Couplings and Strong Approximations to Time Dependent Empirical Processes Based on I.I.D. Fractional Brownian Motions

Péter Kevei* David M. Mason†

Abstract

We define a time dependent empirical process based on n i.i.d. fractional Brownian motions and establish Gaussian couplings and strong approximations to it by Gaussian processes. They lead to functional laws of the iterated logarithm for this process.
Keywords: coupling inequality; fractional Brownian motion; strong approximation; time dependent empirical process.

MSC2010: 62E17, 60G22, 60F15

1 Introduction

The aim in this paper is to derive Gaussian couplings and strong approximations to time dependent empirical processes based on n independent sample continuous fractional Brownian motions, as defined in Subsection 2.1. Our couplings yield surprisingly close *almost sure* approximations of our empirical processes by Gaussian processes defined on sequences of intervals for which weak convergence cannot hold in the limit. As an example of what our strong approximations can do, we show that functional laws of the iterated logarithm [FLIL] for these empirical processes can be derived from those that are known for Gaussian processes.

Our investigations may be thought of as a continuation of those of Kuelbs, Kurtz and Zinn [13], who proved central limit theorems for time dependent empirical processes based on n independent copies of a wide variety of random processes. These include certain self-similar processes of which fractional Brownian motion is a special case. Our results reveal the kind of strong limit theorems that are possible when one turns to the detailed analysis of time dependent empirical processes based on processes which have a fine local random structure, such as fractional Brownian motion.

Kuelbs and Zinn [14, 15] have obtained central limit theorems for a time dependent quantile process based on n independent copies of a wide variety of random processes. In the process they generalized a result of Swanson [25], who used classical weak convergence theory to prove that an appropriately scaled median of n independent Brownian motions converges weakly to a mean

*MTA-SZTE Analysis and Stochastics Research Group, Bolyai Institute, Aradi vértanúk tere 1, 6720 Szeged, Hungary, and Center for Mathematical Sciences, Technische Universität München, Boltzmannstraße 3, 85748 Garching, Germany, e-mail: kevei@math.u-szeged.hu

†Department of Applied Economics and Statistics, University of Delaware, 213 Townsend Hall, Newark, DE 19716, USA, e-mail: davidm@udel.edu

zero Gaussian process. In a sequel to this paper we use the results in the present work to derive strong approximations and FLILs for quantile processes or inverses of these time dependent empirical processes based on n i.i.d. sample continuous fractional Brownian motions. For details see Kevei and Mason [10].

To motivate our work, we point out some implications of a coupling and a strong approximation due to Komlós, Major and Tusnády (KMT) [11]. Let $X_1, X_2, \ldots,$ be i.i.d. F. For each $n \geq 1$, let

$$F_n(x) = n^{-1} \sum_{j=1}^{n} 1\{X_j \leq x\}, \ x \in \mathbb{R},$$

denote the empirical distribution function based on $X_1, \ldots, X_n$, and define the empirical process

$$v_n(x) := \sqrt{n}\{F_n(x) - F(x)\}, \ x \in \mathbb{R}.$$

Using the coupling result given in Theorem 3 of KMT [11] one can construct a probability space on which sit an i.i.d. F sequence X_1, X_2, ..., and a sequence of Brownian bridges $B_1, B_2, \ldots,$ on $[0, 1]$ such that

$$\|v_n - B_n(F)\|_{\mathbb{R}} = O\left(\frac{\log n}{\sqrt{n}}\right), \text{ a.s.,} \tag{1.1}$$

where for a real-valued function Υ defined on a set S we use the notation

$$\|\Upsilon\|_S = \sup_{s \in S} |\Upsilon(s)|. \tag{1.2}$$

The rate $\log n/\sqrt{n}$ in (1.1) is optimal.

Further, by the strong approximation result stated in Theorem 4 of KMT [11] one has on the same probability space an i.i.d. F sequence X_1, X_2, ..., and a sequence of *independent* Brownian bridges $B_1, B_2, \ldots,$ on $[0, 1]$ such that

$$\left\| v_n - \frac{\sum_{j=1}^{n} B_j(F)}{\sqrt{n}} \right\|_{\mathbb{R}} = O\left(\frac{(\log n)^2}{\sqrt{n}}\right), \text{ a.s.} \tag{1.3}$$

It is known that the $n^{-1/2}$ part of the rate in (1.3) is optimal, but not the $(\log n)^2$. It has long been conjectured that the $(\log n)^2$ in (1.3) can be replaced by $\log n$. This is one of the rare cases where any such optimality is known in the rate of strong approximation to an empirical process.

Our goal is to develop analogs of (1.1) and (1.3) for the time dependent empirical processes based on independent copies of sample continuous fractional Brownian motion. These are described in the next section. The rates of coupling and strong approximation that we obtain are unlikely to be anywhere near optimal in the sense just described, however they will be seen to be sufficient to derive from them FLILs for our time dependent empirical processes. We find it noteworthy that useful couplings and strong approximations can be obtained for the kind of complexly formed empirical processes that we consider. Our main results are detailed in Section 2 and they are proved in Section 3. We gather together some needed facts in the Appendix. To prove our main results we use the methodology outlined in Berthet and Mason [3].

2 Coupling and strong approximation to a time dependent empirical process

2.1 A time dependent empirical process

Let $\left\{B^{(H)}\right\} \cup \left\{B_j^{(H)}\right\}_{j\geq 1}$ be a sequence of i.i.d. sample continuous fractional Brownian motions with Hurst index $0 < H < 1$ defined on $[0,\infty)$. Note that $B^{(H)}$ is a continuous mean zero Gaussian process on $[0,\infty)$ with covariance function defined for any $s,t \in [0,\infty)$

$$E\left(B^{(H)}(s)\, B^{(H)}(t)\right) = \frac{1}{2}\left(|s|^{2H} + |t|^{2H} - |s-t|^{2H}\right).$$

By the Lévy modulus of continuity theorem for sample continuous fractional Brownian motion $B^{(H)}$ with Hurst index $0 < H < 1$ (see (3.1) below), we have for any $0 < T < \infty$, w.p. 1,

$$\sup_{0\leq s\leq t\leq T} \frac{\left|B^{(H)}(t) - B^{(H)}(s)\right|}{f_H(t-s)} =: L < \infty, \tag{2.1}$$

where for $u \geq 0$

$$f_H(u) = u^H \sqrt{1 \vee \log u^{-1}} \tag{2.2}$$

and $a \vee b = \max\{a,b\}$. We shall take versions of $\left\{B^{(H)}\right\} \cup \left\{B_j^{(H)}\right\}_{j\geq 1}$ such that (2.1) holds for all of their trajectories.

For any $t \in [0,\infty)$ and $x \in \mathbb{R}$ let $F(t,x) = P\left\{B^{(H)}(t) \leq x\right\}$. Note that

$$F(t,x) = \Phi\left(x/t^H\right), \tag{2.3}$$

where $\Phi(x) = P\{Z \leq x\}$, with Z being a standard normal random variable. For any $n \geq 1$ define the time dependent *empirical distribution function*

$$F_n(t,x) = n^{-1}\sum_{j=1}^{n} 1\left\{B_j^{(H)}(t) \leq x\right\}.$$

Applying Theorem 5 in [13] (also see their Remark 8) one can show for any choice of $0 < \gamma \leq 1 < T < \infty$ that the time dependent *empirical process* indexed by $(t,x) \in \mathcal{T}(\gamma)$,

$$v_n(t,x) = \sqrt{n}\left\{F_n(t,x) - F(t,x)\right\},$$

where

$$\mathcal{T}(\gamma) := [\gamma, T] \times \mathbb{R},$$

converges weakly to a uniformly continuous centered Gaussian process $G(t,x)$ indexed by $(t,x) \in \mathcal{T}(\gamma)$, whose trajectories are bounded, having covariance function

$$\begin{aligned} &E\left(G(s,x)\, G(t,y)\right) \\ &= P\left\{B^{(H)}(s) \leq x, B^{(H)}(t) \leq y\right\} - P\left\{B^{(H)}(s) \leq x\right\} P\left\{B^{(H)}(t) \leq y\right\}. \end{aligned} \tag{2.4}$$

Keeping in mind that $\mathcal{T}(\gamma)$ is equipped with the semimetric

$$\rho\left((s,x),(t,y)\right)=\sqrt{E\left(G\left(s,x\right)-G\left(t,y\right)\right)^{2}}, \tag{2.5}$$

we see by weak convergence that $\mathcal{T}(\gamma)$ is totally bounded and thus separable in the topology induced by this semimetric ρ. Moreover its completion $\mathcal{T}^{c}(\gamma)$ in this topology is compact. Since G is bounded and uniformly continuous on $\mathcal{T}(\gamma)$ it can be extended uniquely to be bounded and uniformly continuous on $\mathcal{T}^{c}(\gamma)$.

Remark 1 *To see how this is done, notice that for each* $t\in[\gamma,T]$*, both* $\{(t,-m)\}_{m\geq1}$ *and* $\{(t,m)\}_{m\geq1}$ *are Cauchy sequences in* $\mathcal{T}(\gamma)$ *with respect to the semimetric* ρ*. Also by the boundedness and uniform continuity of* G *on* $\mathcal{T}(\gamma)$*, the sequences* $\{G(t,-m)\}_{m\geq1}$ *and* $\{G(t,m)\}_{m\geq1}$ *are also bounded Cauchy sequences in* $\mathbb{R}$*. Furthermore, both* $EG^{2}(t,-m)\to0$ *and* $EG^{2}(t,m)\to0$*, as* $m\to\infty$*. Thus we can unambiguously define* $(t,-\infty)$ *as the limit of the sequence* $(t,-m)$ *as* $m\to\infty$ *and* $G(t,-\infty)=0$*, w.p. 1, and* (t,∞) *as the limit of the sequence* (t,m) *as* $m\to\infty$ *and* $G(t,\infty)=0$*, w.p. 1. We see that for any* $t\in[\gamma,T]$ *and* $(s,y)\in\mathcal{T}(\gamma)$*,*

$$\rho\left((t,\pm\infty),(s,y)\right)=\sqrt{E\left(G((t,\pm\infty))-G\left(s,y\right)\right)^{2}}=\sqrt{EG^{2}\left(s,y\right)}$$

and for $s,t\in[\gamma,T]$

$$\rho\left((t,\pm\infty),(s,\pm\infty)\right)=\sqrt{E\left(G((t,\pm\infty))-G\left(s,\pm\infty\right)\right)^{2}}=0.$$

With these definitions ρ *becomes a semimetric on* $[\gamma,T]\times(\mathbb{R}\cup\{-\infty,\infty\})$*. Next consider* $[\gamma,T]\times\{-\infty,\infty\}$ *as an equivalence class, i.e.* $(t,\pm\infty)\sim(s,\pm\infty)$*, whenever* $\rho\left((t,\pm\infty),(s,\pm\infty)\right)=0$*, which always happens, and denote it by* ω *and with some abuse of the previous notation write* $G(\omega)=0$*,* $\rho(\omega,\omega)=0$ *and for any* $(s,y)\in\mathcal{T}(\gamma)$*,* $\rho(\omega,(s,y))=\sqrt{EG^{2}(s,y)}$*, and let* ρ *remain as it was previously defined on* $\mathcal{T}(\gamma)\times\mathcal{T}(\gamma)$*. We define the completion of* $\mathcal{T}^{c}(\gamma)=([\gamma,T]\times\mathbb{R})\cup\{\omega\}$*, which is readily shown to be a complete metric space with semimetric* ρ*.*

Therefore we can consider G as a Gaussian process taking values in the separable Banach space consisting of the continuous functions in the sup-norm on the compact metric space $\mathcal{T}^{c}(\gamma)$. For later use we point out that by Proposition 1 on page 26 of Lifshits [20] we can assume that the Gaussian process $G(t,x)$ is separable.

For future reference we record here that for some finite positive constant $M(\gamma,T,H)$ for all $n\geq1$

$$E\left\|v_{n}\right\|_{\mathcal{T}(\gamma)}\leq M\left(\gamma,T,H\right). \tag{2.6}$$

Assertion (2.6) follows from an application of the Hoffmann–Jørgensen inequality, cf. Ledoux and Talagrand [18], page 156. For the argument see, for instance, Lemma 3.1 of Einmahl and Mason [8].

We restrict ourselves to positive γ, since in Section 8.1 of [13] it is pointed out that the empirical process $v_{n}(t,x)$ indexed by $\mathcal{T}(0):=[0,T]\times\mathbb{R}$ does not converge weakly to a uniformly continuous centered Gaussian process indexed by $(t,x)\in\mathcal{T}(0)$, whose trajectories are bounded. More generally in the sequel, $G(t,x)$ denotes a centered Gaussian process on $\mathcal{T}(0)$ with covariance (2.4) that is uniformly continuous on $\mathcal{T}(\gamma)$ with bounded trajectories for any $0<\gamma\leq1<T<\infty$.

We shall also be using the following empirical process indexed by function notation. Let $X, X_1, X_2, \ldots,$ be i.i.d. random variables from a probability space $(\Omega, \mathcal{A}, P)$ to a measurable space $(S, \mathcal{S})$. Consider an empirical process indexed by a class $\mathcal{G}$ of bounded measurable real valued functions on $(S, \mathcal{S})$ defined by

$$\alpha_n(\varphi) := \sqrt{n}(P_n - P)\varphi = \frac{\sum_{i=1}^n \varphi(X_i) - nE\varphi(X)}{\sqrt{n}}, \ \varphi \in \mathcal{G},$$

where

$$P_n(\varphi) = n^{-1}\sum_{i=1}^n \varphi(X_i) \text{ and } P(\varphi) = E\varphi(X).$$

Keeping this notation in mind, let $\mathcal{C}[0,T]$ be the class of continuous functions g on $[0,T]$ endowed with the topology of uniform convergence and where $\mathcal{B}[0,T]$ denotes the Borel subsets of $\mathcal{C}[0,T]$. Define this subclass of $\mathcal{C}[0,T]$

$$\mathcal{C}_\infty := \left\{ g : \ \sup\left\{ \frac{|g(s) - g(t)|}{f_H(|s-t|)}, \ 0 \le s, t \le T \right\} < \infty \right\}. \tag{2.7}$$

Further, let $\mathcal{F}_{(\gamma,T)}$ be the class of functions of $g \in \mathcal{C}[0,T] \to \mathbb{R}$, indexed by $(t,x) \in \mathcal{T}(\gamma)$, of the form

$$h_{t,x}(g) = 1\{g(t) \le x, g \in \mathcal{C}_\infty\}.$$

Here we permit $\gamma = 0$. Since by (2.1) we can assume that each $B^{(H)}$, $B_j^{(H)}$, $j \ge 1$, is in $\mathcal{C}_\infty$, we see that for any $h_{t,x} \in \mathcal{F}_{(\gamma,T)}$,

$$\alpha_n(h_{t,x}) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \left(1\left\{B_i^{(H)}(t) \le x\right\} - P\left\{B^{(H)}(t) \le x\right\}\right) = v_n(t,x).$$

We shall be using the notation $\alpha_n(h_{t,x})$ and $v_n(t,x)$ interchangeably.

Let $\mathbb{G}_{(\gamma,T)}$ denote the mean zero Gaussian process indexed by $\mathcal{F}_{(\gamma,T)}$, having covariance function defined for $h_{s,x}, h_{t,y} \in \mathcal{F}_{(\gamma,T)}$

$$E\left(\mathbb{G}_{(\gamma,T)}(h_{s,x})\,\mathbb{G}_{(\gamma,T)}(h_{t,y})\right) = P\left\{B^{(H)}(s) \le x, B^{(H)}(t) \le y, B^{(H)} \in \mathcal{C}_\infty\right\}$$
$$-P\left\{B^{(H)}(s) \le x, B^{(H)} \in \mathcal{C}_\infty\right\} P\left\{B^{(H)}(t) \le y, B^{(H)} \in \mathcal{C}_\infty\right\},$$

which since $P\left\{B^{(H)} \in \mathcal{C}_\infty\right\} = 1$,

$$= E\left(G(s,x)\,G(t,y)\right),$$

i.e. $\mathbb{G}_{(\gamma,T)}(h_{t,x})$ defines a probabilistically equivalent version of the Gaussian process $G(t,x)$ for $(t,x) \in \mathcal{T}(\gamma)$. We shall say that a process $\widetilde{\mathcal{Y}}$ is a *probabilistically equivalent version* of $\mathcal{Y}$ if $\widetilde{\mathcal{Y}} \stackrel{\mathrm{D}}{=} \mathcal{Y}$.

Notice that in this notation

$$\begin{aligned}\rho((s,x),(t,y)) &= \sqrt{E\left(\mathbb{G}_{(\gamma,T)}(h_{s,x}) - \mathbb{G}_{(\gamma,T)}(h_{t,y})\right)^2} \\ &= \sqrt{Var\left(h_{s,x}\left(B^{(H)}\right) - h_{t,y}\left(B^{(H)}\right)\right)} \\ &\le \sqrt{E\left(h_{s,x}\left(B^{(H)}\right) - h_{t,y}\left(B^{(H)}\right)\right)^2} =: d_P(h_{s,x}, h_{t,y}).\end{aligned}$$

More generally, for suitable functions f and g we shall write

$$d_P(f,g) = \sqrt{E\left(f\left(B^{(H)}\right) - g\left(B^{(H)}\right)\right)^2}. \tag{2.8}$$

The proofs of a number our results rely on a lemma of Berkes and Philipp [2], which for the convenience of the reader we state here.

Lemma A1 of Berkes and Philipp (1979) *Let $S_i, i = 1,2,3$ be Polish spaces. Let $\mathbf{F}$ be a distribution on $S_1 \times S_2$ and $\mathbf{G}$ be a distribution on $S_2 \times S_3$ such that the second marginal of $\mathbf{F}$ is equal to the first marginal of $\mathbf{G}$. Then there exists a probability space and a random vector (Z_1, Z_2, Z_3) defined on it taking its values in $S_1 \times S_2 \times S_3$ such that (Z_1, Z_2) has distribution $\mathbf{F}$ and (Z_2, Z_3) has distribution $\mathbf{G}$.*

2.2 Our main coupling and strong approximation results for α_n

In the results that follow

$$\nu_0 = 2 + \frac{2}{H} \quad \text{and} \quad H_0 = 1 + H. \tag{2.9}$$

We have the following Gaussian coupling to the empirical process α_n indexed by $\mathcal{F}_{(\gamma_n,T)}$.

