Parameter estimation for the Langevin equation with stationary-increment Gaussian noise

Please cite the original version:

Parameter Estimation for the Langevin Equation with Stationary-Increment Gaussian Noise

Tommi Sottinen[*] and Lauri Viitasaari[†]

January 4, 2017

Abstract

We study the Langevin equation with stationary-increment Gaussian noise. We show the strong consistency and the asymptotic normality with Berry–Esseen bound of the so-called second moment estimator of the mean reversion parameter. The conditions and results are stated in terms of the variance function of the noise. We consider both the case of continuous and discrete observations. As examples we consider fractional and bifractional Ornstein–Uhlenbeck processes. Finally, we discuss the maximum likelihood and the least squares estimators.

1 Introduction

We consider statistical parameter estimation for the unknown parameter $\theta > 0$ in the (generalized) Langevin equation

$$\mathrm{d}U_t^{\theta,\xi} = -\theta U_t^{\theta,\xi}\,\mathrm{d}t + \mathrm{d}G_t, \qquad t \geq 0. \tag{1.1}$$

Here the noise G is Gaussian, centered, and has stationary increments. We assume, without any loss of generality, that $G_0 = 0$. The initial condition $U_0^{\theta,\xi} = \xi$ can be any centered Gaussian random variable. We consider the so-called Second Moment Estimator (SME) and show its strong consistency

---

[*]Department of Mathematics and Statistics, University of Vaasa, P.O.Box 700, FIN-65101 Vaasa, Finland, `tommi.sottinen@iki.fi`.

[†]Department of Mathematics and System Analysis, Aalto University School of Science, Helsinki P.O. Box 11100, FIN-00076 Aalto, Finland, `lauri.viitasaari@aalto.fi`

and asymptotic normality, and provide Berry–Esseen bound for the normal approximation. The SME was called Alternative Estimator by Hu and Nualart [15] to contrast the Least Squares Estimator. We renamed it the SME to emphasize that it is based on the method of moments applied to the second empirical moment.

The Langevin equation is named thus by the pioneering work of Langevin [22]. Sometimes the solutions to the Langevin equation are called Ornstein–Uhlenbeck processes, due to the pioneering work of Ornstein and Uhlenbeck [37]. In these works the noise was the Brownian motion, and in this case the equation has been studied extensively since; see, e.g., Liptser and Shiryaev [23] and the references therein. Recently, the Langevin equation with fractional Brownian noise, i.e., the fractional Ornstein–Uhlenbeck processes, have been studied extensively in, e.g., [3, 10, 11, 19, 20, 30, 31, 35, 36], just to mention a few very recent ones.

The rest of the paper is organized as follows: In Section 2 we consider the Langevin equation is a general setting and provide some general results. Section 3 is the main section of the paper. There we introduce the SME and provide assumptions ensuring its strong consistency and asymptotic normality, or the central limit theorem. We also provide Berry–Esseen bounds for the central limit theorem, and consider the estimation based on discrete observations. In Section 4 we provide examples. We show how some recent results concerning the fractional Ornstein–Uhlenbeck processes follow in a straightforward manner from our results, and extend the previous results. We also study the bifractional Ornstein–Uhlenbeck processes of the second kind. In Section 5 we discuss Least Squares Estimators (LSE) and Maximum Likelihood Estimators (MLE). We argue that the SME is, under the ergodic hypothesis, the most general estimator one could hope for. Moreover, we argue that the LSE is not appropriate in many cases. For the MLE, we point out how it could be used in the general Gaussian setting. In Section 6 we draw some conclusions. Finally, the proofs of all the lemmas of the paper are given in Appendix A.

2 Preliminaries

2.1 General Setting

Let us first consider the Langevin equation (1.1) in a general setting, where G is simply a stochastic process, and the initial condition ξ is any random variable. The solution of (1.1) is

$$U_t^{\theta,\xi} = \mathrm{e}^{-\theta t}\xi + \int_0^t \mathrm{e}^{-\theta(t-s)}\,\mathrm{d}G_s. \tag{2.1}$$

Indeed, nothing is needed here, except the finiteness of the noise: (2.1) is the unique solution to (1.1) in the pathwise sense, and the stochastic integral in (2.1) can be defined pathwise by using the integration by parts as

$$\int_0^t \mathrm{e}^{-\theta(t-s)}\,\mathrm{d}G_s = G_t - \theta \int_0^t \mathrm{e}^{-\theta(t-s)}G_s\,\mathrm{d}s.$$

Any two solutions $U^{\theta,\xi}$ and $U^{\theta,\zeta}$ with the same noise are connected by the relation

$$U_t^{\theta,\zeta} = U_t^{\theta,\xi} + \mathrm{e}^{-\theta t}\big(\zeta - \xi\big).$$

Since our estimation is be based on the solution that starts from zero, we introduce the notation $X^\theta = U^{\theta,0}$.

For the existence of the stationary solution, the noise G must have stationary increments. Then, by extending G to the negative half-line with an independent copy running backwards in time, the stationary solution is

$$U_t^\theta = \int_{-\infty}^t \mathrm{e}^{-\theta(t-s)}\,\mathrm{d}G_s, \quad t \geq 0. \tag{2.2}$$

In other words, the stationary solution is $U^\theta = U^{\theta,\xi_{\mathrm{stat}}}$, with

$$\xi_{\mathrm{stat}} = \int_{-\infty}^0 \mathrm{e}^{\theta t}\,\mathrm{d}G_t.$$

In particular, the stationary solution exists if and only if the integral above converges (almost surely), and in this case

$$X_t^\theta = U_t^\theta - \mathrm{e}^{-\theta t}U_0^\theta. \tag{2.3}$$

Remark 2.1. By [38, Theorem 3.1] all stationary processes are the stationary solutions (2.2) of (1.1) with suitable stationary-increment noise G and parameter θ. Also, by Barndorff and Basse-O'Connor [4, Theorem 2.1] the stationary solution of (1.1) exists for all integrable stationary-increment noises.

2.2 Second Order Stationary-Increment Setting

Assume that the noise G is centered square-integrable process with stationary increments.

Remark 2.2 (Some notation). By v we denote variance of G, by r_θ the autocovariance of U^θ, and by γ_θ the covariance of X^θ. By Φ and $\bar{\Phi}$ we denote the cumulative and complementary cumulative distribution functions of the $\mathscr{N}(0,1)$-distributed variable, respectively; $\mathscr{N}(0,1)$ denotes the standard normal distribution. By C we will denote a universal constant depending only on v; C_θ and $C_{\theta,K}$, and so on, are universal constants depending additionally

3

on θ, and θ and K, and so on. In proofs, the constants may change from line to line and sometimes the dependence on the parameters are suppressed. We use the asymptotic notation $f(T) \sim g(T)$ for

$$\lim_{T \to \infty} \frac{f(T)}{g(T)} = 1.$$

The existence of the stationary covariance r_θ, given by Proposition 2.1, is ensured by the following elementary lemma.

Lemma 2.1. *Let $v \colon \mathbb{R} \to \mathbb{R}$ be a variance function of a process having stationary increments. Then, for all $t \geq 1$,*

$$v(t) \leq C \, t^2.$$

Proposition 2.1.

$$r_\theta(t) = \theta^2 e^{-\theta t} \int_{-\infty}^{t} \int_{-\infty}^{0} e^{\theta(s+u)} g(s, u) \, \mathrm{d}s \mathrm{d}u - \theta \int_{-\infty}^{0} e^{\theta s} g(t, s) \, \mathrm{d}s,$$

where

$$g(t, s) = \frac{1}{2} \Big[v(t) + v(s) - v(t - s) \Big].$$

In particular,

$$r_\theta(0) = \frac{\theta}{2} \int_{0}^{\infty} e^{-\theta t} v(t) \, \mathrm{d}t. \tag{2.4}$$

Proof. By integrating by parts, we obtain

$$r_\theta(t) = \mathbf{E} \left[- \int_{-\infty}^{0} \theta e^{\theta s} G_t G_s \, \mathrm{d}s + \int_{-\infty}^{t} \int_{-\infty}^{0} \theta^2 e^{-\theta(t-s-u)} G_s G_u \, \mathrm{d}s \mathrm{d}u \right].$$

The claim follows from this by the Fubini's theorem, if the integrals above converge. To this end, it is necessary and sufficient that $r_\theta(0)$ is finite. Now,

$$
\begin{aligned}
r_\theta(0) &= \frac{\theta^2}{2} \int_{-\infty}^{0} \int_{-\infty}^{0} e^{\theta t} e^{\theta s} \Big[v(t) + v(s) - v(t - s) \Big] \, \mathrm{d}t \mathrm{d}s \\
&= \theta \int_{-\infty}^{0} e^{\theta t} v(t) \, \mathrm{d}t - \frac{\theta^2}{2} \int_{-\infty}^{0} e^{\theta t} \left[\int_{-\infty}^{-t} e^{\theta(t+s)} v(s) \, \mathrm{d}s \right] \mathrm{d}t.
\end{aligned}
$$

For the latter term we have

$$\frac{\theta^2}{2} \int_{-\infty}^{0} e^{\theta t} \left[\int_{-\infty}^{-t} e^{\theta(t+s)} \, ds \right] dt$$

$$= \frac{\theta^2}{2} \int_{-\infty}^{\infty} v(s) e^{\theta s} \left[\int_{-\infty}^{\min(-s,0)} e^{2\theta t} \, dt \right] ds$$

$$= \frac{\theta^2}{2} \left[\int_{-\infty}^{0} v(s) e^{\theta s} \left[\int_{-\infty}^{0} e^{2\theta t} \, dt \right] ds + \int_{0}^{\infty} v(s) e^{\theta s} \left[\int_{-\infty}^{-t} e^{2\theta t} \, dt \right] ds \right]$$

$$= \frac{\theta}{4} \left[\int_{-\infty}^{0} v(s) e^{\theta s} \, ds + \int_{0}^{\infty} v(s) e^{\theta s} e^{-2\theta s} \, ds \right]$$

$$= \frac{\theta}{2} \int_{0}^{\infty} e^{-\theta s} v(s) \, ds.$$

Consequently, we have shown (2.4). Since, by Lemma 2.1, $v(t) \le Ct^2$, the finiteness of r_θ follows from the representation above. $\qquad\square$

Proposition 2.2.

$$\gamma_\theta(t, s) = r_\theta(t - s) + e^{-\theta(t+s)} r_\theta(0) - e^{-\theta t} r_\theta(s) - e^{-\theta s} r_\theta(t).$$

In particular,

$$k_\theta(t, s) = \left| \gamma_\theta(t, s) - r_\theta(t - s) \right| \le C_\theta e^{-\theta \min(t,s)}.$$

Proof. The formula for γ_θ is immediate from (2.3). As for the estimate, note that $|r_\theta(t)| \le r_\theta(0)$ by the Cauchy–Schwarz inequality. Consequently, assuming $s \le t$,

$$k_\theta(t, s) \le e^{-\theta(t+s)} r_\theta(0) + e^{-\theta t} r_\theta(0) + e^{-\theta s} r_\theta(0)$$

$$= r_\theta(0) \left[e^{-\theta t} + e^{-\theta(t-s)} + 1 \right] e^{-\theta s},$$

from which the estimate follows. $\qquad\square$

2.3   Gaussian Setting

Assume that the stationary-increment noise G in the Langevin equation (1.1) is centered, continuous and Gaussian with $G_0 = 0$. Then the continuous stationary Gaussian solution can be characterized by its autocovariance function r_θ given by Proposition 2.1.

Remark 2.3 (Continuity)**.** In the Gaussian realm the assumption that G is continuous is essential. Indeed, if G were discontinuous at any point, then U^θ would be discontinuous at every point, and also unbounded on every interval by the Belyaev's alternative [5]. Parameter estimation for such a U^θ would be a fools errand, indeed.

3 Second Moment Estimator

For the SME of Definition 3.1 below to be well-defined we need the invertibility of $\psi(\theta) = r_\theta(0)$, which is ensured by the following assumption:

Assumption 3.1 (Invertibility). v is strictly increasing.

Lemma 3.1 (Invertibility). *Suppose Assumption 3.1 holds. Then $\psi\colon \mathbb{R}_+ \to (0, \psi(0+))$ is strictly decreasing infinitely differentiable convex bijection.*

Definition 3.1 (SME). The second moment estimator is

$$\tilde{\theta}_T = \psi^{-1}\left(\frac{1}{T}\int_0^T (X_t^\theta)^2 \, \mathrm{d}t\right),$$

where

$$\psi(\theta) = \frac{\theta}{2}\int_0^\infty \mathrm{e}^{-\theta t} v(t) \, \mathrm{d}t$$

is the variance of the stationary solution.

