion is not Gaussian. Hence we cannot expect to have a direct application of the Stein's method. However, there is no more randomness in the Lévy area that there is in the Brownian motion itself: The Lévy area is adapted to the filtration generated by the underlying Brownian motions. Saying that has two consequences. First, that the probability space we have to consider depends only on the Brownian motion. Moreover, we have to find functional spaces for which the map which sends a Brownian motion to its Lévy area is not only continuous but also Lipschitz. As mentioned in [13], the Besov-Liouville spaces (we used in [5]) are not well fitted to deal with the iterated integral processes we encounter in rough-paths theory. It is much better to work with the Slobodetsky scale of fractional Sobolev spaces.

Once the functional framework is set up, in order to avoid some complicated calculations in infinite dimensional spaces, the idea is to go back to the approach of [2]: Comparing the random walk with the affine interpolation of the Brownian motion in the Slobodetsky scale of fractional Sobolev spaces. This can be done by an application of the Stein's method in finite dimension. The novelty comes from the treatment of the iterated integrals whose existence hugely complicates the computations of the remainders. The final result is obtained by considering the known distance between the enriched affine interpolation and the enriched Brownian motion in fractional Sobolev spaces.

Rough paths theory is essentially a deterministic theory, it is therefore tempting to make the estimate we need in a deterministic setting and then to take the expectation of these bounds. This turns out to be a misleading approach. For instance, consider a sequence of centered, independent identically distributed random variables $(X_n,\, n \ge 1)$ and let $S_n = \sum_{i=1}^n X_i$. If we evaluate the p-th moment of S_n with Hölder inequality, we get that this moment is bounded by a constant times n^p. But if we use, as in the sequel, the Burkholder-Davis-Gundy inequality for discrete time martingale, we get an upper-bound proportional to $n^{p/2}$. This martingale argument which is implicitly used in [14] is the key to our work. For the sake of simplicity, this implies to separate the treatment of the symmetric and anti-symmetric parts of the signature. However, this is of no real importance since, as detailed below, the symmetric part of the signature can be handled as a classical $\mathbf{R}^d$-valued process.

This paper is organized as follows: In Section 2, we give the necessary notions about fractional Sobolev spaces and rough-paths theory. We also give a detailed proof of Lamperti's result in the fractional Sobolev spaces scale for further use and comparison. In Section 3, we then define the Kolmogorov-Rubinstein distance and show that this distance between the random walk and the affine interpolation of the Brownian motion can be reduced to a problem in finite dimension, should we consider a special set of Lipschitz functions. In Section 4, we then present our development of the Stein-Dirichlet method to estimate this distance.

2. Preliminaries

2.1. **Fractional Sobolev spaces.** As in [6, 12], we consider the fractional Sobolev spaces $W_{\eta,p}$ defined for $\eta \in (0,1)$ and $p \geq 1$ as the the closure of $\mathcal{C}^1$ functions with respect to the norm

$$|f|^p_{\eta,p} = \int_0^1 |f(t)|^p \, \mathrm{d}t + \iint_{[0,1]^2} \frac{|f(t)-f(s)|^p}{|t-s|^{1+p\eta}} \, \mathrm{d}t \, \mathrm{d}s.$$

For $\eta = 1$, $W_{1,p}$ is the completion of $\mathcal{C}^1$ for the norm:

$$|f|^p_{1,p} = \int_0^1 |f(t)|^p \, \mathrm{d}t + \int_0^1 |f'(t)|^p \, \mathrm{d}t.$$

They are known to be Banach spaces and to satisfy the Sobolev embeddings [1, 11]:

$$W_{\eta,p} \subset \mathrm{Hol}(\eta - 1/p) \text{ for } \eta - 1/p > 0$$

and

$$W_{\eta,p} \subset W_{\gamma,q} \text{ for } 1 \geq \eta \geq \gamma \text{ and } \eta - 1/p \geq \gamma - 1/q.$$

As a consequence, since $W_{1,p}$ is separable (see [4]), so does $W_{\eta,p}$. We need to compute the $W_{\eta,p}$ norm of primitive of step functions.

Lemma 2.1. *Let* $0 \leq s_1 < s_2 \leq 1$ *and consider*

$$h_{s_1,s_2}(t) = \int_0^t \mathbf{1}_{[s_1,s_2]}(r) \, dr.$$

There exists $c > 0$ *such that for any* s_1, s_2*, we have*

$$\|h_{s_1,s_2}\|_{W_{\eta,p}} \leq c \, |s_2 - s_1|^{1-\eta}. \tag{1}$$

Proof. Remark that for any $s, t \in [0,1]$,

$$|h_{s_1,s_2}(t) - h_{s_1,s_2}(s)| \leq |t-s| \wedge (s_2 - s_1).$$

The result then follows from the definition of the $W_{\eta,p}$ norm. □

We also need to introduce the Riemann-Liouville fractional spaces for the construction of abstract Wiener spaces. For $f \in L^1([0,1];\ dt)$, (denoted by

L^1 for short) the left and right fractional integrals of f are defined by :

$$(I^{\gamma}_{0^+} f)(x) = \frac{1}{\Gamma(\gamma)} \int_0^x f(t)(x-t)^{\gamma-1} \, \mathrm{d}t \ , \ x \geq 0,$$

$$(I^{\gamma}_{1^-} f)(x) = \frac{1}{\Gamma(\gamma)} \int_x^1 f(t)(t-x)^{\gamma-1} \, \mathrm{d}t \ , \ x \leq 1,$$

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

$$\int_0^1 f(s)(I^{\gamma}_{0^+} g)(s) \, \mathrm{d}s = \int_0^1 (I^{\gamma}_{1^-} f)(s) g(s) \, \mathrm{d}s. \tag{2}$$

The Besov-Liouville space $I^{\gamma}_{0^+}(L^p) := \mathcal{I}^+_{\gamma,p}$ is usually equipped with the norm :

$$\|I^{\gamma}_{0^+} f\|_{\mathcal{I}^+_{\gamma,p}} = \|f\|_{L^p}. \tag{3}$$

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

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

It is proved in [10] that for $1 \geq a > b > c > 0$ that the following embeddings are continuous (even compact)

$$W_{a,p} \subset \mathcal{I}^+_{b,p} \subset W_{c,p}.$$

2.2. **Rough paths.** We give a quick introduction to the rough-paths theory. For details, we refer to the monograph [12]. Consider $T^2(\mathbf{R}^d)$, the graded algebra of step two:

$$T^2(\mathbf{R}^d) = \mathbf{R} \oplus \mathbf{R}^d \oplus (\mathbf{R}^d \otimes \mathbf{R}^d).$$

We endow $T^2(\mathbf{R}^d)$ with an algebra structure $(+, ., \otimes)$ where for all (w_0, w_1, w_2), $(z_0, z_1, z_2) \in T^2(\mathbf{R}^d)$, $\lambda \in \mathbf{R}$

$$\begin{aligned} (w_0, w_1, w_2) + (z_0, z_1, z_2) &= (w_0 + z_0, \ w_1 + z_1, \ w_2 + z_2) \\ \lambda.(w_0, w_1, w_2) &= (\lambda w_0, \lambda w_1, \lambda w_2) \\ (w_0, w_1, w_2) \otimes (z_0, z_1, z_2) &= (w_0 z_0, \ w_0 z_1 + z_0 w_1, \ w_0 z_2 + z_0 w_2 + w_1 \otimes z_1). \end{aligned}$$

Introduce the projection maps: For $i = 0, 1, 2$

$$\begin{aligned} \pi_i \, : \, T^2(\mathbf{R}^d) &\longrightarrow (\mathbf{R}^d)^{\otimes i} \\ (w_0, w_1, w_2) &\longmapsto w_i, \end{aligned}$$

