

Donsker's theorem in Wasserstein-1 distance

Laure Coutin, Laurent Decreusefond

▶ **To cite this version:**

Laure Coutin, Laurent Decreusefond. Donsker's theorem in Wasserstein-1 distance. Electronic Communications in Probability, 2020, 25, pp.1–13. hal-02098892

HAL Id: hal-02098892

https://telecom-paris.hal.science/hal-02098892

Submitted on 13 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DONSKER'S THEOREM IN WASSERSTEIN-1 DISTANCE

L. COUTIN AND L. DECREUSEFOND

Abstract. We compute the Wassertein-1 (or Kolmogorov-Rubinstein) distance between a random walk in $\mathbf{R}^d$ and the Brownian motion. The proof is based on a new estimate of the Lipschitz modulus of the solution of the Stein's equation. As an application, we can evaluate the rate of convergence towards the local time at 0 of the Brownian motion.

1. Motivations

For a complete, separable metric space X, the topology of convergence in distribution is metrizable [8] by considering the so-called Kolmogorov-Rubinstein or Wasserstein-1 distance:

$$\mathrm{dist}_{\mathrm{KR}}(\mu,\nu)=\sup_{F\in\mathrm{Lip}_1(X)}\left(\int_X F\ \mathrm{d}\mu-\int_X F\ \mathrm{d}\nu\right) \tag{1}$$

where

$$\mathrm{Lip}_1(X)=\{F\,:\,X\to\mathbf{R},|F(x)-F(y)|\le \mathrm{dist}_X(x,y),\ \forall x,y\in X\}\,.$$

The formulation (1) is well suited to evaluate distance by the Stein's method. When $X=\mathbf{R}$, there is no particular difficulty to evaluate the K-R distance when μ is the Gaussian distribution. When, $X=\mathbf{R}^d$, it is only recently (see [9, 12, 15] and references therein) that some improvement of the standard Stein's method has been proposed to get the K-R distance to the Gaussian measure on $\mathbf{R}^d$. The bottleneck is the estimate of the Lipschitz modulus of the second order derivative of the solution of the Stein's equation when F is only assumed to be Lipschitz continuous. Namely, for $f:\mathbf{R}^d\to\mathbf{R}$, for any $t>0$, consider the function

$$P_tf\,:\,x\in\mathbf{R}^d\longmapsto\int_{\mathbf{R}^d}f(e^{-t}x+\sqrt{1-e^{-2t}}y)\ \mathrm{d}\mu_d(y)$$

where μ_d is the standard Gaussian measure on $\mathbf{R}^d$. In dimension 1, the Stein's equation reads as

$$-xh(x)+h'(x)=f(x)-\int_{\mathbf{R}}f\ \mathrm{d}\mu_1,$$

so that

$$h(x)=\int_0^\infty P_tf(x)\ \mathrm{d}t \tag{2}$$

and the subsequent computations require to evaluate only the Lipschitz modulus of h'. For $f\in L^1(\mu)$, it is classical to see that P_tf is infinitely differentiable and that

$$(P_tf)^{(k)}(x)=\left(\frac{e^{-t}}{\sqrt{1-e^{-2t}}}\right)^k\int_{\mathbf{R}^d}f(e^{-t}x+\sqrt{1-e^{-2t}}y)H_k(y)\ \mathrm{d}\mu_d(y) \tag{3}$$

1991 *Mathematics Subject Classification.* 60F15,60H07,60G15,60G55.

Key words and phrases. Donsker theorem, Malliavin calculus, Stein's method, Wasserstein distance.

The first author is partially supported by ANR MESA.

where H_k is the k-th Hermite polynomial. On the other hand, if f is k-times differentiable, we have

$$(P_t f)^{(k)} = e^{-kt} P_t(f^{(k)}). \tag{4}$$

According to (3), we get

$$h'(x) = \int_0^\infty \frac{e^{-t}}{\sqrt{1-e^{-2t}}} \int_{\mathbf{R}^d} f(e^{-t}x + \sqrt{1-e^{-2t}}y) y \, \mathrm{d}\mu_1(y) \, \mathrm{d}t.$$

It is apparent that the Lipschitz modulus of h' simply depends on the Lipschitz modulus of f. However, in higher dimension, the Stein's equation becomes

$$-x.\nabla h(x) + \Delta h(x) = f(x) - \int_{\mathbf{R}^d} f \, \mathrm{d}\mu_d, \tag{5}$$

whose solution is formally given by (2). The form of (5) entails that we need to estimate the Lipschitz modulus of Δh, which requires to use (3) for $k = 2$. Unfortunately, we have to realize that

$$\left(\frac{e^{-t}}{\sqrt{1-e^{-2t}}}\right)^k \notin L^1([0,+\infty);\ \mathrm{d}t).$$

Hence, until the very recent papers [9, 15], the strategy was to assume that ∇f is Lipschitz, apply once (4) to compute the first derivative of $P_t f$ and then apply (3) to this expression:

$$\Delta h(x) = \int_0^\infty \frac{e^{-t}}{\sqrt{1-e^{-2t}}} \int_{\mathbf{R}^d} \nabla f(e^{-t}x + \sqrt{1-e^{-2t}}y).y \, \mathrm{d}\mu_d(y) \, \mathrm{d}t.$$

This means that instead of computing the supremum in the right-hand-side of (1), over Lipschitz functions, it is computed over functions whose first derivative is Lipschitz. This also defines a distance, which does not change the induced topology but the accuracy of the bound is degraded.

In infinite dimension, a new problem arises which is best explained by going back to the roots of the Stein's method in dimension 1. Consider that we want to estimate the K-R distance in the standard Central Limit Theorem. Let $(X_n,\, n \geq 1)$ be a sequence of independent, identically distributed random variables with $\mathbf{E}[X] = 0$ and $\mathbf{E}\left[X^2\right] = 1$. Let $T_n = n^{-1/2} \sum_{j=1}^n X_j$. The Stein-Dirichlet representation formula [6] states that

$$\mathbf{E}[f(T_n)] - \int_{\mathbf{R}} f \, \mathrm{d}\mu_1 = \mathbf{E}\left[\int_0^\infty L P_t f(T_n) \, \mathrm{d}t\right] \tag{6}$$

where

$$Lf(x) = -xf(x) + f'(x) = L_1 f(x) + L_2 f(x),$$

with obvious notations. Now,

$$L_1 P_t f(T_n) = -T_n (P_t f)'(T_n) = -\frac{1}{\sqrt{n}} \sum_{j=1}^n X_j (P_t f)'(T_n).$$