Proposition 1 *As long as $0 < \gamma_n \leq 1$ satisfies for some $0 \leq \eta < \frac{1}{5H_0}$,*

$$\infty > -\frac{\log \gamma_n}{\log n} \to \eta, \quad \textit{as } n \to \infty, \tag{2.10}$$

for every $\lambda > 1$ there exists a $\rho(\lambda) > 0$ such that for each integer n large enough one can construct on the same probability space random vectors $B_1^{(H)}$, ..., $B_n^{(H)}$ i.i.d. $B^{(H)}$ and a probabilistically equivalent version $\widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}$ of $\mathbb{G}_{(\gamma_n,T)}$ such that,

$$P\left\{\left\|\alpha_n - \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_{(\gamma_n,T)}} > \rho(\lambda)\,(\log n)^{\tau_2}\left(n^{-1/2}\gamma_n^{-5H_0/2}\right)^{2/(2+5\nu_0)}\right\} \leq n^{-\lambda}, \tag{2.11}$$

where $\tau_2 = (19H + 25)/(24H + 20)$ and ν_0 is defined in (2.9). Moreover, in particular, when $\gamma_n = n^{-\eta}$, with $0 \leq \eta < \frac{1}{5H_0}$,

$$P\left\{\left\|\alpha_n - \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_{(\gamma_n,T)}} > \rho(\lambda)\, n^{-\tau_1}(\log n)^{\tau_2}\right\} \leq n^{-\lambda},$$

where $\tau_1 = \tau_1(\eta) = (1 - 5H_0\eta)\,/\,(2 + 5\nu_0) > 0$.

Remark 2 *Notice that Proposition 1 yields the coupling rate*

$$\left\|\alpha_n - \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_{(\gamma_n,T)}} = O_P\left((\log n)^{\tau_2}\left(n^{-1/2}\gamma_n^{-5H_0/2}\right)^{2/(2+5\nu_0)}\right). \tag{2.12}$$

In particular, for any $0 < H < 1$ and $0 < \eta < 1/(5H_0)$ the convergence (2.12) holds with $\gamma_n = n^{-\eta}$, since such γ_n satisfy (2.10). The convergence (2.12) is surprising in light of the results in Section 8.1 in [13], where it is pointed out that the empirical process $v_n(t,x)$ indexed by $[0,T] \times \mathbb{R}$ does not converge weakly to a uniformly continuous centered Gaussian process *indexed by $(t,x) \in [0,T] \times \mathbb{R}$ whose trajectories are bounded. Observe, however, by Theorem 5 in [13] for each $n \geq 1$ there is a version of Gaussian process $G_n(t,x) = \widetilde{\mathbb{G}}_{(\gamma_n,T)}(h_{t,x})$, which is uniformly continuous on $[\gamma_n, T] \times \mathbb{R}$ with bounded trajectories. We shall see that a coupling result following from a special case of Theorem 1.1 of Zaitsev [30] is crucial to establish (2.11) on intervals $[\gamma_n, T]$ such that γ_n goes to zero at the rate (2.10).*

For any $\kappa > 0$ let

$$\mathcal{G}(\kappa) = \{t^\kappa h_{t,x} : (t,x) \in [0,T] \times \mathbb{R}\}. \tag{2.13}$$

For $g \in \mathcal{G}(\kappa)$, with some abuse of notation, we shall write

$$\mathbb{G}_{(0,T)}(g) = t^\kappa \mathbb{G}_{(0,T)}(h_{t,x}). \tag{2.14}$$

Also, in analogy with (1.2), we set

$$\left\|\alpha_n - \mathbb{G}_{(0,T)}^{(n)}\right\|_{\mathcal{G}(\kappa)} := \sup\left\{\left|t^\kappa \alpha_n(h_{t,x}) - t^\kappa \mathbb{G}_{(0,T)}^{(n)}(h_{t,x})\right| : (t,x) \in [0,T] \times \mathbb{R}\right\}.$$

We get the following Gaussian coupling to the empirical process α_n indexed by $\mathcal{G}(\kappa)$.

Proposition 2 *For any $0 < \kappa < \infty$ and every $\lambda > 1$ there exists a $\rho'(\lambda) > 0$ such that for each integer n large enough one can construct on the same probability space random vectors $B_1^{(H)}, \ldots, B_n^{(H)}$ i.i.d. $B^{(H)}$ and a probabilistically equivalent version $\widetilde{\mathbb{G}}_{(0,T)}^{(n)}$ of $\mathbb{G}_{(0,T)}$ such that,*

$$P\left\{\left\|\alpha_n - \widetilde{\mathbb{G}}_{(0,T)}^{(n)}\right\|_{\mathcal{G}(\kappa)} > \rho'(\lambda)\, n^{-\tau_1'} (\log n)^{\tau_2}\right\} \leq n^{-\lambda}, \tag{2.15}$$

where τ_2 is as in Proposition 1 and $\tau_1' = \tau_1'(\kappa) = \kappa/(5H_0 + \kappa(2 + 5\nu_0))$.

Remark 3 *It is shown in Remark 6 that the Gaussian process indexed by $\mathcal{G}(\kappa)$*

$$t^\kappa \mathbb{G}_{(0,T)}(h_{t,x}) = t^\kappa G(t,x), (t,x) \in [0,T] \times \mathbb{R},$$

has a version that is uniformly continuous with bounded trajectories. Therefore Proposition 2 implies that for any $\kappa > 0$ the weighted empirical process $t^\kappa \alpha_n(h_{t,x}) = t^\kappa v_n(t,x)$, $(t,x) \in [0,T] \times \mathbb{R}$, converges weakly to $t^\kappa G(t,x)$. Recall, as pointed out in Remark 2, weak convergence fails if κ is chosen to be zero.

Propositions 1 and 2 lead to the following two strong approximation theorems.

Theorem 1 *As long as $1 \geq \ \gamma = \gamma_n > 0$ is constant, under the assumptions and notation of Proposition 1 for all $1/(2\tau_1(0)) < \alpha < 1/\tau_1(0)$ and $\xi > 1$ there exist a $\rho(\alpha,\xi) > 0$, a sequence of i.i.d. $B_1^{(H)}, B_2^{(H)}, \ldots,$ and a sequence of independent copies $\mathbb{G}_{(\gamma,T)}^{(1)}, \mathbb{G}_{(\gamma,T)}^{(2)}, \ldots,$ of $\mathbb{G}_{(\gamma,T)}$ sitting on the same probability space such that*

$$P\left\{\max_{1\leq m\leq n}\left\|\sqrt{m}\alpha_m - \sum_{i=1}^m \mathbb{G}_{(\gamma,T)}^{(i)}\right\|_{\mathcal{F}_{(\gamma,T)}} > \rho(\alpha,\xi)\, n^{1/2-\tau(\alpha)} (\log n)^{\tau_2}\right\} \leq n^{-\xi}$$

and

$$\max_{1\leq m\leq n}\left\|\sqrt{m}\alpha_m - \sum_{i=1}^m \mathbb{G}_{(\gamma,T)}^{(i)}\right\|_{\mathcal{F}_{(\gamma,T)}} = O\left(n^{1/2-\tau(\alpha)} (\log n)^{\tau_2}\right), \text{ a.s.,}$$

where $\tau(\alpha) = (\alpha\tau_1(0) - 1/2)/(1+\alpha) > 0$.

Theorem 2 *Under the assumptions and notation of Proposition 2 for any $\kappa > 0$, for all $1/(2\tau_1') < \alpha < 1/\tau_1'$, and $\xi > 1$ there exist a $\rho'(\alpha, \xi) > 0$, a sequence of i.i.d. $B_1^{(H)}, B_2^{(H)}, \ldots,$ and a sequence of independent copies $\mathbb{G}_{(0,T)}^{(1)}, \mathbb{G}_{(0,T)}^{(2)}, \ldots,$ of $\mathbb{G}_{(0,T)}$ sitting on the same probability space such that*

$$P\left\{\max_{1\le m\le n}\left\|\sqrt{m}\alpha_m - \sum_{i=1}^{m}\mathbb{G}_{(0,T)}^{(i)}\right\|_{\mathcal{G}(\kappa)} > \rho'(\alpha,\xi)\, n^{1/2-\tau'(\alpha)}(\log n)^{\tau_2}\right\} \le n^{-\xi}$$

and

$$\max_{1\le m\le n}\left\|\sqrt{m}\alpha_m - \sum_{i=1}^{m}\mathbb{G}_{(0,T)}^{(i)}\right\|_{\mathcal{G}(\kappa)} = O\left(n^{1/2-\tau'(\alpha)}(\log n)^{\tau_2}\right),\ \textit{a.s.,} \tag{2.16}$$

where $\tau'(\alpha) = (\alpha\tau_1' - 1/2)/(1+\alpha) > 0$.

Remark 4 *Theorems 1 and 2 are strong approximations, meaning that strong limit theorems can be inferred for the approximated empirical process α_n from those that may hold for the sequence of approximating Gaussian processes as long as the almost sure rate of strong approximation is close enough. This is illustrated in Section 2.4.*

2.3 Comments on the proofs of Theorems 1 and 2

The proofs of Theorems 1 and 2 follow from Propositions 1 and 2 (after some obvious notation translations) exactly as Theorem 1 in [3] follows from their Proposition 1, where a scheme described on pages 236–238 of Philipp [22] is closely followed. (Note that in [3] "$C\rho(\alpha,\gamma)$" should be "$\rho(\alpha,\gamma)$".) The essential ingredients are the maximal Inequalities 1A and 2A. Subsection 4.5.

2.4 Applications to FLIL

Theorem 1 obviously implies that for any fixed choice of $0 < \gamma \le 1 < T$ there exist on the same probability space an i.i.d. sequence $B_1^{(H)}, B_2^{(H)}, \ldots,$ of sample continuous fractional Brownian motions on $[\gamma, T]$ with Hurst index $0 < H < 1$ and a sequence of independent copies $\mathbb{G}_{(\gamma,T)}^{(1)}, \mathbb{G}_{(\gamma,T)}^{(2)}, \ldots,$ of $\mathbb{G}_{(\gamma,T)}$ such that

$$\max_{1\le m\le n}\left\|\sqrt{m}\alpha_m - \sum_{i=1}^{m}\mathbb{G}_{(\gamma,T)}^{(i)}\right\|_{\mathcal{F}_{(\gamma,T)}}$$

$$= \max_{1\le m\le n}\sup_{(t,x)\in\mathcal{T}(\gamma)}\left|\sqrt{m}v_m(t,x) - \sum_{i=1}^{m}G_i(t,x)\right| = o\left(\sqrt{n\log\log n}\right),\ \text{a.s.,} \tag{2.17}$$

where $\mathbb{G}_{(\gamma,T)}^{(i)}(h_{t,x}) =: G_i(t,x)$, for $i \ge 1$. Noting by the comment right after Remark 1, we can consider that each $G_i(t,x)$ is w.p. 1 [with probability 1] in the separable Banach space consisting of continuous functions in the sup-norm on the compact metric space $\mathcal{T}^c(\gamma)$, equipped with the semimetric ρ, we can apply the theorem in LePage [19] (see also Corollary 2.2 of Arcones [1]) to conclude the following FLIL, namely, the sequence of Gaussian processes defined on $\mathcal{T}^c(\gamma)$

$$\left\{ \frac{\sum_{i=1}^n G_i(t,x)}{\sqrt{2n\log\log n}} : (t,x) \in \mathcal{T}^c(\gamma) \right\}$$

is w.p. 1 relatively compact in $\ell_\infty(\mathcal{T}^c(\gamma))$, (the space of bounded functions Υ on $\mathcal{T}^c(\gamma)$ equipped with supremum norm $\|\Upsilon\|_{\ell_\infty(\mathcal{T}^c(\gamma))} = \sup_{\varphi\in\ell_\infty(\mathcal{T}^c(\gamma))} |\Upsilon(\varphi)|$), and its limit set is the unit ball of the reproducing kernel Hilbert space determined by the covariance function $E(G(s,x)G(t,y))$. Note that by continuity of $G(t,x)$ and its covariance function, the same statement holds with $\mathcal{T}^c(\gamma)$ replaced by $\mathcal{T}(\gamma)$. Therefore by (2.17) the same is true for

$$\left\{ \frac{v_n(t,x)}{\sqrt{2\log\log n}} : (t,x) \in \mathcal{T}(\gamma) \right\}. \tag{2.18}$$

This result can also be inferred from the FLIL for the empirical process as stated in Theorem 9 on p. 609 of Ledoux and Talagrand [17] using the fact pointed out above that v_n converges weakly to a bounded uniformly continuous centered Gaussian process $G(t,x)$ indexed by $(t,x) \in \mathcal{T}(\gamma)$. In particular we get that

$$\limsup_{n\to\infty} \frac{\|\alpha_n\|_{\mathcal{F}_{(\gamma,T)}}}{\sqrt{2\log\log n}} = \limsup_{n\to\infty} \sup_{(t,x)\in\mathcal{T}(\gamma)} \left| \frac{v_n(t,x)}{\sqrt{2\log\log n}} \right| = \sigma(\gamma,T), \quad \text{a.s.}$$

where

$$\begin{aligned} \sigma^2(\gamma,T) &= \sup\left\{ E\left(\mathbb{G}^2_{(\gamma,T)}(h_{t,x})\right) : h_{t,x} \in \mathcal{F}_{(\gamma,T)} \right\} \\ &= \sup\left\{ Var(h_{t,x}(B^{(H)})) : (t,x) \in \mathcal{T}(\gamma) \right\} = \frac{1}{4}. \end{aligned}$$

In the same way, on the probability space of Theorem 2, for all $0 < \kappa < \infty$,

$$\begin{aligned} \max_{1\le m\le n} \left\| \sqrt{m}\alpha_m - \sum_{i=1}^m \mathbb{G}^{(i)}_{(0,T)} \right\|_{\mathcal{G}(\kappa)} &= \max_{1\le m\le n} \sup_{(t,x)\in\mathcal{T}(0)} \left| \sqrt{m}t^\kappa v_m(t,x) - \sum_{i=1}^m t^\kappa G_i(t,x) \right| \\ &= o\left(\sqrt{n\log\log n}\right), \text{ a.s.,} \end{aligned} \tag{2.19}$$

where $t^\kappa \mathbb{G}^{(i)}_{(0,T)}(h_{t,x}) =: t^\kappa G_i(t,x)$, for $i \ge 1$. We point out in Remark 6 below that the process $G_\kappa(t,x) := t^\kappa G(t,x)$ has a version that is bounded and uniformly continuous on $\mathcal{T}(0) = [0,T] \times \mathbb{R}$ with respect to the semimetric

$$\rho_\kappa((s,x),(t,y)) = \sqrt{E\left(s^\kappa G(s,x) - t^\kappa G(t,y)\right)^2}. \tag{2.20}$$

From now on we assume that $G_\kappa(t,x)$ is such a version. Denote by $\mathcal{T}^c(0)$ the completion of $\mathcal{T}(0)$ in the topology induced by the semimetric ρ_κ from which we get by arguing as above and applying the LePage theorem that

$$\left\{ \frac{\sum_{i=1}^n t^\kappa G_i(t,x)}{\sqrt{2n\log\log n}} : (t,x) \in \mathcal{T}^c(0) \right\}$$

is, w.p. 1, relatively compact in $\ell_\infty(\mathcal{T}^c(0))$ and its limit set is the unit ball of the reproducing kernel Hilbert space determined by the covariance function $E(s^\kappa t^\kappa G(s,x)G(t,y))$, $(s,x) \in \mathcal{T}^c(0)$. Note that by uniform continuity of $G_\kappa(t,x) = t^\kappa G(t,x)$ and its covariance function, the same statement holds with $\mathcal{T}^c(0)$ replaced by $\mathcal{T}(0)$. Therefore by (2.19) the same is true for the sequence of processes

$$\left\{ \frac{t^\kappa v_n(t,x)}{\sqrt{2\log\log n}} : (t,x) \in \mathcal{T}(0) \right\}. \tag{2.21}$$

This implies that

$$\limsup_{n\to\infty} \sup_{(t,x)\in\mathcal{T}(0)} \left| \frac{t^\kappa v_n(t,x)}{\sqrt{2\log\log n}} \right| = \sigma_\kappa(T), \quad \text{a.s.} \tag{2.22}$$

where

$$\begin{aligned}\sigma_\kappa^2(T) &= \sup\left\{ E\left(\mathbb{G}_{(0,T)}^2(t^\kappa h_{t,x})\right) : t^\kappa h_{t,x} \in \mathcal{G}(\kappa)\right\} \\ &= \sup\left\{ Var(t^\kappa h_{t,x}(B^{(H)})) : (t,x) \in \mathcal{T}(0)\right\} = \frac{T^{2\kappa}}{4}.\end{aligned}$$

FLILs are by no means the only strong limit theorems for α_n that can be derived from Theorems 1 and 2. For instance, one could consider Chung-type LILs.

3 Proofs of Propositions 1 and 2

Before we can prove Propositions 1 and 2 we must first establish Proposition 3 below, which is a version of the coupling given in Proposition 1 that holds on an appropriate class of functions $\mathcal{F}_n$. To do so we must first define this class of functions, derive an entropy bound for it and choose a good grid. Our entropy bound will allow us to fill in the interstices of the empirical and Gaussian processes constructed on $\mathcal{F}_n$ in Proposition 3 by processes defined on all of $\mathcal{F}_{(\gamma_n,T)}$ in such a way as to get useful rates of coupling. The proofs of the bracketing bounds given in Subsection 3.3 form the most technical part of this paper.