The set

$$\begin{aligned} 1 + t^2(\mathbf{R}^d) &= \{w \in T^2(\mathbf{R}^d), \, \pi_0(w) = 1\} \\ &= \Big\{ g = (1, w_1, w_2), \ (w_1, w_2) \in \mathbf{R}^d \oplus (\mathbf{R}^d \otimes \mathbf{R}^d) \Big\} \end{aligned}$$

is a Lie group with respect to the tensor multiplication $\otimes$, [12, Prop. 7.17]. Note that

$$(1, w_1, w_2)^{-1} = (1, -w_1, -w_2 + w_1 \otimes w_1).$$

As usual, a Lie group, like $1+t^2(\mathbf{R}^d)$, leads to a Lie algebra when equipped with notions of product and commutator. Here, the Lie algebra is $(t^2(\mathbf{R}^d),+,.)$ with product $\otimes$ and commutator

$$[g,w] = g\otimes w - w\otimes g = 0 \ \oplus \ \big(\pi_1(g)\otimes\pi_1(w) - \pi_1(w)\otimes\pi_1(g)\big),$$
$$\text{for any } w,g\in t^2(\mathbf{R}^d).$$

Denote by $(e_i,\, 1\le i\le d)$ the canonical basis of $\mathbf{R}^d$, so that $(e_i\otimes e_j,\, 1\le i,j\le d)$ is the canonical basis of $\mathbf{R}^d\otimes\mathbf{R}^d$. Then

$$[\sum_{i=1}^d a_i\, e_i + \sum_{i,j=1}^d c_{i,j}\, e_i\otimes e_j,\ \sum_{i=1}^d b_i\, e_i + \sum_{i,j=1}^d f_{i,j}\, e_i\otimes e_j]$$
$$= \sum_{i<j}(a_ib_j - a_jb_i)\,[e_i,e_j].$$

The exponential and logarithm maps are useful to go back and forth between $t^2(\mathbf{R}^d)$ and $1+t^2(\mathbf{R}^d)$:

$$\begin{aligned}\exp\,:\, t^2(\mathbf{R}^d) &\longrightarrow 1+t^2(\mathbf{R}^d)\\ w &\longmapsto 1+w+\frac{1}{2}(\pi_1 w)^{\otimes 2}\end{aligned}$$

and

$$\begin{aligned}\log\,:\, 1+t^2(\mathbf{R}^d) &\longrightarrow t^2(\mathbf{R}^d)\\ 1+w &\longmapsto w-\frac{1}{2}(\pi_1 w)^{\otimes 2}.\end{aligned}$$

We denote by Σ the set of finite partitions $\sigma=\{t_1,\cdots,t_n\}$ of $[0,1]$. A continuous path z from $[0,1]$ into $\mathbf{R}^d$ is said to have 1-finite variation whenever

$$\sup_{\sigma=\{t_1,\cdots,t_n\}\in\Sigma}\sum_{i=1}^{n-1}|z_{t_{i+1}}-z_{t_i}|<\infty.$$

The set of such functions equipped with this quantity as a norm is denoted by $C^{1-\text{var}}$.

Definition 2.1. The step-2 signature of $z\in C^{1-\text{var}}$ is given by:

$$\begin{aligned}S_2(z)\,:\,[0,1] &\longrightarrow 1+t^2(\mathbf{R}^d)\\ t &\longmapsto \left(1,\ z_t-z_0,\ \int_0^t (z_s-z_0)\otimes dz_s\right).\end{aligned}$$

The free nilpotent group of order 2, $G^2(\mathbf{R}^d)$, is the closed subgroup of $1+t^2(\mathbf{R}^d)$ defined by

$$G^2(\mathbf{R}^d)=\left\{S_2(z),\, z\in C^{1-Var}\right\}.$$

We also consider $\check{\mathrm{G}}_2(\mathbf{R}^d)$ ($\check{\mathrm{G}}_2$ for short since d is fixed), the image of $\mathrm{G}_2(\mathbf{R}^d)$ by the logarithm map. For $z\in C^{1-\text{var}}$, this corresponds to consider only the anti-symmetric part of $\pi_2(S_2(z))$:

$$\begin{aligned}\check{S}_2(z)\,:\,[0,1] &\longrightarrow t^2(\mathbf{R}^d)\\ t &\longmapsto \left(z_t - z_0,\ \int_0^t \Big[(z_s - z_0),\ \mathrm{d}z_s\Big]\right).\end{aligned}$$

Remark 1. If

$$z(t) = \sum_{i=1}^{d}\sum_{k=1}^{m} z_{ik}\, h_k(t)\ e_i$$

where $(h_1, \cdots, h_m)$ are elements of $C^{1-\text{var}}$, we have

$$\log S_2(z)(t) = \Bigg(z_t - z_0,\\ \sum_{1\le i<j\le d}\ \sum_{1\le k<l\le m} (z_{ik}z_{jl} - z_{il}z_{jk}) \left(\int_0^t h_k(s)\ \mathrm{d}h_l(s) - \int_0^t h_l(s)\ \mathrm{d}h_k(s)\right)\ [e_i, e_j]\Bigg). \tag{4}$$

For the sake of notations, we set $\mathcal{A} = \{1, \cdots, d\} \times \{1, \cdots, m\}$ and define the $\prec$ relation by:

$$a = (a_1, a_2) \prec b = (b_1, b_2) \iff (a_1 < b_1) \text{ and } (a_2 < b_2).$$

With these notations, Eqn. (4) then becomes

$$\log S_2(z)(t) = \Bigg(z_t - z_0,\\ \sum_{a\prec b} [z_a,\, z_b] \left(\int_0^t h_{a_2}(s)\ \mathrm{d}h_{b_2}(s) - \int_0^t h_{b_2}(s)\ \mathrm{d}h_{a_2}(s)\right)\ [e_{a_1}, e_{b_1}]\Bigg).$$

The group $G^2(\mathbf{R}^d)$ has the structure of a sub-Riemannian manifold. We will not dwell into the meanders of this very rich but intricate structure. It suffices to say that we can proceed equivalently by considering usual norms as follows.

For $\alpha \in (0, 1)$, a path $w = 1 \oplus w_1 \oplus w_2$ is said to be α-Hölder whenever

$$\rho_\alpha(w) = \max\left(\sup_{s\neq t} \frac{|w_1(t) - w_1(s)|}{|t-s|^\alpha},\ \sup_{s\neq t} \frac{|\pi_2(w(s)^{-1} \otimes w(t))|^{1/2}}{|t-s|^\alpha} \right) < \infty.$$

Note that for $z \in C^{1-\text{var}}$,

$$\rho_\alpha(S_2(z)) = \max\left(\sup_{s\neq t} \frac{|z(t) - z(s)|}{|t-s|^\alpha},\ \sup_{s\neq t} \frac{\left|\int_s^t (z_r - z_s) \otimes\ \mathrm{d}z_r\right|^{1/2}}{|t-s|^\alpha} \right).$$

Definition 2.2. We denote by $H_\alpha(t^2(\mathbf{R}^d))$, the vector space of paths w from $[0, 1]$ into $t^2(\mathbf{R}^d)$ such that $\rho_\alpha(w)$ is finite. It is equipped with the homogeneous norm: For w and v in $H_\alpha(t^2(\mathbf{R}^d))$

$$\|w - v\|_{H_\alpha(t^2(\mathbf{R}^d))} = \rho_\alpha(w - v) = \rho_\alpha\big((w_1 - v_1) \oplus (w_2 - v_2)\big).$$

Unfortunately, as mentioned in [12, Chapter 8.3], this metric space is complete but not separable, which is unacceptable for our purpose (see Definition 3.1 and the remark below). We thus introduce fractional Sobolev spaces as in [13].

Definition 2.3. For any $\eta \in (0,1)$, any $p \geq 2$, $\check{\mathrm{G}}_2 W_{\eta,p}$ is the vector space of paths w from $[0,1]$ into $t^2(\mathbf{R}^d)$ such that

$$\|w_1\|_{\eta,p}^p + \iint_{[0,1]^2} \frac{\left|\pi_2[w_s^{-1},\, w_t]\right|^{p/2}}{|t-s|^{1+\eta p}} \,\mathrm{d}s\ \mathrm{d}t < \infty.$$

The distance on $\check{\mathrm{G}}_2 W_{\eta,p}$ is defined by

$$\|w-v\|_{\check{\mathrm{G}}_2 W_{\eta,p}} = \|\pi_1(w)-\pi_1(v)\|_{W_{\eta,p}} + \left(\iint_{[0,1]^2} \frac{\left|\pi_2[w_s^{-1},\, w_t] - \pi_2[v_s^{-1},\, v_t]\right|^{p/2}}{|t-s|^{1+\eta p}} \,\mathrm{d}s\ \mathrm{d}t\right)^{1/p}.$$

Following [12], we know that $\check{\mathrm{G}}_2 W_{\eta,p}$ is a Banach space included into $H_\alpha(t^2(\mathbf{R}^d))$ for $\alpha = \eta - 1/p$, provided $\alpha > 0$.

