

# Besov-Orlicz moduli of Brownian motion and polygonal partial sum processes

Fabian Mies\*

Delft University of Technology  
Delft, The Netherlands

March 26, 2026

## Abstract

The sample paths of Brownian motion are known to admit the exact Besov-type smoothness exponent  $1/2$  when measured in the sub-Gaussian Orlicz norm. We extend these regularity results by deriving the exact limit of the sub-Gaussian Orlicz modulus for Brownian motion in Banach spaces, and we provide a rate of convergence towards this limiting value. The central technique is a new chaining bound for the Orlicz modulus of a stochastic process. The latter also applies to polygonal partial sum processes of functional random variables and allows us to strengthen Donsker's invariance principle to all function spaces on the Besov-Orlicz scale up to the exact modulus with exponent  $1/2$ . For the critical case, we establish the thresholded weak convergence of the Besov-Orlicz seminorm of the partial sum process. The analytical results find application in a nonparametric statistical testing problem, where Besov-Orlicz statistics are shown to detect a broader range of alternatives compared to Hölderian multiscale statistics.

**Keywords:** functional data, Donsker's Theorem, sub-Gaussian random variables, signal discovery

## 1 Introduction

For iid centered random variables  $X_t$  with unit variance, we consider the interpolated, or polygonal, partial sum process

$$S_n(u) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor un \rfloor} X_t + \frac{un - \lfloor un \rfloor}{\sqrt{n}} X_{\lfloor un \rfloor}.$$

---

\*Correspondence: f.mies@tudelft.nl

This publication is part of the project VI.Veni.242.365 of the NWO Talent Programme, which is financed by the Dutch Research Council (NWO) under the grant <https://doi.org/10.61686/AGAFX90293>.

By Donsker's Theorem, this process converges to a standard Brownian motion  $W$ , weakly in the space  $C[0, 1]$  of continuous functions. This implies that  $T(S_n) \xrightarrow{d} T(W)$  for any continuous functional  $T$  on the space  $C[0, 1]$ . As both processes  $S_n$  and  $W$  admit more path regularity beyond mere continuity, the weak convergence can also be established in smaller function spaces, which allows for the asymptotic treatment of  $T(S_n)$  for a broader class of functionals  $T$ . The bottleneck to derive these invariance principles is typically a tightness conditions. Specifically, to show weak convergence  $S_n \Rightarrow W$  in a function space  $\mathcal{F}$ , one needs to show that for any  $\epsilon > 0$  there exists a compact set  $K \subset \mathcal{F}$  such that  $P(S_n \in K) \geq 1 - \epsilon$ . Compactness in function spaces is closely related to regularity, e.g. via the Arzela-Ascoli Theorem. Thus, extensions of Donsker's invariance principle to smaller function spaces  $\mathcal{F}$  demand sharp oscillations bounds for  $W$  and  $S_n$ .

What is the smallest function space where one can expect an invariance principle to hold? A minimal requirement is that the Brownian paths lie in  $\mathcal{F}$ . Lévy's result on the modulus of continuity of Brownian motion yields the limit

$$\lim_{h \rightarrow 0} \sup_{u \in [0, 1]} \frac{|W_{u+h} - W_u|}{\rho_{1/2, 1/2}(h)} = \sqrt{2}, \quad \text{for } \rho_{\mu, \nu}(h) = h^\mu |\log(e/h)|^\nu, \quad (1)$$

and hence the smallest Hölder-type space for the Brownian paths is  $C^{\rho_{1/2, 1/2}} = B_{\infty, \infty}^{\rho_{1/2, 1/2}}$ . On the other hand, the Brownian motion admits the sharper modulus of continuity  $\rho(h) = \sqrt{h}$  if the regularity is measured not uniformly, but in a weaker norm. Indeed, Ciesielski (1991) showed that the paths belong to the Besov space  $B_{p, q}^{1/2}$  for any  $p \geq 1$ , and even to the Besov-Orlicz space  $B_{\Psi_2, \infty}^{1/2}$ ; see Section 2 for a formal definition. Since  $B_{\Psi_2, \infty}^{1/2} \subsetneq C^{\rho_{1/2, 1/2}}$ , this is a strict improvement on the earlier Hölder regularity, which also holds for Brownian motion in Banach spaces Hytönen and Veraar (2008). However, unlike Lévy's result for the uniform modulus, the exact limiting behavior of the Orlicz modulus  $\omega_{\Psi_2}(h, W)$  as  $h \rightarrow 0$  has not been found so far. Our first contribution is to show that

$$\lim_{h \rightarrow 0} \frac{\omega_{\Psi_2}(h, W)}{\sqrt{h}} \rightarrow \sqrt{\frac{8}{3}},$$

and we also provide the corresponding rate of convergence. The limit is also derived for the vector-valued case.

These regularity results for the limiting Brownian motion provide the reference framework for the possible invariance principles. For smoothness exponent  $s < \frac{1}{2}$ , the weak convergence of the partial sum process has been established in the Hölder spaces  $C^s$  (Lamperti, 1962; Hamadouche, 1998; Račkauskas and Suquet, 2004a,b; Račkauskas and Wendler, 2020), in the Besov space  $B_{p, q}^s$  for  $p, q \in [1, \infty)$ , (Morel, 2004), and in  $B_{p, \infty}^s$  (Giraud and Račkauskas, 2017). Note that  $C^s \subsetneq B_{p, \infty}^s$ , i.e. the switch to Besov spaces does not yield a sharper invariance principle. Stronger smoothness moduli of the form  $\rho_{1/2, \nu}(h) = \sqrt{h} |\log(e/h)|^\nu$  for  $\nu > 1/2$  have been investigated by Račkauskas and Suquet (2004b), establishing weak convergence of the partial sum process in  $C^{\rho_{1/2, \nu}} \subsetneq C^s$ . Recently, it has been shown that Donsker's Theorem holds in the Hölder space  $C^\rho = B_{\infty, \infty}^\rho$

for any modulus  $\rho(h) \gg \rho_{1/2,1/2}(h) = \sqrt{h|\log(e/h)|}$  as  $h \downarrow 0$  (Köhne and Mies, 2025). This in particular includes cases such that  $C^\rho \subsetneq C^{\rho_{1/2,\nu}}$  for all  $\nu > 1/2$ . In view of (1), the latter is essentially optimal. To derive an invariance principle for even sharper moduli, we need to leave the Hölder scale of functions. Currently, the only version of Donsker's Theorem for the Besov-Orlicz spaces is due to Ait Ouahra et al. (2011, 2012) who treat the case,  $B_{\Psi_2,\infty}^{\rho_{1/2,\nu}}$  for  $\nu > 1$ . Since  $C^{\rho_{1/2,1/2}} \hookrightarrow B_{\Psi_2,\infty}^{\rho_{1/2,1}} \hookrightarrow C^{\rho_{1/2,3/2}}$ , see Proposition 2.4 below, the latter Besov invariance principle is in fact not as sharp as the Hölderian invariance principles of Račkauskas and Suquet (2004b) and Köhne and Mies (2025), and hence sub-optimal. It should be noted that all mentioned results require suitable moment conditions on the random variables  $X_t$  which we do not compare here. In particular, we restrict our attention to the iid case and do not allow for nonstationarity or temporal dependence.

In this paper, we establish the weak convergence of  $S_n$  in  $B_{\Psi_2,\infty}^\rho$  for all  $\rho(h) \gg \rho_{1/2,0}(h) = \sqrt{h}$ . This is the sharpest invariance principle to date, and optimal on the Besov and Hölder scale of functions. Moreover, the convergence is shown to hold for random variables in separable Banach spaces of type 2. For the edge case  $\rho(h) = \rho_{1/2,0}(h) = \sqrt{h}$ , i.e. in the space  $B_{\Psi_2,\infty}^{1/2}$ , tightness can not be established, similar the critical Hölder space  $C^{\rho_{1/2,1/2}}$ . Instead, we use the recently developed concept of thresholded weak convergence (Köhne and Mies, 2025) and show that, for some  $\tau_0 > 0$ ,

$$(|S_n|_{B_{\Psi_2,\infty}^{1/2}} \vee \tau) \Rightarrow (|W|_{B_{\Psi_2,\infty}^{1/2}} \vee \tau), \quad \tau \geq \tau_0.$$

The proof builds on a novel bound for the Orlicz modulus of continuity of stochastic processes with sub-Gaussian increments presented in Section 3.1. The latter is applicable to both  $S_n$  and its limit  $W$ , and yields refined path properties of the Brownian motion (Section 3.2) as well as asymptotic equicontinuity of  $S_n$ , and thus weak convergence via Prokhorov's Theorem (Section 4). Our results also apply to random variables  $X_t$  taking values in a separable type 2 Banach space, and to vector-valued Brownian motion.

The new asymptotic theory for the partial sum process has implications in nonparametric statistics, where it allows to improve upon existing multiscale procedures for the signal discovery problem in certain regimes. Based on the idealized white noise model, we demonstrate the advantage of using the Besov-Orlicz norm as test statistic instead of a Hölderian multiscale procedure (Section 5). Subsequently, we employ the invariance principle to transfer this finding to the standard nonparametric regression model. Simulations illustrate the gain in statistical power and highlight the impact of analytical results about the sample path regularity of Brownian motion on statistical methodology.

All technical proofs are gathered in the Appendix.

**Notation** For a random variable  $X$ , we denote the sub-Gaussian Orlicz norm as  $\|X\|_{\Psi_2} = \inf\{K : \mathbb{E} \exp(X^2/K^2) \leq 2\}$ , and  $\Psi_2(x) = \exp(x^2) - 1$ . For real numbers  $a, b$ , we denote  $a \wedge b = \min(a, b)$  and  $a \vee b = \max(a, b)$ . For two sequences  $a_n, b_n$ , we denote  $a_n \ll b_n$  to mean that  $a_n/b_n \rightarrow 0$ .

## 2 Preliminaries on Besov-Orlicz spaces

This section provides a brief background on Besov-Orlicz spaces. Orlicz norms are common tools in probability to quantify the tails of a random variable, see (Vershynin, 2018, Sec. 2.7.1). In functional analysis, Orlicz norms serve a similar purpose to generalize  $L_p$  spaces. Here, we deal with random functions and thus need both types of norms, which we distinguish notationally. In the sequel,  $(\mathcal{X}, |\cdot|)$  denotes a normed vector space, which may be infinite dimensional.

**Definition 2.1.** Let  $\Psi : [0, \infty) \rightarrow [0, \infty)$  be a Young function, i.e. convex and increasing such that  $\Psi(x)/x \rightarrow 0$  as  $x \rightarrow 0$  and  $\Psi(x)/x \rightarrow \infty$  as  $x \rightarrow \infty$ .

For a random variable  $X$  taking values in  $\mathcal{X}$ , define the Orlicz norm

$$\|X\|_\Psi = \|X\|_{\Psi(dP)} = \inf \{K : \mathbb{E}\Psi(|X|/K) \leq 1\}.$$

For a function  $f : [0, 1] \rightarrow \mathcal{X}$ , define the Orlicz norm

$$\|f\|_\Psi = \|f\|_{\Psi(dx)} = \inf \{K : \int \Psi(|f(x)|/K) dx \leq 1\},$$

and the corresponding modulus of continuity

$$\omega_\Psi(h, f) = \inf \left\{ K : \int_0^{1-h} \Psi \left( \frac{|f(x+h) - f(x)|}{K} \right) dx \leq 1 \right\}, \quad h \in (0, 1).$$

As an extension of the classical Besov spaces, Pick and Sickel (1993) introduced the Besov-Orlicz spaces by measuring smoothness not in  $L_p$ , but in an arbitrary Orlicz norm. While Pick and Sickel (1993) only consider polynomial moduli  $\rho(h) = h^s$ , the extension to arbitrary smoothness is straightforward.

**Definition 2.2.** Let  $\rho : [0, 1] \rightarrow [0, \infty)$  be a modulus of continuity, i.e. increasing and continuous with  $\rho(0) = 0$ , and let  $\Psi : [0, \infty) \rightarrow [0, \infty)$  be a Young function. For any  $f : [0, 1] \rightarrow \mathcal{X}$  define the seminorm

$$|f|_{B_{\Psi, \infty}^\rho} = \sup_{h \in (0, 1)} \frac{\omega_\Psi(h, f)}{\rho(h)}.$$

The space  $B_{\Psi, \infty}^\rho$  consists of all functions  $f$  such that  $\|f\|_{B_{\Psi, \infty}^\rho} = \|f\|_\Psi + |f|_{B_{\Psi, \infty}^\rho}$  is finite, identifying functions which are identical almost everywhere. We use the shorthand  $B_{\Psi_2, \infty}^{1/2}$  for  $\Psi_2(x) = e^{x^2} - 1$  and  $\rho(h) = \sqrt{h}$ , and  $B_{\Psi_p, \infty}^p$  for  $\Psi_p(x) = e^{x^p} - 1$  for  $p \geq 1$ .

Note that the Besov-Orlicz scale nests the usual Besov spaces via the polynomial Young function  $\Psi(x) = x^p$ . Moreover, we can contrast this with the classical Hölder-type spaces  $C^\rho$ , which can be put on the Besov scale of function spaces as  $C^\rho = B_{\infty, \infty}^\rho$  by replacing the Orlicz norm  $\|\cdot\|_{\Psi(dx)}$  with the supremum norm  $\|\cdot\|_{L_\infty(dx)}$ , which may formally be interpreted as an Orlicz norm with Young function  $\Psi(x) = \infty \cdot \mathbf{1}(x > 1)$ .

**Definition 2.3.** Let  $\rho : [0, 1] \rightarrow [0, \infty)$  be a modulus of continuity, i.e. increasing and continuous with  $\rho(0) = 0$ . The space  $C^\rho$  consists of all functions  $f : [0, 1] \rightarrow \mathcal{X}$  such that the seminorm

$$|f|_{C^\rho} = \sup_{x, y \in [0, 1]} \frac{|f(x) - f(y)|}{\rho(|x - y|)}$$

is finite. We use the shorthand  $C^s$  for  $\rho(h) = h^s$ .

By the Arzela-Ascoli Theorem, the embedding  $C^\rho \hookrightarrow C^{\rho'}$  is compact whenever the two moduli satisfy  $\rho(h) \ll \rho'(h)$  as  $h \rightarrow 0$ . For Besov-Orlicz spaces, the same property may be derived via a Wavelet basis representation due to Ciesielski et al. (1993). We further extend this embedding to vector-valued Besov-Orlicz spaces.

