Journal of Stochastic Analysis

Volume 3 | Number 3 Article 5

September 2022

A Sharp Rate of Convergence in the Functional Central Limit Theorem with Gaussian Input

S.V. Lototsky
University of Southern California, Los Angeles, CA 90089, USA, lototsky@usc.edu

Recommended Citation
Lototsky, S.V. (2022) "A Sharp Rate of Convergence in the Functional Central Limit Theorem with Gaussian Input," *Journal of Stochastic Analysis*: Vol. 3: No. 3, Article 5.
DOI: 10.31390/josa.3.3.05
Available at: https://digitalcommons.lsu.edu/josa/vol3/iss3/5

Journal of Stochastic Analysis
Vol. 3, No. 3 (2022) Article 5 (17 pages)
DOI: 10.31390/josa.3.3.05

A SHARP RATE OF CONVERGENCE IN THE FUNCTIONAL CENTRAL LIMIT THEOREM WITH GAUSSIAN INPUT

S. V. LOTOTSKY*

Abstract. When the underlying random variables are Gaussian, the classical Central Limit Theorem (CLT) is trivial, but the functional CLT is not. The objective of the paper is to investigate the functional CLT for stationary Gaussian processes in the Wasserstein-1 metric on the space of continuous functions. Matching upper and lower bounds are established, indicating that the convergence rate is slightly faster than in the Lévy-Prokhorov metric.

1. Introduction

By the Central Limit Theorem, given a collection $\{\xi_k,\ k \geq 1\}$ of independent and identically distributed random variables, each with mean zero and variance one, the sequence $S_n = n^{-1/2}\sum_{k=1}^{n}\xi_k$ converges in distribution, as $n \to \infty$, to the standard Gaussian random variable. The Berry-Esseen bound [16, Theorem 15.51] gives the rate of convergence in the Kolmogorov metric:

$$\sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| \leq \frac{\mathbb{E}|\xi_1|^3}{\sqrt{n}}, \tag{1.1}$$

where F_n is the cumulative distribution function of S_n and Φ is the cumulative distribution function of the standard Gaussian random variable. While the rate $1/\sqrt{n}$ is sharp in general, it can be improved by imposing additional conditions on the random variables ξ_k. For example, if $\mathbb{E}\xi_k^3 = 0$ and $\mathbb{E}\xi_k^4 < \infty$, then the left-hand side of (1.1) is of order $1/n$; cf [21, Theorem 5.2.1]. Of course, if each ξ_k is standard normal, then the left-hand side of (1.1) is zero.

The functional version of the Central Limit Theorem, also known as the Donsker invariance principle [16, Theorem 21.43], establishes weak convergence of the sequence of processes

$$S_n(t) = n^{-1/2}\sum_{k=1}^{\lfloor nt\rfloor}\xi_k + \frac{nt-\lfloor nt\rfloor}{\sqrt{n}}\xi_{\lfloor nt\rfloor+1},\ t \geq 0,\ n \geq 1, \tag{1.2}$$

to the standard Brownian motion W. An analog of (1.1) becomes a bound on the distance between the distributions of S_n and W on the space of continuous

Received 2022-9-18; Accepted 2022-10-10; Communicated by the editors.
2020 *Mathematics Subject Classification.* Primary 60F17; Secondary 60G10, 60G15.
Key words and phrases. Brownian Bridge, Gauss-Markov Process, Kantorovich-Rubinstein metric.
* Corresponding author.

functions $\mathcal{C}(0,T)$ in the Lévy-Prokhorov metric. Compared to (1.1), the corresponding rate of convergence depends on integrability properties of ξ_k in a more complicated way: if $\mathbb{E}|\xi_1|^p < \infty$, $p > 2$, then the rate $n^{-(p-2)/(2(p+1))}$ is sharp; if $\mathbb{E}e^{t\xi_1} < \infty$, $|t| < \delta$, $\delta > 0$, then the rate $\ln n/\sqrt{n}$ is sharp. For details, see [7, Chapter 1]; earlier works on the subject include [17, 18, 25].

A more general approach to investigating the rate of convergence is to find a bound on

$$\sup_{\varphi\in\mathcal{G}} |\mathbb{E}\varphi(S_n) - \mathbb{E}\varphi(W)| \tag{1.3}$$

for a suitable class $\mathcal{G}$ of functions $\varphi : \mathcal{C}(0,T) \to \mathbb{R}$. Barbour [3, Theorem 1] used an infinite-dimensional version of Stein's method to establish the benchmark result

$$|\mathbb{E}\varphi(S_n) - \mathbb{E}\varphi(W)| \le \frac{C}{\sqrt{n}}\Big(\sqrt{\ln n} + \mathbb{E}|\xi_1|^3\Big) \tag{1.4}$$

for a certain (rather restrictive) class $\mathcal{G}$; the restrictive nature of this class ensures that there is no contradiction with [7, Chapter 1] or [25]. For various other $\mathcal{G}$, there are bounds of the form

$$|\mathbb{E}\varphi(S_n) - \mathbb{E}\varphi(W)| \le \frac{C}{n^r},\ 0 < r < \frac{1}{2}; \tag{1.5}$$

cf. [6] and references therein.

Unlike (1.1), the left-hand side of (1.4) will not be zero even if the random variables ξ_k are Gaussian, as long as $\mathcal{G}$ is rich enough to capture the infinite-dimensional nature of the problem. In fact, for certain $\mathcal{G}$, one can use (1.3) to define a metric on the space of distributions. For example, if $\mathcal{G}$ is the collection of bounded Lipschitz continuous functions, then convergence in the corresponding bounded Lipschitz metric is equivalent to weak convergence, that is, convergence in the Lévy-Prokhorov metric; cf. [16, Theorem 13.16] or [9, Theorem 11.3.3]. Removing the boundedness condition (for example, to include linear functionals) leads to the Wasserstein-1 metric, which is the subject of this paper.

Recall that, for two probability measures μ, ν on a complete separable metric space E with distance function ρ and the corresponding Borel sigma-algebra $\mathcal{B}(E)$,

- the `bounded Lipschitz metric` is

$$\mathrm{d}_{BL}(\mu,\nu) = \sup_{\varphi}\left|\int_E \varphi d\mu - \int_E \varphi d\nu\right|, \tag{1.6}$$

  with supremum over functions $\varphi : E \to \mathbb{R}$ such that, for all $x, y \in E$, $|\varphi(x) - \varphi(y)| \le \rho(x,y)$ and $|\varphi(x)| \le 1$;
- the `Wasserstein-1 metric` $\mathrm{d}_w(\mu,\nu)$, also known as the `Kantorovich-Rubinstein metric`, is

$$\mathrm{d}_w(\mu,\nu) = \sup_{\varphi}\left|\int_E \varphi d\mu - \int_E \varphi d\nu\right|, \tag{1.7}$$

  with supremum over functions $\varphi : E \to \mathbb{R}$ such that, for all $x, y \in E$, $|\varphi(x) - \varphi(y)| \le \rho(x,y)$;
- the `Lévy-Prokhorov metric` is

$$\mathrm{d}_{LP}(\mu,\nu) = \inf\{\varepsilon > 0 : \mu(A) \le \nu(A^{\varepsilon}) + \varepsilon,\ A \in \mathcal{B}(E)\}, \tag{1.8}$$

where $A^{\varepsilon} = \{x \in E : \inf_{y \in A} \rho(x,y) \le \varepsilon\}$.