Lemma 2.2. *For any $\eta \in (0,1)$, any $p \geq 2$, $\check{\mathrm{G}}_2 W_{\eta,p}$ is separable.*

Proof. Consider the map κ defined as

$$\begin{aligned}\kappa\, :\, ([0,1] \to t_2(\mathbf{R}^d)) &\longrightarrow ([0,1]^2 \to t_2(\mathbf{R}^d)) \\ w &\longmapsto \Big((s,t) \mapsto \Big(\pi_1(w_t) - \pi_1(w_s),\, \pi_2\big[w_s^{-1},\, w_t\big]\Big)\Big).\end{aligned}$$

Consider the measure $\mathrm{d}\mu_{\eta,p}(s,t) = |t-s|^{-1-\eta p}\ \mathrm{d}s\ \mathrm{d}t$. Then,

$$\|w\|_{\check{\mathrm{G}}_2 W_{\eta,p}} = \|\pi_1 \circ \kappa(w)\|_{L^p(\mu_{\eta,p})} + \|\pi_2 \circ \kappa(w)\|^{1/2}_{L^{p/2}(\mu_{2\eta,p/2})}. \tag{5}$$

For any $p \geq 1$ and $\eta \in (0,1)$, $L^p(\mu_{\eta,p})$ is isometrically isomorphic to $L^p(\mathrm{d}s \otimes \mathrm{d}t)$ hence it is separable. This entails that $E = L^p(\mu_{\eta,p}) \times L^{p/2}(\mu_{2\eta,p/2})$ is separable. Equation (5) means that the application T which maps $w \in \check{\mathrm{G}}_2 W_{\eta,p}$ to the couple $(\pi_1 \circ \kappa(w),\ \pi_2 \circ \kappa(w)) \in E$, is an isometry. Thus $\check{\mathrm{G}}_2 W_{\eta,p}$ is isometrically isomorphic to a closed subspace of the separable space E, hence it is separable. □

2.3. **Donsker-Lamperti theorem.** For the sake of completeness and for further comparison, we give the proof of the Donsker-Lamperti theorem in the scale of fractional Sobolev spaces, which induces the convergence in Hölder spaces.

Definition 2.4. The random walk associated to the sequence $(X_k,\, k \geq 1)$ is defined by

$$X^m(t) = \sqrt{m}\, \sum_{k=1}^m X_k\, r_k^m(t) = \sum_{k=1}^m X_k\, h_{s_1,s_2}(t)$$

where

$$r_k^m(t) = \int_0^t \mathbf{1}_{((k-1)/m,\ k/m]}(s)\ \mathrm{d}s \text{ and } h_{s_1,s_2} = \sqrt{m}\ r_k^m. \tag{6}$$

Theorem 2.3. *If for any $k \geq 1$, X_k belongs to L^p for some $p \geq 2$, then there exists $c > 0$ such that*

$$\sup_{m\geq 1} \frac{\mathbf{E}\left[\left|\sum_{k=1}^m X_k\ \sqrt{m}\left(r_k^m(t) - r_k^m(s)\right)\right|^p\right]}{|t-s|^{p/2}} < c\,\mathbf{E}\left[|X_1|^p\right] \tag{7}$$

Proof. For $0 \leq s < t \leq 1$ fixed, the discrete time process

$$Y^{st}\ :\ m \longmapsto Y^{st}_m = \sum_{k=1}^m X_k(r_k^m(t) - r_k^m(s))$$

is a martingale with respect to the filtration $\mathcal{F}_n = \sigma(X_k,\, 1 \leq k \leq m)$. The Burkholder-Davis-Gundy [18] entails that

$$m^p\mathbf{E}\left[\left|\sum_{k=1}^m X_k\left(r_k^m(t) - r_k^m(s)\right)\right|^p\right] \leq c\,m^{p/2}\mathbf{E}\left[\left|\sum_{k=1}^m X_k^2\left(r_k^m(t) - r_k^m(s)\right)^2\right|^{p/2}\right].$$

If $|t-s| \leq 1/m$, there is at most two values of k such that $r_k^m(t) - r_k^m(s)$ is not zero. Furthermore,

$$|r_k^m(t) - r_k^m(s)| \leq |t-s|$$

hence

$$|r_k^m(t) - r_k^m(s)|^2 \leq m^{-1}|t-s|.$$

In this situation,

$$\mathbf{E}\left[\left|\sum_{k=1}^m X_k^2\left(r_k^m(t) - r_k^m(s)\right)^2\right|^{p/2}\right] \leq c\,m^{-p/2}\,\mathbf{E}\left[|X_1|^p\right]\ |t-s|^{p/2},$$

so that (7) holds true for $|t-s| \leq 1/m$. For $|t-s| > 1/m$, we remark that $r_k^m(t) - r_k^m(s)$ is not null for at most $[m(t-s)] + 2$ values of k and since r_k^m is Lipschitz continuous, $|r_k^m(t) - r_k^m(s)| \leq 1/m$ for such value of k. Hence, by convexity inequality,

$$\begin{aligned}\mathbf{E}\left[\left|\sum_{k=1}^m X_k^2\left(r_k^m(t) - r_k^m(s)\right)^2\right|^{p/2}\right] &\leq c\,(m|t-s| + 2)^{p/2}\,m^{-p}\,\mathbf{E}\left[|X_1|^p\right] \\ &\leq cm^{-p/2}\ \mathbf{E}\left[|X_1|^p\right]\ |t-s|^{p/2}.\end{aligned}$$

Hence, (7) is true for $|t-s| \geq 1/m$. □

It is then straightforward that we have:

Corollary 2.4. *Assume that for any $k \geq 1$, X_k belongs to L^p for some $p \geq 2$. Then, for any $\eta < 1/2$,*

$$\sup_{m\geq 1} \mathbf{E}\left[\|X^m\|^p_{W_{\eta,p}}\right] < \infty.$$

Proof. Actually, $(s,t) \mapsto |t-s|^{p/2}$ is $\mu_{\eta,p}$-integrable provided that $p(1/2 - \eta) > 0$, i.e. $\eta < 1/2$. □

Corollary 2.5 (Lamperti). *Assume that for any $k \ge 1$, X_k belongs to L^p for some $p \ge 2$. Then, for any $1/p < \eta < 1/2$, the sequence $(X^m,\, m \ge 1)$ converges in distribution in* $\operatorname{Hol}(\eta - 1/p)$ *to B.*

Proof. It is well-known that the finite dimensional distributions of X^m converge to that of B. From Corollary 2.4, we know that for any $0 < \zeta < 1/2$, for any $\epsilon > 0$, there exists K_ϵ such that

$$\sup_{m \ge 1} \mathbf{P}(\|X^m\|_{W_{\zeta,\,p}} \ge K_\epsilon) \le \epsilon.$$

For $0 < \eta < \zeta < 1/2$, the embedding of $W_{\zeta,p}$ into $W_{\eta,\,p}$ is compact: The $W_{\zeta,\,p}$-ball of radius K_ϵ is compact in $W_{\eta,p}$. Thus, the sequence $(X^m,\, m \ge 1)$ is tight in $W_{\eta,p}$ hence convergent. The result follows by the continuous embedding of $W_{\eta,p}$ into $\operatorname{Hol}(\eta - 1/p)$. □

2.4. **Abstract Wiener spaces.** The construction of the Gaussian measure on a Banach space is a delicate question, we refer to [17, 19] for details. For H a Hilbert space, a cylindrical set is a set of the form

$$Z = \{x \in H,\ (\langle x, h_1\rangle_H, \cdots, \langle x, h_n\rangle) \in B\}$$

for some integer n, where $(h_1, \cdots, h_n)$ is an orthonormal family of H and B a Borelean subset of $\mathbf{R}^n$. Let m_n be the standard Gaussian measure on $\mathbf{R}^n$. By setting, $\mu_H(Z) = m_n(B)$, we get a cylindrical standard Gaussian measure on H. To get a Radon measure on a Banach space W, the usual way is to find a map from H to W which is radonifying: It is a linear map which transforms a cylindrical measure into a true regular Radon measure. We will not dwell into the details of this theory, it suffices to say that we have the following result : (see [19, Proposition XV,4,1] or [17, 25.6.3]),

Lemma 2.6. *Let*

$$\Lambda = \{(\eta, p) \in \mathbf{R}^+ \times \mathbf{R}^+, 0 < \eta - 1/p < 1/2\}.$$

For any $(\eta, p) \in \Lambda$, the embedding

$$\mathcal{I}_{1,2} \xrightarrow{\iota_{\eta,p}} W_{\eta,\,p} \subset \operatorname{Hol}(\eta - 1/p).$$

is a radonifying map.