Remark 3.1. The idea of the SME is to use the ergodicity of the stationary solution directly. Therefore, it would have been more natural to base it on the stationary solution U^θ instead of the zero-initial solution X^θ. However, from the practical point of view, using the solution X^θ makes more sense, since it does not assume that the Ornstein–Uhlenbeck process has reached its stationary state. Moreover, the use of the zero-initial solution instead of the stationary solution makes no difference (except when bias is concerned; see Remark 3.4). Indeed, by virtue of Proposition 3.1 below, we could have used any solution $U^{\theta,\xi}$ with any initial condition ξ.

Proposition 3.1. *Suppose the stationary solution U^θ is ergodic. Then, for all initial distributions ξ*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T (U_t^{\theta,\xi})^2 \, \mathrm{d}t = \psi(\theta) \quad a.s.$$

Proof. Let us write

$$\frac{1}{T}\int_0^T (U_t^{\theta,\xi})^2 \, \mathrm{d}t$$

$$= \frac{1}{T}\int_0^T (U_t^\theta + \mathrm{e}^{-\theta t}(\xi - U_0^\theta))^2 \, \mathrm{d}t$$

$$= \frac{1}{T}\int_0^T (U_t^\theta)^2 \, \mathrm{d}t + \frac{2(\xi - U_0^\theta)}{T}\int_0^T \mathrm{e}^{-\theta t} U_t^\theta \, \mathrm{d}t + \frac{(\xi - U_0^\theta)^2}{T}\int_0^T \mathrm{e}^{-2\theta t} \, \mathrm{d}t.$$

By ergodicity, the first term converges to $\psi(\theta)$ almost surely. Also, it is clear that the third term converges to zero almost surely. As for the second term, note that U^θ is ergodic and centered, which implies that

$$\frac{1}{T} \int_0^T U_t^\theta \, \mathrm{d}t \to 0 \quad a.s..$$

Consequently, the second term converges to zero almost surely. $\qquad \square$

3.1 Strong Consistency

The strong consistency of the SME will follow directly from the ergodicity. For Gaussian processes, the necessary and sufficient conditions for ergodicity are well known and date back to Grenander [13] and Maruyama [24]. We use the following characterization for ergodicity:

Assumption 3.2 (Ergodicity)**.** The autocovariance r_θ satisfies

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T |r_\theta(t)| \, \mathrm{d}t = 0.$$

Remark 3.2 (Gaussian Ergodicity)**.** In addition to Assumption 3.2, other well-known equivalent characterizations for the ergodicity in the Gaussian realm are

(i)
$$\lim_{T \to \infty} \frac{1}{T} \int_0^T r_\theta(t)^2 \, \mathrm{d}t = 0.$$

(ii) The spectral measure μ_θ defined by the Bochner's theorem

$$r_\theta(t) = \int_{-\infty}^\infty \mathrm{e}^{-\mathrm{i}\lambda t} \, \mu_\theta(\mathrm{d}\lambda)$$

has no atoms.

Theorem 3.1 (Strong Consistency)**.** *Suppose Assumption 3.2 and Assumption 3.1 hold. Then*
$$\tilde{\theta}_T \to \theta$$
almost surely as $T \to \infty$.

Proof. By Assumption 3.2, the stationary solution U^θ is ergodic. Consequently, by Proposition 3.1

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T (X_t^\theta)^2 \, \mathrm{d}t = \psi(\theta).$$

Since, by Lemma 3.1, ψ is a continuous bijection, the claim follows from the continuous mapping theorem. $\qquad \square$

Remark 3.3 (Gaussian assumption)**.** The assumption of Gaussianity is not needed in construction of the SME in Definition 3.1. Also, the strong consistency result of Theorem 3.1 does not rely on Gaussianity. However, Assumption 3.2 expresses ergodicity in terms of the autocovariance function r_θ and this is essentially a Gaussian characterization. Theorem 3.1 will remain true for any square-integrable continuous stationary-increment centered noise once Assumption 3.2 is replaced by a suitable assumption that ensures the ergodicity of the stationary solution. On the contrary, later the proof of Theorem 3.2 concerning the asymptotic normality of the SME relies heavily on the assumption of Gaussianity, and cannot be generalized in any straightforward manner to non-Gaussian noises.

Remark 3.4 (Bias)**.** Unbiasedness is a fragile property, as it is not preserved in non-linear transformations. Thus, it is not surprising that the SME is biased. Indeed, suppose we use the stationary solution U^θ instead of X^θ in the SME. Let us call this Stationary Second Moment Estimator (SSME), and denote it by $\ddot{\theta}_T$. Then

$$\mathbf{E}[\psi(\ddot{\theta}_T)] = \frac{1}{T} \int_0^T \mathbf{E}[(U_t^\theta)^2] \, \mathrm{d}t = \frac{1}{T} \int_0^T \psi(\theta) \, \mathrm{d}t = \psi(\theta).$$

So, the SSME is unbiased for $\psi(\theta)$. However, ψ is strictly convex, with makes ψ^{-1} strictly concave. Consequently, $\mathbf{E}[\ddot{\theta}_T] < \theta$. For the estimation based on the zero-initial solution X^θ, even $\psi(\tilde{\theta}_T)$ is biased, but asymptotically unbiased. Indeed, straightforward calculation shows that

$$\mathbf{E}[\psi(\tilde{\theta}_T)] = \psi(\theta) + \frac{1}{T} \left[\int_0^T \mathrm{e}^{-\theta t} r_\theta(t) \, \mathrm{d}t + \frac{\psi(\theta)}{2\theta} \left[1 - \mathrm{e}^{-2\theta T} \right] \right].$$

In principle, since the distribution of $\tilde{\theta}_T$ and the function ψ are known, it is possible to construct an unbiased second moment estimator. However, the formula would be very complicated and, moreover, it would depend on the unknown parameter θ.

3.2   Asymptotic Normality

It turns out that the rate of convergence and the corresponding Berry–Esseen bound for the SME are given by

$$
\begin{aligned}
w_\theta(T) &= \frac{2}{T^2} \int_0^T \!\! \int_0^T r_\theta(t-s)^2 \, \mathrm{d}s \mathrm{d}t, \\
R_\theta(T) &= \frac{\int_0^T |r_\theta(t)| \, \mathrm{d}t}{T \sqrt{w_\theta(T)}}.
\end{aligned}
$$

This leads to the following assumption for the asymptotic normality:

Assumption 3.3 (Normality). $R_\theta(T) \to 0$ as $T \to \infty$.

Our main result, Theorem 3.2 below, shows that the SME satisfies asymptotic normality with asymptotic variance $w_\theta(T)/\psi'(\theta)^2$ and the Berry–Esseen bound for the normal approximation is governed by $R_\theta(T)$.

Theorem 3.2 (Asymptotic Normality with Berry–Esseen Bound). *Suppose Assumption 3.2 and Assumption 3.1 hold. Then there exists a constant C_θ such that*

$$\sup_{x \in \mathbb{R}} \left| \mathbf{P}\left[\frac{|\psi'(\theta)|}{\sqrt{w_\theta(T)}} \left(\tilde{\theta}_T - \theta \right) \le x \right] - \Phi(x) \right| \le C_\theta R_\theta(T).$$

In particular, if Assumption 3.3 holds, then

$$\frac{|\psi'(\theta)|}{\sqrt{w_\theta(T)}} \left(\tilde{\theta}_T - \theta \right) \xrightarrow{\mathrm{d}} \mathcal{N}(0,1).$$

The proof of Theorem 3.2 uses the fourth moment Berry–Esseen bound due to Peccati and Taqqu [28, Theorem 11.4.3] that is stated below as Proposition 3.2. The setting of Proposition 3.2 is as follows: Let $W = (W_t)_{t \in \mathbb{R}_+}$ be the Brownian motion, and let $\mathbf{P}_W$ be its distribution on $L^2(\mathbb{R}_+)$. The qth Wiener chaos is the closed linear subspace of $L^2(\Omega, \mathscr{F}_W, \mathbf{P}_W)$ generated by the random variables $H_q(\xi)$, where H_q is the qth Hermite polynomial

$$H_q(x) = \frac{(-1)^q}{q!} \mathrm{e}^{x^2/2} \frac{\mathrm{d}^q}{\mathrm{d}x^q} \left[\mathrm{e}^{-x^2/2} \right],$$

and $\xi = \int_0^\infty f(t) \, \mathrm{d}W_t$ for some $f \in L^2(\mathbb{R}_+)$.

Proposition 3.2 (Fourth Moment Berry–Esseen Bound). *Let F belong to the qth Wiener chaos with some $q \ge 2$. Suppose $\mathbf{E}[F^2] = 1$. Then*

$$\sup_{x \in \mathbb{R}} \left| \mathbf{P}[F \le x] - \Phi(x) \right| \le 2\sqrt{\frac{q-1}{3q}} \sqrt{\mathbf{E}[F^4] - 3}.$$

The following series of elementary lemmas deal with Gaussian processes in general, not the Gaussian solutions to the Langevin equation in particular. To emphasize this, we drop the parameter θ in the notation. In this general setting, $X = (X_t)_{t \in \mathbb{R}_+}$ is a centered Gaussian process with continuous covariance function $\gamma \colon \mathbb{R}_+^2 \to \mathbb{R}$ and

$$Q_T = \frac{1}{T} \int_0^T \left[X_t^2 - \mathbf{E}[X_t^2] \right] \mathrm{d}t$$

Lemma 3.2. Q_T *belongs to the 2nd Wiener chaos.*

Lemma 3.3.

$$\mathbf{E}\left[Q_T^2\right] \;=\; \frac{2}{T^2} \int_{[0,T]^2} \gamma(t_1, t_2)^2 \, \mathrm{d}t_1 \mathrm{d}t_2$$

$$\mathbf{E}\left[Q_T^4\right] \;=\; 12 \left[\frac{1}{T^2} \int_{[0,T]^2} \gamma(t_1, t_2)^2 \, \mathrm{d}t_1 \mathrm{d}t_2\right]^2$$

$$+ \frac{24}{T^4} \int_{[0,T]^4} \gamma(t_1, t_2)\gamma(t_2, t_3)\gamma(t_3, t_4)\gamma(t_4, t_1) \, \mathrm{d}t_1 \mathrm{d}t_2 \mathrm{d}t_3 \mathrm{d}t_4.$$

Lemma 3.4. *All bounded covariance functions γ satisfy*

$$\int_{[0,T]^4} \gamma(t_1, t_2)\gamma(t_2, t_3)\gamma(t_3, t_4)\gamma(t_4, t_1) \, \mathrm{d}t_1 \mathrm{d}t_2 \mathrm{d}t_3 \mathrm{d}t_4$$

$$\leq \left[\sup_{t \in [0,T]} \int_0^T |\gamma(t, t_1)| \, \mathrm{d}t_1\right]^2 \int_{[0,T]^2} \gamma(t_1, t_2)^2 \, \mathrm{d}t_1 \mathrm{d}t_2.$$

Lemma 3.5. *There exists a constant C such that*

$$\sup_{x \in \mathbb{R}} \left| \mathbf{P}\left[\frac{Q_T}{\sqrt{\mathbf{E}[Q_T^2]}} \leq x\right] - \Phi(x) \right| \;\leq\; C \frac{\sup_{t \in [0,T]} \int_0^T |\gamma(t, s)| \, \mathrm{d}s}{\sqrt{\int_0^T \int_0^T \gamma(t, s)^2 \, \mathrm{d}t \mathrm{d}s}}.$$

Let us then turn back to the special case of the Langevin equation. To this end, we decompose

$$\frac{1}{T} \int_0^T (X_t^\theta)^2 \, \mathrm{d}t - \psi(\theta) = Q_T^\theta + \varepsilon_\theta(T),$$

where

$$Q_T^\theta \;=\; \frac{1}{T} \int_0^T \left[(X_t^\theta)^2 - \mathbf{E}[(X_t^\theta)^2]\right] \mathrm{d}t,$$

$$\varepsilon_\theta(T) \;=\; \frac{1}{T} \int_0^T \left[\mathbf{E}[(X_t^\theta)^2] - \psi(\theta)\right] \mathrm{d}t.$$

Now, the quadratic functional Q_T^θ belongs to the 2nd Wiener chaos, and the idea is to show that Q_T^θ converges to a Gaussian limit with asymptotic variance $w_\theta(T)/\psi'(\theta)^2$ and the associated Berry–Esseen bound $C_\theta R_\theta(T)$, while the remainder $\varepsilon_\theta(T)$ is negligible.