We have $\mathrm{d}_{BL}(\mu,\nu) \le \mathrm{d}_w(\mu,\nu)$ (by definition), $\mathrm{d}_{LP}(\mu,\nu) \le \sqrt{\mathrm{d}_{BL}(\mu,\nu)}$ ([9, Proof of Theorem 11.3.3]), and $\mathrm{d}_{BL}(\mu,\nu) \le 4\mathrm{d}_{LP}(\mu,\nu)$ ([9, Corollary 11.6.5]). In particular, convergence in the Wasserstein-1 metric implies weak convergence, that is, convergence in either Lévy-Prokhorov or bounded Lipschitz metric; the converse is not always true [9, p. 421]; in fact, the diagram in [11] suggests that the Wasserstein-1 metric d_w is the strongest possible for the CLT-type problems in function spaces. Still, a sharp rate of convergence in one metric does not directly lead to a sharp rate in any other metric.

The invariance principle can hold if independence requirement for the random variables ξ_k is relaxed, for example, to a strictly stationary and ergodic martingale difference [20, Theorem 9.1.1], or a stationary Markov process satisfying Doebling's condition [13] [where a bound of the type (1.5) is also established]. In continuous time, if $X = X(t),\ t \in \mathbb{R}$, is a strictly stationary process with mean zero and covariance function $R(t) = \mathbb{E}\big(X(t)X(0)\big)$ satisfying $\int_{-\infty}^{+\infty} R(t)\,dt = 1$, then, under some additional conditions of weak dependence, the sequence of processes

$$S_n(t) = \frac{1}{\sqrt{n}} \int_0^{nt} X(s)\,ds,\ n = 1,2,\ldots, t \in [0,1], \tag{1.9}$$

converges weakly to the standard Brownian motion; cf. [20, Theorem 9.2.1] or [14, Theorem VIII.3.79].

The objective of this paper is to show that if X is a stationary Gauss-Markov process, in either discrete or continuous time, then the Wasserstein-1 distance between S_n and W in the space of continuous functions is of order $\big(n^{-1}\ln n\big)^{1/2}$. In other words, if S_n is Gaussian, then the convergence rate in Wasserstein-1 metric is slightly faster than the Lévy-Prokhorov rate $\ln n/\sqrt{n}$. This difference does not contradict the results from [7, Chapter 1] and [25], and the discrepancy by a $\sqrt{\ln n}$ factor can be explained as follows: in the limit $\sigma \to 0+$, the distance between a Gaussian distribution with mean zero and variance σ^2 and a point mass at zero is of order σ in the Wasserstein-1 metric, but it is of order $\sigma\sqrt{|\ln \sigma|}$ in the Lévy-Prokhorov metric.

Section 2 discusses (the easier) continuous-time case (1.9). Discrete-time case, a generalization of (1.2) for a stationary Gaussian sequence $\{\xi_k,\ k \ge 0\}$, is in Section 3. In Section 4, the results are applied to weak approximation for some ordinary differential equations with additive noise. Section 5 is a summary. Traditionally, models (1.2) and (1.9) are studied on a bounded time interval $[0,T]$, and the index parameter $n = 1,2,\ldots$ is discrete. In this paper, the time interval is $(0,+\infty)$ and the index parameter $\kappa \ge 1$ is not necessarily an integer.

The following two properties of Gaussian processes will be used:

1. **The Borell-TIS inequality** [1, Theorem 2.1.1]: If $X = X(t),\ t \in \mathcal{T}$, is a zero-mean Gaussian process indexed by the set $\mathcal{T}$, and $\mathbb{P}\left(\sup_{t\in\mathcal{T}} X(t) < \infty\right) = 1$,

then $\mathbb{E}\sup_{t\in\mathcal{T}}|X(t)| < \infty$ and, with $X^* = \mathbb{E}\sup_{t\in\mathcal{T}} X(t)$, $\sigma_X^2 = \sup_{t\in\mathcal{T}} \mathbb{E}X^2(t)$,

$$\mathbb{P}\Big(\sup_{t\in\mathcal{T}} X(t) - X^* > x\Big) \leq e^{-x^2/(2\sigma_X^2)},\ x > 0. \tag{1.10}$$

2. **The Fernique-Sudakov inequality** [1, Theorem 2.2.3]: If $X = X(t)$, $Y = Y(t)$ $t \in \mathcal{T}$, are zero-mean Gaussian processes indexed by the set $\mathcal{T}$, and, for all $t, s \in \mathcal{T}$, $\mathbb{E}|X(t) - X(s)|^2 \leq \mathbb{E}|Y(t) - Y(s)|^2$, then

$$\mathbb{E}\sup_{t\in\mathcal{T}} X(t) \leq \mathbb{E}\sup_{t\in\mathcal{T}} Y(t). \tag{1.11}$$

2. Continuous Time

Let $X = X(t)$, $t \in \mathbb{R}$, be a (continuous version of a) stationary Gaussian process with mean zero and covariance $\mathbb{E}X(t)X(s) = e^{-2|t-s|}$. In particular, $X(t)$ is a standard Gaussian random variable for every t. Equivalent characterizations of X are as follows:

$$X(t) = e^{-t}W(e^{2t}); \tag{2.1}$$

$$X(t) = 2\int_{-\infty}^{t} e^{-2(t-s)}dW(s); \tag{2.2}$$

$$dX(t) = -2X(t)dt + 2\,dW(t),\ t \geq 0. \tag{2.3}$$

In (2.1), (2.2), and (2.3), $W = W(t)$, $t \geq 0$, is a standard Brownian motion; in (2.2), when $t < 0$, $W(t) = V(-t)$ for an independent copy V of W. The initial condition $X(0)$ in (2.3) is a standard Gaussian random variable independent of W.

Given a *real number* $\kappa \geq 1$, we define

$$W^\kappa(t) = \frac{1}{\sqrt{\kappa}}\int_0^{\kappa t} X(s)\,ds,\quad t \geq 0. \tag{2.4}$$

Denote by $\mathcal{C}_{(0)}$ the collection of continuous functions $f = f(t)$ on $[0, +\infty)$ such that

$$f(0) = 0,\ \lim_{t\to+\infty}\frac{|f(t)|}{t} = 0.$$

Endowed with the norm

$$\|f\|_{(0)} = \sup_{t>0}\frac{|f(t)|}{1+t},$$

$\mathcal{C}_{(0)}$ becomes a separable Banach space; cf. [8, Section 1.3]. We have $\mathbb{P}(W \in \mathcal{C}_{(0)}) = 1$ (either by the law of large numbers for square integrable martingales or using that $t \mapsto tW(1/t)$ is a standard Brownian motion), and also $\mathbb{P}(W^\kappa \in \mathcal{C}_{(0)}) = 1$, $\kappa \geq 1$, because the ergodic theorem implies

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t X(s)\,ds = \mathbb{E}X(0) = 0$$

with probability one.