Definition 2.5. An abstract Wiener space is a triple (ι, H, W) where H is a separable Hilbert space, W a Banach space and ι the embedding from H into B which has to be to be radonifying.

We have the following diagram:

$$W^* \xrightarrow{\iota^*} H^* \xrightarrow{j_{H^*,H}} H \xrightarrow{\iota} W, \tag{8}$$

where $j_{H^*,H}$ is the bijective isometry between the Hilbert space H^* and its dual H. As a direct consequence of Lemma [2.6], we have

Theorem 2.7. *The triple $(\iota_{\eta,p},\ \mathcal{I}_{1,2},\ W_{\eta,\,p})$ is an abstract Wiener space, for any $(\eta, p) \in \Lambda$.*

The Wiener measure $\mathbf{P}_{\eta,p}$ on $W_{\eta,p}$, is defined by its characteristic function: For all $\eta \in W^*_{\eta,p}$,

$$\int_{W_{\eta,p}} e^{i\,\langle \eta,\, y\rangle_{W^*_{\eta,p},W_{\eta,p}}} \, \mathrm{d}\mathbf{P}_{\eta,p}(y) = \exp(-\frac{1}{2}\|j_{H^*,H}\circ \iota^*_{\eta,p}(\eta)\|^2_{\mathcal{I}_{1,2}}).$$

This means that for any $\eta \in W^*_{\eta,p}$, the random variable $\langle \eta,\, y\rangle_{W^*_{\eta,p},W_{\eta,p}}$ is a centered Gaussian random variable with variance given by

$$\int_{W_{\eta,p}} \langle \eta,\, w\rangle^2_{W^*_{\eta,p},W_{\eta,p}} \, \mathrm{d}\mathbf{P}_{\eta,p}(y) = \|j_{H^*,H}\circ \iota^*_{\eta,p}(\eta)\|^2_{\mathcal{I}_{1,2}}.$$

Remark 2. In what follows, as it is customary, we identify $\mathcal{I}_{1,2}$ and its dual so that the diagram (8) becomes

$$W^*_{\eta,p} \xrightarrow{\mathfrak{i}^*_{\eta,p}=j_{H^*,H}\circ\iota^*_{\eta,p}} \mathcal{I}_{1,2} \xrightarrow{\iota_{\eta,p}} W_{\eta,p}.$$

By construction, $\mathfrak{i}^*_{\eta,p}(W^*_{\eta,p})$ is dense in $\mathcal{I}_{1,2}$ so that we can define the Wiener integral as follows.

Definition 2.6 (Wiener integral)**.** The Wiener integral, denoted as $\delta_{\eta,p}$, is the isometric extension of the map

$$\begin{aligned} \delta_{\eta,p}\ :\ \mathfrak{i}^*_{\eta,p}(W^*_{\eta,p}) \subset \mathcal{I}_{1,2} &\longrightarrow L^2(\mathbf{P}_{\eta,p}) \\ \mathfrak{i}^*_{\eta,p}(\eta) &\longmapsto \langle \eta,\, y\rangle_{W^*_{\eta,p},W_{\eta,p}}. \end{aligned}$$

This means that if $h = \lim_{n\to\infty} \mathfrak{i}^*_{\eta,p}(\eta_n)$ in $\mathcal{I}_{1,2}$,

$$\delta_{\eta,p}h(y) = \lim_{n\to\infty} \langle \eta_n,\, y\rangle_{W^*_{\eta,p},W_{\eta,p}} \text{ in } L^2(\mathbf{P}_{\eta,p}).$$

Remark 3. As the Dirac measure at point $t \in [0,1]$ belongs to any $W^*_{\eta,p}$, we an search for $\mathfrak{i}^*_{\eta,p}(\varepsilon_t)$. For any $h \in \mathcal{I}_{1,2} \subset W_{\eta,p}$, we must have

$$h(t) = \langle \varepsilon_t,\, h\rangle_{W^*_{\eta,p},W_{\eta,p}} = \int_0^1 \dot{\overline{\mathfrak{i}^*_{\eta,p}(\varepsilon_t)}}(s)\dot{h}(s)\,\mathrm{d}s$$

hence

$$\dot{\overline{\mathfrak{i}^*_{\eta,p}(\varepsilon_t)}}(s) = \mathbf{1}_{[0,t]}(s) \text{ and } \mathfrak{i}^*_{\eta,p}(\varepsilon_t)(s) = t\wedge s.$$

This means, that whatever the functional space $W_{\eta,p}$ we are considering,

$$B_{\eta,p} = (\delta_{\eta,p}(t\wedge .),\, t \in [0,1])$$

is a centered Gaussian process of covariance kernel

$$\mathbf{E}\left[B_{\eta,p}(t)B_{\eta,p}(s)\right] = \langle t\wedge .,\, s\wedge .\rangle_{\mathcal{I}_{1,2}} = t\wedge s.$$

Hence, $B_{\eta,p}$ is a standard Brownian motion. Since we work with a sequence of increasing (in the sense of inclusion) spaces, we remove the subscripts when no risk of confusion may happen.

Definition 2.7. A function F from $W_{\eta,p}$ into $\mathbf{R}$ is Lipschitz whenever for any x and y in $W_{\eta,p}$,

$$|F(x) - F(y)| \le \|x-y\|_{W_{\eta,p}}.$$

The set of such functions is denoted by $\mathrm{Lip}(d_{W_{\eta,p}})$.

Definition 2.8 (Ornstein-Uhlenbeck semi-group)**.** For any bounded function on $W_{\eta,p}$, for any $\tau \geq 0$,

$$P_\tau F(x) = \int_{W_{\eta,p}} F(e^{-\tau}x + \beta_\tau y) \, \mathrm{d}\mathbf{P}_{\eta,p}(y)$$

where $\beta_\tau = \sqrt{1 - e^{-2\tau}}$.

The dominated convergence theorem entails that P_τ is ergodic: For any $x \in W_{\eta,p}$,

$$P_\tau F(x) \xrightarrow{\tau \to \infty} \int_{W_{\eta,p}} F \, \mathrm{d}\mathbf{P}_{\eta,p}.$$

Moreover, the invariance by rotation of Gaussian measures implies that

$$\int_{W_{\eta,p}} P_\tau F(x) \, \mathrm{d}\mathbf{P}_{\eta,p}(x) = \int_{W_{\eta,p}} F \, \mathrm{d}\mathbf{P}_{\eta,p}, \text{ for any } \tau \geq 0.$$

Otherwise stated, the Gaussian measure on $W_{\eta,p}$ is the invariant and stationary measure of the semi-group $P = (P_\tau, \, \tau \geq 0)$. For details on the Malliavin gradient, we refer to [16, 21].

Definition 2.9. Let X be a Banach space. A function $F : W_{\eta,p} \to X$ is said to be cylindrical if it is of the form

$$F(y) = \sum_{j=1}^{k} f_j(\delta h_1(y), \cdots, \delta h_k(y)) \, x_j$$

where for any $j \in \{1, \cdots, k\}$, f_j belongs to the Schwartz space on $\mathbf{R}^k$, $(h_1, \cdots, h_k)$ are elements of $\mathcal{I}_{1,2}$ and $(x_1, \cdots, x_j)$ belong to X. The set of such functions is denoted by $\mathfrak{C}(X)$.