3.1 A useful class of functions

To ease the notation from now on we suppress the Hurst index H. As above, let $B(s) = B^{(H)}(s)$, $s \geq 0$, denote a sample continuous fractional Brownian motion with Hurst index $0 < H < 1$. We have

$$E(B(t) - B(s))^2 = |t-s|^{2H}.$$

Note that for all $(s,x), (t,y) \in \mathcal{T}(\gamma)$,

$$\begin{aligned}\rho^2((s,x),(t,y)) &= E(1\{B(s) \leq x\} - F(s,x) - (1\{B(t) \leq y\} - F(t,y)))^2 \\ &\leq E(1\{B(s) \leq x\} - 1\{B(t) \leq y\})^2 = d_P^2(h_{s,x}, h_{t,y}).\end{aligned}$$

For the modulus of continuity of a sample continuous fractional Brownian motion B, with Hurst index H, Wang ([29], Corollary 1.1) proved that

$$\lim_{h\downarrow 0} \sup_{t\in(0,1-h)} \frac{|B(t+h) - B(t)|}{h^H\sqrt{2\log h^{-1}}} = 1, \text{ a.s.} \tag{3.1}$$

Recall the definition of f_H in (2.2). For any $K \geq 1$ denote the class of continuous real-valued functions on $[0,T]$,

$$\mathcal{C}(K) = \{g: \ |g(s) - g(t)| \leq K f_H(|s-t|), 0 \leq s, t \leq T\}. \tag{3.2}$$

One readily checks that $\mathcal{C}(K)$ is closed in $\mathcal{C}[0,T]$. For any $(t,x) \in \mathcal{T}(\gamma)$ let $h_{t,x}^{(K)}$ denote the function of $g \in \mathcal{C}[0,T] \to \{0,1\}$ defined by

$$h_{t,x}^{(K)}(g) = 1\{g(t) \leq x, g \in \mathcal{C}(K)\}.$$

The following class of functions will play an essential role in our proof:

$$\mathcal{F}(K,\gamma) := \left\{h_{t,x}^{(K)} : (t,x) \in \mathcal{T}(\gamma)\right\}. \tag{3.3}$$

It is shown in the Appendix that these classes are *pointwise measurable*, which allows us to take supremums over these classes without the need to worry about measurability problems.

3.2 Bracketing

We shall use the notion of bracketing. Let $\mathcal{G}$ be a class of measurable real-valued functions defined on a measurable space $(S, \mathcal{S})$. A way to measure the size of a class $\mathcal{G}$ is to use $L_2(P)$-brackets. Let l and v be measurable real-valued functions on $(S, \mathcal{S})$ such that $l \leq v$ and $d_P(l,v) < u$, $u > 0$, where

$$d_P(l,v) = \sqrt{E_P\left(l(\xi) - v(\xi)\right)^2}$$

and ξ is a random variable taking values in S defined on a probability space $(\Omega, \mathcal{A}, P)$. The pair of functions l, v form an u-bracket $[l,v]$ consisting of all the functions $f \in \mathcal{G}$ such that $l \leq f \leq v$. Let $N_{[\,]}(u, \mathcal{G}, d_P)$ be the minimum number of u-brackets needed to cover $\mathcal{G}$.

3.3 A useful bracketing bound

Our next aim is to bound the bracketing number $N_{[\,]}(u, \mathcal{F}(K,\gamma), d_P)$, where P is the measure induced on the Borelsets of $\mathcal{C}[0,T]$, by B, with $d_P^2(l,v) = E\left(l(B) - v(B)\right)^2$.

We shall prove the following entropy bound:

Entropy Bound I For some constant C_T (depending on T and H), for $u \in (0, 1/e)$, $\gamma \in (0, 1/e)$ and $K \geq e$,

$$N_{[\,]}(u, \mathcal{F}(K,\gamma), d_P) \leq C_T K^{1/H} u^{-2(1+1/H)} \sqrt{\log u^{-1}} \gamma^{-(1+H)} \left(\log\left(\frac{K}{u\gamma}\right)\right)^{\frac{1}{H}}. \tag{3.4}$$

Proof Choose $\gamma = t_0 < t_1 < \ldots < t_k = T$, such that

$$K f_H(t_i - t_{i-1}) \leq 1, \text{ for } 0 \leq i \leq k, \tag{3.5}$$

and $x_{-m} < x_{-m+1} < \ldots < x_{-1} < x_0 = 0 < x_1 < \ldots < x_m$, with $0 = y_0 < y_1 < \ldots < y_m$, $x_{\pm i} = \pm y_i$, $i = 0, 1, \ldots, m$, such that

$$x_m \geq 2T^H. \tag{3.6}$$

Also put $x_{-(m+1)} = -\infty$, $x_{m+1} = \infty$.

Consider the upper and lower functions: for $g \in \mathcal{C}[0,T]$

$$v_{i,j}(g) = 1\left\{g(t_{i-1}) \le x_j + Kf_H\left(t_i - t_{i-1}\right), g \in \mathcal{C}(K)\right\}$$

and

$$l_{i,j}(g) = 1\left\{g(t_{i-1}) \le x_{j-1} - Kf_H\left(t_i - t_{i-1}\right), g \in \mathcal{C}(K)\right\},$$

for $i = 1, 2, \ldots, k$, $j = -m, \ldots, m, m+1$. Note that $v_{i,m+1}(g) = 1\{g \in \mathcal{C}(K)\}$, and $l_{i,-m}(g) = 0$.

First we show that these functions define a covering. Select any $t_{i-1} < t \le t_i$ (in the case $i = 1$ we allow $t_0 = t$) and $x_{j-1} < x \le x_j$, for $i = 1, \ldots, k$, $j = -m+1, \ldots, m$. Since for any $g \in \mathcal{C}(K)$

$$g\left(t_{i-1}\right) - Kf_H(t_i - t_{i-1}) \le g\left(t\right) \le g\left(t_{i-1}\right) + Kf_H\left(t_i - t_{i-1}\right)$$

we see that for all $g \in \mathcal{C}(K)$, $l_{i,j}(g) \le h_{t,x}^{(K)}(g) \le v_{i,j}(g)$.

Next, for $-\infty < x \le x_{-m}$ and any $t_{i-1} < t \le t_i$, $0 = l_{i,-m}(g) \le h_{t,x}^{(K)}(g) \le v_{i,-m}(g)$, and for $x_m < x < \infty$ and any $t_{i-1} < t \le t_i$, $l_{i,m+1}(g) \le h_{t,x}^{(K)}(g) \le v_{i,m+1}(g) = 1\{g \in \mathcal{C}(K)\}$.

Clearly for $-m+1 \le j \le m$ we get

$$\begin{aligned}
&d_P^2(l_{i,j}, v_{i,j}) = E\left(v_{i,j}(B) - l_{i,j}(B)\right)^2 \\
&= P\Big\{B(t_{i-1}) \in (x_{j-1} - Kf_H(t_i - t_{i-1}), x_j + Kf_H(t_i - t_{i-1})], B \in \mathcal{C}(K)\Big\} \\
&\le P\left\{B(t_{i-1}) \in (x_{j-1} - Kf_H(t_i - t_{i-1}), x_j + Kf_H(t_i - t_{i-1})]\right\} \\
&= \Phi\left(\frac{x_j + Kf_H(t_i - t_{i-1})}{t_{i-1}^H}\right) - \Phi\left(\frac{x_{j-1} - Kf_H(t_i - t_{i-1})}{t_{i-1}^H}\right).
\end{aligned} \tag{3.7}$$

So that for $-m+1 \le j \le m$ we have

$$d_P^2(l_{i,j}, v_{i,j}) \le \frac{1}{\sqrt{2\pi}}\left(x_j - x_{j-1} + 2Kf_H(t_i - t_{i-1})\right) t_{i-1}^{-H}.$$

Inequality (3.7) is also valid for $j = -m$ and $j = m+1$, namely

$$d_P^2(l_{i,-m}, v_{i,-m}) = d_P^2(l_{i,m+1}, v_{i,m+1}) \le 1 - \Phi\left(\frac{x_m - Kf_H(t_i - t_{i-1})}{t_{i-1}^H}\right).$$

Now by $t_{i-1}^H \le T^H$, $2T^H \ge 2$, (3.5) and (3.6) we have

$$\frac{x_m - Kf_H(t_i - t_{i-1})}{t_{i-1}^H} = \frac{2x_m - 2Kf_H(t_i - t_{i-1})}{2t_{i-1}^H} \ge \frac{x_m}{2T^H},$$

which when combined with the standard normal tail bound holding for $z > 0$, $P\{Z \ge z\} \le \frac{1}{z\sqrt{2\pi}}\exp(-z^2/2)$, gives

$$1 - \Phi\left(\frac{x_m - Kf_H(t_i - t_{i-1})}{t_{i-1}^H}\right) \le 1 - \Phi\left(\frac{x_m}{2T^H}\right) \le \frac{1}{\sqrt{2\pi}}\frac{2T^H}{x_m}e^{-\frac{x_m^2}{8T^{2H}}}.$$

From this we see that for $u \in (0, e^{-1})$, the choice $x_m \ge 4T^H\sqrt{\log u^{-1}}$ ensures that

$$d_P^2(l_{i,-m}, v_{i,-m}) = d_P^2(l_{i,m+1}, v_{i,m+1}) \le u^2.$$

Thus to construct our u-covering, it suffices to appropriately partition the intervals

$$[-4T^H\sqrt{\log u^{-1}}, 4T^H\sqrt{\log u^{-1}}] \text{ and } [\gamma, T],$$

so that $x_m \geq 4T^H\sqrt{\log u^{-1}}$, $t_i - t_{i-1}$ satisfies (3.5), and for $0 \leq i \leq k$ and $-m+1 \leq j \leq m$, $d_P^2(l_{i,j}, v_{i,j}) \leq u^2$.

Set

$$\Delta(\gamma, u) = \sqrt{\frac{\pi}{2}}\gamma^H u^2 \text{ and } \Gamma(\gamma, u) = \left(\sqrt{\frac{\pi}{8}}\right)^{1/H} \frac{K^{-1/H}\gamma u^{2/H}}{\left[\log\left(K^{1/H}\gamma^{-1}u^{-2/H}\right)\right]^{1/H}}. \tag{3.8}$$

Let $\lceil x \rceil$ denote here and elsewhere the smallest integer greater than or equal to x. Putting

$$m(\gamma, u) = \left\lceil \frac{4T^H\sqrt{\log u^{-1}}}{\Delta(\gamma, u)} \right\rceil =: m \;\text{ and } k(\gamma, u) = \left\lceil \frac{T-\gamma}{\Gamma(\gamma, u)} \right\rceil =: k,$$

straightforward computations show that for the choice

$$y_i = i\Delta(\gamma, u), i = 0, 1, \ldots, m, \; t_j = \gamma + j\Gamma(\gamma, u), j = 0, 1, \ldots, k-1 \text{ and } t_k = T,$$

by (3.8) we have for $-m+1 \leq j \leq m$

$$d_P^2(l_{i,j}, v_{i,j}) \leq u^2.$$

We also see that $y_m = x_m = 4T^H\sqrt{\log u^{-1}} \geq 2T^H$, and by (3.8) for $0 \leq i \leq k$

$$Kf_H(t_i - t_{i-1}) \leq Kf_H(\Gamma(\gamma, u)) \leq \frac{\gamma^H u^2\sqrt{2\pi}}{4} \leq \gamma^H u^2 \leq 1.$$

Thus (3.5) and (3.6) hold. Hence this choice of t_i and x_j corresponds to a u-covering of $\mathcal{F}(K, \gamma)$. So we have proved the following entropy bound: for $u \in (0, e^{-1}), \gamma \in (0, e^{-1})$ and $K \geq e$

$$N_{[\,]}(u, \mathcal{F}(K, \gamma), d_P) \leq (k(\gamma, u) + 1)(2m(\gamma, u) + 2),$$

thus (3.4) holds for some constant C_T (depending on T and H). □

It will often be convenient to use the following weaker entropy bound, which follows easily from (3.4).

Entropy Bound II For some constant C_T' (depending on T and H), for $u \in (0,1]$, $\gamma \in (0,1]$ and $K \geq e$,

$$N_{[\,]}(u, \mathcal{F}(K, \gamma), d_P) \leq C_T' K^{2/H} u^{-3(1+1/H)} \gamma^{-(1+2/H)}. \tag{3.9}$$

Set

$$\mathcal{F}(K, \gamma, \varepsilon) = \left\{(f, f') \in \mathcal{F}(K, \gamma) \times \mathcal{F}(K, \gamma) : d_P(f, f') < \varepsilon\right\} \tag{3.10}$$

and

$$\mathcal{G}(K, \gamma, \varepsilon) = \left\{f - f' : (f, f') \in \mathcal{F}(K, \gamma, \varepsilon)\right\}, \tag{3.11}$$

that is, $\mathcal{F}(K,\gamma,\varepsilon)$ and $\mathcal{G}(K,\gamma,\varepsilon)$ are the classes of functions on $\mathcal{C}[0,T]$, indexed by $\gamma \le s,t \le T$, $-\infty < x,y < \infty$, defined for $g \in \mathcal{C}[0,T]$ by

$$\left(h_{s,x}^{(K)}(g), h_{t,y}^{(K)}(g)\right) = \left(1\{g(s) \le x, g \in \mathcal{C}(K)\}, 1\{g(t) \le y, g \in \mathcal{C}(K)\}\right)$$

and

$$h_{s,x}^{(K)}(g) - h_{t,y}^{(K)}(g),$$

respectively, and satisfying

$$d_P\left(h_{s,x}^{(K)}, h_{t,y}^{(K)}\right) = \sqrt{E\left(h_{s,x}^{(K)}(B) - h_{t,y}^{(K)}(B)\right)^2} < \varepsilon.$$

We find that independently of ε

$$N_{[\,]}(u, \mathcal{G}(K,\gamma,\varepsilon), d_P) \le \left(N_{[\,]}(u/2, \mathcal{F}(K,\gamma), d_P)\right)^2. \tag{3.12}$$

3.4 Proof of Proposition 1

For any $c > 0$, $n > e$ and $0 < \gamma \le 1 < T$ denote the class of real-valued functions on $[0,T]$,

$$\mathcal{C}_n := \mathcal{C}(\sqrt{c\log n}) = \left\{g :\ |g(s) - g(t)| \le \sqrt{c\log n} f_H(|s-t|),\ 0 \le s,t \le T\right\}, \tag{3.13}$$

and let $\mathcal{C}_\infty$ be as in (2.7). Notice that by (3.1), $P\{B \in \mathcal{C}_\infty\} = 1$. Define the class of functions $\mathcal{C}[0,T] \to \mathbb{R}$ indexed by $[\gamma_n, T] \times \mathbb{R} = \mathcal{T}(\gamma_n)$

$$\mathcal{F}_n = \left\{h_{t,x}^{(\sqrt{c\log n})}(g) = 1\{g(t) \le x, g \in \mathcal{C}_n\} : (t,x) \in \mathcal{T}(\gamma_n)\right\}.$$

To simplify our previous notation we shall write here

$$h_{t,x}^{(n)}(g) = h_{t,x}^{(\sqrt{c\log n})}(g). \tag{3.14}$$

For $h_{t,x}^{(n)} \in \mathcal{F}_n$ let

$$\alpha_n\left(h_{t,x}^{(n)}\right) = n^{-1/2}\sum_{i=1}^{n}\left(1\{B_i(t) \le x, B_i \in \mathcal{C}_n\} - P\{B(t) \le x, B \in \mathcal{C}_n\}\right).$$

Notice that for each $(t,x) \in \mathcal{T}(\gamma_n)$, when $B_i \in \mathcal{C}_n$, for $i = 1,\ldots,n$,

$$\begin{aligned}\alpha_n\left(h_{t,x}^{(n)}\right) &= v_n(t,x) + \sqrt{n}P\{B(t) \le x, B \notin \mathcal{C}_n\} \\ &= \alpha_n(h_{t,x}) + \sqrt{n}P\{B(t) \le x, B \notin \mathcal{C}_n\}.\end{aligned} \tag{3.15}$$

Let $\mathbb{F}_{(\gamma_n,T)}^{(n)}$ denote the mean zero Gaussian process indexed by $\mathcal{F}_n$, having covariance function defined for $h_{s,x}^{(n)}, h_{t,y}^{(n)} \in \mathcal{F}_n$ by

$$\begin{aligned}E\left(\mathbb{F}_{(\gamma_n,T)}^{(n)}\left(h_{s,x}^{(n)}\right)\mathbb{F}_{(\gamma_n,T)}^{(n)}\left(h_{t,y}^{(n)}\right)\right) &= P\{B(s) \le x, B(t) \le y, B \in \mathcal{C}_n\} \\ &- P\{B(s) \le x, B \in \mathcal{C}_n\}P\{B(t) \le y, B \in \mathcal{C}_n\}.\end{aligned}$$

We shall first establish the following auxiliary result.

