

Berry-Esseen bounds of second moment estimators for Gaussian processes observed at high frequency*

Soukaina Douissi

National School of Applied Sciences, Marrakech, Morocco
e-mail: s.douissi@uca.ma

Khalifa Es-Sebaiy

Department of Mathematics, Faculty of Science, Kuwait University, Kuwait
e-mail: khalifa.essebaiy@ku.edu.kw

George Kerchev and Ivan Nourdin

University of Luxembourg, Department of Mathematics, Luxembourg
e-mail: gkerchev@gmail.com; ivan.nourdin@uni.lu

Abstract: Let $Z := \{Z_t, t \geq 0\}$ be a stationary Gaussian process. We study two estimators of $\mathbb{E}[Z_0^2]$, namely $\hat{f}_T(Z) := \frac{1}{T} \int_0^T Z_t^2 dt$, and $\tilde{f}_n(Z) := \frac{1}{n} \sum_{i=1}^n Z_{t_i}^2$, where $t_i = i\Delta_n$, $i = 0, 1, \dots, n$, $\Delta_n \rightarrow 0$ and $T_n := n\Delta_n \rightarrow \infty$. We prove that the two estimators are strongly consistent and establish Berry-Esseen bounds for a central limit theorem involving $\hat{f}_T(Z)$ and $\tilde{f}_n(Z)$. We apply these results to asymptotically stationary Gaussian processes and estimate the drift parameter for Gaussian Ornstein-Uhlenbeck processes.

MSC2020 subject classifications: Primary 60F05; secondary 60G15, 60G10, 62F12, 62M09.

Keywords and phrases: Parameter estimation, strong consistency, rate of normal convergence of the estimators, stationary Gaussian processes, continuous-time observation, high frequency data.

Received February 2021.

Contents

1	Introduction	637
2	Elements of Malliavin calculus on Wiener space	640
3	Parameter estimation for stationary Gaussian processes	643
3.1	Continuous-time observations	643
3.2	Proof of Theorem 1.1.	645
3.3	Discrete-time observations	647

arXiv: [2102.04810](https://arxiv.org/abs/2102.04810)

*G. Kerchev and I. Nourdin were supported by the FNR grant APOGEe (R-AGR-3585-10) at Luxembourg University.

3.4 Proof of Theorem 1.2 649
 4 Parameter estimation for non-stationary Gaussian processes 651
 5 Applications to Gaussian Ornstein-Uhlenbeck processes 654
 5.1 Fractional Ornstein-Uhlenbeck process of the first kind 656
 5.2 Fractional Ornstein-Uhlenbeck process of the second kind 658
 6 Technical results 660
 6.1 Two technical lemmas 660
 6.2 A bound on $\mathbb{E}|\widehat{f}_T(X)|^p$ 663
 Acknowledgments 667
 References 668

1. Introduction

Statistical inference for stochastic processes is of great importance for theoreticians and practitioners alike. While for some processes like Itô-type diffusions and semimartingales, there is extensive literature, the statistical analysis for fractional Gaussian processes is relatively recent.

In this paper, we are interested in the parametric estimation of the variance of stationary Gaussian process which is not necessarily a semimartingale. Let $Z = \{Z_t, t \geq 0\}$ be a continuous centered stationary Gaussian process and $f_Z := E(Z_0^2) > 0$. We consider the following estimators of f_Z :

- When a complete path of the process over a large finite interval is observable, we use the estimator:

$$\widehat{f}_T(Z) := \frac{1}{T} \int_0^T Z_t^2 dt, \quad T > 0. \tag{1.1}$$

- A more practical assumption is that the process Z is observed at discrete time instants $t_i = i\Delta_n$, where $i = 0, \dots, n$ and Δ_n is the step size. Then we consider the following estimator over the *observation window* $T_n := n\Delta_n$:

$$\widetilde{f}_n(Z) := \frac{1}{n} \sum_{i=1}^n Z_{t_i}^2, \quad n \geq 1. \tag{1.2}$$

These estimators are unbiased and we show that they are strongly consistent and admit a central limit theorem. Moreover, we bound the rate of convergence to the normal distribution in terms of total variation distance and Wasserstein distance. Recall that, for two random variables X and Y , the former metrics are respectively given by

$$d_{TV}(X, Y) := \sup_{A \in \mathcal{B}(\mathbb{R})} |\mathbb{P}[X \in A] - \mathbb{P}[Y \in A]|, \tag{1.3}$$

where the supremum is over all Borel sets, and

$$d_W(X, Y) := \sup_{f \in Lip(1)} |\mathbb{E}[f(X)] - \mathbb{E}[f(Y)]|, \tag{1.4}$$

where $Lip(1)$ is the set of all Lipschitz functions with Lipschitz constant ≤ 1 .

Let $\rho(t) = \rho(-t) := \mathbb{E}[Z_0 Z_t]$ for $t \geq 0$. The central result for $\widehat{f}_T(Z)$, whose proof follows the lines of the approach developed in [28, Chapter 7], is the following.

Theorem 1.1. *Assume $\int_{\mathbb{R}} \rho^2(r)dr < \infty$. Let $\mathcal{N} \sim \mathcal{N}(0, 1)$ be the standard normal random variable. Then for all $T > 0$,*

$$d_{TV} \left(\frac{\widehat{f}_T(Z) - f_Z}{\sqrt{\text{Var}(\widehat{f}_T(Z) - f_Z)}}, \mathcal{N} \right) \leq \varphi_T(Z), \tag{1.5}$$

where

$$\varphi_T(Z) = C \max \left\{ \frac{8}{\sqrt{T}} \left(\int_{-T}^T |\rho(t)|^{3/2} dt \right)^2, \frac{48}{T} \left(\int_{-T}^T |\rho(t)|^{4/3} dt \right)^3 \right\}, \tag{1.6}$$

for some absolute constant $C > 0$. The same result holds for the Wasserstein distance.

For the discrete estimator $\widetilde{f}_n(Z)$ we have that:

Theorem 1.2. *Assume $\int_{\mathbb{R}} \rho^2(r)dr < \infty$ and that*

$$\mathbb{E}[|Z_t - Z_s|^2] \leq c|t - s|^{2\alpha}$$

for some $c > 0$ and $\alpha \in (0, 1)$ and when $|t - s|$ is small enough. Let $\mathcal{N} \sim \mathcal{N}(0, 1)$ be the standard normal random variable. If $\Delta_n \rightarrow 0$ and $n\Delta_n \rightarrow \infty$ as $n \rightarrow \infty$, then there is $C > 0$ such that, for every $n \geq 1$,

$$\begin{aligned} & d_{TV} \left(\frac{\widetilde{f}_n(Z) - f_Z}{\sqrt{\text{Var}(\widetilde{f}_n(Z) - f_Z)}}, \mathcal{N} \right) \\ & \leq \varphi_{T_n}(Z) + 2 \left| 1 - \frac{\text{Var}(\widehat{f}_{T_n}(Z) - f_Z)}{\text{Var}(\widetilde{f}_n(Z) - f_Z)} \right| + C [n\Delta_n^{2\alpha+1}]^{1/4}, \end{aligned}$$

where $\varphi_{T_n}(Z)$ satisfies (1.6). The same result holds for the Wasserstein distance.

Thanks to the robustness of our approach, we can extend Theorem 1.1 and Theorem 1.2 to the case when the process is only an asymptotically stationary Gaussian process. In particular we prove the rate of convergence of the second moment estimators for $X := Z + Y$, where Z is the stationary Gaussian process as above and Y is a stochastic process with $\|Y_t\|_{L^1} = \mathcal{O}(t^{-\gamma})$ for some absolute $\gamma > 1$. The bounds will be the same up to an extra term $CT^{(1-\gamma)/2}$ (or $CT_n^{(1-\gamma)/2}$ for the discrete case).

Moreover, we explore applications to drift estimation for the Ornstein-Uhlenbeck process. Let $X^\theta = (X_t^\theta)_{t \geq 0}$ be an ergodic type Gaussian Ornstein-Uhlenbeck process given by the differential equation

$$dX_t^\theta = -\theta X_t^\theta dt + dG_t, \quad X_0^\theta = 0,$$

where θ is the drift parameter and $(G_t)_{t \geq 0}$ is an arbitrary mean-zero Gaussian process. One can show that X^θ is asymptotically stationary and write $X^\theta = Z^\theta + Y^\theta$ where Z^θ is a stationary Gaussian process and $\|Y_t\|_{L^1} = \mathcal{O}(t^{-\gamma})$ for any $\gamma > 1$. Then we provide Berry-Esseen bounds for the estimators $\hat{\theta}_T := g_{Z^\theta}(\hat{f}_T(X^\theta))$ and $\tilde{\theta}_n := g_{Z^\theta}(\tilde{f}_n(X^\theta))$. The function g_{Z^θ} is given via $g_{Z^\theta}^{-1}(\theta) = \mathbb{E}[(Z_0^\theta)^2]$ and \hat{f}_T, \tilde{f}_n are as in (1.1) and (1.2). Concrete bounds are computed for the cases when the process X^θ is of the first and second kind, i.e., when $(G_t)_{t \geq 0}$ is a particular Gaussian process, following the terminology in [22].

Parameter estimation for stationary Gaussian process is usually done via the Maximum likelihood estimator because of its asymptotic optimality, see [35] and [34]. For instance, the MLE estimator of a stationary ARMA process is strongly consistent and asymptotically efficient [5, Section 10.8]. The method of moments is more computationally tractable especially when one considers discrete estimators. Some recent studies include [11, 17, 21] where the mesh in time $\Delta_n = 1$ in (1.2), which is akin to the discretization of a least-squares method for fractional Gaussian processes using fixed-time-step observations. See also [8] for an application of the second moment method to an AR(1) model.

In the last few years the estimators (1.1) and (1.2) have been used, in a number of instance, to study parameter estimation problems in various fractional Gaussian models. Some of these results can be summarized below.

- *The case of continuous-time observations for ergodic-type Gaussian processes, using (1.1):* The work [36] derived a central limit theorem and a Berry-Esseen bound in Kolmogorov distance for the second moment estimator (1.1) of the limiting variance of an Ornstein-Uhlenbeck (OU) process driven by stationary-increment Gaussian noise. In [20], the authors considered the estimator (1.1) (called “Alternative estimator” there) to estimate the drift parameter of an OU process driven by fBm with Hurst parameter $H \in (0, 1)$. They proved a central limit theorem when $H \in (0, 3/4]$ and a noncentral limit theorem for $H \in (3/4, 1)$. However, they did not give speed of convergence for these limit theorems. Their approach used a least squares estimator (LSE) to study the second moment estimator. Berry-Esseen bounds for a central limit theorem with the Kolmogorov distance were given first in [14] and then further improved upon in [18] when $H \in (1/2, 3/4]$.
Moreover, drift parameter estimation using (1.1) was employed for an OU process driven by a Hermite process in [30] and driven by an α -stable Lévy motion in [7]. On the other hand, the consistency and speed of convergence in the TV and Wasserstein norms for the estimator (1.1) of the drift parameter in infinite dimensional linear stochastic equations driven by a fBm are studied by [26].
- *The case of discrete-time observations for ergodic-type Gaussian processes, using (1.2):* In the case when the mesh in time $\Delta_n = 1$, the consistency and speed of convergence in the TV and Wasserstein distance for the estimator (1.2) of the limiting variance of of general Gaussian sequences

were recently developed in the papers [17, 11]. Also, the drift parameter in linear stochastic evolution equation driven by a fBm is considered in [26]. On the other hand, in the case of high frequency data corresponding to $\Delta_n \rightarrow 0$ in (1.2), the statistical inference for several ergodic-type fractional Ornstein-Uhlenbeck (fOU) models, using the estimator (1.2), was recently studied in the papers in [13, 20, 36]. However, these papers did not provide speed of convergence for the asymptotic distribution of (1.2).

For the drift parameter estimation for non-ergodic fractional-noise-driven Ornstein-Uhlenbeck processes we refer the interested readers to [12, 15, 1, 16] and references therein.

For our proofs we employ tools from the analysis on Wiener space, including Wiener chaos calculus and Malliavin calculus. The main theorem regarding normal approximations is the so-called Optimal Fourth Moment Theorem due to Nourdin and Peccati [29]. A review of these tools and results can be found in Section 2. Then the proofs for the stationary and asymptotically stationary Gaussian case are outlined in Sections 3 and 4 respectively. The application to drift estimation for Ornstein-Uhlenbeck processes is carried out in Section 5. Some of our more technical calculations can be seen in Section 6; note that some of results in this section, for instance Proposition 6.3, are of independent interest and use novel techniques.

2. Elements of Malliavin calculus on Wiener space

This section gives a brief overview of some useful facts from the Malliavin calculus on Wiener space. Some of the results presented here are essential for the proofs in the present paper. For our purposes we focus on special cases that are relevant for our setting and omit the general high-level theory. We direct the interested reader to [32, Chapter 1] and [28, Chapter 2].

Fix $(\Omega, \mathcal{F}, \mathbb{P})$ for the Wiener space of a standard Wiener process $W = (W_t)_{t \geq 0}$. The first step is to identify the general centered Gaussian process $(Z_t)_{t \geq 0}$ with an *isonormal Gaussian process* $X = \{X(h), h \in \mathcal{H}\}$ for some Hilbert space \mathcal{H} . Recall that for such processes X , for every $h_1, h_2 \in \mathcal{H}$, one has $\mathbb{E}[X(h_1)X(h_2)] = \langle h_1, h_2 \rangle_{\mathcal{H}}$.

One can define \mathcal{H} as the closure of real-valued step functions on $[0, \infty)$ with respect to the inner product $\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \rangle_{\mathcal{H}} = \mathbb{E}[Z_t Z_s]$. Then the isonormal process X is given by Wiener integral $X(h) := \int_{\mathbb{R}_+} h(s) dW_s$. Note, that, in particular $X(\mathbf{1}_{[0,t]}) \stackrel{d}{=} Z_t$.

The next step involves the *multiple Wiener-Itô integrals*. The formal definition involves the concepts of Malliavin derivative and divergence. We refer the reader to [32, Chapter 1] and [28, Chapter 2]. For our purposes we define the multiple Wiener-Itô integral I_p via the Hermite polynomials H_p . In particular, for $h \in \mathcal{H}$ with $\|h\|_{\mathcal{H}} = 1$, and any $p \geq 1$,

$$H_p(X(h)) = I_p(h^{\otimes p}).$$

For $p = 1$ and $p = 2$ we have the following:

$$H_1(X(\mathbf{1}_{[0,t]})) = X(\mathbf{1}_{[0,t]}) = I_1(\mathbf{1}_{[0,t]}) = Z_t \tag{2.1}$$

$$H_2(X(\mathbf{1}_{[0,t]})) = X(\mathbf{1}_{[0,t]})^2 - \mathbb{E}[X(\mathbf{1}_{[0,t]})^2] = I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) = Z_t^2 - \mathbb{E}[Z_t^2]. \tag{2.2}$$

Note also that I_0 can be taken to be the identity operator.