The trick, which amounts to an integration by parts for a Malliavin structure on independent random variables (see [7]), is to write

$$\mathbf{E}[X_j (P_t f)'(T_n)] = \mathbf{E}\left[X_j \Big((P_t f)'(T_n) - (P_t f)'(T_n - X_j/\sqrt{n}) \Big)\right]$$

in view of the independence of the random variables. Then, we use the fundamental theorem of calculus in this expression around the point $T_n^{\neg j} = T_n - X_j/\sqrt{n}$:

$$\mathbf{E}\left[X_j\Big((P_tf)'(T_n) - (P_tf)'(T_n - X_j/\sqrt{n})\Big)\right] = \frac{1}{\sqrt{n}}\int_0^1 \mathbf{E}\left[X_j^2(P_tf)''(T_n + rX_j/\sqrt{n})\Big)\right]\,\mathrm{d}r.$$

Since,

$$\int_0^1 \mathbf{E}\left[X_j^2(P_tf)''(T_n^{\neg j})\right]\,\mathrm{d}r = \mathbf{E}\left[(P_tf)''(T_n^{\neg j})\right],$$

we get

$$\begin{aligned}(7)\quad LP_tf(T_n) &= -\frac{1}{n}\sum_{j=1}^n \int_0^1 \mathbf{E}\left[X_j^2\Big((P_tf)''(T_n^{\neg j} + rX_j/\sqrt{n}) - (P_tf)''(T_n^{\neg j})\Big)\right]\,\mathrm{d}r \\ &\quad + \frac{1}{n}\sum_{j=1}^n \mathbf{E}\left[(P_tf)''(T_n^{\neg j}) - (P_tf)''(T_n)\right].\end{aligned}$$

This formula confirms that the crux of the matter is now to estimate uniformly the Lipschitz modulus of $(P_tf)''$. It also shows how we get the order of convergence. We have one occurrence of $n^{-1/2}$ in the definition of T_n, which appears in the expression of L_1. The same factor appears a second time when we proceed to the Taylor expansion and then, it will appear a third time when we plug (3) into (7). This means that we have a factor $n^{-3/2}$ which is summed up n times, hence the rate of convergence which is known to be $n^{-1/2}$.

Now, if we are interested in the Donsker theorem, the process whose limit we would like to assess is

$$S_n(t) = \sum_{j=1}^n X_j h_j^n(t)$$

where

$$h_j^n(t) = \sqrt{n}\int_0^t \mathbf{1}_{[j/n,(j+1)/n)}(s)\,\mathrm{d}s.$$

For reasons that will be explained below, the analog of the second order derivatives will involve

$$(8)\qquad \left\langle h_j^{\otimes 2}, \nabla^{(2)}(P_tf)(S_n^{\neg j}S_n + rX_j/\sqrt{n}) - \nabla^{(2)}(P_tf)(S_n^{\neg j})\right\rangle_{I_{1,2}^{\otimes 2}}$$

where ∇ is the Malliavin derivative, $I_{1,2}$ is the Cameron-Martin space

$$I_{1,2} = \left\{f, \exists!\dot{f} \in L^2([0,1],\,\mathrm{d}t) \text{ with } f(t) = \int_0^t \dot{f}(s)\,\mathrm{d}s\right\}$$

and

$$\|f\|_{I_{1,2}} = \|\dot{f}\|_{L^2}.$$

Recall that in the context of Malliavin calculus, this space is identified to its dual which means that the dual of L^2 is not itself. The difficulty is then that we do not have a $n^{-1/2}$ factor in the definition of S_n and it is easily seen that $\|h_j^n\|_{I_{1,2}} = 1$, hence no multiplicative factor will pop up in (8). In [4], we bypassed this difficulty by assuming enough regularity of f so that $\nabla^{(2)}P_tf$ belong to the dual of L^2. Then, in the estimate of terms as those appearing in (8), it is the L^2-norm of h_j^n which appears and it turns out that $\|h_j^n\|_{L^2} \le c\,n^{-1/2}$, hence the presence of a factor n^{-1}, which saves the proof.

The goal of this paper is to weaken the hypothesis on f to be able to upper-bound the true K-R distance between the distribution of S_n and the distribution of a Brownian motion, that is

$$\sup_{f\in \mathrm{Lip}_1(X)} \mathbf{E}\left[f(S_n)\right] - \mathbf{E}\left[f(B)\right].$$

The space X is a Banach space we can choose arbitrarily as far as it can be equipped with the structure of an abstract Wiener space and it contains the sample paths of S_n and B.

The main technical result of this article is Theorem 4.4 which gives a new estimate of the Lipschitz modulus of $\nabla^{(2)}P_t f$ for $t>0$. The main idea is to introduce a hierarchy of approximations. There is a first scale induced by the time discretization coming from the definition of S_n. Then, we consider a coarser discretization onto which we project our approximations in order to benefit from the averaging effect of the ordinary CLT. It turns out that the optimal ratio is obtained when the mesh of the coarser subdivision is roughly the cubic root of the mesh of the reference partition. Moreover, after [3] and [4], we are convinced that it is simpler and as efficient to stick to finite dimension as long as possible. For, we consider the affine interpolation of the Brownian motion as an intermediary process. The distance between the Brownian sample-paths and their affine interpolation is well known. This reduces the problem to estimate the distance between S_n and the affine interpolation of B, a task which can be handled by the Stein's method. It turns out that the bottleneck is in fact the rate of convergence of the Brownian interpolation to the Brownian motion.

This paper is organized as follows. In Section 2, we show how to view fractional Sobolev spaces as Wiener spaces. In Section 3, we explain the line of thoughts we used. The proofs are given in Section 4.

2. Preliminaries

2.1. **Fractional Sobolev spaces.** As in [5, 11], 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

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

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

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

They are known to be Banach spaces and to satisfy the Sobolev embeddings [1, 10]:

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

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 [2]), 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

$$\|h_{s_1,s_2}\|_{\eta,p} \leq c\,|s_2-s_1|^{1/2-\eta}. \tag{9}$$

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 denote by $W_{0,\infty}$ the space of continuous (hence bounded) functions on $[0,1]$ equipped with the uniform norm.

