TOTAL VARIATION BOUNDS FOR GAUSSIAN FUNCTIONALS

LUCA PRATELLI AND PIETRO RIGO

ABSTRACT. Let $X = \{X_t : 0 \leq t \leq 1\}$ be a centered Gaussian process with continuous paths, and $I_n = \frac{a_n}{2} \int_0^1 t^{n-1}(X_1^2 - X_t^2)\,dt$ where the a_n are suitable constants. Fix $\beta \in (0,1)$, $c_n > 0$ and $c > 0$ and denote by N_c the centered Gaussian kernel with (random) variance cX_1^2. Under an Holder condition on the covariance function of X, there is a constant $k(\beta)$ such that

$$\|P(\sqrt{c_n}\, I_n \in \cdot) - E\big[N_c(\cdot)\big]\,\| \leq k(\beta)\,\Big(\frac{a_n}{n^{1+\alpha}}\Big)^\beta + \frac{|c_n - c|}{c} \quad \text{for all } n \geq 1,$$

where $\|\cdot\|$ is total variation distance and α the Holder exponent of the covariance function. Moreover, if $\frac{a_n}{n^{1+\alpha}} \to 0$ and $c_n \to c$, then $\sqrt{c_n}\, I_n$ converges $\|\cdot\|$-stably to N_c, in the sense that

$$\|P_F\big(\sqrt{c_n}\, I_n \in \cdot\big) - E_F\big[N_c(\cdot)\big]\,\| \to 0$$

for every measurable F with $P(F) > 0$. In particular, such results apply to $X =$ fractional Brownian motion. In that case, they strictly improve the existing results in [5] and provide an essentially optimal rate of convergence.

1. Introduction

Malliavin calculus is a powerful tool which leads to effective results in a plenty of frameworks. Being a general approach, however, Malliavin calculus can not be expected to give optimal results in any specific problem. Indeed, it may be that (much) stronger results can be obtained through standard methods designedly shaped to the problem at hand.

This paper provides an example of this fact.

In [5], Nourdin, Nualart and Peccati establish general results on Malliavin operators and then apply such results to some (meaningful) special cases. Following this route, they get stable limit theorems for quadratic functionals of fractional Brownian motion.

Let B be a fractional Brownian motion with Hurst parameter H on the probability space $(\Omega, \mathcal{F}, P)$. Define

$$A_n = \frac{n^{1+H}}{2} \int_0^1 t^{n-1}(B_1^2 - B_t^2)\,dt.$$

The asymptotics of A_n and other analogous functionals (such as weighted power variations) is investigated in various papers. See e.g. [3], [5], [6] and references therein. One more reason for dealing with A_n is that, for $H \geq 1/4$,

$$\int_0^1 t^n B_t\,dB_t = \frac{A_n}{n^H} - \frac{H}{2H+n}$$

2010 *Mathematics Subject Classification.* 60B10, 60F05, 60G15.

Key words and phrases. Fractional Brownian motion, Gaussian process, Stable convergence, Total variation distance.

where the stochastic integral is meant in Skorohod's sense (it reduces to an Ito integral if $H = 1/2$); see [1].

Let $c = H\,\Gamma(2H)$ and let N_c be the Gaussian random probability measure with mean 0 and (random) variance cB_1^2. We denote by $E[N_c(\cdot)]$ the probability measure on $\mathcal{B}(\mathbb{R})$ given by $A \mapsto E[N_c(A)]$. Equivalently, $E[N_c(\cdot)]$ may be regarded as the probability distribution of $\sqrt{c}\,U\,B_1$, where U is a standard normal random variable independent of B_1.

In [5, Theorem 3.6] it is shown that, if $H \in [1/2,\,1)$, then A_n converges stably to N_c and

$$\| P(A_n \in \cdot) - E[N_c(\cdot)]\,\| \le k\,n^{-(1-H)/15} \quad \text{for all } n \ge 1,$$

where $\|\cdot\|$ is total variation distance and k a constant independent of n.

In this paper it is proved that, for *every* $H \in (0,1)$ and $\beta \in (0,1)$, there is a constant $k(H,\beta)$ such that

$$\| P(A_n \in \cdot) - E[N_c(\cdot)]\,\| \le k(H,\beta)\,n^{-\beta\,a(H)} \quad \text{for all } n \ge 1 \tag{1}$$

where

$$a(H) = 1/2 - |1/2 - H|.$$

Furthermore, for fixed $\epsilon > 0$,

$$\| P(A_n \in \cdot) - E[N_c(\cdot)]\,\| \;\text{ is not }\; \mathrm{O}\big(n^{-a(H)-\epsilon}\big). \tag{2}$$

Roughly speaking, $\| P(A_n \in \cdot) - E[N_c(\cdot)]\,\|$ can be estimated for every $H \in (0,1)$ (and not only for $H \in [1/2,\,1)$) with a rate arbitrarily close to n^{-H} or $n^{-(1-H)}$ according to whether $H < 1/2$ or $H \ge 1/2$. Also, in view of (2), such a rate is quite close to be optimal.

In addition, A_n converges stably to N_c in a very strong sense. In fact,

$$\| P_F(A_n \in \cdot) - E_F[N_c(\cdot)]\,\| \to 0 \tag{3}$$

for every $F \in \mathcal{F}$ with $P(F) > 0$, where $P_F(\cdot) = P(\cdot \mid F)$ is the probability measure on $\mathcal{F}$ conditional on F and E_F denotes expectation under P_F.

Both (1) and (3) are proved by standard, elementary tools. Thus, a question is whether they can be obtained by Malliavin calculus.

A last note is that the class of functionals covered by this paper is actually richer than it appears so far. In fact, (1) and (3) are strengthened as follows.

- (i) Up to replacing n^{1+H} and c with appropriate constants, (1) and (3) are still true if B is any centered Gaussian process with a suitable covariance function.
- (ii) Let $K_t = B_t - E(B_tB_1)\,B_1$ and $T = (K_{t_1},\ldots,K_{t_m})$, where $t_1,\ldots,t_m \in [0,1]$. Then, inequality (1) generalizes into
$$\| P\big[(T,\,A_n) \in \cdot\big] - Q_c(\cdot)\,\| \le k\,n^{-\beta\,a(H)} \quad \text{for all } n \ge 1,$$
where the constant k depends on $t_1,\ldots,t_m$, H and β while Q_c is a suitable probability measure on $\mathcal{B}(\mathbb{R}^{m+1})$. Further, the pairs (T, A_n) converge stably (in the strong sense mentioned above) to the product kernel $\delta_T \times N_c$.
- (iii) Our method of proof allows to handle functionals, more general than A_n, such as
$$A_n' = \frac{n^{1+H}}{p}\int_0^1 t^{n-1}(B_1^p - B_t^p)\,\{1 + g(K_t)\}\,dt$$

where K has been defined in (ii), g is a suitable function and $p \geq 2$ any integer. In particular, contrary to Malliavin calculus, the rate of convergence of A'_n does not depend on the degree of g when g is a polynomial.

Points (i)-(ii) are developed in Sections 3-4, while point (iii) is only briefly discussed in Section 5. However, A'_n could be managed by exactly the same techniques used for A_n.

2. Preliminaries

2.1. **Notation.** All random elements involved in this paper are defined on a fixed probability space

$$(\Omega, \mathcal{F}, P).$$

If $F \in \mathcal{F}$ and $P(F) > 0$, we let $P_F(A) = P(A \cap F)/P(F)$ for all $A \in \mathcal{F}$ and we write E_F to denote expectation under P_F.

Let $\mathcal{X}$ be a topological space. Then, $\mathcal{B}(\mathcal{X})$ denotes the Borel σ-field on $\mathcal{X}$ and $\mathbb{P}(\mathcal{X})$ the collection of probability measures on $\mathcal{B}(\mathcal{X})$. Also, $\|\cdot\|$ is the total variation distance on $\mathbb{P}(\mathcal{X})$, namely,

$$\|\mu - \nu\| = \sup_{A \in \mathcal{B}(\mathcal{X})} |\mu(A) - \nu(A)| \quad \text{whenever } \mu, \, \nu \in \mathbb{P}(\mathcal{X}).$$

A measurable map $L : \Omega \to \mathbb{P}(\mathcal{X})$ is called a *random probability measure* on $\mathcal{X}$ or a *kernel* on $\mathcal{X}$. Measurability means that $\omega \mapsto L(\omega)(A)$ is $\mathcal{F}$-measurable for fixed $A \in \mathcal{B}(\mathcal{X})$.

