On the moments of the modulus of continuity of Itô processes[*]

Markus Fischer[†]   Giovanna Nappo[‡]

November 3, 2008 / June 17, 2009

Abstract

The modulus of continuity of a stochastic process is a random element for any fixed mesh size. We provide upper bounds for the moments of the modulus of continuity of Itô processes with possibly unbounded coefficients, starting from the special case of Brownian motion. References to known results for the case of Brownian motion and Itô processes with uniformly bounded coefficients are included. As an application, we obtain the rate of strong convergence of Euler-Maruyama schemes for the approximation of stochastic delay differential equations satisfying a Lipschitz condition in supremum norm.

2000 AMS subject classifications: primary 60J60, 60J65, 60G17, 60G70, 60H99; secondary 34K50, 60H10, 60H35.

Key words and phrases: modulus of continuity; Itô process; extreme values; stochastic differential equation; functional differential equation; delay; Euler-Maruyama scheme.

1 Introduction

A typical trajectory of standard Brownian motion is Hölder continuous of any order less than one half. If such a trajectory is evaluated at two different time points $t_1, t_2 \in [0, 1]$ with $|t_1 - t_2| \leq h$ small, then the difference between the values at t_1 and t_2 is not greater

[†]Corresponding author. Institute for Applied Mathematics, University of Heidelberg, Im Neuenheimer Feld 294, 69120 Heidelberg, Germany. E-Mail: fischer@statlab.uni-heidelberg.de

[‡]Department of Mathematics, University of Rome "La Sapienza", Piazzale Aldo Moro 2, 00185 Roma, Italy. E-Mail: nappo@mat.uniroma1.it

than a multiple of $\sqrt{h \ln(\frac{1}{h})}$, where the proportionality factor depends on the trajectory (and on the time horizon, here equal to one), but not on the choice of the time points. This is a consequence of Lévy's theorem on the uniform modulus of continuity of Brownian motion (cf. Lévy, 1937: p. 172). Let us recall the definition of the modulus of continuity of a deterministic function.

Definition 1. Let $f : [0, \infty) \to \mathbb{R}^d$ and $T > 0$. Then the *modulus of continuity* of f on the interval $[0, T]$ is the function $\mathbf{w}_f(., T)$ defined by

$$[0, \infty) \ni h \mapsto \mathbf{w}_f(h, T) := \sup_{t,s \in [0,T], |t-s| \leq h} |f(t) - f(s)| \in [0, \infty].$$

Here and in what follows, $|.|$ denotes Euclidean distance of appropriate dimension. Let f be any function $[0, \infty) \to \mathbb{R}^d$. Then, according to Definition 1, we have $\mathbf{w}_f(0, T) = 0$, $\mathbf{w}_f(h_1, T) \leq \mathbf{w}_f(h_2, T)$ for all $0 \leq h_1 \leq h_2$, all $T > 0$, and $\mathbf{w}_f(h, T_1) \leq \mathbf{w}_f(h, T_2)$ for all $0 < T_1 \leq T_2$, $h \geq 0$. Moreover, f is continuous on $[0, T]$ if and only if $\mathbf{w}_f(h, T)$ tends to zero as the mesh size h goes to zero.

The modulus of continuity of a stochastic process is a random element for any fixed mesh size $h > 0$. The following results show that the moments of the modulus of continuity of Brownian motion and, more generally, of Itô processes whose coefficients satisfy suitable integrability conditions allow for upper bounds of the form

$$(1) \qquad \mathbf{E}\left[\left(\mathbf{w}_Y(h, T)\right)^p\right] \leq \tilde{C}(p)\left(h \ln(\tfrac{2T}{h})\right)^{\frac{p}{2}} \quad \text{for all } h \in (0, T],$$

where $\tilde{C}(p)$ is a finite positive number depending on the moment p and the coefficients of the Itô process Y. In the case of Brownian motion, the order in h and T of the p-th moments of the modulus of continuity as given by Inequality (1) is exact.

Results to the effect that Inequality (1) holds in case Y is a Brownian motion or a d-dimensional Itô process with uniformly bounded coefficients can be found in various places in the literature, and are linked to various applications, most of them concerning approximation problems.

In Ritter (1990), the problem of recovering a merely continuous univariate function from a finite number of function evalutions is considered in the "average case setting" of information-based complexity theory with respect to the Wiener measure. From Theorem 2 therein it is possible to deduce that the pth-moment of the modulus of continuity of Brownian motion asymptotically behaves like $(\ln(n)/n)^{p/2}$ in the mesh size $h = 1/n$.

In the works by Słomiński (1994, 2001) and Pettersson (1995), Euler schemes for the approximate solution of stochastic differential equations with reflection (in the sense of the Skorohod problem) are studied. Lemma 4.4 of Pettersson (1995) gives the order (in the mesh size) of the second moment of the modulus of continuity of Brownian motion. In

Słomiński (1994), a suboptimal bound of order $h^{p/2-\varepsilon}$ for the pth-moment of the modulus of continuity of Itô processes with bounded coefficients was obtained, while Lemma A.4 of Słomiński (2001) provides bounds of the right order of $(h \ln(1/h))^{p/2}$ for all (integer) moments p. The estimates are extended from the case of Brownian motion (for which an inequality due to Utev (1981) is used) to that of Itô processes with bounded coefficients by means of a time-change argument, which works in a straightforward way thanks to the boundedness assumption.

We too start from the special case of one-dimensional Brownian motion. In Section 2 we collect some classical results from the theory of extreme values in order to show that, for each $p > 0$, there are finite and strictly positive constants $c(p)$ and $C(p)$ such that

$$(2) \quad c(p) \left(h \ln(\tfrac{2T}{h}) \right)^{\frac{p}{2}} \leq \mathbf{E} \left[\left(\mathbf{w}_W(h,T) \right)^p \right] \leq C(p) \left(h \ln(\tfrac{2T}{h}) \right)^{\frac{p}{2}} \quad \text{for all } h \in (0, T],$$

where W is a standard one-dimensional Wiener process. Clearly, the above inequalities imply that

$$(3) \quad c(p) \left(h \ln(\tfrac{T}{h}) \right)^{\frac{p}{2}} \leq \mathbf{E} \left[\left(\mathbf{w}_W(h,T) \right)^p \right] \leq 2^p \, C(p) \left(h \ln(\tfrac{T}{h}) \right)^{\frac{p}{2}} \quad \text{for all } h \in (0, T/2].$$

Observe that neither of the inequalities in (2) and (3) can be deduced from the traditional formulation of Lévy's theorem on the pathwise modulus of continuity of Brownian motion; but cf. Remark 1 below. In addition, we provide explicit bounds on the constant $C(p)$; their derivation, though in the spirit of extreme value theory, is elementary in that it only relies on a few well-known properties of the normal distribution, see Lemma 3 and its proof.

In the appendix, we derive explicit bounds on the constants $C(p)$, again in the case of Brownian motion, using the ideas of Exercise 2.4.8 from Stroock and Varadhan (1979), where an alternative proof of Lévy's result on the pathwise uniform modulus of continuity, based on the Garsia-Rodemich-Rumsey lemma (Garsia et al., 1970), is sketched. The technique also appears in Section 3 of Friz and Victoir (2005), where the pathwise modulus of continuity of enhanced Brownian motion is computed.

In Section 3, we apply the bounds on the moments of the Wiener modulus of continuity in terms of the mesh size, the moment and the time horizon in order to prove Inequality (1), again by a time-change argument, also for Itô processes with possibly unbounded coefficients, see Theorem 1 and Remark 2.

As an application of Theorem 1, we consider the problem of approximately computing the solution to a stochastic delay (or functional) differential equation (SDDE / SFDE). This problem is actually the motivation of our interest in the modulus of continuity, the aim being to generalise some strong convergence results obtained in Calzolari et al. (2007).

Let us first recall the particular case of ordinary stochastic differential equations (SDEs). In this case, it is well known that the pth-moment of the approximation error produced

by a standard Euler-Maruyama discretisation scheme is of order $h^{p/2}$ in the mesh size h of the time grid. Such a result holds on condition that in measuring the error only grid points are taken into account.

Suppose, now, that the approximation error is measured in supremum norm, that is, as the maximal difference over the time interval between the exact solution and the continuous-time approximate solution, where the latter is obtained from the discrete-time approximate solution by piecewise constant or piecewise linear interpolation with respect to the time grid. In the case of SDEs, upper bounds for the p-th moment of the error of order $(h \ln(1/h))^{p/2}$ in the mesh size h can then be derived, see Faure (1992) and Remark B.1.5 in Bouleau and Lépingle (1994: Ch. 5). A similar upper bound in the case of bounded coefficients had been obtained before by Kanagawa (1988); his proof implicitly uses the fact that the moments of the modulus of continuity of Brownian motion are all finite.

In fact, with respect to the above error criterion, the order of the error bound cannot be improved beyond $(h \ln(1/h))^{p/2}$, neither for an Euler scheme nor for any other scheme using the same information, see Hofmann et al. (2000). In Müller-Gronbach (2002), an adaptive Euler scheme is introduced which is optimal also in the sense that it attains, for any moment, the asymptotically optimal constant in the error bound.