2.2. **Fractional spaces $W_{\eta,p}$ as Wiener spaces.** Let

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

In what follows, we always choose η and p in Λ. Consider $(Z_n,\ n \geq 1)$ a sequence of independent, standard Gaussian random variables and let $(z_n,\ n \geq 1)$ be a complete orthonormal basis of $I_{1,2}$. Then, we know from [13] that

$$\sum_{n=1}^{N} Z_n\, z_n \xrightarrow{N\to\infty} B := \sum_{n=1}^{\infty} Z_n\, z_n \text{ in } W_{\eta,p} \text{ with probability 1}, \tag{10}$$

where B is a Brownian motion. We clearly have the diagram

$$W^*_{\eta,p} \xrightarrow{\mathfrak{e}^*_{\eta,p}} (I_{1,2})^* \simeq I_{1,2} \xrightarrow{\mathfrak{e}_{\eta,p}} W_{\eta,p}, \tag{11}$$

where $\mathfrak{e}_{\eta,p}$ is the embedding from $I_{1,2}$ into $W_{\eta,p}$. The space $I_{1,2}$ is dense in $W_{\eta,p}$ since polynomials do belong to $I_{1,2}$. Moreover, Eqn. (10) and the Parseval identity entail that for any $z \in W^*$,

$$\mathbf{E}\left[e^{i\langle z, B\rangle_{W^*_{\eta,p}, W_{\eta,p}}}\right] = \exp\left(-\frac{1}{2}\|\mathfrak{e}^*_{\eta,p}(z)\|^2_{I_{1,2}}\right). \tag{12}$$

We denote by $\mu_{\eta,p}$ the law of B on $W_{\eta,p}$. Then, the diagram (11) and the identity (12) mean that $(I_{1,2}, W_{\eta,p}, \mu_{\eta,p})$ is a Wiener space.

Definition 2.1 (Wiener integral)**.** The Wiener integral, denoted as $\delta_{\eta,p}$, is the isometric extension of the map

$$\begin{aligned} \delta_{\eta,p} \,:\, \mathfrak{e}^*_{\eta,p}(W^*_{\eta,p}) \subset I_{1,2} &\longrightarrow L^2(\mu_{\eta,p}) \\ \mathfrak{e}^*_{\eta,p}(\eta) &\longmapsto \langle \eta,\, y\rangle_{W^*_{\eta,p}, W_{\eta,p}}. \end{aligned}$$

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

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

Definition 2.2 (Ornstein-Uhlenbeck semi-group)**.** For any Lipschitz function on $W_{\eta,p}$, for any $\tau \geq 0$,

$$P_\tau f(x) = \int_{W_{\eta,p}} f(e^{-\tau}x + \beta_\tau y)\, \mathrm{d}\mu_{\eta,p}(y)$$

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

The dominated convergence theorem entails that P_τ is ergodic: For any $x \in W_{\eta,p}$, with probability 1,

$$P_\tau f(x) \xrightarrow{\tau\to\infty} \int_{W_{\eta,p}} f\, \mathrm{d}\mu_{\eta,p}.$$

Moreover, the invariance by rotation of Gaussian measures implies that

$$\int_{W_{\eta,p}} P_\tau f(x)\, \mathrm{d}\mu_{\eta,p}(x) = \int_{W_{\eta,p}} f\, \mathrm{d}\mu_{\eta,p}, \text{ for any } \tau \geq 0.$$

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 [14, 17].

Definition 2.3. 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

$$f(y) = \sum_{j=1}^{k} f_j(\delta_{\eta,p} h_1(y), \cdots, \delta_{\eta,p} h_k(y))\, x_j$$

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 $I_{1,2}$ and $(x_1, \cdots, x_k)$ belong to X. The set of such functions is denoted by $\mathfrak{C}(X)$.

For $h \in I_{1,2}$,

$$\langle \nabla f,\, h\rangle_{I_{1,2}} = \sum_{j=1}^{k}\sum_{l=1}^{k} \partial_l f(\delta_{\eta,p} h_1(y), \cdots, \delta_{\eta,p} h_k(y))\ \langle h_l,\, h\rangle_{I_{1,2}}\ \ x_j,$$

which is equivalent to say

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

It is proved in [16, Theorem 4.8] that

Theorem 2.2. *For $f \in \mathrm{Lip}_1(W_{\eta,p})$, for any $t > 0$, for any $x \in W_{\eta,p}$*

$$\frac{d}{dt} P_t f(x) = -\langle x,\, \nabla(P_t f)(x)\rangle_{W_{\eta,p}, W^*_{\eta,p}} + \sum_{i=1}^{\infty} \left\langle \nabla^{(2)} P_t f(x),\, h_i \otimes h_i \right\rangle_{I_{1,2}} \tag{13}$$

where $(h_n,\, n \geq 1)$ is complete orthonormal basis of H.

Note that a non trivial part of this theorem is to prove that the terms are meaningful: that $\nabla P_t f$ has values in $W^*_{\eta,p}$ instead of $I_{1,2}$ and that $\nabla^{(2)} P_t f(x)$ is trace-class. Actually, we only need a finite dimensional version of this identity in which all these difficulties do not appear.

3. Donsker's theorem in $W_{\eta,p}$

For $m \geq 1$, let $\mathcal{D}^m = \{i/m,\, i = 0, \cdots, m\}$, the regular subdivision of the interval $[0, 1]$. Let

$$\mathcal{A}^m = \{1, \cdots, d\} \times \{0, \cdots, m-1\}$$

and for $a = (a_1,\, a_2) \in \mathcal{A}^m$

$$h^m_a(t) = \sqrt{m} \int_0^t \mathbf{1}_{[a_2, a_2 + 1/m)}(s)\ \mathrm{d}s\ e_{a_1}.$$

Consider

$$S^m = \sum_{a \in \mathcal{A}^m} X_a\, h^m_a$$

where $(X_a,\, a \in \mathcal{A}^m)$ is a family of independent identically distributed, $\mathbf{R}^d$-valued, random variables. We denote by X a random variable which has their common distribution. Moreover, we assume that $\mathbf{E}\,[X] = 0$ and $\mathbf{E}\left[\|X\|^2_{\mathbf{R}^d}\right] = 1$. Remark that $(h^m_a, a \in \mathcal{A}^m)$ is an orthonormal family in $\mathbf{R}^d \otimes I_{1,2} := I^d_{1,2}$. Let

$$\mathcal{V}^m = \mathrm{span}(h^m_a,\, a \in \mathcal{A}^m) \subset I^d_{1,2}.$$