Proposition 2.1. *There exists a constant* C_X *such that, for every* $\kappa \geq 1$ *and every function* $\varphi : \mathcal{C}_{(0)} \to \mathbb{R}$ *satisfying*

$$|\varphi(f) - \varphi(g)| \leq \|f - g\|_{(0)}, \tag{2.5}$$

we have

$$\Big|\mathbb{E}\varphi\big(W^\kappa\big) - \mathbb{E}\varphi(W)\Big| \leq C_X \left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2}. \tag{2.6}$$

Proof. Using (2.3),

$$X(t) = X(0)e^{-2t} + 2\int_0^t e^{-2(t-s)}\, dW(s).$$

Changing the order of integration (stochastic Fubini theorem [23, Theorem IV.46]),

$$W^\kappa(t) = \kappa^{-1/2} W(\kappa t) + \frac{X(0) - X(\kappa t)}{2\sqrt{\kappa}}. \tag{2.7}$$

Define

$$V^\kappa(t) = \kappa^{-1/2} W(\kappa t), \quad X^\kappa(t) = \frac{X(\kappa t) - X(0)}{2}. \tag{2.8}$$

Because $t \mapsto V^\kappa(t)$, $t \geq 0$, is standard Brownian motion for every $\kappa > 0$, we have $\mathbb{E}\varphi\big(W\big) = \mathbb{E}\varphi\big(V^\kappa\big)$ and then, using (2.5),

$$\big|\mathbb{E}\varphi\big(W^\kappa\big) - \mathbb{E}\big(\varphi(W)\big)\Big| \leq \frac{\mathbb{E}\|X^\kappa\|_{(0)}}{\sqrt{\kappa}}.$$

Next, by [5, Proposition 2.1],

$$\lim_{T\to+\infty} \frac{\max_{0\leq t\leq T} X(t)}{\sqrt{2\ln T}} = 1, \tag{2.9}$$

with probability one. Using the same arguments as in [19, Proof of Proposition 10.2], we conclude from (2.9) that

$$\limsup_{T\to+\infty} \frac{\max_{0\leq t\leq T} |X(t)|}{\sqrt{2\ln T}} \leq 2, \tag{2.10}$$

and then continuity of X implies that the random variable

$$\zeta = \sup_{t>0} \frac{|X(t) - X(0)|}{2\sqrt{\ln(2+t)}}$$

is finite with probability one. Indeed, if

$$T^* = \sup\left\{t \geq 0 : \frac{\max_{0\leq t\leq T} |X(t) - X(0)|}{2\sqrt{2\ln(2+T)}} > 5\right\}$$

then, by (2.10), $\mathbb{P}(T^* < \infty) = 1$, so that

$$\zeta \leq \max_{0\leq t\leq T^*} \frac{|X(t) - X(0)|}{2\sqrt{\ln(2+t)}} + 5 < \infty.$$

Moreover, because $t \mapsto X(t) - X(0)$ is a Gaussian process with mean zero, the Borell-TIS inequality (1.10) implies

$$\mathbb{E}\zeta^p < \infty \tag{2.11}$$

for all $p > 0$.

If $t > 0$ and $\kappa \geq 1$, then, by direct computation,

$$(2+\kappa t) \leq (1+\kappa)(1+t),\ \frac{1}{2} \leq \sqrt{\ln(1+\kappa)},\ \frac{\sqrt{\ln(1+t)}}{1+t} \leq \frac{1}{2}.$$

As a result,

$$\begin{aligned}
\frac{\sqrt{\ln(2+\kappa t)}}{1+t} &\leq \frac{\sqrt{\ln(1+t)}}{1+t} + \sqrt{\ln(1+\kappa)} \leq 2\sqrt{\ln(1+\kappa)},\\
\|X^\kappa\|_{(0)} &= \sup_{t>0} \frac{|X(\kappa t) - X(0)|}{2(1+t)}\\
&\leq \left(\sup_{t,\kappa>0} \frac{|X(\kappa t)-X(0)|}{2\sqrt{\ln(2+\kappa t)}}\right)\left(\sup_{t>0}\frac{\sqrt{\ln(2+\kappa t)}}{1+t}\right) \leq 2\zeta\sqrt{\ln(1+\kappa)},
\end{aligned}$$

and (2.6) follows with

$$C_X = \mathbb{E}\left[\sup_{t>0}\frac{|X(t)-X(0)|}{\sqrt{\ln(2+t)}}\right]. \tag{2.12}$$

□

Denote by $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_\kappa$ the measures on $\mathcal{C}_{(0)}$ generated by the processes W and W^κ. The following is the main result of this section, showing that the convergence rate $\kappa^{-1/2}\sqrt{\ln\kappa}$ is sharp for the Wasserstein-1 metric.

Theorem 2.2. *There exist positive constants C_X and c_X such that, for every $\kappa \geq 1$,*

$$c_X\left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2} \leq \mathrm{d}_w(\boldsymbol{\mu}_\kappa, \boldsymbol{\mu}_0) \leq C_X\left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2}. \tag{2.13}$$

Proof. The upper bound in (2.13) follows from (1.7) and Proposition 2.1. To establish the lower bound, we use (1.7) with a particular φ.

If $\varphi : \mathcal{C}_{(0)} \to \mathbb{R}$ is a bounded linear functional, then (2.7) and (2.8) imply

$$\mathbb{E}\varphi(W^\kappa) - \mathbb{E}\varphi(W) = \kappa^{-1/2}\mathbb{E}\varphi(X^\kappa). \tag{2.14}$$

For $f \in \mathcal{C}_{(0)}$, define

$$\varphi_\kappa : f \mapsto \frac{f(t^*_\kappa)}{2},$$

where

$$t^*_\kappa = \arg\max_{0\leq t\leq 1} X^\kappa(t).$$

In particular, $t^*_\kappa \in [0,1]$. Then φ_κ is a bounded linear functional on $\mathcal{C}_{(0)}$:

$$|\varphi_\kappa(f)| \leq \frac{\max_{0\leq t\leq 1}|f(t)|}{2} \leq \max_{0\leq t\leq 1}\frac{|f(t)|}{1+t} \leq \|f\|_{(0)}.$$

Therefore, by (1.7) and (2.14),

$$\mathrm{d}_w(\boldsymbol{\mu}_\kappa, \boldsymbol{\mu}_0) \geq \kappa^{-1/2}|\mathbb{E}\varphi_\kappa(X^\kappa)| = \frac{\mathbb{E}\max_{0\leq t\leq 1} X^\kappa(t)}{2\sqrt{\kappa}}. \tag{2.15}$$