Proposition 3 *As long as* $1 \geq \ \gamma = \gamma_n > 0$ *satisfies (2.10), for every* $\vartheta > 1$ *there exists a* $\eta(\vartheta) > 0$ *such that for each integer* n *large enough one can construct on the same probability space random vectors* $B_1, \ldots, B_n$ *i.i.d.* B *and a probabilistically equivalent version* $\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}$ *of* $\mathbb{F}^{(n)}_{(\gamma_n,T)}$ *such that*

$$P\left\{\left\|\alpha_n - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_n} > \eta(\vartheta)(\log n)^{\tau_2}\left(n^{-1/2}\gamma_n^{-5H_0/2}\right)^{2/(2+5\nu_0)}\right\} \leq n^{-\vartheta}, \tag{3.16}$$

where τ_2 *is given in Proposition 1 and* H_0 *and* ν_0 *are defined as in (2.9). Moreover, in particular, when* $\gamma_n = n^{-\eta}$*, with* $0 < \eta < \frac{1}{5H_0}$ *and is* τ_1 *as in Proposition 1,*

$$P\left\{\left\|\alpha_n - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_n} > \eta(\vartheta)\, n^{-\tau_1}(\log n)^{\tau_2}\right\} \leq n^{-\vartheta}. \tag{3.17}$$

Proof Let B be a sample continuous fractional Brownian motion with Hurst index $0 < H < 1$ restricted to $[0,T]$ taking values in the measurable space $(\mathcal{C}[0,T], \mathcal{B}[0,T])$. As above P denotes the probability measure induced on the Borel sets of $\mathcal{C}[0,T]$ by B. Let $\mathcal{M}$ denote the real-valued measurable functions on the space $(\mathcal{C}[0,T], \mathcal{B}[0,T])$. For any $\varepsilon > 0$ we can choose a grid

$$\mathcal{H}(\varepsilon) = \{h_k : 1 \leq k \leq N(\varepsilon)\} \tag{3.18}$$

of measurable functions $\mathcal{M}$ on $(\mathcal{C}[0,T], \mathcal{B}[0,T])$ such that each $f \in \mathcal{F}_n$ is in a ball $\{f \in \mathcal{M} : d_P(h_k, f) < \varepsilon\}$ around some h_k, $1 \leq k \leq N(\varepsilon)$, where

$$d_P(h_k, f) = \sqrt{E\left(h_k(B) - f(B)\right)^2}.$$

The choice

$$N(\varepsilon) = N_{[\,]}(\varepsilon/2, \mathcal{F}_n, d_P) \tag{3.19}$$

permits us to select such $h_k \in \mathcal{F}_n$. Recalling the previous notation (3.10) and (3.11), set

$$\mathcal{F}_n(\varepsilon) = \mathcal{F}\left(\sqrt{c\log n}, \gamma_n, \varepsilon\right) \text{ and } \mathcal{G}_n(\varepsilon) = \mathcal{G}\left(\sqrt{c\log n}, \gamma_n, \varepsilon\right). \tag{3.20}$$

Fix $n \geq 1$. Let $B_1, \ldots, B_n$ be i.i.d. B, and $\epsilon_1, \ldots, \epsilon_n$ be independent Rademacher random variables mutually independent of $B_1, \ldots, B_n$. Write for $\varepsilon > 0$,

$$\begin{aligned}\mu_n^S(\varepsilon) &= E\left\{\sup_{(f,f')\in\mathcal{F}_n(\varepsilon)}\left|n^{-1/2}\sum_{i=1}^n \epsilon_i\left(f - f'\right)(B_i)\right|\right\} \\ &= E\left\{\sup_{f-f'\in\mathcal{G}_n(\varepsilon)}\left|n^{-1/2}\sum_{i=1}^n \epsilon_i\left(f - f'\right)(B_i)\right|\right\},\end{aligned} \tag{3.21}$$

and

$$\mu_n^G(\varepsilon) = E\left\{\sup_{(f,f')\in\mathcal{F}_n(\varepsilon)}\left|\mathbb{F}^{(n)}_{(\gamma_n,T)}(f) - \mathbb{F}^{(n)}_{(\gamma_n,T)}(f')\right|\right\}. \tag{3.22}$$

Lemma 1 *Given $\varepsilon > 0$, $\delta > 0$, $t > 0$ and $n \geq 1$ large enough, there exist a probability space $(\Omega, \mathcal{A}, P)$ on which sit $B_1, \ldots, B_n$ i.i.d. B and a probabilistically equivalent version $\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}$ of the Gaussian process $\mathbb{F}^{(n)}_{(\gamma_n,T)}$ indexed by $\mathcal{F}_n$ such that for suitable positive constants C_1, C_2, A, A_1 and A_5 with $A_5 \leq 1/2$, independent of $\varepsilon > 0$, $\delta > 0$, $t > 0$ and $n \geq 1$, we have*

$$\begin{aligned} &P\left\{\left\|\alpha_n - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_n} > A\mu_n^S(\varepsilon) + \mu_n^G(\varepsilon) + \delta + (A+1)t\right\} \\ &\leq C_1 N(\varepsilon)^2 \exp\left(-\frac{C_2\sqrt{n}\,\delta}{(N(\varepsilon))^{5/2}}\right) + 2\exp\left(-A_1\sqrt{n}\,t\right) + 4\exp\left(-\frac{A_5 t^2}{\varepsilon^2}\right). \end{aligned} \tag{3.23}$$

Proof of Lemma 1 Our proof applies the procedure detailed in Section 5.1 in [3]. Given $\varepsilon > 0$ and $n \geq 1$, our aim is to construct a probability space $(\Omega, \mathcal{T}, P)$ on which sit $B_1, \ldots, B_n$ i.i.d. B and a version $\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}$ of the Gaussian process $\mathbb{F}^{(n)}_{(\gamma_n,T)}$ indexed by $\mathcal{F}_n$ such that for $\mathcal{H}(\varepsilon)$ and $\mathcal{F}_n(\varepsilon)$ defined as above and for all $A > 0$, $\delta > 0$ and $t > 0$,

$$\begin{aligned} &P\left\{\left\|\alpha_n - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_n} > A\mu_n^S(\varepsilon) + \mu_n^G(\varepsilon) + \delta + (A+1)t\right\} \\ &\leq P\left\{\max_{h\in\mathcal{H}(\varepsilon)}\left|\alpha_n(h) - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}(h)\right| > \delta\right\} \\ &\quad + P\left\{\sup_{(f,f')\in\mathcal{F}_n(\varepsilon)}\left|\alpha_n(f) - \alpha_n(f')\right| > A\mu_n^S(\varepsilon) + At\right\} \\ &\quad + P\left\{\sup_{(f,f')\in\mathcal{F}_n(\varepsilon)}\left|\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}(f) - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}(f')\right| > t + \mu_n^G(\varepsilon)\right\} \\ &=: P_n(\delta) + Q_n(t,\varepsilon) + \widetilde{Q}_n(t,\varepsilon), \end{aligned} \tag{3.24}$$

with all these probabilities simultaneously small for suitably chosen $A > 0$, $\delta > 0$ and $t > 0$. Consider the n i.i.d. mean zero random vectors in $\mathbb{R}^{N(\varepsilon)}$,

$$Y_i := \frac{1}{\sqrt{n}}\left(h_1(B_i) - \mathbb{E}h_1(B), \ldots, h_{N(\varepsilon)}(B_i) - \mathbb{E}h_{N(\varepsilon)}(B)\right),\ 1 \leq i \leq n.$$

First note that by the definition of $h_k \in \mathcal{F}_n$, we have

$$|Y_i|_{N(\varepsilon)} \leq \sqrt{\frac{N(\varepsilon)}{n}},\ 1 \leq i \leq n,$$

where $|\cdot|_N$, $N \geq 1$, denotes the usual Euclidean norm on $\mathbb{R}^N$. Therefore by the coupling inequality (4.1) we can enlarge the probability space on which (3.24) to include $Z_1, \ldots, Z_n$ i.i.d.

$$Z := \left(Z^1, \ldots, Z^{N(\varepsilon)}\right)$$

mean zero Gaussian vectors such that

$$P_n(\delta) \leq \mathbb{P}\left\{\left|\sum_{i=1}^n (Y_i - Z_i)\right|_{N(\varepsilon)} > \delta\right\} \leq C_1 N(\varepsilon)^2 \exp\left(-\frac{C_2\sqrt{n}\,\delta}{(N(\varepsilon))^{5/2}}\right), \tag{3.25}$$

where $Cov(Z^l, Z^k) = Cov(Y^l, Y^k) =: \langle h_l, h_k \rangle$. Moreover by Lemma A1 of Berkes and Philipp this space can be extended to include a probabilistically equivalent version $\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}$ of the Gaussian process $\mathbb{F}^{(n)}_{(\gamma_n,T)}$ indexed by $\mathcal{F}_n$ such that for $1 \le k \le N(\varepsilon)$,

$$\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}(h_k) = \sum_{i=1}^{n} Z_i^k.$$

The $P_n(\delta)$ in (3.24) is defined through this $\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}$. Notice that the probability space on which $Y_1, \ldots, Y_n$, $Z_1, \ldots, Z_n$ and $\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}$ sit depends on $n \ge 1$ and the choice of $\varepsilon > 0$ and $\delta > 0$.

Observe that the class

$$\mathcal{G}_n(\varepsilon) = \left\{ f - f' : (f, f') \in \mathcal{F}_n(\varepsilon) \right\}$$

satisfies

$$\sigma^2_{\mathcal{G}_n(\varepsilon)} = \sup_{(f,f') \in \mathcal{F}_n(\varepsilon)} Var(f(B) - f'(B)) \le \sup_{(f,f') \in \mathcal{F}_n(\varepsilon)} d_P^2(f, f') \le \varepsilon^2.$$

Thus with $A > 0$ as in (4.8) we get by applying Talagrand's inequality,

$$Q_n(t, \varepsilon) = \mathbb{P}\left\{ ||\alpha_n||_{\mathcal{G}_n(\varepsilon)} > A\left(\mu_n^S(\varepsilon) + t\right) \right\} \le 2\exp\left(-\frac{A_1 t^2}{\varepsilon^2} \right) + 2\exp\left(-A_1\sqrt{n}t\right). \tag{3.26}$$

Next, consider the separable centered Gaussian process $\mathbb{Z}_{(f,f')} = \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}(f) - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}(f')$ indexed by $\mathcal{F}_n(\varepsilon)$. We have

$$\begin{aligned}\sigma_T^2(\mathbb{Z}) &= \sup_{(f,f') \in \mathcal{F}_n(\varepsilon)} \mathbb{E}\left(\left(\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}(f) - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}(f') \right)^2 \right) \\ &= \sup_{(f,f') \in \mathcal{F}_n(\varepsilon)} Var(f(B) - f'(B)) \le \sup_{(f,f') \in \mathcal{F}_n(\varepsilon)} d_P^2(f, f') \le \varepsilon^2.\end{aligned}$$

Borell's inequality (4.15) now gives

$$\widetilde{Q}_n(t, \varepsilon) = \mathbb{P}\left\{ \sup_{(f,f') \in \mathcal{F}_n(\varepsilon)} \left| \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}(f) - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}(f') \right| > t + \mu_n^G(\varepsilon) \right\} \le 2\exp\left(-\frac{t^2}{2\varepsilon^2} \right). \tag{3.27}$$

Putting (3.25), (3.26) and (3.27) together we obtain, for some positive constants A, A_1 and A_5 with $A_5 \le 1/2$,

$$\begin{aligned}&\mathbb{P}\left\{ \left\| \alpha_n - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)} \right\|_{\mathcal{F}_n} > A\mu_n^S(\varepsilon) + \mu_n^G(\varepsilon) + \delta + (A+1)t \right\} \\ &\le C_1 N(\varepsilon)^2 \exp\left(-\frac{C_2\sqrt{n}\ \delta}{(N(\varepsilon))^{5/2}} \right) + 2\exp\left(-A_1\sqrt{n}t\right) + 4\exp\left(-\frac{A_5 t^2}{\varepsilon^2} \right).\end{aligned} \tag{3.28}$$

□

Remark 5 *Here are the Polish spaces that allow us to apply the Berkes and Philipp Lemma A1 as in the construction leading to (3.23). By applying the entropy bound (3.9) we can assume via the Dudley entropy condition (4.10) that the Gaussian process* $\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}$ *indexed by* $\mathcal{F}_n$ *in (3.23) is separable, bounded and* d_P *uniformly continuous, where* d_P *is defined as in (2.8). Moreover, since by using (3.9),* $\mathcal{F}_n$ *is readily seen to be totally bounded, its completion* $\mathcal{F}_n^c$ *is compact. (We complete* $\mathcal{F}_n$ *using the procedure described in Remark 1.) Furthermore, the process* $\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}$ *can be readily extended to be a continuous Gaussian process on* $\mathcal{F}_n^c$*. Thus when applying the Berkes and Philipp lemma we can assume that* $\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}$ *is a Gaussian process indexed by* $\mathcal{F}_n^c$ *taking values in the Polish space* S_3 *of bounded real valued functions defined on the compact set* $\mathcal{F}_n^c$ *continuous with respect to* d_P*. Therefore we can assume that* $B_1,\ldots,B_n$ *i.i.d.* B*,* $Y_1,\ldots,Y_n$ *i.i.d.* Y *and* $Z_1,\ldots,Z_n$ *i.i.d.* Z *take values in the Polish space* $S_1\times S_2$*, where* $S_1=\mathcal{C}\left([0,T]\right)^n\times\mathbb{R}^{N(\varepsilon)n}$ *and* $S_2=\mathbb{R}^{N(\varepsilon)n}$*, and* $Z_1,\ldots,Z_n$ *i.i.d.* Z *and* $\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}$ *take values in the Polish space* $S_2\times S_3$*.*

The proof of Proposition 3 will be completed by refining inequality (3.28). Recall the notation $\mathcal{F}\left(K,\gamma\right)$, $\mathcal{G}\left(K,\gamma,\varepsilon\right)$, (3.19) and (3.20). We find that for any $0<\varepsilon,u<e^{-1}$, with $K=M_n=\sqrt{c\log n}$, (3.4) gives the bound, with ν_0 and H_0 as in (2.9), for some $c_1\geq 1$,

$$\begin{aligned} N(u) &= N_{[\,]}(u/2,\mathcal{F}\left(M_n,\gamma\right),d_P)=N_{[\,]}(u/2,\mathcal{F}_n,d_P) \\ &\leq c_1 M_n^{1/H}u^{-\nu_0}\sqrt{\log u^{-1}}\gamma^{-H_0}\left[\log\left(M_n/(u\gamma)\right)\right]^{1/H} \end{aligned} \tag{3.29}$$

and the weaker entropy bound (3.9) combined with (3.12) implies that for some $c_2>0$ and any $u\in(0,1)$, $\gamma\in(0,1)$,

$$\begin{aligned} N_{[\,]}(u,\mathcal{G}\left(M_n,\gamma,\varepsilon\right),d_P) &= N_{[\,]}(u,\mathcal{G}_n\left(\varepsilon\right),d_P) \\ &\leq \left(N_{[\,]}(u/2,\mathcal{F}\left(M_n,\gamma\right),d_P)\right)^2 \\ &\leq c_2 M_n^{4/H}u^{-3\nu_0}\gamma^{-(2+4/H)}. \end{aligned} \tag{3.30}$$

We shall make frequent use of the following elementary inequality. For any $x\geq 1$ and any $\varepsilon\leq 1$ we have

$$\int_0^{\varepsilon}\sqrt{x+\log u^{-1}}du\leq 2\varepsilon\sqrt{x+\log\varepsilon^{-1}}. \tag{3.31}$$

Setting $\sigma=\varepsilon$ in (4.3) and (4.4) below we get using (3.30) and (3.31) that for some $c_3>0$,

$$J\left(\varepsilon,\mathcal{G}_n\left(\varepsilon\right)\right)=\int_{[0,\varepsilon]}\sqrt{1+\log N_{[\,]}(s,\mathcal{G}_n\left(\varepsilon\right),d_P)}\,ds\leq c_3\,\varepsilon\,\sqrt{\log\left[(\log n)/(\varepsilon\gamma)\right]}$$

and for some $b_0>0$

$$a\left(\varepsilon,\mathcal{G}_n\left(\varepsilon\right)\right)=\varepsilon\left(1+\log N_{[\,]}(\varepsilon,\mathcal{G}_n\left(\varepsilon\right),d_P)\right)^{-1/2}\geq b_0\varepsilon\left(\log\left[(\log n)/(\varepsilon\gamma)\right]\right)^{-1/2}.$$

For the $\mu_n^S\left(\varepsilon\right)$ in (3.21) we obtain by inequality (4.7) with measurable envelope $G=1$

$$\mu_n^S\left(\varepsilon\right)\leq c_3\varepsilon A_0\sqrt{\log\left[(\log n)/(\varepsilon\gamma)\right]}+A_0\sqrt{n}\,1\left\{1>\sqrt{n}b_0\varepsilon/\sqrt{\log\left[(\log n)/(\varepsilon\gamma)\right]}\right\},$$

which as long as $1>\varepsilon=\varepsilon_n>0$ and $1\geq\gamma=\gamma_n>0$ satisfy

$$\frac{\sqrt{n}\varepsilon_n}{\sqrt{\log\left[(\log n)/(\varepsilon_n\gamma_n)\right]}}\to\infty,\ \text{as}\ n\to\infty, \tag{3.32}$$

implies that for all large enough n for a suitable $A_1' > 0$

$$\mu_n^S(\varepsilon_n) \le A_1' \varepsilon_n \sqrt{\log[(\log n)/(\varepsilon_n \gamma_n)]}. \tag{3.33}$$

Recall the definition of $\mu_n^G(\varepsilon)$ in (3.21). We get via the Gaussian moment bound (4.11) and inequality (3.31) that for all $0 < \varepsilon_n < 1/e$ and appropriate A_2' and A_3'

$$\mu_n^G(\varepsilon_n) \le A_2' \int_0^{\varepsilon_n} \sqrt{\log[(\log n)/(u\gamma_n)]}\, du \le A_3' \varepsilon_n \sqrt{\log[(\log n)/(\varepsilon_n \gamma_n)]}.$$

Hence, as long as (3.32) is satisfied, for some $D > 0$ we have for all large enough n

$$A\mu_n^S(\varepsilon_n) + \mu_n^G(\varepsilon_n) \le D\varepsilon_n \sqrt{\log[(\log n)/(\varepsilon_n \gamma_n)]}. \tag{3.34}$$

In addition, by (3.29) we have the bound

$$N(\varepsilon_n) \le c_1 (\log n)^{1/(2H)} \varepsilon_n^{-\nu_0} \sqrt{\log \varepsilon_n^{-1}} \gamma_n^{-H_0} \left(\log[(\log n)/(\varepsilon_n \gamma_n)]\right)^{1/H},$$

and also the weaker bound (3.9) gives for some $c_4 > 0$

$$N(\varepsilon_n) \le c_4 (\log n)^{1/H} \varepsilon_n^{-3(1+1/H)} \gamma_n^{-(1+2/H)}.$$

Therefore, in view of (3.34) and (3.23), it is natural to define for suitably large positive γ_1' and γ_2',

$$\delta = \gamma_1' \varepsilon_n \sqrt{\log[(\log n)/(\varepsilon_n \gamma_n)]} \quad \text{and} \quad t = \gamma_2' \varepsilon_n \sqrt{\log[(\log n)/(\varepsilon_n \gamma_n)]}.$$