We denote by δ_x the unit mass at the point x and we write $X \sim \mu$ to mean that μ is the probability distribution of the random variable X. Further, $\mathcal{N}(a, b)$ is the Gaussian law on $\mathcal{B}(\mathbb{R})$ with mean $a \in \mathbb{R}$ and variance $b \geq 0$ (with $\mathcal{N}(a, 0) = \delta_a$). Finally, the standard normal density is denoted by ϕ, namely,

$$\phi(x) = (2\,\pi)^{-1/2} \, \exp\bigl(-(1/2)\, x^2\bigr).$$

2.2. **A (natural) extension of stable convergence.** Let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-field. For each $\mathcal{U} \subset \mathcal{F}$, write

$$\mathcal{U}^+ = \{F \in \mathcal{U} : P(F) > 0\}.$$

Also, let $\mathcal{X}$ be a separable metric space, L a random probability measure on $\mathcal{X}$ and X_n a sequence of $\mathcal{X}$-valued random variables.

Then, X_n converges to L *stably with respect to* $\mathcal{G}$ if

$$P_G(X_n \in \cdot) \overset{weakly}{\longrightarrow} E_G[L(\cdot)]$$

as $n \to \infty$ for all $G \in \mathcal{G}^+$. Here, $E_G[L(\cdot)]$ stands for the probability measure on $\mathcal{B}(\mathcal{X})$ given by $A \mapsto E_G[L(A)]$.

Stable convergence can be extended, consistently with Renyi's original ideas [7], as follows. Fix a distance d on $\mathbb{P}(\mathcal{X})$ and say that X_n converges to L, *d-stably with respect to* $\mathcal{G}$, if

$$d\Big(P_G(X_n \in \cdot), \, E_G[L(\cdot)]\Big) \to 0$$

as $n \to \infty$ for all $G \in \mathcal{G}^+$. Such a definition reduces to the previous one if d is any distance metrizing weak convergence of probability measures.

In the sequel, we let

$$\mathcal{G} = \mathcal{F} \quad \text{and} \quad d = \|\cdot\| = \text{ total variation distance.}$$

Since $\mathcal{G} = \mathcal{F}$, for simplicity, $\mathcal{G}$ is not mentioned at all. Hence, X_n is said *to converge* $\|\cdot\|$*-stably to* L if

$$\lim_n \|P_F(X_n \in \cdot) - E_F[L(\cdot)]\| = 0 \tag{4}$$

for all $F \in \mathcal{F}^+$.

A last note is that, to prove $\|\cdot\|$-stable convergence, it suffices to get (4) for certain conditioning events F.

Lemma 1. *Let* $\mathcal{U} \subset \mathcal{F}$ *be such that: (i)* $\Omega \in \mathcal{U}$*; (ii)* $\mathcal{U}$ *is closed under finite intersections; (iii)* $\sigma(\mathcal{U}) \supset \sigma(L, X_1, X_2, \ldots)$*. If condition* (4) *holds for all* $F \in \mathcal{U}^+$*, then* X_n *converges* $\|\cdot\|$*-stably to* L*.*

Proof. By (i)-(ii) and the inclusion-exclusion formulae, condition (4) holds for each F in the field generated by $\mathcal{U}$. Hence, it can be assumed that $\mathcal{U}$ is a field. Note also that, for any $F, G \in \mathcal{F}^+$,

$$\|P_F(X_n \in \cdot) - E_F[L(\cdot)]\| \le \frac{2P(F\Delta G) + \|P_G(X_n \in \cdot) - E_G[L(\cdot)]\|}{P(F)}.$$

Next, fix $F \in \sigma(\mathcal{U})^+$ and $\epsilon > 0$. Since $\mathcal{U}$ is a field, there is $G \in \mathcal{U}^+$ such that $2\,P(F\Delta G)/P(F) < \epsilon$. Thus, condition (4) holds for all $F \in \sigma(\mathcal{U})^+$. It follows that

$$\sup_{A \in \mathcal{B}(\mathcal{X})} \Big| E\Big(1_{\{X_n \in A\}} V\Big) - E\Big(L(A) V\Big)\Big| \to 0$$

provided the random variable V is bounded and $\sigma(\mathcal{U})$-measurable. Therefore, because of (iii), for each $F \in \mathcal{F}^+$ one obtains

$$\|P_F(X_n \in \cdot) - E_F[L(\cdot)]\| =$$

$$= \frac{1}{P(F)} \sup_{A \in \mathcal{B}(\mathcal{X})} \Big| E\Big(1_{\{X_n \in A\}} E(1_F \mid \sigma(\mathcal{U}))\Big) - E\Big(L(A) E(1_F \mid \sigma(\mathcal{U}))\Big)\Big| \to 0.$$

□

2.3. **Technical lemmas.** A few simple facts are needed to prove our main result (Theorem 6). They are most probably known, but we provide proofs since we are not aware of any reference.

Lemma 2. *Let* X_n *and* X *be real random variables. If*

$$E(|X_n|^p) = O(1) \quad \text{and} \quad \|P(X_n \in \cdot) - P(X \in \cdot)\| = O(n^{-b}),$$

for all $p > 0$ *and some* $b > 0$*, then*

$$|E(X_n) - E(X)| = O\big(n^{-b+\epsilon}\big) \quad \text{for all } \epsilon > 0.$$

Proof. First note that $E(|X|) \le \liminf_n E(|X_n|) < \infty$. Up to considering positive and negative parts, it can be assumed $X_n \ge 0$ and $X \ge 0$. Then, for all $n \ge 1$,

$t_0 > 0$ and $p > 1$, one obtains

$$|E(X_n) - E(X)| = \Big| \int_0^\infty \big\{ P(X_n > t) - P(X > t) \big\}\, dt \Big|$$
$$\le t_0 \, \| P(X_n \in \cdot) - P(X \in \cdot) \| + \int_{t_0}^\infty \big\{ P(X_n > t) + P(X > t) \big\}\, dt$$
$$\le t_0 \, c\, n^{-b} + \int_{t_0}^\infty \frac{E(X_n^p) + E(X^p)}{t^p}\, dt$$
$$\le t_0 \, c\, n^{-b} + \int_{t_0}^\infty \frac{c_p}{t^p}\, dt = t_0 \, c\, n^{-b} + c_p \, \frac{t_0^{1-p}}{p-1}$$

where c and c_p are suitable constants. Thus, given $\epsilon > 0$, it suffices to take $t_0 = n^\epsilon$ and $p > \max(1,\, b/\epsilon)$. □

Lemma 3. *If $a_1,\, a_2 \in \mathbb{R}$, $0 \le b_1 \le b_2$ and $b_2 > 0$, then*

$$\| \mathcal{N}(a_1, b_1) - \mathcal{N}(a_2, b_2) \| \le 1 - \sqrt{\frac{b_1}{b_2}} + \frac{|a_1 - a_2|}{\sqrt{2\,\pi\, b_2}}.$$

Proof. The lemma is trivially true if $b_1 = 0$. Hence, it can be assumed $b_1 > 0$. Let $u = \sqrt{b_1/b_2}$. Since $u \le 1$, then $\phi(x) \le \phi(ux)$ for each $x \ge 0$. Thus,

$$\| \mathcal{N}(a_1, b_1) - \mathcal{N}(a_1, b_2) \| = \| \mathcal{N}(0, b_1) - \mathcal{N}(0, b_2) \|$$
$$= \frac{1}{2} \int_{-\infty}^\infty \Big| \frac{1}{\sqrt{b_1}} \phi\big(\frac{x}{\sqrt{b_1}}\big) - \frac{1}{\sqrt{b_2}} \phi\big(\frac{x}{\sqrt{b_2}}\big) \Big|\, dx = \int_0^\infty |\phi(x) - u\, \phi(ux)|\, dx$$
$$\le (1-u) \int_0^\infty \phi(x)\, dx + u \int_0^\infty \big(\phi(ux) - \phi(x)\big)\, dx$$
$$= \frac{(1-u)}{2} + \frac{1}{2} - \frac{u}{2} = 1 - u.$$