**Proposition 2.4** (Besov-Orlicz embeddings). *Let  $\mathcal{X}$  be a separable Banach space,  $\rho$  a modulus of continuity such that  $ch \leq \rho(h) \leq Ch^r$  for some  $r > 0$  and  $0 < c < C < \infty$ , and  $\rho^*(h) = \rho(h)\sqrt{|\log h|}$ . Then there exists a continuous embedding*

$$B_{\Psi_{2, \infty}}^\rho \hookrightarrow C^{\rho^*}([0, 1]).$$

Moreover, let  $K \subset B_{\Psi_{2, \infty}}^\rho$  be bounded such that  $K_t = \{f(t) : f \in K\} \subset \mathcal{X}$  is compact for every  $t \in [0, 1]$ . Then  $K$  is relatively compact in  $B_{\Psi_{2, \infty}}^{\rho'}$  for every modulus  $\rho'$  such that  $\rho(h) \ll \rho'(h)$  as  $h \downarrow 0$ . In particular,

$$B_{\Psi_{2, \infty}}^\rho \hookrightarrow B_{\Psi_{2, \infty}}^{\rho'} \quad \text{compactly if } \mathcal{X} = \mathbb{R}.$$

Note that functions  $f \in B_{\Psi_{2, \infty}}^\rho$  are only identified up to equivalence. Hence, the embedding  $B_{\Psi_{2, \infty}}^\rho \hookrightarrow C^{\rho^*}$  consist of choosing one representative function in the equivalence class, which is continuous.

## 3 Besov-Orlicz regularity of stochastic processes with sub-Gaussian increments

### 3.1 Modulus of continuity of stochastic processes with sub-Gaussian increments

To establish tightness of the partial sum process  $S_n$  via Prokhorov's Theorem and Proposition 2.4, we need sharp control of the sub-Gaussian modulus of continuity of a stochastic process. Our main technical result is the following theorem. We particularly want to highlight the construction of the chaining grid, see (7) in the Appendix, which is crucial to achieve the sharp constant for the Gaussian case. The grid construction is inspired by the proof of Lévy's limit for the ordinary modulus of Brownian motion as presented in (Schilling, 2021, 10.3).

**Theorem 3.1.** *Let  $(Z_u)_{u \in [0, 1]}$  be a stochastic process taking values in a separable Banach space  $(\mathcal{X}, |\cdot|)$  such that*

- (i)  $\|Z_u - Z_v\|_{\Psi_2} \leq \tau\sqrt{|u - v|}$  for all  $u, v \in [0, 1]$ .
- (ii) For any  $u_1 < v_1$  and  $u_2 < v_2$  such that  $v_1 + h_0 \leq u_2$ , the increments  $(Z_{v_1} - Z_{u_1})$  and  $(Z_{v_2} - Z_{u_2})$  are independent.
- (iii)  $\omega_{\Psi_2}(h, Z) \leq \sqrt{h}\sqrt{\frac{h}{h_0}}R$  for all  $h \leq h_0$ , and a finite random variable  $R$  with  $P(R > r) \leq G(r) \rightarrow 0$  as  $r \rightarrow \infty$ .

Then for any  $\epsilon > 0$ ,

$$P\left(\sup_{h \leq \delta} \frac{\omega_{\Psi_2}(h, Z)}{\sqrt{h}} > \tau(1 + \epsilon)\right) \leq K(\tau, \epsilon, G, \delta) \xrightarrow{\delta \downarrow 0} 0.$$

For the case that  $Z_u$  is a Brownian motion, condition (ii) holds with  $h_0 = 0$  and thus condition (iii) is vacuous. The latter condition is introduced to also cover the case of the partial sum process, where it is satisfied with  $h_0 = 2/n$ . The difference is that for the Brownian case, we can use a chaining construction in time  $u$  up to the finest resolutions due to self-similarity. For the partial sum process, this construction is only fruitful up to the discretization stepsize, below which the extra smoothness due to linear interpolation can be leveraged.

A consequence of Theorem 3.1 is that

$$\limsup_{h \rightarrow 0} \frac{\omega_{\Psi_2}(h, Z)}{\sqrt{h}} \leq \tau, \quad (2)$$

which is sharp for the case of vector-valued Brownian motion, see Theorem 3.4 below. Another question about the limit (2) is the corresponding rate of convergence. To this end, we need to slightly sharpen condition (i) as follows:

- (i)'  $\mathbb{E} \exp(\lambda|Z_u - Z_v|) \leq \kappa \exp\left(\frac{3}{16}\tau^2\lambda^2|u - v|\right)$  for all  $u, v \in [0, 1]$  and  $\lambda \in \mathbb{R}$ .

*Remark 1.* For real-valued processes  $Z_u$ , it might appear more natural to consider the condition

$$\mathbb{E} \exp(\lambda(Z_u - Z_v)) \leq \exp\left(\frac{3}{16}\tau^2\lambda^2|u - v|\right), \quad (3)$$

which is a bound on the the sub-Gaussian variance-proxy of  $Z_u - Z_v$ , which is defined via the moment generating function as

$$\sigma^*(X) = \inf \left\{ \sigma > 0 \mid \mathbb{E} \exp(\lambda X) \leq \exp\left(-\frac{\lambda^2}{2\sigma^2}\right) \text{ for all } \lambda \right\}.$$

This yields a norm which is equivalent to the sub-Gaussian norm  $\|X\|_{\Psi_2}$ , and Leskelä and Zhukov (2026) established the sharp inequality  $\sqrt{3/8}\|X\|_{\Psi_2} \leq \sigma^*(X) \leq \sqrt{\log 2}\|X\|_{\Psi_2}$ . Hence, the factor in the right hand side of (3) is chosen such that it implies condition (i) with the same  $\tau$ , and is indeed slightly stronger. However, for the general vector-valued case, (3) is not sensible and we instead consider the norm, which gives rise to the factor  $\kappa$  in (i)'. Due to this factor, (i)' no longer implies (i), but it nonetheless provides sufficiently strong tail bounds to improve upon Theorem 4.3. Note also that if  $Z$  is a standard univariate Brownian motion, both conditions are satisfied with  $\tau = \sqrt{8/3}$ .

Under the stronger assumption (i)', we can derive the following rate of convergence.

**Theorem 3.2.** *Let  $Z_u$  satisfy conditions (i)', (i), (ii), (iii). Then for any  $s < \frac{1}{10}$ ,*

$$\limsup_{h \rightarrow 0} h^{-s} \left[ \frac{\omega_{\Psi_2}(h, Z)}{\sqrt{h}} - \tau \right] \leq 0 \quad \text{almost surely.}$$

*This implies that*

$$\sup_{h \leq 1} h^{-s} \left[ \frac{\omega_{\Psi_2}(h, Z)}{\sqrt{h}} - \tau \right] < \infty \quad \text{almost surely.}$$

As part of the proof of Theorem 3.2, we derive and use the following tail bound for the modulus for a single  $h$ , which is of independent interest.

**Lemma 3.3.** *Under conditions (i), (ii), (iii), there exists a  $C > 0$  such that for all  $h \in (0, 1)$ , and all  $\eta \in [0, 2]$*

$$P \left( \omega_{\Psi_2}(h, Z) \geq \tau \sqrt{h} (1 + r) \right) \leq C (h \vee h_0)^{\eta r} \left( \frac{1}{r^2 + (2 - \eta)r} \right), \quad r \in [0, \frac{1}{2}].$$

*Under conditions (i)', (ii), and (iii), there exists for any  $p \in (1, \frac{4}{3})$  a  $C_p$  such that for all  $h \in (0, 1)$ ,*

$$P \left( \omega_{\Psi_2}(h, Z) \geq \tau \sqrt{h} (1 + r) \right) \leq C_p \left( \frac{\kappa(1 + r)}{r} \right)^p (h \vee h_0)^{p-1}, \quad r > 0.$$

*Remark 2.* Until the recent contribution of Kempka et al. (2024),  $B_{\Psi_2, \infty}^{1/2}$  was the smallest space to which the sample paths of the limiting Brownian motion  $W$  are known to belong, and thus the smallest space in which we can expect some form of weak convergence of  $S_n$ . Whether any distributional convergence also applies in the smaller space introduced by Kempka et al. (2024) remains open.

## 3.2 Regularity of vector-valued Brownian motion

As a special case of the previous results, we consider  $Z = W$  for a vector-valued Brownian motion  $W$ . To be precise, we suppose that  $W$  is a centered Brownian motion on some separable Banach space  $(\mathcal{X}, \|\cdot\|)$ , i.e.  $v \mapsto l^* W_v$  is a univariate Brownian motion for each continuous linear functional  $l^* : \mathcal{X} \rightarrow \mathbb{R}$ . Extending the work of Ciesielski (1993) on univariate Brownian motion, it was shown by Hytönen and Veraar (2008) that the vector-valued Brownian motion admits paths in  $B_{\Psi_2, \infty}^{1/2}$ . However, the exact behavior of  $\omega_{\Psi_2}(h, W)$  as  $h \rightarrow 0$  has been studied neither for the vector-valued nor for the univariate case. Via Theorem 3.1, we can obtain explicit upper bounds on the latter modulus. In Section 3.1, lower bounds on the modulus could not be obtained because we only impose upper bounds on the stochastic process. For the Brownian case, more structure is available, allowing us to derive matching lower bounds. This yields the exact limit of the modulus of continuity, and also shows that Theorem 3.1 is in general sharp.

**Theorem 3.4.** *Let  $W$  be a centered Brownian motion on some separable Banach space  $(\mathcal{X}, |\cdot|)$ , i.e.  $v \mapsto l^*W(v)$  is a univariate Brownian motion for each continuous linear functional  $l^* : \mathcal{X} \rightarrow \mathbb{R}$ . Then,*

$$\limsup_{h \rightarrow 0} \frac{\omega_{\Psi_2}(h, W)}{\sqrt{h}} = \|W(1)\|_{\Psi_2} \quad \text{almost surely.}$$

For the univariate case  $\mathcal{X} = \mathbb{R}$ , we find

$$\limsup_{h \rightarrow 0} \frac{\omega_{\Psi_2}(h, W)}{\sqrt{h}} = \sqrt{\frac{8}{3}} \quad \text{almost surely.}$$

For comparison, in the real-valued case, Marcus and Rosen (2008b) showed that the  $L_p$  modulus converges as  $\lim_{h \rightarrow 0} \omega_{L_p}(h, W)/\sqrt{h} = \mathbb{E}|W(1)|^p$  for any  $p \geq 1$ , and a corresponding central limit theorem corresponding is derived in Marcus and Rosen (2008a). The same limit for the  $L_p$  modulus of Banach-space valued Brownian motion has been derived by Hytönen and Veraar (2008); see equation (4.1) therein. In contrast, the presented result on the sub-Gaussian modulus of continuity is novel. We note that similar proof techniques could also yield the exact modulus in the exponential-type Orlicz norms  $\Psi_p(x) = \exp(x^p) - 1$ , with  $p \in [1, 2)$ .

For the univariate case, we also obtain rates of convergence for the modulus of continuity as a consequence of Theorem 3.2 and Lemma 3.3.

**Theorem 3.5.** *Let  $W$  be a standard Brownian motion taking values in  $\mathbb{R}$ . Then, for any  $s < \frac{1}{4}$ , and any  $s^* < \frac{1}{10}$ ,*

$$\max \left[ 0, \frac{\omega_{\Psi_2}(h, W)}{\sqrt{h}} - \sqrt{\frac{8}{3}} \right] = \begin{cases} \mathcal{O}_P(h^s), \\ \mathcal{O}_{\text{a.s.}}(h^{s^*}) \end{cases} \quad \text{almost surely.}$$

We conjecture that the faster rate in probability could be sharpened to the boundary case  $s = \frac{1}{4}$ , but whether the second convergence rate can be improved to match the faster rate is unclear.

To derive the rate of convergence for the vector-valued case, we would require a quantitative tail bound as in (i)' for the Brownian motion in Banach spaces. Fernique's Theorem states that  $\mathbb{E} \exp(\eta \|W(1)\|^2) < \infty$  for some  $\eta > 0$ , see e.g. (Da Prato and Zabczyk, 2014, Thm. 2.7), and thus (i)' does hold for some  $\tau$ . However, this  $\tau$  is in general not explicit and it is unknown (to us) whether  $\tau = \|W(1)\|_{\Psi_2}$ . Thus, we leave the extension of Theorem 3.5 to the vector-valued case open.

## 4 Functional central limit theorems

We now return to the asymptotics of the polygonal partial sum process  $S_n$ . For iid real-valued random variables  $Y_t$ , Donsker's Theorem yields weak convergence in  $C[0, 1]$ . The standard approach to prove this is to first establish the finite dimensional weak convergence of  $(S_n(u_1), \dots, S_n(u_m))$  for any  $u_1, \dots, u_m \in [0, 1]$ , and then extend this

to functional weak convergence via tightness of the process  $u \mapsto S_n(u)$ . If the  $Y_t$  are sub-Gaussian, Theorem 3.1 shows that  $S_n \in B_{\Psi_{2,\infty}}^{1/2}$  almost surely, and the compact embedding  $B_{\Psi_{2,\infty}}^{1/2} \hookrightarrow C[0,1]$  yields tightness. In view of the compact Besov-Orlicz embedding  $B_{\Psi_{2,\infty}}^{1/2} \hookrightarrow B_{\Psi_{2,\infty}}^\rho$  for any  $\rho(h) \gg \sqrt{h}$ , see Proposition 2.4, this also yields tightness in even smaller spaces, and thus gives rise to sharper functional central limit theorems.

As the regularity result is established for normed spaces, the corresponding invariance principle also holds for functional data. Here, we specifically assume  $(\mathcal{X}, \|\cdot\|)$  to be a separable Banach space of type 2. That is, there exists a factor  $K = K(\mathcal{X})$  such that for any  $x_1, \dots, x_n \in \mathcal{X}$ , and iid Rademacher random variables  $\epsilon_i$ , we have  $\mathbb{E}\|\sum_{i=1}^n \epsilon_i x_i\|^2 \leq K \sum_{i=1}^n \|x_i\|^2$ . This includes all Hilbert spaces, in particular  $\mathbb{R}^d$ , as well as any  $L_p$  function space with  $p \geq 2$ , see (Ledoux and Talagrand, 2011, Sec. 9.2). This restriction ensures that a sub-Gaussian concentration inequality for sums of iid random variables holds, see Lemma A.3, and that the central limit theorem holds for any iid sequence of square-integrable random variables Hoffmann-Jørgensen and Pisier (1976), see also (Ledoux and Talagrand, 2011, Thm. 10.5). Beyond the class of type 2 Banach spaces, more structural assumptions on the random variables  $X_i$  are required to establish a central limit theorem. For example, the Banach space  $C[0,1]$  is not of type 2, and Kutta et al. (2025) require additional Hölder smoothness assumptions of the random elements  $X_t \in C[0,1]$  to establish weak convergence.