We now have by (3.23), as long as (3.32) holds, that for all large enough n there is a probability space depending on γ_1', γ_2', γ_n and ε_n on which α_n and $\widetilde{\mathbb{F}}_{(\gamma_n,T)}^{(n)}$ sit such that

$$\begin{aligned} &P\left\{\left\|\alpha_n - \widetilde{\mathbb{F}}_{(\gamma_n,T)}^{(n)}\right\|_{\mathcal{F}_n} > \left(D + \gamma_1' + (1+A)\gamma_2'\right)\varepsilon_n \sqrt{\log[(\log n)/(\varepsilon_n \gamma_n)]}\right\} \\ &\le \frac{C_1 c_4^2 (\log n)^{2/H}}{\gamma_n^{2+4/H} \varepsilon_n^{3\nu_0}} \exp\left(-\frac{c_5 \sqrt{n} \gamma_1' \varepsilon_n^{1+5\nu_0/2} \gamma_n^{5H_0/2} \sqrt{\log[(\log n)/(\varepsilon_n \gamma_n)]}}{(\log n)^{\frac{5}{4H}} (\log \varepsilon_n^{-1})^{5/4} \left(\log[(\log n)/(\varepsilon_n \gamma_n)]\right)^{\frac{5}{2H}}}\right) \\ &\quad + 2\exp\left(-A_1 \gamma_2' \sqrt{n} \varepsilon_n \sqrt{\log[(\log n)/(\varepsilon_n \gamma_n)]}\right) + 4\exp\left(-A_5 (\gamma_2')^2 \log[(\log n)/(\varepsilon_n \gamma_n)]\right), \end{aligned}$$

for some $c_5 > 0$. Choose ε_n such that

$$\sqrt{n} \varepsilon_n^{1+5\nu_0/2} \gamma_n^{5H_0/2} = (\log n)^{\frac{1}{2} + 5\left(\frac{3}{4H} + \frac{1}{4}\right)}.$$

Then by (2.10)

$$\frac{\log \varepsilon_n^{-1}}{\log n} \to \frac{1 - 5H_0\eta}{2 + 5\nu_0} > 0,$$

and

$$\frac{\log(\varepsilon_n \gamma_n)^{-1}}{\log n} \to \frac{1 - 5H_0\eta}{2 + 5\nu_0} + \eta =: \zeta > 0.$$

An easy computation shows that the exponent of the first term satisfies with a positive constant χ

$$\frac{\sqrt{n} \varepsilon_n^{1+5\nu_0/2} \gamma_n^{5H_0/2} \sqrt{\log[(\log n)/(\varepsilon_n \gamma_n)]}}{(\log n)^{\frac{5}{4H}} (\log \varepsilon_n^{-1})^{5/4} \left(\log[(\log n)/(\varepsilon_n \gamma_n)]\right)^{\frac{5}{2H}}} \sim \chi \log n$$

and

$$\varepsilon_n \sqrt{\log\left[(\log n)/(\varepsilon_n \gamma_n)\right]} \sim \sqrt{\zeta}\,(\log n)^{\tau_2}\left(n^{-1/2}\gamma_n^{-5H_0/2}\right)^{1/(1+5\nu_0/2)},$$

where τ_2 is given in Proposition 1. We readily obtain from these last bounds that for every $\vartheta > 1$ there exist $D > 0$, $\gamma_1' > 0$ and $\gamma_2' > 0$ such that for all $n \geq 1$ large enough, α_n and $\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}$ can be defined on the same probability space so that

$$P\Big\{ \left\|\alpha_n - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_n} > \left(D + \gamma_1' + (1+A)\,\gamma_2'\right) \\ \times \sqrt{2\zeta}(\log n)^{\tau_2}\left(n^{-1/2}\gamma_n^{-5H_0/2}\right)^{1/(1+5\nu_0/2)} \Big\} \leq n^{-\vartheta},$$

which in the special case when $\gamma_n = n^{-\eta}$, with $0 \leq \eta < \frac{1}{5H_0}$, gives

$$P\left\{\left\|\alpha_n - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_n} > \left(D + \gamma_1' + (1+A)\,\gamma_2'\right)\sqrt{2\zeta} n^{-\tau_1}\,(\log n)^{\tau_2}\right\} \leq n^{-\vartheta},$$

where $\tau_1 = (1 - 5H_0\eta)\,/\,(2 + 5\nu_0) > 0$. It is clear now that there exists a $\eta(\vartheta) > 0$ such that (3.16) and (3.17) hold. This completes the proof of Proposition 3. $\square$

We are now ready to complete the proof of Proposition 1. This will be accomplished in two steps.

Step 1 We shall construct the needed version $\widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}$ of the Gaussian process $\mathbb{G}_{(\gamma_n,T)}$ as required in Proposition 1. Set for a fixed n for any $m \geq e$

$$\mathcal{F}_{m,n} = \left\{h^{(m)}_{t,x}(g) = 1\left\{g(t) \leq x, g \in \mathcal{C}_m\right\} : (t,x) \in \mathcal{T}(\gamma_n)\right\}$$

and let

$$\mathcal{F}_{\infty,n} = \left\{h^{(\infty)}_{t,x}(g) = 1\left\{g(t) \leq x, g \in \mathcal{C}_\infty\right\} : (t,x) \in \mathcal{T}(\gamma_n)\right\}.$$

(Note that $\mathcal{F}_{\infty,n} = \mathcal{F}_{(\gamma_n,T)}$.) Set

$$\mathcal{F}_\infty(\gamma_n) = \mathcal{F}_{\infty,n} \cup \cup_{m\geq e}\mathcal{F}_{m,n}.$$

Let $\mathbb{H}_{(\gamma_n,T)}$ be the mean zero Gaussian process indexed by $\mathcal{F}_\infty(\gamma_n)$ such that for $h^{(k)}_{s,x}, h^{(m)}_{t,y} \in \mathcal{F}_\infty(\gamma_n)$ with $e < k \leq m \leq \infty$

$$Cov\left(\mathbb{H}_{(\gamma_n,T)}\left(h^{(k)}_{s,x}\right), \mathbb{H}_{(\gamma_n,T)}\left(h^{(m)}_{t,y}\right)\right) = P\left\{B(s) \leq x, B(t) \leq y, B \in \mathcal{C}_k\right\} \\ - P\left\{B(s) \leq x, B \in \mathcal{C}_k\right\} P\left\{B(t) \leq y, B \in \mathcal{C}_m\right\}.$$

In particular

$$Cov\left(\mathbb{H}_{(\gamma_n,T)}\left(h^{(\infty)}_{s,x}\right), \mathbb{H}_{(\gamma_n,T)}\left(h^{(\infty)}_{t,y}\right)\right) = Cov\left(\mathbb{G}_{(\gamma_n,T)}(h_{s,x}), \mathbb{G}_{(\gamma_n,T)}(h_{t,y})\right).$$

Thus $\mathbb{H}_{(\gamma_n,T)}\left(h^{(\infty)}_{s,x}\right)$ is a version of the Gaussian process $\mathbb{G}_{(\gamma_n,T)}(h_{s,x})$. The process $\widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}$ required in the statement of Proposition 1 will be a version of the process $\mathbb{H}_{(\gamma_n,T)}\left(h^{(\infty)}_{s,x}\right)$.

Notice that for $e < k \leq m \leq \infty$,

$$E\left(\mathbb{H}_{(\gamma_n,T)}\left(h_{s,x}^{(k)}\right) - \mathbb{H}_{(\gamma_n,T)}\left(h_{s,x}^{(m)}\right)\right)^2 \leq P\left\{B\left(s\right) \leq x, B \in \mathcal{C}_m - \mathcal{C}_k\right\} \leq P\left\{B \notin \mathcal{C}_k\right\}. \tag{3.35}$$

In the following lemma using Dudley's entropy condition (4.14) we show that $\mathbb{H}_{(\gamma_n,T)}$ has a continuous modification. To do so, we introduce further notation. For a set $\mathbb{T}$ equipped with a semimetric ρ let $N\left(\varepsilon, \mathbb{T}, \rho\right)$ denote the minimal number of ρ-balls of radius ε needed to cover $\mathbb{T}$.

Lemma 2 *The Gaussian process* $\mathbb{H}_{(\gamma_n,T)}$ *has a bounded uniformly continuous modification* $\widehat{\mathbb{H}}_{(\gamma_n,T)}$.

Proof Using the definition of $\mathcal{C}_k$ in (3.13) and the Landau–Shepp inequality (4.17) we obtain

$$P\{B \notin \mathcal{C}_k\} = P\{L > \sqrt{c \log k}\} \leq Ce^{-Dc\log k} = Ck^{-Dc}. \tag{3.36}$$

Let us fix $1 \geq \varepsilon > 0$ and choose $k = \lceil (4C/\varepsilon^2)^{1/(cD)} \rceil$, where $\lceil\cdot\rceil$ stands for the upper integer part. Then from (3.35) and (3.36) follow for any $m \geq k$ (allowing $m = \infty$) that

$$d_P^2(h_{t,x}^{(m)}, h_{t,x}^{(k)}) \leq P\{B \notin \mathcal{C}_k\} \leq \varepsilon^2/4.$$

For each $\ell \leq k$ choose a $d_P - \varepsilon/2$ grid $\{h_{t_i,x_i}^{(\ell)}\}_{i=1}^{N_\ell}$ in $\mathcal{F}(\sqrt{c\log \ell}, \gamma_n) = \mathcal{F}_{\ell,n}$. The entropy bound II (3.9) and the choice of k shows that

$$\begin{aligned} N_\ell &\leq C(\log k)^{1/H} \varepsilon^{-3(1+1/H)} \gamma_n^{-(1+2/H)} \\ &\leq C' \log \varepsilon^{-1}\, \varepsilon^{-3(1+1/H)} \gamma_n^{-(1+2/H)}. \end{aligned} \tag{3.37}$$

Consider the finite set of functions

$$\mathcal{G} = \cup_{\ell \leq k}\{h_{t_i,x_i}^{(\ell)} : i = 1, 2, \ldots, N_\ell\}.$$

We claim that $\mathcal{G}$ is a $d_P - \varepsilon$ grid in $\mathcal{F}_\infty(\gamma_n)$. Indeed, let $h_{t,x}^{(m)} \in \mathcal{F}_\infty(\gamma_n)$ be arbitrary. If $m \leq k$ then there is an $h_{t_i,x_i}^{(m)} \in \mathcal{G}$ such that $d_P(h_{t,x}^{(m)}, h_{t_i,x_i}^{(m)}) \leq \varepsilon/2$. For $m > k$ we have

$$d_P(h_{t,x}^{(m)}, h_{t_i,x_i}^{(k)}) \leq d_P(h_{t,x}^{(m)}, h_{t,x}^{(k)}) + d_P(h_{t,x}^{(k)}, h_{t_i,x_i}^{(k)}) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

where $h_{t_i,x_i}^{(k)}$ is chosen such that $d_P(h_{t,x}^{(k)}, h_{t_i,x_i}^{(k)}) \leq \varepsilon/2$.

Thus $\mathcal{G}$ is indeed a $d_P - \varepsilon$ grid in $\mathcal{F}_\infty(\gamma_n)$, for which by (3.37)

$$N(\varepsilon, \mathcal{F}_\infty(\gamma_n), d_P) \leq |\mathcal{G}| = \sum_{\ell=1}^{k} N_\ell \leq C\varepsilon^{-a}\gamma_n^{-(1+2/H)}, \tag{3.38}$$

with $a = 2/(cD)+6/H$, say. Thus Dudley's condition (4.14) is satisfied, and a bounded uniformly continuous modification $\widehat{\mathbb{H}}_{(\gamma_n,T)}$ exists. □

From now on to reduce notation we shall assume that $\mathbb{H}_{(\gamma_n,T)}$ is its bounded uniformly continuous modification. Consider the class of functions $\mathcal{C}\left[0,T\right] \to \mathbb{R}^2$ indexed by $(t,x) \in \mathcal{T}\left(\gamma_n\right)$ given by

$$\mathcal{D}_n = \left\{\left(h_{t,x}^{(n)}, h_{t,x}^{(\infty)}\right) : (t,x) \in \mathcal{T}\left(\gamma_n\right)\right\}.$$

Define the mean zero Gaussian process on $\mathcal{D}_n$

$$\mathbb{D}_n^{(n)}\left(h_{t,x}^{(n)},h_{t,x}^{(\infty)}\right)=\mathbb{H}_{(\gamma_n,T)}\left(h_{t,x}^{(n)}\right)-\mathbb{H}_{(\gamma_n,T)}\left(h_{t,x}^{(\infty)}\right). \tag{3.39}$$

Introduce the semimetric on $\mathcal{D}_n$

$$\rho_P^{(1)}\left(\left(h_{s,x}^{(n)},h_{s,x}^{(\infty)}\right),\left(h_{t,y}^{(n)},h_{t,y}^{(\infty)}\right)\right)=\sqrt{E\left(\mathbb{D}_n^{(n)}\left(h_{s,x}^{(n)},h_{s,x}^{(\infty)}\right)-\mathbb{D}_n^{(n)}\left(h_{t,y}^{(n)},h_{t,y}^{(\infty)}\right)\right)^2}.$$

Notice that

$$\begin{aligned}\rho_P^{(1)}\left(\left(h_{s,x}^{(n)},h_{s,x}^{(\infty)}\right),\left(h_{t,y}^{(n)},h_{t,y}^{(\infty)}\right)\right)&\leq\sqrt{2}\,d_P\left(h_{s,x}^{(n)},h_{t,y}^{(n)}\right)+\sqrt{2}\,d_P\left(h_{t,y}^{(\infty)},h_{y,y}^{(\infty)}\right)\\&=:d_P^{(1)}\left(\left(h_{s,x}^{(n)},h_{s,x}^{(\infty)}\right),\left(h_{t,y}^{(n)},h_{t,y}^{(\infty)}\right)\right).\end{aligned}$$

Thus $\rho_P^{(1)}$ is bounded by the semimetric $d_P^{(1)}$.

With the view towards applying the Gaussian moment inequality (4.12) let

$$\begin{aligned}diam\left(\mathcal{D}_n\right)&=\sup\left\{\rho_P^{(1)}\left(\left(h_{s,x}^{(n)},h_{s,x}^{(\infty)}\right),\left(h_{t,y}^{(n)},h_{t,y}^{(\infty)}\right)\right):\left(h_{s,x}^{(n)},h_{s,x}^{(\infty)}\right),\left(h_{t,y}^{(n)},h_{t,y}^{(\infty)}\right)\in\mathcal{D}_n\right\}\\&=\sup\left\{\rho_P^{(1)}\left(\left(h_{s,x}^{(n)},h_{s,x}^{(\infty)}\right),\left(h_{t,y}^{(n)},h_{t,y}^{(\infty)}\right)\right):\left(s,x\right),\left(t,y\right)\in\mathcal{T}(\gamma_n)\right\}\end{aligned}$$

denote the diameter of the set $\mathcal{D}_n$. Observe that

$$diam\left(\mathcal{D}_n\right)\leq 2\sup_{(s,x)\in\mathcal{T}(\gamma_n)}\sqrt{E\left(\mathbb{H}_{(\gamma_n,T)}\left(h_{s,x}^{(n)}\right)-\mathbb{H}_{(\gamma_n,T)}\left(h_{s,x}^{(\infty)}\right)\right)^2},$$

which by (3.35) is

$$\leq 2\sqrt{P\left\{B\notin\mathcal{C}_n\right\}}.$$

This last bound, in turn, by the Landau–Shepp inequality (4.17) below is

$$=2\sqrt{P\left\{L>\sqrt{c\log m}\right\}}\leq 2\sqrt{C}\exp\left(-\frac{Dc\log n}{2}\right). \tag{3.40}$$

Thus for any $\Delta>1$ there exists a $c>0$ such that

$$diam\left(\mathcal{D}_n\right)\leq n^{-\Delta}. \tag{3.41}$$

Next notice that by the definition of $\mathcal{D}_n$ and by (3.38) for some constant $c_6\geq 1$,

$$N(u,\mathcal{D}_n,d_P^{(1)})\leq\left(N(u/2,\mathcal{F}_\infty\left(\gamma_n\right),d_P)\right)^2\leq c_6u^{-2a}\gamma_n^{-(2+4/H)}. \tag{3.42}$$

Write

$$\left\|\mathbb{D}_n^{(n)}\right\|_{\mathcal{D}_n}=\sup\left\{\left|\mathbb{D}_n^{(n)}\left(h_{t,x}^{(n)},h_{t,x}^{(\infty)}\right)\right|:(t,x)\in\mathcal{T}\left(\gamma_n\right)\right\}.$$

Combining (3.41) and (3.42) with the Gaussian moment inequality (4.12) we have

$$\begin{aligned}\left\|\mathbb{D}_n^{(n)}\right\|_{\mathcal{D}_n}&\leq E\left|\mathbb{H}_{(\gamma_n,T)}\left(h_{\gamma_n,0}^{(n)}\right)-\mathbb{H}_{(\gamma_n,T)}\left(h_{\gamma_n,0}^{(\infty)}\right)\right|+A_4\int_0^{n^{-\Delta}}\sqrt{\log N(u,\mathcal{D}_n,d_P^{(1)})}\,du\\&\leq n^{-\Delta}+A_4\int_0^{n^{-\Delta}}\sqrt{\log c_6-2a\log u-\left(2+\frac{4}{H}\right)\log\gamma_n}\,du,\end{aligned}$$

which by using (2.10) and (3.31) gives for some $b > 0$,

$$\left\| \mathbb{D}_n^{(n)} \right\|_{\mathcal{D}_n} \le b n^{-\Delta} \sqrt{\log n}. \tag{3.43}$$

We have by using the Landau–Shepp inequality (4.17) that for a sufficiently large $c > 0$

$$\begin{aligned} \sigma^2_{\mathcal{D}_n}\left(\mathbb{D}_n^{(n)}\right) &= \sup\left\{ E\left(\mathbb{D}_n^{(n)}\left(h_{t,x}^{(n)}, h_{t,x}^{(\infty)}\right)\right)^2 : (t,x) \in \mathcal{T}\left(\gamma_n\right)\right\} \\ &\le P\left\{B \notin \mathcal{C}_n\right\} \le n^{-\Delta}. \end{aligned} \tag{3.44}$$

Hence by Borell's inequality (4.15), for all $z > 0$,

$$P\left\{ \left| \left\| \mathbb{D}_n^{(n)} \right\|_{\mathcal{D}_n} - E \left\| \mathbb{D}_n^{(n)} \right\|_{\mathcal{D}_n} \right| > z \right\} \le 2 \exp\left(- \frac{z^2}{2\sigma^2_{\mathcal{D}_n}\left(\mathbb{D}_n^{(n)}\right)} \right),$$

which on account of (3.43) and (3.44) gives for all $\theta > 1$

$$P\left\{ \left\| \mathbb{D}_n^{(n)} \right\|_{\mathcal{D}_n} > b n^{-\Delta}\sqrt{\log n} + 2n^{-\Delta/2}\sqrt{\theta \log n} \right\} \le n^{-\theta}. \tag{3.45}$$

Returning to the construction of $\widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}$ in Proposition 1, for each $n > e$, let $\mathbb{F}^{(n)}_{(\gamma_n,T)}$ denote the restriction of $\mathbb{H}_{(\gamma_n,T)}$ to $\mathcal{F}_n$ and $\mathbb{G}_{(\gamma_n,T)}$ the restriction of $\mathbb{H}_{(\gamma_n,T)}$ to $\mathcal{F}_{(\gamma_n,T)}$. Notice by (3.37), $\mathcal{F}_\infty\left(\gamma_n\right)$ is totally bounded in the d_P semimetric, as are $\mathcal{F}_n$ and $\mathcal{F}_{\infty,n} = \mathcal{F}_{(\gamma,T)}$. By the discussion given in Remark 5 for $\mathbb{F}^{(n)}_{(\gamma_n,T)}$ and $\mathcal{F}^{(n)}_{(\gamma_n,T)}$, $\mathbb{G}_{(\gamma_n,T)}$ can be extended to a continuous function on the completion $\mathcal{F}^c_{(\gamma_n,T)}$ of $\mathcal{F}_{(\gamma_n,T)}$, which is compact. Therefore we can argue that

$$\left(\mathbb{F}^{(n)}_{(\gamma_n,T)}, \mathbb{G}_{(\gamma_n,T)}, \left\| \mathbb{D}_n^{(n)} \right\|_{\mathcal{D}_n}\right)$$

takes values in the Polish space $S_3 \times S_4$, where S_3 is as in Remark 5 and $S_4 = S_3' \times \mathbb{R}$, with S_3' being the Banach space of bounded real valued functions defined on the compact set $\mathcal{F}^c_{(\gamma_n,T)}$ continuous with respect to d_P. Hence Lemma A1 of Berkes and Philipp applies here and we can enlarge the probability space on which inequality (3.16) holds to include a version

$$\left(\widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}, \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}, \left\| \widetilde{\mathbb{D}}_n^{(n)} \right\|_{\mathcal{D}_n}\right)$$

of the process $\left(\mathbb{F}^{(n)}_{(\gamma_n,T)}, \mathbb{G}_{(\gamma_n,T)}, \left\| \mathbb{D}_n^{(n)} \right\|_{\mathcal{D}_n}\right)$ so that besides (3.16), (3.45) is also valid.