Lemma 3.6 (Equivalence of Variance)**.** *In general,*

$$\mathbf{E}[(Q_T^\theta)^2] \sim w_\theta(T) = \frac{4}{T} \int_0^T r_\theta(t)^2 (T - t) \, \mathrm{d}t.$$

In particular, if $\int_0^\infty r_\theta(t)^2 \, \mathrm{d}t < \infty$, we obtain the rate

$$\mathbf{E}[(Q_T^\theta)^2] \sim \frac{4 \int_0^\infty r_\theta(t)^2 \, \mathrm{d}t}{T}.$$

Lemma 3.7 (Berry–Esseen Bound). *There exists a constant C_θ such that*

$$\sup_{x \in \mathbb{R}} \left| \mathbf{P}\left[\frac{Q_T^\theta}{\sqrt{w_\theta(T)}} \leq x \right] - \Phi(x) \right| \leq C_\theta R_\theta(T).$$

Proof of Theorem 3.2. Suppose first that

$$x \leq -\frac{|\psi'(\theta)|\theta}{\sqrt{w_\theta(T)}}.$$

Since $\tilde{\theta}_T > 0$ almost surely, we then have

$$\mathbf{P}\left[\frac{|\psi'(\theta)|}{\sqrt{w_\theta(T)}} \left(\tilde{\theta}_T - \theta \right) \leq x \right] = 0$$

and a standard estimate for the tail of a normal random variable yields

$$\Phi\left(x \right) \leq \frac{C}{x} \leq C_\theta \sqrt{w_\theta(T)}.$$

Suppose then that

$$x > -\frac{|\psi'(\theta)|\theta}{\sqrt{w_\theta(T)}}.$$

Since ψ is strictly decreasing and continuous, we have

$$\mathbf{P}\left[\frac{|\psi'(\theta)|}{\sqrt{w_\theta(T)}} \left(\tilde{\theta}_T - \theta \right) \leq x \right]$$

$$= \mathbf{P}\left[\tilde{\theta}_T \leq \frac{\sqrt{w_\theta(T)}}{|\psi'(\theta)|} x + \theta \right]$$

$$= \mathbf{P}\left[\psi(\tilde{\theta}_T) \geq \psi\left(\frac{\sqrt{w_\theta(T)}}{|\psi'(\theta)|} x + \theta \right) \right]$$

$$= \mathbf{P}\left[\psi(\tilde{\theta}_T) - \psi(\theta) \geq \psi\left(\frac{\sqrt{w_\theta(T)}}{|\psi'(\theta)|} x + \theta \right) - \psi(\theta) \right]$$

$$= \mathbf{P}\left[Q_T^\theta + \varepsilon_\theta(T) \geq \psi\left(\frac{\sqrt{w_\theta(T)}}{|\psi'(\theta)|} x + \theta \right) - \psi(\theta) \right]$$

Let us then introduce the short-hand notation

$$\nu = \frac{\psi\left(\frac{\sqrt{w_\theta(T)}}{|\psi'(\theta)|} x + \theta \right) - \psi(\theta)}{\sqrt{w_\theta(T)}}.$$

11

By using the calculation and the short-hand notation above, we split

$$\left| \mathbf{P}\left[\frac{|\psi'(\theta)|}{\sqrt{w_\theta(T)}} \left(\tilde{\theta}_T - \theta \right) \leq x \right] - \Phi(x) \right|$$

$$= \left| \mathbf{P}\left[\frac{Q_T^\theta + \varepsilon_\theta(T)}{\sqrt{w_\theta(T)}} \geq \nu \right] - \Phi(x) \right|$$

$$\leq \left| \mathbf{P}\left[\frac{Q_T^\theta + \varepsilon_\theta(T)}{\sqrt{w_\theta(T)}} \geq \nu \right] - \bar{\Phi}(\nu) \right| + \left| \bar{\Phi}(\nu) - \Phi(x) \right|$$

$$= A_1 + A_2.$$

For the term A_1, we split again

$$A_1 = \left| \mathbf{P}\left[\frac{Q_T^\theta + \varepsilon_\theta(T)}{\sqrt{w_\theta(T)}} \geq \nu \right] - \bar{\Phi}(\nu) \right|$$

$$\leq \left| \mathbf{P}\left[\frac{Q_T^\theta}{\sqrt{w_\theta(T)}} \geq \nu - \frac{\varepsilon_\theta(T)}{\sqrt{w_\theta(T)}} \right] - \bar{\Phi}\left(\nu - \frac{\varepsilon_\theta(T)}{\sqrt{w_\theta(T)}} \right) \right|$$

$$+ \left| \bar{\Phi}\left(\nu - \frac{\varepsilon_\theta(T)}{\sqrt{w_\theta(T)}} \right) - \bar{\Phi}(\nu) \right|$$

$$= A_{1,1} + A_{1,2}.$$

By the Berry–Esseen bound of Lemma 3.7, $A_{1,1} \leq C_\theta R_\theta(T)$. Consider then $A_{1,2}$. Since $|\bar{\Phi}(x) - \bar{\Phi}(y)| \leq |x - y|$, we have

$$A_{1,2} \leq \frac{\varepsilon_\theta(T)}{\sqrt{w_\theta(T)}}.$$

By the Cauchy–Schwarz inequality $|r_\theta(t)| \leq r_\theta(0) = \psi(\theta)$. Consequently,

$$\varepsilon_\theta(T) = \psi(\theta)\frac{1}{T} \int_0^T e^{2\theta t}\,dt - \frac{2}{T} \int_0^T e^{-\theta t} r_\theta(t)\,dt$$

$$\leq \psi(\theta)\frac{1}{T} \int_0^T \left[e^{-2\theta t} + e^{-\theta t} \right]\,dt$$

$$\leq \frac{C_\theta}{T}.$$

Therefore,

$$A_{1,2} \leq \frac{C_\theta/T}{\sqrt{1/T^2 \int_0^T \int_0^T r_\theta(t-s)^2\,ds dt}}$$

$$= \frac{C_\theta}{\sqrt{\int_0^T \int_0^T r_\theta(t-s)^2\,ds dt}}$$

$$\leq C_\theta \frac{\int_0^T |r_\theta(t)|\,dt}{\sqrt{\int_0^T \int_0^T r_\theta(t-s)^2\,ds dt}},$$

12

where the last inequality follows from the fact that $r_\theta(0) > 0$ and we can assume that T is greater than some absolute constant.

Finally, it remains to consider the term A_2. For this, recall that ψ is smooth. Therefore, by the mean value theorem, there exists some number $\eta \in [\theta, \theta + \frac{\sqrt{w_\theta(T)}}{|\psi'(\theta)|}x]$ such that

$$
\begin{aligned}
\nu &= \frac{1}{\sqrt{w_\theta(T)}} \left[\psi\left(\frac{\sqrt{w_\theta(T)}}{|\psi'(\theta)|}x + \theta \right) - \psi(\theta) \right] \\
&= \frac{1}{\sqrt{w_\theta(T)}} \psi'(\eta) \left[\frac{\sqrt{w_\theta(T)}}{|\psi'(\theta)|}x \right] \\
&= \frac{\psi'(\eta)}{|\psi'(\theta)|}x.
\end{aligned}
$$

Furthermore, since ψ is decreasing, we have

$$
\frac{\psi'(\eta)}{|\psi'(\theta)|} = -\frac{\psi'(\eta)}{\psi'(\theta)}.
$$

Consequently,

$$
\bar{\Phi}(\nu) = \Phi\left(\frac{\psi'(\eta)}{\psi'(\theta)}x \right).
$$

Note also that, since ψ is convex, for any x we have

$$
x \le \frac{\psi'(\eta)}{\psi'(\theta)}x.
$$

Then

$$
\begin{aligned}
A_2 &= \left| \bar{\Phi}(\nu) - \Phi(x) \right| \\
&= \left| \Phi\left(\frac{\psi'(\eta)}{\psi'(\theta)}x \right) - \Phi(x) \right| \\
&= \frac{1}{\sqrt{2\pi}} \int_x^{\frac{\psi'(\eta)}{\psi'(\theta)}x} e^{-\frac{y^2}{2}}\,dy.
\end{aligned}
$$

Suppose then first that

$$
-\frac{|\psi'(\theta)|\theta}{\sqrt{w_\theta(T)}} < x < -\frac{|\psi'(\theta)|\theta}{2\sqrt{w_\theta(T)}}. \tag{3.1}
$$

13

Then

$$A_2 \leq \int_x^{\frac{\psi'(\eta)}{\psi'(\theta)}x} e^{-\frac{y^2}{2}} \, dy$$

$$\leq \frac{|x|}{|\psi'(\theta)|} \left|\psi'(\eta) - \psi'(\theta)\right| e^{-\frac{|\psi'(\eta)|^2 x^2}{2|\psi'(\theta)|^2}}$$

$$\leq \frac{C_\theta}{|x|}$$

$$\leq C_\theta \sqrt{w_\theta(T)}$$

where the last inequalities follows from (3.1) together with the fact that a function $f(x,y) = x^2 e^{-\frac{z^2 x^2}{2}}|z - 1|$ is uniformly bounded. Finally, let

$$x > -\frac{|\psi'(\theta)|\theta}{2\sqrt{w_\theta(T)}}. \tag{3.2}$$

By the proof of Lemma 3.1 we have

$$\psi''(\theta) = \frac{1}{2} \int_0^\infty e^{-\theta s} s^2 \, dv(s),$$

and hence $\psi''\left(\theta + \frac{\sqrt{w_\theta(T)}}{|\psi'(\theta)|}x\right)$ is uniformly bounded for any x satisfying (3.2). By using the change of variable $y = x^2 z$ together with the fact that $f_2(x,y) = x^2 e^{-\frac{x^4 z^2}{2}}$ is also uniformly bounded, we observe

$$A_2 \leq \int_x^{\frac{\psi'(\eta)}{\psi'(\theta)}x} e^{-\frac{y^2}{2}} \, dy$$

$$= \int_{\frac{1}{x}}^{\frac{\psi'(\eta)}{x\psi'(\theta)}} x^2 e^{-\frac{x^4 z^2}{2}} \, dz$$

$$\leq C_\theta \frac{1}{|x\psi'(\theta)|} \left|\psi'(\eta) - \psi'(\theta)\right|.$$

By using the mean value theorem again, we find some $\tilde{\eta} \in [\theta, \eta]$ such that

$$\frac{1}{|x|}\left|\psi'(\eta) - \psi'(\theta)\right| = |\psi''(\tilde{\eta})|\frac{\sqrt{w_\theta(T)}}{|\psi'(\theta)|} \leq C_\theta \sqrt{w_\theta(T)}$$

by the fact that $\psi''(\tilde{\eta})$ is bounded. Therefore, it remains to show that

$$\sqrt{w_\theta(T)} \leq C_\theta R_\theta(T),$$

which translates into showing that

$$\frac{2}{T} \int_0^T \int_0^t r_\theta(t-s)^2 \, ds dt \leq C_\theta \int_0^T |r_\theta(t)| \, dt.$$

Since $r_\theta(t)^2 \leq \psi(\theta)|r_\theta(t)|$, the inequality above follows by applying l'Hôpital's rule to it. This finishes the proof of Theorem 3.2 $\qquad\qquad\square$

Next we state some corollaries that make Theorem 3.2 somewhat easier to use in applications. Corollary 3.1 deals with the classical $\sqrt{T}$ rate of convergence and Corollary 3.2 deals with mixed models.

Corollary 3.1 (Classical Rate). *Suppose Assumption 3.1 holds. Assume* $\int_0^\infty r_\theta(t)^2 \, dt < \infty$. *Denote*

$$\sigma^2(\theta) = 4\frac{\int_0^\infty r_\theta(t)^2 \, dt}{|\psi'(\theta)|}.$$

Then there exists a constant C_θ such that

$$\sup_{x \in \mathbb{R}} \left| \mathbf{P}\left[\frac{\sqrt{T}}{\sigma(\theta)} \left(\tilde{\theta}_T - \theta \right) \leq x \right] - \Phi(x) \right|$$

$$\leq C_\theta \left(\frac{1}{\sqrt{T}} \int_0^T |r_\theta(t)| \, dt + \sqrt{\int_T^\infty r_\theta(t)^2 \, dt} + \sqrt{\frac{1}{T} \int_0^T r_\theta(t)^2 t \, dt} \right).$$

Proof. First note that Assumption 3.2 is implied by the assumption that $\int_0^\infty r_\theta(t)^2 \, dt < \infty$. Then, let us split

$$\left| \mathbf{P}\left[\frac{\sqrt{T}(\tilde{\theta}_T - \theta)}{\sigma(\theta)} \leq x \right] - \Phi(x) \right|$$

$$\leq \left| \mathbf{P}\left[\frac{|\psi'(\theta)|(\tilde{\theta}_T - \theta)}{\sqrt{w_\theta(T)}} \leq \frac{x 2\sqrt{\int_0^\infty r_\theta(t)^2 \, dt}}{\sqrt{T w_\theta(T)}} \right] - \Phi\left(\frac{x 2\sqrt{\int_0^\infty r_\theta(t)^2 \, dt}}{\sqrt{T w_\theta(T)}} \right) \right|$$

$$+ \left| \Phi\left(\frac{x 2\sqrt{\int_0^\infty r_\theta(t)^2 \, dt}}{\sqrt{T w_\theta(T)}} \right) - \Phi(x) \right|$$

$$= A_1 + A_2.$$

Now

$$T w_T(\theta) \sim 4 \int_0^T r_\theta(t)^2 \, dt \sim 4 \int_0^\infty r_\theta(t)^2 \, dt,$$

i.e., $T w_T(\theta)$ is asymptotically a positive constant. Consequently, we can take the supremum over x on a compact interval, and Theorem 3.2 implies that the term A_1 is dominated by

$$C_\theta R_\theta(T) \leq C_\theta \frac{\int_0^T |r_\theta(t)| \, dt}{\sqrt{T \int_0^T r_\theta(t)^2 \, dt}} \leq C_\theta \frac{1}{\sqrt{T}} \int_0^T |r_\theta(t)| \, dt.$$