Next, define

$$\zeta_\kappa = \frac{\max_{0\le t\le 1} X^\kappa(t)}{\sqrt{\ln(1+\kappa)}} = \frac{\max_{0\le t\le \kappa} X(t) - X(0)}{2\sqrt{\ln(1+\kappa)}}, \quad \kappa \ge 1;$$

the second equality follows from (2.8). By (2.11), the family $\{\zeta_\kappa,\ \kappa \ge 1\}$ is uniformly integrable, so that (2.9) implies

$$\lim_{\kappa\to\infty} \mathbb{E}\zeta_\kappa = \frac{1}{\sqrt{2}}, \tag{2.16}$$

which, in turn, means

$$\inf_{\kappa\ge 1} \mathbb{E}\zeta_\kappa > 0. \tag{2.17}$$

The lower bound in (2.13), with

$$c_X = \frac{1}{2}\inf_{\kappa\ge 1} \frac{\mathbb{E}\left[\max_{0\le t\le\kappa} X(t)\right]}{\sqrt{\ln(1+\kappa)}}, \tag{2.18}$$

now follows from (2.15) and (2.17), because $\mathbb{E}X(0) = 0$. □

To get a better idea about numerical values of C_X and c_X, we need

Proposition 2.3. *Let $X = X(t),\ t \in \mathbb{R}$, be a stationary Gaussian process with mean zero and covariance $e^{-2|t-s|}$. Define the random variable*

$$\bar{\eta} = \max_{0\le t\le 1} X(t). \tag{2.19}$$

Then

$$1.2 < \mathbb{E}\bar{\eta} < 3.2; \tag{2.20}$$

$$\mathbb{P}(\bar{\eta} > x) \le e^{-(x-3.2)^2/2},\ x \ge 3.2. \tag{2.21}$$

Proof. Let $B = B(t)$, $0 \le t \le 1$, be the standard Brownian bridge and $W = W(t)$, $0 \le t \le 1$, a standard Brownian motion. Then

$$\mathbb{E}|B(t) - B(s)|^2 = |t-s| - |t-s|^2,\ \mathbb{E}|X(t) - X(s)|^2 = 2(1 - e^{-2|t-s|}),$$
$$\mathbb{E}|W(t) - W(s)|^2 = |t-s|,$$

and, because

$$x - x^2 \le 1 - e^{-2x} \le 2x,\ x \ge 0,$$

inequality (1.11) implies

$$2\,\mathbb{E}\max_{0\le t\le 1} B(t) \le \mathbb{E}\bar{\eta} \le 4\,\mathbb{E}\max_{0\le t\le 1} W(t).$$

It is well known (e.g. [9, Section 12.3]) that

$$\mathbb{P}\left(\max_{0\le t\le 1} B(t) > x\right) = e^{-2x^2},\ \ \mathbb{P}\left(\max_{0\le t\le 1} W(t) > x\right) = \frac{2}{\sqrt{2\pi}}\int_x^{+\infty} e^{-t^2/2}dt. \tag{2.22}$$

Then

$$\mathbb{E}\max_{0\le t\le 1} B(t) = \int_0^{+\infty} e^{-2x^2}\,dx = \frac{\sqrt{2\pi}}{4} > 0.6,$$
$$\mathbb{E}\max_{0\le t\le 1} W(t) = \frac{2}{\sqrt{2\pi}}\int_0^{+\infty} xe^{-x^2/2}\,dx = \sqrt{\frac{2}{\pi}} < 0.8,$$

and (2.20) follows. After that, (2.21) is a re-statement of the Borell-TIS inequality (1.10). □

We can now show that the number C_X defined in (2.12) satisfies

$$1.4 < C_X < 14. \tag{2.23}$$

For the lower bound, note that

$$C_X \ge \mathbb{E}\left[\sup_{t>0}\frac{X(t)-X(0)}{\sqrt{\ln(2+t)}}\right],$$

whereas

$$\sup_{t>0}\frac{X(t)-X(0)}{\sqrt{\ln(2+t)}} \ge \limsup_{T\to\infty}\frac{\max_{0\le t\le T}\big(X(t)-X(0)\big)}{\sqrt{\ln(2+T)}},$$

and it remains to apply (2.16).

For the upper bound in (2.23), start by writing

$$C_X \le \mathbb{E}\left[\sup_{t>0}\frac{|X(t)|}{\sqrt{\ln(2+t)}}\right] + \frac{\mathbb{E}|X(0)|}{\sqrt{\ln 2}};\quad \frac{\mathbb{E}|X(0)|}{\sqrt{\ln 2}} = \sqrt{\frac{2}{\pi\ln 2}} \approx 0.96. \tag{2.24}$$

Next, let $\bar X$ be the process consisting of iid copies of $X(t),\ t\in[0,1)$, on each of the intervals $[k-1,k)$, $k=1,2,\ldots$. Then $\mathbb{E}X^2(t)=\mathbb{E}\bar X^2(t)$, $\mathbb{E}X(t)X(s)\ge \mathbb{E}\bar X(t)\bar X(s)$, and so

$$\mathbb{E}\left[\sup_{t>0}\frac{|X(t)|}{\sqrt{\ln(2+t)}}\right] \le 2\mathbb{E}\left[\sup_{t>0}\frac{X(t)}{\sqrt{\ln(2+t)}}\right] \le 2\mathbb{E}\left[\sup_{t>0}\frac{\bar X(t)}{\sqrt{\ln(2+t)}}\right], \tag{2.25}$$

where the first inequality follows from [19, Proposition 10.2], and the second, from (1.11). On the other hand, if $\bar\eta_k,\ k\ge 1$, are iid copies of the random variable $\bar\eta$ from (2.19), then

$$\mathbb{E}\left[\sup_{t>0}\frac{\bar X(t)}{\sqrt{\ln(2+t)}}\right] \le \mathbb{E}\left[\sup_{k\ge 1}\frac{\bar\eta_k}{\sqrt{\ln(1+k)}}\right]. \tag{2.26}$$

Using (2.21) with $x>5$,

$$\mathbb{P}\Big(\sup_{k\ge1}\frac{\bar\eta_k}{\sqrt{\ln(1+k)}} > x\Big) \le \sum_{k\ge1}\mathbb{P}\Big(\bar\eta_k > x\sqrt{\ln(1+k)}\Big)$$
$$\le \sum_{k\ge1}\frac{1}{(1+k)^{-(x-3.2)^2/2}} \le \frac{2}{(x-3.2)^2-2},$$

and therefore

$$\mathbb{E}\left[\sup_{k\geq 1}\frac{\bar{\eta}_k}{\sqrt{\ln(1+k)}}\right]\leq 5+2\int_5^{+\infty}\frac{dx}{(x-3.2)^2-2}<6.5.$$

Then (2.23) follows from (2.24) – (2.26).

Next, we will show that the number c_X defined in (2.18) satisfies

$$0.2<c_X<0.8. \tag{2.27}$$

Indeed, the upper bound follows immediately from (2.16). For the lower bound, start by noting that, for $N\leq\kappa<N+1$,

$$\mathbb{E}\left[\max_{0\leq t\leq\kappa}X(t)\right]\geq\mathbb{E}\left[\max_{k=1,\dots,N}X(k)\right]$$

and $\mathbb{E}|X(k)-X(m)|^2=2(1-e^{-2})>1$. Now take iid Gaussian Y_k, $k=1,\dots,N$ with mean zero and variance $1/2$. Then $\mathbb{E}|Y(k)-Y(m)|^2=1$ and so

$$\mathbb{E}|X(k)-X(m)|^2\geq\mathbb{E}|Y(k)-Y(m)|^2.$$

By (1.11),

$$\mathbb{E}\left[\max_{k=1,\dots,N}X(k)\right]\geq\mathbb{E}\left[\max_{k=1,\dots,N}Y(k)\right],$$

and, by [19, Lemma 10.2],

$$\mathbb{E}\left[\max_{k=1,\dots,N}Y(k)\right]\geq 0.4\sqrt{\ln N},$$

leading to the lower bound in (2.27).