Step 2 We shall show that inequality (2.11) holds for $\left\| \alpha_n - \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)} \right\|_{\mathcal{F}_{(\gamma_n,T)}}$, which will complete the proof of Proposition 1. Define for $\left(h_{t,x}^{(n)}, h_{t,x}^{(\infty)}\right) \in \mathcal{D}_n$,

$$\widetilde{\mathbb{D}}_n^{(n)}\left(h_{t,x}^{(n)}, h_{t,x}^{(\infty)}\right) = \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}\left(h_{t,x}^{(n)}\right) - \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}\left(h_{t,x}\right).$$

Clearly

$$\left\| \mathbb{D}_n^{(n)} \right\|_{\mathcal{D}_n} \stackrel{\mathrm{D}}{=} \left\| \widetilde{\mathbb{D}}_n^{(n)} \right\|_{\mathcal{D}_n} = \sup\left\{ \left| \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}\left(h_{t,x}^{(n)}\right) - \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}\left(h_{t,x}\right) \right| : (t,x) \in \mathcal{T}\left(\gamma_n\right) \right\}.$$

Notice that by (3.15)

$$\left\|\alpha_n - \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_{(\gamma_n,T)}} \leq \sup\left\{\left|\alpha_n\left(h^{(n)}_{t,x}\right) - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}\left(h^{(n)}_{t,x}\right)\right| : (t,x) \in \mathcal{T}(\gamma_n)\right\}$$
$$+ \sum_{i=1}^{n} \frac{1\{B_i \notin \mathcal{C}_n\}}{\sqrt{n}} + \sqrt{n}P\{B \notin \mathcal{C}_n\} + \left\|\widetilde{\mathbb{D}}^{(n)}_n\right\|_{\mathcal{D}_n}.$$

Let

$$\delta_n(\Delta) = bn^{-\Delta}\sqrt{\log n} + 2n^{-\Delta/2}\sqrt{\theta \log n}.$$

Recalling that $P\{B \notin \mathcal{C}_n\} \leq n^{-\Delta}$, we get by (3.16) and (3.45)

$$P\left\{\left\|\alpha_n - \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_{(\gamma_n,T)}} > \eta(\vartheta)(\log n)^{\tau_2}\left(n^{-1/2}\gamma_n^{-5H_0/2}\right)^{2/(2+5\nu_0)} + \sqrt{n}n^{-\Delta} + \delta_n(\Delta)\right\}$$
$$\leq P\left\{\left\|\alpha_n - \widetilde{\mathbb{F}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_n} > \eta(\vartheta)(\log n)^{\tau_2}\left(n^{-1/2}\gamma_n^{-5H_0/2}\right)^{2/(2+5\nu_0)}\right\} \tag{3.46}$$
$$+ P\left\{\sum_{i=1}^{n} 1\{B_i \notin \mathcal{C}_n\} > 0\right\} + P\left\{\left\|\mathbb{D}^{(n)}_n\right\|_{\mathcal{D}_n} > \delta_n(\Delta)\right\}$$
$$\leq n^{-\vartheta} + n^{1-\Delta} + n^{-\theta}.$$

Noting that the Δ in the above inequalities can be made as large as desired by choosing c large enough, we see that for every $\lambda > 1$, for sufficiently large $c > 0$ (that is large $\Delta > 0$), $\vartheta > 0$, $\theta > 0$ and all large n

$$n^{-\theta} + n^{1-\Delta} + n^{-\vartheta} < n^{-\lambda}, \tag{3.47}$$

and for any choice of $\vartheta > 0$, $\theta > 0$ and large enough $c > 0$ (large $\Delta > 0$), for all large n

$$\eta(\vartheta)\left(n^{-1/2}\gamma_n^{-5H_0/2}\right)^{2/(2+5\nu_0)} \geq \sqrt{n}n^{-\Delta} + \delta_n(\Delta).$$

Thus there is a $\rho(\lambda) > 0$, and $c > 0$, $\vartheta > 0$ and $\theta > 0$ such that for all large enough n,

$$\rho(\lambda)(\log n)^{\tau_2}\left(n^{-1/2}\gamma_n^{-5H_0/2}\right)^{2/(2+5\nu_0)}$$
$$> \eta(\vartheta)(\log n)^{\tau_2}\left(n^{-1/2}\gamma_n^{-5H_0/2}\right)^{2/(2+5\nu_0)} + \sqrt{n}n^{-\Delta} + \delta_n(\Delta)$$

and (3.47) holds, which by (3.46) implies that

$$P\left\{\left\|\alpha_n - \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_{(\gamma_n,T)}} > \rho(\lambda)(\log n)^{\tau_2}\left(n^{-1/2}\gamma_n^{-5H_0/2}\right)^{2/(2+5\nu_0)}\right\} < n^{-\lambda},$$

that is, for all such large n there exists a suitable probability space such that (2.11) holds. This completes the proof of Proposition 1. □

3.5 Proof of Proposition 2

Put $\gamma_n = n^{-\eta}$, with $\eta = (5H_0 + \kappa(2+5\nu_0))^{-1}$. Note that for this choice of η

$$\tau_1(\eta) = \tau_1' = \tau_1'(\kappa) = \kappa/(5H_0 + \kappa(2+5\nu_0)) = \kappa\eta. \tag{3.48}$$

Applying Proposition 1, for every $\lambda' > \lambda > 1$ there exists a $\rho(\lambda') > 0$ such that for each integer n large enough one can construct on the same probability space random vectors $B_1, \dots, B_n$ i.i.d. B and a probabilistically equivalent version $\widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}$ of $\mathbb{G}_{(\gamma_n,T)}$ such that,

$$P\left\{\left\|\alpha_n - \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}\right\|_{\mathcal{F}_{(\gamma_n,T)}} > \rho\left(\lambda'\right) n^{-\tau_1} (\log n)^{\tau_2}\right\} \le n^{-\lambda'},$$

with $\tau_1 = \tau_1(\eta) = (1 - 5H_0\eta)/(2+5\nu_0)$, which, since $T^\kappa/t^\kappa \ge 1$ for $t \in [\gamma_n, T]$, implies that

$$P\left\{\sup_{(t,x)\in[\gamma_n,T]\times\mathbb{R}} t^\kappa \left|\alpha_n(h_{t,x}) - \widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}(h_{t,x})\right| > T^\kappa \rho\left(\lambda'\right) n^{-\tau_1} (\log n)^{\tau_2}\right\} \le n^{-\lambda'}. \tag{3.49}$$

Using Lemma A1 of Berkes and Philipp, we can enlarge the probability on which (3.49) holds to include a Gaussian process $\mathbb{G}_{(0,T)}$ indexed by $\mathcal{G}(\kappa)$, so that $\mathbb{G}_{(0,T)}$ and $\widetilde{\mathbb{G}}^{(n)}_{(\gamma_n,T)}$ agree on

$$\left\{t^\kappa h_{t,x} : (t,x) \in [\gamma_n, T] \times \mathbb{R}\right\}.$$

(The validity of the application of the Berkes and Philipp lemma can be argued as in Remark 5.) Further we have, using inequality (4.18) below with $\delta = \kappa$, that for a suitable $d_1 > 0$ for all large n

$$P\left\{\sup\left\{t^\kappa \left|G(t,x)\right| : (t,x) \in \left[0, n^{-\eta}\right] \times \mathbb{R}\right\} > d_1 n^{-\eta\kappa}\sqrt{\log n}\right\} \le n^{-\lambda'}, \tag{3.50}$$

where $G(t,x) = \mathbb{G}_{(0,T)}(h_{t,x})$ for $(t,x) \in [0, n^{-\eta}] \times \mathbb{R}$.

Next by using inequality (4.25) below with $\delta = \kappa$ we get that for a suitable $d_2 > 0$ for all large n

$$P\left\{\sup\left\{t^\kappa \left|\alpha_n(h_{t,x})\right| : (t,x) \in \left[0, n^{-\eta}\right] \times \mathbb{R}\right\} > d_2 n^{-\eta\kappa}\sqrt{\log n}\right\} \le n^{-\lambda'}. \tag{3.51}$$

Recall the notation (2.14). Combining inequalities (3.49), (3.50) and (3.51), and noting that $\tau_2 > 1/2$, we get for all large enough n

$$P\left\{\left\|\alpha_n - \widetilde{\mathbb{G}}_{(0,T)}\right\|_{\mathcal{G}(\kappa)} > \left(d_1 + d_2 + T^\kappa \rho(\lambda')\right) n^{-\min\{\tau_1, \eta\kappa\}} (\log n)^{\tau_2}\right\} \le 3n^{-\lambda'}.$$

It is clear now that the optimal choice for η satisfies $\tau_1(\eta) = \kappa\eta$, which by (3.48) our chosen value fulfills. Thus by choosing λ' so that $3n^{-\lambda'} < n^{-\lambda}$, setting $\rho'(\lambda) = d_1 + d_2 + T^\kappa \rho(\lambda')$, we conclude that (2.15) holds. □

Remark 6 *Here we discuss the continuity of the Gaussian process* $\mathbb{G}_{(0,T)}$ *indexed by* $\mathcal{G}(\kappa)$*. A straightforward argument based on Inequality 1 in the Appendix shows that, w.p. 1, for all* $\varepsilon > 0$ *there exists a* $0 < \gamma < 1$ *such that*

$$\sup_{(t,x)\in[0,\gamma]\times\mathbb{R}} t^\kappa \left|G(t,x)\right| < \varepsilon. \tag{3.52}$$

Moreover, as pointed out above, for any $0 < \gamma < 1$, $G(t,x)$ is almost surely bounded and uniformly continuous on $[\gamma, T] \times \mathbb{R}$, when equipped with the semimetric (2.5), which implies the same for $t^{\kappa} G(t,x)$, which when combined with (3.52), readily implies that $t^{\kappa} G(t,x)$ is almost surely bounded and uniformly continuous on $\mathcal{T}(0) = [0,T] \times \mathbb{R}$ with respect to the semimetric ρ_κ (2.20). Also by applying Proposition 1 on page 26 of Lifshits [20] we can assume that the Gaussian process $t^{\kappa} G(t,x)$ is separable. Thus there exists a version of $t^{\kappa} G(t,x)$ that is bounded and uniformly continuous on $\mathcal{T}(0)$.

4 Appendix

4.1 A Gaussian coupling inequality

Einmahl and Mason [8] pointed out in their Fact 2.2 that the Strassen–Dudley theorem (see Theorem 11.6.2 in Dudley [7]) in combination with a special case of Theorem 1.1 of Zaitsev [30] (also see the discussion after its statement) yields the following Gaussian coupling. Here $|\cdot|_N$, $N \geq 1$, denotes the usual Euclidean norm on $\mathbb{R}^N$.

Coupling inequality. *Let $Y_1, \dots, Y_n$ be independent mean zero random vectors in $\mathbb{R}^N$, $N \geq 1$, such that for some $b > 0$,*

$$|Y_i|_N \leq b, \ i = 1, \dots, n.$$

If $(\Omega, \mathcal{T}, P)$ is rich enough then for each $\delta > 0$, one can define independent normally distributed mean zero random vectors $Z_1, \dots, Z_n$ with Z_i and Y_i having the same covariance matrix for $i = 1, \dots, n$, such that for universal constants $C_1 > 0$ and $C_2 > 0$,

$$P\left\{ \left| \sum_{i=1}^{n} (Y_i - Z_i) \right|_N > \delta \right\} \leq C_1 N^2 \exp\left(-\frac{C_2 \delta}{N^2 b} \right). \tag{4.1}$$

Remark 7 *Actually Einmahl and Mason did not specify the N^2 in (4.1) and they applied a less precise result given Theorem 1.1 in [31] with N^2 replaced by $N^{5/2}$, however their argument is equally valid when based upon Theorem 1.1 in [30]. Zaitsev [30] remarks that the assumptions of Theorem 1.1 of [31] imply those of Theorem 1.1 of [30]. See, in particular, the paragraph right above Remark 1.1 in [30]. Also see equation (18) in [32].*

4.2 Pointwise measurable classes

Definition. A class $\mathcal{G}$ of measurable real-valued functions defined on a measurable space $(S, \mathcal{S})$ is *pointwise measurable* if there exists a countable subclass $\mathcal{G}_\infty$ of $\mathcal{G}$ such that we can find for any function $f \in \mathcal{G}$ a sequence of functions $\{f_m\}$ in $\mathcal{G}_\infty$ for which $\lim_{m \to \infty} f_m(x) = f(x)$ for all $x \in S$. For more about pointwise measurability see pages 109-110 and Example 2.3.4 of van der Vaart and Wellner [28], as well as Section 8.2 of Kosorok [12].

We shall show here that the classes of functions $\mathcal{F}(K, \gamma)$, $K \geq 1$, of the form (3.3), where $0 \leq \gamma < 1 < T < \infty$ are pointwise measurable. Let $\mathbb{Q}$ denote the set of rational numbers. For any $K \geq 1$ consider the countable class $\mathcal{F}_{\infty,K}$ of functions of $g \in \mathcal{C}[0,T] \to \{0,1\}$ indexed by $u, v \in [\gamma, T] \cap \mathbb{Q} \cup \{\gamma, T\}, y \in \mathbb{Q}$ defined by

$$1\{g(v) - K f_H(|v-u|) \leq y, \ g \in \mathcal{C}(K)\},$$

where $\mathcal{C}(K)$ is as in (3.2). Clearly for each $(t,x) \in \mathcal{T}(\gamma) = [\gamma, T] \times \mathbb{R}$ we can choose sequences s_m and $t_m \in [\gamma, T] \cap \mathbb{Q} \cup \{\gamma, T\}$ such that $t_m \searrow t$ and $s_m \nearrow t$. Also we can select a sequence $y_m \in \mathbb{Q} \searrow x$. We see that each

$$1\left\{g(t_m) - Kf_H(|t_m - s_m|) \le y_m, g \in \mathcal{C}(K)\right\} \in \mathcal{F}_{\infty,K}.$$

Moreover, if $g \in \mathcal{C}(K)$, then $g(t_m) - Kf_H(|t_m - s_m|) \le g(t)$ and $g(t_m) - Kf_H(|t_m - s_m|) \to g(t)$. Thus if $g(t) \le x$ and $g \in \mathcal{C}(K)$ then

$$1\left\{g(t_m) - Kf_H(|t_m - s_m|) \le y_m, g \in \mathcal{C}(K)\right\} = 1 \to 1 = h_{t,x}^{(K)}(g).$$

Whereas if $g(t) > x$ then for some $\delta > 0$, $g(t) > x + \delta$ and all large enough m,

$$g(t_m) - Kf_H(|t_m - s_m|) > x + \delta/2 \text{ and } x + \delta/4 > y_m.$$

This says that eventually $g(t_m) - Kf_H(|t_m - s_m|) > y_m$ and thus

$$1\left\{g(t_m) - Kf_H(|t_m - s_m|) \le y_m, g \in \mathcal{C}(K)\right\} = 0 = h_{t,x}^{(K)}(g).$$

Hence $\mathcal{F}(K,\gamma)$ is pointwise measurable with countable subclass $\mathcal{F}_{\infty,K}$.

For any $\kappa > 0$ and $K \ge 1$ let $\mathcal{G}(\kappa, K)$ denote the class of functions $g \in \mathcal{C}[0,T] \to [0, T^\kappa]$ indexed by $(t,x) \in \mathcal{T}(0) = [0,T] \times \mathbb{R}$ defined by

$$t^\kappa h_{t,x}^{(K)}(g) = t^\kappa 1\left\{g(t) \le x, g \in \mathcal{C}(K)\right\}. \tag{4.2}$$

Clearly by a slight modification of the above argument $\mathcal{G}(\kappa, K)$ is pointwise measurable.