For a discussion of error criteria in the approximation of SDEs see Sections 10.1 and 10.2 in Asmussen and Glynn (2007). Proposition 2.1 there also gives the asymptotic order of the expected value of the Euler modulus of continuity (see Definition 2 below) in the case of Brownian motion.

Observe that trajectories of the continuous-time approximate solution can be actually computed – up to machine precision – at any point in time, not only at the grid points, by a simple polygonal interpolation. Values of the driving Wiener process, in particular, are needed at grid points only. Measuring the error in supremum norm in this way allows to approximate in a strong sense functionals which depend on the trajectories of the exact solution.

In some works, a continuous-time version of the approximate solution to an SDE is obtained by "continuous interpolation": On any time interval between two neighbouring grid points, the approximate solution is a diffusion with constant coefficients driven by the original Wiener process; see, for instance, Section 10.2 in Kloeden and Platen (1999). The expected error in supremum norm between the approximate solution so obtained and the exact solution is of order $h^{1/2}$ (or $h^{p/2}$ for the p-th moment) in the mesh size h. This way of constructing the continuous-time approximate solution, while convenient for the error analysis, does not allow to approximate path-dependent functionals. In order to evaluate the continuous-time approximate solution at any point not belonging to the time grid,

one has to know the value of the driving Wiener process also at that point. In order to determine the value of a path-dependent functional, for instance the maximum over a time interval, one would have to know the exact values of the driving Wiener process over an entire time interval.

In Section 4, we consider the problem of strong uniform convergence for piecewise linear Euler-Maruyama approximation schemes in the case of SDDEs / SFDEs whose coefficients satisfy a functional Lipschitz condition in supremum norm. We show that the rate of convergence is of order $(h \ln(1/h))^{p/2}$, thus extending the above mentioned results. When the continuous-time approximate solution is built from the underlying Wiener process by continuous interpolation, then the approximation error of the Euler-Maruyama scheme is of order $h^{p/2}$ also in the case of SDDEs / SFDEs, see Section 5 in Mao (2003). Notice that the Lipschitz condition there is more restrictive than the one adopted below. In Hu et al. (2004), the strong rate of convergence of a Milstein scheme for SDDEs with point delay is shown to be of first order (in contrast to the order $\frac{1}{2}$ of the Euler scheme). The continuous-time approximate solution employed there is not fully implementable, as it is constructed in the same way as in the work by Kloeden and Platen (1999) mentioned above.

2   Moments of the modulus of continuity of one-dimensional Brownian motion

Let W be a standard one-dimensional Wiener process living on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$, and denote by $\mathbf{w}_W$ its modulus of continuity, that is, the random element $\Omega \ni \omega \mapsto \mathbf{w}_{W(\omega)}$, where $\mathbf{w}_{W(\omega)}$ is the modulus of continuity of the path $[0, \infty) \ni t \mapsto W(t, \omega)$ in the sense of Definition 1.

The aim of this section is to prove the inequalities in (2). The proof is based on classical results in extreme value theory; it also yields a representation of the constants $c(p)$ and $C(p)$. We start by introducing another modulus of continuity, which we will call the Euler modulus of continuity.

Definition 2. Let f be a deterministic function. The *Euler modulus of continuity* of f on the interval $[0, T]$ is the function $\mathbf{w}_f^E(., T)$ defined by

$$[0, \infty) \ni h \mapsto \mathbf{w}_f^E(h, T) := \sup_{t \in [0, T]} |f(t) - f(h \lfloor t/h \rfloor)| \in [0, \infty].$$

In view of the above definition, it is immediate to check that

(4) $\mathbf{w}_f^E(h, T) \leq \mathbf{w}_f(h, T) \leq 3\, \mathbf{w}_f^E(h, T)$ for all $h \in (0, T]$,

whence we may concentrate on the Euler modulus of continuity.

Lemma 1. *Let W be a standard one-dimensional Wiener process, and let $(Z_i)_{i \in \mathbb{N}}$ be a sequence of independent random variables with standard normal distribution. Then, for all $p > 0$, all $h \in (0, T]$,*

$$(5) \quad \mathbf{E}\left[\max_{i=1,\ldots,\lfloor T/h \rfloor} |Z_i|^p\right] h^{p/2} \;\leq\; \mathbf{E}\left[\left(\mathbf{w}_W^E(h, T)\right)^p\right] \;\leq\; 2\,\mathbf{E}\left[\max_{i=1,\ldots,\lceil T/h \rceil} |Z_i|^p\right] h^{p/2}.$$

Proof. It is clear that, for any continuous function f, any $n \in \mathbb{N}$,

$$\mathbf{w}_f^E(h, n\,h) \;=\; \max_{i=1,\ldots,n} \;\sup_{t \in [(i-1)h, ih]} \big|f(t) - f((i-1)\,h)\big|$$

$$= \max_{i=1,\ldots,n} \max\big\{\Delta_f(i, h),\, \Delta_{-f}(i, h)\big\},$$

where $\Delta_f(i, h) := \sup\big\{f(t) - f((i-1)h);\; t \in [(i-1)h, i\,h]\big\}$. It is easily seen that

$$\big(\max_{i=1,\ldots,\lfloor T/h \rfloor} \Delta_f(i, h)\big)^p \;\leq\; \big(\mathbf{w}_f^E(h, T)\big)^p$$

$$\leq\; \big(\max_{i=1,\ldots,\lceil T/h \rceil} \Delta_f(i, h)\big)^p \;+\; \big(\max_{i=1,\ldots,\lceil T/h \rceil} \Delta_{-f}(i, h)\big)^p.$$

By the symmetry of Brownian motion, it follows that

$$(6) \quad \mathbf{E}\left[\big(\max_{i=1,\ldots,\lfloor T/h \rfloor} \Delta_W(i, h)\big)^p\right] \;\leq\; \mathbf{E}\left[\big(\mathbf{w}_W^E(h, T)\big)^p\right] \;\leq\; 2\,\mathbf{E}\left[\big(\max_{i=1,\ldots,\lceil T/h \rceil} \Delta_W(i, h)\big)^p\right].$$

Then, by means of a rescaling argument, we immediately get that

$$\mathbf{E}\left[\big(\max_{i=1,\cdots,\lfloor T/h \rfloor} \Delta_W(i, h)\big)^p\right] \;=\; h^{p/2}\, \mathbf{E}\left[\big(\max_{i=1,\ldots,\lceil T/h \rceil} \Delta_W(i, 1)\big)^p\right].$$

Finally, the well known fact that the random variables

$$\Delta_W(i, 1) \;=\; \sup\big\{W(t) - W(i-1);\; t \in [i-1, i]\big\}$$

are independent, with the same distribution as $|Z_i|$, ends the proof of the inequalities (5).
□

Lemma 2. *Let $(Z_i)_{i \in \mathbb{N}}$ be a sequence of independent standard Gaussian random variables. Then, for all $p > 0$,*

$$(7) \quad \lim_{n \to \infty} \mathbf{E}\left[\left(\frac{\max_{i=1,\ldots,n} |Z_i|}{\sqrt{2 \ln(2n)}}\right)^p\right] \;=\; 1.$$

Before giving a proof of the above result, let us show how that it implies the inequalities in (2). Indeed, by (7),

$$(8) \quad c_0(p)\,(2\ln(2n))^{p/2} \;\leq\; \mathbf{E}\left[\max_{i=1,\ldots,n} |Z_i|^p\right] \;\leq\; C_0(p)\,(2\ln(2n))^{p/2},$$

6

for every $n \in \mathbb{N}$, where

$$(9) \qquad c_0(p) := \inf_{n \in \mathbb{N}} \frac{\mathbf{E}\left[\max_{i=1,\ldots,n} |Z_i|^p\right]}{(2\ln(2n))^{p/2}}, \qquad C_0(p) := \sup_{n \in \mathbb{N}} \frac{\mathbf{E}\left[\max_{i=1,\ldots,n} |Z_i|^p\right]}{(2\ln(2n))^{p/2}}.$$

Therefore, taking into account that

$$\mathbf{w}_f^E\big(h, \lfloor \tfrac{T}{h} \rfloor h\big) \;\leq\; \mathbf{w}_f^E(h,T) \;\leq\; \mathbf{w}_f^E\big(h, \lceil \tfrac{T}{h} \rceil h\big),$$

and that, for $x \geq 1$, $\frac{1}{2}\ln(2x) \leq \ln(2\lfloor x \rfloor) \leq \ln(2\lceil x \rceil) \leq 2\ln(2x)$, we find that

$$(10) \qquad c_0(p)\,\big(h\,\ln(\tfrac{2T}{h})\big)^{p/2} \;\leq\; \mathbf{E}\left[\big(\mathbf{w}_W^E(h,T)\big)^p\right] \;\leq\; 2\,C_0(p)\,\big(4h\,\ln(\tfrac{2T}{h})\big)^{p/2},$$

where the bounds are valid for all $h \in (0,T]$. The inequalities in (2) then follow from those in (4) where

$$(11) \qquad c(p) \;=\; c_0(p), \qquad\qquad C(p) \;=\; 2 \cdot 6^p \cdot C_0(p).$$

The proof of Lemma 2 is based on some classical results of extreme value theory concerning the sequence

$$M_n := \max_{i=1,\ldots,n} X_i,$$

where $(X_i)_{i \in \mathbb{N}}$ is a sequence of i.i.d. random variables with common distribution function F. Simplified versions of these results are stated below in Propositions 1 and 2.