To summarize, we can write (2.13) in a more explicit form

$$0.2\left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2}\leq \mathrm{d}_w(\boldsymbol{\mu}_\kappa,\boldsymbol{\mu}_0)\leq 14\left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2},\ \kappa\geq 1.$$

Theorem 2.2 can be used to study convergence on a bounded interval. For $T>0$, let $\mathcal{C}(0,T)$ be the space of continuous functions on $[0,T]$ with the sup norm, and denote by $\boldsymbol{\mu}_0^T$ and $\boldsymbol{\mu}_\kappa^T$ the measures on $\mathcal{C}(0,T)$ generated by the processes W and W^κ.

Theorem 2.4. *There exist positive constants $C_{X,T}$ and $c_{X,T}$ such that, for every $\kappa\geq 1$,*

$$c_{X,T}\left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2}\leq \mathrm{d}_w(\boldsymbol{\mu}_\kappa^T,\boldsymbol{\mu}_0^T)\leq C_{X,T}\left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2}. \tag{2.28}$$

Proof. If $f\in\mathcal{C}_{(0)}$, then, for $0<t<T$, $|f(t)|\leq(1+T)|f(t)|/(1+t)$, so that

$$\|f\|_{\mathcal{C}(0,T)}=\max_{0\leq t\leq T}|f(t)|\leq(1+T)\|f\|_{(0)}$$

and the upper bound in (2.28) follows from the upper bound in (2.13), with $C_{X,T}=(1+T)C_X$.

The lower bound, with

$$c_{X,T}=\frac{1}{2}\inf_{\kappa\geq 1}\frac{\mathbb{E}\left[\max_{0\leq t\leq T}X^\kappa(t)\right]}{\sqrt{\ln(1+\kappa)}},$$

follows after repeating the corresponding steps in the proof of Theorem 2.2. □

3. Discrete Time

Consider a stationary Gaussian sequence $X = \{X_n,\ n \geq 0\}$, with $\mathbb{E}X_n = 0$ and $\mathbb{E}X_{k+n}X_k = (1-a)a^n/(1+a),\ n,k \geq 0$, where $a \in (-1,1)$ and $a = 0$ corresponds to the sequence of iid standard Gaussian random variables. The variance of X_n is chosen so that, for all $a \in (-1,1)$, the covariance function $R(n) = \mathbb{E}X_{k+n}X_n$ of X satisfies $R(0) + 2\sum_{n=1}^{\infty} R(n) = 1$.

Using a collection $\xi_k,\ k = 0, \pm 1, \pm 2, \dots$ of iid standard normal random variables, we get the discrete-time analogs of (2.2) and (2.3):

$$X_n = (1-a)\sum_{k=-\infty}^{n} a^{n-k}\xi_k, \tag{3.1}$$

$$X_{n+1} = aX_n + \xi_{n+1}; \tag{3.2}$$

in (3.2), the initial condition X_0 is independent of $\xi_k,\ k \geq 1$, and is a normal random variable with mean 0 and variance $(1-a)/(1+a)$.

For $x > 0$, let $\lfloor x \rfloor$ denote the largest integer that is less than or equal to x. Define the processes W^κ by

$$W^\kappa(t) = \frac{1}{\sqrt{\kappa}}\sum_{n=1}^{\lfloor \kappa t \rfloor} X_n + \frac{\kappa t - \lfloor \kappa t \rfloor}{\sqrt{\kappa}} X_{\lfloor \kappa t \rfloor + 1},\ t \geq 0,\ \kappa \geq 1. \tag{3.3}$$

The second term on the right-hand side of (3.3) ensures that W^κ is a continuous function of t.

The case $a = 0$, that is, the Gaussian version of the original Donsker theorem, is of special interest; the corresponding process W^κ will be denoted by S_κ:

$$S_\kappa(t) = \frac{1}{\sqrt{\kappa}}\sum_{n=1}^{\lfloor \kappa t \rfloor} \xi_n + \frac{\kappa t - \lfloor \kappa t \rfloor}{\sqrt{\kappa}} \xi_{\lfloor \kappa t \rfloor + 1},\ t \geq 0,\ \kappa \geq 1. \tag{3.4}$$

We have $\mathbb{P}(W^\kappa \in \mathcal{C}_{(0)}) = 1$ for every $\kappa \geq 1$, and $a \in (-1,1)$, because the ergodic theorem implies

$$\lim_{t\to\infty} \frac{1}{\lfloor \kappa t \rfloor}\sum_{n=1}^{\lfloor \kappa t \rfloor} X_n = \mathbb{E}X_0 = 0$$

with probability one.

Let $W = W(t),\ t \geq 0$, be a standard Brownian motion, and denote by $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_\kappa$ the measures on $\mathcal{C}_{(0)}$ generated by the processes W and W^κ. The following is the discrete-time analog of Theorem 2.2.

Theorem 3.1. *There exist positive constants C_a and c_a such that, for every $\kappa \geq 1$,*

$$c_a\left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2} \leq \mathrm{d}_w(\boldsymbol{\mu}_\kappa, \boldsymbol{\mu}_0) \leq C_a\left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2}. \tag{3.5}$$

Proof. The steps are the same as in the proof of Theorem 2.2.

Substituting (3.1) in (3.3) and changing the order of summation,

$$W^{\kappa}(t) = S_{\kappa}(t) + \frac{X^{\kappa}(t)}{\sqrt{\kappa}}, \tag{3.6}$$

where S_κ is from (3.4) and

$$X^{\kappa}(t) = \left((a - a^{\lfloor \kappa t \rfloor}) \sum_{n=-\infty}^{0} a^{-n}\xi_n - \sum_{n=1}^{\lfloor \kappa t \rfloor} a^{\lfloor \kappa t \rfloor - n}\xi_n \right).$$

As a result, it is enough to establish (3.5) when $a = 0$:

$$c_0 \left(\frac{\ln(1+\kappa)}{\kappa} \right)^{1/2} \leq \mathrm{d}_w(\boldsymbol{\mu}_{\kappa}, \boldsymbol{\mu}_0) \leq C_0 \left(\frac{\ln(1+\kappa)}{\kappa} \right)^{1/2}. \tag{3.7}$$

Then, similar to the continuous time case, we see that $\mathbb{E}\|X^{\kappa}\|_{(0)} \leq \bar{C}_a\sqrt{\ln(1+\kappa)}$, with a suitable constant $\bar{C}_a$, and then (3.5) follows from (3.7) with $C_a = C_0 + \bar{C}_a$ and $c_a = c_0$.