**Theorem 4.1.** *Let  $X_t$  be an iid sequence of centered random variables with values in a separable Banach space  $(\mathcal{X}, \|\cdot\|)$  of type 2, such that  $\|X_t\|_{\Psi_2} < \infty$ . Then  $S_n \Rightarrow W$  in  $B_{\Psi_{2,\infty}}^\rho$  for any modulus  $\rho(h) \gg \sqrt{h}$  as  $h \rightarrow \infty$ , where  $W(u)_{u \in [0,1]}$  is a  $\mathcal{X}$ -valued Brownian motion with  $\text{Cov}(W(1)) = \text{Cov}(X_1)$ .*

Because of the continuous embedding  $B_{\Psi_{2,\infty}}^\rho \hookrightarrow C^{\rho^*}$  with  $\rho^*(h) = \rho(h)\sqrt{|\log h|}$ , see Proposition 2.4, we also obtain a corresponding Hölderian invariance principle.

**Corollary 4.2.** *Under the conditions of Theorem 4.1, the weak convergence  $S_n \Rightarrow W$  holds in  $C^{\rho^*}$  for any modulus  $\rho^*(h) \gg \sqrt{h|\log h|}$ .*

This invariance principle in Besov spaces is the first result for moduli of continuity in the regime  $\sqrt{h} \ll \rho(h) \ll \sqrt{h|\log h|}$ . In comparison to the existing functional limit theorems in the literature, which require moduli  $\rho(h) \gg \sqrt{h|\log h|}$ , our stronger convergence result imposes slightly stricter assumptions. Specifically, we consider partial sums of iid sub-Gaussian random variables, whereas Morel (2004) and Giraudo and Račkauskas (2017) have weaker moment assumptions, and Ait Ouahra et al. (2011, 2012) show tightness of any stochastic process with sub-Gaussian increments, of which the partial sum process is an example. While these results only cover real-valued random variables  $X_t$ , some authors consider Hölderian invariance principles for vector-valued (functional) random variables  $X_t$ . In particular, Račkauskas and Suquet (2009) study linear time series in Hilbert spaces and subsequently extend this to Banach spaces (Račkauskas and Suquet, 2010), though only for moduli  $\rho(h) \gg \rho_{1/2,\nu}$  for some  $\nu > 1/2$ . We also refer to Kutta and Dörnemann (2025) for  $m$ -approximable time series with values in Banach

spaces, and to Kutta et al. (2025) for mixing time series taking values in the space  $C[0, 1]$ , both for moduli  $\rho(h) = h^\mu$  with  $\mu < 1/2$ .

The modulus  $\rho(h) = \rho_{1/2,0}(h) = \sqrt{h}$  is critical as this is exactly the Besov-Orlicz regularity of the limiting Brownian motion. In this edge case, we can not establish weak convergence of the partial sum process, but it is nonetheless possible to obtain the distributional limit of the Besov-Orlicz seminorm.

**Theorem 4.3.** *Under the conditions of Theorem 4.1, there exists a  $K = K(\mathcal{X}) > 0$  such that*

$$\left( |S_n|_{B_{\Psi_2, \infty}^{1/2}} \vee \tau \right) \Rightarrow \left( |W|_{B_{\Psi_2, \infty}^{1/2}} \vee \tau \right), \quad \text{for all } \tau > K \|X\|_{\Psi_2}. \quad (4)$$

If  $\mathcal{X} = \mathbb{R}$  and  $\mathbb{E} \exp(\lambda X_i) \leq \exp(\sigma^2 \lambda^2 / 2)$  for all  $\lambda$ , then (4) holds for  $\tau > \sigma \sqrt{8/3}$ .

Theorem 4.3 shows that the Besov-Orlicz seminorm of  $S_n$  converges in distribution, though not on its full support but only beyond a certain value  $\tau$ . Note that the  $\tau$  does not depend on  $n$ . This type of thresholded weak convergence has been discovered recently (Köhne and Mies, 2025) when studying the Hölderian seminorm  $|S_n|_{C^{\rho_{1/2,1/2}}}$  for  $\rho_{1/2,1/2}(h) = \sqrt{h(\log e/h)}$ , which is the regularity of the limiting Brownian motion in the Hölderian framework. The present result allows for a sharper modulus of continuity as smoothness is measured in the functional sub-Gaussian norm, rather than uniformly. Moreover, the prior work Köhne and Mies (2025) did not study infinite-dimensional random variables  $X_i$ .

**Example 4.4.** Standard Gaussian random variables  $X_t \sim \mathcal{N}(0, 1)$  satisfy the condition of Theorem 4.3 with  $\sigma = 1$ , and the threshold  $\tau = \sqrt{8/3}$  matches exactly the limit in Theorem 3.4. In other words,  $\|W\|_{B_{\Psi_2, \infty}^{1/2}} \geq \tau$  almost surely, and hence the convergence holds indeed on the whole support. In this sense, Theorem 4.3 is sharp. This also justifies to approximate  $\|W\|_{B_{\Psi_2, \infty}^{1/2}}$  by  $\|S_n\|_{B_{\Psi_2, \infty}^{1/2}}$  with iid standard Gaussian  $X_i$ .

**Example 4.5.** The asymptotic results for the partial sum process can be applied to various functionals of interest in mathematical statistic. We consider a recent example in the context of nonparametric changepoint detection. Let  $Y_i$  be iid real-valued random vectors, and let  $X_i = \mathbb{1}(\cdot \leq Y_i)$ . Then  $\frac{1}{n} \sum_{i=1}^n X_i$  is the empirical distribution function, and  $S_n$  the corresponding partial sum process. If we consider  $X_i$  as elements in the weighted  $L_2$  space  $L_2(\mu)$  for a finite measure  $\mu$  on  $\mathbb{R}^d$ , then  $X_i$  are bounded by one. Since  $L_2(\mu)$  is a Hilbert space, Theorem 4.1 and 4.3 apply. This partial sum process has been employed by Kutta et al. (2025) to design the multiscale changepoint test statistic

$$T_\rho(S_n) = \sup_{h \in [0, \frac{1}{2}]} \sup_{u \in [h, 1-h]} \frac{\| [S_n(u) - S_n(u-h)] - [S_n(u+h) - S_n(u)] \|}{\rho(h)}.$$

For any  $\rho \gg \rho_{1/2,1/2}$ , Theorem 4.1 yields the weak convergence  $S_n \Rightarrow W$  in  $C^\rho$  and thus  $T_\rho(S_n) \Rightarrow T_\rho(W)$ . This also covers moduli in the regime  $\rho_{1/2,1/2}(h) \ll \rho(h) \ll \rho_{1/2,\nu}(h)$ ,

which was previously excluded in the literature. Inspired by the new Besov-Orlicz limit theory, we may also use the statistic

$$T_{\rho, \Psi_2}(S_n) = \inf \left\{ K : \sup_{0 < h < \frac{1}{2}} \int_h^{1-h} \Psi_2 \left( \frac{\| [S_n(u) - S_n(u-h)] - [S_n(u+h) - S_n(u)] \|}{\rho(h)} \right) du \leq 1 \right\}.$$

As  $T_{\rho, \Psi_2}$  is a continuous functional on  $B_{\Psi_2, \infty}^\rho$ , Theorem 4.1 yields  $T_{\rho, \Psi_2}(S_n) \Rightarrow T_{\rho, \Psi_2}(W)$  for any  $\rho(h) \gg \sqrt{h} = \rho_{1/2, 0}$ . Moreover, arguments similar to Theorem 4.3 show that there exists some  $\tau$  such that

$$\left( T_{\rho_{1/2, 0}, \Psi_2}(S_n) \vee \tau \right) \Rightarrow \left( T_{\rho_{1/2, 0}, \Psi_2}(W) \vee \tau \right).$$

Unlike the methodology presented by Kutta et al. (2025), our asymptotic theory currently does not allow for temporally dependent noise.

**Example 4.6.** Exact evaluation of the Besov-Orlicz seminorm of  $S_n$  is not possible in practice and one needs to resort to discretization of the involved integrals. Instead, it is practically more feasible to consider the non-interpolated partial sum process  $\tilde{S}_n(u) = \frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor un \rfloor} X_t$ , and the statistic

$$\begin{aligned} D_{\rho, n}(\tilde{S}_n) &= \max_{m=1, \dots, n} \omega_{\Psi_2} \left( \frac{m}{n}, \tilde{S}_n \right) / \rho(m/n) \\ &= \max_{m=1, \dots, n} \inf \left\{ K_m > 0 : \frac{1}{n} \sum_{a=0}^{n-m} \Psi_2 \left( \frac{\| \sum_{t=a+1}^{a+m} X_t \|}{K_m \cdot \sqrt{n}} \right) \leq 1 \right\} / \rho(m/n), \end{aligned}$$

or the dyadic version

$$D_{\rho, n, \text{DYADIC}}(\tilde{S}_n) = \max_{m=\lceil 2^{-j} n \rceil, j \in \mathbb{N}_0} \omega_{\Psi_2} \left( \frac{m}{n}, \tilde{S}_n \right) / \rho(m/n).$$

The discretized but interpolated versions  $D_{\rho, n}(S_n)$  and  $D_{\rho, n, \text{DYADIC}}(S_n)$  can be defined analogously. Despite the discretization, these statistics admit the same limit distribution as the continuous variants

$$D_\rho(S_n) = |S_n|_{B_{\Psi_2, \infty}^\rho} \quad \text{and} \quad D_{\rho, \text{DYADIC}}(S_n) = \sup_{j \in \mathbb{N}_0} \omega_{\Psi_2}(2^{-j}, S_n) / \rho(2^{-j}).$$

**Theorem 4.7.** *Under the conditions of Theorem 4.1,*

$$D_{\rho, n}(S_n), D_{\rho, n}(\tilde{S}_n), D_\rho(S_n) \Rightarrow D_\rho(W),$$

*i.e. all three statistics converge to the same limit distribution. Similarly*

$$D_{\rho, n, \text{DYADIC}}(S_n), D_{\rho, n, \text{DYADIC}}(\tilde{S}_n), D_{\rho, \text{DYADIC}}(S_n) \Rightarrow D_{\rho, \text{DYADIC}}(W).$$

*Moreover, for the same  $K = K(\mathcal{X})$  as in Theorem 4.3, and for the critical modulus  $\rho(h) = \sqrt{h}$ , these statistics converge beyond any threshold  $\tau > (K + 2)\|X\|_{\Psi_2}$ .*

## 5 Relation to the signal discovery problem

As a particular application of our analytical results in mathematical statistic, we consider the nonparametric regression model

$$Y_t = f\left(\frac{t}{n}\right) + \eta_t, \quad t = 1, \dots, n, \quad (5)$$

for centered, sub-Gaussian random variables  $\eta_t$  with unit variance. When performing inference on the unknown function  $f$ , a fundamental problem to test the simple null hypothesis  $\mathcal{H}_0 : f = f_0$  for a given  $f_0$ . We focus without loss of generality on the canonical case  $f_0 \equiv 0$ , which is also known as the signal discovery problem, see Spokoiny (1998); Walther and Perry (2022) and (Giné and Nickl, 2016, Sec. 6.2.3). Via test inversion for all candidate signals  $f$  in nonparametric class of functions  $\mathcal{F}$ , statistical methods for this fundamental signal discovery problem readily enable goodness-of-fit testing, nonparametric confidence intervals, and localization of structural breaks, as outlined elsewhere (Köhne and Mies, 2025). Hence, the goal is to develop a test which is consistent against a rich nonparametric class of alternative signals  $f$ .

As a motivating example, consider the sequence of local alternatives  $f_n(u) = \delta_n \mathbb{1}(u \in [a_n, b_n])$  for  $0 \leq a_n < b_n \leq 1$  and  $\delta_n > 0$ ,  $|b_n - a_n| \geq \frac{1}{n}$ . This alternative can be detected with the global scan statistic  $\max_I T_n(I)$  where  $T_n(I) = |\sum_{t \in I} Y_t|/\sqrt{I}$  for any integer interval  $I \subset [1, n]$ . Assuming standard Gaussian errors, and upon choosing a suitable critical value of order  $\sqrt{\log n}$  based on an extreme value limit distribution, the latter test is consistent in the regime  $n\delta_n^2|b_n - a_n| > 2 \log(n)$  (Sharpnack and Arias-Castro, 2016). A better detection threshold is achieved by the multiscale scan statistic of Dümbgen and Spokoiny (2001) of the form  $T_n^{\text{DS}} = \max_I T_n(I) - \sqrt{2 \log(en/|I|)}$ , which yields a consistent test if  $n\delta_n^2|b_n - a_n| > 2 \log(e/|b_n - a_n|)$ . Since its inception, the concept of multiscale scan statistics has been exploited and developed by various researchers, e.g. for changepoint detection (Frick et al., 2014) and density estimation (Dümbgen and Walther, 2008). For non-Gaussian errors, the distribution of  $T_n^{\text{DS}}$  can be approximated by its Gaussian analogue, but only if a lower bound on the interval length  $I$  is imposed (Frick et al., 2014; Dette et al., 2020). Thus, as discussed in Köhne and Mies (2025), the method does not achieve the minimax rate of detection against very short signals. Instead, it is argued therein that the multiplicatively weighted multiscale statistic

$$T_n^{\text{KM}} = \max_{I \subset [1, n]} \frac{T_n(I)}{\sqrt{\log(en/|I|)}},$$

is more robust, as the central limit theorem applies to  $T_n^{\text{KM}}$  without a lower bound on the interval length. Indeed, the statistic  $T_n^{\text{KM}}$  yields a consistent test if  $n\delta_n^2|b_n - a_n| \gg \log(e/|b_n - a_n|)$ , where the slightly stronger constraint is the price of not assuming Gaussianity.

A fundamental limitation of these supremum-type statistics is that they only use local information. If there are many disjoint intervals  $I$  for which  $T_n(I)$  almost exceeds the critical value, the multiscale test will still not reject the null.

**Proposition 5.1.** *Let the errors  $\eta_t$  be standard Gaussian, and consider the sequence of alternatives  $f_n = \delta_n \sin(u/l_n)$  with  $\delta_n, l_n \rightarrow 0$  such that  $1 \ll n\delta_n^2 l_n \ll \log(1/l_n)$ . Then  $T_n^{\text{DS}}$  and  $T_n^{\text{KM}}$  are stochastically bounded, i.e. the test fails to be consistent.*

To connect this statistical problem to the regularity of Brownian motion, we note that model (5) is equivalent to observing  $\frac{1}{\sqrt{n}} \sum_{t=1}^k Y_t = \frac{1}{\sqrt{n}} \sum_{t=1}^k f(t) + \frac{1}{\sqrt{n}} \sum_{t=1}^k \eta_t$  for  $k = 1, \dots, n$ , which closely resembles the idealized Gaussian white noise model

$$Y_n(u) = \sqrt{n} \int_0^u f(v) dv + W_v, \quad u \in [0, 1], \quad (6)$$

where  $W_v$  is a Brownian motion; see Tsybakov (2008, Sec. 1.10). In this setting, the statistic  $T_n^{\text{KM}}$  corresponds to the Hölder seminorm  $|Y|_{C^{\rho_{1/2}}} = \sup_{u,v \in [0,1]} |Y(u) - Y(v)| / \rho_{1/2}(|u - v|)$  with  $\rho_{1/2}(h) = \sqrt{h \log(e/h)}$ , and the respective hypothesis test is consistent against any sequence of alternatives  $f_n$  such that  $|\sqrt{n}F_n|_{C^{\rho_{1/2}}} \rightarrow \infty$ , where  $F_n(u) = \int_0^u f_n(v) dv$ . For  $f_n = \delta_n \sin(u/l_n)$  as in Proposition 5.1, we have  $|F_n|_{C^{\rho_{1/2}}} \propto \sqrt{\delta_n^2 l_n / \log(1/l_n)}$ , which reveals the regime where the test is inconsistent.