15

For the second term, we use the estimate

$$\sup_{x\in\mathbb{R}} |\Phi(\rho x) - \Phi(x)| \le |\rho - 1|.$$

Setting

$$\rho = \frac{2\sqrt{\int_0^\infty r_\theta(t)^2\,\mathrm{d}t}}{\sqrt{Tw_\theta(T)}} = \frac{\sqrt{\int_0^\infty r_\theta(t)^2\,\mathrm{d}t}}{\sqrt{\frac{1}{T}\int_0^T\int_0^t r_\theta(s)^2\,\mathrm{d}s\mathrm{d}t}},$$

we obtain for the term A_2 the upper bound

$$\left| \frac{\sqrt{\int_0^\infty r_\theta(t)^2\,\mathrm{d}t} - \sqrt{\frac{1}{T}\int_0^T\int_0^t r_\theta(s)^2\,\mathrm{d}s\mathrm{d}t}}{\sqrt{\frac{1}{T}\int_0^T\int_0^t r_\theta(s)^2\,\mathrm{d}s\mathrm{d}t}} \right|$$

$$\le C_\theta \sqrt{\left| \int_0^\infty r_\theta(t)^2\,\mathrm{d}t - \frac{1}{T}\int_0^T\int_0^t r_\theta(s)^2\,\mathrm{d}s\mathrm{d}t \right|}$$

$$= C_\theta \sqrt{\left| \int_0^\infty r_\theta(t)^2\,\mathrm{d}t - \frac{1}{T}\int_0^T r_\theta(t)^2(T-t)\,\mathrm{d}t \right|}$$

$$= C_\theta \sqrt{\left| \int_T^\infty r_\theta(t)^2\,\mathrm{d}t + \frac{1}{T}\int_0^T r_\theta(t)^2 t\,\mathrm{d}t \right|}$$

$$\le C_\theta \left(\sqrt{\int_T^\infty r_\theta(t)^2\,\mathrm{d}t} + \sqrt{\frac{1}{T}\int_0^T r_\theta(t)^2 t\,\mathrm{d}t} \right),$$

since $|\sqrt{a} - \sqrt{b}| \le \sqrt{|a - b|}$ and $\sqrt{a + b} \le \sqrt{a} + \sqrt{b}$. $\qquad\square$

Corollary 3.2 (Mixed Models)**.** *Let G^i, $i = 1, \ldots, n$, be independent continuous stationary-increment Gaussian processes with zero mean each satisfying Assumption 3.2 and Assumption 3.1. Let $r_{\theta,i}$ be the autocovariance of the stationary solution corresponding the noise G^i. Assume that $r_{\theta,i} \ge 0$ for all i. Then, for the noise $G = \sum_{i=1}^n G^i$, there exists a constant C_θ such that*

$$\sup_{x\in\mathbb{R}} \left| \mathbf{P}\left[\frac{|\psi'(\theta)|}{\sqrt{w_\theta(T)}} \left(\tilde{\theta}_T - \theta \right) \le x \right] - \Phi(x) \right|$$

$$\le C_\theta \max_{i=1,\ldots,n} \frac{n\int_0^T r_{\theta,i}(t)\,\mathrm{d}t}{\sqrt{\int_0^T\int_0^T r_{\theta,i}(t-s)^2\,\mathrm{d}s\mathrm{d}t}}.$$

Proof. Since the G^i's are independent, the autocovariance for the mixed model with noise G is $r_\theta = \sum_{i=1}^n r_{\theta,i}$. Consequently, Assumption 3.2 and Assumption 3.1 hold. It remains to show that

$$\frac{\int_0^T \sum_{i=1}^n r_{\theta,i}(t)\,\mathrm{d}t}{\sqrt{\int_0^T\int_0^T \left(\sum_{i=1}^n r_{\theta,i}(t-s)\right)^2\,\mathrm{d}s\mathrm{d}t}} \le \frac{\max_{i=1,\ldots,n} n\int_0^T r_{\theta,i}(t)\,\mathrm{d}t}{\sqrt{\int_0^T\int_0^T r_{\theta,j}(t-s)^2\,\mathrm{d}s\mathrm{d}t}}$$

16

for any $j = 1, \ldots, n$. The case for the nominator is clear. For the denominator, we use the fact that the $r_{\theta,j}$'s are non-negative. Indeed, then

$$\left(\sum_{i=1}^{n} r_{\theta,i}(t-s) \right)^2 \geq r_{\theta,j}(t-s)^2$$

for any j, as the cross-terms $r_{\theta,i}(t-s)r_{\theta,k}(t-s)$ are positive. $\square$

We end this section by providing the following result on the convergence of the moments of the estimator.

Theorem 3.3. *Suppose that the variance function v satisfies*

$$v(s) \sim Cs^{2H}$$

for some $H \in (0,1)$ as $s \to 0$. Assume further that there exists $T_0 > 0$ such that for any $p \geq 1$ we have

$$\sup_{T \geq T_0} \mathbf{E} \left[\frac{1}{\left[\frac{1}{T} \int_0^T (X_u^\theta)^2 \mathrm{d}u \right]^p} \right] < \infty. \tag{3.3}$$

If also Assumption 3.3 holds, then for any $p \geq 1$ we have

$$\mathbf{E} \left[\left(\frac{|\psi'(\theta)|}{\sqrt{w_\theta(T)}} \left(\tilde{\theta}_T - \theta \right) \right)^p \right] \to \mathbf{E}\left[N^p \right],$$

where $N \sim \mathcal{N}(0,1)$.

Proof. By mean value theorem we have

$$\frac{|\psi'(\theta)|}{\sqrt{w_\theta(T)}} \left(\tilde{\theta}_T - \theta \right) = \frac{|\psi'(\theta)|}{\psi'\left(\psi^{-1}(\xi) \right)} \cdot \frac{1}{\sqrt{w_\theta(T)}} \left(Q_T^\theta + \varepsilon_\theta(T) \right),$$

where ξ is some random point between $\psi(\theta)$ and $\frac{1}{T} \int_0^T (X_t^\theta)^2 \mathrm{d}t$. Moreover, by continuous mapping theorem, Slutsky's theorem and Theorem 3.2, we know that

$$\left[\frac{|\psi'(\theta)|}{\psi'\left(\psi^{-1}(\xi) \right)} - \frac{|\psi'(\theta)|}{\psi'(\theta)} \right] \cdot \frac{1}{\sqrt{w_\theta(T)}} \left(Q_T^\theta + \varepsilon_\theta(T) \right)$$

converges to zero in distribution, and hence also in probability. Thus it suffices to show that $\frac{1}{\psi'(\psi^{-1}(\xi))}$ is bounded in L^p for any $p \geq 1$. Indeed, this implies that

$$\left\{ \left[\frac{|\psi'(\theta)|}{\psi'\left(\psi^{-1}(\xi) \right)} - \frac{|\psi'(\theta)|}{\psi'(\theta)} \right] \cdot \frac{1}{\sqrt{w_\theta(T)}} \left(Q_T^\theta + \varepsilon_\theta(T) \right) \right\}^p$$

is uniformly integrable for any $p \geq 1$, and thus converges to zero also in L^p. From this the claim follows since all the moments of $\frac{Q_T^\theta}{\sqrt{w_\theta(T)}}$ converge to the moments of standard normal random variable (see [26, Proposition 5.2.2.]).

First we estimate

$$
\begin{aligned}
\psi'(\theta) &= \frac{1}{2} \int_0^\infty s e^{-\theta s} \, \mathrm{d}v(s) \\
&\geq C \int_{\frac{1}{\theta}}^{\frac{2}{\theta}} s e^{-\theta s} \, \mathrm{d}v(s) \\
&\geq \frac{C}{\theta} \left[v\left(\frac{2}{\theta}\right) - v\left(\frac{1}{\theta}\right) \right] \\
&\geq C\theta^{-1-2H}
\end{aligned}
$$

for large enough θ.

We next prove that $\psi(\theta) \leq C\theta^{-2H\gamma}$ for any $\gamma \in (0,1)$ and large enough θ, which in turn implies $\psi^{-1}\left(\theta^{-2H\gamma}\right) \leq C\theta$. For this we write

$$
\int_0^\infty e^{-s} v\left(\frac{s}{\theta}\right) \mathrm{d}s = \int_0^{\theta^{1-\gamma}} e^{-s} v\left(\frac{s}{\theta}\right) \mathrm{d}s + \int_{\theta^{1-\gamma}}^\infty e^{-s} v\left(\frac{s}{\theta}\right) \mathrm{d}s.
$$

For the first term we have

$$
\begin{aligned}
\int_0^{\theta^{1-\gamma}} e^{-s} v\left(\frac{s}{\theta}\right) \mathrm{d}s &\leq v\left(\theta^{-\gamma}\right) \int_0^\infty e^{-s} \, \mathrm{d}s \\
&\leq C\theta^{-2H\gamma},
\end{aligned}
$$

and for the second term we have

$$
\begin{aligned}
\int_{\theta^{1-\gamma}}^\infty e^{-s} V\left(\frac{s}{\theta}\right) \mathrm{d}s &\leq e^{-\frac{\theta^{1-\gamma}}{2}} \int_0^\infty e^{-\frac{s}{2}} V\left(\frac{s}{\theta}\right) \mathrm{d}s \\
&\leq \theta e^{-\frac{\theta^{1-\gamma}}{2}} \psi(1) \\
&\leq C\theta^{-2H\gamma}.
\end{aligned}
$$

Together these estimates imply

$$
\frac{1}{|\psi'\left(\psi^{-1}(\xi)\right)|^p} \leq C\xi^{-\frac{(2H+1)p}{2H\gamma}}.
$$

This, together with (3.3), proves the claim. □

3.3 Discrete Observations

In practice continuous observations are rarely available. Therefore, it is important to consider the case of discrete observations. To control the

18

error introduced by the unobserved time-points, we assume that the driving noise G is Hölder continuous with some index $H \in (0, 1)$, i.e., G is Hölder continuous with parameter γ for all $\gamma < H$. The general idea is, that the smaller the H the more care must be taken in choosing the time-mesh of the observations. This gives rise to the condition (3.4) in Theorem 3.4.

Note that from the form of the Langevin equation it is immediate that any solution is Hölder continuous with index H if and only if the driving noise is Hölder continuous with the same index H. Due to [2, Corollary 1], the following assumption is not only sufficient, but also necessary, for the Hölder continuity with index H:

Assumption 3.4 (Hölder continuity). Let $H \in (0, 1)$. For all $\varepsilon > 0$ there exists a constant C_ε such that

$$v(t) \leq C_\varepsilon \, t^{2H-\varepsilon}.$$

For notational simplicity, we assume equidistant observation times $t_k = k\Delta_N$, $k = 0, \ldots, N$. Denote $T_N = N\Delta_N$ and assume that $\Delta_N \to 0$ with $T_N \to \infty$. The SME based on the discrete observations is

$$\tilde{\theta}_N = \psi^{-1} \left(\frac{1}{T_N} \sum_{k=1}^N (X_{k\Delta_N}^\theta)^2 \Delta_N \right).$$

Theorem 3.4 (Discrete Observations). *Suppose Assumption 3.2, Assumption 3.1 and Assumption 3.4 hold. Assume further that*

$$N\Delta_N^\beta \to 0, \tag{3.4}$$

where

$$\beta = \beta(H) = \frac{2H + \frac{1}{2}}{H + \frac{1}{2}} - \delta$$

for some $\delta > 0$. Then,

$$\tilde{\theta}_N \to \theta \quad a.s.$$

Moreover, if Assumption 3.3 holds, then

$$\frac{|\psi'(\theta)|}{\sqrt{w_\theta(T_N)}} \left(\tilde{\theta}_N - \theta \right) \xrightarrow{d} \mathcal{N}(0, 1).$$

Proof. Following the proof of [3, Theorem 3.2], it is enough to show that

$$\frac{1}{\sqrt{w_\theta(T_N)}} \left(\frac{1}{T_N} \sum_{k=1}^N (X_{k\Delta_N}^\theta)^2 \Delta_N - \frac{1}{T_N} \int_0^{T_N} (X_t^\theta)^2 \, \mathrm{d}t \right) \to 0 \quad \text{a.s.} \tag{3.5}$$

Let

$$Y_k = \sup_{t,s \in [t_{k-1}, t_k]} \frac{|X_t^\theta - X_s^\theta|}{|t - s|^{H-\varepsilon}}$$