Proposition 1 (Gnedenko (1943)). *Assume that $F(x) < 1$ for every $x \in \mathbb{R}$ and that the sequence $(\beta_n) \subset \mathbb{R}$ converges to infinity. Then the sequence $(\frac{M_n}{\beta_n})$ converges to 1 in probability if and only if for any $\epsilon > 0$,*

$$\lim_{n \to \infty} n\big(1 - F\big(\beta_n(1+\epsilon)\big)\big) \;=\; 0,$$
$$\lim_{n \to \infty} n\big(1 - F\big(\beta_n(1-\epsilon)\big)\big) \;=\; +\infty.$$

Proposition 2 (Theorem 3.2 in Pickands (1968), Exercise 2.1.3 in Resnick (1987)). *Assume that $F(x) < 1$ for every $x \in \mathbb{R}$ and that $\mathbf{E}[(X_1)_{-}^p] < \infty$, where $p > 0$ and $(a)_{-} := \max\{-a, 0\}$. Assume further that the sequence $(\frac{M_n}{\beta_n})$ converges to 1 in probability. Then $\lim_{n \to \infty} \mathbf{E}\left[(M_n/\beta_n)^p\right] = 1$.*

Proof of Lemma 2. In our case $X_i = |Z_i|$, so that $M_n := \max_{i=1,\ldots,n} |Z_i|$ and, if $\Phi(x)$ denotes the distribution function of Z_i, then the validity of (7) follows from Proposition 2, because **(i)** $1 - F(x) = 2(1 - \Phi(x)) < 1$ for every $x \in \mathbb{R}$, **(ii)** the negative part of X_1 is zero, **(iii)** $(\frac{M_n}{\beta_n})$ converges to 1 in probability, where $\beta_n := \sqrt{2\ln(2n)}$.

Properties **(i)** and **(ii)** are obvious, and we only need to check property **(iii)**. Let $\varphi = \Phi'$ denote the density function of Z_i, so that asymptotically $1 - \Phi(x) \sim \varphi(x)/x$. Then, for any $\alpha = 1 \pm \epsilon > 0$, we have $n(1 - F(\beta_n \alpha)) = 2n(1 - \Phi(\beta_n \alpha)) \sim 2n\frac{\varphi(\beta_n \alpha)}{\beta_n \alpha}$, whence

$$n\left(1 - F(\beta_n \,\alpha)\right) \sim \sqrt{\frac{2}{\alpha^2\,\pi}}\,\frac{n}{\sqrt{\ln 2n}}\,e^{-\frac{1}{2}\,2\alpha^2\,\ln(2n)} = \sqrt{\frac{2}{\alpha^2\,\pi}}\,\frac{n}{\sqrt{\ln 2n}}\,(2n)^{-\alpha^2} =: \hat{C}(\alpha)\,\frac{n^{1-\alpha^2}}{\sqrt{\ln 2n}}.$$

The sequence $\frac{n^{1-\alpha^2}}{\sqrt{\ln 2n}}$ converges to zero or to infinity as $\alpha = 1 + \epsilon > 1$ or $\alpha = 1 - \epsilon < 1$, whence we can apply Proposition 1 in order to obtain property **(iii)**, i.e. the relative stability of (M_n). $\qquad\square$

Remark 1. As already observed, neither of the inequalities in (2) follows from Lévy's theorem on the pathwise uniform modulus of Brownian motion, which states, in the notation of this section, that

$$(12) \qquad \mathbf{P}\left(\limsup_{h \to 0+}\ \sup_{t,s \in [0,1], |t-s|=h}\ \frac{|W(t) - W(s)|}{\sqrt{2h \ln(\frac{1}{h})}} = 1\right)\ =\ 1,$$

see Lévy (1937: pp. 168-172) and Itô and McKean (1974: pp. 36-38). Equation (12) implies that there is a finite, non-negative random variable M such that, for **P**-almost all $\omega \in \Omega$,

$$\sup_{t,s \in [0,1], |t-s| \le h}\ |W(t,\omega) - W(s,\omega)|\ \le\ M(\omega)\,\sqrt{h \ln(\tfrac{1}{h})}\quad \text{for all } h \in (0, \tfrac{1}{2}],$$

while nothing can be said about the moments of M (clearly, M cannot be bounded from above). Consequently, no upper bounds on the moments of $\mathbf{w}_W(.,1)$ can be obtained from Lévy's result.

As regards the lower bounds, Equation (12) does not allow to conclude that the constants $c(p)$ in (2) are strictly positive, because it does not guarantee that the lower limit in (12) is positive. "A closer examination of the proof" of Lévy's theorem (Taylor, 1974: §3), however, shows that

$$(13) \qquad \mathbf{P}\left(\liminf_{h \to 0+}\ \sup_{t,s \in [0,1], |t-s|=h}\ \frac{|W(t) - W(s)|}{\sqrt{2h \ln(\frac{1}{h})}} = 1\right)\ =\ 1.$$

Indeed, the first part of the proof as given in Itô and McKean (1974: p. 37) still works if, instead of using 2^n subintervals of length 2^{-n}, we consider increments of the Brownian path over $\lfloor 1/h \rfloor$ subintervals of length h.

Equation (13) guarantees the existence of strictly positive constants $c(p)$ in (2), although it does not yield any representation of these constants.

The next lemma provides explicit upper bounds on the moments of the modulus of continuity of one-dimensional Brownian motion.

Lemma 3. *Let W be a standard one-dimensional Wiener process. Then for any $p > 0$, any $T > 0$,*

$$(14) \quad \mathbf{E}\left[\left(\mathbf{w}_W(h,T)\right)^p\right] \ \leq \ \frac{5}{\sqrt{\pi}} \cdot (6/\sqrt{\ln(2)})^p \cdot \Gamma\left(\tfrac{p+1}{2}\right) \cdot \left(h\ln(\tfrac{2T}{h})\right)^{\frac{p}{2}} \quad \textit{for all } h \in (0,T].$$

Proof. Let $p > 0$, $T > 0$, $h \in (0,T]$. Let $(Z_i)_{i\in\mathbb{N}}$ be a sequence of independent random variables with standard normal distribution. Denote by $\mathbf{w}_W^E(.,T)$ the Euler modulus of continuity according to Definition 2. By the upper bound in (2), taking into account (11) and (9), we see that

$$\begin{aligned}
\mathbf{E}\left[\left(\mathbf{w}_W(h,T)\right)^p\right] \ &\leq \ 2 \cdot 6^p \cdot \sup_{n\in\mathbb{N}} \frac{\mathbf{E}\left[\max_{i=1,\ldots,n}|Z_i|^p\right]}{(2\ln(2n))^{p/2}} \cdot \left(h\ln(\tfrac{2T}{h})\right)^{\frac{p}{2}}. \\
&= \ 2 \cdot 6^p \cdot \left(h\ln(\tfrac{2T}{h})\right)^{\frac{p}{2}} \cdot \sup_{n\in\mathbb{N}} \mathbf{E}\left[\left(\frac{M_n}{\beta_n}\right)^p\right],
\end{aligned}$$

where $\beta_n := \sqrt{2\ln(2n)}$ and $M_n := \max_{i=1,\ldots,n}|Z_i|$, $n \in \mathbb{N}$, as in the proof of Lemma 2. It is therefore sufficient to obtain a uniform upper bound for the expectations $\mathbf{E}[(M_n/\beta_n)^p]$. Indeed, for all $n \in \mathbb{N}$ it holds that

$$\begin{aligned}
\mathbf{E}\left[\left(\frac{M_n}{\beta_n}\right)^p\right] \ &= \ \int_0^\infty \left(1 - \mathbf{P}\left(M_n \leq x^{1/p}\beta_n\right)\right) dx \\
&\leq \ 2^{-p/2} + \int_{2^{-p/2}}^\infty \left(1 - \left(F\left(x^{1/p}\beta_n\right)\right)^n\right) dx \\
&\leq \ 2^{-p/2} + \int_{2^{-p/2}}^\infty n\left(1 - F\left(x^{1/p}\beta_n\right)\right) dx,
\end{aligned}$$

where $F(.)$ is the common distribution function of the i.i.d. sequence $(|Z_i|)_{i\in\mathbb{N}}$, that is, $F(x) = 2\Phi(x) - 1$ with $\Phi(.)$ the standard normal distribution function. The last of the above inequalities holds, because $(1 - a^n) \leq n(1-a)$ for any $a \in [0,1]$.