For any $m > 0$, the map π^m is the orthogonal projection from $H := I^d_{1,2}$ onto $\mathcal{V}^m$. Let $0 < N < m$, for $f \in \mathrm{Lip}_1(W_{\eta,p})$, we write

$$\mathbf{E}\,[f(S^m)] - \mathbf{E}\,[f(B)] = \sum_{i=1}^{3} A_i \tag{14}$$

where

$$A_1 = \mathbf{E}\left[f(S^m)\right] - \mathbf{E}\left[f(\pi^N(S^m)\right]$$
$$A_2 = \mathbf{E}\left[f\circ\pi^N(S^m)\right] - \mathbf{E}\left[f\circ\pi^N(B^m)\right],$$
$$A_3 = \mathbf{E}\left[f\circ\pi^N(B^m)\right] - \mathbf{E}\left[f(B)\right],$$

where B^m is the affine interpolation of the Brownian motion:

$$B^m(t) = \sum_{a\in\mathcal{A}^m} \sqrt{m}\,\left(B_{a_1}(a_2+\frac{1}{m}) - B_{a_1}(a_2)\right)\,h_a^m(t).$$

The two terms A_1 and A_3 are of the same nature: We have to compare two processes which live on the same probability space. Since f is Lipschitz, we can proceed by comparison of their sample-paths. The term A_2 is different as the two processes involved live on different probability spaces. This is for this term that the Stein's method will be used.

We know from [11] that

Theorem 3.1. *For any $(\eta,p)\in\Lambda$, there exists $c>0$ such that*

$$\sup_N N^{1/2-\eta}\,\mathbf{E}\left[\|B^N-B\|_{\eta,p}^p\right]^{1/p} \le c. \tag{15}$$

Moreover, we have

Theorem 3.2. *Let $(\eta,p)\in\Lambda$. Assume that $X\in L^p(W;\mathbf{R}^d,\mu_{\eta,p})$. There exists a constant $c>0$ such that*

$$\sup_{m,N} N^{\frac{1}{2}-\eta}\,\mathbf{E}\left[\|S^m-\pi^N(S^m)\|_{\eta,p}^p\right]^{1/p} \le c\,\|X\|_{L^p}.$$

This upper-bound is far from being optimal and it is likely that it could be improved to obtain a factor $N^{1-\eta}$. However, in view of (15), it would bring no improvement to our final result.

Theorem 3.3. *Let $(\eta,p)\in\Lambda$. Let X_a belong to $L^p(W;\mathbf{R}^d,\mu_{\eta,p})$ for some $p\ge 3$. Then, there exists $c>0$ such that for any $f\in\mathrm{Lip}_1(W_{\eta,p})$,*

$$\mathbf{E}\left[f(\pi^N(S^m))\right] - \mathbf{E}\left[f(\pi^N(B^m))\right] \le c\,\|X\|_{L^p}\,\frac{N^{1+\eta}}{\sqrt{m}}\,\ln(\frac{N^{1+\eta}}{\sqrt{m}}). \tag{16}$$

The global upper-bound for (14) is proportional to

$$N^{-1/2+\eta} + N^{1+\eta}m^{-1/2}\ln(N^{1+\eta}m^{-1/2}).$$

See N as a function of m and note that this expression is minimal for $N\sim m^{1/3}$. Plug this into the previous expressions to obtain the main result of this paper:

Theorem 3.4. *Assume that $X\in L^p(W;\mathbf{R}^d,\mu_{\eta,p})$. Then, there exists a constant $c>0$ such that*

$$\sup_{f\in\mathrm{Lip}_1(W_{\eta,p})} \mathbf{E}\left[f(S^m)\right] - \mathbf{E}\left[f(B)\right] \le c\,\|X\|_{L^p(\Omega)}^p\; m^{-\frac{1}{6}+\frac{\eta}{3}}\;\ln(m). \tag{17}$$

As an application of the previous considerations, we obtain as a corollary an approximation theorem for the local time of the Brownian motion.

The reflected Brownian motion is defined as

$$R_t = B_t + \sup_{0\le s\le t}\max\left(0,-B_s\right)$$

and the reflected linear interpolation of random walk is

$$R_t^m = X_t^m + \sup_{0\le s\le t}\max\left(0,-X_s^m\right) := X_t^m + L_0^m(t).$$

The process $L_0(t) := \sup_{0\leq s\leq t} \max(0,-B_s)$ is an expression of the local time of the Brownian motion at 0. Note that the map $f \mapsto \left(t \mapsto f(t) + \sup_{0\leq s\leq t} \max(0,-f(s))\right)$ is Lipschitz continuous from any $W_{\eta,p}$ into $W_{0,\infty}$. One of the interest of our new result is that we can then apply the previous theorem in $W_{0,\infty}$ to L_0^m and L_0. We get

Corollary 3.5. *Assume that the hypothesis of Theorem 3.4 hold. There exists a constant $c > 0$ such that*

$$\sup_{f\in \mathrm{Lip}_1(W_{0,\infty})} \mathbf{E}\left[f(L_0^m)\right] - \mathbf{E}\left[f(L_0)\right] \leq c\|X\|_{L^3}\, m^{-\frac{1}{6}} \ln(m).$$

4. Proofs

In what follows, c denote a non significant constant which may vary from line to line. We borrow from the current usage in rough path theory the notation

$$f_{s,t} = f(t) - f(s). \tag{18}$$

As a preparation to the proof of Theorem 3.2, we need the following lemma.

Lemma 4.1. *For all $p \geq 2$, there exists a constant c_p such that for any sequence of independent, identically distributed random variables $(X_i, i \in \mathbf{N})$ with $X \in L^p$ and any sequence $(\alpha_i,\, i \in \mathbf{N})$.*

$$\mathbf{E}\left[\left|\sum_{i=1}^n \alpha_i X_i\right|^p\right] \leq c_p \left|\{i \leq n, \alpha_i \neq 0\}\right|^{p/2} \left(\sum_{i\leq n} |\alpha_i|^p\right)\, \mathbb{E}(|X|^p),$$

where $|A|$ is the cardinality of the set A.