4.3 Inequalities for empirical processes

In this subsection $\mathcal{G}$ is a pointwise measurable class of measurable real-valued functions defined on a measurable space $(S, \mathcal{S})$. For any $0 < \sigma < 1$, set

$$J(\sigma, \mathcal{G}) = \int_{[0,\sigma]} \sqrt{1 + \log N_{[\,]}(s, \mathcal{G}, d_P)}\, ds \tag{4.3}$$

and

$$a(\sigma, \mathcal{G}) = \sigma \left[1 + \log N_{[\,]}(\sigma, \mathcal{G}, d_P)\right]^{-1/2}. \tag{4.4}$$

Lemma 19.34 in van der Vaart [27] gives the following moment bound. (Note the needed " +1" in the definition of $J(\sigma,\mathcal{G})$ and $a(\sigma,\mathcal{G})$.)

Moment inequality. *Let $\xi, \xi_1, \ldots, \xi_n$ be i.i.d. and assume that $\mathcal{G}$ has a measurable envelope function G and $E\left(g^2(\xi)\right) < \sigma^2 < 1$ for every $g \in \mathcal{G}$. We have, for a universal constant A_0',*

$$E\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \left(g(\xi_i) - Eg(\xi_i)\right)\right\|_{\mathcal{G}} \le A_0'\left[J(\sigma,\mathcal{G}) + \sqrt{n}\, E\left(G(\xi)\, 1\left\{G(\xi) > \sqrt{n} a(\sigma,\mathcal{G})\right\}\right)\right]. \tag{4.5}$$

Let ϵ be a Rademacher variable, i.e. $P\{\epsilon = 1\} = P\{\epsilon = -1\} = 1/2$, and consider independent Rademacher variables $\epsilon_1, \ldots, \epsilon_n$ independent of $\xi_1, \ldots, \xi_n$. From a special case of a well-known symmetrization lemma, we have for any class of functions $\mathcal{G}$ in $L_1(P)$

$$\frac{1}{2}E\left\|\sum_{i=1}^{n}\epsilon_i\left(g(\xi_i)-Eg(\xi)\right)\right\|_{\mathcal{G}} \leq E\left\|\sum_{i=1}^{n}\left(g(\xi_i)-Eg(\xi)\right)\right\|_{\mathcal{G}} \leq 2E\left\|\sum_{i=1}^{n}\epsilon_i g(\xi_i)\right\|_{\mathcal{G}}.$$

(See Lemma 6.3 of Ledoux and Talagrand [18].) In particular we get

$$\begin{aligned} E\left\|\sum_{i=1}^{n}\epsilon_i g(\xi_i)\right\|_{\mathcal{G}} &\leq E\left\|\sum_{i=1}^{n}\epsilon_i\left(g(\xi_i)-Eg(\xi)\right)\right\|_{\mathcal{G}} + E\left|\sum_{i=1}^{n}\epsilon_i\right|\left\|Eg(\xi)\right\|_{\mathcal{G}} \\ &\leq 2E\left\|\sum_{i=1}^{n}\left(g(\xi_i)-Eg(\xi)\right)\right\|_{\mathcal{G}} + \sigma\sqrt{n}. \end{aligned} \tag{4.6}$$

Thus we readily get from (4.5) with $A_0 = 2A_0' + 1$ and noting that the integrand of $J(\sigma, \mathcal{G})$ is greater than or equal to 1,

$$E\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\epsilon_i g(X_i)\right\|_{\mathcal{G}} \leq A_0\left[J(\sigma,\mathcal{G}) + \sqrt{n}\;E\left(G(\xi)\,1\left\{G(\xi) > \sqrt{n}\;a(\sigma,\mathcal{G})\right\}\right)\right]. \tag{4.7}$$

We shall be using the moment bound (4.7) in conjunction with the following exponential inequality due to Talagrand [26]. This maximal version is pointed out by Einmahl and Mason [9, Inequality A.1 on p.31].

Talagrand inequality. *Let $\mathcal{G}$ be a pointwise measurable class of measurable real-valued functions defined on a measurable space $(S, \mathcal{S})$ satisfying $||g||_\infty \leq M,\ g \in \mathcal{G}$, for some $0 < M < \infty$. Let X, X_n, $n \geq 1$, be a sequence of i.i.d. random variables defined on a probability space $(\Omega, \mathcal{A}, P)$ and taking values in S, then for all $t > 0$ we have for suitable finite constants $A, A_1 > 0$,*

$$P\left\{\max_{1\leq m\leq n}||\sqrt{m}\alpha_m||_{\mathcal{G}} \geq A\left(E\left\|\sum_{i=1}^{n}\epsilon_i g(X_i)\right\|_{\mathcal{G}} + t\right)\right\} \leq 2\exp\left(-\frac{A_1 t^2}{n\sigma_{\mathcal{G}}^2}\right) + 2\exp\left(-\frac{A_1 t}{M}\right), \tag{4.8}$$

where $\sigma_{\mathcal{G}}^2 = \sup_{g\in\mathcal{G}} Var(g(X))$.

4.4 Inequalities for Gaussian processes

Let $\mathbb{Z}$ be a separable mean zero Gaussian process on a probability space $(\Omega, \mathcal{A}, P)$ indexed by a set $\mathbb{T}$, equipped with a semimetric

$$\rho(s,t) = \sqrt{E\left(\mathbb{Z}(t) - \mathbb{Z}(s)\right)^2}. \tag{4.9}$$

For each $\varepsilon > 0$ let $N(\varepsilon, \mathbb{T}, \rho)$ denote the minimal number of ρ-balls of radius ε needed to cover $\mathbb{T}$. Write $\|\mathbb{Z}\|_{\mathbb{T}} = \sup_{t\in\mathbb{T}}|\mathbb{Z}_t|$ and $\sigma_{\mathbb{T}}^2(\mathbb{Z}) = \sup_{t\in\mathbb{T}} E\left(\mathbb{Z}_t^2\right)$.

According to Dudley [6], the entropy condition

$$\int_{[0,1]} \sqrt{\log N(\varepsilon, \mathbb{T}, \rho)}\, d\varepsilon < \infty \tag{4.10}$$

ensures the existence of a separable, bounded, ρ-uniformly continuous modification of $\mathbb{Z}$. The following moment bound is a version of Corollary 2.2.8 in van der Vaart and Wellner [28]. (Also see their Problem 2.2.14.)

Gaussian moment inequality. *For some universal constant $A_4 > 0$ and all $\sigma > 0$ we have*

$$E\left(\sup_{\rho(s,t)<\sigma} |\mathbb{Z}_t - \mathbb{Z}_s|\right) \le A_4 \int_{[0,\sigma]} \sqrt{\log N(\varepsilon, \mathbb{T}, \rho)}\, d\varepsilon \tag{4.11}$$

and for any $t_0 \in \mathbb{T}$,

$$E\left(\|\mathbb{Z}\|_{\mathbb{T}}\right) \le E\,|\mathbb{Z}_{t_0}| + A_4 \int_{[0,\mathbb{D}]} \sqrt{\log N(\varepsilon, \mathbb{T}, \rho)}\, d\varepsilon, \tag{4.12}$$

with

$$\mathbb{D} = \sup_{s,t\in\mathbb{T}} \rho(s,t) \tag{4.13}$$

denoting the diameter of $\mathbb{T}$.

Notice that if d is a semimetric on $\mathbb{T}$ such that for all $s, t \in T$, $d(s,t) \ge \rho(s,t)$, then

$$\sup_{\{s:\rho(s,t)<\sigma\}} |\mathbb{Z}_t - \mathbb{Z}_s| \ge \sup_{\{s:\ d(s,t)<\sigma\}} |\mathbb{Z}_t - \mathbb{Z}_s|$$

and $N(\varepsilon, \mathbb{T}, d) \ge N(\varepsilon, \mathbb{T}, \rho)$. Thus

$$\int_{[0,1]} \sqrt{\log N(\varepsilon, \mathbb{T}, d)}\, d\varepsilon < \infty \tag{4.14}$$

implies by the Dudley result the existence of a separable, bounded, d-uniformly continuous modification of $\mathbb{Z}$. (Here note that ρ-uniformly continuous implies d-uniformly continuous.) Moreover the moment inequalities in (4.11) and (4.12) hold when ρ is replaced by d and in the definition of $\mathbb{D}$.

In particular, these inequalities hold when $\mathbb{Z} = \mathbb{G}_{(\gamma,T)}$, the Gaussian process defined at the end of Subsection 2.1, where $\mathbb{T} = \mathcal{F}_{(\gamma,T)}$ and $d = d_P$ is as defined in (2.8), and $\mathbb{D} = \sup\left\{d_P(f,g) : f, g \in \mathcal{F}_{(\gamma,T)}\right\}$ is the diameter $\mathbb{D}$ of $\mathbb{T} = \mathcal{F}_{(\gamma,T)}$.

The following large deviation probability estimate for $\|\mathbb{Z}\|_{\mathbb{T}}$ is due to Borell [4]. (Also see Proposition A.2.1 in [28].) Let $M(X)$ denote the a median of $\|\mathbb{Z}\|_{\mathbb{T}}$, i.e. $P\{\|\mathbb{Z}\|_{\mathbb{T}} \ge M(X)\} \ge 1/2$ and $P\{\|\mathbb{Z}\|_{\mathbb{T}} \le M(X)\} \ge 1/2$. We shall assume that $M(X)$ is finite.

Borell's inequality. *For all $z > 0$,*

$$P\left\{\left|\|\mathbb{Z}\|_{\mathbb{T}} - E\left(\|\mathbb{Z}\|_{\mathbb{T}}\right)\right| > z\right\} \le 2\exp\left(-\frac{z^2}{2\sigma_{\mathbb{T}}^2(\mathbb{Z})}\right). \tag{4.15}$$

4.4.1 Application of Landau–Shepp Theorem

We shall be using the following version of the Landau and Shepp [LS] [16] theorem:

Theorem [LS] *Let X_t, $t \in T$, be a real valued separable Gaussian process such that w.p.* 1, $\sup_{t\in T}|X_t| < \infty$, *then for any* $0 < \beta < 1/\left(2\sigma^2\right)$, *where* $\sigma^2 = \sup_{t\in T} Var\left(X_t\right)$, *for all* y *sufficiently large*

$$P\left\{\sup_{t\in T}|X_t| > y\right\} < \exp\left(-\beta y^2\right). \tag{4.16}$$

Refer to Landau and Shepp [16] theorem (also see Sato [23], Theorem 2.5 of Marcus and Shepp [21] and Proposition A.2.3 in [28]).

Recall the definition of L in (2.1). Since L is finite, w.p. 1, we can apply the Landau and Shepp theorem to infer that for appropriate constants $C > 0$ and $D > 0$, for all $t > 0$,

$$P\left\{L > t\right\} \le C\exp\left(-Dt^2\right). \tag{4.17}$$

4.5 Four maximal inequalities

For the following inequalities recall the mean zero Gaussian process G with covariance function defined in (2.4). Inequalities 1 and 2 are required for the proof of Proposition 2, and Inequalities 1A and 2A are needed in the proofs of Theorems 1 and 2.

Inequality 1. *For all* $0 < \varrho < \infty$ *and* $\delta > 0$ *we have for some constant* $\mu(\delta)$ *and all* $z > 0$

$$P\left\{\sup_{(t,x)\in[0,\varrho]\times\mathbb{R}} t^\delta\left|G\left(t,x\right)\right| > \varrho^\delta 2^\delta\mu\left(\delta\right) + z\right\} \le 2\exp\left(-\frac{z^2\varrho^{-2\delta}}{2^{2\delta+1}}\right) \tag{4.18}$$

and for each $n \ge 1$ *and for* $t^\delta G^{(1)}\left(t,x\right),\dots,t^\delta G^{(n)}\left(t,x\right)$ *i.i.d.* $t^\delta G\left(t,x\right)$

$$P\left\{\max_{1\le m\le n}\sup_{(t,x)\in[0,\varrho]\times\mathbb{R}}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{m}t^\delta G^{(i)}\left(t,x\right)\right| > \varrho^\delta 2^\delta\mu\left(\delta\right) + z\right\} \le 4\exp\left(-\frac{z^2\varrho^{-2\delta}}{2^{2\delta+1}}\right). \tag{4.19}$$

Proof Define for any integer $k \ge 0$,

$$\mathcal{T}_k = \left[2^{-k}, 2^{-k+1}\right]\times\mathbb{R}.$$

Theorem 5 in [13] implies that, w.p. 1, for each integer k,

$$\sup\left\{\left|G\left(t,x\right)\right| : \left(t,x\right)\in\mathcal{T}_k\right\} < \infty. \tag{4.20}$$

Notice that for any $k \ge 0$

$$\sup\left\{\left|G\left(t,x\right)\right| : \left(t,x\right)\in\mathcal{T}_k\right\} \overset{\mathrm{D}}{=} \sup\left\{\left|G\left(t,x\right)\right| : \left(t,x\right)\in\mathcal{T}_0\right\}.$$

Furthermore (4.20) and separability of $G\left(t,x\right)$ permits us to apply the Landau–Shepp theorem (see (4.16)) to get

$$\mu_0 := E\left(\sup\left\{\left|G\left(t,x\right)\right| : \left(t,x\right)\in\mathcal{T}_0\right\}\right) < \infty.$$

Thus for any integer K

$$E\left(\sup\left\{t^{\delta}\left|G\left(t,x\right)\right| : \left(t,x\right)\in\left[0,2^{-K}\right]\times\mathbb{R}\right\}\right)$$

$$\leq \mu_0\sum_{k=K}^{\infty}2^{-\delta k} = 2^{-\delta K}\mu_0/\left(1-2^{-\delta}\right) =: 2^{-\delta K}\mu\left(\delta\right).$$

This implies that, w.p. 1,

$$\sup\left\{t^{\delta}\left|G\left(t,x\right)\right| : \left(t,x\right)\in\left[0,2^{-K}\right]\times\mathbb{R}\right\} < \infty.$$

Also

$$\sup\left\{Var\left(t^{\delta}G\left(t,x\right)\right) : \left(t,x\right)\in\left[0,2^{-K}\right]\times\mathbb{R}\right\}\leq 2^{-2\delta K}.$$

Applying Borell's inequality (4.15) with $\mathbb{Z}\left(t,x\right) = t^{\delta}G\left(t,x\right)$, $\mathbb{T} = \left[0,2^{-K}\right]\times\mathbb{R}$, $E\left(\|\mathbb{Z}\|_{\mathbb{T}}\right)\leq 2^{-\delta K}\mu\left(\delta\right)$ and $\sigma_{\mathbb{T}}^{2}\left(\mathbb{Z}\right)\leq 2^{-2\delta K}$, we get for all $z>0$ and integers K

$$P\left\{\sup_{(t,x)\in[0,2^{-K}]\times\mathbb{R}}t^{\delta}\left|G\left(t,x\right)\right| > 2^{-\delta K}\mu\left(\delta\right)+z\right\}\leq 2\exp\left(-\frac{z^{2}2^{2\delta K}}{2}\right).$$

Choose any $0<\varrho<\infty$ and integer K such that $2^{-K}\geq\varrho>2^{-K-1}$. We see that

$$2^{K+1}>\varrho^{-1}\geq 2^{K}\geq\varrho^{-1}/2.$$

Hence $[0,\varrho]\times\mathbb{R}\subset\left[0,2^{-K}\right]\times\mathbb{R}$. Therefore

$$P\left\{\sup\left\{t^{\delta}\left|G\left(t,x\right)\right| : \left(t,x\right)\in[0,\varrho]\times\mathbb{R}\right\}>\varrho^{\delta}2^{\delta}\mu\left(\delta\right)+z\right\}$$

$$\leq P\left\{\sup\left\{t^{\delta}\left|G\left(t,x\right)\right| : \left(t,x\right)\in\left[0,2^{-K}\right]\times\mathbb{R}\right\}>2^{-\delta K}\mu\left(\delta\right)+z\right\}$$

$$\leq 2\exp\left(-\frac{z^{2}2^{2\delta K}}{2}\right)\leq 2\exp\left(-\frac{z^{2}\varrho^{-2\delta}}{2^{2\delta+1}}\right).$$

Inequality (4.19) follows from Lévy's inequality (see Proposition A.1.2 in van der Vaart and Wellner [28]) along with separability of the Gaussian process $t^{\delta}G\left(t,x\right)$. $\square$

Inequality 1A. *For all $0<\gamma<1<T<\infty$ we have for some constant μ and all $z>0$*

$$P\left\{\sup_{(t,x)\in\mathcal{T}(\gamma)}\left|G\left(t,x\right)\right|>\mu+z\right\}\leq 2\exp\left(-\frac{z^{2}}{2}\right) \tag{4.21}$$

and for each $n\geq 1$ and $G^{(1)}\left(t,x\right),\ldots,G^{(n)}\left(t,x\right)$ i.i.d. $G\left(t,x\right)$

$$P\left\{\max_{1\leq m\leq n}\sup_{((t,x)\in\mathcal{T}(\gamma)}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{m}G^{(i)}\left(t,x\right)\right|>\mu+z\right\}\leq 4\exp\left(-\frac{z^{2}}{2}\right). \tag{4.22}$$

Proof Theorem 5 in [13] implies that, w.p. 1,

$$\sup\left\{\left|G\left(t,x\right)\right| : \left(t,x\right)\in\mathcal{T}\left(\gamma\right)\right\}<\infty. \tag{4.23}$$

Furthermore (4.23) permits us to apply the Landau–Shepp theorem to get

$$\mu := E\left(\sup\left\{\left|G\left(t,x\right)\right| : \left(t,x\right) \in \mathcal{T}\left(\gamma\right)\right\}\right) < \infty.$$

Also

$$\sup\left\{Var\left(G\left(t,x\right)\right) : \left(t,x\right) \in \mathcal{T}\left(\gamma\right)\right\} \leq 1.$$

Applying Borell's inequality (4.15) with $\mathbb{Z}\left(t,x\right) = G\left(t,x\right)$, $\mathbb{T} = \mathcal{T}\left(\gamma\right)$, $E\left(\|\mathbb{Z}\|_{\mathbb{T}}\right) = \mu$ and $\sigma_{\mathbb{T}}^2\left(\mathbb{Z}\right) \leq 1$, we get for all $z > 0$

$$P\left\{\sup_{(t,x)\in\mathcal{T}(\gamma)} \left|G\left(t,x\right)\right| > \mu + z\right\} \leq 2\exp\left(-\frac{z^2}{2}\right).$$

Inequality (4.22) follows from Lévy's inequality and separability of the Gaussian process $G\left(t,x\right)$.