At this point, let $\beta(.)$ be the function $[1,\infty) \ni t \mapsto \sqrt{2\ln(2t)} \in [0,\infty)$ and observe that, for fixed $y \geq 1$, the mapping $[1,\infty) \ni t \mapsto t \cdot \left(1 - F\left(y\,\beta(t)\right)\right)$ is non-increasing. To see this, notice that $t \cdot \beta'(t) = \frac{2}{\beta(t)}$ for $t \geq 1$, observe that

$$F'(x) \ = \ 2\,\varphi(x), \qquad 1 - F(x) \ = \ 2\left(1 - \Phi(x)\right) \ \leq \ 2\,\frac{\varphi(x)}{x}, \qquad x > 0,$$

where $\varphi = \Phi'$, and check that, for all $y \geq 1/\sqrt{2}$, $t \geq 1$,

$$\begin{aligned}
\frac{d}{dt}\left(t \cdot \left(1 - F\left(y\,\beta(t)\right)\right)\right) \ &= \ 1 - F\left(y\,\beta(t)\right) - 2\,\varphi\left(y\,\beta(t)\right)\frac{2y}{\beta(t)} \\
&\leq \ \frac{2\varphi(y\beta(t))}{y\beta(t)}(1 - 2\,y^2) \ \leq \ 0.
\end{aligned}$$

9

Since $x^{1/p} \geq 1/\sqrt{2}$ whenever $x \geq 2^{-p/2}$, it follows that

$$
\mathbf{E}\left[\left(\frac{M_n}{\beta_n}\right)^p\right] \quad \leq \quad 2^{-p/2} + \int_{2^{-p/2}}^{\infty} \left(1 - F\left(x^{1/p}\beta_1\right)\right) dx \quad \leq \quad 2^{-p/2} + \mathbf{E}\left[\left(\frac{|Z_1|}{\beta_1}\right)^p\right]
$$

$$
= \quad 2^{-p/2} + \left(2\ln(2)\right)^{-p/2}\mathbf{E}\left[(|Z_1|)^p\right] \quad = \quad 2^{-p/2} + (\ln(2))^{-p/2}\frac{\Gamma(\frac{p+1}{2})}{\sqrt{\pi}},
$$

because Z_1 has standard normal distribution, whence $|Z_1|^2$ has Gamma distribution with shape parameter $\frac{1}{2}$ and scale parameter 2. Consequently, we find that

$$
\mathbf{E}\left[\left(\mathbf{w}_W(h,T)\right)^p\right] \leq \quad 2 \cdot \left(6/\sqrt{\ln(2)}\right)^p \cdot \left((\ln(2)/2)^{p/2} + \frac{\Gamma(\frac{p+1}{2})}{\sqrt{\pi}}\right) \cdot \left(h\ln\left(\frac{2T}{h}\right)\right)^{\frac{p}{2}}
$$

$$
\leq \quad \frac{5}{\sqrt{\pi}} \cdot \left(6/\sqrt{\ln(2)}\right)^p \cdot \Gamma\left(\frac{p+1}{2}\right) \cdot \left(h\ln\left(\frac{2T}{h}\right)\right)^{\frac{p}{2}}.
$$

$\square$

3   Upper bounds for Itô processes

In this section, we turn to deriving upper bounds on the moments of the modulus of continuity for quite general Itô processes.

Theorem 1. *Let W be a d_1-dimensional Wiener process adapted to a filtration $(\mathcal{F}_t)$ satisfying the usual assumptions and defined on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Let $Y = (Y^{(1)}, \ldots, Y^{(d)})^\mathsf{T}$ be an Itô process of the form*

$$
Y(t) = y_0 + \int_0^t b(s)ds + \int_0^t \sigma(s)dW(s), \quad t \geq 0,
$$

where $y_0 \in \mathbb{R}^d$ and b, σ are $(\mathcal{F}_t)$-adapted processes with values in $\mathbb{R}^d$ and $\mathbb{R}^{d \times d_1}$, respectively. Let $T > 0$, and let $\zeta = \zeta_T$, $\xi = \xi_T$ be $[0, \infty]$-valued $\mathcal{F}_T$-measurable random variables such that for all $t, s \in [0, T]$, all $i \in \{1, \ldots, d\}$, $j \in \{1, \ldots, d_1\}$, all $\omega \in \Omega$,

$$
\int_s^t \left|b_i(\tilde{s}, \omega)\right| d\tilde{s} \quad \leq \quad \zeta(\omega) \cdot \sqrt{|t-s|\ln\left(\frac{2T}{|t-s|}\right)},
$$

$$
\int_s^t \sigma_{i,j}^2(\tilde{s}, \omega) d\tilde{s} \quad \leq \quad \xi(\omega) \cdot |t-s|.
$$

Let $p \geq 1$. If the processes b, σ are such that

(H1) $\mathbf{E}[\zeta^p] < \infty$,

(H2) *there is $\varepsilon > 0$ such that $\mathbf{E}\left[\xi^{\frac{p}{2}+\varepsilon}\right] < \infty$,*

then there is a finite constant $\tilde{C}(p) > 0$ such that (1) holds, that is,

(15) $\qquad \mathbf{E}\left[\left(\mathbf{w}_Y(h,T)\right)^p\right] \leq \tilde{C}(p)\left(h\ln(\frac{2T}{h})\right)^{\frac{p}{2}} \quad$ *for all $h \in (0, T]$.*

Proof. With $T > 0$, $p \geq 1$, it holds for all $t, s \in [0, T]$ that

$$\left|Y(t) - Y(s)\right|^p \quad \leq \quad d^{\frac{p}{2}} \left(\left|Y^{(1)}(t) - Y^{(1)}(s)\right|^p + \ldots + \left|Y^{(d)}(t) - Y^{(d)}(s)\right|^p\right),$$

and for the i-th component of Y we have

$$\left|Y^{(i)}(t) - Y^{(i)}(s)\right|^p \quad = \quad \left|\int_s^t b_i(\tilde{s})d\tilde{s} + \sum_{j=1}^{d_1} \int_s^t \sigma_{ij}(\tilde{s})dW^j(\tilde{s})\right|^p$$

$$\leq \quad (d_1 + 1)^p \left(\zeta^p \left(|t - s| \ln\left(\frac{2T}{|t-s|}\right)\right)^{\frac{p}{2}} + \sum_{j=1}^{d_1} \left|\int_s^t \sigma_{ij}(\tilde{s})dW^j(\tilde{s})\right|^p\right).$$

Hence, by Hypothesis (H1), for any $h \in (0, T]$,

$$\mathbf{E}\left[\left(\mathbf{w}_Y(h, T)\right)^p\right] \quad = \quad \mathbf{E}\left[\sup_{t, s \in [0,T], |t-s| \leq h} \left|Y(t) - Y(s)\right|^p\right]$$

$$\leq \quad d^{\frac{p}{2}}(d_1 + 1)^p \left(d \cdot \mathbf{E}\left[\zeta^p\right] \cdot \left(h \ln\left(\frac{2T}{h}\right)\right)^{\frac{p}{2}} + \sum_{i=1}^{d} \sum_{j=1}^{d_1} \mathbf{E}\left[\sup_{t, s \in [0,T], |t-s| \leq h} \left|\int_s^t \sigma_{ij}(\tilde{s})dW^j(\tilde{s})\right|^p\right]\right).$$

To prove the assertion, it is enough to show that the $d \cdot d_1$ expectations on the right-hand side of the last inequality are finite and of the right order in h. Let $i \in \{1, \ldots, d\}$, $j \in \{1, \ldots, d_1\}$, and define the one-dimensional process $M = M^{(i,j)}$ by

$$M(t) := \int_0^{t \wedge T} \sigma_{ij}(\tilde{s})\, dW^j(\tilde{s}) + \left(W^j(t) - W^j(T)\right) \cdot \mathbf{1}_{\{t > T\}}, \quad t \geq 0.$$

Since $\int_0^T \sigma_{ij}^2(\tilde{s})d\tilde{s}$ is $\mathbf{P}$-almost surely finite as a consequence of Hypothesis (H2), the process M is a (continuous) local martingale vanishing at zero, and M can be represented as a time-changed Brownian motion. More precisely, by the Dambis-Dubins-Schwarz theorem, for instance Theorem 3.4.6 in Karatzas and Shreve (1991: p. 174), there is a standard one-dimensional Brownian motion $\tilde{W}$ living on $(\Omega, \mathcal{F}, \mathbf{P})$ such that, $\mathbf{P}$-almost surely,

$$M(t) = \tilde{W}(\langle M \rangle_t) \quad \text{for all } t \geq 0,$$

where $\langle M \rangle$ is the quadratic variation process associated with M, that is,

$$\langle M \rangle_t = \int_0^{t \wedge T} \sigma_{ij}^2(\tilde{s})\, d\tilde{s} + (t - T) \vee 0, \quad t \geq 0.$$