Letting $\alpha = (a_1 - a_2)/\sqrt{b_2}$, one also obtains

$$\| \mathcal{N}(a_1, b_2) - \mathcal{N}(a_2, b_2) \| = \| \mathcal{N}(\alpha,\, 1) - \mathcal{N}(0, 1) \|$$
$$= \frac{1}{2} \int_{-\infty}^\infty |\phi(x - \alpha) - \phi(x)|\, dx = \int_{-\frac{|\alpha|}{2}}^{\frac{|\alpha|}{2}} \phi(x)\, dx \le \frac{|\alpha|}{\sqrt{2\,\pi}}.$$

Therefore,

$$\| \mathcal{N}(a_1, b_1) - \mathcal{N}(a_2, b_2) \| \le \| \mathcal{N}(a_1, b_1) - \mathcal{N}(a_1, b_2) \| + \| \mathcal{N}(a_1, b_2) - \mathcal{N}(a_2, b_2) \|$$
$$\le 1 - u + \frac{|\alpha|}{\sqrt{2\,\pi}}.$$

□

If $a_1 = a_2$, Lemma 3 is well known; see e.g. [4, Proposition 3.6.1]. In this case, one trivially obtains

$$\| \mathcal{N}(a, b_1) - \mathcal{N}(a, b_2) \| \le \frac{|b_1 - b_2|}{b_i} \quad \text{for each } i \text{ such that } b_i > 0.$$

As already pointed out, Lemma 3 is most probably known even if $a_1 \ne a_2$, but we do not know of any reference.

In the next result, $U, T_1, \ldots, T_m$ are real random variables and $T = (T_1, \ldots, T_m)$.

Lemma 4. *Let λ and μ be the probability distributions of T and (T,U), respectively, and*

$$\nu = \lambda \times \mathcal{N}(0,b) \quad \textit{where } b > 0.$$

If (T,U) has a centered Gaussian distribution, then

$$\|\mu - \nu\| \le \frac{r}{\sqrt{b}} \Big\{ |\sqrt{b} - \sqrt{E(U^2)}| + \sum_{i=1}^m |E(UT_i)| \Big\}$$

where the constant r depends only on λ.

Proof. Throughout this proof, each $x \in \mathbb{R}^m$ is regarded as a column vector, x' is the transpose of x and

$$\|x\|^* = \sum_{i=1}^m |x_i|.$$

Let Σ be the covariance matrix of T. If $\operatorname{rank}(\Sigma) = 0$, then $T_1 = \ldots = T_m = 0$ a.s. and the lemma is trivially true because of Lemma 3. If $0 < \operatorname{rank}(\Sigma) < m$, there is a subvector T_0 of T and a (linear) function h such that $T = h(T_0)$ a.s. and T_0 has non-singular covariance matrix. In that case, denoting by V a random variable independent of T such that $V \sim \mathcal{N}(0,b)$, one obtains

$$\begin{aligned}\|\mu - \nu\| = \|P\big((T,U)\in\cdot\big) - P\big((T,V)\in\cdot\big)\| &= \|P\big((h(T_0),U)\in\cdot\big) - P\big((h(T_0),V)\in\cdot\big)\| \\ &\le \|P\big((T_0,U)\in\cdot\big) - P\big((T_0,V)\in\cdot\big)\|.\end{aligned}$$

Thus, up to replacing T with T_0, it can be assumed $\operatorname{rank}(\Sigma) = m$.

Let $C' = \big(E(UT_1), \ldots, E(UT_m)\big)$. For $t \in \mathbb{R}^m$, define

$$\mu_t = \mathcal{N}\Big(C'\Sigma^{-1}t,\, E(U^2) - C'\Sigma^{-1}C\Big).$$

Then, $\{\mu_t : t \in \mathbb{R}^m\}$ is a version of the conditional distribution of U given T. By Lemma 3,

$$\begin{aligned}\|\mu_t - \mathcal{N}(0,b)\| &\le \frac{1}{\sqrt{b}}\Big(\Big|\sqrt{b} - \sqrt{E(U^2) - C'\Sigma^{-1}C}\Big| + |C'\Sigma^{-1}t|\Big) \\ &\le \frac{1}{\sqrt{b}}\Big(|\sqrt{b} - \sqrt{E(U^2)}| + \sqrt{C'\Sigma^{-1}C} + |C'\Sigma^{-1}t|\Big).\end{aligned}$$

Hence, there is a constant r_0 depending only on Σ such that

$$\|\mu_t - \mathcal{N}(0,b)\| \le \frac{r_0}{\sqrt{b}}\Big(|\sqrt{b} - \sqrt{E(U^2)}| + \|C\|^* + \|C\|^*\|\Sigma^{-1}t\|^*\Big).$$

Define

$$r = r_0\,\big(1 + \int \|\Sigma^{-1}t\|^*\,\lambda(dt)\big)$$

(recall that λ is the probability distribution of T). Then,

$$\begin{aligned}\|\mu - \nu\| &\le \int \|\mu_t - \mathcal{N}(0,b)\|\,\lambda(dt) \\ &\le \frac{r_0}{\sqrt{b}}\Big(|\sqrt{b} - \sqrt{E(U^2)}| + \|C\|^* + \|C\|^* \int \|\Sigma^{-1}t\|^*\,\lambda(dt)\Big) \\ &\le \frac{r}{\sqrt{b}}\Big(|\sqrt{b} - \sqrt{E(U^2)}| + \|C\|^*\Big).\end{aligned}$$

□

The last lemma plays a key role in proving Theorem 6.

Lemma 5. *Let U be a standard normal random variable and $D=\{U>x\}$, where $x\in\mathbb{R}\cup\{-\infty\}$. Then, there is a constant $q\geq 1$ (independent of x) such that*

$$\|P_D\big(rU^2+yU+z\in\cdot\big)-P_D\big(yU\in\cdot\big)\|\leq \min\Big(1,\ \frac{q}{P(D)}\,\frac{|r|+|z|}{|y|}\Big)$$

for all r, y, $z\in\mathbb{R}$ with $y\neq 0$.

Proof. On noting that

$$\|P_D\big(rU^2+yU+z\in\cdot\big)-P_D\big(yU\in\cdot\big)\|=\|P_D\Big(\frac{r}{y}\,U^2+U+\frac{z}{y}\in\cdot\Big)-P_D(U\in\cdot)\|,$$

it can be assumed $y=1$.

By Lemma 3, $\|P(U+z\in\cdot)-P(U\in\cdot)\|\leq|z|$. Hence,

$$\begin{aligned}\|P\big(rU^2+U+z\in\cdot\big)-P\big(U\in\cdot\big)\|&\leq\|P\big(rU^2+U+z\in\cdot\big)-P\big(U+z\in\cdot\big)\|+|z|\\&=\|P\big(rU^2+U\in\cdot\big)-P\big(U\in\cdot\big)\|+|z|.\end{aligned}$$

Define

$$\eta=P\Big(\{U>x\}\Delta\{rU^2+U+z>x\}\Big).$$

Then,

$$\begin{aligned}&P(D)\,\|P_D\big(rU^2+U+z\in\cdot\big)-P_D\big(U\in\cdot\big)\|\\&\leq\eta+\|P\big(rU^2+U+z\in\cdot\big)-P\big(U\in\cdot\big)\|\\&\leq\eta+|z|+\|P\big(rU^2+U\in\cdot\big)-P\big(U\in\cdot\big)\|.\end{aligned}$$

Since $\{U+z>x\}\subset\{rU^2+U+z>x\}$ if $r>0$ and $\{rU^2+U+z>x\}\subset\{U+z>x\}$ if $r<0$, one also obtains

$$\begin{aligned}\eta&\leq P\Big(\{U>x\}\Delta\{U+z>x\}\Big)+P\Big(\{U+z>x\}\Delta\{rU^2+U+z>x\}\Big)\\&\leq|z|+\Big|P\big(U+z>x\big)-P\big(rU^2+U+z>x\big)\Big|\\&\leq|z|+\|P(rU^2+U\in\cdot)-P(U\in\cdot)\|.\end{aligned}$$