By moving from supremum-type statistics to functionals based on Orlicz norms, it is possible to make better use of the global structure of the signal. Thus, we suggest to test for  $f \equiv 0$  via the test statistic  $|Y|_{B_{\Psi_2, \infty}^{1/2}}$ , and use the quantiles of the random variable  $|W|_{B_{\Psi_2, \infty}^{1/2}}$  as critical values. This yields a consistent test if  $|\sqrt{n}F_n|_{B_{\Psi_2, \infty}^{1/2}} \rightarrow \infty$ . Because the embedding  $C^{\rho_{1/2}} \hookrightarrow B_{\Psi_2, \infty}^{1/2}$  is strict, this statistic is powerful against a broader range of alternatives  $f_n$  where  $|\sqrt{n}F_n|_{B_{\Psi_2, \infty}^{1/2}} \rightarrow \infty$  while  $|\sqrt{n}F_n|_{C^{\rho_{1/2}}} \rightarrow 0$ . This specifically helps for the signals introduced above, without making explicit use of the periodicity. It turns out that the nonparametric Besov-Orlicz test detects the sinusoidal alternative if  $n\delta_n^2 l_n \gg 1$ , where classical multiscale tests fail. This is a consequence of the following more general result about lower bounds on Besov-Orlicz norms.

**Proposition 5.2.** *Let  $f : [0, 1] \rightarrow \mathbb{R}$  and  $I_1, \dots, I_m \subset [0, 1]$  be disjoint intervals such that  $|I_k| \geq 2h$  for each  $k$ , and either  $f(u) \geq \delta$  for all  $u \in I_k$ , or  $f(u) \leq -\delta$  for all  $u \in I_k$ . Then*

$$|F|_{B_{\Psi, \infty}^{\rho}} \geq \frac{\delta h}{\rho(h) \Psi^{-1}(\frac{1}{mh})}.$$

In particular, for  $\rho_{\nu}(h) = \sqrt{h} |\log h|^{1/\nu}$ ,

$$\begin{aligned} |F|_{B_{\Psi_2, \infty}^{1/2}} &\geq \delta \sqrt{h} \left| \log \frac{e}{mh} \right|^{-\frac{1}{2}}, & |F|_{B_{\infty, \infty}^{\rho_{\nu}}} &\geq \delta \sqrt{h} \left| \log \frac{e}{h} \right|^{-\frac{1}{\nu}}, \\ |F|_{B_{p, \infty}^{1/2}} &\geq \delta \sqrt{h} (mh)^{\frac{1}{p}}, & |F|_{B_{\Psi_p, \infty}^{\rho_{\nu}}} &\geq \delta \sqrt{h} \left| \log \frac{e}{h} \right|^{-\frac{1}{\nu}} \left| \log \frac{e}{mh} \right|^{-\frac{1}{p}}. \end{aligned}$$

To compare the Besov-Orlicz statistic with the usual multiscale statistic, we contrast the lower bounds for the seminorms of  $F$  in  $B_{\Psi_2, \infty}^{1/2}$  and  $B_{\infty, \infty}^{\rho_2}$  and find that the former is bigger if  $m$  is large. For  $m \propto 1/h$ , the difference amounts to a factor  $\sqrt{|\log h|}$  which

enables the detectability of the sine signal in the regime  $1 \ll n\delta_n^2 l_n \ll \log(1/l_n)$ . We highlight that the sine signal only serves as illustration, while Proposition 5.2 illustrates the consistency of the procedure against a broader range of alternatives. In particular, our test is nonparametric and does not use the specific sinusoidal shape.

To apply these ideas beyond the idealized white noise model, and rather to the original regression model (5), we consider the interpolated partial sum process  $S_n(u)$  and use the Besov-Orlicz test statistic  $T_n^{\text{BO}} = |\sqrt{n}S_n|_{B_{\Psi_2, \infty}^{1/2}}$ . By virtue of Theorem 4.3,  $T_n^{\text{BO}}$  converges in distribution to the Besov-Orlicz seminorm of  $W$ . That is, our new functional central limit theorem makes the connection between the regression model (5) and the idealized white noise model (6) precise.

We illustrate the difference of using the Besov-Orlicz versus the Hölder as statistic via a simulation. First, we consider the power of the two test statistics in the white-noise model. For the simulation, we choose a discretization grid of mesh size  $10^{-5}$  for the discretization of the Brownian motion, the integrals for the Besov-Orlicz norm, and as evaluation points of the Hölder norm. Moreover, to save computational time, we evaluate the supremum  $h \in (0, 1)$  only on a dyadic grid. With this setup, each test statistic is simulated multiple times under the null, i.e. without a signal, and under the alternative, i.e. with a signal. The p-values under the alternative are computed based on the simulation of the null case, and thus corrected for discretization effects.

To show that the benefits of the Besov-Orlicz statistic are not restricted to the sine function, we consider the following signal, inspired by Proposition 5.2: For a length scale  $l$  and a magnitude  $\delta$ , we set

$$f_{\text{FLIP}}(x) = \sum_{k=0}^{\lceil 1/(2l) \rceil} \eta_k [\mathbf{1}(x \in [2kl, (2k+1)l]) - \mathbf{1}(x \in [(2k+1)l, 2(k+1)l])],$$

where  $\eta_k$  are iid Rademacher random variables, see Figure 1(a). In every simulation round, the signs  $\eta_k$  are sampled independently. For each lengthscale, the magnitude is chosen such that  $n\delta^2 l = |\log_2 l|^{1/4}$ . Thus, as  $l \rightarrow 0$ , Proposition 5.1 and 5.2 suggests that the Besov-Orlicz statistic should be consistent, while the Hölder statistic is inconsistent. This is in line with the empirical performance as shown in Figure 1, which demonstrates the power of the test as a function of the significance level, i.e. the ROC curve.

As a second example, we consider the Doppler function  $f_{\text{DOPPLER}}(x) = \delta \sin(4/x)$  for different  $\sqrt{n}\delta \in \{2, 4, 8\}$ , see Figure 2(a). In this case, both statistics perform comparably. This shows that the Besov-Orlicz statistic yields stronger tests in certain regimes, and generally comparable tests, which reflects the fact that  $B_{\Psi_2, \infty}^{1/2} \hookrightarrow C^{\rho_{1/2}}$  is a strict, continuous embedding.

## 6 Discussion

While the functional central limit theorem is formulated for iid random variables, an extension to independent but nonstationary data is possible because the tightness result, Theorem 3.2, does not require stationarity. For example, one might study a locally

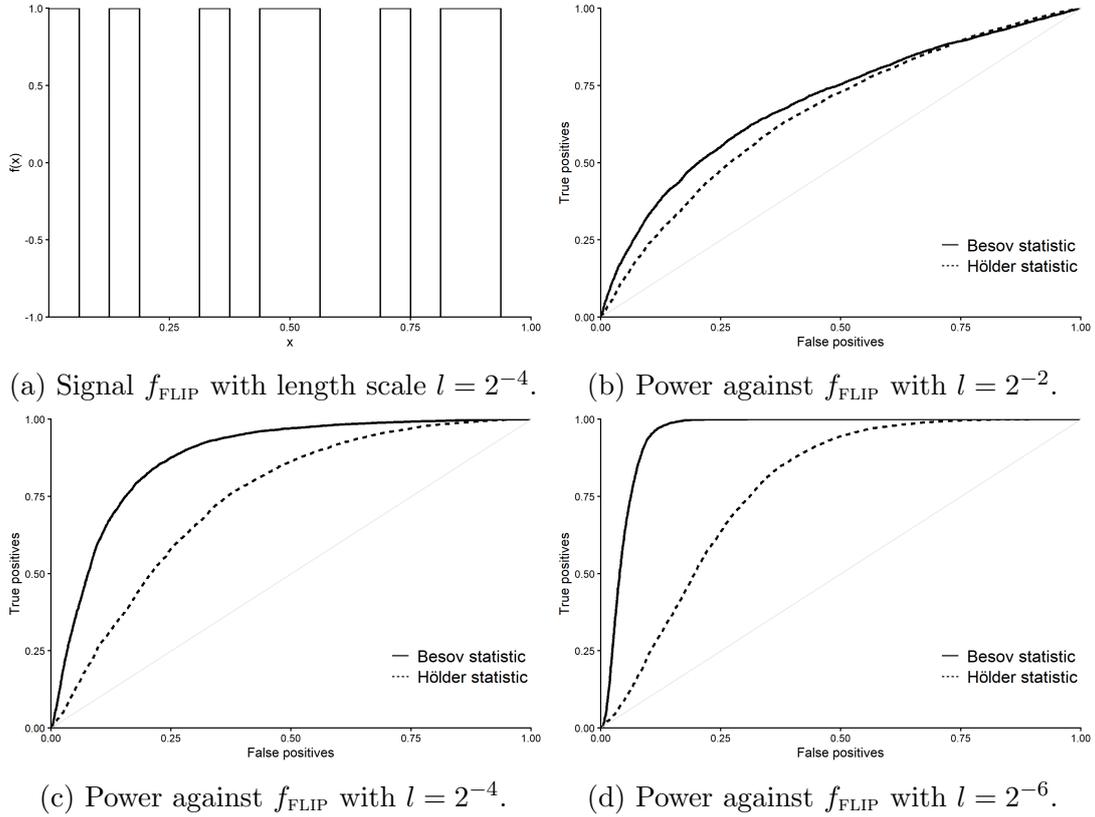
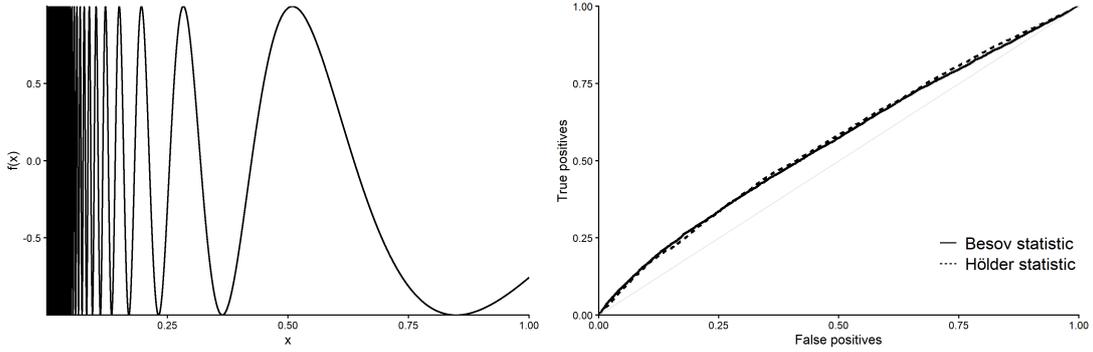
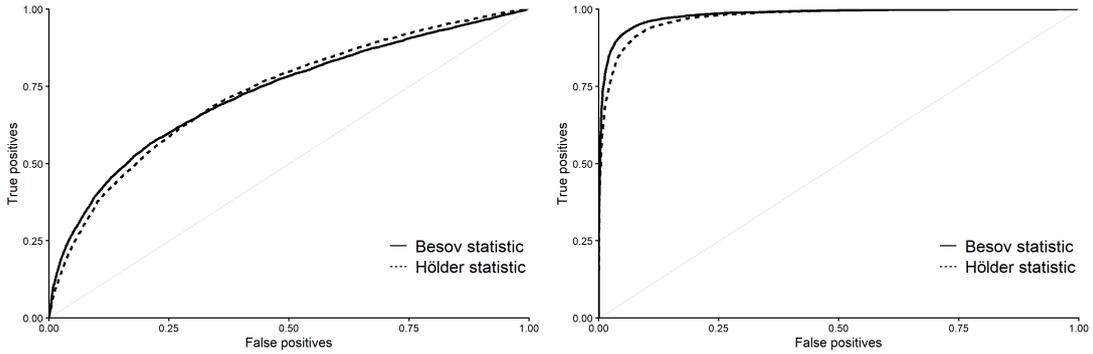


Figure 1: Example test signal (a) and ROC curves (b-d) for the Besov-Orlicz statistic  $\|Y_n\|_{B_{\Psi_{2,\infty}}^{1/2}}$  and the Hölder statistic  $\|Y_n\|_{C^{\rho_{1/2}}}$  in the white noise model, with alternatives  $f_{\text{FLIP}}$ . Reported quantities are based on  $10^4$  Monte Carlo simulations.



(a) Signal  $f_{\text{DOPPLER}}$  with  $\delta = 1$ .

(b) Power against  $f_{\text{DOPPLER}}$  with  $\sqrt{n}\delta = 2$ .



(c) Power against  $f_{\text{DOPPLER}}$  with  $\sqrt{n}\delta = 4$ .

(d) Power against  $f_{\text{DOPPLER}}$  with  $\sqrt{n}\delta = 8$ .

Figure 2: Example test signal (a) and ROC curves (b-d) for the Besov-Orlicz statistic  $\|Y_n\|_{B_{\Psi_2, \infty}^{1/2}}$  and the Hölder statistic  $\|Y_n\|_{C^{\rho_{1/2}}}$  in the white noise model, with alternatives  $f_{\text{DOPPLER}}$ . Reported quantities are based on  $10^4$  Monte Carlo simulations.

stationary array  $X_{t,n} \sim P_{t/n}$  of rows of independent random variables, where  $P_u$  is a probability distribution on  $\mathcal{X}$  for each  $u \in [0, 1]$  such that  $\|X_{t,n}\|_{\Psi_2}$  admits an upper bound. Existing invariance principles for locally stationary data show that  $S_n \Rightarrow W$  in  $C[0, 1]$  for an inhomogeneous Brownian motion  $W$ , see e.g. Mies (2023), and Theorem 3.2 allows to conclude (thresholded) weak convergence in  $B_{\Psi_2, \infty}^p$ .

The independence of the random variables is only needed for condition (ii) of Theorem 3.2, and the partial sum process would still satisfy this condition if the  $X_t$  were  $m$ -dependent instead of independent. However, a tightness result for more general temporal dependence structures would require an adaptation of the proof tailored to the specific ergodicity assumptions, such as mixing conditions (Bradley, 2005) or a decaying functional dependence measure in a nonlinear Bernoulli shift model (Wu, 2005).