19

be the $(H-\varepsilon)$-Hölder constant of the process X^θ on the subinterval $[t_{k-1}, t_k]$. Similarly, let (with slight abuse of notation) Y_N be the $(H-\varepsilon)$-Hölder constant of the process X^θ on the entire interval $[0, T_N]$ Then, by the identity $a^2 - b^2 = (a+b)(a-b)$, the Hölder continuity of X^θ, and the triangle inequality,

$$
\frac{1}{\sqrt{w_\theta(T_N)}} \left| \frac{1}{T_N} \sum_{k=1}^N (X_{t_k}^\theta)^2 \Delta_N - \frac{1}{T_N} \int_0^{T_N} (X_t^\theta)^2 \, \mathrm{d}t \right|
$$

$$
\leq \quad \frac{1}{T_N \sqrt{w_\theta(T_N)}} \sum_{k=1}^N \int_{t_{k-1}}^{t_k} \left| (X_{t_k}^\theta)^2 - (X_t^\theta)^2 \right| \mathrm{d}t
$$

$$
\leq \quad \frac{2}{T_N \sqrt{w_\theta(T_N)}} \sum_{k=1}^N \sup_{t \in [t_{k-1}, t_k]} \left| X_t^\theta \right| \int_{t_{k-1}}^{t_k} \left| X_{t_k}^\theta - X_t^\theta \right| \mathrm{d}t
$$

$$
\leq \quad \frac{2}{T_N \sqrt{w_\theta(T_N)}} \sum_{k=1}^N \sup_{t \in [t_{k-1}, t_k]} \left| X_t^\theta \right| Y_k \int_{t_{k-1}}^{t_k} \left| t - t_{k-1} \right|^{H-\varepsilon} \mathrm{d}t
$$

$$
= \quad \frac{C \Delta_N^{H-\varepsilon+1}}{T_N \sqrt{w_\theta(T_N)}} \sum_{k=1}^N \sup_{t \in [t_{k-1}, t_k]} \left| X_t^\theta \right| Y_k.
$$

Note that

$$
\sum_{k=1}^N \sup_{t \in [t_{k-1}, t_k]} \left| X_t^\theta \right| \leq T_N^{H-\varepsilon} Y_N,
$$

$$
\sum_{k=1}^N Y_k^2 \leq N Y_N^2,
$$

$$
w_\theta(T) \geq C T^{-1},
$$

where the last estimate follows, e.g., from Lemma 3.6. Consequently, it remains to show that

$$
N^{H-\varepsilon+\frac{1}{2}} \Delta_N^{2H-2\varepsilon+\frac{1}{2}} Y_N^2 \to 0 \quad \text{a.s.} \tag{3.6}
$$

By [2, Theorem 1 and Lemma 2] (see also [3, Remark 2.3]) we have for all $p \geq 1$ a constant $C = C_{\theta, H, \varepsilon, p}$ such that

$$
\mathbf{E}\left[Y_N^{2p} \right] \leq C T_N^{2\varepsilon p}.
$$

From this estimate and from the Markov's inequality it follows that for all $y > 0$ and $p \geq 1$,

$$
\mathbf{P}\left[\frac{Y_N^2}{N^\varepsilon} > y \right] \leq \frac{C}{y^p} \left(\Delta_N^{2\varepsilon} N^\varepsilon \right)^p.
$$

Now, by choosing p large enough, we obtain

$$\sum_{N=1}^{\infty} \mathbf{P}\left[\frac{Y_N^2}{N^{\varepsilon}} > y\right] < \infty$$

if

$$\Delta_N^{2\varepsilon} N^{\varepsilon} \leq N^{-\alpha}$$

for some $\alpha > 0$. By (3.4), we may choose $\alpha = 2\varepsilon/\beta - \varepsilon$. Indeed, since $\beta < 2$, it follows that $\alpha > 0$. Consequently, by the Borel–Cantelli lemma $N^{-\varepsilon} Y_N^2 \to 0$ almost surely. By applying this to (3.6) it remains to show that

$$N^{H+\frac{1}{2}} \Delta_N^{2H-2\varepsilon+\frac{1}{2}} \to 0.$$

But this follows from (3.4) by choosing $\varepsilon < \min\{H + 1/4, \delta(H + 1/2)/2\}$. The details are left to the reader. $\qquad\square$

Remark 3.5. The Berry–Esseen bound for Theorem 3.4 can be obtained as in the proof above by analyzing the speed of convergence in (3.5). We leave the details for the reader.

4 Examples

4.1 Fractional Ornstein–Uhlenbeck Process of the First Kind

The fractional Brownian motion B^H with Hurst index $H \in (0, 1)$ is the stationary-increment Gaussian process with variance function $v_H(t) = t^{2H}$. Actually, it is the (upto a multiplicative constant) unique stationary-increment Gaussian process that is H-self-similar meaning that

$$B^H \stackrel{\mathrm{d}}{=} a^{-H} B_a^H.$$

for all $a > 0$. For the fractional Brownian motion the Hurst index H is both the index of self-similarity and the Hölder index. We refer to Biagini et al. [6] and Mishura [25] for more information on the fractional Brownian motion. The fractional Ornstein–Uhlenbeck process (of the first kind) is the stationary solution to the Langevin equation

$$\mathrm{d}U_t^{H,\theta} = -\theta U_t^{H,\theta}\,\mathrm{d}t + \mathrm{d}B_t^H, \quad t \geq 0. \tag{4.1}$$

The fractional Ornstein–Uhlenbeck processes (of different kinds) and related parameter estimations have been studied extensively recently, see, e.g., [3, 8, 15, 16, 17, 18, 32]. By Cheridito et al. [8, Theorem 2.3] the autocovariance $r_{H,\theta}$ of the stationary solution satisfies, for $H \neq 1/2$, the asymptotic expansion

$$r_{H,\theta}(t) \sim \frac{H(2H-1)}{\theta^2} t^{2H-2} \tag{4.2}$$

as $t \to \infty$. Also, by Hu and Nualart [15],

$$\psi_H(\theta) = \frac{H\Gamma(2H)}{\theta^{2H}},$$

where Γ is the Gamma function. Consequently, Assumption 3.2 and Assumption 3.1 are satisfied for all H, and Assumption 3.3 is satisfied for $H \le 3/4$. Also, Assumption 3.4, required for discrete observations, is satisfied for all H. Finally, we observe that Corollary 3.1 is applicable for $H \in (0, 3/4)$, and by using the self-similarity of the fractional Brownian motion it is clear that

$$\int_0^\infty r_{H,\theta}(t)^2 \, \mathrm{d}t = \theta^{-2H} \sigma_H^2,$$

where we have denoted

$$\sigma_H^2 = \int_0^\infty r_{H,1}(t)^2 \, \mathrm{d}t. \tag{4.3}$$

Let $\tilde{\theta}_T^H$ be the SME associated with the equation (4.1). Proposition 4.1 below extends the result of Hu and Nualart [15, Theorem 4.1] both by extending the range of H and by providing the Berry–Esseen bounds. We note, however, that the result of Proposition 4.1 is far from optimal. Indeed, in this case the maximum likelihood estimator with optimal rate for all $H \in (0, 1)$ can be constructed as in Kleptsyna and Le Breton [18].

Proposition 4.1 (Fractional Ornstein–Uhlenbeck Process of the First Kind). *Let σ_H be given by (4.3).*

(i) *Let $H \in (0, 1/2]$. Then*

$$\sup_{x \in \mathbb{R}} \left| \mathbf{P} \left[\sqrt{\frac{T}{\theta \sigma_H^2}} \left(\tilde{\theta}_T^H - \theta \right) \le x \right] - \Phi(x) \right| \le \frac{C_{H,\theta}}{\sqrt{T}}.$$

(ii) *Let $H \in (1/2, 3/4)$. Then*

$$\sup_{x \in \mathbb{R}} \left| \mathbf{P} \left[\sqrt{\frac{T}{\theta \sigma_H^2}} \left(\tilde{\theta}_T^H - \theta \right) \le x \right] - \Phi(x) \right| \le \frac{C_{H,\theta}}{\sqrt{T^{3-4H}}}.$$

(iii) *Let $H = 3/4$. Then*

$$\sup_{x \in \mathbb{R}} \left| \mathbf{P} \left[\sqrt{\frac{T}{\theta \sigma^2 \log T}} \left(\tilde{\theta}_T^{3/4} - \theta \right) \le x \right] - \Phi(x) \right| \le \frac{C_{3/4,\theta}}{\sqrt{\log T}},$$

where σ is an absolute constant.

Proof. Consider first the case $H \in (0, 1/2)$. By Corollary 3.1, it is enough to show that

$$\frac{1}{\sqrt{T}} \int_0^T |r_{H,\theta}(t)| \, \mathrm{d}t + \sqrt{\int_T^\infty r_{H,\theta}(t)^2 \, \mathrm{d}t} + \sqrt{\frac{1}{T} \int_0^T r_{H,\theta}(t)^2 t \, \mathrm{d}t} \le \frac{C_{H,\theta}}{\sqrt{T}}.$$

Here the first term is the dominating one. Indeed, by (4.2),

$$\frac{1}{\sqrt{T}} \int_0^\infty |r_{H,\theta}(t)| \, \mathrm{d}t \quad \sim \quad \frac{1}{\sqrt{T}} C_{H,\theta} \int_1^\infty t^{2H-2} \, \mathrm{d}t$$

$$\le \quad \frac{C_{\theta,H}}{\sqrt{T}},$$

$$\sqrt{\int_T^\infty r_{H,\theta}(t)^2 \, \mathrm{d}t} \quad \sim \quad C_{H,\theta} \sqrt{\int_T^\infty t^{4H-4} \, \mathrm{d}t}$$

$$= \quad C_{H,\theta} \sqrt{T^{4H-3}},$$

$$\sqrt{\frac{1}{T} \int_0^T r_{H,\theta}(t)^2 t \, \mathrm{d}t} \quad \sim \quad C_{H,\theta} \sqrt{\frac{1}{T} \int_1^T t^{4H-3} \, \mathrm{d}t}$$

$$= \quad C_{H,\theta} \sqrt{T^{4H-3}}$$

The case $H = 1/2$ is classical and well-known, and stated here only for the sake of completeness.

The case $H \in (1/2, 3/4)$ can be analyzed exactly the same way as the case $H \in (0, 1/2)$, except now it is the second and third terms that dominate.

Consider then the case $H = 3/4$. Now Corollary 3.1 is not applicable. Consequently, we have to use Theorem 3.2 directly. Let us first calculate the asymptotic rate. By applying l'Hôpital's rule twice and then the asymptotic expansion (4.2), we obtain

$$w_{3/4,\theta}(T) \quad = \quad \frac{2}{T^2} \int_0^T \int_0^T r_{3/4,\theta}(t-s)^2 \, \mathrm{d}s \mathrm{d}t$$

$$\sim \quad \frac{2 \log T}{T} \frac{1}{\log T} \int_0^T r_{3/4,\theta}(t)^2 \, \mathrm{d}t$$

$$\sim \quad \frac{2(3/8)^2}{\theta^4} \frac{\log T}{T} \frac{1/T}{1/T}$$

$$= \quad \frac{2(3/8)^2}{\theta^4} \frac{\log T}{T},$$

Consequently,

$$\frac{|\psi'_{3/4}(\theta)|}{\sqrt{w_{3/4,\theta}(T)}} \sim \frac{\frac{3}{4}\Gamma(\frac{3}{2})\frac{3}{2}\theta^{-5/2}}{\sqrt{2}\frac{3}{8}\theta^{-2}} \sqrt{\frac{T}{\log T}} = \sqrt{\frac{T}{\theta\sigma^2 \log T}},$$

by setting σ^2 appropriately. For the Berry–Esseen upper bound, we estimate

$$\frac{\int_0^T |r_{3/4,\theta}(t)|\,\mathrm{d}t}{T\sqrt{w_{3/4,\theta}(T)}} \le C_\theta \frac{\int_0^T t^{-1/2}\,\mathrm{d}t}{\sqrt{T}\sqrt{\log T}} \le C_\theta \frac{\sqrt{T}}{\sqrt{T}\sqrt{\log T}}$$

The claim follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 4.1. In the case $H > 3/4$ our method does not provide asymptotic normality. Indeed, due to the results in Breton and Nourdin [7] it is expected that asymptotic normality cannot hold in this case.