Remark 2.1. Some notation for Hilbert spaces. Let \mathcal{H} be a Hilbert space. Given an integer $q \geq 2$ the Hilbert spaces $\mathcal{H}^{\otimes q}$ and $\mathcal{H}^{\odot q}$ correspond to the q th tensor product and q th symmetric tensor product of \mathcal{H} . If $f \in \mathcal{H}^{\otimes q}$ is given by $f = \sum_{j_1, \dots, j_q} a(j_1, \dots, j_q) e_{j_1} \otimes \dots \otimes e_{j_q}$, where $(e_{j_i})_{i \in [1, q]}$ form an orthonormal basis of $\mathcal{H}^{\otimes q}$, then the symmetrization \tilde{f} is given by

$$\tilde{f} = \frac{1}{q!} \sum_{\sigma} \sum_{j_1, \dots, j_q} a(j_1, \dots, j_q) e_{\sigma(j_1)} \otimes \dots \otimes e_{\sigma(j_q)},$$

where the first sum runs over all permutations σ of $\{1, \dots, q\}$. Then \tilde{f} is an element of $\mathcal{H}^{\odot q}$. We also make use of the concept of contraction. The r th contraction of two tensor products $e_{j_1} \otimes \dots \otimes e_{j_p}$ and $e_{k_1} \otimes \dots \otimes e_{k_q}$ is an element of $\mathcal{H}^{\otimes(p+q-2r)}$ given by

$$\begin{aligned} & (e_{j_1} \otimes \dots \otimes e_{j_p}) \otimes_r (e_{k_1} \otimes \dots \otimes e_{k_q}) \\ &= \left[\prod_{\ell=1}^r \langle e_{j_\ell}, e_{k_\ell} \rangle \right] e_{j_{r+1}} \otimes \dots \otimes e_{j_q} \otimes e_{k_{r+1}} \otimes \dots \otimes e_{k_q}. \end{aligned} \tag{2.3}$$

The main motivation for introducing the multiple integrals comes from the following properties:

- *Isometry property of integrals* [28, Proposition 2.7.5]. Fix integers $p, q \geq 1$ as well as $f \in \mathcal{H}^{\odot p}$ and $g \in \mathcal{H}^{\odot q}$.

$$\mathbb{E}[I_p(f)I_q(g)] = \begin{cases} p! \langle f, g \rangle_{\mathcal{H}^{\otimes p}} & \text{if } p = q \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

- *Product formula* [28, Proposition 2.7.10]. Let $p, q \geq 1$. If $f \in \mathcal{H}^{\odot p}$ and $g \in \mathcal{H}^{\odot q}$ then

$$I_p(f)I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}(f \tilde{\otimes}_r g). \tag{2.5}$$

- *Hypercontractivity in Wiener Chaos.* For every $q \geq 1$, \mathcal{H}_q denotes the q th Wiener chaos of W , defined as the closed linear subspace of $L^2(\Omega)$ generated by the random variables $\{H_q(W(h)), h \in \mathcal{H}, \|h\|_{\mathcal{H}} = 1\}$ where H_q is the q th Hermite polynomial. For any $F \in \oplus_{l=1}^q \mathcal{H}_l$ (i.e. in a fixed sum of Wiener chaoses), we have

$$(\mathbb{E}[|F|^p])^{1/p} \leq c_{p,q} (\mathbb{E}[|F|^2])^{1/2} \quad \text{for any } p \geq 2. \tag{2.6}$$

It should be noted that the constants $c_{p,q}$ above are known with some precision when F is a single chaos term: indeed, by [28, Corollary 2.8.14], $c_{p,q} = (p-1)^{q/2}$.

The second part of important results we borrow from Malliavin calculus concerns estimates on the distance between random variables. There are two key estimates linking total variation distance and the Malliavin calculus, which were both obtained by Nourdin and Peccati. The first one is an observation relating an integration-by-parts formula on Wiener space with a classical result of Ch. Stein. The second is a quantitatively sharp version of the famous fourth moment theorem of Nualart and Peccati.

Let \mathcal{N} denote the standard normal law. For each integer n , let $F_n \in \mathcal{H}_q$. Assume $\text{Var}[F_n] = 1$ and $(F_n)_n$ converges in distribution to a normal law. It is known (the *fourth moment theorem* in [33]) that this convergence is equivalent to $\lim_n \mathbb{E}[F_n^4] = 3$. The following optimal estimate for $d_{TV}(F_n, \mathcal{N})$, known as the optimal fourth moment theorem, was proved in [29]: with the sequence F as above, assuming convergence, there exist two constants $c, C > 0$ depending only on the type of the process F but not on n , such that

$$c \max \{ \mathbb{E}[F_n^4] - 3, |\mathbb{E}[F_n^3]| \} \leq d_{TV}(F_n, \mathcal{N}) \leq C \max \{ \mathbb{E}[F_n^4] - 3, |\mathbb{E}[F_n^3]| \}. \quad (2.7)$$

Recall that for a standardized random variable F , i.e., with $\mathbb{E}[F] = 0$ and $\mathbb{E}[F^2] = 1$, the third and fourth cumulants are respectively

$$\begin{aligned} \kappa_3(F) &:= \mathbb{E}[F^3], \\ \kappa_4(F) &:= \mathbb{E}[F^4] - 3. \end{aligned}$$

Throughout the paper we use the notation $\mathcal{N} \sim \mathcal{N}(0, 1)$. We also use the notation C for any positive real constant, independently of its value which may change from line to line when this does not lead to ambiguity.

Remark 2.2. We note that the optimal bound (2.7) holds with d_{TV} replaced by d_W . Indeed, by [28, Theorem 3.5.2]:

$$d_W(F, \mathcal{N}) \leq \sup_{\varphi \in \mathcal{F}} |\mathbb{E}[\varphi'(F)] - \mathbb{E}[F\varphi(F)]|$$

where \mathcal{F} is the set of C^1 functions φ such that $|\varphi'|_\infty \leq \sqrt{2/\pi}$. Now using the concepts of the Malliavin derivative operator D and Ornstein-Uhlenbeck generator L , see [27, Theorem 4.15], one has $|\mathbb{E}[\varphi'(F)] - \mathbb{E}[F\varphi(F)]| = |\mathbb{E}[\varphi'(F)\mathbb{E}[(1 - \langle DF, -DL^{-1}F \rangle | \mathcal{F})]]|$, and thus:

$$d_W(F, \mathcal{N}) \leq \sqrt{2/\pi} \mathbb{E}|\mathbb{E}[1 - \langle DF, -DL^{-1}F \rangle | \mathcal{F}]|.$$

However, this is the same bound one has (up to the constant $\sqrt{2/\pi}$) in the proof of the optimal fourth moment theorem [29, Proof of Theorem 1.2].

3. Parameter estimation for stationary Gaussian processes

In this section we present a general framework for the parameter estimation of the variance of a stationary Gaussian process. We prove the consistency and provide upper bounds in the total variation and Wasserstein distances for the rate of normal convergence of the MC estimators (1.1) and (1.2).

Let $Z := \{Z_t, t \geq 0\}$ be a continuous centered stationary Gaussian process that can be represented as a Wiener-Itô (multiple) integral $Z_t = I_1(\mathbf{1}_{[0,t]})$ for every $t \geq 0$, as in (2.1). Let $\rho(r) = E(Z_r Z_0)$ denote the covariance of Z for every $r \geq 0$, and let $\rho(r) = \rho(-r)$ for all $r < 0$. Our main assumption throughout the paper is that

$$\sigma_Z^2 := 4 \int_{\mathbb{R}} \rho^2(r) dr < \infty.$$

3.1. Continuous-time observations

We estimate the variance $f_Z := \mathbb{E}(Z_0^2)$, when the whole trajectory of Z is observed up to time $T > 0$. We consider the estimator (1.1) given by

$$\hat{f}_T(Z) = \frac{1}{T} \int_0^T Z_t^2 dt, \quad T > 0$$

as a statistic to estimate f_Z , based on the continuous-time observation of Z . Our goal is to establish Theorem 1.1, i.e., a Berry-Esseen bound on the convergence of $\hat{f}_T(Z) - f_Z$. First we show some simpler properties of $\hat{f}_T(Z)$.

Lemma 3.1. *The estimator $\hat{f}_T(Z)$ is unbiased and strongly consistent. In particular,*

$$\sqrt{T} \left\| \hat{f}_T(Z) - f_Z \right\|_{L^2} \uparrow \sigma_Z \text{ as } T \rightarrow \infty. \tag{3.1}$$

Proof. By stationarity, $\mathbb{E}[Z_0^2] = \mathbb{E}[Z_s^2]$ for every $s \geq 0$ and thus $\mathbb{E}[\hat{f}_T(Z)] = f_Z$, so the estimator $\hat{f}_T(Z)$ is unbiased.

The next step is to show that the estimator $\hat{f}_T(Z)$ is strongly consistent, i.e.,

$$\hat{f}_T(Z) \rightarrow f_Z \text{ almost surely as } T \rightarrow \infty.$$

Let

$$V_T(Z) := \sqrt{T} \left(\hat{f}_T(Z) - f_Z \right) = \frac{1}{\sqrt{T}} \int_0^T (Z_t^2 - E[Z_t^2]) dt = \frac{1}{\sqrt{T}} \int_0^T I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) dt, \tag{3.2}$$

where we have used (2.2). Then

$$\mathbb{E}[V_T(Z)^2] = \frac{1}{T} \mathbb{E} \left[\left(\int_0^T I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) dt \right)^2 \right] = \frac{1}{T} \int_{[0,T]^2} \mathbb{E} \left[I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) I_2(\mathbf{1}_{[0,s]}^{\otimes 2}) \right] dt ds. \tag{3.3}$$

By (2.4),

$$\mathbb{E} \left[I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) I_2(\mathbf{1}_{[0,s]}^{\otimes 2}) \right] = 2! \langle \mathbf{1}_{[0,t]}^{\otimes 2}, \mathbf{1}_{[0,s]}^{\otimes 2} \rangle_{\mathcal{H}^{\otimes 2}} = 2 (\mathbb{E}[Z_t Z_s])^2 = 2\rho(t-s)^2.$$

Therefore,

$$\begin{aligned} \mathbb{E}[V_T(Z)^2] &= \frac{2}{T} \int_{[0,T]^2} \rho(t-s)^2 dt ds = \frac{4}{T} \int_0^T \int_u^T \rho(u)^2 dt du \\ &= 4 \int_0^\infty \left(1 - \frac{u}{T}\right)_+ \rho(u)^2 \leq 4 \int_0^T \rho(u)^2 du < \infty, \end{aligned} \tag{3.4}$$

where $(x)_+ = \max\{x, 0\}$. Note that the above also implies that $\mathbb{E}[V_T(Z)^2] \uparrow \sigma_Z^2$, establishing (3.1). Alternatively, $\|\widehat{f}_T(Z) - f_Z\|_{L^2} \leq \sigma_Z/\sqrt{T}$. In particular, this shows that $\widehat{f}_T(Z)$ converges to f_Z in L^2 . At this point we recall [25, Lemma 2.1]):

Lemma 3.2. *Let $\gamma > 0$. Let $(Z_n)_{n \in \mathbb{N}}$ be a sequence of random variables. If for every $p \geq 1$ there exists a constant $c_p > 0$ such that for all $n \in \mathbb{N}$,*

$$\|Z_n\|_{L^p(\Omega)} \leq c_p \cdot n^{-\gamma},$$

then for all $\varepsilon > 0$ there exists a random variable α_ε which is almost surely finite such that

$$|Z_n| \leq \alpha_\varepsilon \cdot n^{-\gamma+\varepsilon} \quad \text{almost surely}$$

for all $n \in \mathbb{N}$. Moreover, $\mathbb{E}|\alpha_\varepsilon|^p < \infty$ for all $p \geq 1$.

We can apply Lemma 3.2 as soon as $\|\widehat{f}_T(Z) - f_Z\|_{L^p} \leq C_p/\sqrt{T}$ for every $p \geq 1$. We have shown that for $p = 2$. However, using the inequality between L^p norms and L^q norms, one has that the same bound holds for $p \in [1, 2)$. Finally, one can apply the hypercontractivity property (2.6) in Wiener chaos to get the result for all $p \geq 2$.

Therefore, $\widehat{f}_n(Z)$ converges almost surely to f_Z as $n \rightarrow \infty$ (and $n \in \mathbb{N}$). We now use the following more technical result. While it is of course not new, we could not find a reference where it was stated exactly as we needed it to be. This is why we give its proof in Section 6 for the sake of completeness.

Lemma 3.3. *Let $\{u_t, t \geq 0\}$ be a continuous stochastic process such that for any $p \geq 1$, there is a positive constant $C_p > 0$ such that $\sup_{t \geq 0} \mathbb{E}[|u_t|^p] < C_p$. In addition, we assume*

$$\frac{1}{n} \int_0^n u_t dt \longrightarrow 0 \quad \text{almost surely as } n \rightarrow \infty.$$

Then,

$$\frac{1}{T} \int_0^T u_t dt \longrightarrow 0 \quad \text{almost surely as } T \rightarrow \infty.$$

Note that by stationarity $\sup_{t \geq 0} \mathbb{E}[|Z_t|^p] = \mathbb{E}[|Z_0|^p] < C_p$ for some $C_p > 0$. Therefore, by Lemma 3.3, $\widehat{f}_T(Z)$ converges almost surely to f_Z as $T \rightarrow \infty$ (and $T \in \mathbb{R}_+$). \square

Now, we turn to the proof of Theorem 1.1.