To prove (3.7), we choose the random variables ξ_k in (3.4) as the increments of the Brownian motion W:

$$\frac{\xi_n}{\sqrt{\kappa}} = W(n/\kappa) - W((n-1)/\kappa). \tag{3.8}$$

Then, for $(n-1)/\kappa \leq t \leq n/\kappa$, the process $t \mapsto S_{\kappa} - W$ is a Brownian bridge, and, for every function $\varphi : \mathcal{C}_{(0)} \to \mathbb{R}$ satisfying (2.5),

$$|\mathbb{E}\varphi(S_{\kappa}) - \mathbb{E}\varphi(W)| \leq \mathbb{E}\|B^{\kappa}\|_{(0)}, \tag{3.9}$$

where B^{κ} is a collection of independent Brownian bridges on $[(n-1)/\kappa, n/\kappa]$, $n = 1, 2, \ldots$.

Direct computations show that, for $N = 1, 2, \ldots$,

$$\sqrt{\kappa}\, \mathbb{E} \max_{0 \leq t \leq N/\kappa} |B^{\kappa}(t)| \leq 8\sqrt{\ln(N+1)}, \tag{3.10}$$

$$\sqrt{\kappa}\, \mathbb{E} \max_{0 \leq t \leq N/\kappa} B^{\kappa}(t) \geq 0.3\sqrt{\ln(N+1)}; \tag{3.11}$$

the numbers 0.3 and 8 do not necessarily provide optimal bounds. Indeed, let $B = B(t),\ t \in [0,1]$, be the standard Brownian bridge, and let

$$\eta = \max_{0 \leq t \leq 1} B(t).$$

Then

$$\sqrt{\kappa}\, \mathbb{E} \max_{0 \leq t \leq N/\kappa} |B^{\kappa}(t)| = \mathbb{E} \max_{k=1,\ldots,N} |\eta_k|,$$

where $\eta_k,\ k = 1, \ldots, N$, are iid copies of η. Also,

$$\mathbb{E} \max_{k=1,\ldots,N} \eta_k \leq \mathbb{E} \max_{k=1,\ldots,N} |\eta_k| \leq 2\mathbb{E} \max_{k=1,\ldots,N} \eta_k.$$

To derive (3.10), we repeat the arguments from the proof of Lemma 10.1 in [19] using (2.22) and conclude that $\mathbb{E} \max_{k=1,\ldots,N} \eta_k \leq 4\sqrt{\ln(N+1)}$. Similarly, for (3.11), we repeat the proof of Lemma 10.2 in [19].

Next, denote by $\bar{B}$ the process B^κ corresponding to $\kappa = 1$, that is, the collection of independent standard Brownian bridges on $[n-1, n]$, $n \geq 1$. Then we get the upper bound in (3.7), with

$$C_0 = \mathbb{E}\left[\sup_{t>0} \frac{|\bar{B}(t)|}{\sqrt{\ln(2+t)}}\right] < 14,$$

by combining (3.9) and (3.10); the upper bound on C_0 is from (2.23), because, by (1.11), $C_0 \leq C_X$. The lower bound in (3.7), with

$$c_0 = \frac{1}{2} \inf_{\kappa \geq 1} \frac{\mathbb{E}\left[\max_{0 \leq t \leq \kappa} \bar{B}(t)\right]}{\sqrt{\ln(1+\kappa)}} > 0.2,$$

follows from (3.11) after the same arguments as in the proof of Theorem 2.2; the lower bound on c_0 follows from (3.11). ☐

For $T > 0$, let $\mathcal{C}(0,T)$ be the space of continuous functions on $[0,T]$ with the sup norm, and denote by $\boldsymbol{\mu}_0^T$ and $\boldsymbol{\mu}_\kappa^T$ the measures on $\mathcal{C}(0,T)$ generated by the processes W and W^κ. The discrete-time version of Theorem 2.4 is obvious. When $a = 0$, and there is no continuous-time analog, we also have the following result (cf. [2, Proposition 2.1]).

Proposition 3.2. *If $\boldsymbol{\mu}_\kappa^T$ is the measure on $\mathcal{C}(0,T)$ generated by the process S_κ from* (3.4)*, then*

$$\lim_{\kappa\to\infty} \sqrt{\frac{\kappa}{\ln \kappa}}\, \mathrm{d}_w(\boldsymbol{\mu}_\kappa^T, \boldsymbol{\mu}_0^T) = \sqrt{2}. \tag{3.12}$$

Proof. Using the random variables η_k from the proof of Theorem 3.1,

$$\limsup_{\kappa\to\infty} \sqrt{\frac{\kappa}{\ln \kappa}}\, \mathrm{d}_w(\boldsymbol{\mu}_\kappa^T, \boldsymbol{\mu}_0^T) \leq \lim_{N\to\infty} \frac{\mathbb{E}\left[\max\limits_{k=1,\ldots,N} |\eta_k|\right]}{\sqrt{\ln N}},$$

and

$$\liminf_{\kappa\to\infty} \sqrt{\frac{\kappa}{\ln \kappa}}\, \mathrm{d}_w(\boldsymbol{\mu}_\kappa^T, \boldsymbol{\mu}_0^T) \geq \lim_{N\to\infty} \frac{\mathbb{E}\left[\max\limits_{k=1,\ldots,N} \eta_k\right]}{\sqrt{\ln N}}.$$

By (2.22) and [24, Theorem 1],

$$\lim_{N\to\infty} \frac{\max\limits_{k=1,\ldots,N} \eta_k}{\sqrt{\ln N}} = \lim_{N\to\infty} \frac{\max\limits_{k=1,\ldots,N} |\eta_k|}{\sqrt{\ln N}} = \sqrt{2}$$

with probability 1; then uniform integrability [22, Theorem 2.1] implies (3.12). ☐

4. Applications

Let W be a standard Brownian motion and let W^κ be the process from (2.4) or (3.3). Consider a continuous mapping $\Psi : \mathcal{C}_{(0)} \to \mathcal{C}_{(0)}$. Denote by $\boldsymbol{\mu}_{0,\psi}$ and $\boldsymbol{\mu}_{\kappa,\psi}$ the measures on $\mathcal{C}_{(0)}$ generated by the processes $\Psi(W)$ and $\Psi(W^\kappa)$.

Proposition 4.1. *We have weak convergence* $\lim_{\kappa\to\infty} \boldsymbol{\mu}_{\kappa,\psi} = \boldsymbol{\mu}_{0,\psi}$. *Moreover, if there exists a number* C_ψ *such that, for all* $f, g \in \mathcal{C}_{(0)}$,

$$\|\Psi(f) - \Psi(g)\|_{(0)} \leq C_\psi \|f - g\|_{(0)}, \tag{4.1}$$

then

$$\mathrm{d}_w(\boldsymbol{\mu}_{\kappa,\psi}, \boldsymbol{\mu}_{0,\psi}) \leq \bar{C}_X C_\psi \left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2}, \tag{4.2}$$

with $\bar{C}_X = C_X$ *from* (2.12) *in continuous time and* $\bar{C}_X = C_a$ *from* (3.5) *in discrete time.*

Proof. Weak convergence follows by the continuous mapping theorem (e.g. [4, Theorem 2.7]). To establish (4.2), we use either (2.6) or the upper bound in (3.5) and note that if $\varphi : \mathcal{C}_{(0)} \to \mathbb{R}$ satisfies (2.5), then $|\varphi(\Psi(f)) - \varphi(\Psi(g))| \leq C_\psi \|f - g\|_{(0)}$. □

Example 1. Given $\alpha > 0$, let Y^κ, Y be the solutions of

$$Y^\kappa(t) = -\alpha \int_0^t Y^\kappa(s)\, ds + W^\kappa(t), \quad Y(t) = -\alpha \int_0^t Y(s)\, ds + W(t), \quad t \geq 0.$$

Then $Y^\kappa, Y \in \mathcal{C}_{(0)}$, and the corresponding measures $\boldsymbol{\nu}_\kappa, \boldsymbol{\nu}$ on $\mathcal{C}_{(0)}$ satisfy

$$\mathrm{d}_w(\boldsymbol{\nu}_\kappa, \boldsymbol{\nu}) \leq 2\bar{C}_X \left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2}, \ \kappa \geq 1;$$

the constant $\bar{C}_X$ is from Proposition 4.1.