4.2 Noises Arising from Self-Similar Processes

The examples in the next two subsections deal with self-similar processes. The motivation comes from the result of Doob [9] stating that the classical Ornstein–Uhlenbeck process can be viewed as the inverse Lamperti transform of the Brownian motion that is $1/2$-self-similar. Therefore, let us start with an H-self-similar Gaussian process Y^H. (For a representation of such processes in terms of the Brownian motion see Yazigi [39]). The inverse Lamperti transform with self-similarity parameters H and scale parameter θ is

$$(\mathscr{L}_{H,\theta}^{-1} Y^H)_t = \mathrm{e}^{-\theta t} Y_{a(t)}^H,$$

where

$$a(t) = a_{H,\theta}(t) = \frac{H}{\theta} \mathrm{e}^{\frac{\theta}{H} t}.$$

If Y^H is H-self-similar, then $\mathscr{L}_{H,\theta}^{-1} Y^H$ is stationary, and vice versa. Furthermore, and all stationary solutions $U^\theta = U^{H,\theta}$ of the Langevin equation are inverse Lamperti transforms of some H-self-similar Y^H, see [21, 38]. Therefore we have, on the one hand,

$$U_t^{H,\theta} = \mathrm{e}^{-\theta t} Y_{a(t)}^H$$

for some H-self-similar Y^H, and, on the other hand,

$$U_t^{H,\theta} = \int_{-\infty}^t \mathrm{e}^{-\theta(t-s)}\,\mathrm{d}G_s^{H,\theta}$$

for some stationary-increment noise $G^{H,\theta}$ arising from the H-self-similar process Y^H. Actually, we have

$$G_t^{H,\theta} = \int_0^t \mathrm{e}^{-\theta s}\,\mathrm{d}Y_{a(t)}^H.$$

Moreover, the family $\{G^{H,\theta}\,;\,\theta > 0\}$ satisfies the scaling property

$$\theta^H G_{\cdot/\theta}^{H,\theta} \overset{\mathrm{d}}{=} G^H,$$

24

where we have denoted $G^H = G^{H,1}$. This motivates the study of the following Langevin equation where the noise arises from the H-self-similar process Y^H that depends on θ, but with a noise G^H that is independent of θ:

$$\mathrm{d}U_t^{H,\theta} = -\theta U_t^{H,\theta}\mathrm{d}t + \mathrm{d}G_t^H. \tag{4.4}$$

The solutions of these Langevin equations are called Ornstein–Uhlenbeck processes of the second kind.

4.3 Fractional Ornstein–Uhlenbeck Process of the Second Kind

The fractional Ornstein–Uhlenbeck process of the second kind arises by setting $Y^H = B^H$, the fractional Brownian motion. This process has been studied e.g. in [1, 3, 17]. By Kaarakka and Salminen [17, Proposition 3.11] the autocovariance $r_{H,\theta}$ of the fractional Ornstein–Uhlenbeck process of the second kind has exponential decay. Consequently, Corollary 3.1 implies the following:

Proposition 4.2. *For the fractional Ornstein–Uhlenbeck processes of the second find*

$$\sup_{x\in\mathbb{R}}\left|\mathbf{P}\left[\frac{\sqrt{T}}{\sigma_{H,\theta}(\theta)}\left(\tilde{\theta}_T^H - \theta\right) \le x\right] - \Phi(x)\right| \le \frac{C_{H,\theta}}{\sqrt{T}},$$

where

$$\sigma_{H,\theta}(\theta) = \sqrt{4\int_0^\infty r_{H,\theta}(t)^2\,\mathrm{d}t}.$$

4.4 Bifractional Ornstein–Uhlenbeck Process of the Second Kind

The bifractional Brownian motion $B^{H,K}$, introduced by Houdré and Villa [14], with parameters $H \in (0,1)$ and $K \in (0,1]$ is the Gaussian process with covariance

$$\mathbf{E}\left[B_t^{H,K}B_t^{H,K}\right] = \frac{1}{2^K}\left[(t^{2H} + s^{2H})^K - |t - s|^{2HK}\right].$$

The bifractional Brownian motion is a generalization of the fractional Brownian motion, but it does not have stationary increments, except in the fractional case $K = 1$. Consequently, there does not seem to be a natural way to define the bifractional Ornstein–Uhlenbeck process of the first kind that would have a stationary version. The bifractional Brownian motion is, however, HK-self-similar, see Russo and Tudor [29]. Consequently, we can define the bifractional Ornstein–Uhlenbeck process by setting $Y^{HK} = B^{H,K}$ in equation (4.4).

Lemma 4.1. *The autocovariance $r_{H,K,\theta}$ of the bifractional Ornstein–Uhlenbeck process of the second kind has exponential decay.*

Lemma 4.1 combined with Corollary 3.1 immediately yields:

Proposition 4.3. *For the bifractional Ornstein–Uhlenbeck processes of the second find*

$$\sup_{x\in\mathbb{R}}\left|\mathbf{P}\left[\frac{\sqrt{T}}{\sigma_{H,K,\theta}}\left(\tilde{\theta}_T^{H,K,\theta}-\theta\right)\leq x\right]-\Phi(x)\right|\leq\frac{C_{H,K,\theta}}{\sqrt{T}},$$

where

$$\sigma_{H,K,\theta}=\sqrt{4\int_0^\infty r_{H,K,\theta}(t)^2\,\mathrm{d}t}.$$

5 Discussion on Other Estimators

It is a celebrated result by Gauss [12] that for multivariate Gaussian distributions the Least Squares Estimator (LSE) and the Maximum Likelihood Estimator (MLE) coincide, and this is a characterizing property of the Gaussian distribution. Indeed, this is why Gaussian distributions are named thus. In the infinite-dimensional case of Gaussian processes, the situation is more delicate, as the following discussion shows. The discussion is based on Hu and Nualart [15] in the case of the LSE and on Kleptsyna and Le Breton [18] in the case of the MLE. To make the discussion short, we do not present explicit assumptions in terms of the variance v, although this would be possible. Instead, we confine ourselves in presenting the general ideas and implicit assumptions.

5.1 Least Squares Estimator

In this subsection, δ denotes the Skorohod integral. We refer to Nualart [27] for details on Skorohod integrals.

One the one hand, the LSE

$$\hat{\theta}_T=-\frac{\int_0^T X_t^\theta\,\delta X_t^\theta}{\int_0^T (X_t^\theta)^2\,\mathrm{d}t}\tag{5.1}$$

arises heuristically by minimizing

$$\int_0^T |\dot{X}_t^\theta+\theta X_t^\theta|^2\,\mathrm{d}t.$$

On the other hand, by using the Langevin equation of the solution X^θ, one would hope that

$$\int_0^T X_t^\theta \, \delta X_t^\theta = -\theta \int_0^T (X_t^\theta)^2 \, \mathrm{d}t + \int_0^T X_t^\theta \, \delta G_t. \qquad (5.2)$$

This would lead to the LSE

$$\widehat{\theta}_T = \theta - \frac{\int_0^T X_t^\theta \, \delta G_t}{\int_0^T (X_t^\theta)^2 \, \mathrm{d}t}. \qquad (5.3)$$

Unfortunately, the Skorohod integral is not (bi)linear. In particular, the equation (5.2) does not hold. Indeed, this is obvious from the fact that Skorohod integrals have zero mean. Consequently, the LSE's defined by (5.1) and (5.3), respectively, are not the same. The LSE defined by (5.3) has been shown to be consistent for some fractional Ornstein–Uhlenbeck processes, see [1, 15]. However, the LSE (5.3) depends on θ, the parameter we want to estimate! Therefore, the LSE (5.3) is not an estimator at all. Moreover, to show that the LSE (5.3) is consistent, one has to show that the term $\int_0^T X_t^\theta \, \delta G_t / \int_0^T (X_t^\theta)^2 \, \mathrm{d}t$ converges to zero. This suggest that it would be more natural to define the LSE by (5.1). However, Proposition 5.1 below shows that the LSE (5.1) will fail under rather general assumptions.

Proposition 5.1. *Assume that U^θ is ergodic, and that the Skorohod integral $\int_0^T X_t^\theta \, \delta X_t^\theta$ exists. Let $\hat{\theta}_T$ be defined by (5.1). If $(X_T^\theta)^2/T \to 0$ in $L^1(\Omega)$ and almost surely, then*

$$\hat{\theta}_T \to 0 \quad a.s.$$

Proof. By the Itô formula in [33],

$$\int_0^T X_t^\theta \, \delta X_t^\theta = \frac{1}{2}(X_T^\theta)^2 - \frac{1}{2}\mathbf{E}[(X_T^\theta)^2].$$

Since U^θ is ergodic, Proposition 3.1 implies that

$$\frac{1}{T}\int_0^T (X_t^\theta)^2 \, \mathrm{d}t \to \psi(\theta) > 0 \quad \text{a.s.}$$

Consequently,

$$\hat{\theta}_T = \frac{1}{2}\frac{(X_T^\theta)^2/T - \mathbf{E}[(X_T^\theta)^2/T]}{\frac{1}{T}\int_0^T (X_t^\theta)^2 \, \mathrm{d}t} \to 0 \quad \text{a.s,}$$

and the claim follows. $\qquad \square$

5.2 Maximum Likelihood Estimator

It is well-known stat the Maximum Likelihood Estimator (MLE) is the best estimator in many ways: it is e.g. typically optimal in the local minimax sense. The MLE is known for example for the fractional Ornstein–Uhlenbeck process of the first kind [18], and in this case it outperforms the SME. We discuss shortly, how the MLE could be constructed for more general cases.

In this subsection, the integrals are abstract Wiener integral as defined in e.g. [34], or, equivalently, Skorohod integrals as defined e.g. in [33].

To construct the MLE, we assume the following Volterra representation for the noise G: There exists a Gaussian martingale M with bracket $\langle M \rangle$ and a kernel $k \in L^2_{\mathrm{loc}}(\mathbb{R}^2_+, \mathrm{d}\langle M \rangle \times \mathrm{d}\langle M \rangle)$ such that

$$G_t = \int_0^t k(t,s) \, \mathrm{d}M_s. \tag{5.4}$$

Furthermore, we assume the following inverse Volterra representation:

$$M_t = \int_0^t k^*(t,s) \, \mathrm{d}G_t. \tag{5.5}$$

Next, we define

$$M_t^\theta = \int_0^t k^*(t,s) \, \mathrm{d}X_t^\theta,$$

$$\Xi_t^\theta = \frac{\mathrm{d}}{\mathrm{d}\langle M \rangle_t} \int_0^t k^*(t,s) X_s^\theta \, \mathrm{d}s,$$

implicitly assuming their existence.

Proposition 5.2 (MLE). *Assume representations (5.4)–(5.5) and assume that $\Xi^\theta \in L^2(\Omega \times [0,T], \mathrm{d}\mathbf{P} \times \mathrm{d}\langle M \rangle)$. Then the MLE based on the observations X_t^θ, $t \in [0,T]$, is*

$$\bar{\theta}_T = -\frac{\int_0^T \Xi_t^\theta \, \mathrm{d}M_t^\theta}{\int_0^T (\Xi_t^\theta)^2 \, \mathrm{d}\langle M \rangle_t}.$$

Moreover, if $\int_0^T (\Xi_t^\theta)^2 \, \mathrm{d}\langle M \rangle_t \to \infty$ almost surely, then the MLE is strongly consistent.

Proof. By integrating the Langevin equation (1.1) against the kernel k^* on both sides, we obtain

$$M_t^\theta = -\theta \int_0^t k^*(t,s) X_s^\theta \, \mathrm{d}s + M_t.$$

By plugging in Ξ^θ, this translates into

$$\mathrm{d}M_t^\theta = -\theta \Xi_t^\theta \, \mathrm{d}\langle M \rangle_t + \mathrm{d}M_t.$$

Consequently, we can use the Girsanov theorem for Gaussian martingales, which states that the log-likelihood $\ell_T(\theta) = \log \mathrm{d}\mathbf{P}_T^\theta / \mathrm{d}\mathbf{P}_T$ can be written as

$$\ell_T(\theta) = -\theta \int_0^T \Xi_t \, \mathrm{d}M_t^\theta - \frac{\theta^2}{2} \int_0^T (\Xi_t^\theta)^2 \, \mathrm{d}\langle M \rangle_t.$$

The MLE $\bar{\theta}_T$ follows from this by maximizing with respect to θ.

The strong consistency follows from the equation

$$\bar{\theta}_T - \theta = -\frac{\int_0^T \Xi_t^\theta \, \mathrm{d}M_t}{\int_0^T (\Xi_t^\theta)^2 \, \mathrm{d}\langle M \rangle_t}$$

by using the martingale convergence theorem. $\qquad\square$

6   Conclusions

We have considered the Langevin equation with general stationary-increment continuous centered Gaussian noise. We have stated mild conditions on the variance function of the noise ensuring strong consistency and asymptotic normality of the so-called second moment estimator of the mean-reversion parameter. We have also provided Berry–Esseen bounds for the normal approximation. We have shown that the second moment estimator works for discrete observations provided that the noise is Hölder continuous and the observation-mesh is dense enough with respect to the Hölder index of the noise. We have also shown that our results work in examples rising from fractional and bifractional Brownian noises, thus extending some recent results. Finally, we discussed least squares estimators and maximum likelihood estimators. We argue that the second moment estimator is, under the stationarity assumption, the most general estimator, i.e., it works under the mildest assumptions.