3.2. Proof of Theorem 1.1.

The random variable

$$\frac{\widehat{f}_T(Z) - f_Z}{\sqrt{\text{Var}(\widehat{f}_T(Z) - f_Z)}} = \frac{V_T(Z)}{\sqrt{\mathbb{E}[V_T(Z)^2]}}$$

is centered and normalized. Then the fourth moment theorem (2.7) applies and then

$$\begin{aligned} d_{TV} \left(\frac{\widehat{f}_T(Z) - f_Z}{\sqrt{\text{Var}(\widehat{f}_T(Z) - f_Z)}}, \mathcal{N} \right) &= d_{TV} \left(\frac{V_T(Z)}{\sqrt{\mathbb{E}[V_T(Z)^2]}}, \mathcal{N} \right) \\ &\leq C \max \left\{ \frac{\kappa_3(V_T(Z))}{\mathbb{E}[V_T(Z)^2]^{3/2}}, \frac{\kappa_4(V_T(Z))}{\mathbb{E}[V_T(Z)^2]^2} \right\}. \end{aligned} \tag{3.5}$$

We are left to study the third and fourth cumulants of $V_T(Z)$. The following technical result is a slight modification of [4, Propositions 6.3, 6.4]:

Lemma 3.4. *For every $T > 0$,*

$$|\kappa_3(V_T(Z))| \leq \frac{8}{\sqrt{T}} \left(\int_{-T}^T |\rho(t)|^{3/2} dt \right)^2, \tag{3.6}$$

$$|\kappa_4(V_T(Z))| \leq \frac{48}{T} \left(\int_{-T}^T |\rho(t)|^{4/3} dt \right)^3. \tag{3.7}$$

Proof. The proof follows the same approach as in [4] and is included in Section 6. We note that our proof is more detailed than the one in [4]. \square

The corresponding bound for the Wasserstein distance follows from Remark 2.2. Thus Theorem 1.1 is established. At this point we present two corollaries. First, we study the asymptotic behavior of the bound (1.6) under the additional assumption of the decay of correlations of Z_t . We have the following:

Corollary 3.5. Assume there exists $0 < \beta < \frac{3}{4}$ such that, $|\rho(t)| = \mathcal{O}(t^{2\beta-2})$. Then,

$$d_{TV} \left(\frac{\widehat{f}_T(Z) - f_Z}{\sqrt{\text{Var}(\widehat{f}_T(Z) - f_Z)}}, \mathcal{N} \right) \leq C \begin{cases} T^{-1/2} & \text{if } 0 < \beta < \frac{2}{3}, \\ \log^2(T)T^{-1/2} & \text{if } \beta = \frac{2}{3}, \\ T^{6\beta-\frac{9}{2}} & \text{if } \frac{2}{3} < \beta < \frac{3}{4}. \end{cases} \quad (3.8)$$

Proof. Note that there is a constant $C > 0$ such that $|\rho(s)| < C$ for $s \in [-1, 1]$. Then, since

$$\int_{-T}^T |\rho(t)|^p dt \leq C_p \left(1 + \int_1^T |\rho(t)|^p dt \right),$$

for $p = 3/2$ and $p = 4/3$, one can establish the following bounds on $\kappa_3(V_T(Z))$ and $\kappa_4(V_T(Z))$:

$$|\kappa_3(V_T(Z))| \leq C \begin{cases} T^{-1/2} & \text{if } 0 < \beta < \frac{2}{3}, \\ \log^2(T)T^{-1/2} & \text{if } \beta = \frac{2}{3}, \\ T^{6\beta-\frac{9}{2}} & \text{if } \frac{2}{3} < \beta < \frac{3}{4}, \end{cases} \quad (3.9)$$

and

$$|\kappa_4(V_T(Z))| \leq C \begin{cases} T^{-1} & \text{if } 0 < \beta < \frac{5}{8}, \\ \log^3(T)T^{-1} & \text{if } \beta = \frac{5}{8}, \\ T^{8\beta-6} & \text{if } \frac{5}{8} < \beta < \frac{3}{4}. \end{cases} \quad (3.10)$$

The result follows by a direct application of Theorem 1.1. □

Next, using the convergence of $\mathbb{E}[V_T(Z)^2]$, one can also establish the following corollary to Theorem 1.1.

Corollary 3.6. There exists a constant $C > 0$ such that, for all $T > 0$,

$$d_{TV} \left(\frac{\sqrt{T}}{\sigma_Z} (\widehat{f}_T(Z) - f_Z), \mathcal{N} \right) \leq \varphi_T(Z) + 2 \left| 1 - \frac{\sigma_Z^2}{E(V_T(Z)^2)} \right|$$

where $\varphi_T(Z)$ is as in Theorem 1.1. Moreover, if there exists $0 < \beta < \frac{3}{4}$ such that, $|\rho(t)| = \mathcal{O}(t^{2\beta-2})$,

$$d_{TV} \left(\frac{\sqrt{T}}{\sigma_Z} (\widehat{f}_T(Z) - f_Z), \mathcal{N} \right) \leq C \begin{cases} T^{-1/2} & \text{if } 0 < \beta \leq \frac{5}{8}, \\ T^{4\beta-3} & \text{if } \frac{5}{8} < \beta < \frac{3}{4}. \end{cases} \quad (3.11)$$

The same result holds for the Wasserstein distance.

Proof. The first part follows from the following technical result:

Lemma 3.7 ([11, Lemma 5.1]). *Let $\mu \in \mathbb{R}$ and $\sigma > 0$. Then, for every integrable real-valued random variable F ,*

$$d_{TV}(\mu + \sigma F, \mathcal{N}) \leq d_{TV}(F, \mathcal{N}) + \sqrt{\frac{\pi}{2}}|\mu| + 2\left|1 - \frac{1}{\sigma^2}\right|.$$

Therefore,

$$\begin{aligned} & d_{TV}\left(\frac{\sqrt{T}}{\sigma_Z}(\widehat{f}_T(Z) - f_Z), \mathcal{N}\right) \\ & \leq d_{TV}\left(\frac{\widehat{f}_T(Z) - f_Z}{\sqrt{\text{Var}(\widehat{f}_T(Z) - f_Z)}}, \mathcal{N}\right) + 2\left|1 - \frac{\sigma_Z^2}{\mathbb{E}[V_T(Z)^2]}\right|, \end{aligned}$$

by the definition of $V_T(Z)$ and the fact that $\mathbb{E}[\widehat{f}_T(Z)] = f_Z$. Recall that $\mathbb{E}[V_T(Z)^2] \uparrow \sigma_Z^2$. Then the second term on the right-hand side above is bounded by $C|\mathbb{E}[V_T(Z)^2 - \sigma_Z^2]|$ for some $C > 0$.

Next, since $|\rho(t)| = \mathcal{O}(t^{2\beta-2})$, and using the representation (3.4),

$$\begin{aligned} |\mathbb{E}[V_T(Z)^2 - \sigma_Z^2]| &= 4 \int_T^\infty \rho(u)^2 du + 4 \int_0^T \frac{u}{T} \rho(u)^2 du \\ &\leq 4C \left(\int_T^\infty u^{4\beta-4} du + \frac{1}{T} \int_0^1 \rho(u)^2 du + \frac{1}{T} \int_1^T u^{4\beta-3} du \right). \end{aligned}$$

A direct computation yields that,

$$|\mathbb{E}[V_T(Z)^2 - \sigma_Z^2]| \leq C \begin{cases} T^{-1} & \text{if } 0 < \beta < \frac{1}{2}, \\ \log(T)T^{-1} & \text{if } \beta = \frac{1}{2}, \\ T^{4\beta-3} & \text{if } \frac{1}{2} < \beta < \frac{3}{4}. \end{cases} \quad (3.12)$$

Finally, the bound (3.11) follows from (3.8) and (3.12). \square

3.3. Discrete-time observations

In this section we estimate the limiting variance f_Z based on discrete high-frequency data in time of Z , by considering the discrete version $\widetilde{f}_n(Z)$ of the estimator $\widehat{f}_T(Z)$:

$$\widetilde{f}_n(Z) := \frac{1}{n} \sum_{i=1}^n Z_{t_i}^2,$$

where $t_i = i\Delta_n$, $i = 0, \dots, n$, $\Delta_n \rightarrow 0$ and $T_n := n\Delta_n \rightarrow \infty$.

We will assume an additional property of the process Z . It would allow us to compare quantitatively the two estimators $\widehat{f}_T(Z)$ and $\widetilde{f}_n(Z)$:

Assumption 1. For all $s, t \in \mathbb{R}_+$ such that $|s - t|$ is small enough,

$$\mathbb{E}[|Z_t - Z_s|^2] \leq C|t - s|^{2\alpha}, \tag{3.13}$$

for some constant $0 < \alpha < 1$.

Similarly to the continuous alternative $\widehat{f}_T(Z)$, we first show the following:

Lemma 3.8. *The estimator $\widetilde{f}_n(Z)$ is unbiased. Assume that Assumption 1 holds. Then,*

$$\mathbb{E}|\widehat{f}_T(Z) - \widetilde{f}_n(Z)|^2 \leq C_\alpha \Delta_n^{2\alpha}, \tag{3.14}$$

where $C_\alpha > 0$ is a constant that depends only on α . Moreover, if $n\Delta_n^\eta \rightarrow 0$, as $n \rightarrow \infty$ for some $\eta > 1$, then $\widetilde{f}_n(Z)$ is strongly consistent.

Proof. The first property follows from the fact that the process Z is stationary. To show strong consistency define, similarly to (3.2),

$$U_n(Z) := \sqrt{T_n} \left(\widetilde{f}_n(Z) - f_Z \right) = \frac{\Delta_n}{\sqrt{T_n}} \sum_{i=1}^n (Z_{t_i}^2 - \mathbb{E}Z_{t_i}^2). \tag{3.15}$$

Then

$$\widetilde{f}_n(Z) - f_Z = \frac{U_n(Z)}{\sqrt{T_n}} = \frac{V_{T_n}(Z)}{\sqrt{T_n}} + \frac{U_n(Z) - V_{T_n}(Z)}{\sqrt{T_n}}. \tag{3.16}$$

We know from (3.1) and Lemma 3.2 that $V_{T_n}/\sqrt{T_n}$ converges almost surely to 0. Let $\delta_n(Z) := U_n(Z) - V_{T_n}(Z) = \sqrt{T_n}(\widetilde{f}_n(Z) - \widehat{f}_T(Z))$. We estimate the second moment of $\delta_n(Z)$:

$$\begin{aligned} \mathbb{E}[\delta_n(Z)^2] &\leq \mathbb{E} \left[\left(\frac{1}{\sqrt{T_n}} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} |Z_{t_i}^2 - Z_t^2| dt \right)^2 \right] \\ &\leq \frac{n}{T_n} \sum_{i=1}^n \mathbb{E} \left[\left(\int_{t_{i-1}}^{t_i} |Z_{t_i}^2 - Z_t^2| dt \right)^2 \right], \end{aligned}$$

where we have applied the inequality between arithmetic mean and quadratic mean. Next, by the Cauchy-Schwarz inequality, for $s, t \in [t_{i-1}, t_i]$,

$$\begin{aligned} &\mathbb{E}[|Z_{t_i}^2 - Z_t^2| |Z_{t_i}^2 - Z_s^2|] \\ &\leq \left(\mathbb{E}[(Z_{t_i}^2 - Z_t^2)^2] \mathbb{E}[(Z_{t_i} - Z_s)^2] \right)^{1/2} \\ &\leq \left(\mathbb{E}[(Z_{t_i} - Z_t)^4] \mathbb{E}[(Z_{t_i} - Z_s)^4] \mathbb{E}[(Z_{t_i} + Z_t)^4] \mathbb{E}[(Z_{t_i} + Z_s)^4] \right)^{1/4}. \end{aligned}$$

Now, by the hypercontractivity property (2.6) and (3.13)

$$\mathbb{E}[(Z_{t_i} - Z_t)^4]^{1/4} \leq C \mathbb{E}[(Z_{t_i} - Z_t)^2]^{1/2} \leq C|t_i - t|^\alpha.$$

Next, using stationarity

$$\mathbb{E}[(Z_{t_i} + Z_t)^4] \leq 4\mathbb{E}[(Z_{t_i}^2 + Z_t^2)^2] \leq 16\mathbb{E}[Z_0^4].$$

Therefore, for some constant $C > 0$ depending on the fourth moments of Z_0 ,

$$\begin{aligned} \mathbb{E}[\delta_n(Z)^2] &\leq \frac{n}{T_n} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} C|t_i - t|^\alpha |t_i - s|^\alpha ds dt \\ &= \frac{C\Delta_n^{2\alpha+2}}{T_n} \sum_{i=1}^n \int_0^1 \int_0^1 (uv)^\alpha dudv \\ &\leq Cn\Delta_n^{2\alpha+1}, \end{aligned}$$

with $u = \frac{t-t_{i-1}}{\Delta_n}$, $v = \frac{s-t_{i-1}}{\Delta_n}$. Thus, (3.14) is established.

Now, using that $n\Delta_n^\eta \rightarrow 0$ for some $\eta > 1$,

$$\mathbb{E} \left[\left(\frac{\delta_n(Z)}{\sqrt{T_n}} \right)^2 \right] \leq C\Delta_n^{2\alpha} \leq Cn^{-2\alpha/\eta} (n\Delta_n^\eta)^{2\alpha/\eta} \leq Cn^{-2\alpha/\eta}. \quad (3.17)$$

By the hypercontractivity property (2.6) and Lemma 3.2, we obtain $\frac{\delta_n(Z)}{\sqrt{T_n}} \rightarrow 0$, and thus $\hat{f}_{T_n}(Z) - \tilde{f}_n(Z) \rightarrow 0$ almost surely. By Lemma 3.1 $\hat{f}_{T_n} - f_Z \rightarrow 0$ almost surely. Therefore, $\tilde{f}_n(Z) - f_Z \rightarrow 0$ almost surely and the discrete estimator is strongly consistent. \square

We now turn to the proof of Theorem 1.2.

3.4. Proof of Theorem 1.2.

The discrete estimator $\tilde{f}_n(Z)$ is unbiased, so $Var(\tilde{f}_n(Z) - f_Z) = \mathbb{E}[(\tilde{f}_n(Z) - f_Z)^2]$. Next, by the triangle inequality one has,

$$\begin{aligned} &d_{TV} \left(\frac{\tilde{f}_n(Z) - f_Z}{\sqrt{\mathbb{E}[(\tilde{f}_n(Z) - f_Z)^2]}}, \mathcal{N} \right) \\ &\leq d_{TV} \left(\frac{\tilde{f}_n(Z) - f_Z}{\sqrt{\mathbb{E}[(\tilde{f}_n(Z) - f_Z)^2]}}, \frac{\hat{f}_{T_n}(Z) - f_Z}{\sqrt{\mathbb{E}[(\tilde{f}_n(Z) - f_Z)^2]}} \right) \\ &\quad + d_{TV} \left(\frac{\hat{f}_{T_n}(Z) - f_Z}{\sqrt{\mathbb{E}[(\hat{f}_{T_n}(Z) - f_Z)^2]}}, \mathcal{N} \right). \end{aligned} \quad (3.18)$$

By Lemma 3.7,

$$d_{TV} \left(\frac{\hat{f}_n(Z) - f_Z}{\sqrt{\mathbb{E}[(\hat{f}_{T_n}(Z) - f_Z)^2]}}, \mathcal{N} \right)$$

$$\leq d_{TV} \left(\frac{\widehat{f}_n(Z) - f_Z}{\sqrt{\mathbb{E}[(\widehat{f}_{T_n}(Z) - f_Z)^2]}}, \mathcal{N} \right) + 2 \left| 1 - \frac{\text{Var}(\widehat{f}_{T_n}(Z) - f_Z)}{\text{Var}(\widehat{f}_n(Z) - f_Z)} \right|.$$

Then, by Theorem 1.1 the first term on the right-hand side is further bounded by $\varphi(T_n)$ as defined in (1.6).