In addition, since $-U\sim U$,

$$\|P\big(-rU^2+U\in\cdot\big)-P\big(U\in\cdot\big)\|=\|P\big(rU^2+U\in\cdot\big)-P\big(U\in\cdot\big)\|.$$

To summarize, to prove the lemma, it suffices to see that there is a constant $q\geq 1$ such that

$$\|P\big(r\,U^2+U\in\cdot\big)-P\big(U\in\cdot\big)\|\leq q\,r\quad\text{ for each } r\in(0,1]. \tag{5}$$

We finally prove (5). Define

$$q=7\,+\,6\int_0^\infty u^3\,\phi\Big(\frac{2u}{1+\sqrt{1+4u}}\Big)\,du$$

and fix $r\in(0,1]$. Let g be the density of $r\,U^2+U$ with respect to Lebesgue measure. Then, $g(u)=0$ if $u\leq-1/4r$ and $g(u)=\phi_1(u)+\phi_2(u)$ if $u>-1/4r$, where

$$\phi_1(u)=\phi\Big(\frac{-1+\sqrt{1+4ru}}{2r}\Big)\,\frac{1}{\sqrt{1+4ru}},\qquad \phi_2(u)=\phi\Big(-\frac{1+\sqrt{1+4ru}}{2r}\Big)\,\frac{1}{\sqrt{1+4ru}}.$$

On noting that $P\big(U<-(1/c)\big)\le c$ for all $c>0$ and

$$\int_{-\frac{1}{4r}}^{\infty}\phi_2(u)\,du=\int_{-\infty}^{-\frac{1}{2r}}\phi(u)\,du=P\big(U<-\frac{1}{2r}\big),$$

one obtains

$$2\,\|P\big(r\,U^2+U\in\cdot\big)-P\big(U\in\cdot\big)\|=\int_{-\infty}^{\infty}|g(u)-\phi(u)|\,du\le 6r+\int_{-\frac{1}{4r}}^{\infty}|\phi_1(u)-\phi(u)|\,du$$

$$\le 6r+\int_{-\frac{1}{4r}}^{\infty}\phi\Big(\frac{-1+\sqrt{1+4ru}}{2r}\Big)\,\frac{4r|u|}{\sqrt{1+4ru}}\,du+\int_{-\frac{1}{4r}}^{\infty}\Big|\,\phi\Big(\frac{-1+\sqrt{1+4ru}}{2r}\Big)-\phi(u)\Big|\,du.$$

Since $r\le 1$,

$$\int_{-\frac{1}{4r}}^{\infty}\phi\Big(\frac{-1+\sqrt{1+4ru}}{2r}\Big)\,\frac{4r|u|}{\sqrt{1+4ru}}\,du=4r\int_{-\frac{1}{2r}}^{\infty}|ru^2+u|\,\phi(u)\,du\le 4r\,E\big\{U^2+|U|\big\}\le 8r.$$

Since

$$\frac{-1+\sqrt{1+4ru}}{2r}=\frac{2u}{1+\sqrt{1+4ru}}\quad\text{and}\quad\Big|\frac{2u}{1+\sqrt{1+4ru}}-u\Big|\le 4ru^2,$$

the Lagrange theorem yields

$$\int_{-\frac{1}{4r}}^{\infty}\Big|\,\phi\Big(\frac{-1+\sqrt{1+4ru}}{2r}\Big)-\phi(u)\Big|\,du\le$$

$$\le\int_{-\frac{1}{4r}}^{0}4ru^2\,\frac{2|u|}{1+\sqrt{1+4ru}}\,\phi(u)\,du\;+\;\int_0^{\infty}4ru^2\,u\,\phi\Big(\frac{2u}{1+\sqrt{1+4ru}}\Big)\,du$$

$$\le 8r\int_{-\frac{1}{4r}}^{0}|u|^3\,\phi(u)\,du\;+\;4r\int_0^{\infty}u^3\,\phi\Big(\frac{2u}{1+\sqrt{1+4u}}\Big)\,du$$

$$\le 12r\int_0^{\infty}u^3\,\phi\Big(\frac{2u}{1+\sqrt{1+4u}}\Big)\,du.$$

Collecting all these facts together,

$$\|P\big(r\,U^2+U\in\cdot\big)-P\big(U\in\cdot\big)\|\le 3r\;+\;4r\;+\;6r\int_0^{\infty}u^3\phi\Big(\frac{2u}{1+\sqrt{1+4u}}\Big)\,du=q\,r.$$

Thus, condition (5) holds, and this concludes the proof. □

Even if conceptually simple, the above proof is quite cumbersome. As suggested by an anonymous referee, such a proof could be notably shortened by exploiting Stein's method to get (5). However, as noted in Section 1, one goal of this paper is to use elementary tools only. Thus, we have not adopted this shorter proof.

3. Results

From now on, $X=\{X_t:0\le t\le 1\}$ is a centered Gaussian process with continuous paths and covariance function

$$f(s,t)=E\big(X_sX_t\big).$$

We assume $f(1,1)=E(X_1^2)=1$. More importantly, we assume that there are two constants $\alpha>0$ and $\gamma>0$ such that

$$|f(s,t)-f(s,1)|\le\gamma\,(1-t)^{\alpha}\quad\text{for all } s,\,t\in[0,1]. \tag{6}$$

Note that, since the X-paths are continuous, condition (6) implies continuity of f.

Let

$$K_t = X_t - f(t,1)X_1.$$

The process K is a sort of Brownian bridge based on X. In particular, K is independent of X_1. In the sequel, we fix $m \geq 1$, $t_1, \ldots, t_m \in [0,1]$, and we let

$$T = (K_{t_1}, \ldots, K_{t_m}).$$

We aim to investigate

$$I_n = \frac{a_n}{2} \int_0^1 t^{n-1}(X_1^2 - X_t^2)\, dt,$$

where $n \geq 1$ and the constant a_n is given by

$$a_n = \Big(E\big[\big(\int_0^1 t^{n-1} f(t,1)\, K_t\, dt\big)^2\big]\Big)^{-1/2}.$$

The expectation involved in a_n is finite but may be null. Thus, the convention $1/0 = \infty$ is adopted.

For each $c \geq 0$,

$$N_c = \mathcal{N}\big(0,\, cX_1^2\big)$$

denotes the centered Gaussian kernel with variance cX_1^2. Also, Q_c is the product probability measure on $\mathcal{B}\big(\mathbb{R}^{m+1}\big)$ given by

$$Q_c = P(T \in \cdot) \times E[N_c(\cdot)].$$

We are now able to state our main result.

Theorem 6. *Let X be a centered Gaussian process with continuous paths and covariance function satisfying* (6)*. Fix $m \geq 1$, $t_1, \ldots, t_m \in [0,1]$, and define $T = (K_{t_1}, \ldots, K_{t_m})$. Then, for all $\beta \in (0,1)$, $c_n > 0$ and $c > 0$, one obtains*

$$\| P\big[(T,\, \sqrt{c_n}\, I_n) \in \cdot\big] - Q_c(\cdot)\, \| \leq \frac{|c_n - c|}{c} + k\left(\frac{a_n}{n^{1+\alpha}}\right)^{\beta} \quad \text{for all } n \geq 1$$

where k is a constant depending on $t_1, \ldots, t_m$ and β. Moreover, if $\frac{a_n}{n^{1+\alpha}} \to 0$ and $c_n \to c$, then

$$(T,\, \sqrt{c_n}\, I_n) \quad \textit{converges } \|\cdot\|\textit{-stably to the product kernel} \quad \delta_T \times N_c.$$

Based on Theorem 6, if $\frac{a_n}{n^{1+\alpha}} \to 0$ and $c_n \to c$, a conjecture is that $(K,\, \sqrt{c_n}\, I_n)$ converges $\|\cdot\|$-stably to $\delta_K \times N_c$, where K is regarded as a random element with values continuous functions on $[0,1]$. This is not true, however, as shown by Example 9.

Theorem 6 is actually a consequence of the following lemma, which has possible independent interest.