Regarding the regularity results for the Brownian motion, it is not clear whether the rates established in Theorem 3.4 are sharp, and whether the difference between the almost sure rate and the rate in probability is indeed a property of Brownian motion, or merely a limitation of our proof. One approach could be to derive a distributional limit of  $h^{-1/4}[\omega_{\Psi_2}(h, W)/\sqrt{h} - \sqrt{8/3}]$  for  $h \rightarrow 0$ , as done by Marcus and Rosen (2008a) for the  $L_p$  modulus, which provides a promising avenue for future investigations.

## A Proofs

### Proof of Proposition 2.4: Embeddings

Consider  $\mathcal{X} = \mathbb{R}$  first. We start by showing that the Besov-Orlicz space can be embedded in  $C[0, 1]$ . Consider  $\mathcal{X} = \mathbb{R}$ , and let  $r > 0$  such that  $\rho(h) \ll \rho_r(h) := h^r$  as  $h \rightarrow 0$ . Then  $B_{\Psi_2, \infty}^p \subset B_{p, \infty}^r$  for any  $p > 1$ , and the latter space embeds into  $C[0, 1]$  by standard Besov embeddings. That is, for any  $f : [0, 1] \rightarrow \mathbb{R}$  such that  $\|f\|_{B_{\Psi_2, \infty}^p} \leq 1$ , there exists a  $f^* \in C[0, 1]$  with  $f^* = f$  almost everywhere. Moreover, by the continuous embedding  $B_{p, \infty}^r \hookrightarrow C^s$  for any  $s < r - \frac{1}{p}$  (Giné and Nickl, 2016, Prop. 4.3.2), this  $f^*$  is uniformly continuous with modulus of continuity  $\omega(h) \leq Ch^s \|f\|_{B_{\Psi_2, \infty}^p}$ , for some  $C$  not depending on  $f$ . In the sequel, we identify  $f$  with  $f^*$ .

Now let  $x^* \in [0, 1]$  and  $h \in [0, 1]$  such that  $|f(x^*+h) - f(x^*)| = \sup_{x \in [0, 1]} \sup_{\tilde{h} \leq h} |f(x+\tilde{h}) - f(x)| = \omega(f, h)$ . Then  $|f^*(x+h) - f(x)| \geq \omega(f, h)/2$  for all  $|x - x^*| \leq \Delta := (\omega(f, h)/(4C))^{1/s}$ , and thus

$$\begin{aligned} \omega_{\Psi_2}(f, h) &= \|f(\cdot + h) - f(\cdot)\|_{\Psi_2} \\ &\geq \inf \left\{ K > 0 : \int_{x^* \pm \Delta} \Psi_2 \left( \frac{|f(x+h) - f(x)|}{K} \right) dx \leq 1 \right\} \\ &\geq \inf \left\{ K > 0 : \Delta \cdot \Psi_2 \left( \frac{\omega(f, h)}{2K} \right) dx \leq 1 \right\} \\ &= \frac{\omega(f, h)}{2\Psi_2^{-1}(\frac{1}{\Delta})} = \frac{\omega(f, h)}{2\sqrt{\log(1 + \frac{1}{\Delta})}} \end{aligned}$$

$$\geq \frac{\omega(f, h)}{C \sqrt{\log(e + \frac{1}{\omega(f, h)})}},$$

for a sufficiently big universal  $C$ . Now note that the function  $\varphi : [0, \omega(f, 1)] \rightarrow \mathbb{R}$ ,  $z \mapsto z/\sqrt{\log(e + 1/z)}$  is strictly increasing, and thus

$$\omega(f, h) \leq \varphi^{-1}(C\omega_{\Psi_2}(f, h)).$$

Moreover, observe that

$$\varphi\left(z\sqrt{\log(e + 1/z)}\right) = z \cdot \sqrt{\frac{\log(e + \frac{1}{z})}{\log\left(e + \frac{1}{z\sqrt{\log(e + \frac{1}{z})}}\right)}} \geq z.$$

This allows us to conclude that  $\varphi^{-1}(y) \leq y\sqrt{\log(e + 1/y)}$ . Thus, for a potentially bigger  $C$ ,

$$\begin{aligned} \omega(f, h) &\leq \varphi^{-1}(C\omega_{\Psi_2}(f, h)) \leq C\omega_{\Psi_2}(f, h)\sqrt{\log(e + 1/\omega_{\Psi_2}(f, h))} \\ &\leq C\rho(h)\sqrt{\log(e + 1/\rho(h))} \\ &\leq C\rho(h)\sqrt{\log(e/h)} = C\rho^*(h), \end{aligned}$$

because  $\rho(h) \geq ch$ . This proves the continuous embedding into  $C^{\rho^*}$  for real-valued functions.

Now let  $\mathcal{X}$  be a separable Banach space with countable dense subset  $\mathcal{D}$ . For any  $f \in \overline{B_{\Psi_2, \infty}^{\rho}}$ , and  $y \in \mathcal{D}$ , define the real-valued function  $f_y(x) = \|x - y\|$  such that  $|f_y|_{B_{\Psi_2, \infty}^{\rho}} \leq |f|_{B_{\Psi_2, \infty}^{\rho}}$  by the triangle inequality. As a consequence of the real-valued embedding established above, all  $f_y$  admit a joint, uniform modulus of continuity  $\omega(h) \leq C\rho^*(h)$ . Hence, there exists a  $f_y^* \in C[0, 1]$  and a set  $\mathcal{A}_y$  with  $|\mathcal{A}_y| = 1$  and  $f_y^*(x) = f_y(x)$  for all  $x \in \mathcal{A}_y$ . Let  $\mathcal{A} = \bigcap_{y \in \mathcal{D}} \mathcal{A}_y$  such that  $|\mathcal{A}| = 1$ . The restricted function  $f : \mathcal{A} \rightarrow \mathcal{X}$  is uniformly continuous, because for any  $y \in \mathcal{D}$ ,

$$\begin{aligned} \|f(x+h) - f(x)\| &\leq \inf_{y \in \mathcal{D}} \|f(x+h) - y\| + \|f(x) - y\| \\ &\leq \inf_{y \in \mathcal{D}} \left| \|f(x+h) - y\| - \|f(x) - y\| \right| + 2\|f(x) - y\| \\ &= \inf_{y \in \mathcal{D}} |f_y(x+h) - f_y(x)| + 2\|f(x) - y\| \\ &\leq \omega(h) + 2 \inf_{y \in \mathcal{D}} \|f(x) - y\| = \omega(h). \end{aligned}$$

Hence, we can extend  $f$  continuously from  $\mathcal{A}$  to  $[0, 1]$  such that  $\|f(x+h) - f(x)\| \leq C\rho^*(h)$ . This establishes the embedding into  $C^{\rho^*}$  in the infinite dimensional case.

Regarding the compactness of the embedding, suppose that  $K$  is bounded by 1 without loss of generality. Then its closure  $\overline{K}$  is also bounded by one. Since  $K_t$  is compact and all  $f \in K$  are uniformly equicontinuous as shown above, we find that  $K$  is compact

in  $C[0, 1]$  by Arzela-Ascoli. Thus, if  $f_n \in K$  such that  $f_n \rightarrow f$  in  $B_{\Psi_2, \infty}^{\rho'}$ , then also  $f_n \rightarrow f$  uniformly, and hence  $f_n(t) \rightarrow f(t)$ . This yields that  $\overline{K}_t = \overline{(K_t)}$ , and the latter is also compact. That is, the closure  $\overline{K}$  satisfies the same qualitative properties as  $K$ , and thus to establish relative compactness, we may assume without loss of generality that  $K$  is closed.

Using again the compactness of  $K$  in  $C[0, 1]$ , let  $f_n \in K$  be a sequence such that  $f_n \rightarrow f$  in uniformly, and thus also  $\|f_n - f\|_{\Psi_2} \rightarrow 0$ . We want to show that  $f_n \rightarrow f$  in  $B_{\Psi_2, \infty}^{\rho'}$ . Because  $g \mapsto \omega_{\Psi_2}(g, h)$  is continuous on the domain  $C[0, 1]$  for each fixed  $h > 0$ , we find that  $\omega_{\Psi_2}(f, h) = \lim \omega_{\Psi_2}(f_n, h) \leq \rho(h)$ . Hence,  $\|f\|_{B_{\Psi_2, \infty}^{\rho}} \leq \sup_{g \in K} \|g\|_{B_{\Psi_2, \infty}^{\rho}} \leq 1$ . Then, for any  $\delta > 0$ ,

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \sup_{h \in (0, 1)} \frac{\omega_{\Psi_2}(f_n - f, h)}{\rho'(h)} \\ & \leq \limsup_{n \rightarrow \infty} \sup_{h \in (\delta, 1)} \frac{\omega_{\Psi_2}(f_n - f, h)}{\rho'(h)} + \limsup_{n \rightarrow \infty} \sup_{h \in (0, \delta]} \frac{\omega_{\Psi_2}(f_n - f, h)}{\rho'(h)} \\ & \leq \limsup_{n \rightarrow \infty} \sup_{h \in (\delta, 1)} \frac{\omega_{\Psi_2}(f_n - f, h)}{\rho'(h)} + \sup_{h \in (0, \delta]} \frac{2\rho(h)}{\rho'(h)} \\ & = \sup_{h \in (0, \delta]} \frac{2\rho(h)}{\rho'(h)}. \end{aligned}$$

The latter term tends to zero as  $\delta \rightarrow 0$  as  $\rho' \gg \rho$ . Since  $\delta > 0$  is arbitrary, we conclude that  $|\sup_{h \in (0, 1)} \frac{\omega_{\Psi_2}(f_n - f, h)}{\rho'(h)}| \rightarrow 0$  as  $n \rightarrow \infty$ , and thus  $f_n \rightarrow f$  in  $B_{\Psi_2, \infty}^{\rho'}$ .

Returning to the special case  $\mathcal{X} = \mathbb{R}$ , boundedness of  $K$  in  $B_{\Psi_2, \infty}^{\rho}$  implies that each  $K_t$  is bounded and hence compact. Otherwise, if there exists a sequence  $f_n \in K$  with  $f_n(t) \rightarrow \infty$ , then the equicontinuity of  $K$  would imply that  $\|f_n\|_{\Psi_2} \rightarrow \infty$  as well, which contradicts the boundedness of  $f_n$  in  $B_{\Psi_2, \infty}^{\rho}$ .  $\square$

## Proof of Theorem 3.1 and 3.2

The major part of the argument to establish Theorem 3.1 and 3.2 is the same, hence we present it jointly.

Denote  $h_{i,j} = j2^{-i}$ . Suppose we can establish that for some increasing function  $H(h)$  with  $H(0) > 0$ , and some  $\zeta \in (0, 1)$ ,

$$\omega_{\Psi_2}(h_{i,j}, Z) \leq H(h_{i,j})\tau\sqrt{h_{i,j}}, \quad i \geq \lceil \log_2 \delta \rceil, \quad j = 1, \dots, 2^{\lfloor i\zeta \rfloor}. \quad (7)$$

Then for any  $h \in (0, 1)$ , choose  $m = m(h) \in \mathbb{N}$  such that  $2^{-(m+1)(1-\zeta)} < h \leq 2^{-m(1-\zeta)}$ . We can then find a  $j(h) \leq 2^{\lfloor m\zeta \rfloor}$  such that  $j(h)2^{-m} \leq h < [j(h) + 1]2^{-m}$ . Denote  $[h] = j(h)2^{-m}$ , so that  $[h]$  is a left neighbor of  $h$  on the dyadic grid  $\{k2^{-m(h)} : k = 0, \dots, 2^{m(h)}\}$ . Thus, we may write  $h$  as the series

$$h = [h] + \sum_{k=m}^{\infty} a_k(h)2^{-k}, \quad a_k(h) \in \{0, 1\}.$$

Hence, provided that (7) holds, we find for any  $h \leq \delta$  that

$$\begin{aligned}
\omega_{\Psi_2}(h, Z) &\leq \omega_{\Psi_2}([h], Z) + \sum_{k=m(h)}^{\infty} \omega_{\Psi_2}(2^{-k}, Z) \\
&\leq H([h])\tau\sqrt{[h]} + H(2^{-m}) \sum_{k=m(h)}^{\infty} \tau 2^{-\frac{k}{2}} \\
&\leq H(h)\tau\sqrt{h} + \tau H(2^{-m(h)}) 2^{-\frac{m(h)}{2}} \frac{\sqrt{2}}{\sqrt{2}-1} \\
&\leq H(h)\tau\sqrt{h} + H(h^{\frac{1}{1-\zeta}})\tau\sqrt{h}^{\frac{1}{1-\zeta}} \frac{\sqrt{2}}{\sqrt{2}-1}, \tag{8}
\end{aligned}$$

which implies

$$\limsup_{h \rightarrow 0} \frac{\omega_{\Psi_2}(h, Z)}{\sqrt{h}} \leq \tau H(0).$$

In view of the derivations above, it is sufficient to establish (7). That is, we need to control the probability

$$\begin{aligned}
&P\left(\omega_{\Psi_2}(h_{i,j}, Z) \geq H(h_{i,j})\tau\sqrt{h_{i,j}} \text{ for some } i \geq \lceil \log_2 \delta \rceil, j \leq 2^{i\zeta}\right) \\
&\leq P\left(\omega_{\Psi_2}(h_{i,j}, Z) \geq H(h_{i,j})\tau\sqrt{h_{i,j}} \text{ for some } \lceil \log_2 \delta \rceil \leq i \leq \left\lfloor \frac{\lceil \log_2(h_0/M) \rceil}{1-\zeta} \right\rfloor, j \leq 2^{i\zeta}\right) \\
&\quad + P\left(R > H(0)\tau\sqrt{M}\right) \\
&\leq \sum_{i=\lceil \log_2 \delta \rceil}^{\left\lfloor \frac{\lceil \log_2(h_0/M) \rceil}{1-\zeta} \right\rfloor} \sum_{j=1}^{2^{\lceil i\zeta \rceil}} P\left(\omega_{\Psi_2}(h_{i,j}, Z) \geq H(h_{i,j})\tau\sqrt{h_{i,j}}\right) + P\left(R > H(0)\tau\sqrt{M}\right), \tag{9}
\end{aligned}$$

where we used condition (iii) in the last step, and  $M \in \mathbb{N}$  will be specified later. To bound (9), we proceed differently for condition (i), i.e. Theorem 3.1, and condition (i)', i.e. Theorem 3.2. Specifically, we need to choose a different function  $H$ .