Thus, we are left to bound the first term in (3.18). By the definition of total variation distance one has

$$d_{TV} \left(\frac{\widetilde{f}_n(Z) - f_Z}{\sqrt{\mathbb{E}[(\widetilde{f}_n(Z) - f_Z)^2]}}, \frac{\widehat{f}_{T_n}(Z) - f_Z}{\sqrt{\mathbb{E}[(\widehat{f}_{T_n}(Z) - f_Z)^2]}} \right) = d_{TV}(U_n(Z), V_{T_n}(Z)),$$

with the two quantities given in (3.2) and (3.15). We recall the following technical result:

Lemma 3.9 ([24, Theorem 3.5]). *There is a positive constant $C > 0$ such that, for all multiple integrals F and G of order 2,*

$$d_{TV}(F, G) \leq C \left(\frac{\mathbb{E}[(F - G)^2]}{\mathbb{E}[F^2]} \right)^{1/4}.$$

Therefore, there is a positive constant $C > 0$ such that, for every $n \geq 1$

$$d_{TV}(U_n(Z), V_{T_n}(Z)) \leq C \left(\frac{\mathbb{E}[\delta_n(Z)^2]}{\mathbb{E}[V_{T_n}(Z)^2]} \right)^{1/4} \leq C(n\Delta_n^{2\alpha+1})^{1/4}, \quad (3.19)$$

where we have used (3.1) and (3.14). The corresponding bound for the Wasserstein distance follows from Remark 2.2. Thus, the proof of Theorem 1.2 is completed.

Remark 3.10. Let us present an example where the bounds on total variation in Theorem 1.2 decrease to 0 and thus a central limit theorem holds. Let $\Delta_n = 1/n^\lambda$ with $\frac{1}{2\alpha+1} < \lambda < 1$. Then, $n\Delta_n \rightarrow \infty$ and $n\Delta_n^{2\alpha+1} \rightarrow 0$. Moreover,

$$2 \left| 1 - \frac{\text{Var}(\widehat{f}_{T_n}(Z) - f_Z)}{\text{Var}(\widehat{f}_n(Z) - f_Z)} \right| = 2 \left| \frac{\mathbb{E}[U_n(Z)^2] - \mathbb{E}[V_{T_n}(Z)^2]}{\mathbb{E}[U_n(Z)^2]} \right| \leq 2 \frac{\mathbb{E}[\delta_n(Z)^2]}{\mathbb{E}[U_n(Z)^2]}.$$

Now, if $\mathbb{E}[\delta_n(Z)^2] \rightarrow 0$ and $\mathbb{E}[V_{T_n}(Z)] \uparrow \sigma_Z^2$, then $\mathbb{E}[U_n(Z)^2] \uparrow \sigma_Z^2$. Thus,

$$d_{TV} \left(\frac{\widetilde{f}_n(Z) - f_Z}{\sqrt{\mathbb{E}[(\widetilde{f}_n(Z) - f_Z)^2]}}, \mathcal{N} \right) \leq \varphi_{T_n}(Z) + Cn\Delta_n^{2\alpha+1}.$$

Further bounds on $\varphi_{T_n}(Z)$ are presented, for instance, in Corollary 3.5 when $|\rho(t)| = \mathcal{O}(t^{2\beta-2})$ for some $\beta \in (0, 3/4)$.

Finally, note that our approach for the proof of Theorem 1.2 can be applied directly to Corollary 3.6 and then the following holds:

Corollary 3.11. *Let $\sigma_Z^2 = 4 \int_0^\infty \rho(u)^2 du$. Under the same assumptions as in Theorem 1.2, for all $n \geq 1$,*

$$d_{TV} \left(\frac{\sqrt{T}}{\sigma_Z} (\tilde{f}_n(Z) - f_Z), \mathcal{N} \right) \leq \varphi_{T_n}(Z) + 2 \left| 1 - \frac{\sigma_Z^2}{E(V_{T_n}(Z)^2)} \right| + C(n\Delta_n^{2\alpha+1})^{1/4},$$

where $\varphi_{T_n}(Z)$ is as in Theorem 1.1 and $C > 0$ is an absolute constant depending on $\mathbb{E}[Z_0^4]$. The same result holds for the Wasserstein distance.

4. Parameter estimation for non-stationary Gaussian processes

In practical applications, the data rarely comes from a stationary process. This is reflected on the modeling side by for instance studying stochastic systems started at a point mass rather than the system’s stationary distribution. Some more specific examples are included in Section 5.

The present section is devoted to the general treatment of a process that is asymptotically stationary. In particular, let Z be a centered stationary Gaussian process (as in Section 3), and let Y be a stochastic process satisfying the following: there exists a constant $\gamma > 1$ such that for every $p \geq 1$ and for all $T > 0$,

$$\|Y_T\|_{L^p} = \mathcal{O}(T^{-\gamma}). \tag{4.1}$$

Then we consider second moment estimators for the process $X := Z + Y$. Our goal is to estimate the limiting variance $f_X := \lim_{T \rightarrow \infty} \mathbb{E}[X_T^2]$. The limit exists and in fact $f_X = f_Z$. Indeed,

$$\mathbb{E}[X_T^2] = \mathbb{E}[(Z_T + Y_T)^2] = \mathbb{E}[Z_T^2] + 2\mathbb{E}[Z_T Y_T] + \mathbb{E}[Y_T^2] \rightarrow f_Z,$$

since $\mathbb{E}[Z_T^2] = f_Z$, $\mathbb{E}[Y_T^2] = \mathcal{O}(T^{-2\gamma})$ and by the Cauchy-Schwarz inequality $|\mathbb{E}[Z_T Y_T]| = \mathcal{O}(T^{-\gamma})$. As in Section 3 we consider the second moment estimators $\hat{f}_T(X)$ and $\tilde{f}_n(X)$, based on continuous-time and discrete-time observations of X :

$$\hat{f}_T(X) := \frac{1}{T} \int_0^T X_t^2 dt, \quad T > 0, \tag{4.2}$$

$$\tilde{f}_n(X) := \frac{1}{n} \sum_{i=1}^n X_{t_i}^2, \quad n \geq 1, \tag{4.3}$$

where $t_i = i\Delta_n$, $i = 0, \dots, n$, $\Delta_n \rightarrow 0$ and $T_n := n\Delta_n \rightarrow \infty$. We establish some basic properties of the two estimators.

Proposition 4.1. *The estimators $\hat{f}_T(X)$ and $\tilde{f}_n(X)$ are asymptotically unbiased. Moreover, there is a constant $C > 0$, such that for all $T > 0$,*

$$|\mathbb{E}[\hat{f}_T(X) - f_X]| \leq CT^{-\gamma}, \text{ and } |\mathbb{E}[\tilde{f}_n(X) - f_X]| \leq CT^{-\gamma}. \tag{4.4}$$

Proof. We established that for all $T > 0$, $\mathbb{E}[X_T^2] = f_X + err(t)$, where the error term satisfies $|err(t)| = \mathcal{O}(t^{-\gamma})$. Therefore, for the continuous estimator,

$$\|\mathbb{E}[\widehat{f}_T(X) - f_X]\| \leq \frac{1}{T} \int_0^T |err(t)| dt \leq CT^{-\gamma} \rightarrow 0 \quad \text{as } T \rightarrow \infty,$$

and thus $\widehat{f}_T(X)$ is asymptotically unbiased. A similar computation yields the same for $\widetilde{f}_n(Z)$. □

We next show strong consistency.

Proposition 4.2. *If $\int_{\mathbb{R}} \rho^2(r) dr < \infty$ and the condition (4.1) holds, then*

$$\left\| \widehat{f}_T(X) - f_X \right\|_{L^2} \leq CT^{-1/2}. \tag{4.5}$$

Moreover,

$$\widehat{f}_T(X) \longrightarrow f_X \quad \text{almost surely as } T \rightarrow \infty. \tag{4.6}$$

In addition, if Z satisfies the helix property (3.13), and $n\Delta_n^{2\alpha+1} \rightarrow 0$, as $n \rightarrow \infty$, then

$$\widetilde{f}_n(X) \longrightarrow f_X \quad \text{almost surely as } T \rightarrow \infty. \tag{4.7}$$

Proof. The first step, as before, is to establish a bound on $\left\| \widehat{f}_T(X) - f_X \right\|_{L^2}$. We have that

$$\left\| \widehat{f}_T(X) - f_X \right\|_{L^2} \leq \left\| \widehat{f}_T(X) - \widehat{f}_T(Z) \right\|_{L^2} + \left\| \widehat{f}_T(Z) - f_X \right\|_{L^2}.$$

Note, that by the triangle inequality and the Cauchy-Schwarz inequality,

$$\left\| \widehat{f}_T(X) - \widehat{f}_T(Z) \right\|_{L^2} \leq \sqrt{\mathbb{E} \left| \frac{1}{T} \int_0^T Y_t dt \right|^2} \leq \frac{1}{T} \left(\int_{[0,T]^2} (\mathbb{E}[Y_t^2] \mathbb{E}[Y_s^2])^{1/2} ds dt \right)^{1/2}. \tag{4.8}$$

Therefore, using $\|Y_t\|_{L^p} = \mathcal{O}(t^{-\gamma})$,

$$\left\| \widehat{f}_T(X) - f_X \right\|_{L^2} \leq C(T^{-\gamma} + T^{-1/2}) \leq CT^{-1/2},$$

where we have applied (3.1).

By hypercontractivity (2.6) we can extend the bound to L^p norms for $p \geq 1$. Then applying Lemma 3.2 and Lemma 3.3 consecutively establishes strong consistency for $\widehat{f}_T(X)$.

The proof for $\widetilde{f}_n(X)$ is similar and a main ingredient is the bound from (3.14). □

We now turn to the alternative formulations for our main results Theorem 1.1 and Theorem 1.2 when the stochastic process is asymptotically stationary. Recall that $\rho(t) = \rho(-t) := \mathbb{E}[Z_0 Z_t]$ for $t \geq 0$.

Theorem 4.3. *Assume that $\int_{\mathbb{R}} \rho^2(r) dr < \infty$ and that the condition (4.1) holds. Let $\mathcal{N} \sim \mathcal{N}(0, 1)$ be the standard normal random variable. Then, there exists a constant $C > 0$ such that, for all $T > 0$,*

$$d_{TV} \left(\frac{\widehat{f}_T(X) - f_X - \mathbb{E}[\widehat{f}_T(X) - f_X]}{\sqrt{\text{Var}(\widehat{f}_T(X) - f_X)}}, \mathcal{N} \right) \leq \varphi_T(Z) + CT^{\frac{1-\gamma}{4}}, \quad (4.9)$$

where $\varphi_T(Z)$ is defined in (1.6). The same result holds for the Wasserstein distance.

Proof. By the triangle inequality and Lemma (3.7):

$$\begin{aligned} & d_{TV} \left(\frac{\widehat{f}_T(X) - f_X - \mathbb{E}[\widehat{f}_T(X) - f_X]}{\sqrt{\text{Var}(\widehat{f}_T(X) - f_X)}}, \mathcal{N} \right) \\ & \leq d_{TV} \left(\frac{\widehat{f}_T(Z) - f_Z}{\sqrt{\text{Var}(\widehat{f}_T(Z) - f_Z)}}, \mathcal{N} \right) \\ & \quad + d_{TV} \left(\frac{\widehat{f}_T(X) - f_Z}{\sqrt{\text{Var}(\widehat{f}_T(Z) - f_Z)}}, \frac{\widehat{f}_T(Z) - f_Z}{\sqrt{\text{Var}(\widehat{f}_T(Z) - f_Z)}} \right) \\ & \quad + \sqrt{\frac{\pi}{2}} \left| \frac{\mathbb{E}[\widehat{f}_T(X) - f_X]}{\sqrt{\text{Var}(\widehat{f}_T(X) - f_X)}} \right| + 2 \left| 1 - \frac{\text{Var}(\widehat{f}_T(X) - f_X)}{\text{Var}(\widehat{f}_T(Z) - f_Z)} \right|. \end{aligned} \quad (4.10)$$

We analyze each term on the right hand side of (4.10). First, by Theorem 1.1,

$$d_{TV} \left(\frac{\widehat{f}_T(Z) - f_Z}{\sqrt{\text{Var}(\widehat{f}_T(Z) - f_Z)}}, \mathcal{N} \right) \leq \varphi_T(Z),$$

with $\varphi_T(Z)$ as defined in (1.6).

Then, using standard properties of the total variation distance and Lemma 3.9:

$$\begin{aligned} & d_{TV} \left(\frac{\widehat{f}_T(X) - f_Z}{\sqrt{\text{Var}(\widehat{f}_T(Z) - f_Z)}}, \frac{\widehat{f}_T(Z) - f_Z}{\sqrt{\text{Var}(\widehat{f}_T(Z) - f_Z)}} \right) \\ & = d_{TV} \left(\widehat{f}_T(X) - f_Z, \widehat{f}_T(Z) - f_Z \right) \leq C \left(\frac{\mathbb{E}|\widehat{f}_T(X) - \widehat{f}_T(Z)|^2}{\mathbb{E}|\widehat{f}_T(Z) - f_Z|^2} \right)^{1/4} \end{aligned}$$

$$\leq CT^{(1-\gamma)/4}.$$

Indeed, $\mathbb{E}|\widehat{f}_T(X) - \widehat{f}_T(Z)|^2 = \mathcal{O}(T^{-\gamma})$ and $T\mathbb{E}|\widehat{f}_T(Z) - f_Z|^2 \rightarrow \sigma_Z^2 < \infty$, see (3.1). Moreover since $|\mathbb{E}[\widehat{f}_T(X) - f_X]| = \mathcal{O}(T^{-\gamma})$, one has

$$\left| \text{Var}(\widehat{f}_T(X) - f_X) - \text{Var}(\widehat{f}_T(Z) - f_Z) \right| = \mathcal{O}(T^{-\gamma}). \quad (4.11)$$

Then, there exists a constant $C > 0$ such that for all $T > 0$,

$$\sqrt{\frac{\pi}{2}} \left| \frac{\mathbb{E}[\widehat{f}_T(X) - f_X]}{\sqrt{\text{Var}(\widehat{f}_T(X) - f_X)}} \right| \leq CT^{1/2-\gamma}$$

and

$$2 \left| 1 - \frac{\text{Var}(\widehat{f}_T(X) - f_X)}{\text{Var}(\widehat{f}_T(Z) - f_Z)} \right| \leq CT^{1-\gamma}.$$

Therefore,

$$d_{TV} \left(\frac{\widehat{f}_T(X) - f_X - \mathbb{E}[\widehat{f}_T(X) - f_X]}{\sqrt{\text{Var}(\widehat{f}_T(X) - f_X)}}, \mathcal{N} \right) \leq \varphi_T(Z) + CT^{\frac{1-\gamma}{4}},$$

as desired. The corresponding bound for the Wasserstein distance follows from Remark 2.2. \square

We now turn to the discrete version of the previous result.