Lemma 7. *Let $D = \{X_1 > x\}$, where $x \in \mathbb{R} \cup \{-\infty\}$, and let $Q_{1,D}$ be the product probability measure on $\mathcal{B}\big(\mathbb{R}^{m+1}\big)$ given by*

$$Q_{1,D} = P(T \in \cdot) \times E_D[N_1(\cdot)].$$

(Note that $Q_{1,\Omega} = Q_1$). For every $\beta \in (0,1)$ there is constant k, depending on $x, t_1, \ldots, t_m$ and β, such that

$$\| P_D\big[(T, I_n) \in \cdot\big] - Q_{1,D}(\cdot)\, \| \leq k\left(\frac{a_n}{n^{1+\alpha}}\right)^{\beta} \quad \text{for all } n \geq 1.$$

Finally, Theorem 6 implies the results mentioned in Section 1, which in turn largely improve [5, Theorem 3.6].

Recall that a fractional Brownian motion is a centered Gaussian process with covariance function

$$f(s,t) = \frac{|s|^{2H} + |t|^{2H} - |s-t|^{2H}}{2},$$

where the Hurst parameter H ranges in $(0,1)$. Such f satisfies condition (6) with

$$\alpha = \min(1,\, 2H).$$

Corollary 8. *Let B be a fractional Brownian motion with Hurst parameter H. Define*

$$c = H\,\Gamma(2H), \quad a(H) = 1/2 - |1/2 - H|, \quad A_n = \frac{n^{1+H}}{2}\int_0^1 t^{n-1}(B_1^2 - B_t^2)\,dt,$$
$$T = (K_{t_1}, \ldots, K_{t_m}) \quad \textit{for some } m \geq 1 \textit{ and } t_1, \ldots, t_m \in [0,1].$$

Then, (T, A_n) converges $\|\cdot\|$-stably to the product kernel $\delta_T \times N_c$. In addition, for each $H \in (0,1)$ and $\beta \in (0,1)$, one obtains

$$\|P\big[(T,\, A_n) \in \cdot\big] - Q_c(\cdot)\,\| \leq k\, n^{-\beta\, a(H)} \quad \textit{for all } n \geq 1$$

where k is a constant depending on $H, t_1, \ldots, t_m$ and β.

As noted in Section 1 (see condition (2)) the rate of convergence provided by Corollary 8 is nearly optimal.

4. Proofs

Our proofs rely on two simple facts. First, I_n can be written as

$$I_n = r_n X_1^2 + Y_n X_1 + Z_n, \tag{7}$$

where the r_n are constants and (Y_n, Z_n) random variables independent of X_1. Second, because of Lemma 5, the asymptotic behavior of I_n essentially agrees with that of $Y_n X_1$.

To prove (7), just note that, since $X_t = K_t + f(t,1)X_1$,

$$I_n = (a_n/2)\int_0^1 t^{n-1}\Big((1 - f(t,1)^2)X_1^2 - 2f(t,1)K_t X_1 - K_t^2\Big)\,dt.$$

Thus, one obtains (7) and (Y_n, Z_n) independent of X_1 by letting

$$r_n = (a_n/2)\int_0^1 t^{n-1}(1 - f(t,1)^2)\,dt,$$
$$Y_n = -a_n \int_0^1 t^{n-1} f(t,1)\,K_t\,dt,$$
$$Z_n = -(a_n/2)\int_0^1 t^{n-1} K_t^2\,dt.$$

Two more remarks are in order. First, up to replacing k with $\max\big(k,\, 1\big)$, the inequalities in Theorem 6 and Lemma 7 are trivially true if $a_n > n^{1+\alpha}$ (in particular, if $a_n = \infty$). Accordingly, in the sequel, we assume that

$$a_n \leq n^{1+\alpha} \quad \text{for each } n \geq 1.$$

Second, for each $\epsilon > 0$, Stirling formula implies

$$\frac{n^{1+\epsilon}}{\Gamma(\epsilon+1)} \int_0^1 t^{n-1}(1-t)^\epsilon \, dt = \frac{n^\epsilon \, n!}{\Gamma(n+\epsilon+1)} = 1 \, + \, \mathrm{O}(n^{-1}). \tag{8}$$

4.1. **Proof of Lemma 7.** Let $\beta \in (0,1)$. Denote by λ_n and λ_n^* the probability distributions of (Y_n, Z_n) and (T, Y_n, Z_n), respectively. Since $I_n = r_n X_1^2 + Y_n X_1 + Z_n$ and X_1 is independent of (T, Y_n, Z_n), one obtains

$$\begin{aligned}
&\|P_D\big[(T, I_n) \in \cdot\big] - P_D\big[(T,\, Y_n X_1) \in \cdot\big]\, \| \leq \\
&\leq \int \|P_D\big[(t,\, r_n X_1^2 + y X_1 + z) \in \cdot\big] - P_D\big[(t,\, y X_1) \in \cdot\big]\, \|\, \lambda_n^*(dt, dy, dz) \\
&\leq \int \|P_D\big(r_n X_1^2 + y X_1 + z \in \cdot\big) - P_D\big(y X_1 \in \cdot\big)\, \|\, \lambda_n(dy, dz) \\
&\leq \int \min\big(1,\ \frac{q}{P(D)}\, \frac{|r_n| + |z|}{|y|}\big)\, \lambda_n(dy, dz) \\
&= E\Big(\min\big(1,\ \frac{q}{P(D)}\, \frac{|r_n| + |Z_n|}{|Y_n|}\big)\Big)
\end{aligned}$$

where q is a constant and the last inequality depends on Lemma 5. Thus, to prove Lemma 7, it suffices to see that

$$E\Big(\min\big(1,\ \frac{q}{P(D)}\, \frac{|r_n| + |Z_n|}{|Y_n|}\big)\Big) = \mathrm{O}\Big(\big(\frac{a_n}{n^{1+\alpha}}\big)^\beta\Big) \tag{9}$$

and

$$\|P_D\big[(T,\, Y_n X_1) \in \cdot\big] - Q_{1,D}(\cdot)\| = \mathrm{O}\Big(\big(\frac{a_n}{n^{1+\alpha}}\big)^\beta\Big). \tag{10}$$

We begin with (9). Fix an integer $m \geq 1$ and define

$$\delta = \sup_t (1 + |f(t,1)|) \quad \text{and} \quad \mu_n(dt) = n\, 1_{[0,1]}(t)\, t^{n-1}\, dt.$$

By condition (6), $|1 - f(t,1)| \leq \gamma(1-t)^\alpha$. By Stirling formula (8),

$$n \int_0^1 t^{n-1}(1-t)^{m\alpha}\, dt \leq b\, n^{-m\alpha}$$

for some constant b. Hence,

$$\begin{aligned}
|2\, r_n|^m &= a_n^m \left|(1/n) \int_0^1 (1 - f(t,1)^2)\, \mu_n(dt)\right|^m \leq (a_n/n)^m \int_0^1 |1 - f(t,1)^2|^m\, \mu_n(dt) \\
&\leq (a_n/n)^m\, \delta^m \int_0^1 |1 - f(t,1)|^m\, \mu_n(dt) \leq (a_n/n)^m\, (\gamma\, \delta)^m \int_0^1 (1-t)^{m\alpha}\, \mu_n(dt) \\
&= (a_n/n)^m\, (\gamma\, \delta)^m\, n \int_0^1 t^{n-1}(1-t)^{m\alpha}\, dt \leq b\, (\gamma\, \delta)^m \left(\frac{a_n}{n^{1+\alpha}}\right)^m.
\end{aligned}$$

As to Z_n, first note that K_t is a centered Gaussian random variable with variance

$$\begin{aligned}
E(K_t^2) &= f(t,t) - f(t,1)^2 \leq |1 - f(t,t)| + |1 - f(t,1)^2| \\
&\leq \delta\, \big(|f(t,1) - f(t,t)| + 2\, |1 - f(t,1)|\big) \leq 3\, \gamma\, \delta\, (1-t)^\alpha,
\end{aligned}$$

where the last inequality is by (6). Since $X_1 \sim \mathcal{N}(0,1)$,

$$E(K_t^{2m}) = E(X_1^{2m})\, E(K_t^2)^m \leq E(X_1^{2m})\, (3\, \gamma\, \delta)^m\, (1-t)^{m\alpha}.$$