For $h \in \mathcal{I}_{1,2}$,

$$\langle \nabla F, \, h \rangle_{\mathcal{I}_{1,2}} = \sum_{j=1}^{k} \sum_{l=1}^{k} \partial_l f(\delta h_1(y), \cdots, \delta h_k(y)) \; \langle h_l, \, h \rangle_{\mathcal{I}_{1,2}} \; x_j,$$

which is equivalent to say

$$\nabla F = \sum_{j,l=1}^{k} \partial_j f(\delta h_1(y), \cdots, \delta h_k(y)) \, h_l \otimes \; x_j.$$

The space $\mathbb{D}_{1,2}(X)$ is the closure of the space of cylindrical functions with respect to the norm

$$\|F\|_{1,2}^2 = \|F\|^2_{L^2(\mathbf{P}_{\eta,p};X)} + \|\nabla F\|^2_{L^2(\mathbf{P}_{\eta,p};\mathcal{I}_{1,2} \otimes X)}.$$

By induction, higher order gradients are defined similarly. For any $k \geq 1$, the norm on the space $\mathbb{D}_{k,2}(X)$ is given by

$$\|F\|_{k,2}^2 = \|F\|^2_{L^2(\mathbf{P}_{\eta,p};X)} + \sum_{j=1}^{k} \|\nabla^{(j)} F\|^2_{L^2(\mathbf{P}_{\eta,p};\mathcal{I}_{1,2}^{\otimes j} \otimes X)}.$$

According to [20], we have the following properties of $\mathbf{P}_t F$.

Proposition 1. *Let $F \in L^1(W_{\eta,p}, \mathbf{P}_{\eta,p})$ and $x, y \in W_{\eta,p}$. For any $t > 0$, $P_tF(x)$ belongs to $\mathbb{D}_{k,2}$ for any $k \geq 1$. Moreover, the operator $\nabla^{(2)}P_tF(x)$ is trace-class. Let L be the formal operator defined by*

$$LG(x) = -\langle x, \nabla G(x)\rangle_{W_{\eta,p},W^*_{\eta,p}} + \operatorname{trace}_{W_{1,2}}(\nabla^{(2)}G(x)).$$

Then, for any $t > 0$, P_tF belongs to the domain of L and

$$\frac{d}{dt}P_tF(x) = LP_tF(x).$$

3. Rate of convergence

3.1. **Kolmogorov-Rubinstein distance.** In [14], the proof of Lamperti's Theorem is given for one dimensional processes but it can be straightforwardly adapted to $\mathbf{R}^d$-valued random walks and Brownian motion: $X^{\mathcal{A}}$ becomes the $\mathbf{R}^d$-valued process

$$X^{\mathcal{A}}(t) = \sqrt{m}\sum_{a\in\mathcal{A}} X_a\ r_{a_2}(t)\, e_{a_1} = \sum_{a\in\mathcal{A}} X_a\, h_a(t)$$

where $h_a(t) = \sqrt{m}\ r_{a_2}(t)\, e_{a_1}$ and $(X_a,\, a \in \mathcal{A})$ is a family of independent identically distributed random variables of mean 0 and variance 1. Furthermore, B is the d-dimensional Brownian motion:

$$B(t) = \sum_{i=1}^{d} B_i(t)\, e_i.$$

The enriched Brownian motion $\mathbb{B}$, is the $\mathrm{G}_2(\mathbf{R}^d)$-value process defined by

$$\mathbb{B}(t) = 1 \oplus B(t) \oplus \sum_{i,j=1}^{d}\int_0^t B_i(s)\circ\ \mathrm{d}B_j(s)\ e_i\otimes e_j,\ \text{ for any } t\in[0,1],$$

where the stochastic integrals are to be understood in the Stratonovitch sense. Theorem 13.32 of [12] says that $S_2(X^{\mathcal{A}})$ converges to $\mathbb{B}$ in some Hölder type spaces. Our primary goal is to give the rate of this convergence. For, we need to define a distance between probability measures over Hölder spaces. There are several possibilities of such a definition, the best suited for an estimate by the Stein method is the Kolmogorov-Rubinstein[1] distance:

Definition 3.1 (Kolmogorov-Rubinstein distance)**.** For μ and ν two probability measures on a metric space (W, d_W), their Kolmogorov-Rubinstein distance is given by

$$\operatorname{dist}_{\mathrm{KR}}(\mu,\, \nu) = \sup_{F\in \mathrm{Lip}(d_W)} \int_W F\ \mathrm{d}\mu - \int_W F\ \mathrm{d}\nu,$$

where

$$\mathrm{Lip}(d_W) = \Big\{F\,:\, W\to\mathbf{R},\ \forall w,\, v\in W,\ |F(w)-F(v)|\leq d_W(w,\, v)\Big\}.$$

[1]We stick to the denomination suggested in [22] even if this distance is often called the Wasserstein distance.

Theorem 11.3.3 of [9] states that the topology induced by this distance on the set of probability measures on W is the same as the topology of the convergence in law whenever the metric space W is separable. Unfortunately, as we already mentioned, Hölder spaces are not separable, thus to have a meaningful result, we turn to work on fractional Sobolev spaces. It is of no importance since Sobolev embeddings ensure that convergence in fractional Sobolev spaces induces convergence in Hölder spaces.

Our new goal is then to estimate the Kolmogorov-Rubinstein distance in $\mathrm{G}_2 W_{\eta,p}$ between $\mathbb{B}$ and $S_2(X^{\mathcal{A}})$. Remark that

$$\begin{aligned}\pi_2 S_2(X^{\mathcal{A}}) &= \pi_2 \check{\mathrm{S}}_2(X^{\mathcal{A}}) \\ &\quad + 2\sum_{a\in\mathcal{A}}(X_a^2-1)\int_0^t h_a^m(s)\otimes \,\mathrm{d}h_a^m(s) + \sum_{a\in\mathcal{A}} h_a^m(t)\otimes h_a^m(t) \\ &= U_1^m + U_2^m + U_3^m.\end{aligned}$$

On the other hand,

$$\begin{aligned}\mathbb{B}(t) &= \sum_{1\le i<j\le d}\left(\int_0^t B_i(s)\circ \,\mathrm{d}B_j(s) - \int_0^t B_j(s)\circ \,\mathrm{d}B_i(s)\right)\,[e_i,\, e_j] \\ &\quad + 2\sum_{i=1}^d \int_0^t B_i(s)\,\mathrm{d}B_i(s)\; e_i\otimes e_i + \sum_{i=1}^d t\, e_i \\ &= \log \mathbb{B}(t) + U_2 + U_3.\end{aligned}$$

where in U_2, the stochastic integral is taken in the Itô sense. Direct computations show that $U_3^m - U_3$ tends to 0 as $1/m$. The sequence U_2^m converges to U_2 as fast as $X^{\mathcal{A}}$ converges to B; a rate which is expected and which will turn out to be much slower than $1/m$. In summary, the Kolmogorov-Rubinstein distance between $\mathbb{B}$ and $S_2(X^{\mathcal{A}})$ has the same asymptotic behavior as the distance between $\log\mathbb{B}$ and $\check{\mathrm{S}}_2(X^{\mathcal{A}})$. Our final objective is then to estimate the Kolmogorov-Rubinstein distance between the distributions of $\log\mathbb{B}$ and $\check{\mathrm{S}}_2(X^{\mathcal{A}})$ in $\check{\mathrm{G}}_2 W_{\eta,p}$.