A   Proofs of Lemmas

Proof of Lemma 2.1. Let $t > 0$ and let $\lfloor t \rfloor$ be the greatest integer not exceeding t. Then

$$|G_t| \le |G_t - G_{\lfloor t \rfloor}| + \sum_{k=1}^{\lfloor t \rfloor} |G_k - G_{k-1}|.$$

By the Minkowski's inequality and stationary of the increments,

$$\sqrt{\mathbf{E}[G_t^2]} \le \sqrt{\mathbf{E}[G_{t-\lfloor t \rfloor}^2]} + \lfloor t \rfloor \sqrt{\mathbf{E}[G_1^2]}.$$

The claim follows from this. $\qquad\square$

Proof of Lemma 3.1. By changing the variable in (2.4) we obtain

$$\psi(\theta) = \frac{1}{2} \int_0^\infty e^{-t} \, v\left(\frac{t}{\theta}\right) dt. \tag{A.1}$$

Since v is strictly increasing, this shows that ψ is also strictly decreasing. Furthermore, $\psi(\theta) \to 0$ as $\theta \to \infty$ by the monotone convergence theorem.

By the Lebesgue's dominated convergence theorem, the function

$$\theta \mapsto \int_0^\infty e^{-\theta t} v(t) \, dt$$

is smooth. Consequently, ψ is smooth.

Finally, let us show that ψ is convex. Differentiating ψ we observe

$$\psi'(\theta) = \frac{1}{2} \int_0^\infty e^{-\theta s} v(s) \, ds - \frac{\theta}{2} \int_0^\infty s e^{-\theta s} v(s) \, ds$$

from which it follows, by integration by parts, that

$$\psi'(\theta) = -\frac{1}{2} \int_0^\infty s e^{-\theta s} \, dv(s),$$

where dv is the measure associated to the increasing function v. Similarly, differentiating and using integration by parts again we get

$$\psi''(\theta) = \frac{1}{2} \int_0^\infty s^2 e^{-\theta s} \, dv(s) \geq 0.$$

Consequently, ψ is convex. $\qquad\square$

Proof of Lemma 3.2. Let $T \geq 0$ be fixed. Then, by [33, Theorem 12], there exists a Brownian motion W and a kernel $K_T \in L^2([0,T]^2)$ such that we have the representation

$$X_t = \int_0^T K_T(t,s) \, dW_s,$$

for all $t \leq T$. Consequently X_t belongs to the 1st Wiener chaos for all $t > 0$. Then note that

$$X_t^2 - \mathbf{E}[X_t^2] = 2H_2(X_t).$$

Consequently, it belongs to the 2nd Wiener chaos. Finally, note that, because γ is continuous, the integral

$$\frac{1}{T} \int_0^T 2H_2(X_t) \, dt$$

can be defined as a limit in $L^2(\Omega)$ in the 2nd Wiener chaos. The claim follows from this. $\qquad\square$

Proof of Lemma 3.3. The claim is an obvious consequence of Isserlis' theorem. $\qquad\square$

Proof of Lemma 3.4. By the symmetry $\gamma(s,t) = \gamma(t,s)$

$$\int_{[0,T]^4} \gamma(t_1,t_2)\gamma(t_2,t_3)\gamma(t_3,t_4)\gamma(t_4,t_1)\,\mathrm{d}t_1\mathrm{d}t_2\mathrm{d}t_3\mathrm{d}t_4$$

$$= \int_0^T\int_0^T \left(\int_0^T \gamma(t_1,t_2)\gamma(t_2,t_3)\,\mathrm{d}t_2\right)^2 \mathrm{d}t_1\mathrm{d}t_3$$

$$\leq \int_0^T\int_0^T \left(\int_0^T |\gamma(t_1,t_2)|\,\mathrm{d}t_2\right)\left(\int_0^T |\gamma(t_1,t_2)|\gamma(t_2,t_3)^2\,\mathrm{d}t_2\right)\mathrm{d}t_1\mathrm{d}t_3$$

$$\leq \int_0^T \left(\sup_{t\in[0,T]}\int_0^T |\gamma(t,t_2)|\,\mathrm{d}t_2 \int_0^T\int_0^T |\gamma(t_1,t_2)|\,\mathrm{d}t_1\,\gamma(t_2,t_3)^2\,\mathrm{d}t_2\right)\mathrm{d}t_3$$

$$\leq \left[\sup_{t\in[0,T]}\int_0^T |\gamma(t,t_2)|\,\mathrm{d}t_2\right]^2 \int_0^T\int_0^T \gamma(t_2,t_3)^2\,\mathrm{d}t_2\mathrm{d}t_3.$$

$\qquad\square$

Proof of Lemma 3.5. Note that $Q_T/\sqrt{\mathbf{E}[Q_T^2]}$ belongs to the 2nd Wiener chaos and has unit variance. Consequently, by the fourth-moment Proposition 3.2, it suffices to show that

$$\frac{\mathbf{E}\left[Q_T^4\right]}{\mathbf{E}\left[Q_T^2\right]^2} - 3 \leq C\frac{\left(\sup_{t\in[0,T]}\int_{[0,T]}|\gamma(t,s)|\,\mathrm{d}s\right)^2}{\int_{[0,T]^2}\gamma(t,s)^2\,\mathrm{d}s\mathrm{d}t}.$$

Denote

$$I_2(T) = \int_{[0,T]^2} \gamma(t_1,t_s)^2\,\mathrm{d}t_1\mathrm{d}t_2,$$

$$I_4(T) = \int_{[0,T]^4} \gamma(t_1,t_2)\gamma(t_2,t_3)\gamma(t_3,t_4)\gamma(t_4,t_1)\,\mathrm{d}t_1\mathrm{d}t_2\mathrm{d}t_3\mathrm{d}t_4.$$

Then, by Lemma 3.3,

$$\mathbf{E}[Q_T^4] = \frac{12}{T^4}I_2(T)^2 + \frac{24}{T^4}I_4(T),$$

$$\mathbf{E}[Q_T^2]^2 = \frac{4}{T^4}I_2(T)^2,$$

and, by Lemma 3.4

$$I_4(T) \leq \left[\sup_{t\in[0,T]}\int_0^T |\gamma(t,s)|\,\mathrm{d}s\right]^2 I_2(T).$$

31

Consequently,

$$\frac{\mathbf{E}\left[Q_T^4\right]}{\left(\mathbf{E}\left[Q_T^2\right]\right)^2} - 3 = \frac{12I_2(T)^2 + 24I_4(T) - 12I_2(T)^2}{4I_2(T)^2}$$

$$= 6\frac{I_4(T)}{I_2(T)^2}$$

$$\leq 6\frac{\left[\sup_{t\in[0,T]}\int_0^T |\gamma(t,s)|\,\mathrm{d}s\right]^2}{I_2(T)}.$$

The claim follows from this. □

Proof of Lemma 3.6. We have

$$\frac{\mathbf{E}[(Q_T^\theta)^2]}{w_\theta(T)} = 1 + \frac{\frac{1}{T}\int_0^T \int_0^T [\mathbf{E}(X_t X_s)]^2 - r^2(t-s)\mathrm{d}s\mathrm{d}t}{Tw_\theta(T)}. \tag{A.2}$$

By Proposition 2.2 and ergodicity condition $\frac{1}{T}\int_0^T |r(t)|\mathrm{d}t \to 0$ it follows that

$$\frac{1}{T}\int_0^T \int_0^T [\mathbf{E}(X_t X_s)]^2 - r^2(t-s)\mathrm{d}s\mathrm{d}t \to 0. \tag{A.3}$$

Moreover, by symmetry, change-of-variables and the Fubini theorem

$$\int_0^T \int_0^T r_\theta(t-s)^2\,\mathrm{d}s\mathrm{d}t = 2\int_0^T \int_t^T r_\theta(s)^2\,\mathrm{d}t\mathrm{d}s = 2\int_0^T r_\theta(t)^2(T-t)\,\mathrm{d}t.$$

Consequently, by assuming that $T > 1$,

$$\int_0^T r_\theta(t)^2(T-t)\,\mathrm{d}t \geq \int_0^1 r_\theta(t)^2(T-t)\,\mathrm{d}t \geq C(T-1)$$

This shows $Tw_\theta(T) \geq C$ which, together with (A.2) and (A.3), implies $\mathbf{E}[(Q_T^\theta)^2] \sim w_\theta(T)$. Also, we have shown the equality

$$w_\theta(T) = \frac{4}{T^2}\int_0^T r_\theta(t)^2(T-t)\,\mathrm{d}t.$$

Consider then the case $\int_0^\infty r_\theta(t)^2\,\mathrm{d}t < \infty$. By the equivalence $\mathbf{E}[(Q_T^\theta)^2] \sim \frac{4}{T^2}\int_0^T r_\theta(t)^2(T-t)\,\mathrm{d}t$, we have

$$\frac{1}{T}\int_0^T r_\theta(t)^2(T-t)\,\mathrm{d}t = \int_0^T r_\theta(t)^2\,\mathrm{d}t - \frac{1}{T}\int_0^T r_\theta(t)^2 t\,\mathrm{d}t.$$

Here the first term converges to $\int_0^\infty r_\theta(t)^2\,\mathrm{d}t < \infty$. For the second term we have

$$\frac{1}{T}\int_0^T r_\theta(t)^2 t\,\mathrm{d}t \leq \frac{1}{\sqrt{T}}\int_0^{\sqrt{T}} r_\theta(t)^2 t\,\mathrm{d}t + \int_{\sqrt{T}}^\infty r_\theta(t)^2 t\,\mathrm{d}t \to 0$$

which completes the proof. □

Proof of Lemma 3.7. Let us first split

$$
\sup_{x\in\mathbb{R}}\left|\mathbf{P}\left[\frac{Q_T^\theta}{\sqrt{w_\theta(T)}}\le x\right]-\Phi(x)\right|
$$

$$
\le\ \sup_{x\in\mathbb{R}}\left|\mathbf{P}\left[\frac{Q_T^\theta}{\sqrt{\mathbf{E}[(Q_T^\theta)^2]}}\le\sqrt{\frac{w_\theta(T)}{\mathbf{E}[(Q_T^\theta)^2]}}x\right]-\Phi\left(\sqrt{\frac{w_\theta(T)}{\mathbf{E}[(Q_T^\theta)^2]}}x\right)\right|
$$

$$
+\sup_{x\in\mathbb{R}}\left|\mathbf{P}\left[\Phi\left(\sqrt{\frac{w_\theta(T)}{\mathbf{E}[(Q_T^\theta)^2]}}x\right)\right]-\Phi(x)\right|
$$

$$
\le\ \sup_{x\in\mathbb{R}}\left|\mathbf{P}\left[\frac{Q_T^\theta}{\sqrt{\mathbf{E}[(Q_T^\theta)^2]}}\le x\right]-\Phi(x)\right|
$$

$$
+\sup_{x\in\mathbb{R}}\left|\Phi\left(\sqrt{\frac{w_\theta(T)}{\mathbf{E}[(Q_T^\theta)^2]}}x\right)-\Phi(x)\right|
$$

$$
=\ A_1+A_2.
$$

For the term A_1, let us first estimate

$$
\int_0^T|\gamma_\theta(t,s)|\,\mathrm{d}s
$$

$$
=\ \int_0^T\left|r_\theta(t-s)+\mathrm{e}^{-\theta(t+s)}r_\theta(0)-\mathrm{e}^{-\theta t}r_\theta(s)-\mathrm{e}^{-\theta s}r_\theta(t)\right|\,\mathrm{d}s
$$

$$
\le\ \int_0^T|r_\theta(t-s)|\,\mathrm{d}s+\frac{\mathrm{e}^{-\theta t}|r_\theta(0)|}{\theta}+\mathrm{e}^{-\theta t}\int_0^T|r_\theta(s)|\,\mathrm{d}s+\frac{|r_\theta(t)|}{\theta}
$$

$$
\le\ \int_0^T|r_\theta(t-s)|\,\mathrm{d}s+\mathrm{e}^{-\theta t}\int_0^T|r_\theta(s)|\,\mathrm{d}s+\frac{\left(\mathrm{e}^{-\theta t}+1\right)|r_\theta(0)|}{\theta}
$$

$$
\le\ \int_0^T|r_\theta(t-s)|\,\mathrm{d}s+\int_0^T|r_\theta(s)|\,\mathrm{d}s+C,
$$

Consequently,

$$
\sup_{t\in[0,T]}\int_0^T|\gamma_\theta(t,s)|\,\mathrm{d}s\ \le\ 2\sup_{t\in[0,T]}\int_0^T|r_\theta(t-s)|\,\mathrm{d}s+C
$$

$$
=\ 2\sup_{t\in[0,T]}\int_{-t}^{T-t}|r_\theta(u)|\,\mathrm{d}u+C
$$

$$
\le\ 2\int_{-T}^T|r_\theta(u)|\,\mathrm{d}u+C
$$

$$
=\ 4\int_0^T|r_\theta(u)|\,\mathrm{d}u+C.
$$

Since we are interested in the case $T \to \infty$, we can assume that T is bigger than some absolute positive constant. Consequently, since r_θ continuous with $r_\theta(0) > 0$, it follows from the estimate above that

$$\sup_{t \in [0,T]} \int_0^T |\gamma_\theta(t,s)| \, \mathrm{d}s \le C \int_0^T |r_\theta(u)| \, \mathrm{d}u.$$

Therefore, by applying Lemma 3.5 and Lemma 3.6, it follows that $A_1 \le C_\theta R_\theta(T)$.