Indeed, by direct computation,

$$Y^\kappa(t) = \Psi_\alpha(W^\kappa)(t), \quad Y(t) = \Psi_\alpha(W)(t),$$

where

$$\Psi_\alpha : f(t) \mapsto f(t) - \alpha \int_0^t e^{-\alpha(t-s)} f(s)\, ds, \quad f \in \mathcal{C}_{(0)}, \tag{4.3}$$

is a linear operator. To see that Ψ_α maps $\mathcal{C}_{(0)}$ to itself, note that, for every $t > T > 0$,

$$\begin{aligned}
\frac{|\Psi_\alpha(f)(t)|}{1+t} &\leq \frac{|f(t)|}{1+t} + \frac{\alpha}{1+t} \int_0^t e^{-\alpha(t-s)} |f(s)|\, ds \\
&\leq \frac{|f(t)|}{1+t} + \frac{\alpha}{1+t} \int_0^T e^{-\alpha(t-s)} |f(s)|\, ds \\
&+ \frac{\alpha e^{-\alpha t}}{1+t} \int_T^t (1+s) e^{\alpha s} \frac{|f(s)|}{1+s}\, ds.
\end{aligned}$$

If $\lim_{t\to\infty} |f(t)|/(1+t) = 0$, then, for every $\varepsilon > 0$, we can find T so that $|f(s)|/(1+s) < \varepsilon$, $s > T$. As a result, keeping in mind that

$$\frac{\alpha}{1+t} \int_T^t (1+s) e^{\alpha s}\, ds \leq \alpha \int_0^t e^{\alpha s}\, ds \leq e^{\alpha t},$$

we compute

$$\limsup_{t\to\infty} \frac{|\Psi_\alpha(f)(t)|}{1+t} \leq \varepsilon$$

and conclude that $\lim_{t\to\infty} |\Psi_\alpha(f)(t)|/(1+t) = 0$. Similarly,

$$\|\Psi_\alpha(f)\|_{(0)} \leq \|f\|_{(0)} \left(1 + \alpha \int_0^{+\infty} e^{-\alpha t}\, dt\right) = 2\|f\|_{(0)}, \tag{4.4}$$

so that (4.1) holds with $C_\psi = 2$.

The analog of Proposition 4.1 on a bounded interval is as follows. Let Ψ^T be a continuous mapping of $\mathcal{C}(0,T)$ to itself. Denote by $\boldsymbol{\mu}^T_{0,\psi}$ and $\boldsymbol{\mu}^T_{\kappa,\psi}$ the measures on $\mathcal{C}(0,T)$ generated by the processes $\Psi^T(W)$ and $\Psi^T(W^\kappa)$.

Proposition 4.2. *We have weak convergence* $\lim_{\kappa\to\infty} \boldsymbol{\mu}^T_{\kappa,\psi} = \boldsymbol{\mu}^T_{0,\psi}$. *Moreover, if there exists a number* C^T_ψ *such that, for all* $f, g \in \mathcal{C}(0,T)$,

$$\|\Psi^T(f) - \Psi^T(g)\|_{\mathcal{C}(0,T)} \leq C^T_\psi \|f - g\|_{\mathcal{C}(0,T)}, \tag{4.5}$$

then

$$\mathrm{d}_w(\boldsymbol{\mu}^T_{\kappa,\psi}, \boldsymbol{\mu}^T_{0,\psi}) \leq (1+T)\bar{C}_X C^T_\psi \left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2},$$

with $\bar{C}_X = C_X$ *from* (2.12) *in continuous time and* $\bar{C}_X = C_a$ *from* (3.5) *in discrete time.*

Example 2. Let the function $b = b(x),\ x \in \mathbb{R}$, satisfy

$$|b(x) - b(y)| \leq K|x-y|,\ x, y \in \mathbb{R}, \tag{4.6}$$

and let Y^κ, Y be the solutions of

$$Y^\kappa(t) = \int_0^t b\big(Y^\kappa(s)\big)\, ds + W^\kappa(t),\quad Y(t) = \int_0^t b\big(Y(s)\big)\, ds + W(t),\quad 0 \leq t \leq T.$$

If $\boldsymbol{\nu}^T_\kappa, \boldsymbol{\nu}^T$ are the corresponding measures on $\mathcal{C}(0,T)$, then

$$\mathrm{d}_w(\boldsymbol{\nu}^T_\kappa, \boldsymbol{\nu}^T) \leq (1+T)\bar{C}_X e^{KT} \left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2}.$$

Indeed, for $f \in \mathcal{C}(0,T)$, define $\Psi^T(f)(t) = y(t)$ as the solution of

$$y(t) = \int_0^t b\big(y(s)\big)\, ds + f(t),\ 0 \leq t \leq T.$$

By direct computation ([10, Chapter 4, Lemma 1.1]), we have (4.5) with $C^T_\psi = e^{KT}$.

Example 3. Let us combine Examples 1 and 2. Take a positive number α and a function $b = b(x)$ satisfying (4.6), and let Y^κ, Y be the solutions of

$$\begin{aligned} Y^\kappa(t) &= -\alpha \int_0^t Y^\kappa(s)\, ds + \int_0^t b\big(Y^\kappa(s)\big)\, ds + W^\kappa(t), \\ Y(t) &= -\alpha \int_0^t Y(s)\, ds + \int_0^t b\big(Y(s)\big)\, ds + W(t),\quad t \geq 0. \end{aligned}$$

If $\alpha > K$, then $Y^\kappa, Y \in \mathcal{C}_{(0)}$ and, for the corresponding measures $\boldsymbol{\nu}_\kappa, \boldsymbol{\nu}$,

$$\mathrm{d}_w(\boldsymbol{\nu}_\kappa, \boldsymbol{\nu}) \leq \frac{2\alpha \bar{C}_X}{\alpha - K} \left(\frac{\ln(1+\kappa)}{\kappa}\right)^{1/2}. \tag{4.7}$$

Indeed, for $f \in \mathcal{C}_{(0)}$, define $\Psi(f)(t) = y(t)$ as the solution of

$$y(t) = -\alpha \int_0^t y(s)\,ds + \int_0^t b\big(y(s)\big)\,ds + f(t),\ t \geq 0.$$

Using variation of parameters formula and (4.3),

$$\Psi(f)(t) = \int_0^t e^{-\alpha(t-s)} b\big(\Psi(f)(s)\big)\,ds + \Psi_a(f)(t).$$

Then, similar to Example 1, we conclude that Ψ maps $\mathcal{C}_{(0)}$ to itself. In particular, using (4.6) and (4.4),

$$\Psi(f)(t) - \Psi(g)(t) = \int_0^t e^{-\alpha(t-s)} \Big(b\big(\Psi(f)(s)\big) - b\big(\Psi(g)(s)\big)\Big)\,ds + \Psi_\alpha(f-g)(t)$$

so that

$$\|\Psi(f) - \Psi(g)\|_{(0)} \leq K\|\Psi(f) - \Psi(g)\|_{(0)} \int_0^\infty e^{-as}\,ds + 2\|f-g\|_{(0)}.$$

As a result, if $\alpha > K$, then $\|\Psi(f) - \Psi(g)\|_{(0)} \leq 2\alpha\|f-g\|_{(0)}/(\alpha - K)$, and (4.7) follows from (4.2).