3.2. **Reduction to finite dimension.** Should we follow the same procedure as the one we used in [5], we would face the same complications to compute the trace term in some infinite dimensional space. We remark that $X^{\mathcal{A}}$ belongs to the finite dimensional space

$$V = \operatorname{span}\{h_a^m,\, a\in\mathcal{A}\} \subset \mathcal{I}_{1,2}.$$

Consider $(g_n,\, n\ge 1)$ a complete orthonormal basis of $V^\perp$ in $\mathcal{I}_{1,2}$. The Itô-Nisio Theorem says that B can be represented as the $W_{\eta,p}$-convergent sum

$$B = \sum_{a\in\mathcal{A}} \delta h_a^m\, h_a^m + \sum_{n=1}^\infty \delta g_n\, g_n = B_V + B_V^\perp,$$

where δh is the Malliavin divergence (or Wiener integral) associated to B (see [16]). It turns out that B_V is also the affine interpolation of B so that we may use the results of [12] to estimate the distance between $\check{\mathrm{S}}_2(B_V)$ and $\log\mathbb{B}$. We can then resort to the Stein's method in finite dimension to estimate only the distance between $\check{\mathrm{S}}_2(X^{\mathcal{A}})$ and $\check{\mathrm{S}}_2(B_V)$. We can always write

$$\begin{aligned}\sup_{F\in \mathrm{Lip}(d_{\check{\mathrm{G}}_2 W_{\eta,p}})} & \mathbf{E}\left[F(\log \mathbb{B})\right]-\mathbf{E}\left[F(\check{\mathrm{S}}_2(X^{\mathcal{A}}))\right]\\ \le \sup_{F\in \mathrm{Lip}(d_{\check{\mathrm{G}}_2 W_{\eta,p}})} & \mathbf{E}\left[F(\log \mathbb{B})\right]-\mathbf{E}\left[F(\check{\mathrm{S}}_2(B_V))\right]\\ & + \sup_{F\in \mathrm{Lip}(d_{\check{\mathrm{G}}_2 W_{\eta,p}})} \mathbf{E}\left[F(\check{\mathrm{S}}_2(B_V))\right]-\mathbf{E}\left[F(\check{\mathrm{S}}_2(X^{\mathcal{A}}))\right].\end{aligned}$$

On the one hand, since $\log \mathbb{B}$ and $\check{\mathrm{S}}_2(B_V)$ live on the same probability space, for $F\in \mathrm{Lip}(d_{\check{\mathrm{G}}_2 W_{\eta,p}})$, according to [12, Proposition 13.20],

$$\begin{aligned}\mathbf{E}\left[F(\log \mathbb{B})\right]-\mathbf{E}\left[F(\check{\mathrm{S}}_2(B_V))\right] &\le \mathbf{E}\left[\| \log \mathbb{B}-\check{\mathrm{S}}_2(B_V)\|_{\check{\mathrm{G}}_2 W_{\eta,p}}\right]\\ &\le c\, m^{-(1/2-\eta)}.\end{aligned} \tag{9}$$

It remains to estimate

$$\sup_{F\in \mathrm{Lip}(d_{\check{\mathrm{G}}_2 W_{\eta,p}})} \mathbf{E}\left[F(\check{\mathrm{S}}_2(B_V))\right]-\mathbf{E}\left[F(\check{\mathrm{S}}_2(X^{\mathcal{A}}))\right].$$

Actually, for technical reasons, we could not make this estimate for F only Lipschitz. As for the multivariate Gaussian approximation, we must have a condition on the regularity of the second derivative of test functions.

Definition 3.2. Let $\mathcal{I}^{\pm}_{2,2}=(I^1_{0^+}\circ I^1_{1^-})(L^2)$. For $F\,:\,\check{\mathrm{G}}_2 W_{\eta,p}\longrightarrow \mathbf{R}$, let $\check{F}=F\circ \check{\mathrm{S}}_2$ as described in Figure 1.

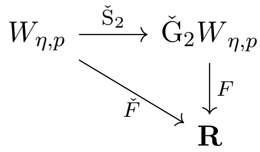

FIGURE 1. Definition of $\check{F}$.

Let $\Sigma_{\eta,p}$ be the set of functions $F\in \mathrm{Lip}(\check{\mathrm{G}}_2 W_{\eta,p})$ such that $\nabla^{(2)}\check{F}$ belongs to $L^2(W_{\eta,p};\,\mathcal{I}^{\pm}_{2,2}\otimes \mathcal{I}^{\pm}_{2,2})$ and

$$\begin{aligned}\left|\left\langle (\nabla^{(2)}\check{F})(x)-(\nabla^{(2)}\check{F})(x+g),\, h\otimes k\right\rangle_{\mathcal{I}_{1,2}^{\otimes 2}}\right| \\ \le \|h\|_{L^2}\,\|k\|_{L^2}\,\|\check{\mathrm{S}}_2(x)-\check{\mathrm{S}}_2(x+g)\|_{\check{\mathrm{G}}_2 W_{\eta,p}},\end{aligned}$$

for any $x\in W_{\eta,p}$, for any $g\in \mathcal{I}_{1,2}$, for any $h,k\in L^2$.

Remark 4. If $\check{F}$ is thrice differentiable in the direction of $\mathcal{I}_{1,2}$ with for any $x\in W_{\eta,p}$, $\nabla^{(j)}\check{F}(x)\in (\mathcal{I}^{\pm}_{2,2})^{\otimes j}$ for any $j=1,2,3$ and

$$\|\nabla^{(3)}F\|_{L^\infty(W;(\mathcal{I}^{\pm}_{2,2})^{\otimes 3})}<\infty$$

then the fundamental theorem of calculus entails that

$$\left|\left\langle (\nabla^{(2)}\check{F})(x) - (\nabla^{(2)}\check{F})(x+g),\, h\otimes k\right\rangle_{\mathcal{I}_{1,2}^{\otimes 2}}\right| \\ \le \|\nabla^{(3)}F\|_{L^\infty(W;(\mathcal{I}_{2,2}^{\pm})^{\otimes 3})}\|h\|_{L^2}\,\|k\|_{L^2}\;\|g\|_{L^2},$$

so that $\|\nabla^{(3)}F\|^{-1}_{L^\infty(W;(\mathcal{I}_{2,2}^{\pm})^{\otimes 3})}\check{F}$ belongs to $\Sigma_{\eta,p}$.

Section 4 is devoted to prove our main theorem:

Theorem 3.1. *Let $(\eta,p)\in\Lambda$ and $p\ge 3$. If X_a belongs to L^p, then*

$$\sup_{F\in\Sigma_{\eta,p}} \mathbf{E}\left[F(\check{\mathrm{S}}_2(B_V))\right] - \mathbf{E}\left[F(\check{\mathrm{S}}_2(X^{\mathcal{A}}))\right] \le c\,\|X_a\|^3_{L^p}\, m^{-(1/2-\eta)}. \tag{10}$$

Remark 5. Note that the integrability of the X_a's does enlarge the spaces in which the convergence holds, as in the Lamperti Theorem, but it does not modify the rate of convergence.

We now detail the proofs of the main estimates.

4. Stein method

We have to estimate

$$\sup_{F\in\Sigma_{\eta,p}} \mathbf{E}\left[F(\check{\mathrm{S}}_2(B_V))\right] - \mathbf{E}\left[F(\check{\mathrm{S}}_2(X^{\mathcal{A}}))\right].$$

Recall that

$$X^{\mathcal{A}} = \sum_{a\in\mathcal{A}} X_a h_a^m,\ B_V = \sum_{a\in\mathcal{A}} \delta h_a^m\ h_a^m \text{ and } V = \operatorname{span}\{h_a^m,\, a\in\mathcal{A}\}.$$

Let W (respectively H) be the space V equipped with the norm of $W_{\eta,p}(\mathbf{R}^d)$ (respectively of $\mathcal{I}_{1,2}(\mathbf{R}^d)$). Since V is finite dimensional, the difference between W and H is tenuous but still of some importance to clarify the situation. We have the following situation

$$\begin{array}{ccc} W^*_{\eta,p}\subset W^* \xrightarrow{\ \mathrm{i}^*_{\eta,p}\ } & H\subset\mathcal{I}_{1,2} & \\ & \big\downarrow \iota_{\eta,p} & \\ & W\subset W_{\eta,p} \xrightarrow{\ \check{\mathrm{S}}_2\ } & \check{\mathrm{G}}_2 W_{\eta,p} \xrightarrow{\ F\ } \mathbf{R}. \end{array}$$

For the sake of notations, we set $\check{F} = F\circ\check{\mathrm{S}}_2$.The Stein-Dirichlet representation formula (see [7]) then stands that

$$\mathbf{E}\left[\check{F}(B_V)\right] - \mathbf{E}\left[\check{F}(X^{\mathcal{A}})\right] = \mathbf{E}\left[\int_0^\infty LP_t\check{F}(X^{\mathcal{A}})\,\mathrm{d}t\right]. \tag{11}$$