Let us then consider the term A_2. Now, by the mean value theorem,

$$
\begin{aligned}
A_2 &= \sup_{x \in \mathbb{R}} \left| \Phi\left(\sqrt{\frac{w_\theta(T)}{\mathbf{E}[(Q_T^\theta)^2]}} x \right) - \Phi(x) \right| \\
&\le \frac{1}{\sqrt{2\pi}} \sup_{x \in \mathbb{R}} \left(\mathrm{e}^{-1/2 \eta_\theta(T)^2} |x| \right) \left| \sqrt{\frac{w_\theta(T)}{\mathbf{E}[(Q_T^\theta)^2]}} - 1 \right|,
\end{aligned}
$$

where

$$\eta_\theta(T,x) \in \left[x, x + \sqrt{\frac{w_\theta(T)}{\mathbf{E}[(Q_T^\theta)^2]}} x \right].$$

Since $\sqrt{\frac{w_\theta(T)}{\mathbf{E}[(Q_T^\theta)^2]}} \sim 1$, it follows that

$$A_2 \le \left| \sqrt{\frac{w_\theta(T)}{\mathbf{E}\left[(Q_T^\theta)^2 \right]}} - 1 \right| = \frac{\left| \sqrt{w_\theta(T)} - \sqrt{\mathbf{E}\left[(Q_T^\theta)^2 \right]} \right|}{\sqrt{\mathbf{E}\left[(Q_T^\theta)^2 \right]}}.$$

Consequently, by the asymptotic equivalence of $w_\theta(T) \sim \mathbf{E}[(Q_T^\theta)^2]$, it remains to show that

$$T \left| \sqrt{w_\theta(T)} - \sqrt{\mathbf{E}[(Q_T^\theta)^2]} \right| \le C_\theta \int_0^T |r_\theta(t)| \, \mathrm{d}t.$$

(Actually, we show that the left hand side is bounded.) For this purpose, we estimate, by using the inequality $|\sqrt{a} - \sqrt{b}| \le \sqrt{|a-b|}$ and the identity $a^2 - b^2 = (a+b)(a-b)$, that

$$
\begin{aligned}
&T \left| \sqrt{w_\theta(T)} - \sqrt{\mathbf{E}[(Q_T^\theta)^2]} \right| \\
&= \sqrt{2} \left| \sqrt{\int_0^T \int_0^T r_\theta(t-s)^2 \, \mathrm{d}s \mathrm{d}t} - \sqrt{\int_0^T \int_0^T \gamma_\theta(t,s)^2 \, \mathrm{d}s \mathrm{d}t} \right| \\
&\le \sqrt{2} \sqrt{\left| \int_0^T \int_0^T [r_\theta(t-s)^2 - \gamma_\theta(t,s)^2] \, \mathrm{d}s \mathrm{d}t \right|} \\
&= \sqrt{2} \sqrt{\left| \int_0^T \int_0^T \left(r_\theta(t-s) + \gamma_\theta(t,s) \right) \left(r_\theta(t-s) - \gamma_\theta(t,s) \right) \mathrm{d}s \mathrm{d}t \right|}.
\end{aligned}
$$

By applying Proposition 2.2 to the estimate above, we obtain

$$T\left|\sqrt{w_\theta(T)} - \sqrt{\mathbf{E}[(Q_T^\theta)^2]}\right|$$

$$\leq \quad C_\theta\sqrt{\left|\int_0^T\int_0^T \big(r_\theta(t-s) + \gamma_\theta(t,s)\big)\mathrm{e}^{-\theta\min(s,t)}\,\mathrm{d}s\mathrm{d}t\right|}.$$

Now, $|r_\theta(t-s)| \leq r_\theta(0)$ and $|\gamma_\theta(t,s)| \leq r_\theta(0) + 1$, by Proposition 2.2. Consequently, the integral above is bounded, and the proof is finished. $\quad\square$

Proof of Lemma 4.1. Let

$$a(t) = a_{H,1}(t) = H\mathrm{e}^{t/H}.$$

Then

$$r_{H,K,\theta}(t) = \frac{1}{2^K}\mathrm{e}^{-\theta t}\left[\big(a(t)^{2H}+1\big)^K - \big(a(t)-1\big)^{2HK}\right].$$

By the Taylor's theorem

$$\begin{aligned}
\big(a(t)^{2H}+1\big)^K &= a(t)^{2HK} + K\xi(t)^{K-1}, \\
\big(a(t)-1\big)^{2HK} &= a(t)^{2HK} - 2HK\eta(t)^{2HK-1},
\end{aligned}$$

for some $\xi(t) \in [a(t), a(t)+1]$ and $\eta(t) \in [a(t)-1, a(t)]$. Consequently,

$$\begin{aligned}
r_{H,K,\theta}(t) &\sim C_{H,K,\theta}\mathrm{e}^{-\theta t}\left[a(t)^{K-1} + a(t)^{2HK-1}\right] \\
&\sim C_{H,K,\theta}\mathrm{e}^{-\theta t\max\left\{\frac{1}{HK}-1, 1+\frac{1}{HK}-\frac{1}{H}\right\}},
\end{aligned}$$

which shows the exponential decay. $\quad\square$

References

[1] E. AZMOODEH AND J. I. MORLANES, *Drift parameter estimation for fractional Ornstein-Uhlenbeck process of the second kind*, Statistics, 49 (2015), pp. 1–18.

[2] E. AZMOODEH, T. SOTTINEN, L. VIITASAARI, AND A. YAZIGI, *Necessary and sufficient conditions for Hölder continuity of Gaussian processes*, Statist. Probab. Lett., 94 (2014), pp. 230–235.

[3] E. AZMOODEH AND L. VIITASAARI, *Parameter estimation based on discrete observations of fractional Ornstein-Uhlenbeck process of the second kind*, Stat. Inference Stoch. Process., 18 (2015), pp. 205–227.

[4] O. E. BARNDORFF-NIELSEN AND A. BASSE-O'CONNOR, *Quasi Ornstein-Uhlenbeck processes*, Bernoulli, 17 (2011), pp. 916–941.

[5] Y. BELYAEV, *Local properties of the sample functions of stationary Gaussian processes.*, Teor. Veroyatn. Primen., 5 (1960), pp. 128–131.

[6] F. BIAGINI, Y. HU, B. ØKSENDAL, AND T. ZHANG, *Stochastic calculus for fractional Brownian motion and applications*, Probability and its Applications (New York), Springer-Verlag London, Ltd., London, 2008.

[7] J.-C. BRETON AND I. NOURDIN, *Error bounds on the non-normal approximation of Hermite power variations of fractional Brownian motion*, Electron. Commun. Probab., 13 (2008), pp. 482–493.

[8] P. CHERIDITO, H. KAWAGUCHI, AND M. MAEJIMA, *Fractional Ornstein-Uhlenbeck processes*, Electron. J. Probab., 8 (2003), pp. no. 3, 14 pp. (electronic).

[9] J. L. DOOB, *The Brownian movement and stochastic equations*, Ann. of Math. (2), 43 (1942), pp. 351–369.

[10] K. ES-SEBAIY AND D. NDIAYE, *On drift estimation for non-ergodic fractional Ornstein-Uhlenbeck process with discrete observations*, Afr. Stat., 9 (2014), pp. 615–625.

[11] K. ES-SEBAIY AND C. A. TUDOR, *Fractional Ornstein-Uhlenbeck processes mixed with a gamma distribution*, Fractals, 23 (2015), pp. 1550032, 10.

[12] C. F. GAUSS, *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*, Cambridge Library Collection, Cambridge University Press, Cambridge, 2011. Reprint of the 1809 original.

[13] U. GRENANDER, *Stochastic processes and statistical inference*, Ark. Mat., 1 (1950), pp. 195–277.

[14] C. HOUDRÉ AND J. VILLA, *An example of infinite dimensional quasi-helix*, in Stochastic models (Mexico City, 2002), vol. 336 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2003, pp. 195–201.

[15] Y. HU AND D. NUALART, *Parameter estimation for fractional Ornstein-Uhlenbeck processes*, Statist. Probab. Lett., 80 (2010), pp. 1030–1038.

[16] Y. HU AND J. SONG, *Parameter estimation for fractional Ornstein-Uhlenbeck processes with discrete observations*, in Malliavin calculus and stochastic analysis, vol. 34 of Springer Proc. Math. Stat., Springer, New York, 2013, pp. 427–442.

[17] T. KAARAKKA AND P. SALMINEN, *On fractional Ornstein-Uhlenbeck processes*, Commun. Stoch. Anal., 5 (2011), pp. 121–133.

[18] M. L. Kleptsyna and A. Le Breton, *Statistical analysis of the fractional Ornstein-Uhlenbeck type process*, Stat. Inference Stoch. Process., 5 (2002), pp. 229–248.

[19] Y. Kozachenko, A. Melnikov, and Y. Mishura, *On drift parameter estimation in models with fractional Brownian motion*, Statistics, 49 (2015), pp. 35–62.

[20] K. Kubilius, Y. Mishura, K. Ralchenko, and O. Seleznjev, *Consistency of the drift parameter estimator for the discretized fractional Ornstein-Uhlenbeck process with Hurst index $H \in (0, \frac{1}{2})$*, Electron. J. Stat., 9 (2015), pp. 1799–1825.

[21] J. Lamperti, *Semi-stable stochastic processes*, Trans. Amer. Math. Soc., 104 (1962), pp. 62–78.

[22] P. Langevin, *Sur la théorie du mouvement brownien*, C. R. Acad. Sci. Paris, 146 (1908), pp. 530–533.

[23] R. S. Liptser and A. N. Shiryaev, *Statistics of random processes. II*, vol. 6 of Applications of Mathematics (New York), Springer-Verlag, Berlin, expanded ed., 2001. Applications, Translated from the 1974 Russian original by A. B. Aries, Stochastic Modelling and Applied Probability.

[24] G. Maruyama, *The harmonic analysis of stationary stochastic processes*, Mem. Fac. Sci. Kyūsyū Univ. A., 4 (1949), pp. 45–106.

[25] Y. S. Mishura, *Stochastic calculus for fractional Brownian motion and related processes*, vol. 1929 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2008.

[26] I. Nourdin and G. Peccati, *Normal approximations with Malliavin calculus: From Stein's method to universality*, vol. 192 of Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 2012.

[27] D. Nualart, *The Malliavin calculus and related topics*, Probability and its Applications (New York), Springer-Verlag, Berlin, second ed., 2006.

[28] G. Peccati and M. S. Taqqu, *Wiener chaos: moments, cumulants and diagrams*, vol. 1 of Bocconi & Springer Series, Springer, Milan; Bocconi University Press, Milan, 2011. A survey with computer implementation, Supplementary material available online.

[29] F. Russo and C. A. Tudor, *On bifractional Brownian motion*, Stochastic Process. Appl., 116 (2006), pp. 830–856.

[30] G. Shen, X. Yin, and L. Yan, *Least squares estimation for Ornstein–Uhlenbeck processes driven by the weighted fractional Brownian motion*, Acta Math. Sci. Ser. B Engl. Ed., 36 (2016), pp. 394–408.

[31] L. Shen and Q. Xu, *Asymptotic law of limit distribution for fractional Ornstein-Uhlenbeck process*, Adv. Difference Equ., (2014), pp. 2014:75, 7.

[32] T. Sottinen and C. A. Tudor, *Parameter estimation for stochastic equations with additive fractional Brownian sheet*, Stat. Inference Stoch. Process., 11 (2008), pp. 221–236.

[33] T. Sottinen and L. Viitasaari, *Stochastic Analysis of Gaussian Processes via Fredholm Representation*, Int. J. Stoch. Anal., (2016), pp. Art. ID 8694365, 15.

[34] T. Sottinen and A. Yazigi, *Generalized Gaussian bridges*, Stochastic Process. Appl., 124 (2014), pp. 3084–3105.

[35] X. Sun and F. Guo, *On integration by parts formula and characterization of fractional Ornstein–Uhlenbeck process*, Statist. Probab. Lett., 107 (2015), pp. 170–177.

[36] K. Tanaka, *Maximum likelihood estimation for the non-ergodic fractional Ornstein-Uhlenbeck process*, Stat. Inference Stoch. Process., 18 (2015), pp. 315–332.

[37] G. E. Uhlenbeck and L. S. Ornstein, *On the theory of the brownian motion*, Phys. Rev., 36 (1930), pp. 823–841.

[38] L. Viitasaari, *Representation of stationary and stationary increment processes via Langevin equation and self-similar processes*, Statist. Probab. Lett., 115 (2016), pp. 45–53.

[39] A. Yazigi, *Representation of self-similar Gaussian processes*, Statist. Probab. Lett., 99 (2015), pp. 94–100.