Consequently, it holds $\mathbf{P}$-almost surely that

$$\sup_{t, s \in [0,T], |t-s| \leq h} \left|\int_s^t \sigma_{ij}(\tilde{s})dW^{(j)}(\tilde{s})\right|^p \quad = \quad \sup_{t, s \in [0,T], |t-s| \leq h} \left|\tilde{W}(\langle M \rangle_t) - \tilde{W}(\langle M \rangle_s)\right|^p$$

$$\leq \quad \sup\left\{\left|\tilde{W}(u) - \tilde{W}(v)\right|^p ; u, v \in [0, \langle M \rangle_T], |u - v| \leq \sup_{t \in [h,T]} \langle M \rangle_t - \langle M \rangle_{t-h}\right\}$$

$$\leq \quad \left(\mathbf{w}_{\tilde{W}}(\delta_h, \tau)\right)^p,$$

where τ and δ_s, $s \in [0,T]$, are the random elements defined by

$$\tau(\omega) := \langle M \rangle_T(\omega), \qquad \delta_s(\omega) := \sup_{t \in [s,T]} \langle M \rangle_t(\omega) - \langle M \rangle_{t-s}(\omega), \qquad \omega \in \Omega.$$

Notice that $\tau(\omega) = \langle M \rangle_T(\omega) = \delta_T(\omega)$, $\delta_h(\omega) \le \xi(\omega)h$ and $\tau(\omega) \le \xi(\omega)T$ for all $\omega \in \Omega$, $h \in (0,T]$. By the monotonicity of the modulus of continuity it follows that

$$\mathbf{w}_{\tilde{W}(\omega)}\big(\delta_h(\omega), \tau(\omega)\big) \ \le \ \mathbf{w}_{\tilde{W}(\omega)}\big(\xi(\omega)h, \tau(\omega)\big) \ \le \ \mathbf{w}_{\tilde{W}(\omega)}\big(\xi(\omega)h, \xi(\omega)T\big), \quad \omega \in \Omega.$$

Let $\alpha > 1$. Then, by Hölder's inequality and Lemma 3, for all $h \in (0,T]$ it holds that

$$
\begin{aligned}
\mathbf{E}\left[\big(\mathbf{w}_{\tilde{W}}(\delta_h, \tau)\big)^p\right] \quad &\le \quad \mathbf{E}\left[\big(\mathbf{w}_{\tilde{W}}(\xi\,h, \xi\,T)\big)^p\right] \\
&\le \quad \sum_{n=1}^{\infty} \mathbf{E}\left[\mathbf{1}_{\{\xi \in (n-1,n]\}}\big(\mathbf{w}_{\tilde{W}}(nh, nT)\big)^p\right] \\
&\le \quad \sum_{n=1}^{\infty} \mathbf{P}\{\xi \in (n-1,n]\}^{\frac{\alpha-1}{\alpha}} \cdot \mathbf{E}\left[\big(\mathbf{w}_{\tilde{W}}(nh, nT)\big)^{\alpha \cdot p}\right]^{\frac{1}{\alpha}} \\
&\le \quad \left(\sum_{n=1}^{\infty} \mathbf{P}\{\xi \in (n-1,n]\}^{\frac{\alpha-1}{\alpha}} \cdot n^{\frac{p}{2}}\right) C_{p,\alpha} \cdot \big(h\ln(\tfrac{2T}{h})\big)^{\frac{p}{2}} \\
&\le \quad \left(\sum_{n=1}^{\infty} \mathbf{P}\{\xi \in (n-1,n]\} \cdot n^{\frac{(p+\varepsilon)\alpha}{2(\alpha-1)}}\right)^{\frac{\alpha-1}{\alpha}} \left(\sum_{n=1}^{\infty} n^{-\frac{\varepsilon \cdot \alpha}{2}}\right)^{\frac{1}{\alpha}} C_{p,\alpha} \cdot \big(h\ln(\tfrac{2T}{h})\big)^{\frac{p}{2}} \\
&\le \quad \mathbf{E}\left[(\xi+1)^{\frac{(p+\varepsilon)\alpha}{2(\alpha-1)}}\right]^{\frac{\alpha-1}{\alpha}} \left(\sum_{n=1}^{\infty} n^{-\frac{\varepsilon \cdot \alpha}{2}}\right)^{\frac{1}{\alpha}} C_{p,\alpha} \cdot \big(h\ln(\tfrac{2T}{h})\big)^{\frac{p}{2}},
\end{aligned}
$$

where $C_{p,\alpha} := (5/\sqrt{\pi})^{1/\alpha} \cdot (6/\sqrt{\ln(2)})^p \cdot \Gamma(\tfrac{\alpha p+1}{2})^{1/\alpha}$ and $\varepsilon > 0$ is as in Hypothesis (H2). If we choose α greater than $\max\{\tfrac{2}{\varepsilon}, 2\}$, then Hypothesis (H2) implies that the expectation and the infinite sums in the last two lines above are finite. The assertion follows. $\qquad\square$

Remark 2. The proof of Theorem 1 also shows how to obtain bounds on the moments of the modulus of continuity when the diffusion coefficient is not pathwise essentially bounded, but instead satisfies Hypotheses (H1) and (H2) with

$$
\begin{aligned}
\int_s^t \big|b_i(\tilde{s},\omega)\big|d\tilde{s} \quad &\le \quad \zeta(\omega) \cdot \sqrt{|t-s|^\beta \ln\big(\tfrac{2T}{|t-s|}\big)}, \\
\int_s^t \sigma_{i,j}^2(\tilde{s},\omega)d\tilde{s} \quad &\le \quad \xi(\omega) \cdot |t-s|^\beta
\end{aligned}
$$

for some $\beta \in (0,1)$. The order of the p-th moment changes accordingly, namely from $(h\ln(\tfrac{2T}{h}))^{p/2}$ to $(h^\beta \ln(\tfrac{2T}{h}))^{p/2}$. Moreover, in virtue of Lemma 3, upper bounds for the constant $\tilde{C}(p)$ can be computed.

4 Euler approximation for SDDEs and the modulus of continuity

In this section, we consider the approximation of solutions to stochastic delay differential equations (SDDEs). More precisely, let $\boldsymbol{X} = (X(t))_{t \in [-\tau, T]}$ be a continuous process satisfying

$$
(16) \quad \begin{cases} X(t) = \eta(0) + \int_0^t \mu(u, \Pi_u X) du + \int_0^t \sigma(u, \Pi_u X) dW(u), & 0 \le t \le T, \\ X(t) = \eta(t), & -\tau \le t \le 0, \end{cases}
$$

where τ is a positive constant, $(W(t))_{t \in [0,T]}$ a standard d_1-dimensional Wiener process living on the filtered probability space $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}_t)_{t \in [0,T]})$, $\eta = (\eta(s))_{s \in [-\tau, 0]}$ is a $\mathcal{F}_0$-measurable $C([-\tau, 0], \mathbb{R}^d)$-valued random variable, and $(\Pi_t X)_{t \in [0,T]}$ is the $C([-\tau, 0], \mathbb{R}^d)$-valued process defined by

$$
\Pi_t X(s) := X(t+s), \quad -\tau \le s \le 0.
$$

The process $(\Pi_t X)$ is called the *segment process* associated with the state process $\boldsymbol{X}$, and $\mathcal{C} := C([-\tau, 0], \mathbb{R}^d)$ is called the *segment space*. The segment space $\mathcal{C}$ is equipped, as usual, with the supremum norm, denoted by $\|.\|$.

As an example, the functions $\mu(t, \theta)$ and $\sigma(t, \theta)$ with $\theta \in \mathcal{C}$ can be taken to be of the form

$$
(17) \quad g\left(t, \max_{u \in [\tau_{i-1}, \tau_i]} \theta(u); i = 1, \dots, r\right)
$$

where $-\tau = \tau_0 < \tau_1 < \dots < \tau_r = 0$, or

$$
(18) \quad g\left(t, \int_{-\tau}^0 \psi_i(u, \theta(u)) \gamma_i(du); i = 1, \cdots, r\right),
$$

where γ_i are finite measures on $[-\tau, 0]$, and g and ψ_i are appropriate continuous functions. The fixed point delay model

$$
dX(t) = g_\mu\big(X(t), X(t-\tau)\big) dt + g_\sigma\big(X(t), X(t-\tau)\big) dW(t)
$$

can be recovered by choosing, in (18), $r = 2$, $\psi_1(u, x) = \psi_2(u, x) = x$, and γ_1, γ_2 as the Dirac measures δ_0, $\delta_{-\tau}$.