Theorem 4.4. *Assume $\int_{\mathbb{R}} \rho^2(r) dr < \infty$, Z verifies (3.13), and the condition (4.1) holds. Let $\mathcal{N} \sim \mathcal{N}(0, 1)$. Then, there is $C > 0$ such that, for every $n \geq 1$,*

$$d_{TV} \left(\frac{\sqrt{T_n}}{\sigma_Z} (\widetilde{f}_n(X) - f_X), \mathcal{N} \right) \leq \varphi_{T_n}(Z) + C [n\Delta_n^{2\alpha+1}]^{1/4} + CT_n^{\frac{1-\gamma}{4}}.$$

Proof. We follow the same steps as in the proof of Theorem 4.4. The extra term $Cn[\Delta_n^{2\alpha+1}]^{1/4}$ comes from the bound in Theorem 1.2. \square

5. Applications to Gaussian Ornstein-Uhlenbeck processes

In this section we consider the Gaussian Ornstein-Uhlenbeck process $X^\theta := \{X_t^\theta, t \geq 0\}$ defined via the following linear stochastic differential equation

$$dX_t^\theta = -\theta X_t^\theta dt + dG_t, \quad X_0^\theta = 0. \quad (5.1)$$

Here $\theta > 0$ is an unknown parameter and G is an arbitrary mean-zero Gaussian process such that $Z_t^\theta := \int_{-\infty}^t e^{-\theta(t-s)} dG_s$, for $t \geq 0$, is a stationary Gaussian process. The equation (5.1) has the following explicit solution (see [10])

$$X_t^\theta = e^{-\theta t} \int_0^t e^{\theta s} dG_s, \quad t \geq 0,$$

where the integral can be understood in the Wiener sense. Then $X_t^\theta = Z_t^\theta + e^{-\theta t}Z_0^\theta$ thus satisfies (4.1) with $Y := e^{-\theta t}Z_0^\theta$. Moreover, $\|Y_T\|_{L^p} = \mathcal{O}(T^{-\gamma})$ is satisfied for every $p \geq 1$ and for every constant $\gamma > 1$.

As in the previous sections our goal is to establish a Berry-Esseen theorem involving estimators of the parameters of the process. Here we want to estimate θ . Our approach is based on writing

$$\theta = g_{Z^\theta}(\mathbb{E}[(Z_0^\theta)^2]),$$

for some invertible function $g_{Z^\theta} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$. There are certain cases when there is an explicit expression for g_{Z^θ} (or its inverse). We explore those in the coming sections. For now, we describe the general treatment and make only one assumption for g_{Z^θ} :

Assumption 2. The function g_{Z^θ} is a diffeomorphism, and is twice continuously differentiable.

We estimate θ based the continuous and discrete observations of X :

$$\hat{\theta}_T := g_{Z^\theta} \left(\frac{1}{T} \int_0^T (X_t^\theta)^2 dt \right) = g_{Z^\theta} \left(\hat{f}_T(X^\theta) \right), \quad T > 0, \tag{5.2}$$

$$\tilde{\theta}_n := g_{Z^\theta} \left(\frac{1}{n} \sum_{i=1}^n (X_{t_i}^\theta)^2 \right) = g_{Z^\theta} \left(\tilde{f}_n(X^\theta) \right), \quad n \geq 1, \tag{5.3}$$

where $t_i = i\Delta_n$, $i = 0, \dots, n$, $\Delta_n \rightarrow 0$ and $T_n = n\Delta_n \rightarrow \infty$, whereas $\hat{f}_T(X^\theta)$ and $\tilde{f}_n(X^\theta)$ are given by (4.2) and (4.3), respectively.

The following holds for the continuous estimator $\hat{\theta}_T$.

Theorem 5.1. Assume that $\int_{\mathbb{R}} \rho^2(r)dr < \infty$ and that Assumption 2 holds. Then for every $\gamma > 1$, and $p \geq 1$,

$$\begin{aligned} & d_W \left(\frac{\hat{\theta}_T - \theta - \mathbb{E}[\hat{\theta}_T - \theta]}{g'_{Z^\theta}(f_{X^\theta}) \sqrt{\text{Var}(\hat{f}_T(X^\theta) - f_{X^\theta})}}, \mathcal{N} \right) \\ & \leq C \frac{(\mathbb{E}|g''(\zeta_T)|^p)^{1/p}}{\sqrt{T}} + \varphi_T(Z^\theta) + CT^{-1/2}, \end{aligned} \tag{5.4}$$

where ζ_T is a random variable in $[\hat{f}_T(X^\theta), f_{X^\theta}]$ and for some absolute constant $C > 0$.

Proof. Recall that by definition $\theta = g_{Z^\theta}(f_{X^\theta})$. Under Assumption 2

$$\left(\hat{\theta}_T - \theta \right) = g'_{Z^\theta}(f_{X^\theta}) \left(\hat{f}_T(X^\theta) - f_{X^\theta} \right) + \frac{1}{2} g''_{Z^\theta}(\zeta_T) \left(\hat{f}_T(X^\theta) - f_{X^\theta} \right)^2$$

for some random point ζ_T between $\hat{f}_T(X^\theta)$ and f_{X^θ} . Denote $V_f^2 := \text{Var}(\hat{f}_T(X^\theta) - f_{X^\theta})$. Then,

$$d_W \left(\frac{\hat{\theta}_T - \theta - \mathbb{E}[\hat{\theta}_T - \theta]}{g'_{Z^\theta}(f_{X^\theta})V_f}, \mathcal{N} \right) \leq \frac{|\mathbb{E}[\hat{\theta}_T - \theta]| + \frac{1}{2} \mathbb{E} \left| (\hat{f}_T(X^\theta) - f_{X^\theta})^2 g''_{Z^\theta}(\zeta_T) \right|}{|g'_{Z^\theta}(f_{X^\theta})V_f|}$$

$$+ d_W \left(\frac{1}{\sqrt{V_f}} (\widehat{f}_T(X^\theta) - f_{X^\theta}), \mathcal{N} \right),$$

where we have used that $d_W(x_1 + x_2, y) \leq \mathbb{E}[|x_2|] + d_W(x_1, y)$ for any random variables x_1, x_2, y . This property of the Wasserstein distance is the main reason our results in Section 5 concern d_W and not d_{TV} .

The second term in the inequality above is bounded in Theorem 4.3. By Hölder's inequality, and the hypercontractivity property, for $p, q > 1$ with $1/p + 1/q = 1$,

$$\begin{aligned} \mathbb{E} \left| (\widehat{f}_T(X^\theta) - f_{X^\theta})^2 g''_{Z^\theta}(\zeta_T) \right| &\leq (\mathbb{E}|g''(\zeta_T)|^p)^{1/p} \left(\mathbb{E} \left| \widehat{f}_T(X^\theta) - f_{X^\theta} \right|^{2q} \right)^{1/q} \\ &\leq C (\mathbb{E}|g''(\zeta_T)|^p)^{1/p} \mathbb{E} \left| \widehat{f}_T(X^\theta) - f_{X^\theta} \right|^2, \end{aligned}$$

for some constant $C > 0$ depending on p . Moreover,

$$|\mathbb{E}[\widetilde{\theta}_n - \theta]| \leq |g'_{Z^\theta}(f_{X^\theta})| \left| \mathbb{E}[\widehat{f}_T(X^\theta) - f_{X^\theta}] \right| + \mathbb{E} \left| (\widehat{f}_T(X^\theta) - f_{X^\theta})^2 g''_{Z^\theta}(\zeta_T) \right|.$$

Recall, that by (4.5), $\mathbb{E}|\widehat{f}_T(X^\theta) - f_{X^\theta}|^2 \leq C/T$, and by (4.4) $|\mathbb{E}[\widehat{f}_T(X^\theta) - f_{X^\theta}]| \leq CT^{-\gamma}$. Therefore,

$$\begin{aligned} &|\mathbb{E}[\widetilde{\theta}_n - \theta]| + \frac{1}{2} \mathbb{E} \left| (\widehat{f}_T(X^\theta) - f_{X^\theta})^2 g''_{Z^\theta}(\zeta_T) \right| \\ &\leq C \left(|g'_{Z^\theta}(f_{X^\theta})| + \frac{3}{2} (\mathbb{E}|g''(\zeta_T)|^p)^{1/p} \right) T^{-1}. \end{aligned}$$

Now, by (3.1) and (4.11), $TV_f^2 \rightarrow \sigma_z^2$. The bound (5.4) follows. \square

Similarly, we obtain the rate of convergence in law of $\sqrt{T_n}(\widetilde{\theta}_n - \theta)$ as follows.

Theorem 5.2. *Suppose that the conditions of Theorem 4.4 hold. If $g''(\xi_n)$, with $\xi_n \in [|\widehat{f}_n(X^\theta)|, f_{X^\theta}]$, has a moment of order greater than 1 which is bounded in n ,*

$$d_W \left(\frac{\sqrt{T_n}}{\sigma_{Z^\theta} g'_{Z^\theta}(f_{X^\theta})} (\widetilde{\theta}_n - \theta), \mathcal{N} \right) \leq \frac{C}{\sqrt{T_n}} + \varphi_{T_n}(Z^\theta) + C [n\Delta_n^{2\alpha+1}]^{1/4}.$$

5.1. Fractional Ornstein-Uhlenbeck process of the first kind

Here we consider the Ornstein-Uhlenbeck process $X^\theta := \{X_t^\theta, t \geq 0\}$ driven by a fractional Brownian motion $\{B_t^H, t \geq 0\}$ of Hurst index $H \in (0, 1)$. More precisely, X^θ is the solution of the following linear stochastic differential equation

$$X_0^\theta = 0; \quad dX_t^\theta = -\theta X_t^\theta dt + dB_t^H, \quad t \geq 0, \quad (5.5)$$

where $\theta > 0$ is an unknown parameter. The process X^θ is called a fractional Ornstein-Uhlenbeck process of the first kind, following the notation in [22]. As in the general case described above there is an explicit solution to (5.5):

$$X_t^\theta = \int_0^t e^{-\theta(t-s)} dB_s^H. \tag{5.6}$$

Moreover,

$$Z_t^\theta = \int_{-\infty}^t e^{-\theta(t-s)} dB_s^H \tag{5.7}$$

is a stationary Gaussian process, see [10, 17]. The process Z^θ is also the stationary solution of equation (5.5) when $X_0^\theta = Z_0^\theta$. We are thus in the setup of Section 4 with $X^\theta = Z^\theta + Y^\theta$, where $Y_t^\theta = -e^{-\theta t} Z_0^\theta$. Again, $\|Y_t\|_{L^p} = \mathcal{O}(t^{-\gamma})$ for every $p \geq 1$ and any $\gamma > 1$. Note that by integration by parts and the formula for covariance for fractional Brownian motion

$$\begin{aligned} \mathbb{E}[(Z_0^\theta)^2] &= \mathbb{E} \left(\int_{-\infty}^0 e^{\theta s} dB_s^H \right)^2 = \mathbb{E} \int_{-\infty}^0 \int_{-\infty}^0 B_s^H B_r^H e^{\theta(s+r)} ds dr \\ &= \theta^2 \int_{-\infty}^0 \int_{-\infty}^0 \frac{1}{2} e^{\theta(s+r)} (|r|^{2H} + |s|^{2H} - |r-s|^{2H}) dr ds \\ &= \theta^{-2H} H \Gamma(2H). \end{aligned}$$

Therefore, $\theta = g_{Z^\theta}(\mathbb{E}[(Z_0^\theta)^2])$ where the invertible function $g_{Z^\theta} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is given by

$$g_{Z^\theta}(x) = \left(\frac{H \Gamma(2H)}{x} \right)^{\frac{1}{2H}}, \quad x > 0. \tag{5.8}$$

The estimators $\hat{\theta}_T$ and $\tilde{\theta}_T$ given by (5.2) and (5.3) were carefully studied in [20]. In particular, due to [20, Theorem 9] and [20, Theorem 11] the following holds:

Proposition 5.3. *Let $H \in (0, 3/4)$. The estimators $\hat{\theta}_T$ and $\tilde{\theta}_n$ are strongly consistent. Denote*

$$\delta_H^2 := \frac{\theta}{(2H)^2} \times \begin{cases} (4H - 1) + \frac{2\Gamma(2-4H)\Gamma(4H)}{\Gamma(2H)\Gamma(1-2H)} & \text{if } H \in (0, \frac{1}{2}), \\ (4H - 1) \left(1 + \frac{\Gamma(3-4H)\Gamma(4H-1)}{\Gamma(2-2H)\Gamma(2H)} \right) & \text{if } H \in [\frac{1}{2}, \frac{3}{4}). \end{cases} \tag{5.9}$$

The following limit theorems hold:

1. $\sqrt{T}(\hat{\theta}_T - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \delta_H^2)$ as $T \rightarrow \infty$.
2. Assume there is $p \in (1, \frac{3+2H}{1+2H} \wedge (1 + 2H))$ such that $n\Delta_n^p \rightarrow 0$. Then $\sqrt{T_n}(\tilde{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \delta_H^2)$ as $n \rightarrow \infty$.

We can extend these results and obtain Berry-Esseen bounds using Theorems 5.1 and 5.2.

Theorem 5.4. *Let $H \in (0, \frac{3}{4})$, and δ_H be given by (5.9). Then*

$$d_W \left(\frac{\sqrt{T}}{\delta_H} (\hat{\theta}_T - \theta), \mathcal{N} \right) \leq C \begin{cases} \frac{1}{\sqrt{T}} & \text{if } 0 < H \leq \frac{5}{8}, \\ \frac{1}{T^{3-4H}} & \text{if } \frac{5}{8} < H < \frac{3}{4}. \end{cases} \tag{5.10}$$

Moreover,

$$d_W \left(\frac{\sqrt{T_n}}{\delta_H} (\tilde{\theta}_n - \theta), \mathcal{N} \right) \leq C [n\Delta_n^{2H+1}]^{1/2} + C \begin{cases} \frac{1}{\sqrt{n\Delta_n}} & \text{if } 0 < H \leq \frac{5}{8}, \\ \frac{1}{(n\Delta_n)^{3-4H}} & \text{if } \frac{5}{8} < H < \frac{3}{4}. \end{cases} \tag{5.11}$$

Proof. Recall that $\mathbb{E}[Z_0 Z_r] = \mathcal{O}(r^{2H-2})$ and $H \in (0, 3/4)$, see [11, Proposition 4.2]. Thus we are in the setting of Corollary 3.6. Then $\varphi_T(Z)$ and $\varphi_{T_n}(Z)$ are respectively bounded via (3.11) with $\beta = H$.