Proof. The Burkholder-Davis Gundy inequality applied to the discrete martingale $(\sum_{i=1}^n \alpha_i X_i,\, n \geq 0)$ yields

$$\mathbf{E}\left[\left|\sum_{i=1}^n \alpha_i X_i\right|^p\right] \leq c_p \mathbf{E}\left[\left|\sum_{i=1}^n \alpha_i^2 X_i^2\right|^{p/2}\right].$$

Using Jensen inequality we obtain

$$\mathbf{E}\left[\left|\sum_{i=1}^n \alpha_i X_i\right|^p\right] \leq c_p \left|\{i \leq n, \alpha_i \neq 0\}\right|^{p/2-1}\, \mathbf{E}\left[\sum_{i=1}^n |\alpha_i|^p X_i^p\right].$$

The proof is thus complete. □

Proof of Theorem 3.2. Actually, we already proved in [4] that

$$\mathbf{E}\left[\|S^m_{s,t}\|^p\right] \leq c\|X\|_{L^p}\, \left(\sqrt{t-s} \wedge m^{-1/2}\right). \tag{19}$$

Assume that s and t belongs to the same sub-interval: There exists $l \in \{1, ..., N\}$ such that

$$\frac{l-1}{N} \leq s < t \leq \frac{l}{N}.$$

Then we have (see (18))

$$\pi^N(S^m)_{s,t} = \sqrt{N}\, \left(\sum_{k=1}^m X_k\, (h_k^m, h_l^N)_{I_{1,2}}\right)\, (t-s).$$

Using Lemma 4.1, there exists a constant c such that

$$\frac{\|\pi^N(S^m)_{s,t}\|_{L^p}}{\sqrt{N}\,|t-s|} \leq c\,\|X\|_{L^p}\, \left|\{k,\, (h_k^m, h_l^N)_{I_{1,2}} \neq 0\}\right|^{1/2} \sup_k \left|(h_k^m, h_l^N)_{I_{1,2}}\right|.$$

Note that $|(h_k^m, h_l^N)_{I_{1,2}}| \leq \sqrt{\frac{N}{m}}$ and there is at most $\frac{m}{N}+2$ terms such that $(h_k^m, h_N^l)_{I_{1,2}}$ is non zero. Thus,

$$\frac{\|\pi^N(S^m)_{s,t}\|_{L^p}}{\sqrt{N}\,|t-s|} \leq c\,\|X\|_{L^p}\,\left(\frac{m}{N}+2\right)^{1/2}\sqrt{\frac{N}{m}} \leq c\,\|X\|_{L^p},$$

as m/N tends to infinity. Since $|t-s| \leq 1/N$,

$$\|\pi^N(S^m)_{s,t}\|_{L^p} \leq c\,\|X\|_{L^p}\,\sqrt{|t-s|}. \tag{20}$$

For $0 \leq s \leq t \leq 1$ let $s_+^N := \min\{l,\, s \leq \frac{l}{N}\}$ and $t_-^N := \sup\{l,\, t \geq \frac{l}{N}\}$. We have

$$\begin{aligned}\pi^N(S^m)_{s,t} - S^m_{s,t} = &\left(\pi^N(S^m)_{s,s_+^N} - S^m_{s,s_+^N}\right)\\ &+ \left(\pi^N(S^m)_{s_+^N,t_-^N} - S^m_{s_+^N,t_-^N}\right) + \left(\pi^N(S^m)_{t_-^N,t} - S^m_{t_-^N,t}\right).\end{aligned}$$

Note that for all $f \in W_{\eta,p}$, $\pi^N(f)$ is the linear interpolation of f along the subdivision $\mathcal{D}_N$; hence, for $s,t \in \mathcal{D}_N$, $\pi^N(S^m)_{s,t} = S^m_{s,t}$. Thus the median term vanishes and we obtain

$$\begin{aligned}\mathbf{E}\left[\|\pi^N(S^m)_{s,t} - S^m_{s,t}\|^p\right] \leq c\Big(&\mathbf{E}\left[\|\pi^N(S^m)_{s,s_+^N}\|^p\right] + \mathbf{E}\left[\|S^m_{s,s_+^N}\|^p\right]\\ &+ \mathbf{E}\left[\|\pi^N(S^m)_{t_-^N,t}\|^p\right] + \mathbf{E}\left[\|S^m_{t_-^N,t}\|^p\right]\Big).\end{aligned} \tag{21}$$

From (20), we deduce that

$$\mathbf{E}\left[\|\pi^N(S^m)_{s,s_+^N}\|^p\right]^{1/p} \leq c\,\|X\|_{L^p}\,\sqrt{s_+^N - s} \leq c\,\|X\|_{L^p}\,N^{-1/2}, \tag{22}$$

and the same holds for $\mathbf{E}\left[\|\pi^N(S^m)_{t_-^N,t}\|^p\right]$. We infer from (19), (20) and (22) that

$$\mathbf{E}\left[\|\pi^N(S^m)_{s,t} - S^m_{s,t}\|^p\right]^{1/p} \leq c\,\|X\|_{L^p}\left(\sqrt{t-s} \wedge N^{-1/2}\right). \tag{23}$$

A straightforward computation shows that

$$\iint_{[0,1]^2} \frac{[|t-s| \wedge N^{-1}]^{p/2}}{|t-s|^{1+\eta p}}\,\mathrm{d}s\,\mathrm{d}t \leq c\,N^{-p(1/2-\eta)}. \tag{24}$$

The result follows (23) and (24). □

4.1. **Stein method.** We wish to estimate

$$\mathbf{E}\left[f(\pi^N(S^m))\right] - \mathbf{E}\left[f(\pi^N(B^m))\right],$$

using the Stein's method. For the sake of simplicity, we set

$$f_N = f \circ \pi^N.$$

The Stein-Dirichlet representation formula [6] stands that, for any $\tau_0 > 0$,

$$\begin{aligned}\mathbf{E}\left[f_N(B^m)\right] - \mathbf{E}\left[f_N(S^m)\right] &= \mathbf{E}\left[\int_0^\infty \frac{d}{dt} P_\tau f_N(S^m)\,\mathrm{d}\tau\right]\\ &= \mathbf{E}\left[P_{\tau_0} f_N(S^m) - f_N(S^m)\right] + \mathbf{E}\left[\int_{\tau_0}^\infty L P_\tau f_N(S^m)\,\mathrm{d}\tau\right],\end{aligned}$$

where

$$L P_\tau f_N(S^m) = -\left\langle S^m, \nabla P_\tau f_N(S^m)\right\rangle_H + \sum_{a \in \mathcal{A}^m}\left\langle \nabla^{(2)} P_\tau f_N(S^m),\, h_a^m \otimes h_a^m\right\rangle_{I_{1,2}{}^{\otimes 2}}.$$

It is straightforward (see [4, Lemma 4.1]):