Let us assume that the following conditions hold:

(A1) η is a $\mathcal{F}_0$-measurable $\mathcal{C}$-valued random variable such that

$$
\mathbf{E}\left[\|\Pi_0 X\|^{2k}\right] = \mathbf{E}\left[\sup_{s \in [-\tau, 0]} |\eta(s)|^{2k}\right] < \infty, \quad k = 1, 2.
$$

(A2) The functionals $\mu(t,\theta)$ and $\sigma(t,\theta)$ on $[0,T] \times \mathcal{C}$ are globally Lipschitz continuous in space and α-Hölder continuous in time for some constant $\alpha \in [\frac{1}{2}, 1]$, that is,

$$|\mu(t,\theta) - \mu(t',\bar{\theta})|^2 + |\sigma(t,\theta) - \sigma(t',\bar{\theta})|^2 \leq K\left(|t-t'|^{2\alpha} + \|\theta - \bar{\theta}\|^2\right),$$

and satisfy the growth condition

$$|\mu(t,\theta)|^2 + |\sigma(t,\theta)|^2 \leq K\left(1 + \|\theta\|^2\right)$$

for some constant $K > 0$.

Conditions (A1) and (A2) guarantee the existence and uniqueness of solutions to Equation (16) as well as the bound

$$(19) \qquad \mathbf{E}\left[\sup_{u \in [0,T]} \|\Pi_u X\|^{2k}\right] < \infty, \quad k = 1, 2,$$

see, for instance, Theorem II.2.1 and Lemma III.1.2 in Mohammed (1984) and Theorem I.2 in Mohammed (1996). The bounds in (19) together with the sublinearity of μ and σ imply that conditions (H1) and (H2) of Theorem 1 are satisfied with $p = 2$ and $\varepsilon = 1$. As a consequence, there exists a finite constant $\tilde{K}$ such that

$$(20) \qquad \mathbf{E}\left[\mathbf{w}_X^2(h; [0,T])\right] \leq \tilde{K}\, h\, \ln(\tfrac{T}{h}) \quad \text{for all } h \in (0, T/2].$$

The above upper bound is the key point in the proof of our convergence result, see Proposition 3 below.

The approximation scheme proposed here is defined by

$$(21) \qquad (\Pi_{\lfloor t/h \rfloor \cdot h} X^n)_{t \in [0,T]},$$

where the sequence of processes $\boldsymbol{X^n} = (X^n(t))_{t \in [-\tau, T]}$ is defined according to the *piecewise linear Euler-Maruyama scheme*, that is, $\boldsymbol{X^n}$ is the piecewise linear interpolation of the Euler discretization scheme with step size $h = h_n = T/n$, where $\tau = mh$ for some $m \in \mathbb{N}$ (for the sake of simplicity, we assume that T/τ is rational):

$$(22) \qquad \begin{cases} X^n((\ell+1)h) = X^n(\ell h) + \mu(\ell h, \Pi_{\ell h} X^n)h \\ \qquad\qquad\quad + \sigma(\ell h, \Pi_{\ell h} X^n)\big[W((\ell+1)h) - W(\ell h)\big], & 0 \leq \ell \leq n-1, \\ X^n(\ell h) = \eta(\ell h), & -m \leq \ell \leq 0. \end{cases}$$

The *piecewise-constant $\mathcal{C}$-valued process* $(\Pi_{\lfloor t/h \rfloor \cdot h} X^n)_{t \in [0,T]}$ defined in (21) can be interpreted as an approximation of the $\mathcal{C}$-valued process $(\Pi_t X)_{t \in [0,T]}$.

Proposition 3. *Assume that conditions (A1) and (A2) are satisfied and that the initial condition η satisfies, for all $h \in (0, \tau/2]$,*

$$(23) \qquad\qquad \mathbf{E}\left[\mathbf{w}_\eta^2\big(h; [-\tau, 0]\big)\right] \leq C_\eta\, h\, \ln(\tfrac{\tau}{h}).$$

Then there exists a finite constant C_X such that, for all $h \in (0, T \wedge \tau/2]$,

$$(24) \qquad\qquad \mathbf{E}\left[\sup_{t\in[0,T]}\left\|\Pi_{h\cdot\lfloor t/h\rfloor}X^n - \Pi_t X\right\|^2\right] \leq C_X\, h\, \ln(\tfrac{T}{h}).$$

The above result generalizes Proposition 4.2 in Calzolari et al. (2007), since it does not require the boundedness of the coefficients μ and σ (there denoted by a and b, respectively). We point out that Proposition 3 requires condition (A1) to hold also for $k = 2$, while $k = 1$ is sufficient when $|\sigma(t, \theta)|$ is bounded by a deterministic constant.

By (20), the above result implies that a similar upper bound also holds for the expectation of $\sup_{u\in[-\tau,T]}|X^n(u) - X(u)|^2$. In this respect, Proposition 3 can be seen as a generalization to SDDEs of the result by Faure (1992) mentioned in Section 1.

Proposition 3 can be proved in much the same way as the above quoted Proposition 4.2 from Section 5 of Calzolari et al. (2007), which deals with the one-dimensional case; the boundedness assumption on the coefficients there is made only to get the upper bound (20). In the proof, which we briefly sketch below for the ease of the reader, we also consider the *continuous Euler-Maruyama scheme*, that is, the processes $\boldsymbol{Z^n} = \big(Z^n(t)\big)_{t\in[0,T]}$ defined by $Z^n(t) := \eta(t)$ for $-\tau \leq t \leq 0$ and, for $0 \leq t \leq T$,

$$dZ^n(t) := \mu(h\cdot\lfloor s/h\rfloor, \Pi_{h\cdot\lfloor s/h\rfloor}X^n)ds + \sigma(h\cdot\lfloor s/h\rfloor, \Pi_{h\cdot\lfloor s/h\rfloor}X^n)dW(s).$$

The processes $\boldsymbol{Z^n}$ can be interpreted as intermediate continuous-time approximations.

Notice that the process $\boldsymbol{X^n}$ is not adapted to the filtration of the driving Wiener process, while the processes $\boldsymbol{Z^n}$ and $(\Pi_{\lfloor t/h\rfloor\cdot h}X^n)_{t\in[0,T]}$ are both adapted.

Proof of Proposition 3 (sketch). The processes $\boldsymbol{Z^n}$ have the property that

$$Z^n(\ell h) = X^n(\ell h) \qquad \text{for all } \ell \geq -m,$$

which implies that the piecewise linear interpolation of $\boldsymbol{Z^n}$ coincides with $\boldsymbol{X^n}$. In other words,

$$(25) \qquad\qquad P^h Z^n(s) = X^n(s) \qquad \text{for all } s \in [-\tau, T],$$

where P^h denotes the operator which gives the piecewise linear interpolation with step size h of a function $f : [-\tau, T] \to \mathbb{R}$, that is,

$$P^h f(v) := \lambda(v)\, f(h\cdot\lfloor v/h\rfloor + h) + (1 - \lambda(v))\, f(h\cdot\lfloor v/h\rfloor),$$

where $\lambda(v) := v/h - \lfloor v/h \rfloor$. By rewriting $f(v) = \lambda(v)\, f(v) + (1 - \lambda(v))\, f(v)$, we get

$$
\begin{aligned}
\left| P^h f(v) - f(v) \right| &= \left| \lambda(v) \left(f(h \cdot \lfloor v/h \rfloor + h) - f(v) \right) + (1 - \lambda(v)) \left(f(h \cdot \lfloor v/h \rfloor) - f(v) \right) \right| \\
&\leq \lambda(v)\, \mathbf{w}_f(h) + (1 - \lambda(v)) \mathbf{w}_f(h) \; = \; \mathbf{w}_f(h).
\end{aligned}
$$

Furthermore, taking into account (25), we see that

$$
\begin{aligned}
\sup_{t \in [0,T]} \| \Pi_t X^n - \Pi_t P^h X \| &= \sup_{k: kh \in [-\tau, T]} |X^n(kh) - X(kh)| \;=\; \sup_{k: kh \in [-\tau, T]} |Z^n(kh) - X(kh)| \\
&\leq \sup_{t \in [-\tau, T]} |Z^n(t) - X(t)| \;=\; \sup_{t \in [0,T]} \| \Pi_t Z^n - \Pi_t X \|.
\end{aligned}
$$

Hence it holds that

$$
\begin{aligned}
&\left\| \Pi_{h \cdot \lfloor t/h \rfloor} X^n - \Pi_t X \right\|^2 \\
(26) \qquad &\leq \; 2 \left\| \Pi_{h \cdot \lfloor t/h \rfloor} X^n - \Pi_{h \cdot \lfloor t/h \rfloor} P^h X \right\|^2 \;+\; 2 \left\| \Pi_{h \cdot \lfloor t/h \rfloor} P^h X - \Pi_t X \right\|^2 \\
&\leq \; 2 \sup_{t \in [0,T]} \| \Pi_t Z^n - \Pi_t X \|^2 \;+\; 2\, \mathbf{w}_X^2(h; [-\tau, T]).
\end{aligned}
$$

The result is now a consequence of the inequality

$$
(27) \qquad \mathbf{E} \left[\sup_{u \in [0,T]} \| \Pi_u Z^n - \Pi_u X \|^2 \right] \;\leq\; C_1(T) \left(\mathbf{E} \left[\mathbf{w}_X^2(h; [-\tau, T]) \right] + h^{2\alpha} \right),
$$

(cf. (5.12) in the proof of Lemma 5.3 in Calzolari et al. (2007)), the moment bound (20), and the assumption on the moments of the modulus of continuity of η. $\qquad\square$