Therefore,

$$E\big(|2\,Z_n|^m\big) \le (a_n/n)^m\, E\Big(\int_0^1 K_t^{2m}\,\mu_n(dt)\Big) = (a_n/n)^m \int_0^1 E(K_t^{2m})\,\mu_n(dt)$$

$$\le E(X_1^{2m})\,(3\,\gamma\,\delta)^m\,(a_n/n)^m\, n\int_0^1 t^{n-1}(1-t)^{m\alpha}\,dt \le b\,E(X_1^{2m})\,(3\,\gamma\,\delta)^m\,\Big(\frac{a_n}{n^{1+\alpha}}\Big)^m.$$

Hence, for each fixed $m \ge 1$, there is a constant $b(m)$ such that

$$E\Big((|r_n|+|Z_n|)^m\Big) \le (1/2)\,\Big(|2\,r_n|^m + E(|2\,Z_n|^m)\Big) \le b(m)\,\Big(\frac{a_n}{n^{1+\alpha}}\Big)^m.$$

Next, take $m > \beta/(1-\beta)$. On noting that $\min(1,\,u) \le u^\beta$ for every $u \ge 0$ (since $0<\beta<1$) one obtains

$$E\Big(\min\big(1,\ \frac{q}{P(D)}\,\frac{|r_n|+|Z_n|}{|Y_n|}\big)\Big) \le \Big(\frac{q}{P(D)}\Big)^\beta\, E\Big((|r_n|+|Z_n|)^\beta\,|Y_n|^{-\beta}\Big)$$

$$\le \Big(\frac{q}{P(D)}\Big)^\beta\, E\Big((|r_n|+|Z_n|)^m\Big)^{\frac{\beta}{m}}\, E\big(|Y_n|^{\frac{-m\beta}{m-\beta}}\big)^{\frac{m-\beta}{m}}$$

where the second inequality depends on Schwartz-Holder inequality with exponent $p = m/\beta$. Further,

$$E\big(|Y_n|^{\frac{-m\beta}{m-\beta}}\big)^{\frac{m-\beta}{m}} = E\big(|X_1|^{\frac{-m\beta}{m-\beta}}\big)^{\frac{m-\beta}{m}} < \infty$$

since $Y_n \sim \mathcal{N}(0,1) \sim X_1$ and $m\beta < m-\beta$. Thus, one finally obtains

$$E\Big(\min\big(1,\ \frac{q}{P(D)}\,\frac{|r_n|+|Z_n|}{|Y_n|}\big)\Big) \le b^*\,\Big(\frac{a_n}{n^{1+\alpha}}\Big)^\beta$$

for some constant b^*.

It remains to prove (10). For each $u \in \mathbb{R}\setminus\{0\}$, let ν_u be the product probability measure on $\mathcal{B}(\mathbb{R}^{m+1})$ given by

$$\nu_u = P(T\in\cdot)\times\mathcal{N}(0,u^2).$$

Recalling that $T=(K_{t_1},\ldots,K_{t_m})$ and $Y_n u \sim \mathcal{N}(0,u^2)$, Lemma 4 implies

$$\|P\big[(T,\,Y_n u)\in\cdot\big] - \nu_u(\cdot)\| \le \frac{r}{|u|}\,\sum_{i=1}^m |E(K_{t_i}Y_n u)| = r\,\sum_{i=1}^m |E(K_{t_i}Y_n)|$$

where the constant r depends only on the distribution of T. Conditioning on X_1, it follows that

$$\|P_D\big[(T,\,Y_nX_1)\in\cdot\big] - Q_{1,D}(\cdot)\| \le \frac{1}{P(D)}\int_x^\infty \|P\big[(T,\,Y_n u)\in\cdot\big]-\nu_u(\cdot)\|\,\phi(u)\,du$$

$$\le \frac{r}{P(D)}\,\sum_{i=1}^m |E(K_{t_i}Y_n)|.$$

On the other hand,

$$-E(K_{t_i}Y_n) = a_n\,E\Big(\int_0^1 t^{n-1}f(t,1)\,K_{t_i}\,K_t\,dt\Big) = a_n\int_0^1 t^{n-1}f(t,1)\,E\big(K_{t_i}\,K_t\big)\,dt$$

$$= a_n\int_0^1 t^{n-1}f(t,1)\,\big\{f(t_i,t)-f(t_i,1)f(t,1)\big\}\,dt.$$

By condition (6),

$$\Big| f(t,1)\,\big\{f(t_i,t)-f(t_i,1)f(t,1)\big\}\Big| \le \gamma\,\delta\,(\delta-1)\,(1-t)^{\alpha}.$$

Therefore, using (8) again, there is a constant b^{**} such that

$$\|P_D\big[(T,\,Y_nX_1)\in\cdot\big]-Q_{1,D}(\cdot)\| \le \frac{m\,r\,\gamma\,\delta\,(\delta-1)}{P(D)}\,a_n\int_0^1 t^{n-1}\,(1-t)^{\alpha}\,dt \le b^{**}\,\frac{a_n}{n^{1+\alpha}}.$$

Finally, since $a_n\le n^{1+\alpha}$ and $\beta\in(0,1)$, one also obtains

$$\|P_D\big[(T,\,Y_nX_1)\in\cdot\big]-Q_{1,D}(\cdot)\| \le b^{**}\left(\frac{a_n}{n^{1+\alpha}}\right)^{\beta}.$$

Thus, condition (10) holds, and this concludes the proof of Lemma 7.

4.2. **Lemma 7 implies Theorem 6.** First note that $\|\mu-\nu\|=\|\mu\circ h^{-1}-\nu\circ h^{-1}\|$ whenever μ and ν are probability laws on $\mathcal{B}(\mathbb{R}^{m+1})$ and $h:\mathbb{R}^{m+1}\to\mathbb{R}^{m+1}$ is bijective and Borel. Given a constant $b>0$, take

$$h(u,v)=(u,\,\frac{v}{\sqrt{b}})\quad\text{for } u\in\mathbb{R}^m \text{ and } v\in\mathbb{R}.$$

On noting that $Q_b\circ h^{-1}=Q_1$, one obtains

$$\|P\big[(T,\,\sqrt{b}\,I_n)\in\cdot\big]-Q_b(\cdot)\|=\|P\big[(T,\,I_n)\in\cdot\big]-Q_1(\cdot)\|.$$

Fix $\beta\in(0,1)$, $c_n>0$ and $c>0$. By Lemma 7,

$$\|P\big[(T,\,\sqrt{c_n}\,I_n)\in\cdot\big]-Q_{c_n}(\cdot)\|=\|P\big[(T,\,I_n)\in\cdot\big]-Q_1(\cdot)\|\le k\left(\frac{a_n}{n^{1+\alpha}}\right)^{\beta}$$

for some constant k and every $n\ge1$. By Lemma 3,

$$\|Q_{c_n}(\cdot)-Q_c(\cdot)\|\le E\Big\{\|\mathcal{N}(0,\,c_nX_1^2)-\mathcal{N}(0,\,cX_1^2)\|\Big\}\le\frac{|c_n-c|}{c}.$$

Therefore,

$$\|P\big[(T,\,\sqrt{c_n}\,I_n)\in\cdot\big]-Q_c(\cdot)\|\le\frac{|c_n-c|}{c}+k\left(\frac{a_n}{n^{1+\alpha}}\right)^{\beta}\quad\text{for all } n\ge1.$$

Next, suppose $\frac{a_n}{n^{1+\alpha}}\to0$ and $c_n=c=1$. We have to show that (T,I_n) converges $\|\cdot\|$-stably to $\delta_T\times N_1$. To this end, because of Lemma 1, we can restrict to those conditioning events F of the type

$$F=\big\{S\in A,\,X_1>x\big\}$$

where $S=(K_{s_1},\ldots,K_{s_p})$ for some $p\ge1$ and $s_1,\ldots,s_p\in[0,1]$, $A\in\mathcal{B}(\mathbb{R}^p)$ and $x\in\mathbb{R}\cup\{-\infty\}$. Take one such F, with $P(F)>0$, and write $D=\{X_1>x\}$. Then,

$$P_F\big[(T,\,I_n)\in\cdot\big]=\frac{P_D\big[S\in A,\,(T,\,I_n)\in\cdot\big]}{P(S\in A)}\quad\text{and}$$

$$E_F\big[\delta_T\times N_1(\cdot)\big]=\frac{E_D\big[1_{\{S\in A\}}\,\delta_T\times N_1(\cdot)\big]}{P(S\in A)}.$$

Let $T^*=(S,T)$ and $Q^*_{1,D}=P(T^*\in\cdot)\times E_D[N_1(\cdot)]$. By Lemma 7, applied with T^* and $Q^*_{1,D}$ in the place of T and $Q_{1,D}$, there is a constant k such that

$$\|P_F\big[(T,\,I_n)\in\cdot\big]-E_F\big[\delta_T\times N_1(\cdot)\big]\|\le\frac{\|P_D\big[(T^*,\,I_n)\in\cdot\big]-Q^*_{1,D}(\cdot)\|}{P(S\in A)}\le\frac{k\left(\frac{a_n}{n^{1+\alpha}}\right)^{\beta}}{P(S\in A)}.$$

Hence, (T, I_n) converges $\|\cdot\|$-stably to $\delta_T \times N_1$ since $\frac{a_n}{n^{1+\alpha}} \to 0$.