To establish (5.10) and (5.11) it is left to show that $\mathbb{E}|g''(\zeta_T)|^p < \infty$ for some $p \geq 1$. Using the monotonicity of g'' and the fact that $\zeta_T \in [\hat{f}_T(X^\theta), f_{X^\theta}]$, it is enough to show that $\mathbb{E}|\hat{f}_T(X^\theta)|^p < \infty$ for some $p \geq 1$. This follows as an application of the technical Proposition 6.3 in Section 6.2. Indeed, since $X_t^\theta = Z_t^\theta - e^{\theta t} Z_0^\theta$ for a centered stationary Gaussian process $(Z_t^\theta)_{t \geq 0}$, one needs only to check that Z^θ satisfies the condition (6.4), i.e.,

$$\lim_{t \rightarrow 0} \mathbb{E}[(Z_t^\theta - Z_0^\theta)^2] = 0, \quad \text{and} \quad \bigcap_{t \in \mathbb{R}} \overline{\text{sp}}\{Z_s^\theta : -\infty < s \leq t\} = \{0\}, \tag{5.12}$$

where $\overline{\text{sp}}$ denotes the L^2 -closure of the linear span of a set of square-integrable random variables. Note, that by (5.7), and integration by parts,

$$\begin{aligned} \mathbb{E}[(Z_t^\theta - Z_0^\theta)^2] &= \mathbb{E} \int_0^t \int_0^t B_s^H B_r^H e^{\theta(s+r)} ds dr \\ &= \int_0^t \int_0^t \frac{\theta^2}{2} e^{\theta(s+r)} (r^{2H} + s^{2H} - |r - s|^{2H}) ds dr, \end{aligned}$$

which in turn approaches 0 as $t \rightarrow 0$ and the first part of (5.12) is satisfied. For the second part, note that $\{Z_s^\theta : -\infty < s \leq t\} \subset \sigma(B_s^H : -\infty < s \leq t)$ and the intersection of the sigma algebras is empty. \square

5.2. Fractional Ornstein-Uhlenbeck process of the second kind

The last example we consider is the so-called fractional Ornstein-Uhlenbeck process of the second kind, defined via the stochastic differential equation

$$S_0^\mu = 0, \text{ and } dS_t^\mu = -\mu S_t^\mu dt + dY_t^{(1)}, \quad t \geq 0, \tag{5.13}$$

where $Y_t^{(1)} = \int_0^t e^{-s} dB_{a_s}^H$ with $a_s = He^{\frac{s}{H}}$ and $\{B_t^H, t \geq 0\}$ is a fractional Brownian motion with Hurst parameter $H \in (\frac{1}{2}, 1)$, and where $\mu > 0$ is the

unknown real parameter which we would like to estimate. The equation (5.13) admits an explicit solution, see [22, Equation 3.9]:

$$S_t^\mu = e^{-\mu t} \int_0^t e^{\mu s} dY_s^{(1)} = e^{-\mu t} \int_0^t e^{(\mu-1)s} dB_{a_s}^H = H^{(1-\mu)H} e^{-\mu t} \int_{a_0}^{a_t} r^{(\mu-1)H} dB_r^H.$$

Hence we can also write

$$S_t^\mu = Z_t^\mu - e^{-\mu t} Z_0^\mu,$$

where

$$Z_t^\mu = e^{-\mu t} \int_{-\infty}^t e^{(\mu-1)s} dB_{a_s}^H = H^{(1-\mu)H} e^{-\mu t} \int_0^{a_t} r^{(\mu-1)H} dB_r^H.$$

From [17, Lemma 37], for every $H \in (\frac{1}{2}, 1)$,

$$g_{Z^\mu}^{-1}(\mu) = f_{X^\mu} = f_{Z^\mu} = E \left[(Z_0^\mu)^2 \right] = \frac{(2H-1)H^{2H}}{\mu} \mathcal{B}(1-H+\mu H, 2H-1),$$

where here $\mathcal{B}(\cdot)$ is the usual beta function. The function $\mu \mapsto g_{Z^\mu}^{-1}(\mu)$ is monotone (decreasing) and convex from \mathbb{R}_+ to \mathbb{R}_+ . Now, the following Berry-Esseen result holds:

Theorem 5.5. *Assume $H \in (\frac{1}{2}, 1)$. Then*

$$d_W \left(\frac{\sqrt{T}}{\sigma_{Z^\mu} g'_{Z^\mu}(f_{X^\mu})} (\hat{\mu}_T - \mu), \mathcal{N} \right) \leq C/\sqrt{T}.$$

Also,

$$d_W \left(\frac{\sqrt{T_n}}{\sigma_{Z^\mu} g'_{Z^\mu}(f_{X^\mu})} (\tilde{\mu}_n - \mu), \mathcal{N} \right) \leq C [n\Delta_n^{2H+1}]^{1/2} + C/\sqrt{n\Delta_n}.$$

Proof. From [22], there exist $c, C > 0$ such that for all large $|t|$,

$$\rho_{Z^\mu}(t) = \mathbb{E}[Z_0^\mu Z_t^\mu] \leq C e^{-c|t|}.$$

Thus, by a straightforward calculation,

$$\varphi_T(Z^\mu) \leq C/\sqrt{T}.$$

Moreover, according to [3, Lemma 4], for all $H \in (0, 1)$, and $|t-s|$ small enough, $\mathbb{E} \left| Y_t^{(1)} - Y_s^{(1)} \right|^2 = C|t-s|^{2H}$ and we are in the setting of Corollary 3.6.

Finally, one needs to bound $\mathbb{E}|g''(\zeta_T)|^p$. As in the proof of Theorem 5.4 we use the technical proposition Proposition 6.3 in Section 6.2. Indeed, since $S_t^\mu = Z_t^\mu - e^{-\mu t} Z_0^\mu$ for a centered stationary Gaussian process $(Z_t^\mu)_{t \geq 0}$, one needs only to check that Z^μ satisfies the condition (6.4), i.e.,

$$\lim_{t \rightarrow 0} \mathbb{E}[(Z_t^\mu - Z_0^\mu)^2] = 0, \quad \text{and} \quad \bigcap_{t \in \mathbb{R}} \overline{\text{sp}}\{Z_s^\mu : -\infty < s \leq t\} = \{0\}, \quad (5.14)$$

where $\overline{\text{sp}}$ denotes the L^2 -closure of the linear span of a set of square-integrable random variables. Recall [2, Lemma 2.1]. One has $\mathbb{E}[(Z_t^\mu - Z_0^\mu)^2] \sim Ht^{2H}$ as $t \rightarrow 0^+$. Thus the first part of (5.14) is satisfied. For the second part, note that $\{Z_s^\theta : -\infty < s \leq t\} \subset \sigma(B_{He^s/H}^H : -\infty < s \leq t)$ and the intersection of the sigma algebras is empty. \square

6. Technical results

6.1. Two technical lemmas

Proof of Lemma 3.3. We employ similar arguments to [30, Proposition 4.1]. First, write

$$\frac{1}{T} \int_0^T u_t dt = \frac{1}{[T]} \int_0^{[T]} u_t dt + \frac{1}{T} \int_{[T]}^T u_t dt + \left(\frac{1}{T} - \frac{1}{[T]}\right) \int_0^{[T]} u_t dt.$$

Notice that

$$\left| \left(\frac{1}{T} - \frac{1}{[T]}\right) \int_0^{[T]} u_t dt \right| = \left(1 - \frac{[T]}{T}\right) \left| \frac{1}{[T]} \int_0^{[T]} u_t dt \right| \leq \left| \frac{1}{[T]} \int_0^{[T]} u_t dt \right|.$$

We know that $\frac{1}{[T]} \int_0^{[T]} u_t dt \rightarrow 0$ almost surely as $T \rightarrow 0$. Thus, we are left to show that

$$\left| \frac{1}{T} \int_{[T]}^T u_t dt \right| \rightarrow 0, \text{ almost surely as } T \rightarrow \infty. \tag{6.1}$$

Recall, that $\sup_{t \geq 0} \mathbb{E}|u_t|^p < C_p$ for any $p \geq 1$. Thus, by Minkowski's inequality

$$\left(\mathbb{E} \left| \frac{1}{[T]} \int_{[T]}^{[T]+1} |u_t| dt \right|^p \right)^{1/p} \leq \frac{1}{[T]} \int_{[T]}^{[T]+1} (\mathbb{E}|u_t|^p)^{1/p} dt \leq \frac{C_p^{1/p}}{[T]}.$$

Now, by Lemma 3.2,

$$\left| \frac{1}{[T]} \int_{[T]}^{[T]+1} |u_t| dt \right| \rightarrow 0, \text{ almost surely as } T \rightarrow \infty,$$

and (6.1) follows. \square

Proof of Lemma 3.4. According to (3.2) and since $\mathbb{E}[V_T(Z)] = 0$, we have

$$\begin{aligned} \kappa_3(V_T(Z)) &= \mathbb{E}[V_T(Z)^3] = \mathbb{E} \left[\left(\frac{1}{\sqrt{T}} \int_0^T I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) \right)^3 \right] \\ &= \frac{1}{T^{3/2}} \int_{[0,T]^3} \mathbb{E} \left[I_2(\mathbf{1}_{[0,r]}^{\otimes 2}) I_2(\mathbf{1}_{[0,s]}^{\otimes 2}) I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) \right] dr ds dt. \end{aligned}$$

Applying the product formula (2.5)

$$I_2(\mathbf{1}_{[0,s]}^{\otimes 2})I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) = I_4(\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_0 \mathbf{1}_{[0,t]}^{\otimes 2}) + 4I_2(\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_1 \mathbf{1}_{[0,t]}^{\otimes 2}) + I_0(\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_2 \mathbf{1}_{[0,t]}^{\otimes 2}). \tag{6.2}$$

Now, by isometry (2.4)

$$\begin{aligned} \mathbb{E} \left[I_2(\mathbf{1}_{[0,r]}^{\otimes 2})I_2(\mathbf{1}_{[0,s]}^{\otimes 2})I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) \right] &= 4\mathbb{E} \left[I_2(\mathbf{1}_{[0,r]}^{\otimes 2})I_2(\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_1 \mathbf{1}_{[0,t]}^{\otimes 2}) \right] \\ &= 8\langle \mathbf{1}_{[0,r]}^{\otimes 2}, (\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_1 \mathbf{1}_{[0,t]}^{\otimes 2}) \rangle_{\mathcal{H}^{\otimes 2}}. \end{aligned}$$

Recall the formula for contraction (2.3). Then,

$$\mathbf{1}_{[0,s]}^{\otimes 2} \otimes_1 \mathbf{1}_{[0,t]}^{\otimes 2} = \langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]} \rangle_{\mathcal{H}} \mathbf{1}_{[0,s]} \otimes \mathbf{1}_{[0,t]}.$$

The symmetrization is then

$$\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_1 \mathbf{1}_{[0,t]}^{\otimes 2} = \frac{1}{2} \langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]} \rangle_{\mathcal{H}} (\mathbf{1}_{[0,s]} \otimes \mathbf{1}_{[0,t]} + \mathbf{1}_{[0,t]} \otimes \mathbf{1}_{[0,s]}).$$

Therefore,

$$\mathbb{E} \left[I_2(\mathbf{1}_{[0,r]}^{\otimes 2})I_2(\mathbf{1}_{[0,s]}^{\otimes 2})I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) \right] = 8\langle \mathbf{1}_{[0,r]}, \mathbf{1}_{[0,s]} \rangle_{\mathcal{H}} \langle \mathbf{1}_{[0,r]}, \mathbf{1}_{[0,t]} \rangle_{\mathcal{H}} \langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]} \rangle_{\mathcal{H}},$$

and then

$$\begin{aligned} \kappa_3(V_T(Z)) &= \frac{8}{T^{3/2}} \int_{[0,T]^3} \langle \mathbf{1}_{[0,r]}, \mathbf{1}_{[0,s]} \rangle_{\mathcal{H}} \langle \mathbf{1}_{[0,r]}, \mathbf{1}_{[0,t]} \rangle_{\mathcal{H}} \langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]} \rangle_{\mathcal{H}} dr ds dt \\ &= \frac{8}{T^{3/2}} \int_{[0,T]^3} \mathbb{E}[Z_r Z_s] \mathbb{E}[Z_r Z_t] \mathbb{E}[Z_s Z_t] dr ds dt \\ &= \frac{8}{T^{3/2}} \int_{[0,T]^3} \rho(r-s)\rho(r-t)\rho(s-t) dr ds dt \\ &= \frac{8}{T^{3/2}} \int_0^T \int_{-t}^{T-t} \int_{-t}^{T-t} \rho(x-y)\rho(x)\rho(y) dx dy dt \\ &= \frac{8}{T^{3/2}} \int_{-T}^T \int_{-T}^T \rho(x-y)\rho(x)\rho(y) dx dy \int_{(-x) \vee (-y) \vee 0}^{(T-x) \wedge (T-y) \wedge T} dt. \end{aligned}$$

Let $\rho_T(x) := |\rho(x)|\mathbf{1}_{|x| \leq T}$. Then,

$$\begin{aligned} \kappa_3(V_T(Z)) &\leq \frac{8}{\sqrt{T}} \int_{\mathbb{R}} \int_{\mathbb{R}} \rho_T(x-y)\rho(x)\rho(y) dx dy \\ &= \frac{8}{\sqrt{T}} \int_{\mathbb{R}} (\rho_T * \rho_T)(y)\rho_T(y) dy \\ &\leq \frac{8}{\sqrt{T}} \|\rho_T * \rho_T\|_{L^3(\mathbb{R})} \|\rho_T\|_{L^{3/2}(\mathbb{R})}, \end{aligned}$$

where $\rho_T * \rho_T$ is the convolution of the two functions and we have applied Hölder’s inequality in the last line. Now, recall Young’s inequality: If $p, q, r \geq 1$, $f \in L^p(\mathbb{R})$, $g \in L^q(\mathbb{R})$ and $1/p + 1/q = 1/r + 1$, then

$$\|f * g\|_{L^r(\mathbb{R})} \leq \|f\|_{L^p(\mathbb{R})} \|g\|_{L^q(\mathbb{R})}.$$

Hence,

$$\kappa_3(V_T(Z)) \leq \|\rho_T\|_{L^{3/2}(\mathbb{R})}^3 = \left(\int_{-T}^T |\rho(t)|^{3/2} dt \right)^2,$$

and (3.6) is established.