Lemma 4.2. *For any $(\eta, p) \in \Lambda$, there exists a constant $c > 0$ such that for any sequence of independent, centered random vectors $(X_a, a \in \mathcal{A}^m)$ such that $\mathbf{E}[\|X\|^p] < \infty$, for any $f \in \mathrm{Lip}_1(W_{\eta,p})$, we have*

$$\mathbf{E}[f(S^m)] - \mathbf{E}[P_{\tau_0} f(S^m)] \leq c \, \|X\|_{L^p} \, \sqrt{1 - e^{\tau_0}}.$$

We now show, that as usual, the rate of convergence in the Stein's method is related to the Lipschitz modulus of the second order derivative of the solution of the Stein's equation. Namely, we have

Lemma 4.3. *For any $f \in \mathrm{Lip}_1(W_{\eta,p})$, we have*

$$\begin{aligned}
&\mathbf{E}[LP_\tau f_N(S^m)] \\
&= -\mathbf{E}\left[\sum_{a\in\mathcal{A}^m} \left\langle \nabla^{(2)} P_\tau f_N(S^m_{\neg a}) - \nabla^{(2)} P_\tau f_N(S^m),\, h^m_a \otimes h^m_a \right\rangle_{I_{1,2}{}^{\otimes 2}}\right] \\
&+ \mathbf{E}\left[\sum_{a\in\mathcal{A}} X_a^2 \int_0^1 \left\langle \nabla^{(2)} P_\tau f_N(S^m_{\neg a} + r X_a h^m_a) - \nabla^{(2)} P_\tau f_N(S^m_{\neg a}),\, h^{m\otimes 2}_a \right\rangle_{I_{1,2}{}^{\otimes 2}} dr\right].
\end{aligned}$$

Proof of Lemma 4.3. Let $S^m_{\neg a} = S^m - X_a h^m_a$. Since the X_a's are independent,

$$\begin{aligned}
&\mathbf{E}\left[\langle \nabla P_\tau f_N,\, S^m\rangle_{I_{1,2}}\right] \\
&= \mathbf{E}\left[\sum_{a\in\mathcal{A}^m} X_a \, \langle \nabla P_\tau f_N(S^m), h^m_a\rangle_{I_{1,2}}\right] \\
&= \mathbf{E}\left[\sum_{a\in\mathcal{A}^m} X_a \, \langle \nabla P_\tau f_N(S^m) - \nabla P_\tau f_N(S^m_{\neg a}),\, h^m_a\rangle_{I_{1,2}}\right] \\
&= \mathbf{E}\left[\sum_{a\in\mathcal{A}^m} X_a^2 \left\langle \nabla^{(2)} P_\tau f_N(S^m_{\neg a}),\, h^m_a \otimes h^m_a \right\rangle_{I_{1,2}{}^{\otimes 2}}\right] \\
&+ \mathbf{E}\left[\sum_{a\in\mathcal{A}} X_a^2 \int_0^1 \left\langle \nabla^{(2)} P_\tau f_N(S^m_{\neg a} + r\, X_a h^m_a) - \nabla^{(2)} P_\tau f_N(S^m_{\neg a}), h^{m\otimes 2}_a \right\rangle_{I_{1,2}{}^{\otimes 2}} \mathrm{d}r\right],
\end{aligned}$$

according to the Taylor formula. Since $\mathbf{E}\left[X_a^2\right] = 1$, we have

$$\begin{aligned}
&\mathbf{E}\left[\sum_{a\in\mathcal{A}^m} X_a^2 \left\langle \nabla^{(2)} P_\tau f_N(S^m_{\neg a}),\, h^m_a \otimes h^m_a \right\rangle_{I_{1,2}{}^{\otimes 2}}\right] \\
&\qquad = \mathbf{E}\left[\sum_{a\in\mathcal{A}^m} \left\langle \nabla^{(2)} P_\tau f_N(S^m_{\neg a}),\, h^m_a \otimes h^m_a \right\rangle_{I_{1,2}{}^{\otimes 2}}\right].
\end{aligned}$$

The result follows by difference. □

The main difficulty and then the main contribution of this paper is to find an estimate of

$$\sup_{v\in\mathcal{V}^m} \left\langle \nabla^{(2)} P_\tau f_N(v) - \nabla^{(2)} P_\tau f_N(v + \varepsilon h^m_a),\, h^m_a \otimes h^m_a \right\rangle_{I_{1,2}{}^{\otimes 2}},$$

for any ε.

Theorem 4.4. *There exists a constant c such that for any $\tau > 0$, for any $v \in \mathcal{V}^m$, for any $f \in \mathrm{Lip}(W_{\eta,p})$,*

$$\left|\left\langle \nabla^{(2)} P_\tau^m f_N(v + \varepsilon h_a^m) - \nabla^{(2)} P_\tau^m f_N(v),\, h_a^m \otimes h_a^m \right\rangle_{I_{1,2}{}^{\otimes 2}}\right| \leq c\, \frac{e^{-5\tau/2}}{\beta_{\tau/2}^2}\, \varepsilon N^{\eta-\frac{1}{2}} \sqrt{\frac{N^3}{m^3}}. \tag{25}$$

Proof of Theorem 4.4. We know from [16, 4] that we have the following representation: for any $h \in I_{1,2}$,

$$\left\langle \nabla^{(2)} P_\tau^m f(v),\, h \otimes h \right\rangle_{I_{1,2}{}^{\otimes 2}} = \frac{e^{-3\tau/2}}{\beta_{\tau/2}^2}\, \mathbf{E}\left[f\Big(w_\tau(v, B^m, \hat{B}^m) \Big) \delta_{\eta,p} h(B) \delta_{\eta,p} h(\hat{B}) \right] \tag{26}$$

where

$$w_\tau(v, y, z) = e^{-\tau/2}(e^{-\tau/2} v + \beta_{\tau/2} y) + \beta_{\tau/2} z$$

and $\hat{B}$ is an independent copy of B. Since the map v is linear with respect to its three arguments,

$$f_N\Big(w_\tau(v,\, B^m,\, \hat{B}^m) \Big) = f_N\Big(w_\tau(\pi^N v,\, \pi^N B^m,\, \pi^N \hat{B}^m) \Big).$$

Hence,

$$\begin{aligned} &\left(\frac{e^{-3\tau/2}}{\beta_{\tau/2}^2} \right)^{-1} \left\langle \nabla^{(2)} P_\tau^m f_N(v),\, h \otimes h \right\rangle_{I_{1,2}{}^{\otimes 2}} \\ &\quad = \mathbf{E}\left[f_N\Big(w_\tau(\pi^N v, \pi^N B, \pi^N \hat{B}) \Big) \mathbf{E}\left[\delta_{\eta,p} h(B) \,|\, \pi^N B^m \right] \mathbf{E}\left[\delta_{\eta,p} h(\hat{B}) \,|\, \pi^N \hat{B}^m \right] \right] \end{aligned} \tag{27}$$