Finally, suppose $\frac{a_n}{n^{1+\alpha}} \to 0$ and $c_n \to c$. Fix $F \in \mathcal{F}^+$. Arguing as above,

$$\|P_F\big[(T,\, \sqrt{c_n}\, I_n) \in \cdot\big] - E_F\big[\delta_T \times N_{c_n}(\cdot)\big]\| = \|P_F\big[(T,\, I_n) \in \cdot\big] - E_F\big[\delta_T \times N_1(\cdot)\big]\|.$$

Thus,

$$\begin{aligned}
&\|P_F\big[(T,\, \sqrt{c_n}\, I_n) \in \cdot\big] - E_F\big[\delta_T \times N_c(\cdot)\big]\| \le \\
&\le \|P_F\big[(T,\, I_n) \in \cdot\big] - E_F\big[\delta_T \times N_1(\cdot)\big]\| + \|E_F\big[\delta_T \times N_{c_n}(\cdot)\big] - E_F\big[\delta_T \times N_c(\cdot)\big]\| \\
&\le \|P_F\big[(T,\, I_n) \in \cdot\big] - E_F\big[\delta_T \times N_1(\cdot)\big]\| + \frac{|c_n - c|}{c} \to 0.
\end{aligned}$$

Therefore, $(T,\, \sqrt{c_n}\, I_n)$ converges $\|\cdot\|$-stably to $\delta_T \times N_c$, and this concludes the proof of the implication "Lemma 7 $\Rightarrow$ Theorem 6".

4.3. **Theorem 6 implies Corollary 8.** Define $c = H\,\Gamma(2H)$, $a(H) = 1/2 - |1/2 - H|$ and recall that $\alpha = \min(1,\, 2H)$ and $A_n = \frac{n^{1+H}}{a_n}\, I_n$. In view of Theorem 6, it suffices to show that

$$\Big(\frac{n^{1+H}}{a_n}\Big)^2 = c\, +\, \mathrm{O}\big(n^{-a(H)}\big). \tag{11}$$

To prove (11), first note that

$$\begin{aligned}
a_n^{-2} = E\big[\big(\int_0^1 t^{n-1} f(t,1)\, K_t\, dt\big)^2\big] &= E\Big(\int_0^1 \int_0^1 (s\,t)^{n-1} f(s,1)\, f(t,1)\, K_s\, K_t\, ds\, dt\Big) \\
&= \int_0^1 \int_0^1 (s\,t)^{n-1} f(s,1)\, f(t,1)\, E(K_s\, K_t)\, ds\, dt \\
&= \int_0^1 \int_0^1 (s\,t)^{n-1} f(s,1)\, f(t,1)\, \big\{f(s,t) - f(s,1) f(t,1)\big\}\, ds\, dt.
\end{aligned}$$

On the other hand,

$$f(s,t) - f(s,1)f(t,1) = f(s,t) - f(s,1) - f(t,1) + 1 - (1 - f(s,1))(1 - f(t,1))$$

and conditions (6) and (8) yield

$$\begin{aligned}
&\int_0^1 \int_0^1 (s\,t)^{n-1} f(s,1)\, f(t,1)\, (1 - f(s,1))(1 - f(t,1))\, ds\, dt \\
&= \Big(\int_0^1 t^{n-1} f(t,1)\, (1 - f(t,1))\, dt\Big)^2 \le \gamma^2 \Big(\int_0^1 t^{n-1} (1-t)^\alpha\, dt\Big)^2 \le b\, \gamma^2\, n^{-2-2\alpha}
\end{aligned}$$

for some constant b. Similarly,

$$\begin{aligned}
&\Big|\int_0^1 \int_0^1 (s\,t)^{n-1} \big\{f(s,1)f(t,1) - 1\big\}\, \big\{f(s,t) - f(s,1) - f(t,1) + 1\big\}\, ds\, dt\Big| \\
&\le 2\,\gamma \int_0^1 \int_0^1 (s\,t)^{n-1}\, |(1 - f(s,1)) - f(s,1)(f(t,1) - 1)|\, (1-t)^\alpha\, ds\, dt \\
&\le 2\,\gamma^2 \int_0^1 \int_0^1 (s\,t)^{n-1}\, \big\{(1-s)^\alpha + (1-t)^\alpha\big\}\, (1-t)^\alpha\, ds\, dt \le b^*\, \gamma^2\, n^{-2-2\alpha}
\end{aligned}$$

for some constant b^*. It follows that

$$a_n^{-2} = \int_0^1 \int_0^1 (s\,t)^{n-1} \big\{ f(s,t) - f(s,1) - f(t,1) + 1 \big\}\, ds\, dt \,+\, \mathrm{O}(n^{-2-2\alpha})$$

$$= (1/2) \int_0^1 \int_0^1 (s\,t)^{n-1} \big\{ (1-s)^{2H} + (1-t)^{2H} - |s-t|^{2H} \big\}\, ds\, dt \,+\, \mathrm{O}(n^{-2-2\alpha})$$

$$= \int_0^1 t^{n-1}\, dt \int_0^1 s^{n-1} (1-s)^{2H}\, ds - \int_0^1 t^{n-1} \int_0^t s^{n-1} (t-s)^{2H}\, ds\, dt \,+\, \mathrm{O}(n^{-2-2\alpha})$$

$$= \frac{(n-1)!\,\Gamma(2H+1)}{n\,\Gamma(n+2H+1)} - \int_0^1 t^{2n+2H-1}\, dt \int_0^1 s^{n-1}(1-s)^{2H}\, ds \,+\, \mathrm{O}(n^{-2-2\alpha})$$

$$= \frac{(n-1)!\,\Gamma(2H+1)}{n\,\Gamma(n+2H+1)} - \frac{(n-1)!\,\Gamma(2H+1)}{(2n+2H)\,\Gamma(n+2H+1)} \,+\, \mathrm{O}(n^{-2-2\alpha}).$$

On noting that $\Gamma(2H+1) = 2\,c$ and $a(H) = \alpha - H$, one finally obtains

$$n^{2+2H} a_n^{-2} = n^{2+2H} \frac{(n-1)!\,\Gamma(2H+1)}{\Gamma(n+2H+1)} \,\frac{n+2H}{2\,n\,(n+H)} \,+\, \mathrm{O}(n^{-2(\alpha-H)})$$

$$= c\, n^{2H} \frac{n!}{\Gamma(n+2H+1)} \,\frac{n+2H}{n+H} \,+\, \mathrm{O}(n^{-2a(H)})$$

$$= \frac{c\, n^{2H}\, n!}{\Gamma(n+2H+1)} + \frac{c\, H\, n^{2H}\, n!}{(n+H)\,\Gamma(n+2H+1)} \,+\, \mathrm{O}(n^{-2a(H)})$$

$$= c \,+\, c\Big(\frac{n^{2H}\, n!}{\Gamma(n+2H+1)} - 1 \Big) \,+\, \mathrm{O}(n^{-1}) \,+\, \mathrm{O}(n^{-2a(H)})$$

$$= c \,+\, \mathrm{O}(n^{-1}) \,+\, \mathrm{O}(n^{-2a(H)}) = c \,+ \mathrm{O}(n^{-2a(H)}).$$

This concludes the proof of (11). Thus, Corollary 8 is actually a consequence of Theorem 6.