We end this section by observing that the rate of convergence obtained in (24) cannot be improved. Indeed, in (16), take $\mu = 0$ and $\sigma = 1$, and $\eta(s) = 0$ for all $s \in [-\tau, 0]$, i. e. the case $X(t) = W(t)$ for all $t \in [0,T]$. In this case, $Z^n(t) = W(t)$ for all $t \in [-\tau, T]$, so that the continuous Euler approximation is useless, while $W^n(t) = P^h W(t)$ for all $t \in [-\tau, T]$. Since, for any process $\boldsymbol{X}$,

$$
\begin{aligned}
\sup_{t \in [0,T]} \left\| \Pi_{h \cdot \lfloor t/h \rfloor} X^n - \Pi_t X \right\| &= \sup_{t \in [0,T]} \sup_{s \in [-\tau, 0]} \left| X^n(s + h \cdot \lfloor t/h \rfloor) - X(s + t) \right| \\
&\geq \max_{ih + \frac{h}{2} \in [0,T]} \; \max_{kh \in [-\tau, 0]} \left| X^n(kh + ih) - X(kh + ih + \tfrac{h}{2}) \right| \\
&\geq \max_{j: jh, jh + \frac{h}{2} \in [0,T]} \left| X^n(jh) - X(jh + \tfrac{h}{2}) \right|,
\end{aligned}
$$

it follows that

$$
\sup_{t \in [0,T]} \left\| \Pi_{h \cdot \lfloor t/h \rfloor} W^n - \Pi_t W \right\| \;\geq\; \max_{j: jh, jh + \frac{h}{2} \in [0,T]} \left| W(jh) - W(jh + \tfrac{h}{2}) \right| \;=\; \sqrt{\frac{h}{2}} \max_{j = 0, \ldots, n-1} |Z_j^h|,
$$

where Z_j^h are independent standard Gaussian random variables. With the same extreme value technique used in Section 2 it can be shown that the moments of

$$\max_{0 \le j \le n-1} |Z_j^h|/\sqrt{2\ln(2n)} \; = \; \max_{0 \le j \le n-1} |Z_j^h|/\sqrt{2\ln(2T/h)}$$

converge to 1. Thus, we can obtain a result of the form

$$\mathbf{E}\left[\sup_{t \in [0,T]} \left\| \Pi_{h \cdot \lfloor t/h \rfloor} W^n - \Pi_t W \right\|^p\right] \; = \; O\big((h \ln(T/h))^{p/2}\big).$$

Similar results hold for the piecewise linear approximation $\boldsymbol{W}^n$ as well as for the piecewise constant approximation $\bar{\boldsymbol{W}}^n$, i.e. $\bar{W}^n(t) := W(h\lfloor t/h \rfloor)$, $t \in [0,T]$:

$$\sup_{t \in [-\tau,T]} |W^n(t) - W(t)| \; \le \; \mathbf{w}_W^E(h; [0,T]) \; = \; \sup_{t \in [-\tau,T]} |\bar{W}^n(t) - W(t)|$$

and

$$\sup_{t \in [-\tau,T]} |W^n(t) - W(t)| \; \ge \; \max_{1 \le i \le n; i \text{ even}} \left| \tfrac{1}{2}\left[W(ih+h) - W(\tfrac{h}{2}+ih)\right] + \tfrac{1}{2}\left[W(ih) - W(\tfrac{h}{2}+ih)\right] \right|,$$

where the random variables inside the absolute value are independent Gaussian random variables with mean zero and variance $\frac{1}{4}\frac{h}{2} + \frac{1}{4}\frac{h}{2} = \frac{h}{4}$.

Appendix

Here, we derive bounds on the moments of the modulus of continuity of one-dimensional Brownian motion in an alternative way with respect to Section 2. For simplicity, we let the mesh size h be in $(0, \frac{T}{e}]$ and choose a slightly different expression for the logarithmic factor appearing in the moment bounds.

Lemma 4. *Let W be a standard one-dimensional Wiener process. Then there is a constant $K > 0$ such that for any $p > 0$, any $T > 0$,*

$$(28) \qquad \mathbf{E}\left[\big(\mathbf{w}_W(h,T)\big)^p\right] \; \le \; K^p \cdot p^{\frac{p}{2}} \cdot \big(h \ln(\tfrac{T}{h})\big)^{\frac{p}{2}} \quad \text{for all } h \in (0, \tfrac{T}{e}].$$

The approach we take in proving the lemma should be compared to the derivation of Lévy's exact modulus of continuity for Brownian motion described in Exercise 2.4.8 of Stroock and Varadhan (1979). The main ingredient is an inequality due to Garsia, Rodemich, and Rumsey Jr., see Garsia et al. (1970) and also Theorem 2.1.3 in Stroock and Varadhan (1979: p. 47). Their inequality allows to get an upper bound for $|W(t,\omega) - W(s,\omega)|^p$ in terms of T, the distance $|t-s|$ and $\xi(\omega)$, where ξ is a suitable random variable.

Proof of Lemma 4. Let $p \geq 1$. Let us first suppose that $T = 1$. Inequality (28) for $T \neq 1$ will be derived from the self-similarity of Brownian motion. In order to prepare for the application of the Garsia-Rodemich-Rumsey lemma, define on $[0, \infty)$ the strictly increasing functions Ψ and μ by

$$\Psi(x) := \exp\left(\tfrac{x^2}{2}\right) - 1, \qquad \mu(x) := \sqrt{c}\,x, \qquad x \in [0, \infty),$$

where $c > p$. Clearly,

$$\Psi(0) = 0 = \mu(0), \quad \Psi^{-1}(y) = \sqrt{2\ln(y+1)} \quad \text{for all } y \geq 0, \quad d\mu(x) = \mu(dx) = \frac{\sqrt{c}}{2\sqrt{x}}\,dx.$$

Define the $\mathcal{F}$-measurable random variable $\xi = \xi_c$ with values in $[0, \infty]$ by letting

$$(29) \qquad \xi(\omega) := \int_0^1 \int_0^1 \Psi\left(\frac{|W(t,\omega) - W(s,\omega)|}{\mu(|t-s|)}\right) ds\,dt, \quad \omega \in \Omega.$$

Notice that μ and ξ depend on the choice of the parameter c. Since $\frac{W(t)-W(s)}{\sqrt{t-s}}$ has standard normal distribution $N(0,1)$, we see that

$$\begin{aligned}
\mathbf{E}\left[\xi^p\right] &\leq \mathbf{E}\left[(\xi+1)^p\right] = \mathbf{E}\left[\left(\int_0^1 \int_0^1 \exp\left(\frac{|W(t)-W(s)|^2}{2c|t-s|}\right) ds\,dt\right)^p\right] \\
&\leq \mathbf{E}\left[\int_0^1 \int_0^1 \exp\left(\frac{|W(t)-W(s)|^2}{2c|t-s|}\right)^p ds\,dt\right] \\
&= \int_0^1 \int_0^1 \mathbf{E}\left[\exp\left(\frac{p}{2c}\left(\frac{W(t)-W(s)}{\sqrt{t-s}}\right)^2\right)\right] ds\,dt = \sqrt{\frac{c}{c-p}}.
\end{aligned}$$

In particular, $\xi(\omega) < \infty$ for $\mathbf{P}$-almost all $\omega \in \Omega$. The Garsia-Rodemich-Rumsey inequality now implies that for all $\omega \in \Omega$, all $t, s \in [0,1]$,

$$\left|W(t,\omega) - W(s,\omega)\right| \leq 8 \int_0^{|t-s|} \Psi^{-1}\left(\frac{4\xi(\omega)}{x^2}\right) \mu(dx) = 8 \int_0^{|t-s|} \sqrt{2\ln\left(\frac{4\xi(\omega)}{x^2}+1\right)}\,\frac{\sqrt{c}}{2\sqrt{x}}\,dx.$$

Notice that if $\xi(\omega) = \infty$ then the above inequality is trivially satisfied. With $h \in (0, \tfrac{1}{e}]$, we have

$$\begin{aligned}
(30) \quad \sup_{t,s\in[0,1],|t-s|\leq h}\left|W(t,\omega) - W(s,\omega)\right| &\leq 4\sqrt{2c}\int_0^h \sqrt{\ln\left(4\xi(\omega)+x^2\right) + 2\ln\left(\tfrac{1}{x}\right)}\,\frac{dx}{\sqrt{x}} \\
&\leq 4\sqrt{2c}\left(\sqrt{\ln(4\xi(\omega)+1)}\int_0^h \frac{dx}{\sqrt{x}} + \sqrt{2}\int_0^h \left(\sqrt{\ln(\tfrac{1}{x})} - \frac{1}{\sqrt{\ln(\tfrac{1}{x})}} + \frac{1}{\sqrt{\ln(\tfrac{1}{x})}}\right)\frac{dx}{\sqrt{x}}\right) \\
&\leq 8\sqrt{2c}\left(\sqrt{\ln(4\xi(\omega)+1)} + 2\sqrt{2}\right)\sqrt{h\ln(\tfrac{1}{h})} \leq 32\sqrt{c}\left(\sqrt{\xi(\omega)}+1\right)\sqrt{h\ln(\tfrac{1}{h})}.
\end{aligned}$$