From Lemma 4.7, we know that

$$\mathrm{Var}\left(\mathbf{E}\left[\delta_{\eta,p} h(\hat{B}) \,|\, \pi^N \hat{B}^m \right] \right) \leq c\, \frac{N}{m} \tag{28}$$

for $m > 8\, N$, and the same holds for the other conditional expectation. Use Cauchy-Schwarz inequality in (27) and take (28) into account to obtain

$$\begin{aligned} &\left(\frac{e^{-3\tau/2}}{\beta_{\tau/2}^2} \right)^{-1} \left| \left\langle \nabla^{(2)} P_\tau^m f_N(v + \varepsilon h_a^m) - \nabla^{(2)} P_\tau^m f_N(v),\, h_a^m \otimes h_a^m \right\rangle_{I_{1,2}{}^{\otimes 2}} \right| \\ &\leq c \left(\frac{N}{m} \right)^2 \left\| w_\tau(\pi^N v, \pi^N B^m, \pi^N \hat{B}^m) - w_\tau(\pi^N v + \varepsilon \pi^N h_a^m, \pi^N B^m, \pi^N \hat{B}^m) \right\|_{W_{\eta,p}} \\ &= c e^{-\tau} \varepsilon \left(\frac{N}{m} \right)^2 \left\| \pi^N h_a^m \right\|_{W_{\eta,p}} \end{aligned} \tag{29}$$

since f_N belongs to $\mathrm{Lip}_1(W_{\eta,p})$. Furthermore,

$$\pi^N(h_a^m) = \sum_{b \in \mathcal{A}^N} \left\langle h_a^m, h_b^N \right\rangle_{I_{1,2}} h_b^N.$$

We already know that

$$0 \leq \left\langle h_a^m, h_b^N \right\rangle_{I_{1,2}} \leq \sqrt{\frac{N}{m}}$$

and that at most two terms $\left\langle h_a^m, h_b^N \right\rangle_{I_{1,2}}$ are non zero. Moreover, according to Lemma 2.1

$$\|h_b^N\|_{W_{\eta,p}} \leq c\, N^{\eta-\frac{1}{2}}.$$

Thus,

$$\|\pi^N(h_a^m)\|_{W_{\eta,p}} \leq c\sqrt{\frac{N}{m}}\, N^{\eta-\frac{1}{2}}. \tag{30}$$

Plug estimation (30) into estimation (29) yields estimate (25). □

According to (25) and Lemma 4.3, since the cardinality of $\mathcal{A}^m$ is dm, we obtain the following theorem.

Theorem 4.5. *If X_a belongs to L^p, for any $\tau_0 > 0$, there exists $c > 0$ such that*

$$\mathbf{E}\left[\int_{\tau_0}^{\infty} LP_\tau f_N(S^m)\ d\tau\right] \leq c\,\|X\|_{L^p}\ \frac{N^{1+\eta}}{\sqrt{m}}\ \int_{\tau_0}^{\infty} \frac{e^{-5\tau/2}}{1-e^{-\tau/2}}\ d\tau. \tag{31}$$

If we combine Lemma 4.2 and (31), we get

$$|\mathbf{E}\left[f_N(S^m)\right] - \mathbf{E}\left[f_N(B^m)\right]| \leq c\|X\|_{L^p}\left(\sqrt{1-e^{-\tau_0}} + \frac{N^{1+\eta}}{\sqrt{m}}\ \int_{\tau_0}^{\infty} \frac{e^{-5\tau/2}}{1-e^{-\tau/2}}\ \mathrm{d}\tau\right).$$

Optimizing with respect to τ_0 yields Theorem 3.3.

It remains to prove (28). For the sake of simplicity, we give the proof for $d = 1$. The general situation is similar but with more involved notations.

We recall that

$$\pi^N(B^m) = \sum_{b=0}^{N-1} G_b^{m,N} h_b^N.$$

where

$$G_b^{m,N} = \sum_{a=0}^{m-1} \left\langle h_a^m, h_b^N \right\rangle_{I_{1,2}} \delta_{\eta,p}(h_a^m). \tag{32}$$

Lemma 4.6. *The covariance matrix Γ of the Gaussian vector $(G_b^{m,N},\ b = 0, \cdots, N-1)$ is invertible and satisfies*

$$\|\Gamma^{-1}\|_\infty \leq 2. \tag{33}$$

Proof. Since the h_a^m are orthogonal in L^2, for any $b, c \in \{0, \cdots, N-1\}$,

$$\Gamma_{b,c} = \sum_{a=0}^{m-1} \left\langle h_a^m, h_b^N \right\rangle_{I_{1,2}} \left\langle h_a^m, h_c^N \right\rangle_{I_{1,2}}. \tag{34}$$

Since a sub-interval of $\mathcal{D}_m$ intersects at most two sub-intervals of $\mathcal{D}_N$, the matrix Γ is tridiagonal. Furthermore, we know that

$$0 \leq \left\langle h_a^m, h_b^N \right\rangle_{I_{1,2}} \leq \sqrt{\frac{N}{m}}, \tag{35}$$

and for each b, there are at least $(N/m - 3)$ terms of this kind which are equal to $(N/m)^{-1/2}$. Hence,

$$\Gamma_{b,b} \geq (\frac{m}{N} - 3)(\sqrt{\frac{N}{m}})^2 \geq \frac{3}{4}.$$

Since Γ is tridiagonal, this implies that it is invertible. Moreover, let D be the diagonal matrix extracted from Γ. We have proved that $\|D\|_\infty \geq 3/4$.