5. Concluding remarks

The techniques of Section 4 yield something more than the results stated in Section 3. The latter could be actually generalized as follows.

- $T = (K_{t_1}, \ldots, K_{t_m})$ could be replaced by

  $$T = \Big(\int K_t\, \nu_1(dt), \ldots, \int K_t\, \nu_m(dt) \Big)$$

  where $\nu_1, \ldots, \nu_m$ are probability measures on $\mathcal{B}([0,1])$. In this way, even if in a different framework, it is possible to obtain results formally analogous to [2], [3] and [5, Sections 5-6].
- I_n could be replaced by

  $$I_n = \frac{a_n}{2} \int_0^1 t^{n-1} (X_1^2 - X_t^2)\, \{1 + g(K_t)\}\, dt$$

  where $g : \mathbb{R} \to \mathbb{R}$ is locally Lipschitz with at most exponential growth and $g(0) = 0$. In particular, if g is a polynomial, the rate of convergence of I_n is not affected by the degree of g. Using Malliavin calculus, instead, such a rate depends on the degree of g.

- Up to replacing $N_c = \mathcal{N}\big(0,\, cX_1^2\big)$ with $\mathcal{N}\big(0,\, cX_1^{2p-2}\big)$, where $p \geq 2$ is any integer, I_n could be replaced by

$$I_n = \frac{a_n}{p} \int_0^1 t^{n-1} (X_1^p - X_t^p)\, dt.$$

We close the paper by making precise a couple of points raised in the previous sections. We begin with the conjecture stated after Theorem 6.

Example 9. Let $\mathcal{C}$ be the set of real continuous functions on $[0,1]$, equipped with uniform distance, and Q the product probability measure on $\mathcal{B}\big(\mathcal{C} \times \mathbb{R}\big)$ given by

$$Q = P(K \in \cdot) \times E[N_1(\cdot)].$$

Suppose $\frac{a_n}{n^{1+\alpha}} \to 0$. Toward a contradiction, suppose also that $P\big[(K,\, I_n) \in \cdot\big]$ converges to Q in total variation distance.

Using Lemma 5 and arguing as in the proof of Lemma 7,

$$\|P\big[(K,\, I_n) \in \cdot\big] - P\big[(K,\, Y_n X_1) \in \cdot\big]\| \to 0.$$

Hence, $P\big[(K,\, Y_n X_1) \in \cdot\big]$ also converges to Q in total variation. Write $\mathcal{K} = \sigma(K)$ and take a standard normal random variable U, independent of (K, X_1), and a bounded non-negative Borel function $h : \mathbb{R} \to \mathbb{R}$. On noting that $(K,\, UX_1) \sim Q$,

$$\begin{aligned}
&\sup_{A \in \mathcal{B}(\mathcal{C})} \Big| \int 1_A(K) \big\{ E[h(Y_n X_1) \mid \mathcal{K}] - E[h(UX_1) \mid \mathcal{K}] \big\}\, dP \Big| \\
&\quad = \sup_{A \in \mathcal{B}(\mathcal{C})} \Big| \int 1_A(K) \big\{ h(Y_n X_1) - h(UX_1) \big\}\, dP \Big| \\
&\quad \leq \sup h\, \|P\big[(K,\, Y_n X_1) \in \cdot\big] - Q(\cdot)\| \to 0.
\end{aligned}$$

Hence,

$$\lim_n E \Big| E\big[h(Y_n X_1) \mid \mathcal{K}\big] - E\big[h(UX_1) \mid \mathcal{K}\big] \Big| = 0.$$

On the other hand, letting $h_b(x) = |x|\, 1_{[-b,b]}(x)$, one obtains

$$\begin{aligned}
E\Big| E\big[|Y_n X_1| \mid \mathcal{K}\big] - E\big[|UX_1| \mid \mathcal{K}\big] \Big| &\leq E\Big| E\big[h_b(Y_n X_1) \mid \mathcal{K}\big] - E\big[h_b(UX_1) \mid \mathcal{K}\big] \Big| + \\
&\quad + E\Big(|Y_n X_1|\, 1_{\{|Y_n X_1| > b\}} \Big) + E\Big(|UX_1|\, 1_{\{|UX_1| > b\}} \Big) \\
&= E\Big| E\big[h_b(Y_n X_1) \mid \mathcal{K}\big] - E\big[h_b(UX_1) \mid \mathcal{K}\big] \Big| + 2\, E\Big(|UX_1|\, 1_{\{|UX_1| > b\}} \Big)
\end{aligned}$$

for all $b > 0$. This fact implies

$$\lim_n E \Big| E[|Y_n X_1| \mid \mathcal{K}] - E[|UX_1| \mid \mathcal{K}] \Big| = 0.$$

But this is a contradiction. In fact, Y_n is $\mathcal{K}$-measurable, and thus

$$E\big[|Y_n X_1| \mid \mathcal{K}\big] = |Y_n|\, E[|X_1|] \quad \text{and} \quad E\big[|UX_1| \mid \mathcal{K}\big] = E[|UX_1|] \quad \text{a.s.}$$

To sum up, $P\big[(K,\, I_n) \in \cdot\big]$ does not converge to Q in total variation. As a consequence, $(K,\, I_n)$ does not converge $\|\cdot\|$-stably to $\delta_K \times N_1$.

Finally, in (2), we claimed that $\|P(A_n \in \cdot) - E[N_c(\cdot)]\|$ is not $\mathrm{O}\big(n^{-a(H)-\epsilon}\big)$ for every fixed $\epsilon > 0$. To prove this fact, in view of Lemma 2, it suffices to see that

$$\Big|E(A_n) - E\Big(\int u\,N_c(du)\Big)\Big| = |E(A_n)| \geq b\,n^{-a(H)}$$

for some constant $b > 0$ and every $n \geq 1$. In fact, if $H \geq 1/2$, then

$$E(A_n) = \frac{n^{1+H}}{2}\int_0^1 t^{n-1}E(B_1^2 - B_t^2)\,dt = \frac{n^{1+H}}{2}\int_0^1 t^{n-1}(1-t^{2H})\,dt$$
$$= \frac{H\,n^H}{n+2H} \geq \frac{H}{1+2H}\,n^{-(1-H)} = \frac{H}{1+2H}\,n^{-a(H)}.$$

If $H < 1/2$, the proof of (2) is not straightforward and we omit the explicit calculations. We just note that, if $D = \{B_1 > x\}$ for some $x \in (1,\infty)$, one obtains

$$\Big|E_D(A_n) - E_D\Big(\int u\,N_c(du)\Big)\Big| = |E_D(A_n)| \geq b_x\,n^{-H} = b_x\,n^{-a(H)}$$

for every $H \in (0,\,1/2)$ and some constant $b_x > 0$.

References

[1] Cheridito P., Nualart D. (2005) Stochastic integral of divergence type with respect to fractional Brownian motion with Hurst parameter $H \in (0,\,1/2)$, *Ann. I.H.P.*, 41, 1049-1081.

[2] Nourdin I., Peccati G. (2008) Weighted variations of iterated Brownian motion, *Elect. J. Probab.*, 13, 1229-1256.

[3] Nourdin I., Nualart D., Tudor C.A. (2010) Central and non-central limit theorems for weighted power variations of fractional Brownian motion, *Ann. I.H.P.*, 46, 1055-1079.

[4] Nourdin I., Peccati G. (2012) *Normal approximations with Malliavin calculus: From Stein's method to universality*, Cambridge University Press.

[5] Nourdin I., Nualart D., Peccati G. (2016) Quantitative stable limit theorems on the Wiener space, *Ann. Probab.*, 44, 1-41.

[6] Peccati G., Yor M. (2004) Four limit theorems for quadratic functionals of Brownian motion and Brownian bridge, In: *Asymptotic Methods in Stochastics, AMS, Fields Institute Communication Series*, 75-87.

[7] Renyi A. (1963) On stable sequences of events, *Sankhya* A, 25, 293-302.

Luca Pratelli, Accademia Navale, viale Italia 72, 57100 Livorno, Italy
E-mail address: pratel@mail.dm.unipi.it

Pietro Rigo (corresponding author), Dipartimento di Matematica "F. Casorati", Universita' di Pavia, via Ferrata 1, 27100 Pavia, Italy
E-mail address: pietro.rigo@unipv.it