*Proof of Theorem 3.1.* Choose  $H(h) = 1 + \epsilon$  constant, for any  $\epsilon > 0$ . Via Lemma 3.3, we obtain

$$\begin{aligned}
&\sum_{i=\lceil \log_2 \delta \rceil}^{\left\lfloor \frac{\lceil \log_2(h_0/M) \rceil}{1-\zeta} \right\rfloor} \sum_{j=1}^{2^{\lceil i\zeta \rceil}} P\left(\omega_{\Psi_2}(h_{i,j}, Z) \geq H(h_{i,j})\tau\sqrt{h_{i,j}}\right) \\
&\leq \sum_{i=\lceil \log_2 \delta \rceil}^{\left\lfloor \frac{\lceil \log_2(h_0/M) \rceil}{1-\zeta} \right\rfloor} \sum_{j=1}^{2^{\lceil i\zeta \rceil}} \frac{C}{\epsilon^2} (h_{i,j} \vee h_0)^{2\epsilon} \leq \frac{CM^{2\epsilon}}{\epsilon^2} \sum_{i=\lceil \log_2 \delta \rceil}^{\infty} \sum_{j=1}^{2^{\lceil i\zeta \rceil}} h_{i,j}^{2\epsilon} \\
&\leq \frac{CM^{2\epsilon}}{\epsilon^2} \sum_{i=\lceil \log_2 \delta \rceil}^{\infty} 2^{i(\zeta-2\epsilon)}.
\end{aligned}$$

Upon choosing  $\zeta < 2\epsilon$ , the latter series is summable. Now let  $M = M(\delta) \rightarrow \infty$  sufficiently slow such that the latter term tends to zero as  $\delta \rightarrow 0$ , while at the same time  $P\left(R > H(0)\tau\sqrt{M}\right) \leq G(H(0)\tau\sqrt{M}) \rightarrow 0$ . We find that the probability in (9) vanishes as  $\delta \rightarrow 0$ , hence (7) holds almost surely as  $\delta \rightarrow 0$ , and (8) yields simultaneously for all  $h \leq \delta$

$$\omega_{\Psi_2}(h, Z) \leq \tau\sqrt{h}(1 + \epsilon) \left( \frac{2\sqrt{2}}{\sqrt{2} - 1} h^{\frac{1}{2(1-\zeta)} - \frac{1}{2}} \right) \leq \tau\sqrt{h}(1 + \epsilon) \left( \frac{2\sqrt{2}}{\sqrt{2} - 1} \delta^{\frac{1}{2(1-\zeta)} - \frac{1}{2}} \right).$$

For  $\delta$  small enough, the latter bound is smaller than  $\tau\sqrt{h}(1+2\epsilon)$ . Since  $\epsilon > 0$  is arbitrary, this establishes the claim of Theorem 3.1.  $\square$

*Proof of Theorem 3.2.* Via Lemma 3.3, we obtain for any  $p \in (1, \frac{4}{3})$ , and for  $H(h) = 1 + h^s$ ,

$$\begin{aligned} & \sum_{i=\lceil \log_2 \delta \rceil}^{\lfloor \frac{\log_2(h_0/M)}{1-\zeta} \rfloor} \sum_{j=1}^{2^{\lfloor i\zeta \rfloor}} P\left(\omega_{\Psi_2}(h_{i,j}, Z) \geq H(h_{i,j})\tau\sqrt{h_{i,j}}\right) \\ & \leq C(p, \kappa) \sum_{i=\lceil \log_2 \delta \rceil}^{\lfloor \frac{\log_2(h_0/M)}{1-\zeta} \rfloor} \sum_{j=1}^{2^{\lfloor i\zeta \rfloor}} \left( \frac{H(h_{i,j})}{H(h_{i,j}) - 1} \right)^p (h_{i,j} \vee h_0)^{p-1} \\ & \leq CM^{p-1} \sum_{i=\lceil \log_2 \delta \rceil}^{\lfloor \frac{\log_2(h_0/M)}{1-\zeta} \rfloor} \sum_{j=1}^{2^{\lfloor i\zeta \rfloor}} \left( \frac{H(h_{i,j})}{H(h_{i,j}) - 1} \right)^p h_{i,j}^{p-1} \\ & \leq CH(\delta)^p M^{p-1} \sum_{i=\lceil \log_2 \delta \rceil}^{\infty} \sum_{j=1}^{2^{\lfloor i\zeta \rfloor}} (H(h_{i,j}) - 1)^{-p} h_{i,j}^{p-1} \\ & \leq C2^p M^{p-1} \sum_{i=\lceil \log_2 \delta \rceil}^{\infty} \sum_{j=1}^{2^{\lfloor i\zeta \rfloor}} h_{i,j}^{p(1-s)-1} = 2^p M^{p-1} \sum_{i=\lceil \log_2 \delta \rceil}^{\infty} 2^{-i(p(1-s)-1)} \sum_{j=1}^{2^{\lfloor i\zeta \rfloor}} j^{p(1-s)-1} \\ & \leq C2^p M^{p-1} \sum_{i=\lceil \log_2 \delta \rceil}^{\infty} 2^{i[1-p(1-s)(1-\zeta)]}. \end{aligned}$$

Upon choosing  $p < \frac{4}{3}$  big enough, the latter series is summable for any  $s < 1 - \frac{3}{4(1-\zeta)}$ . Now let  $M = M(\delta) \rightarrow \infty$  sufficiently slowly such that the latter term tends to zero as  $\delta \rightarrow 0$ , while at the same time  $P\left(R > H(0)\tau\sqrt{M}\right) \leq G(H(0)\tau\sqrt{M}) \rightarrow 0$ . We find that the probability in (9) vanishes as  $\delta \rightarrow 0$ , hence (7) holds almost surely as  $\delta \rightarrow 0$ , and (8) yields simultaneously for all  $h \leq \delta$

$$\omega_{\Psi_2}(h, Z) \leq \tau\sqrt{h} \left( 1 + h^s + \frac{2\sqrt{2}}{\sqrt{2} - 1} h^{\frac{1}{2(1-\zeta)} - \frac{1}{2}} \right).$$

To balance the rates, we choose  $\zeta = \frac{1}{6}$  and  $s < \frac{1}{4} - \zeta = \frac{1}{10}$ . This yields

$$\omega_{\Psi_2}(h, Z) \leq \tau\sqrt{h} \left( 1 + \frac{4\sqrt{2}}{\sqrt{2}-1} h^s \right),$$

which implies

$$\limsup_{h \rightarrow 0} h^{-s} \left[ \frac{\omega_{\Psi_2}(h, Z)}{\sqrt{h}} - \tau \right] \leq \tau \frac{\sqrt{2}}{\sqrt{2}-1} \quad \text{almost surely.}$$

Since  $s < \frac{1}{10}$  is arbitrary, we conclude that the limit is indeed zero.  $\square$

### Proof of Lemma 3.3

Denote  $\xi = \xi(r) = \sqrt{h}\tau(1+r)$  for brevity, and assume  $\tau = 1$  without loss of generality. We observe that

$$\begin{aligned} & P(\omega_{\Psi_2}(h, Z) > \xi) \\ & \leq P \left[ \int_0^{1-h} \Psi_2 \left( \frac{|Z(u+h)-Z(u)|}{\xi} \right) du > 1 \right] \\ & = P \left[ \int_0^{1-h} \Psi_2 \left( \frac{|Z_{u+h}-Z_u|}{\xi} \right) - \mathbb{E} \Psi_2 \left( \frac{|Z_{u+h}-Z_u|}{\xi} \right) du > 1 - \int_0^{1-h} \mathbb{E} \Psi_2 \left( \frac{|Z_{u+h}-Z_u|}{\xi} \right) du \right]. \end{aligned}$$

Now observe that  $x^2 \exp(x^2) \geq \exp(x^2) - 1$  because  $x^2 \geq 1 - \exp(-x^2) = x^2 \exp(-\tilde{x}^2)$  for some  $|\tilde{x}| \leq |x|$ , and thus

$$\begin{aligned} \frac{d}{dz} \mathbb{E} \exp(z^2 |Z_{u+h} - Z_u|^2) &= 2\mathbb{E} \{ z |Z_{u+h} - Z_u|^2 \exp(z^2 |Z_{u+h} - Z_u|^2) \} \\ &= \frac{2}{z} \mathbb{E} \{ z^2 |Z_{u+h} - Z_u|^2 \exp(z^2 |Z_{u+h} - Z_u|^2) \} \\ &\geq \frac{2}{z} \mathbb{E} [\exp(z^2 |Z_{u+h} - Z_u|^2) - 1] = \frac{2}{z} \mathbb{E} \Psi_2(z |Z_{u+h} - Z_u|). \end{aligned}$$

Hence, for some  $\tilde{r} \in [0, r]$ ,

$$\begin{aligned} \mathbb{E} \exp \left( \frac{|Z_{u+h}-Z_u|}{\xi(r)} \right)^2 &= \mathbb{E} \exp \left( \frac{|Z_{u+h}-Z_u|}{\xi(0)} \right)^2 - r \frac{d}{dr} \mathbb{E} \exp \left( \frac{|Z_{u+h}-Z_u|}{\xi(r)} \right)^2 \Big|_{r=\tilde{r}} \\ &= \mathbb{E} \exp \left( \frac{|Z_{u+h}-Z_u|}{\xi(0)} \right)^2 - r \frac{\xi'(\tilde{r})}{\xi(\tilde{r})^2} \frac{d}{dz} \mathbb{E} \exp(z^2 |Z_{u+h} - Z_u|^2) \Big|_{z=1/\xi(\tilde{r})} \\ &\leq \mathbb{E} \exp \left( \frac{|Z_{u+h}-Z_u|}{\xi(0)} \right)^2 - r \frac{\xi'(\tilde{r})}{\xi(\tilde{r})^2} 2\xi(\tilde{r}) \mathbb{E} \Psi_2 \left( \frac{|Z_{u+h} - Z_u|}{\xi(\tilde{r})} \right) \\ &\leq 2 - \frac{2r}{1+r}, \end{aligned}$$

and thus  $\mathbb{E}\Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi(r)}\right) \leq 1 - \frac{2r}{1+r}$ . We conclude that

$$1 - \int_0^{1-h} \mathbb{E}\Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi}\right) du \geq h + (1-h)\frac{2r}{1+r} \geq \frac{r}{1+r}.$$

We thus found

$$P(\omega_{\Psi_2}(h, Z) \geq \xi(r)) \leq P\left[\int_0^{1-h} \Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi(r)}\right) - \mathbb{E}\Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi(r)}\right) du > \frac{r}{1+r}\right] \quad (10)$$

For some  $M$  such that  $h > h_0/M$ , i.e.  $M > h_0/h$ , we decompose

$$\begin{aligned} & \left| \int_0^{1-h} \Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi}\right) - \mathbb{E}\Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi}\right) du \right| \\ & \leq \sum_{i=0}^M \left| \sum_{j \bmod (M+1)=i} \int_{(j-1)h}^{jh} \Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi}\right) - \mathbb{E}(\dots) du \right|. \end{aligned}$$

The  $i$ -th term is the sum of  $(hM)^{-1}$  independent random variables. Using the inequality  $\|\sum_{i=1}^n X_i\|_{L_p}^p \leq 2\sum_{i=1}^n \|X_i\|_{L_p}^p$  for any  $p \in [1, 2]$  and independent centered random variables  $X_i$ , see (Von Bahr and Esseen, 1965, Thm. 2), we find

$$\begin{aligned} & \left\| \int_0^{1-h} \Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi}\right) - \mathbb{E}(\dots) du \right\|_{L_p} \\ & \leq \sum_{i=0}^M \left\| \sum_{j \bmod (M+1)=i} \int_{(j-1)h}^{jh} \Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi}\right) - \mathbb{E}(\dots) du \right\|_{L_p} \\ & \leq 2^{\frac{1}{p}} M (hM)^{-\frac{1}{p}} \max_j \int_{(j-1)h}^{jh} \left\| \Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi}\right) - \mathbb{E}(\dots) \right\|_{L_p} du \\ & \leq 2(hM)^{1-\frac{1}{p}} \max_{u \in [0, 1-h]} \left\| \Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi}\right) - \mathbb{E}(\dots) \right\|_{L_p} \\ & \leq 4(hM)^{1-\frac{1}{p}} \max_{u \in [0, 1-h]} \left[ \mathbb{E}\Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi}\right)^p \right]^{\frac{1}{p}} \\ & \leq 16(h \vee h_0)^{1-\frac{1}{p}} \max_{u \in [0, 1-h]} \left[ \mathbb{E}\Psi_2\left(\frac{|Z_{u+h}-Z_u|}{\xi}\right)^p \right]^{\frac{1}{p}}. \end{aligned} \quad (11)$$

In the last step, we specify  $M = 1$  if  $h > h_0$ , and otherwise  $M = \lceil h_0/h \rceil \geq 2h_0/h$ . In both cases,  $hM \leq 2(h \vee h_0)$ .

To bound (11), we use the conditions (i) resp. (i)'.  
Under condition (i), Lemma A.2 below yields

$$(11) \leq 16(h \vee h_0)^{1-\frac{1}{p}} \left( 2 + \frac{2}{\frac{(1+r)^2}{p} - 1} \right)^{\frac{1}{p}}$$

In combination with (10), we obtain

$$\begin{aligned} P(\omega_{\Psi_2}(h, Z) > \xi(r)) &\leq 16^2(h \vee h_0)^{p-1} \left( 2 + \frac{2}{\frac{(1+r)^2}{p} - 1} \right) \\ &= C(h \vee h_0)^{p-1} \left( \frac{(1+r)^2}{(1+r)^2 - p} \right). \end{aligned}$$

Letting  $p = 1 + \eta r \leq 2$  for some  $\eta \in (0, 2)$  and  $r \in [0, \frac{1}{2}]$ , we find

$$P(\omega_{\Psi_2}(h, Z) > \xi(r)) = C(h \vee h_0)^{\eta r} \left( \frac{1}{r^2 + (2 - \eta)r} \right), \quad r \in [0, \frac{1}{2}].$$

Under condition (i)', Lemma A.1 below implies that

$$(11) \quad \leq 16(h \vee h_0)^{1-\frac{1}{p}} \left[ 1 - \frac{3}{4} \frac{p}{(1+r)^2} \right]^{-\frac{1}{2p}},$$

which is finite for  $p < \frac{4}{3}$ . In combination with (10), we obtain

$$\begin{aligned} P(\omega_{\Psi_2}(h, Z) > \xi(r)) &\leq 4^p \kappa^p \left( \frac{r}{1+r} \right)^{-p} (h \vee h_0)^{p-1} \left[ 1 - \frac{3}{4} \frac{p}{(1+r)^2} \right]^{-\frac{1}{2}} \\ &\leq 16^p (1 + \kappa)^p \left( \frac{r}{1+r} \right)^{-p} (h \vee h_0)^{p-1} \left[ 1 - \frac{3}{4} p \right]^{-\frac{1}{2}}. \quad \square \end{aligned}$$

**Lemma A.1.** *Condition (i)' implies*

$$\mathbb{E} \exp(\|Z_u - Z_v\|^2 / K^2) \leq \kappa \left[ 1 - \frac{3}{4} \frac{\tau^2 |u - v|}{K^2} \right]^{-\frac{1}{2}}. \quad (12)$$