For $|b - c| = 1$, there is at most one term of the sum (34) which yields a non zero scalar product, hence

$$|\Gamma_{b,c}| \leq \frac{N}{m}.$$

Set $S = \Gamma - D$. The matrix $D^{-1}S$ has at most two non null entries and

$$\|D^{-1}S\|_\infty \leq \frac{8}{3}\frac{N}{m} \leq \frac{1}{3},$$

if $m > 8N$. By iteration, we get for any $k \geq 1$,

$$\|(D^{-1}S)^k\|_\infty \leq \frac{1}{3^k}\cdot$$

Moreover,

$$\sum_{k=0}^{\infty}(-D^{-1}S)^k = (\mathrm{Id} + D^{-1}S)^{-1} = \Gamma^{-1}D.$$

Thus,

$$\|\Gamma^{-1}\|_\infty \leq \frac{4}{3}\sum_{k=0}^{\infty}\frac{1}{3^k} = 2.$$

The proof is thus complete. □

Lemma 4.7. *There exists a constant c which depends only on the dimension d such that for all m, N with $m > 8N$, for any $a \in \mathcal{A}^N$*

$$\mathrm{Var}\left[\mathbf{E}\left[\delta_{\eta,p}(h_a^m)\,|\,\pi^N(B^m)\right]\right] \leq c\,\frac{N}{m}\cdot$$

Proof. Using the framework of Gaussian vectors, for all $a \in \{0, \cdots, m-1\}$

$$\mathbf{E}\left[\delta_{\eta,p}(h_a^m)\,|\,\pi^N(B^m)\right] = \sum_{b\in\mathcal{A}^N} C_{a,b}^{m,N} G_b^{m,N}. \tag{36}$$

For any $c \in \{0, \cdots, N-1\}$, on the one hand

$$\begin{aligned}\mathbf{E}\left[\mathbf{E}\left[\delta_{\eta,p}(h_a^m)\,|\,\pi^N(B^m)\right]\ G_c\right] &= \sum_{b=0}^{N-1}\sum_{\tau=0}^{m-1} C_{a,b}^{m,N}\left\langle h_\tau^m,\ h_b^N\right\rangle_{I_{1,2}}\left\langle h_\tau^m,\ h_c^N\right\rangle_{I_{1,2}}\\ &= \sum_{b=0}^{N-1} C_{a,b}^{m,N}\Gamma_{b,c}.\end{aligned}$$

and on the other hand,

$$\mathbf{E}\left[\mathbf{E}\left[\delta_{\eta,p}(h_a^m)\,|\,\pi^N(B^m)\right]\ G_c\right] = \mathbf{E}\left[\delta_{\eta,p}(h_a^m)\,G_c\right] = \left\langle h_a^m,\ h_c^N\right\rangle_{I_{1,2}}.$$

This means that

$$\left(\left\langle h_a^m,\ h_c^N\right\rangle_{I_{1,2}},\, c = 0, \cdots, N-1\right) = \left(C_{a,b}^{m,N},\, b = 0, \cdots, N-1\right)\Gamma.$$

In view of Lemma 4.6, this entails that

$$\left(C_{a,b}^{m,N},\, b = 0, \cdots, N-1\right) = \left(\left\langle h_a^m,\ h_c^N\right\rangle_{I_{1,2}},\, c = 0, \cdots, N-1\right)\Gamma^{-1}.$$

Once again we invoke (35) and the fact that at most two of the terms $\left\langle h_a^m,\ h_c^N\right\rangle_{I_{1,2}}$ are non zero for a fixed a, to deduce that

$$\sup_{a,b}|C_{a,b}^{m,N}| \leq 2\|\Gamma^{-1}\|_\infty\sqrt{\frac{N}{m}} = 4\sqrt{\frac{N}{m}}\cdot \tag{37}$$

Now then, according to the very definition of the conditional expectation

$$\begin{aligned}\mathrm{Var}\left[\mathbf{E}\left[\delta_{\eta,p}(h_a^m)|\pi^N(B^m)\right]\right] &= \mathbf{E}\left[\delta_{\eta,p}(h_a^m)\ \mathbf{E}\left[\delta_{\eta,p}(h_a^m)|\pi^N(B^m)\right]\right]\\ &= \sum_{b=0}^{N-1} C_{a,b}^{m,N}\left\langle h_a^m, h_b^N\right\rangle_{I_{1,2}}.\end{aligned}$$

Hence,

$$\mathrm{Var}\left[\mathbf{E}\left[\delta_{\eta,p}(h_a^m)|\pi^N(B^m)\right]\right] \leq 2\sup_{a,b}|C_{a,b}^{m,N}|\sqrt{\frac{N}{m}} \leq 8\,\frac{N}{m},$$

according to (37). The constant 8 has to be modified when $d > 1$. □

References

[1] R. A. Adams and J; J. F. Fournier. *Sobolev Spaces*. Academic Press, jun 2003.

[2] H. Brézis. *Analyse fonctionnelle*. Masson edition, 1987.

[3] L. Coutin and L. Decreusefond. Stein's method for Brownian approximations. *Communications on Stochastic Analysis*, 7(3):349–372, September 2013. 00000.

[4] L. Coutin and L. Decreusefond. Convergence rate in the rough donsker theorem. *arXiv:1707.01269 [math]*, July 2017.

[5] L. Decreusefond. Stochastic integration with respect to Volterra processes. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 41(2):123–149, mar 2005.

[6] L. Decreusefond. The Stein-Dirichlet-Malliavin method. *ESAIM: Proceedings*, page 11, 2015.

[7] L. Decreusefond and H. Halconruy. Malliavin and Dirichlet structures for independent random variables. *Stochastic Processes and their Applications*, aug 2018.

[8] R. M. Dudley. *Real Analysis and Probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002.

[9] X. Fang, Q.-M. Shao, and L. Xu. Multivariate approximations in Wasserstein distance by Stein's method and Bismut's formula.

[10] D. Feyel and A. de la Pradelle. On fractional Brownian processes. *Potential Analysis*, 10(3):273–288, may 1999.

[11] P. Friz and N. Victoir. *Multidimensional Stochastic Processes as Rough Paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. 00382 Theory and applications.

[12] T. Gallouët, G. Mijoule, and Y. Swan. Regularity of solutions of the stein equation and rates in the multivariate central limit theorem.

[13] K. Itô and M. Nisio. On the convergence of sums of independent Banach space valued random variables. *Osaka Journal of Mathematics*, 5:35–48, 1968.

[14] D. Nualart. *The Malliavin Calculus and Related Topics*, volume 17. Springer–Verlag, 1995.

[15] M. Raič. A multivariate central limit theorem for Lipschitz and smooth test functions.

[16] H.-H. Shih. On Stein's method for infinite-dimensional Gaussian approximation in abstract Wiener spaces. *Journal of Functional Analysis*, 261(5):1236–1283, September 2011. 00000.

[17] A. S. Üstünel. Analysis on Wiener space and applications. *arXiv:1003.1649*, 12(1):85–90, mar 2010.

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

LTCI, Télécom Paris, Institut polytechnique de Paris, Paris, France
E-mail address — : Laurent.Decreusefond@mines-telecom.fr