Similarly, we have

$$\begin{aligned} & \kappa_4(V_T(Z)) \\ &= \mathbb{E}[V_T(Z)^4] - 3\mathbb{E}[V_T(Z)^2]^2 \\ &= \frac{1}{T^2} \mathbb{E} \left[\left(\int_0^T I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) \right)^4 \right] - \frac{3}{T^2} \mathbb{E} \left[\left(\int_0^T I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) \right)^2 \right]^2 \\ &= \frac{1}{T^2} \int_{[0,T]^4} \mathbb{E} \left[I_2(\mathbf{1}_{[0,s]}^{\otimes 2}) I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) I_2(\mathbf{1}_{[0,u]}^{\otimes 2}) I_2(\mathbf{1}_{[0,v]}^{\otimes 2}) \right] ds dt du dv \\ &\quad - \frac{3}{T^2} \left(\int_{[0,T]^2} \mathbb{E} \left[I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) I_2(\mathbf{1}_{[0,s]}^{\otimes 2}) \right] dt ds \right)^2. \end{aligned}$$

By (6.2) and the isometry property

$$\begin{aligned} & \mathbb{E} \left[I_2(\mathbf{1}_{[0,s]}^{\otimes 2}) I_2(\mathbf{1}_{[0,t]}^{\otimes 2}) I_2(\mathbf{1}_{[0,u]}^{\otimes 2}) I_2(\mathbf{1}_{[0,v]}^{\otimes 2}) \right] \\ &= \mathbb{E} \left[I_4(\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_0 \mathbf{1}_{[0,t]}^{\otimes 2}) I_4(\mathbf{1}_{[0,u]}^{\otimes 2} \tilde{\otimes}_0 \mathbf{1}_{[0,v]}^{\otimes 2}) \right] \\ &\quad + 16 \mathbb{E} \left[I_2(\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_1 \mathbf{1}_{[0,t]}^{\otimes 2}) I_2(\mathbf{1}_{[0,u]}^{\otimes 2} \tilde{\otimes}_1 \mathbf{1}_{[0,v]}^{\otimes 2}) \right] \\ &\quad + \mathbb{E} \left[I_0(\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_2 \mathbf{1}_{[0,t]}^{\otimes 2}) I_0(\mathbf{1}_{[0,u]}^{\otimes 2} \tilde{\otimes}_2 \mathbf{1}_{[0,v]}^{\otimes 2}) \right]. \end{aligned}$$

For the contractions the following holds:

$$\begin{aligned} \mathbf{1}_{[0,s]}^{\otimes 2} \otimes_0 \mathbf{1}_{[0,t]}^{\otimes 2} &= \mathbf{1}_{[0,s]} \otimes \mathbf{1}_{[0,s]} \otimes \mathbf{1}_{[0,t]} \otimes \mathbf{1}_{[0,t]}, \\ \mathbf{1}_{[0,s]}^{\otimes 2} \otimes_1 \mathbf{1}_{[0,t]}^{\otimes 2} &= \langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]} \rangle_{\mathcal{H}} \mathbf{1}_{[0,s]} \otimes \mathbf{1}_{[0,t]}, \\ \mathbf{1}_{[0,s]}^{\otimes 2} \otimes_2 \mathbf{1}_{[0,t]}^{\otimes 2} &= \langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]} \rangle_{\mathcal{H}} \langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]} \rangle_{\mathcal{H}}. \end{aligned}$$

After taking symmetrization into account and by the isometry formula (2.4),

$$\begin{aligned} \mathbb{E} \left[I_4(\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_0 \mathbf{1}_{[0,t]}^{\otimes 2}) I_4(\mathbf{1}_{[0,u]}^{\otimes 2} \tilde{\otimes}_0 \mathbf{1}_{[0,v]}^{\otimes 2}) \right] &= 4 (\mathbb{E}[Z_s Z_u] \mathbb{E}[Z_t Z_v])^2 \\ &\quad + 4 (\mathbb{E}[Z_s Z_u] \mathbb{E}[Z_t Z_v])^2 \end{aligned}$$

$$\begin{aligned}
 &+ 16\mathbb{E}[Z_s Z_u]\mathbb{E}[Z_s Z_v]\mathbb{E}[Z_t Z_u]\mathbb{E}[Z_t Z_v], \\
 \mathbb{E} \left[I_2(\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_1 \mathbf{1}_{[0,t]}^{\otimes 2}) I_2(\mathbf{1}_{[0,u]}^{\otimes 2} \tilde{\otimes}_1 \mathbf{1}_{[0,v]}^{\otimes 2}) \right] &= \mathbb{E}[Z_s Z_t]\mathbb{E}[Z_u Z_v]\mathbb{E}[Z_s Z_u]\mathbb{E}[Z_t Z_v] \\
 &+ \mathbb{E}[Z_s Z_t]\mathbb{E}[Z_u Z_v]\mathbb{E}[Z_s Z_v]\mathbb{E}[Z_t Z_u], \\
 \mathbb{E} \left[I_0(\mathbf{1}_{[0,s]}^{\otimes 2} \tilde{\otimes}_2 \mathbf{1}_{[0,t]}^{\otimes 2}) I_0(\mathbf{1}_{[0,u]}^{\otimes 2} \tilde{\otimes}_2 \mathbf{1}_{[0,v]}^{\otimes 2}) \right] &= (\mathbb{E}[Z_s Z_t]\mathbb{E}[Z_u Z_v])^2.
 \end{aligned}$$

Therefore by symmetry and (3.3),

$$\begin{aligned}
 \kappa_4(V_T(Z)) &= \frac{48}{T^2} \int_{[0,T]^4} \mathbb{E}[Z_s Z_u]\mathbb{E}[Z_s Z_v]\mathbb{E}[Z_t Z_u]\mathbb{E}[Z_t Z_v] ds dt du dv \\
 &+ \frac{4 + 4 + 1 - 3 * 4}{T^2} \int_{[0,T]^4} \mathbb{E}[Z_s Z_t]^2 \mathbb{E}[Z_u Z_v]^2 ds dt du dv \\
 &\leq \frac{48}{T^2} \int_{[0,T]^4} \rho(s-u)\rho(s-v)\rho(t-u)\rho(t-v) ds dt du dv \\
 &= \frac{48}{T^2} \int_{[0,T]^2} \int_{\mathbb{R}^2} \rho_T(u-s)\rho_T(v-s)\rho_T(t-u)\rho_T(t-v) du dv ds dt \\
 &= \frac{48}{T^2} \int_{[0,T]^2} \int_{\mathbb{R}^2} \rho_T(t-s-x)\rho_T(t-s-y)\rho_T(x)\rho_T(y) dx dy ds dt \\
 &= \frac{48}{T^2} \int_{[0,T]^2} ((\rho_T * \rho_T)(t-s))^2 ds dt \leq \frac{48}{T} \|\rho_T * \rho_T\|_{L^2(\mathbb{R})}^2,
 \end{aligned}$$

where we have also applied the change of variables $x = t - u$, $y = t - v$ and used the formula for convolution. Next, by Young’s inequality,

$$\|\rho_T * \rho_T\|_{L^2(\mathbb{R})}^2 \leq \|\rho_T\|_{L^{4/3}(\mathbb{R})}^4,$$

and this establishes (3.7). □

6.2. A bound on $\mathbb{E}|\hat{f}_T(\mathbf{X})|^p$

First, we recall an important representation for stationary Gaussian processes.

Definition 6.1. Let ϕ_S denote the set of functions $\xi \in L^2(\mathbb{R})$ such that $\xi(t) = 0$ for all $t < 0$. If $\xi \in \phi_S$, we can define for all $t \in \mathbb{R}$,

$$Z_t^\xi := \int_{\mathbb{R}} \xi(t-u) dW_u, \tag{6.3}$$

where $(W_t)_{t \geq 0}$ is the Wiener process. The process $(Z_t^\xi)_{t \in \mathbb{R}}$ is a stationary centered Gaussian process.

The following result is key for our approach.

Theorem 6.2 (Karhunen [23]). *Let $\{Z_t, t \in \mathbb{R}\}$ be a stationary centered Gaussian process such that*

$$\lim_{t \rightarrow 0} \mathbb{E}[(Z_t - Z_0)^2] = 0, \quad \bigcap_{t \in \mathbb{R}} \overline{sp}\{Z_s : -\infty < s \leq t\} = \{0\}, \tag{6.4}$$

where $\overline{\text{sp}}$ denotes the L^2 -closure of the linear span of a set of square-integrable random variables. Then there exists $\xi \in \phi_S$ such that

$$\{Z_t, t \in \mathbb{R}\} \stackrel{(d)}{=} \{Z_t^\xi, t \in \mathbb{R}\}, \tag{6.5}$$

where (d) denotes equality of all finite-dimensional distributions, and $\{Z_t^\xi, t \in \mathbb{R}\}$ is the stationary centered Gaussian process defined in (6.3).

The following proposition asserts that, for Z a stationary centered Gaussian process satisfying (6.4), the estimator $\widehat{f}_T(Z)$ given by (1.1) admits uniformly bounded negative moments of all order. To prove it, we mainly rely on the equality in law (6.5), that allows us to make use of conditional expectations with respect to the filtration generated by the Wiener process W of (6.3). Another more classic approach could have been to use a Laplace transform argument instead, but we have not figured out how to do it.

Proposition 6.3. *Let $\{Z_t, t \geq 0\}$ be a stationary centered Gaussian process satisfying (6.4), and $\mathbb{E}[Z_0^2] = 1$. Then, for every $p > 0$, there is $T_0 > 0$ such that*

$$\sup_{T \geq T_0} \mathbb{E} \left[\left(\frac{1}{T} \int_0^T Z_t^2 dt \right)^{-p} \right] < \infty \tag{6.6}$$

$$\text{and } \sup_{T \geq T_0} \mathbb{E} \left[\left(\frac{1}{T} \int_0^T (Z_t - e^{\theta t} Z_0)^2 dt \right)^{-p} \right] < \infty.$$

Moreover, for every $p > 0$, there is $n_0 \geq 1$ such that

$$\sup_{n \geq n_0} \mathbb{E} \left[\left(\frac{1}{n} \sum_{i=1}^n Z_{t_i}^2 \right)^{-p} \right] < \infty \tag{6.7}$$

$$\text{and } \sup_{n \geq n_0} \mathbb{E} \left[\left(\frac{1}{n} \sum_{i=1}^n (Z_{t_i} - e^{\theta t_i} Z_0)^2 \right)^{-p} \right] < \infty,$$

where $t_i = i\Delta_n$, for $i = 1, \dots, n$, with $n\Delta_n \rightarrow \infty$ and $\Delta_n \rightarrow 0$, as $n \rightarrow \infty$.

Remark 6.4. We note the main idea for the proof is inspired by the approach in [31, Theorem 1.1], which provides a bound on the some negative moments of the Malliavin derivative.

Proof. Let $p > 0$. Condition (6.4) is satisfied and thus Theorem 6.2 implies that $\{Z_t, t \in \mathbb{R}_+\} \stackrel{(d)}{=} \{Z_t^\xi, t \in \mathbb{R}_+\}$ where $\xi \in L^2(\mathbb{R})$ with $\xi(t) = 0$, for $t \leq 0$, and Z_t^ξ is given by (6.3). Then, for a positive integer $m > 2p$,

$$\mathbb{E} \left[\left(\frac{1}{T} \int_0^T Z_t^2 dt \right)^{-p} \right] = \mathbb{E} \left[\left(\frac{1}{T} \int_0^T (Z_t^\xi)^2 dt \right)^{-p} \right]$$

$$= \mathbb{E} \left[\left(\sum_{k=1}^m \frac{1}{T} \int_{(k-1)T/m}^{kT/m} (Z_t^\xi)^2 dt \right)^{-p} \right].$$

Recall, that by the inequality between arithmetic and geometric means, for any positive reals x_1, \dots, x_m , $\sum_{k=1}^m x_k \geq m \prod_{k=1}^m x_k^{1/m}$. Then,

$$\mathbb{E} \left[\left(\frac{1}{T} \int_0^T Z_t^2 dt \right)^{-p} \right] \leq m^{-p} \mathbb{E} \left[\prod_{k=1}^m \left(\frac{1}{T} \int_{(k-1)T/m}^{kT/m} (Z_t^\xi)^2 dt \right)^{-p/m} \right]. \tag{6.8}$$

We proceed by conditioning. Let $\mathcal{F}_t := \sigma(W_u, u \leq t)$. By definition, Z_t^ξ is \mathcal{F}_t -measurable. Thus,

$$\begin{aligned} & \mathbb{E} \left[\left(\frac{1}{T} \int_0^T Z_t^2 dt \right)^{-p} \right] \\ & \leq m^{-p} \mathbb{E} \left[\prod_{k=1}^m \mathbb{E} \left[\left(\frac{1}{T} \int_{(k-1)T/m}^{kT/m} (Z_t^\xi)^2 dt \right)^{-p/m} \middle| \mathcal{F}_{(k-1)T/m} \right] \right]. \end{aligned}$$

Next, note that

$$\begin{aligned} & \mathbb{E} \left[\left(\frac{1}{T} \int_{(k-1)T/m}^{kT/m} (Z_t^\xi)^2 dt \right)^{-p/m} \middle| \mathcal{F}_{(k-1)T/m}^W \right] \\ & = \int_0^\infty \mathbb{P} \left(\frac{1}{T} \int_{(k-1)T/m}^{kT/m} (Z_t^\xi)^2 dt \leq x^{-m/p} \middle| \mathcal{F}_{(k-1)T/m}^W \right) dx \\ & \leq 1 + \int_1^\infty \mathbb{P} \left(\frac{1}{T} \int_{(k-1)T/m}^{kT/m} (Z_t^\xi)^2 dt \leq x^{-m/p} \middle| \mathcal{F}_{(k-1)T/m}^W \right) dx. \end{aligned} \tag{6.9}$$

Applying the Carbery-Wright Inequality [6], there is a universal constant $c > 0$ such that, for any $\varepsilon > 0$ we can write

$$\begin{aligned} & \mathbb{P} \left(\frac{1}{T} \int_{(k-1)T/m}^{kT/m} (Z_t^\xi)^2 dt \leq \varepsilon \middle| \mathcal{F}_{(k-1)T/m}^W \right) \\ & \leq \frac{c\sqrt{\varepsilon}}{\mathbb{E} \left[\frac{1}{T} \int_{(k-1)T/m}^{kT/m} (Z_t^\xi)^2 dt \middle| \mathcal{F}_{(k-1)T/m}^W \right]}. \end{aligned} \tag{6.10}$$

Next, note that for any $0 \leq a < b$,

$$\begin{aligned} \mathbb{E} \left[\int_a^b (Z_t^\xi)^2 dt \middle| \mathcal{F}_a \right] & = \int_a^b \mathbb{E} \left[(Z_t^\xi)^2 \middle| \mathcal{F}_a \right] dt \\ & \geq \int_a^b 2Z_a^\xi \mathbb{E} [Z_t^\xi - Z_a^\xi | \mathcal{F}_a] + \mathbb{E} [(Z_t^\xi - Z_a^\xi)^2 | \mathcal{F}_a] dt \end{aligned}$$

$$\geq \int_a^b \int_a^t \xi^2(t-u) du dt = \int_0^{b-a} (b-a-v)\xi^2(v) dv, \tag{6.11}$$

where we have used Itô isometry and the fact that $Z_t^\xi - Z_a^\xi$ is independent of \mathcal{F}_a .