*Proof of Lemma A.1.* We generalize the argument presented by Leskelä and Zhukov (2026) to the extended tail condition (i)' for Banach spaces. Let  $Z$  be a random variable taking values in  $(\mathcal{X}, \|\cdot\|)$  such that  $\mathbb{E} \exp(\lambda \|Z\|) \leq \exp(C^2 \lambda^2 / 2)$  for all  $\lambda \in \mathbb{R}$ . For any  $K > 0$ , and a standard normal random variable  $\eta$ , independent of  $Z$ ,

$$\exp(\|Z\|^2 / K^2) = \int \exp(\eta \cdot \sqrt{2} \|Z\| / K) dP(\eta),$$

and thus

$$\begin{aligned} \mathbb{E} \exp(\|Z\|^2 / K^2) &= \mathbb{E} \exp(\eta \cdot \sqrt{2} \|Z\| / K) \\ &\leq \kappa \mathbb{E} \exp(C^2 \eta^2 / K^2) \\ &= \kappa \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \exp(C^2 x^2 / K^2) dx = \frac{\kappa}{\sqrt{1 - 2C^2 / K^2}}, \end{aligned}$$

for any  $K > \sqrt{2}C$ . With  $C^2 = \frac{3}{8}\tau^2|u-v|$ , we obtain (12). In particular, for  $K = C\sqrt{8/3}$ , we find that (i)' implies (ii).  $\square$

**Lemma A.2.** For a sub-Gaussian random variable  $X$  taking values in a normed space  $(\mathcal{X}, \|\cdot\|)$ , and  $1 \leq p < K^2$ ,

$$\begin{aligned} \mathbb{E}\Psi_2\left(\frac{\|X\|}{K\|X\|_{\Psi_2}}\right)^p &\leq \int_0^\infty \frac{2}{(1+v^{1/p})^{K^2}} dv \\ &\leq \begin{cases} 2 + \frac{2}{\frac{K^2}{p}-1}, & p > 1, \\ \frac{2}{K^2-1}, & p = 1. \end{cases} \end{aligned}$$

*Proof of Lemma A.2.* By rescaling, let  $\|X\|_{\Psi_2} = 1$  without loss of generality. Then  $\mathbb{E}\Psi_2(\|X\|) \leq 1$ , i.e.  $\mathbb{E}\exp(\|X\|^2) \leq 2$ , and thus

$$\begin{aligned} \mathbb{E}\Psi_2(\|X\|/K)^p &= \int_0^\infty P(\exp(\|X\|^2/K^2) > 1 + v^{1/p}) dv \\ &= \int_0^\infty P(\exp(\|X\|^2) > (1 + v^{1/p})^{K^2}) dv \\ &\leq \int_0^\infty \frac{2}{(1 + v^{1/p})^{K^2}} dv \\ &\leq 2 + 2 \int_1^\infty v^{-K^2/p} dv = 2 + \frac{2}{\frac{K^2}{p} - 1}. \end{aligned}$$

Note that this bound is very loose. For  $p = 1$ , the integral can be computed explicitly by shifting  $v \mapsto v + 1$ .  $\square$

### Proof of Theorem 3.4: Regularity of Brownian motion

First, note that  $\|W(1)\|_{\Psi_2} < \infty$  by virtue of Fernique's Theorem, which states that  $\mathbb{E}\exp(\eta|W(1)|^2) < \infty$  for some  $\eta > 0$ ; see e.g. (Da Prato and Zabczyk, 2014, Thm. 2.7).

The upper bound  $\limsup_{h \rightarrow 0} \omega_{\Psi_2}(h, W)/\sqrt{h} \leq \|W(1)\|_{\Psi_2}$  is a consequence of Theorem 3.1 with  $h_0 = 0$  and  $\tau = \|W(1)\|_{\Psi_2}$ , and the Brownian self-similarity  $W(u+h) - W(u) \stackrel{d}{=} \sqrt{h}W(1)$ .

For the respective lower bound, choose any  $K < \|W(1)\|_{\Psi_2}$  and write

$$\begin{aligned} \int_0^{1-\frac{1}{n}} \Psi_2\left(\frac{|W(u+\frac{1}{n}) - W(u)|}{K\sqrt{\frac{1}{n}}}\right) &= \sum_{i=0}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \Psi_2\left(\frac{|W(u+\frac{1}{n}) - W(u)|}{K\sqrt{\frac{1}{n}}}\right) du \\ &\stackrel{d}{=} \frac{1}{n} \sum_{i=0}^{n-1} \int_0^1 \Psi_2\left(\frac{|W(i+v+1) - W(i+v)|}{K}\right) dv, \end{aligned}$$

using the self-similarity of the Brownian motion. Since the processes  $[W(i+v+1) - W(i+v)]_{v \in [0,1]}$  and  $[W(j+v+1) - W(j+v)]_{v \in [0,1]}$  are standard Brownian motions, and independent for  $|i-j| \geq 2$ , the standard law of large numbers yields

$$\frac{1}{n} \sum_{i=0}^{n-1} \int_0^1 \Psi_2\left(\frac{|W(i+v+1) - W(i+v)|}{K}\right) dv$$

$$\xrightarrow{a.s.} \mathbb{E} \left[ \int_0^1 \Psi_2 \left( \frac{|W(v+1) - W(v)|}{K} \right) dv \right] = \mathbb{E} \Psi_2(|W(1)|/K).$$

If  $K < \|W(1)\|_{\Psi_2}$ , then this limit is larger than 1. Hence, for any  $K < \|W(1)\|_{\Psi_2}$ , almost surely there is a (random)  $N$  such that  $\omega_{\Psi_2}(\frac{1}{n}, W)\sqrt{n} \geq K$  for all  $n \geq N$ . Hence, almost surely,  $\limsup_{h \rightarrow 0} \omega_{\Psi_2}(h, W)/\sqrt{h} \geq \|W(1)\|_{\Psi_2}$ .  $\square$

## Proof of Theorem 4.1 and Theorem 4.3: Functional weak convergence

**Lemma A.3.** *Let  $X_1, \dots, X_n$  be independent centered random variables taking values in a type-2 Banach space  $(\mathcal{X}, \|\cdot\|)$ . Then there exists a constant  $K = K(\mathcal{X})$  such that*

$$\left\| \sum_{i=1}^n X_i \right\|_{\Psi_2} \leq K \sqrt{\sum_{i=1}^n \|X_i\|_{\Psi_2}^2}.$$

*Proof of Lemma A.3.* For iid Rademacher random variables  $\epsilon_i$ , i.e.  $P(\epsilon_i = 1) = P(\epsilon_i = -1) = 1/2$ , we may use the symmetrization bound (Ledoux and Talagrand, 2011, Lemma 6.3) and a sub-Gaussian concentration inequality (Ledoux and Talagrand, 2011, Theorem 6.21) to obtain

$$\begin{aligned} \left\| \sum_{t=1}^n X_t \right\|_{\Psi_2} &\leq K \left\| \sum_{t=1}^n X_t \right\|_{L_1} + K \sqrt{\sum_{t=1}^n \|X_t\|_{\Psi_2}^2} \\ &\leq 2K \left\| \sum_{t=1}^n \epsilon_t X_t \right\|_{L_1} + K \sqrt{\sum_{t=1}^n \|X_t\|_{\Psi_2}^2}. \end{aligned}$$

Since  $\mathcal{X}$  is of type 2,

$$\left\| \sum_{t=1}^n \epsilon_t X_t \right\|_{L_1} \leq \left\| \sum_{t=1}^n \epsilon_t X_t \right\|_{L_2} \leq K(\mathcal{X}) \sqrt{\sum_{t=1}^n \|\epsilon_t X_t\|_{L_2}^2} \leq K_{\mathcal{X}} \sqrt{\sum_{t=1}^n \|\epsilon_t X_t\|_{\Psi_2}^2}.$$

$\square$

**Lemma A.4.** *Let  $X_i$  be iid centered, sub-Gaussian random variables such that  $\|X_i\|_{\Psi_2} < \infty$ . Then the interpolated partial sum process  $S_n$  satisfies conditions (ii) and (iii) of Theorem 3.1 with  $h_0 = \frac{2}{n}$ , and  $G(r) = 2 / \left( \frac{r^2}{32\tau^2} - 1 \right)$ .*

*Proof of Lemma A.4.* Condition (ii) is obvious. For condition (iii), let

$$R^* = \inf \left\{ r : \frac{1}{n} \sum_{i=1}^n \Psi_2 \left( \frac{\|X_i\|}{r} \right) \leq 1 \right\},$$

with the convention  $X_0 = 0$ . For  $h \leq \frac{1}{n}$ , and  $u \in [\frac{k}{n}, \frac{k+1}{n})$ , the construction of  $S_n$  yields  $\|S_n(u+h) - S_n(u)\| \leq h\sqrt{n} \max(\|X_{k+2}\|, \|X_{k+1}\|)$ . Thus, with the notational convention  $X_{n+1} = X_n$  and  $X_0 = 0$ , and using  $h_0 = 2/n$ ,

$$\begin{aligned} & \int_0^{1-h} \Psi_2 \left( \frac{\|S_n(u+h) - S_n(u)\|}{\sqrt{8}\sqrt{h}\sqrt{\frac{h}{h_0}}R^*} \right) du \\ & \leq \frac{1}{n} \sum_{k=0}^{n-1} \Psi_2 \left( \frac{\sqrt{nh}}{\sqrt{8}\sqrt{h}\sqrt{\frac{h}{h_0}}R^*} (\|X_{k+2}\| + \|X_{k+1}\|) \right) \\ & = \frac{1}{n} \sum_{k=0}^{n-1} \Psi_2 \left( \frac{1}{2R^*} (\|X_{k+2}\| + \|X_{k+1}\|) \right) \leq 1. \end{aligned}$$

Hence, condition (iii) holds with  $R = \sqrt{8}R^*$ . Moreover,

$$\begin{aligned} P(R^* > r) & \leq P \left( \frac{1}{n} \sum_{i=1}^n \Psi_2 \left( \frac{\|X_i\|}{r} \right) > \frac{1}{2} \right) \\ & \leq 2 \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^n \Psi_2 \left( \frac{\|X_i\|}{r} \right) \right] \\ & \leq \frac{2}{\left(\frac{r}{2\tau}\right)^2 - 1} =: G(r) \xrightarrow{r \downarrow 0} 0, \end{aligned}$$

by virtue of Lemma A.2.  $\square$

*Proof of Theorem 4.1.* Since  $\mathcal{X}$  is a Banach space of type 2, and the random variables  $X_t$  have finite second moments, they satisfy the central limit theorem, see Hoffmann-Jørgensen and Pisier (1976). From condition (ii) of Theorem 4.1 and the independence of  $X_t$ , we may readily conclude that  $(S_n(u_i))_{i=1, \dots, m} \Rightarrow (W(u_i))_{i=1, \dots, m}$  for any  $u_1, \dots, u_m \in [0, 1]$ . To establish tightness for the invariance principle, we use Theorem 4.3, and conditions (ii) and (iii) therein hold by virtue of Lemma A.4. Moreover, observe that  $S_n(u) - S_n(v) = \sum_{i=1}^n X_i w_{i,n}(u, v)$  for weights with  $\sum_{i=1}^n w_{i,n}^2 \leq |u - v|$ . Hence, via Lemma A.3,

$$\|S_n(u) - S_n(v)\|_{\Psi_2} \leq C \sqrt{\sum_{i=1}^n w_{i,n}(u, v)^2 \|X_i\|_{\Psi_2}^2} \leq C\tau \sqrt{|u - v|},$$

for some  $C = C(\mathcal{X})$ . This establishes condition (i) of Theorem 4.3. Hence,  $\|S_n\|_{B_{\Psi_2, \infty}^{1/2}}$  is stochastically bounded. Since  $B_{\Psi_2, \infty}^{1/2} \hookrightarrow B_{\Psi_2, \infty}^\rho$  is compact by Proposition 2.4, the sequence  $S_n$  is relatively compact in  $B_{\Psi_2, \infty}^\rho$ . We obtain weak convergence by virtue of Prokhorov's Theorem.  $\square$

*Proof of Theorem 4.3.* By Theorem 4.1,  $S_n \Rightarrow W$  in  $C[0, 1]$ . For any  $\delta > 0$ , the mapping  $T_{[\delta, 1]} : f \mapsto \sup_{h \in [\delta, 1]} \omega_{\Psi_2}(h, f)/\sqrt{h}$  is a continuous functional on  $C[0, 1]$ , and thus  $T_{[\delta, 1]}(S_n) \Rightarrow T_{[\delta, 1]}(W)$ . Now define  $T_{(0, \delta)}(f) = \sup_{h \in (0, \delta)} \omega_{\Psi_2}(h, f)/\sqrt{h}$ , such that  $|f|_{B_{\Psi_2, \infty}^{1/2}} = \max(T_{(0, \delta)}(f), T_{[\delta, 1]}(f))$ .

In the proof Theorem 4.1, we established the sufficient conditions of Theorem 3.1 for some  $\tau = K(\mathcal{X})\|X\|_{\Psi_2}$ . Thus for any  $t > \tau$ , we find  $\limsup_n P(T_{(0, \delta)}(S_n) > \tau + \epsilon) \rightarrow 0$  as  $\delta \rightarrow 0$ . On the other hand,  $T_{[\delta, 1]}(W) \rightarrow |W|_{B_{\Psi_2, \infty}^{1/2}}$  as  $\delta \rightarrow 0$ . Thus, for any  $t > \tau$ ,

$$\begin{aligned} \limsup_{n \rightarrow \infty} P\left(|S_n|_{B_{\Psi_2, \infty}^{1/2}} \geq t\right) &\leq \limsup_{n \rightarrow \infty} P(T_{(0, \delta)}(S_n) \geq t) + \limsup_{n \rightarrow \infty} P(T_{[\delta, 1]}(S_n) \geq t) \\ &\leq P(T_{[\delta, 1]}(W) \geq t) \xrightarrow{\delta \downarrow 0} P(T_{(0, 1)}(W) \geq t). \end{aligned}$$

This establishes the thresholded weak convergence by the Portmanteau Lemma.