4.1. **Technical lemmas.** *In what follows, c is a constant which may vary from line to line. We denote $\mathbf{P}^V_{\eta,p}$ the restriction of $\mathbf{P}_{\eta,p}$ to V: It is the distribution of B_V.*

With the notations of Proposition 1, we have:

Lemma 4.1. *For any $F \in \operatorname{Lip}(d_{W_{\eta,p}})$, for any $t > 0$,*

$$\begin{aligned}
&\mathbf{E}\left[LP_tF(X^{\mathcal{A}})\right]\\
&= \sum_{a\in\mathcal{A}} \int_0^1 \mathbf{E}\left[X_a^2 \left\langle \nabla^{(2)}P_tF(X^{\mathcal{A}}_{\neg a}) - \nabla^{(2)}P_tF(X^{\mathcal{A}}_{\neg a} + rX_a h_a^m),\right.\right.\\
&\qquad\qquad\qquad \left.\left. h_a^m \otimes h_a^m \right\rangle_{\mathcal{I}_{1,2}}\right] \, dr\\
&\qquad + \sum_{a\in\mathcal{A}} \mathbf{E}\left[\left\langle \nabla^{(2)}P_tF(X^{\mathcal{A}}) - \nabla^{(2)}P_tF(X^{\mathcal{A}}_{\neg a}), h_a^m \otimes h_a^m \right\rangle_{\mathcal{I}_{1,2}}\right].
\end{aligned}$$

Proof. Recall that

$$\begin{aligned}
LP_tF(X^{\mathcal{A}}) = &-\left\langle X^{\mathcal{A}}, \nabla P_tF(X^{\mathcal{A}})\right\rangle_{\mathcal{I}_{1,2}}\\
&+ \sum_{a\in\mathcal{A}} \left\langle h_a^m \otimes h_a^m,\, \nabla^{(2)}P_tF(X^{\mathcal{A}})\right\rangle_{\mathcal{I}_{1,2}\otimes\mathcal{I}_{1,2}}.
\end{aligned}$$

By independence,

$$\begin{aligned}
&\mathbf{E}\left[\left\langle X^{\mathcal{A}}, \nabla P_tF(X^{\mathcal{A}})\right\rangle_{\mathcal{I}_{1,2}}\right]\\
&\qquad = \sum_{a\in\mathcal{A}} \mathbf{E}\left[X_a \left\langle h_a^m,\, \nabla P_tF(X^{\mathcal{A}}) - \nabla P_tF(X^{\mathcal{A}}_{\neg a})\right\rangle_{\mathcal{I}_{1,2}}\right].
\end{aligned}$$

The fundamental theorem of calculus now states that

$$\begin{aligned}
&\mathbf{E}\left[\left\langle X^{\mathcal{A}}, \nabla P_tF(X^{\mathcal{A}})\right\rangle_{\mathcal{I}_{1,2}}\right]\\
&\qquad = \sum_{a\in\mathcal{A}} \int_0^1 \mathbf{E}\left[X_a^2 \left\langle h_a^m \otimes h_a^m,\, \nabla^{(2)}P_tF(X^{\mathcal{A}}_{\neg a} + rX_a h_a^m)\right\rangle_{\mathcal{I}_{1,2}}\right] \mathrm{d}r.
\end{aligned}$$

Since $\mathbf{E}\left[X_a^2\right] = 1$, the result follows by successive cancellations. □

Introduce for any $a \prec b \in \mathcal{A}$ and any $0 \le s < t \le 1$,

$$\begin{aligned}
&h_{a,b}^m(s,t)\\
&= \left\{ \int_s^t \left(h_{a_2}^m(r) - h_{a_2}^m(s)\right) \mathrm{d}h_{b_2}^m(r) - \int_s^t \left(h_{b_2}^m(r) - h_{b_2}^m(s)\right) \mathrm{d}h_{a_2}^m(r) \right\} [e_{a_1},\, e_{b_1}].
\end{aligned}$$

Lemma 4.2. *Let $(U_a,\, a \in \mathcal{A})$ be a family of independent identically distributed random variables which belong to L^p. Then,*

$$\mathbf{E}\left[\left| U_a \sum_{b\prec a} U_b\, h_{a,b}^m \right|_{W_{\eta,p}}^p\right] \le c\, m^{-(1/2-\eta)}\, \mathbf{E}\left[|U_a|^p\right]^2.$$

Proof. By independence and as in the proof of Theorem 2.3,

$$\mathbf{E}\left[\left| U_a \sum_{b\prec a} U_b\, h_{a,b}^m(s,t) \right|^p\right] \le c\, \mathbf{E}\left[|U_a|^p\right] \mathbf{E}\left[\left(\sum_{b\prec a} |U_b|^2 |h_{a,b}^m(s,t)|^2\right)^{p/2}\right].$$

Since a is fixed and $b \prec a$, if $|t-s| \le 1/m$, $h^m_{a,b}$ is not zero only for $b_2 = a_2 - 1$ and then, it is bounded by $m|t-s|^2$. If $|t-s| \ge 1/m$, $h^m_{a,b}$ is not zero for at most $[dm|t-s|]$ values of b and then, each term is bounded by $1/m$. Hence

$$\mathbf{E}\left[\left(\sum_{b\prec a} |U_b|^2 |h^m_{a,b}(s,t)|^2\right)^{p/2}\right] \le c\, m^{-p/2}(m|t-s|)^{p/2}\mathbf{E}\left[|U_b|^p\right].$$

The result follows by integration with respect to $\mu_{\eta,p}$. □

4.2. **Proof of the main theorem.** The result of Lemma 4.1 raises a problem which did not exist in finite dimension: There is no apparent $m^{-1/2}$ factor which gives the rate of convergence after applying a Taylor expansion of the convenient order. Said otherwise, there is no clue that the difference between $X^{\mathcal{A}}$ and $X^{\mathcal{A}}_{\neg a}$ should be small. Actually, the $m^{-1/2}$ factor is hidden in the h^m_a's whose L^∞ norm is exactly $m^{-1/2}$. But the scalar product we have introduced involves their $\mathcal{I}_{1,2}$ norm, which is 1. The necessary degree of freedom is given here by the possibility to consider h^m_a as an element of another functional space. We borrowed this idea from [20], our presentation being hopefully more straightforward.

Proof of Theorem 3.1. For $t \ge 0$, $r \in [0,1]$, $y \in W_{\eta,p}$, let

$$X^{\mathcal{A}}_{\neg a}(t,r,y) = e^{-t}(X^{\mathcal{A}}_{\neg a} + rX_a h^m_a) + \beta_t y. \tag{12}$$

Lemma 4.1 and Proposition 1 imply that

$$\begin{aligned}
&\mathbf{E}\left[LP_tF(X^{\mathcal{A}})\right]\\
&= e^{-2t}\sum_{a\in\mathcal{A}}\int_0^1 \mathbf{E}\left[X_a^2\int_{W_{\eta,p}}\Big\langle \nabla^{(2)}F(X^{\mathcal{A}}_{\neg a}(t,0,y))\right.\\
&\qquad\qquad\left. -\nabla^{(2)}F\big(X^{\mathcal{A}}_{\neg a}(t,r,y)\big),\, h^m_a\otimes h^m_a\Big\rangle_{\mathcal{I}_{1,2}^{\otimes 2}}\, \mathrm{d}\mathbf{P}^V_{\eta,p}(y)\right]\\
&= e^{-2t}\sum_{a\in\mathcal{A}}\int_0^1 \mathbf{E}\left[X_a^2\int_{W_{\eta,p}}\Big\langle \nabla^{(2)}F(X^{\mathcal{A}}_{\neg a}(t,1,y))\right.\\
&\qquad\qquad\left. -\nabla^{(2)}F\big(X^{\mathcal{A}}_{\neg a}(t,0,y)\big),\, h^m_a\otimes h^m_a\Big\rangle_{\mathcal{I}_{1,2}^{\otimes 2}}\, \mathrm{d}\mathbf{P}^V_{\eta,p}(y)\right]
\end{aligned}$$