By isometry, $\int_0^\infty \xi^2(v) dv = \mathbb{E}[Z_0^2] = 1$. As a result, there is $T_0 > 0$ such that $\int_0^{\frac{T_0}{2m}} \xi^2(v) dv \geq \frac{1}{2}$. Thus, by (6.11), for every $T \geq T_0$,

$$\mathbb{E} \left[\frac{1}{T} \int_{(k-1)T/m}^{kT/m} (Z_t^\xi)^2 dt \mid \mathcal{F}_{(k-1)T/m} \right] \geq \frac{1}{2m} \int_0^{\frac{T}{2m}} \xi^2(v) dv \geq \frac{1}{4m}. \tag{6.12}$$

Therefore, combining (6.9), (6.10) and (6.12), we obtain for every $T \geq T_0$,

$$\mathbb{E} \left[\left(\frac{1}{T} \int_{(k-1)T/m}^{kT/m} Z_t^2 dt \right)^{-p/m} \mid \mathcal{F}_{(k-1)T/m} \right] \leq 1 + 4cm \int_1^\infty x^{-\frac{m}{2p}} dx. \tag{6.13}$$

Hence,

$$\gamma_{m,T_0} := \sup_{\substack{T \geq T_0 \\ 1 \leq k \leq m}} \mathbb{E} \left[\left(\frac{1}{T} \int_{(k-1)T/m}^{kT/m} Z_t^2 dt \right)^{-p/m} \mid \mathcal{F}_{(k-1)T/m} \right] < \infty. \tag{6.14}$$

Consequently, it follows from (6.8) and (6.14) that, for all $T \geq T_0$,

$$\mathbb{E} \left[\left(\frac{1}{T} \int_0^T Z_t^2 dt \right)^{-p} \right] \leq m^{-p} (\gamma_{m,T_0})^m < \infty,$$

which completes the proof of the first part of (6.6). To establish the second part of (6.6) one need only replace Z_t by $Z_t - e^{\theta t} Z_0$ in the proof above. Indeed, the key inequalities (6.8), (6.9) and (6.10) are the same (with the change $Z_t^\xi \rightarrow Z_t^\xi - e^{\theta t} Z_0^\xi$) and the equivalent of (6.11) is

$$\begin{aligned} & E \left[\int_a^b (Z_t^\xi - e^{\theta t} Z_0^\xi)^2 dt \mid \mathcal{F}_a \right] \\ & \geq \int_a^b 2(Z_a^\xi - e^{\theta t} Z_0^\xi) \mathbb{E}[Z_t^\xi - Z_a^\xi \mid \mathcal{F}_a] + \mathbb{E}[(Z_t^\xi - Z_a^\xi)^2 \mid \mathcal{F}_a] dt \\ & \geq \int_a^b \int_a^t \xi^2(t-u) du dt = \int_0^{b-a} (b-a-v)\xi^2(v) dv. \end{aligned} \tag{6.15}$$

Now, let us prove the discrete version (6.8). First we prove it for the case $n = m^2$. Define, for every $m \geq 1$, $T_m := m^2 \Delta_m$ such that $\Delta_m \rightarrow 0$, and $\frac{T}{m} = m \Delta_m \rightarrow \infty$ as $m \rightarrow \infty$. Fix $p > 0$, and let m_0 be a positive integer such

that, for every $m \geq m_0$, $m > 2p$ and $\int_0^{\frac{T_m}{m}} \xi^2(v)dv \geq \frac{1}{2}$. We have $t_i = i\Delta_m$ for $i = 1, \dots, m$. Write

$$\frac{1}{m^2} \sum_{i=1}^{m^2} Z_{t_i}^2 = \frac{1}{T_m} \int_0^{T_m} Y_t dt,$$

where $Y_t := \sum_{i=1}^{m^2} Z_{t_i} 1_{(t_{i-1}, t_i]}(t)$. Also, denote $Y_t^\xi := \sum_{i=1}^{m^2} Z_{t_i}^\xi 1_{(t_{i-1}, t_i]}(t)$.

We follow the same techniques as in the proof of (6.6). Notice that the inequalities (6.8)-(6.11) hold with Y_t^ξ instead of Z_t^ξ .

Thus, it suffices to prove the following equivalent of (6.12): for every $m \geq m_0$,

$$\mathbb{E} \left[\frac{1}{T_m} \int_{(k-1)T_m/m}^{kT_m/m} (Y_t^\xi)^2 dt \middle| \mathcal{F}_{(k-1)T_m/m} \right] \geq \frac{1}{2m^2}.$$

Notice that for every $k = 1, \dots, m$,

$$\frac{1}{T_m} \int_{(k-1)T_m/m}^{kT_m/m} (Y_t^\xi)^2 dt = \frac{1}{m^2} \sum_{j=1}^m (Z_{t_{(k-1)m+j}}^\xi)^2.$$

Moreover,

$$\mathbb{E} \left[(Z_{t_{(k-1)m+j}}^\xi)^2 \middle| \mathcal{F}_{t_{(k-1)m}} \right] \geq \int_{t_{(k-1)m}}^{t_{(k-1)m+j}} \xi^2(t_{(k-1)m+j} - u) du = \int_0^{t_j} \xi^2(v) dv.$$

Therefore, for every $m \geq m_0$,

$$\begin{aligned} & \mathbb{E} \left[\frac{1}{T_m} \int_{(k-1)T_m/m}^{kT_m/m} (Y_t^\xi)^2 dt \middle| \mathcal{F}_{(k-1)T_m/m} \right] \\ & \geq \frac{1}{m^2} \sum_{j=1}^m \int_0^{t_j} \xi^2(v) dv \geq \frac{1}{m^2} \int_0^{t_m} \xi^2(v) dv \\ & = \frac{1}{m^2} \int_0^{\frac{T}{m}} \xi^2(v) dv \geq \frac{1}{2m^2}, \end{aligned}$$

which yields the proof of (6.8) for $n = m^2$.

For general n a simple computation yields

$$\frac{1}{n} \sum_{i=1}^n Z_{t_i}^2 \geq \frac{1}{n} \sum_{i=1}^{\lfloor \sqrt{n} \rfloor^2} Z_{t_i}^2 \geq C \frac{1}{\lfloor \sqrt{n} \rfloor^2} \sum_{i=1}^{\lfloor \sqrt{n} \rfloor^2} Z_{t_i}^2, \tag{6.16}$$

for some absolute constant $C > 0$, and thus the first part of (6.8) is established. The second part follows using the same techniques as above and (6.15). \square

Acknowledgments

We would like to thank the referee for his/her careful readings, constructive remarks and useful suggestions.

References

- [1] Alazemi, F., Alsenafi, A., Es-Sebaïy, K. (2020). Parameter estimation for Gaussian mean-reverting Ornstein-Uhlenbeck processes of the second kind: non-ergodic case. *Stochastics and Dynamics* 20(2), 2050011 (25 pages). [MR4080159](#)
- [2] Azmoodeh, E., Viitasaari, L. (2015) Parameter estimation based on discrete observations of fractional Ornstein-Uhlenbeck process of the second kind. *Stat Inference Stoch Process*, 18, 205-227. [MR3395605](#)
- [3] Bajja, S., Es-Sebaïy, K., Viitasaari, L. (2020). Volatility estimation in fractional Ornstein-Uhlenbeck models. *Stochastic Models*, 36(1), 94-111. [MR4067888](#)
- [4] Biermé, H., Bonami, A., Nourdin, I., Peccati, G. (2012). Optimal Berry-Esseen rates on the Wiener space: the barrier of third and fourth cumulants. *ALEA* 9, no. 2, 473-500. [MR3069374](#)
- [5] Brockwell, P., Davis, R. (2009) *Time Series: Theory and Methods* (Springer series in Statistics), second ed. Springer. [MR2839251](#)
- [6] Carbery, A., Wright, J. (2001) Distributional and L^q norm inequalities for polynomials over convex bodies in \mathbb{R}^n . *Math. Res. Lett.* 8, 3, 233-248. [MR1839474](#)
- [7] Chen, Y., Hu, Y., Long, H. (2019) Generalized moment estimators for α -stable Ornstein-Uhlenbeck motions from discrete observations. *Stat. Inference Stoch. Process.* 23, 1, 53-81. [MR4072252](#)
- [8] Chen, Y., Tian, L., Ying, L. (2020) Second moment estimator for an $AR(1)$ model driven by a long memory Gaussian noise. arXiv preprint [arXiv:2008.12443](#).
- [9] Cheridito, P. (2004). Gaussian moving averages, semimartingales and option pricing. *Stochastic processes and their applications*, 109(1), 47-68. [MR2024843](#)
- [10] Cheridito, P., Kawaguchi, H., Maejima, M. (2003). Fractional Ornstein-Uhlenbeck processes, *Electr. J. Probab.* 8, 1-14. [MR1961165](#)
- [11] Douissi, S., Es-Sebaïy, K., Viens, F. (2019). Berry-Esseen bounds for parameter estimation of general Gaussian processes. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 16, 633-664. [MR3949273](#)
- [12] El Machkouri, M., Es-Sebaïy, K., Ouknine, Y. (2016). Least squares estimator for non-ergodic Ornstein-Uhlenbeck processes driven by Gaussian processes. *Journal of the Korean Statistical Society* 45, 329-341. [MR3527650](#)
- [13] El Onsy, B., Es-Sebaïy, K., Viens, F. (2017). Parameter Estimation for a partially observed Ornstein-Uhlenbeck process with long-memory noise. *Stochastics*, 89(2), 431-468. [MR3590429](#)
- [14] Es-Sebaïy, K. (2013). Berry-Esseen bounds for the least squares estimator for discretely observed fractional Ornstein-Uhlenbeck processes. *Statist. Probab. Lett.* 83., 10, 2372-2385. [MR3093828](#)
- [15] Es-Sebaïy, K., Alazemi, F., Al-Foraih, M. (2019). Least squares type estimation for discretely observed non-ergodic Gaussian Ornstein-Uhlenbeck processes. *Acta Mathematica Scientia*, 39(4), 989-1002. [MR4066516](#)

- [16] Es-Sebaiy, K. and Es-Sebaiy, M. (2020). Estimating drift parameters in a non-ergodic Gaussian Vasicek-type model. *Statistical Methods and Applications*. [MR4266406](#)
- [17] Es-Sebaiy, K., Viens, F. (2019). Optimal rates for parameter estimation of stationary Gaussian processes. *Stochastic Processes and their Applications*, 129(9), 3018-3054. [MR3985554](#)
- [18] Jiang, H., Liu, J., Wang, S. (2018) Self-normalized asymptotic properties for the parameter estimation in fractional Ornstein-Uhlenbeck process. *Stochastics and Dynamics*. vol. 18, 5 (29 pages). [MR3956358](#)
- [19] Hu, Y., Nualart, D. (2010). Parameter estimation for fractional Ornstein-Uhlenbeck processes. *Statist. Probab. Lett.* **80**, 1030-1038. [MR2638974](#)
- [20] Hu, Y., Nualart, D., Zhou, H. (2019). Parameter estimation for fractional Ornstein-Uhlenbeck processes of general Hurst parameter. *Statistical Inference for Stochastic Processes*, 22(1), 111-142. [MR3918739](#)
- [21] Hu, Y., Song, J. (2013). Parameter estimation for fractional Ornstein-Uhlenbeck processes with discrete observations. F. Viens et al (eds), *Malliavin Calculus and Stochastic Analysis: A Festschrift in Honor of David Nualart*, 427-442, Springer. [MR3070455](#)
- [22] Kaarakka, T., Salminen, P. (2011). On Fractional Ornstein-Uhlenbeck process. *Communications on Stochastic Analysis*, 5 (1), 121-133. [MR2808539](#)
- [23] Karhunen, K., (1950). Über die Struktur stationärer zufälliger Funktionen. *Ark. Mat.* 1 (3), 141-160. [MR0034557](#)
- [24] Kosov, E. (2019). Total variation distance estimates via L2-norm for polynomials in log-concave random vectors. *International Mathematics Research Notices*, rnz278. <https://doi.org/10.1093/imrn/rnz278> [MR4338224](#)
- [25] Kloeden, P. and Neuenkirch, A. (2007). The pathwise convergence of approximation schemes for stochastic differential equations. *LMS J. Comp. Math.* **10**, 235-253. [MR2320830](#)
- [26] Křáží, P., Maslowski, B. (2019). Central limit theorems and minimum-contrast estimators for linear stochastic evolution equations. *Stochastics*, 91, 1109-1140. [MR4023510](#)
- [27] Nourdin, I. (2013) Lectures on Gaussian approximations with Malliavin calculus. *Séminaire de Probabilités XLV. Lecture Notes in Math*, 2078, Springer, Cham. [MR3185909](#)
- [28] Nourdin, I., Peccati, G. (2012). Normal approximations with Malliavin calculus: from Stein's method to universality. *Cambridge Tracts in Mathematics* 192. Cambridge University Press, Cambridge. [MR2962301](#)
- [29] Nourdin, I., Peccati, G. (2015). The optimal fourth moment theorem. *Proc. Amer. Math. Soc.* 143, 3123-3133. [MR3336636](#)
- [30] Nourdin, I., Tran, T. D. (2019). Statistical inference for Vasicek-type model driven by Hermite processes. *Stochastic Processes and their Applications*, 129(10), 3774-3791. [MR3997661](#)
- [31] Nourdin, I., Nualart, D. (2013) Fisher Information and the Fourth Moment Theorem. *Ann. Inst. H. Poincaré Probab. Statist.*, 52, 2, 849-867. [MR3498012](#)
- [32] Nualart, D. (2006). *The Malliavin calculus and related topics*. Springer-

- Verlag, Berlin.* [MR2200233](#)
- [33] Nualart, D., Peccati, G. (2005). Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* 33, no. 1, 177-193. [MR2118863](#)
 - [34] Pham-Dinh-Tuan (1977). Estimation of parameters of a continuous time Gaussian stationary process with rational spectral density, *Biometrika* 64, 2, 385-399. [MR0501676](#)
 - [35] Schervish, M. (1997). *Theory of Statistics* (Springer series in Statistics) Springer. [MR1354146](#)
 - [36] Sottinen, T., Viitasaari, L. (2018). Parameter estimation for the Langevin equation with stationary-increment Gaussian noise. *Statistical Inference for Stochastic Processes*, 21(3), 569-601. [MR3846993](#)
 - [37] Wood, A., Chan, G. (1994). Simulation of stationary Gaussian processes. *Journal of computational and graphical statistics*, 3(4):409-432. [MR1323050](#)