Consequently, for all $h \in (0, \frac{1}{e}]$,

$$
\mathbf{E}\left[\sup_{t,s\in[0,1],|t-s|\leq h}|W(t)-W(s)|^p\right] \leq 32^p \cdot c^{\frac{p}{2}} \cdot \mathbf{E}\left[\left(\sqrt{\xi}+1\right)^p\right]\left(h\ln(\tfrac{1}{h})\right)^{\frac{p}{2}}
$$

$$
\leq 64^p \cdot c^{\frac{p}{2}} \cdot \left(\sqrt{\mathbf{E}[\xi^p]}+1\right)\left(h\ln(\tfrac{1}{h})\right)^{\frac{p}{2}}.
$$

Choosing the parameter c to be equal to $\frac{9}{8}p$, we find that for all $h \in (0, \frac{1}{e}]$,

$$
(31) \qquad \mathbf{E}\left[\left(\mathbf{w}_W(h,1)\right)^p\right] = \mathbf{E}\left[\sup_{t,s\in[0,1],|t-s|\leq h}|W(t)-W(s)|^p\right]
$$

$$
\leq (96/\sqrt{2})^p \cdot p^{\frac{p}{2}} \cdot (\sqrt{3}+1)\left(h\ln(\tfrac{1}{h})\right)^{\frac{p}{2}} < 192^p \cdot p^{\frac{p}{2}} \cdot \left(h\ln(\tfrac{1}{h})\right)^{\frac{p}{2}}.
$$

The asserted inequality thus holds for any $K \geq 192$ in case $T = 1$. To derive the assertion for arbitrary $T > 0$, recall that by letting $\tilde{W}(t) := \frac{1}{\sqrt{T}}W(T \cdot t)$, $t \geq 0$, we obtain a second standard one-dimensional Wiener process $\tilde{W}$. Therefore,

$$
\mathbf{E}\left[\left(\mathbf{w}_W(h,T)\right)^p\right] = \mathbf{E}\left[\sup_{t,s\in[0,T],|t-s|\leq h}|W(t)-W(s)|^p\right]
$$

$$
= \mathbf{E}\left[\sup_{t,s\in[0,T],|t-s|\leq h}\left|\sqrt{T}\tilde{W}\left(\tfrac{t}{T}\right)-\sqrt{T}\tilde{W}\left(\tfrac{s}{T}\right)\right|^p\right]
$$

$$
= T^{\frac{p}{2}}\,\mathbf{E}\left[\sup_{t,s\in[0,1],|t-s|\leq \frac{h}{T}}|\tilde{W}(t)-\tilde{W}(s)|^p\right] = T^{\frac{p}{2}}\,\mathbf{E}\left[\left(\mathbf{w}_{\tilde{W}}\left(\tfrac{h}{T},1\right)\right)^p\right].
$$

Since W and $\tilde{W}$ have the same distribution, estimate (31) implies that for all $h \in (0, \frac{T}{e}]$,

$$
\mathbf{E}\left[\left(\mathbf{w}_W(h,T)\right)^p\right] \leq T^{\frac{p}{2}} \cdot 192^p \cdot p^{\frac{p}{2}} \cdot \left(\tfrac{h}{T}\ln(\tfrac{T}{h})\right)^{\frac{p}{2}},
$$

which yields Inequality (28). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 3. The proof of Lemma 4 shows that the constant K need not be greater than 192. The assertion of the lemma remains valid with h from the interval $(0, \alpha \cdot T]$ for any $\alpha \in (\frac{1}{e}, 1)$, but the constant K such that Inequality (28) holds will be different.

Remark 4. From the chain of inequalities (30) in the proof of Lemma 4 it is easy to see that higher than polynomial moments of the Wiener modulus of continuity exist. More specifically, let $\xi = \xi_c$ be the random variable defined by (29), and let $\lambda > 0$. By the second but last line in (30) we have for all $h \in (0, \frac{T}{e}]$,

$$
(32) \qquad \mathbf{E}\left[\exp\left(\lambda(\mathbf{w}_W(h,T))^2\right)\right] \leq \mathbf{E}\left[\left(e \cdot (4\xi+1)\right)^{2048c\,\lambda\,h\ln(\frac{1}{h})}\right].
$$

The expectation on the right-hand side of (32) is finite if $c > 2048c\,\lambda\,h\ln(\frac{1}{h})$, that is, the above exponential-quadratic moment exists if $\lambda\,h\ln(\frac{1}{h}) < \frac{1}{2048}$. The situation here should

19

be compared to the case of standard Gaussian random variables. The constant $\frac{1}{2048}$ is, of course, not optimal.

Remark 5. The p-dependent factors in Inequalities (28) and (14) are asymptotically equivalent in the moment p up to a factor $\tilde{K}^p$ for some $\tilde{K} > 0$. This is a consequence of Stirling's formula for the Gamma function. When the p-dependent factors in Inequality (14) are bounded by an expression of the form $K^p \cdot p^{p/2}$, then K need not be greater than four for $p > 1$ big enough.

References

S. Asmussen and P. W. Glynn. *Stochastic Simulation*, volume 57 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2007.

N. Bouleau and D. Lépingle. *Numerical Methods for Stochastic Processes*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, New York, 1994.

A. Calzolari, P. Florchinger, and G. Nappo. Convergence in nonlinear filtering for stochastic delay systems. *SIAM J. Control Optim.*, 46(5):1615–1636, 2007.

O. Faure. *Simulation du mouvement brownien et des diffusions*. PhD thesis, ENPC, Paris, 1992.

P. Friz and N. Victoir. Approximations of the Brownian rough path with applications to stochastic analysis. *Ann. Inst. H. Poincaré, Probab. Statist.*, 41(4):703–724, 2005.

A. M. Garsia, E. Rodemich, and H. Rumsey, Jr. A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Math. J.*, 20(6):565–578, 1970.

B. Gnedenko. Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math. (2)*, 44:423–453, 1943.

N. Hofmann, T. Müller-Gronbach, and K. Ritter. Step size control for the uniform approximation of systems of stochastic differential equations with additive noise. *Ann. Appl. Probab.*, 10(2):616–633, 2000.

Y. Hu, S.-E. A. Mohammed, and F. Yan. Discrete-time approximations of stochastic delay equations: the Milstein scheme. *Ann. Prob.*, 32(1A):265–314, 2004.

K. Itô and McKean. *Diffusion Processes and their Sample Paths*, volume 125 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1974.

S. Kanagawa. On the rate of convergence for Maruyama's approximate solutions of stochastic differential equations. *Yokohama Math. J.*, 36(1):79–86, 1988.

I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2nd edition, 1991.

P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, corrected 3rd printing, 1999.

P. Lévy. *Théorie de l'addition des variables aléatoires*. Gauthier-Villars, Paris, 1937.

X. Mao. Numerical solutions of stochastic functional differential equations. *LMS J. Comput. Math.*, 6:141–161, 2003.

S.-E. A. Mohammed. *Stochastic Functional Differential Equations*. Pitman Publishing, London, 1984.

S.-E. A. Mohammed. Stochastic differential systems with memory: Theory, examples and applications. In L. Decreusefond, J. Gjerde, B. Øksendal, and A. Üstünel, editors, *Stochastic Analysis and Related Topics VI. Papers from the Sixth Oslo-Silivri Workshop on Stochastic Analysis held in Geilo*, pages 1–77, Boston, 1996. Birkhäuser.

T. Müller-Gronbach. The optimal uniform approximation of systems of stochastic differential equations. *Ann. Appl. Probab.*, 12(2):664–690, 2002.

R. Pettersson. Approximations for stochastic differential equations with reflecting convex boundary. *Stochastic Process. Appl.*, 59(2):295–308, 1995.

J. Pickands, III. Moment convergence of sample extremes. *Ann. Math. Stat.*, 39:881–889, 1968.

S. I. Resnick. *Extreme Values, Regular Variation, and Point Processes*, volume 4 of *Applied Probability. A Series of the Applied Probability Trust*. Springer-Verlag, New York, 1987.

K. Ritter. Approximation and optimization on the Wiener space. *J. Complexity*, 6(4): 337–364, 1990.

L. Słomiński. Euler's approximations of solutions of SDEs with reflecting boundary. *Stochastic Processes Appl.*, 94:317–337, 2001.

L. Słomiński. On approximation of solutions of multidimensional SDEs with reflecting boundary conditions. *Stochastic Process. Appl.*, 50(2):197–219, 1994.

D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes*, volume 233 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1979.

S. J. Taylor. Regularity of irregularities on a Brownian path. *Ann. Inst. Fourier*, 24(2): 195–203, 1974.

S. A. Utev. Remark on the rate of convergence in the invariance principle. (Russian). *Sibirsk. Mat. Zh.*, 22(5):206–209, 1981.