Since $\|h^m_a\|_{L^2} = m^{-1/2}$, for $F \in \Sigma_{\eta,p}$,

$$\begin{aligned}
&\left|\mathbf{E}\left[LP_t\check{F}(X^{\mathcal{A}})\right]\right|\\
&\quad\le c\,e^{-2t}\sup_{\substack{r\in[0,1]\\ y,z\in W_{\eta,p}}}\mathbf{E}\left[(1+|X_a|^2)\left\|X^{\mathcal{A}}_{\neg a}(t,0,y,z) - X^{\mathcal{A}}_{\neg a}(t,r,y)\right\|_{\check{\mathrm{G}}_2W_{\eta,p}}\right],
\end{aligned} \tag{13}$$

where a is any chosen index of $\mathcal{A}$. By the definition of the norm on $\check{\mathrm{G}}_2 W_{\eta,p}$, we have

$$\begin{aligned}
&\sup_{\substack{r\in[0,1]\\ y,z\in W_{\eta,p}}} \mathbf{E}\left[(1+|X_a|^2)\left\|X^{\mathcal{A}}_{\neg a}(t,0,y,z)-X^{\mathcal{A}}_{\neg a}(t,r,y)\right\|_{\check{\mathrm{G}}_2 W_{\eta,p}}\right]\\
&= \sup_{\substack{r\in[0,1]\\ y,z\in W_{\eta,p}}} \mathbf{E}\left[(1+|X_a|^2)\left\|X^{\mathcal{A}}_{\neg a}(t,0,y,z)-X^{\mathcal{A}}_{\neg a}(t,r,y)\right\|_{W_{\eta,p}}\right]\\
&+ \sup_{\substack{r\in[0,1]\\ y,z\in W_{\eta,p}}} \mathbf{E}\Bigg[(1+|X_a|^2)\bigg(\iint\Big(\pi_2\check{\mathrm{S}}_2(X^{\mathcal{A}}_{\neg a}(t,r,y))\\
&\qquad\qquad -\pi_2\check{\mathrm{S}}_2(X^{\mathcal{A}}_{\neg a}(t,0,y))\Big)_{u,v}^{p/2}\,\mathrm{d}\mu_{\eta,p}(u,v)\bigg)^{1/p}\Bigg]\\
&\qquad\qquad = A_1+A_2.
\end{aligned}$$

According to (1) and (12),

$$|A_1|\le e^{-t}\mathbf{E}\left[(1+|X_a|^2)|X_a|\right]\|h^m_a\|_{W_{\eta,p}}\le 2\,\mathbf{E}\left[|X_a|^3\right]\,m^{-(1/2-\eta)}. \tag{14}$$

Furthermore,

$$\begin{aligned}
&\Big(\pi_2\check{\mathrm{S}}_2(X^{\mathcal{A}}_{\neg a}(t,r,y))-\pi_2\check{\mathrm{S}}_2(X^{\mathcal{A}}_{\neg a}(t,0,y))\Big)_{u,v}\\
&\qquad\qquad = r\,X_a(t,y)\sum_{b\prec a}X_b(t,y)\,h^m_{a,b}(u,v),
\end{aligned}$$

where

$$X_a(t,y)=e^{-t}X_a+\beta_t y.$$

Hence, according to Hölder inequality,

$$\begin{aligned}
&|A_2|\le c\,\mathbf{E}\left[|X_a|^p\right]^{2/p}\\
&\times\mathbf{E}\left[\left(\iint\Big|X_a(t,y)\sum_{b\prec a}X_b(t,y)\,h^m_{a,b}(u,v)\Big|^{p/2}\,\mathrm{d}\mu_{\eta,p}(u,v)\right)^{1/(p-2)}\right]^{(p-2)/p}\\
&\le c\,\mathbf{E}\left[|X_a|^p\right]^{2/p}\\
&\qquad\times\mathbf{E}\left[\iint\Big|X_a(t,y)\sum_{b\prec a}X_b(t,y)\,h^m_{a,b}(u,v)\Big|^{p/2}\,\mathrm{d}\mu_{\eta,p}(u,v)\right]^{1/p}.
\end{aligned}$$

Lemma 4.2 implies that there exists $c>0$ such that for any $y\in W_{\eta,p}$,

$$\begin{aligned}
|A_2|&\le c\,\mathbf{E}\left[|X_a|^p\right]^{2/p}\mathbf{E}\left[|X_a|^{p/2}\right]^{2/p}\,m^{-(1/2-\eta)}\\
&\le c\,\mathbf{E}\left[|X_a|^p\right]^{3/p}\,m^{-(1/2-\eta)}.
\end{aligned} \tag{15}$$

Since $|\mathcal{A}|=d.m$, plug (14) and (15) into (13) to obtain the existence of $c>0$ such that for any $t>0$,

$$\Big|\mathbf{E}\left[LP_tF(X^{\mathcal{A}})\right]\Big|\le c\,e^{-2t}\mathbf{E}\left[|X_a|^p\right]^{3/p}\,m^{-(1/2-\eta)}. \tag{16}$$

In view of (11), by integration over $\mathbf{R}^+$, we get (10) and the proof is complete. □

References

[1] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, second ed., Pure and Applied Mathematics (Amsterdam), vol. 140, Elsevier/Academic Press, Amsterdam, 2003.

[2] A. D. Barbour, *Stein's method for diffusion approximations*, Probability Theory and Related Fields **84** (1990), no. 3, 297–322.

[3] A. D. Barbour and L. H. Y. Chen, *An introduction toStein's method*, Lecture Notes Series, vol. 4, National University of Singapore, 2005.

[4] H. Brézis, *Analyse fonctionnelle*, Masson ed., 1987.

[5] L. Coutin and L. Decreusefond, *Stein's method for Brownian approximations*, Communications on Stochastic Analysis **7** (2013), no. 3, 349–372.

[6] L. Decreusefond, *Stochastic calculus with respect to Volterra processes*, Annales de l'Institut Henri Poincaré (B) Probability and Statistics **41** (2005), 123–149.

[7] L. Decreusefond, *The Stein-Dirichlet-Malliavin method*, ESAIM: Proceedings (2015), 11.

[8] M. D. Donsker, *An invariance principle for certain probability limit theorems*, Mem. Amer. Math. Soc. **6** (1951).

[9] R. M. Dudley, *Real analysis and probability*, Cambridge Studies in Advanced Mathematics, vol. 74, Cambridge University Press, Cambridge, 2002.

[10] D. Feyel and A. De La Pradelle, *Fractional integrals and Brownian processes*, Comm. Pure Appl. Math. **51** (1998), no. 1, 23–45.

[11] D. Feyel and A. de La Pradelle, *On fractional Brownian processes*, Potential Anal. **10** (1999), no. 3, 273–288.

[12] P. Friz and N. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010.

[13] P. Friz and N. Victoir, *A variation embedding theorem and applications*, Journal of Functional Analysis **239** (2006), no. 2, 631637.

[14] J. Lamperti, *On convergence of stochastic processes*, Transactions of the American Mathematical Society **104** (1962), 430–435.

[15] I. Nourdin and G. Peccati, *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality*, Cambridge University Press, 2012.

[16] D. Nualart, *The Malliavin Calculus and Related Topics*, vol. 17, SpringerVerlag, 1995.

[17] A. Pietsch, *Operator ideals*, North-Holland Mathematical Library, vol. 20, North-Holland Publishing Co., Amsterdam-New York, 1980.

[18] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes, and martingales. Vol. 2*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2000.

[19] L. Schwartz (ed.), *Séminaire Laurent Schwartz 1969–1970: Applications radonifiantes*, Centre de Mathématiques, École Polytechnique, Paris, 1970.

[20] H.-H. Shih, *On Steins method for infinite-dimensional Gaussian approximation in abstract Wiener spaces*, Journal of Functional Analysis **261** (2011), no. 5, 1236–1283.

[21] A. S. Üstünel, *Analysis on Wiener Space and Applications*, arXiv:1003.1649 **12** (2010), no. 1, 85–90.

[22] C. Villani, *Topics in optimal transportation*, vol. 58, Graduate Studies in Mathematics, no. 2, American Mathematical Society, Providence, RI, 2003.

Institut Mathématique de Toulouse, Université P. Sabatier, Toulouse, France
E-mail address: coutin@math.univ-toulouse.fr

LTCI, Telecom ParisTech, Université Paris-Saclay, 75013, Paris, France
E-mail address: laurent.decreusefond@telecom-paristech.fr