5. Concluding Remarks

A proof of the functional Central Limit Theorem for processes of the type (2.4) or (3.3) usually includes the following steps:

(1) A Gordin-type decomposition [12], when W^κ is written as a sum of a martingale and an a "small" correction;
(2) A coupling argument, when W^κ is constructed on the same probability space as W;
(3) A Skorokhod embedding for the martingale component of W^κ.

Each step leads to an approximation error; in particular, [17, 18] developed a systematic procedure, now known as the KMT approximation, to minimize the error due to the Skorokhod embedding. When the underlying processes are Gaussian, some of the approximation errors are not present.

In continuous time, the first two steps are the equality (2.7). There is no need for Skorokhod embedding because the martingale component is the Brownian motion. In discrete time, the first step is the equality (3.6), whereas (3.8) represents coupling and the Skorokhod embedding. For convergence in the space of continuous functions, the $\sqrt{\ln \kappa}$ correction to the classical rate $1/\sqrt{\kappa}$ comes from the growth of the maximum of iid standard Gaussian random variables.

Keeping in mind that rate of convergence in the functional CLT can depend both on the underlying functional space and on the distance between the measures on that space, the rate $1/\sqrt{\kappa}$ is possible to achieve. For example, by considering W and S_κ [from (3.4)] as processes in $L_1(0,T)$, as opposed to $\mathcal{C}(0,T)$, direct computations [2, Proposition 2.1] yield

$$\mathbb{E}\int_0^1 |S_\kappa(t) - W(t)|\,dt = \frac{1}{\sqrt{\kappa}} \int_0^1 \mathbb{E}|B(t)|\,dt = \sqrt{\frac{2}{\pi\,\kappa}} \int_0^1 \sqrt{t(1-t)}\,dt = \sqrt{\frac{\pi}{32\,\kappa}},$$

that is, the Wasserstein-1 distance between S_κ and W in $L_1(0,T)$ is of order $1/\sqrt{\kappa}$; see also [3, Remark 1].

Given the variety of spaces that can support W and W^κ and the different ways to measure the distance between the corresponding probability distributions [11], identifying all situations with a sharp $\kappa^{-1/2}$ bound becomes an interesting challenge. In the space of continuous functions with the sup norm, there is strong evidence that the Wasserstein distance between W and W^κ cannot converge to zero faster than $\sqrt{\ln\kappa/\kappa}$: the results of this paper demonstrate it in the Gaussian case, and, by [15, Corollary 4.4], the simple symmetric random walk cannot beat this rate either.

Acknowledgment. The work was supported by AFOSR grant FA9550-21-1-0015.

References

1. Adler, R. J., and Taylor, J. E.: *Random fields and geometry*, Springer, New York, 2007.
2. Asmussen, S., and Glynn, P. W.: *Stochastic Simulation: Algorithms and Analysis*, Springer, New York, 2007.
3. Barbour, A. D.: Stein's method for diffusion approximations, *Probab. Theory Related Fields* **84** (1990), no. 3, 297–322.
4. Billingsley, P.: *Convergence of probability measures*, second ed. John Wiley & Sons, Inc., New York, 1999.
5. Chen, X.: The limit law of the iterated logarithm, *J. Theor. Probab.* **28** (2015), no. 2, 721–725.
6. Coutin, L., and Decreusefond, L.: Donsker's theorem in Wasserstein-1 distance, *Electron. Commun. Probab* **25** (2020), no. 27, 1–13.
7. Csörgő, M., and Horváth, L.: *Weighted approximations in probability and statistics*, John Wiley & Sons, Chichester, 1993.
8. Deuschel, J.-D., and Stroock, D. W.: *Large deviations,* Academic Press, Boston, 1989.
9. Dudley, R. M.: *Real Analysis and Probability*, Cambridge University Press, Cambridge, 2004.
10. Freidlin, M. I., and Wentzell, A. D.: *Random perturbations of dynamical systems*, second ed. Springer, New York, 1998.
11. Gibbs, A. L., and Su, F. E.: On choosing and bounding probability metrics, *International Statistical Review* **70** (2002), no. 3, 419–435.
12. Gordin, M. I.: The central limit theorem for stationary processes, *Dokl. Akad. Nauk SSSR* **188** (1969), 739–741.
13. Haeusler, E.: On the rate of convergence in the invariance principle for real-valued functions of Doeblin processes, *Journal Of Multivariate Analysis* **15** (1984), 73–90.
14. Jacod, J., and Shiryaev, A. N.: *Limit Theorems for Stochastic Processes*, second ed. Springer, 2003.
15. Kjos-Hanssen, B., and Szabados, T.: Kolmogorov complexity and strong approximation of Brownian motion, *Proc. Amer. Math. Soc.* **139** (2011), no. 9, 3307–3316.
16. Klenke, A.: *Probability theory—a comprehensive course*, third ed. Springer, 2020.
17. Komlós, J., Major, P., and Tusnády, G.: An approximation of partial sums of independent RV's and the sample DF, I, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32** (1975), 111–131.
18. Komlós, J., Major, P., and Tusnády, G.: An approximation of partial sums of independent RV's, and the sample DF, II, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **34** (1976), no. 1, 33–58.
19. Lifshits, M.: *Lectures on Gaussian processes*, Springer, Heidelberg, 2012.
20. Liptser, R. S., and Shiryayev, A. N.: *Theory of Martingales*, Kluwer Academic Publishers, Boston, 1989.
21. Petrov, V. V.: *Limit Theorems of Probability Theory*, Oxford University Press, Oxford, 1995.

22. Pickands, J.: Moment convergence of sample extremes, *Ann. Math. Statist.* **39** (1968), 881–889.
23. Protter, P. E.: *Stochastic integration and differential equations*, Springer, Berlin, 2005.
24. Resnick, S. I., and Tomkins, R. J.: Almost sure stability of maxima, *J. Appl. Prob.* **10** (1973), no. 2, 387–401.
25. Sakhanenko, A. I.: Estimates of the rate of convergence in the invariance principle, in: *Limit theorems of probability theory and related questions* (1982) 72–78, “Nauka” Sibirsk. Otdel., Novosibirsk,

S. V. Lototsky: Department of Mathematics, USC, Los Angeles, CA 90089, USA
Email address: lototsky@usc.edu
URL: https://dornsife.usc.edu/sergey-lototsky/