□

Inequality 2. *For all $0 < \varrho < \infty$ and $\delta > 0$ we have for some $E(\delta)$ and for suitable finite positive constants $A, A_1 > 0$, for all $z > 0$*

$$\begin{aligned} &P\left\{\max_{1\leq m\leq n} \sup_{(t,x)\in[0,\varrho]\times\mathbb{R}} \left|\sqrt{m} t^{\delta} \alpha_m\left(h_{t,x}\right)\right| > \sqrt{n} A\left(E(\delta) 2^{\delta} \varrho^{\delta} + z\right)\right\} \\ &\leq 2\left\{\exp\left(-z^2 A_1\left(2\varrho\right)^{-2\delta}\right) + \exp\left(-z\sqrt{n} A_1\left(2\varrho\right)^{-\delta}\right)\right\}. \end{aligned} \tag{4.24}$$

Note, in particular, Inequality 2 implies that for all $\lambda > 1$ there is a $d > 1$ such that

$$P\left\{\sup\left\{\left|t^{\delta}\alpha_n\left(h_{t,x}\right)\right| : \left(t,x\right) \in [0,\varrho]\times\mathbb{R}\right\} \geq d\varrho^{\delta}\sqrt{\log n}\right\} < n^{-\lambda}. \tag{4.25}$$

Proof For any $k \geq 1$ and $g \in \mathcal{C}\left[0,\varrho\right]$, let

$$g_k\left(t\right) = 2^{kH} g\left(t 2^{-k}\right),\ t \in \left[0,\varrho\right],$$

and for any $k \geq 1$, $t \in [0,\varrho]$, $x \in \mathbb{R}$ and $g \in \mathcal{C}\left[0,\varrho\right]$ set

$$h_{t,x,k}\left(g\right) = h_{t,x}\left(g_k\right) = 1\left\{g_k\left(t\right) \leq x\right\}.$$

Clearly w.p. 1

$$\sup_{(t,x)\in\mathcal{T}_k} \left|\sum_{i=1}^n \epsilon_i h_{t,x}(B_i)\right| = \sup_{(t,x)\in\mathcal{T}_0} \left|\sum_{i=1}^n \epsilon_i h_{t,x,k}(B_i)\right|.$$

Moreover, since

$$\left\{B_j\right\}_{j\geq1} \stackrel{\mathrm{D}}{=} \left\{2^{kH} B_j\left(\cdot/2^k\right)\right\}_{j\geq1},$$

we see that

$$\sup_{(t,x)\in\mathcal{T}_k} \left|\sum_{i=1}^n \epsilon_i h_{t,x,k}(B_i)\right| \stackrel{\mathrm{D}}{=} \sup_{(t,x)\in\mathcal{T}_0} \left|\sum_{i=1}^n \epsilon_i h_{t,x}(B_i)\right|$$

and thus

$$E\sup_{(t,x)\in\mathcal{T}_k} \left|\sum_{i=1}^n \epsilon_i h_{t,x}(B_i)\right| = E\sup_{(t,x)\in\mathcal{T}_0} \left|\sum_{i=1}^n \epsilon_i h_{t,x}(B_i)\right|. \tag{4.26}$$

We readily see by inequality (4.6)

$$E \sup_{(t,x)\in\mathcal{T}_0} \left|\sum_{i=1}^{n} \epsilon_i h_{t,x}(B_i)\right| \leq 2\sqrt{n} E \left\|v_n\right\|_{\mathcal{T}_0} + \sqrt{n},$$

which by (2.6) is $\leq 2\left(M\left(1,2,H\right)+1\right)\sqrt{n} =: E_0\sqrt{n}$. Thus

$$E \sup_{(t,x)\in\mathcal{T}_0} \left|\sum_{i=1}^{n} \epsilon_i h_{t,x}(B_i)\right| \leq E_0\sqrt{n}. \tag{4.27}$$

Next, for all $\delta > 0$

$$E \sup \left\{\left|t^{\delta}\sum_{i=1}^{n} \epsilon_i h_{t,x}(B_i)\right| : 0 \leq t \leq 2^{-K}, x \in \mathbb{R}\right\}$$

is by (4.26) and (4.27)

$$\leq \sum_{k=K}^{\infty} 2^{-k\delta} E \sup_{(t,x)\in\mathcal{T}_k} \left|\sum_{i=1}^{n} \epsilon_i h_{t,x}(B_i)\right| \leq E\left(\delta\right) 2^{-K\delta}\sqrt{n}\,, \tag{4.28}$$

where $E\left(\delta\right) = E_0/\left(1-2^{-\delta}\right)$.

Let

$$\mathcal{H}\left(\delta,K\right) = \left\{t^{\delta}h_{t,x} : (t,x) \in \left[0,2^{-K}\right]\times\mathbb{R}\right\}.$$

From (4.28) we get

$$E \sup \left\{\left|\sum_{i=1}^{n} \epsilon_i g(B_i)\right| : g \in \mathcal{H}\left(\delta,K\right)\right\} \leq E\left(\delta\right) 2^{-K\delta}\sqrt{n}\,. \tag{4.29}$$

Also observe that each $g \in \mathcal{H}\left(\delta,K\right)$ satisfies $|g| \leq 2^{-K\delta}$. Applying Talagrand's inequality (4.8) with $M = 2^{-K\delta}$, $\sigma^2_{\mathcal{H}(\delta,K)} = 2^{-2K\delta}$ and the bound (4.29), we get that for any $\delta > 0$ we have for suitable finite positive constants $A, A_1 > 0$, for all $z > 0$

$$\begin{aligned} &P\left\{\max_{1\leq m\leq n} ||\sqrt{m}\alpha_m||_{\mathcal{H}(\delta,K)} \geq \sqrt{n}A(E\left(\delta\right)2^{-K\delta}+z)\right\} \\ &\leq 2(\exp(-z^2A_1 2^{2K\delta}) + \exp(-z\sqrt{n}A_1 2^{K\delta})). \end{aligned} \tag{4.30}$$

Inequality (4.24) follows from inequality (4.30). To see this choose any $0 < \varrho < \infty$ and integer K such that $2^{-K} \geq \varrho > 2^{-K-1}$. We see that

$$2^{K+1} > \varrho^{-1} \geq 2^{K} \geq \varrho^{-1}/2.$$

Hence $\left\{t^{\delta}h_{t,x} : (t,x) \in [0,\varrho]\times\mathbb{R}\right\} \subset \mathcal{H}\left(\delta,K\right)$. Thus

$$\begin{aligned} &P\left\{\max_{1\leq m\leq n} \sup_{(t,x)\in[0,\varrho]\times\mathbb{R}} |\sqrt{m}t^{\delta}\alpha_m\left(h_{t,x}\right)| \geq \sqrt{n}A(E\left(\delta\right)2^{\delta}\varrho^{\delta}+z)\right\} \\ &\leq P\left\{\max_{1\leq m\leq n} ||\sqrt{m}\alpha_m||_{\mathcal{H}(\delta,K)} \geq \sqrt{n}A(E\left(\delta\right)2^{-K\delta}+z)\right\} \\ &\leq 2(\exp(-z^2A_1 2^{2K\delta}) + \exp(-z\sqrt{n}A_1 2^{K\delta})) \\ &\leq 2\left\{\exp\left(-z^2A_1(2\varrho)^{-2\delta}\right) + \exp\left(-z\sqrt{n}A_1(2\varrho)^{-\delta}\right)\right\}. \end{aligned}$$

□

Inequality 2A. *For all $0<\gamma<1<T<\infty$, we have for some some $L(\gamma,T)$ and all $z>0$ for suitable finite positive constants $A, A_1>0$, for all $z>0$*

$$P\left\{\max_{1\le m\le n}\sup_{(t,x)\in\mathcal{T}(\gamma)}\left|\sqrt{m}\alpha_m\left(h_{t,x}\right)\right|\ge\sqrt{n}A\left(L(\gamma,T)+z\right)\right\} \le 2\left\{\exp\left(-z^2A_1\right)+\exp\left(-z\sqrt{n}A_1\right)\right\}. \tag{4.31}$$

Proof We see by inequality (4.6)

$$E\sup_{(t,x)\in\mathcal{T}(\gamma)}\left|\sum_{i=1}^n\epsilon_i h_{t,x}(B_i)\right|\le 2\sqrt{n}E\left\|v_n\right\|_{\mathcal{T}(\gamma)}+\sqrt{n},$$

which by (2.6) is $\le 2\left(M\left(\gamma,T,H\right)+1\right)\sqrt{n}=:L(\gamma,T)\sqrt{n}$. Thus

$$E\sup_{(t,x)\in\mathcal{T}(\gamma)}\left|\sum_{i=1}^n\epsilon_i h_{t,x}(B_i)\right|\le L(\gamma,T)\sqrt{n}. \tag{4.32}$$

Applying Talagrand's inequality (4.8) with $M=1$, $\sigma^2_{\mathcal{F}_{(\gamma,T)}}=1$ and the bound (4.32), give (4.31).
□

Remark 8 *Actually, to apply Talagrand's inequality in the proofs of Inequalities 2 and 2A, as it is stated in (4.8), the classes of functions $\mathcal{H}\left(\delta,K\right)$ and $\mathcal{F}_{(\gamma,T)}$ should be pointwise measurable. Here we shall discuss how to take care of this detail in the proof of Inequality 2. A similar discussion works for the proof of Inequality 2A.*

For any $k\ge1$ let

$$\mathcal{H}\left(\delta,K,k\right)=\left\{g1\left\{g\in\mathcal{C}\left(k\right)\right\}:g\in\mathcal{H}\left(\delta,K\right)\right\}.$$

The class $\mathcal{H}\left(\delta,K,k\right)$ is pointwise measurable. Applying Talagrand's inequality we get with $M=2^{-K\delta}$ and $\sigma^2_{\mathcal{H}(\delta,K,k)}=2^{-2K\delta}$

$$P\left\{\max_{1\le m\le n}||\sqrt{m}\alpha_m||_{\mathcal{H}(\delta,K,k)}\ge A\left(E\left\|\sum_{i=1}^n\epsilon_i g(B_i)\right\|_{\mathcal{H}(\delta,K,k)}+t\right)\right\} \le 2\exp\left(-\frac{2^{2K\delta}A_1t^2}{n}\right)+2\exp\left(-2^{K\delta}A_1t\right).$$

Obviously by the Wang [29] result (3.1), w.p. 1, $B\in\cup_{k=1}^\infty\mathcal{C}\left(k\right)$. Therefore, w.p. 1, for any $n\ge1$, $B_1,\dots,B_n$, i.i.d. B there exists a $k\ge1$ such that uniformly in $(t,x)\in[0,\varrho]\times\mathbb{R}$, $h^{(k)}_{t,x}\left(B_i\right)=h_{t,x}\left(B_i\right)$ and $t^\delta h^{(k)}_{t,x}\left(B_i\right)=t^\delta h_{t,x}\left(B_i\right)$, for $i=1,\dots,n$. This says that, w.p. 1, for any $n\ge1$, there exists a $k\ge1$, such that uniformly in $(t,x)\in[0,\varrho]\times\mathbb{R}$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n t^\delta h_{t,x}\left(B_i\right)1\left\{B_i\notin\mathcal{C}\left(k\right)\right\}=0.$$

Furthermore

$$\sup_{(t,x)\in[0,\varrho]\times\mathbb{R}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} t^{\delta} E h_{t,x}\left(B_i\right) 1\left\{B_i \notin \mathcal{C}\left(k\right)\right\} \leq \sqrt{n}\varrho^{\delta} P\left\{B \notin \mathcal{C}\left(k\right)\right\},$$

which converges to zero for each fixed $n \geq 1$, as $k \to \infty$. By passing to the limit, as $k \to \infty$, we get for any $\delta > 0$ and $t > 0$

$$P\left\{\max_{1\leq m\leq n} ||\sqrt{m}\alpha_m||_{\mathcal{H}(\delta,K)} \geq A\left(E\left\|\sum_{i=1}^{n}\epsilon_i g(B_i)\right\|_{\mathcal{H}(\delta,K)} + t\right)\right\}$$
$$\leq 2\exp\left(-\frac{2^{2K\delta}A_1 t^2}{n}\right) + 2\exp\left(-2^{K\delta}A_1 t\right).$$

Similarly one can argue the validity of the Talagrand inequality using the index class $\mathcal{F}_{(\gamma,T)}$.

Acknowledgement The authors thank the Associate Editor for a comment that led to Remark 1. PK was partially supported by the Hungarian Scientific Research Fund OTKA PD106181, by the European Union and co-funded by the European Social Fund under the project 'Telemedicine-focused research activities on the field of Mathematics, Informatics and Medical sciences' of project number TÁMOP-4.2.2.A-11/1/KONV-2012-0073, and by a postdoctoral fellowship of the Alexander von Humboldt Foundation.

References

[1] Arcones, M. A.: On the law of the iterated logarithm for Gaussian processes. J. Theor. Probab. **8**, 877–903 (1995)

[2] Berkes, I., Philipp, W.: Approximation theorems for independent and weakly dependent random vectors. Ann. Probab. **7** 29–54 (1979)

[3] Berthet, P., Mason, D. M.: Revisiting two strong approximation results of Dudley and Philipp. High dimensional probability, IMS Lecture Notes Monogr. Ser., 51, Inst. Math. Statist., Beachwood, OH, 155–172 (2006)

[4] Borell, C.: The Brunn-Minkowski inequality in Gauss space. Invent. Math. **30**, 207–216 (1975)

[5] de la Peña, V. H. and Giné, E.: Decoupling. From dependence to independence. Randomly stopped processes. $U-$statistics and processes. Martingales and beyond. Probability and its Applications. Springer–Verlag, New York (1999)

[6] Dudley, R. M.: The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. J. Functional Analysis **1**, 290–330 (1967)

[7] Dudley, R. M.: Real Analysis and Probability. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA (1989)

[8] Einmahl, U., Mason, D. M.: Gaussian approximation of local empirical processes indexed by functions. Probab. Theory Relat. Fields **107**, 283–311 (1997)

[9] Einmahl, U., Mason, D. M.: An empirical process approach to the uniform consistency of kernel-type function estimators. J. Theor. Probab. **13**, 1–37 (2000)

[10] Kevei, P. and Mason, D. M.: Strong approximations to time dependent empirical and quantile processes based on independent fractional Brownian motions, arXiv:1308.4939.

[11] Komlós, J., Major, P., Tusnády, G.: An approximation of partial sums of independent rv's and the sample df. I, Z. Wahrsch verw Gebiete **32**, 111–131 (1975)

[12] Kosorok, M. R.: Introduction to empirical processes and semiparametric inference. Springer Series in Statistics. Springer, New York (2008)

[13] Kuelbs, J., Kurtz, T., Zinn, J.: A CLT for empirical processes involving time-dependent data. Ann. Probab. **41**, 785–816 (2013)

[14] Kuelbs, J., Zinn, J.: Empirical quantile-clts for time-dependent data. High Dimensional Probability VI, Banff, AB, 2011, Progr. in Probab., vol. 66, pp. 167-194 (2013)

[15] Kuelbs, J., Zinn, J.: Empirical quantile central limit theorems for some self-similar processes. J. Theoret. Probab. **28**, 313–336 (2015)

[16] Landau, H. J., Shepp, L. A.: On the supremum of a Gaussian process. Sankhyā Ser. A **32**, 369–378 (1970)

[17] Ledoux, M., Talagrand, M.: Comparison theorems, random geometry and some limit theorems for empirical processes. Ann. Probab. **17**, 596–631 (1989)

[18] Ledoux, M. and Talagrand, M.: Probability in Banach spaces. Isoperimetry and processes. Ergebnisse der Mathematik und ihrer Grenzgebiete (3), 23, Springer-Verlag, Berlin (1991)

[19] LePage, R. D.: Log log law for Gaussian processes. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **25**, 103–108 (1972/73)

[20] Lifshits, M. A.: Gaussian random functions. Mathematics and its Applications, 322, Kluwer Academic Publishers, Dordrecht (1995)

[21] Marcus, M. B., Shepp, L. A.: Sample behavior of Gaussian processes. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory, pp. 423–441. Univ. California Press, Berkeley, Calif., (1972)

[22] Philipp, W.: Invariance principles for independent and weakly dependent random variables. Dependence in probability and statistics (Oberwolfach, 1985), 225–268, *Progr. Probab. Statist.*, **11**, Birkhäuser Boston, Boston, MA, (1986)

[23] Satô, H.: A remark on Landau–Shepp's theorem. Sankhyā Ser. A **33**, 227–228 (1971)

[24] Shorack, G. R., Wellner, Jon A.: Empirical processes with applications to statistics. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, (1986)

[25] Swanson, J.: Weak convergence of the scaled median of independent Brownian motions. Probab. Theory Relat. Fields **138**, 269–304 (2007)

[26] Talagrand, M.: Sharper bounds for Gaussian and empirical processes. Ann. Probab. **22**, 28–76 (1994)

[27] van der Vaart, A. W.: Asymptotic statistics. Cambridge Series in Statistical and Probabilistic Mathematics, 3. Cambridge University Press, Cambridge (1998)

[28] van der Vaart, A. W., Wellner, J. A.: Weak convergence and empirical processes. With applications to statistics. Springer Series in Statistics. Springer, New York (1996)

[29] Wang, W.: On a functional limit result for increments of a fractional Brownian motion. Acta Math. Hungar. **93** 153–170 (2001)

[30] Zaitsev, A. Yu.: Estimates of the Lévy-Prokhorov distance in the multivariate central limit theorem for random variables with finite exponential moments. Theory Probab. Appl. **31**, 203–220 (1987a)

[31] Zaitsev, A. Yu.: On the Gaussian approximation of convolutions under multidimensional analogues of S. N. Bernstein's inequality conditions. Probab. Theory Relat. Fields 74, 534–566 (1987b)

[32] Zaitsev, A. Yu.: The accuracy of strong Gaussian approximation for sums of independent random vectors Russian Math. Surveys 68, 721–761(2013)