Now consider the special case  $\mathcal{X} = \mathbb{R}$  with the additional condition that  $\mathbb{E} \exp(\lambda X_t) \leq \exp(\lambda^2 \sigma^2 / 2)$ . It suffices to establish condition (i) of Theorem 4.3 with  $\tau = \sigma \sqrt{8/3}$ . With the same weights  $w_{i,n}(u, v)$  as above, we find

$$\begin{aligned} \mathbb{E} \exp(\lambda[S_n(u) - S_n(v)]) &= \mathbb{E} \exp\left(\lambda \sum_{i=1}^n w_{i,n}(u, v) X_i\right) = \prod_{i=1}^n \mathbb{E} \exp(\lambda w_{i,n}(u, v) X_i) \\ &\leq \exp\left(\frac{\sigma^2}{2} \sum_{i=1}^n w_{i,n}(u, v)^2\right) \leq \exp\left(\frac{\sigma^2 |u - v|}{2}\right). \end{aligned}$$

As in Remark 1, we conclude that  $\|S_n(u) - S_n(v)\|_{\Psi_2} \leq \sigma \sqrt{8/3} \sqrt{|u - v|}$ .  $\square$

## Proof of Theorem 4.7: Discretized statistics

First, note that

$$\begin{aligned} \|S_n - \tilde{S}_n\|_{\Psi_2(dx)} &\leq \inf \left\{ K > 0 : \sum_{t=1}^n \int_{(t-1)/n}^{t/n} \Psi_2\left(\frac{\|S_n(u) - \tilde{S}_n(u)\|}{K}\right) du \right\} \\ &\leq \inf \left\{ K > 0 : \frac{1}{n} \sum_{t=1}^n \Psi_2\left(\frac{\|X_t\|}{K \cdot \sqrt{n}}\right) \right\} \\ &= \frac{1}{\sqrt{n}} \inf \left\{ K > 0 : \frac{1}{n} \sum_{t=1}^n \Psi_2\left(\frac{\|X_t\|}{K}\right) \right\}. \end{aligned}$$

The latter infimum is an empirical Orlicz norm of iid random variables, which converges almost surely to  $\|X\|_{\Psi_2(dP)}$ , see Mies (2026). Thus,

$$\limsup_{n \rightarrow \infty} \sqrt{n} \|S_n - \tilde{S}_n\|_{\Psi_2(dx)} \leq \|X\|_{\Psi_2} \quad \text{almost surely,}$$

and hence

$$\limsup_{n \rightarrow \infty} \sqrt{n} \sup_{h \in (0,1)} \left| \omega_{\Psi_2}(h, S_n) - \omega_{\Psi_2}(h, \tilde{S}_n) \right| \leq 2 \|X\|_{\Psi_2} \quad \text{almost surely.} \quad (13)$$

As  $\rho(1/n) \gg 1/\sqrt{n}$ , we conclude that  $D_{\rho,n}(\tilde{S}_n) = D_{\rho,n}(S_n) + o(1)$ . Moreover, for any  $\delta > 0$ ,

$$\begin{aligned} & |D_{\rho,n}(S_n) - D_{\rho}(S_n)| \\ & \leq \sup_{h \in [\delta, 1]} \left| \frac{\omega_{\Psi_2}(\frac{[nh]}{n}, S_n)}{\rho(\frac{[nh]}{n})} - \frac{\omega_{\Psi_2}(h, S_n)}{\rho(h)} \right| + \sup_{h \leq \delta} \frac{\omega_{\Psi_2}(h, S_n)}{\rho(h)} \\ & \leq \sup_{h \in [\delta, 1]} \left| \frac{\omega_{\Psi_2}(\frac{[nh]}{n}, S_n) - \omega_{\Psi_2}(h, S_n)}{\rho(\delta)} \right| \\ & \quad + \sup_{h \in [\delta, 1]} \frac{\omega_{\Psi_2}(h, S_n)}{\rho(h)} \left| \frac{\rho(h) - \rho(\frac{[nh]}{n})}{\rho(\delta)} \right| + \sup_{h \leq \delta} \frac{\omega_{\Psi_2}(h, S_n)}{\rho(h)} \\ & = \mathcal{O}_P \left( \frac{\sqrt{1/n}}{\rho(\delta)} \right) + o_P \left( \frac{1}{\rho(\delta)} \right) + \mathcal{O}_P \left( \frac{\sqrt{\delta}}{\rho(\delta)} \right) \quad \rightarrow 0. \end{aligned}$$

In the last step, we used (13) for the first term, the regularity of  $S_n$  for the third term, and the continuity of  $\rho$  for the second term. Now, letting  $\delta = \delta_n \rightarrow 0$  sufficiently slowly, the whole expression tends to zero. This shows that  $D_{\rho,n}(S_n)$  and  $D_{\rho,n}(\tilde{S}_n)$  have the same limit as  $D_{\rho}(S_n)$ , which converges to  $D_{\rho}(W)$  by Theorem 4.1. The same argument applies to the dyadic statistic, where  $D_{\rho, \text{DYADIC}}(S_n)$  converges as a continuous functional of  $S_n$ .

Now consider the critical case  $\rho(h) = \sqrt{h}$ . For any  $\delta > 0$ , define

$$\begin{aligned} D_{\rho,n}^{(0,\delta)}(\tilde{S}_n) &= \max_{m=1, \dots, n, \frac{m}{n} < \delta} \omega_{\Psi_2}(\frac{m}{n}, \tilde{S}_n) / \rho(\frac{m}{n}), \\ D_{\rho,n}^{[\delta,1]}(\tilde{S}_n) &= \max_{m=1, \dots, n, \frac{m}{n} \geq \delta} \omega_{\Psi_2}(\frac{m}{n}, \tilde{S}_n) / \rho(\frac{m}{n}) \end{aligned}$$

and analogously  $D_{\rho}^{(0,\delta)}$  and  $D_{\rho}^{[\delta,1]}$ . Then  $D_{\rho,n}^{[\delta,1]}(\tilde{S}_n) \Rightarrow D_{\rho}^{[\delta,1]}(W)$ , which in turn converges to  $D_{\rho}(W)$  as  $\delta \rightarrow 0$ . At the same time, for any  $\kappa > 2 \|X\|_{\Psi_2}$ ,

$$\begin{aligned} \limsup_{n \rightarrow \infty} P \left( D_{\rho,n}^{(0,\delta)}(\tilde{S}_n) > \tau \right) &\leq \limsup_{n \rightarrow \infty} P \left( D_{\rho,n}^{(0,\delta)}(S_n) > \tau - \kappa \right) \\ &\quad + \limsup_{n \rightarrow \infty} P \left( \sup_{h \in (0,1)} \left| \omega_{\Psi_2}(h, S_n) - \omega_{\Psi_2}(h, \tilde{S}_n) \right| / \rho(1/n) > \kappa \right). \end{aligned}$$

The second term tends to zero by (13), and the first term tends to zero for any  $\tau > K \|X\|_{\Psi_2} + \kappa$  as  $\delta \rightarrow 0$ , as shown in the proof of Theorem 4.3. We may thus conclude the proof as in Theorem 4.3. The dyadic statistic may be handled analogously.  $\square$

## Proof of Proposition 5.2

For any  $K > 0$ , and  $h \leq \min |I_k|/2$ ,

$$\int_0^{1-h} \Psi \left( \frac{|F(v+h) - F(v)|}{K\rho(h)} \right) dv \geq \left( \sum_{k=1}^m |I_k|/2 \right) \Psi \left( \frac{\delta h}{K\rho(h)} \right) \geq mh \Psi \left( \frac{\delta h}{K\rho(h)} \right),$$

which equals 1 for  $K = \delta h/\rho(h)/\Psi^{-1}(\frac{1}{mh})$ . Hence,

$$|F|_{B_{\Psi, \infty}^{\rho}} \geq \frac{\omega_{\Psi}(h, F)}{\rho(h)} \geq \frac{\delta h}{\rho(h)\Psi^{-1}(\frac{1}{mh})}.$$

The special cases are a consequence of the identity  $\Psi_p^{-1}(y) = \log(1+y)^{1/p}$ , such that

$$|F|_{B_{\Psi_p^{\rho\nu}, \infty}^{\rho\nu}} \geq \delta\sqrt{h} \log(e/h)^{-\frac{1}{\nu}} \log(1 + \frac{1}{mh})^{-\frac{1}{p}} \geq \delta\sqrt{h} |\log \frac{e}{h}|^{-\frac{1}{\nu}} |\log \frac{e}{mh}|^{-\frac{1}{p}}.$$

The norms in  $B_{\infty, \infty}^{\rho\nu}$  and  $B_{\Psi_{2, \infty}}^{1/2}$  are limiting cases as  $p \rightarrow \infty$  and  $\nu \rightarrow \infty$ , respectively.  $\square$

## References

- Ait Ouahra, M., Kissami, A., and Sghir, A. (2011). Un critère de tension dans les espaces de Besov-Orlicz et applications au problème du temps d'occupation. Annales mathématiques Blaise Pascal, 18(2):301–321.
- Ait Ouahra, M., Kissami, A., and Sghir, A. (2012). Un principe d'invariance de type Donsker dans une classe d'espaces de Besov-Orlicz. Annales mathématiques Blaise Pascal, 19(1):263–269.
- Bradley, R. C. (2005). Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions. Probability Surveys, 2(none).
- Ciesielski, Z. (1991). Modulus of smoothness of the Brownian paths in the  $L_p$  norm. Constructive theory of functions (Varna, Bulgaria, 1991), pages 71–75.
- Ciesielski, Z. (1993). Orlicz spaces, spline systems, and brownian motion. Constructive Approximation, 9(2-3):191–208.
- Ciesielski, Z., Kerkycharian, G., and Roynette, B. (1993). Quelques espaces fonctionnels associés à des processus gaussiens. Studia Mathematica, 107(2):171–204.
- Da Prato, G. and Zabczyk, J. (2014). Stochastic Equations in Infinite Dimensions. Cambridge University Press, 2nd edition.
- Detle, H., Eckle, T., and Vetter, M. (2020). Multiscale change point detection for dependent data. Scandinavian Journal of Statistics, 47(4):1243–1274.

- Dümbgen, L. and Spokoiny, V. G. (2001). Multiscale Testing of Qualitative Hypotheses. The Annals of Statistics, 29(1).
- Dümbgen, L. and Walther, G. (2008). Multiscale inference about a density. The Annals of Statistics, 36(4).
- Frick, K., Munk, A., and Sieling, H. (2014). Multiscale change point inference. Journal of the Royal Statistical Society: Series B: Statistical Methodology, pages 495–580.
- Giné, E. and Nickl, R. (2016). Mathematical Foundations of Infinite-Dimensional Statistical Models. Cambridge University Press.
- Giraud, D. and Račkauskas, A. (2017). Weak invariance principle in some Besov spaces for stationary martingale differences. Lithuanian Mathematical Journal, 57(4):441–467.
- Hamadouche, D. (1998). Weak convergence of smoothed empirical process in Hölder spaces. Statistics & Probability Letters, 36(4):393–400.
- Hoffmann-Jørgensen, J. and Pisier, G. (1976). The law of large numbers and the central limit theorem in Banach spaces. The Annals of Probability, pages 587–599.
- Hytönen, T. and Veraar, M. (2008). On Besov regularity of Brownian motions in infinite dimensions. Probability and Mathematical Statistics, 28(1):143–162. arXiv:0801.2959.
- Kempka, H., Schneider, C., and Vybiral, J. (2024). Path Regularity of the Brownian Motion and the Brownian Sheet. Constructive Approximation, 59(2):485–539.
- Kutta, T., Dette, H., and Wang, S. (2025). Multiscale Change Point Detection for Functional Time Series. arXiv:2511.06870.
- Kutta, T. and Dörnemann, N. (2025). Monitoring time series with short detection delay. Electronic Journal of Statistics, 19(1).
- Köhne, J. and Mies, F. (2025). At the edge of Donsker’s Theorem: Asymptotics of multiscale scan statistics. arXiv:2506.05112.
- Lamperti, J. (1962). On convergence of stochastic processes. Transactions of the American Mathematical Society, 104(3):430–435.
- Ledoux, M. and Talagrand, M. (2011). Probability in Banach spaces: isoperimetry and processes. Classics in mathematics. Springer, Berlin ; London.
- Leskelä, L. and Zhukov, M. (2026). Sharp constants relating the sub-Gaussian norm and the sub-Gaussian parameter. Electronic Communications in Probability, 31:1–11.
- Marcus, M. B. and Rosen, J. (2008a). CLT for  $L_p$  moduli of continuity of Gaussian processes. Stochastic Processes and their Applications, 118(7):1107–1135.

- Marcus, M. B. and Rosen, J. (2008b).  $L_p$  moduli of continuity of Gaussian processes and local times of symmetric Lévy processes. The Annals of Probability, 36(2).
- Mies, F. (2023). Functional estimation and change detection for nonstationary time series. Journal of the American Statistical Association, 118(542):1011–1022.
- Mies, F. (2026). Empirical Orlicz norms. Statistics and Probability Letters, to appear. arXiv:2510.25408.
- Morel, B. (2004). Weak convergence of summation processes in Besov spaces. Studia Mathematica, 165(1):19–37.
- Pick, L. and Sickel, W. (1993). Several Types of Intermediate Besov–Orlicz Spaces. Mathematische Nachrichten, 164(1):141–165.
- Račkauskas, A. and Suquet, C. (2004a). Hölder norm test statistics for epidemic change. Journal of Statistical Planning and Inference, 126(2):495–520.
- Račkauskas, A. and Suquet, C. (2004b). Necessary and Sufficient Condition for the Functional Central Limit Theorem in Hölder Spaces. Journal of Theoretical Probability, 17(1):221–243.
- Račkauskas, A. and Suquet, C. (2009). Hölderian invariance principle for Hilbertian linear processes. ESAIM: Probability and Statistics, 13:261–275.
- Račkauskas, A. and Suquet, C. (2010). On limit theorems for Banach-space-valued linear processes. Lithuanian Mathematical Journal, 50(1):71–87.
- Račkauskas, A. and Wendler, M. (2020). Convergence of U-processes in Hölder spaces with application to robust detection of a changed segment. Statistical Papers, 61(4):1409–1435.
- Schilling, R. L. (2021). Brownian Motion: a guide to random processes and stochastic calculus. De Gruyter, Berlin ; Boston, 3rd edition edition.
- Sharpnack, J. and Arias-Castro, E. (2016). Exact asymptotics for the scan statistic and fast alternatives. Electronic Journal of Statistics, 10(2).
- Spokoiny, V. (1998). Adaptive and spatially adaptive testing of a nonparametric hypothesis. Math. Methods Statist., 7(3):245–273.
- Tsybakov, A. B. (2008). Introduction to Nonparametric Estimation, volume 53. Springer.
- Vershynin, R. (2018). High-dimensional probability: an introduction with applications in data science. Number 47 in Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, Cambridge New York Melbourne New Delhi Singapore.

- Von Bahr, B. and Esseen, C.-G. (1965). Inequalities for the  $r$ th Absolute Moment of a Sum of Random Variables,  $1 \leq r \leq 2$ . The Annals of Mathematical Statistics, 36(1):299–303.
- Walther, G. and Perry, A. (2022). Calibrating the Scan Statistic: Finite Sample Performance Versus Asymptotics. Journal of the Royal Statistical Society Series B: Statistical Methodology, 84(5):1608–1639.
- Wu, W. B. (2005). Nonlinear system theory: Another look at dependence. Proceedings of the National Academy of Sciences, 102(40):14150–14154.