Drift Estimation for Discretely Sampled SPDEs

3 authors:

Igor Cialenco
Illinois Institute of Technology
49 PUBLICATIONS   **284** CITATIONS

SEE PROFILE

Francisco J Delgado Vences
Universidad Nacional Autónoma de México
9 PUBLICATIONS   **6** CITATIONS

SEE PROFILE

Hyun-Jung Kim
University of Southern California
8 PUBLICATIONS   **9** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   MLEs for Stochastic PDEs View project

Drift estimation for discretely sampled SPDEs

Igor Cialenco [*]   Francisco Delgado-Vences [†]   Hyun-Jung Kim [‡]

ABSTRACT: The aim of this paper is to study the asymptotic properties of the maximum likelihood estimator (MLE) of the drift coefficient for fractional stochastic heat equation driven by an additive space-time noise. We consider the traditional for stochastic partial differential equations statistical experiment when the measurements are performed in the spectral domain, and in contrast to the existing literature, we study the asymptotic properties of the maximum likelihood (type) estimators (MLE) when both, the number of Fourier modes and the time go to infinity. In the first part of the paper we consider the usual setup of continuous time observations of the Fourier coefficients of the solutions, and show that the MLE is consistent, asymptotically normal and optimal in the mean-square sense. In the second part of the paper we investigate the natural time discretization of the MLE, by assuming that the first N Fourier modes are measured at M time grid points, uniformly spaced over the time interval $[0, T]$. We provide a rigorous asymptotic analysis of the proposed estimators when $N \to \infty$ and/or $T, M \to \infty$. We establish sufficient conditions on the growth rates of N, M and T, that guarantee consistency and asymptotic normality of these estimators.

1 Introduction

Undoubtedly, the stochastic partial differential equations (SPDEs) serve as a modern powerful modeling tool in describing the evolution of dynamical systems in the presence of spatial-temporal uncertainties with particular applications in fluid mechanics, oceanography, temperature anomalies, finance, economics, biological and ecological systems, and many other applied disciplines. Major breakthrough results have been established on the general analytical theory for SPDEs, such as existence, uniqueness and regularity properties of the solutions. For an in depth discussion of the theory of SPDEs and their various applications, we refer to recent monographs [LR17, LR18]. In contrast, the investigation of inverse problems for SPDEs, and in particular parameter estimation problems, are still in their emerging phase. We refer to the survey papers [Lot09, Cia18] and the monograph [LR17, Chapter 6] for an overview of the literature and existing methodologies on

[*]Department of Applied Mathematics, Illinois Institute of Technology
10 W 32nd Str, Building REC, Room 208, Chicago, IL 60616, USA
Email: cialenco@iit.edu, URL: http://math.iit.edu/~igor

[†]Instituto de Matemáticas, UNAM, Oaxaca de Juarez, México 61000
Email: delgado@im.unam.mx

[‡]Department of Applied Mathematics, Illinois Institute of Technology
10 W 32nd Str, Building REC, Room 208, Chicago, IL 60616, USA
Email: hkim129@iit.edu, URL: https://sites.google.com/view/hyun-jungkim

statistical inference for parabolic SPDEs. Most of the existing results are obtained within the so-called spectral approach, when it is assumed that the observer measures the values of one realization of the first N Fourier modes of the solution continuously over a finite time interval $[0, T]$. In such cases, usually the statistical problems are addressed via maximum likelihood estimators (MLEs), and the asymptotic properties of the estimators are studied in the large number of Fourier modes regime, $N \to \infty$, while time horizon T is fixed. The large time asymptotics regime $T \to \infty$, while N being fixed, usually falls in the realm of finite dimensional stochastic differential equations, which is a well-established research area. This asymptotic regime in the context of SPDEs was only briefly discussed in [LR17, CX15]. Only several works have been dedicated to parameter estimation problems for SPDEs in discrete sampling setup. In [PR97, PR02, PR03, Mar03], the authors investigate some version of the discretized MLEs for some particular equations. In [PT07], and more recently in [CH17, BT17, Cho19, BT19], using various approaches the authors study the estimation of the drift and/or volatility coefficients when the solution is sampled discretely in physical domain.

The aim of this work is to provide a rigorous and comprehensive asymptotic analysis of the time discretized MLE for the drift coefficient of a fractional heat equation driven by an additive space-time noise. The precise form of the considered equations and their well posedness are presented in Section 2. The main results of this work can be summarized as follows:

▷ In Section 3 we assume the same sampling scheme as in the existing literature on spectral approach, namely continuous time observations of the N Fourier modes for $t \in [0, T]$, and study the asymptotic properties of the MLE, when *both*, $N, T \to \infty$. We prove that this estimator is (strongly) consistent, and asymptotically normal. We give two proofs of the asymptotic normality, one based on Malliavin calculus, which we believe can be used, with slight modifications to other similar problems. Another proof uses classical results from general probability and explodes the particular structure and properties of the underlying problem. In particular, we show that the estimator is optimal in the mean-square sense.

▷ In Section 4 we consider the *natural time discretization of the MLE*, by assuming that the first N Fourier modes are measured at M time grid points, uniformly spaced over the time interval $[0, T]$. We study the asymptotic properties of the proposed estimator when $N, T, M \to \infty$. In particular, we prove that the estimator is consistent if $N, M, T \to \infty$, or $N, M \to \infty$ while T is fixed, and if $T^2 N^{\frac{4\beta}{d} - 1}/M^2 \to 0$, where d is the space dimension, and β is the power of the Laplacian. Moreover, if $4\beta < d$, then consistency holds true, when $N \to \infty$, when M, T are fixed. This, in particular implies that to estimate efficiently the drift parameter it is enough to observe the Fourier modes at one instant of time - a result that agrees with recent discoveries in [CH17, BT17] where the solution is sampled in physical domain. Under some additional technical assumptions on the growth rates of N, M and T, we also prove that the proposed estimator is also asymptotically normal, with the same rate of convergence $\sqrt{T} N^{\frac{\beta}{d} + \frac{1}{2}}$ as the MLE from continuous time observation setup.

Some technical proofs, auxiliary results and relevant elements of Malliavin calculus are deferred to Appendix.

Open problems and future work. A reasonable extension of the present work is to investigate similar estimators and problems given that the solution is observed discretely in the physical domain, in which case, one has to additionally approximate the Fourier modes by a sum. While analogous asymptotic properties are expected to hold true, rigourous proofs remain to be established. As already mentioned, most of the existing literature on parameter estimation for SPDEs is

focused on sampling the Fourier modes in continuous time. In particular, the MLE approach was successfully applied to nonlinear equations [CGH11, PS19], and to equations driven by a fractional noise [CLP09]. Besides MLEs, in [CGH18] the authors propose an alternative class of estimators, called trajectory fitting estimators, and a Bayesian approach to estimating drift coefficients for a class of SPDEs driven by multiplicative noise is considered in [CCG19]. It is imperative, from theoretical and practical point of view, to study the asymptotic properties of the discretized versions of the estimators proposed in the above mentioned works, especially by looking at various asymptotic regimes (large time-space sampling, small mesh size, etc).

2 Setup of the problem and some auxiliary results

Let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$ be a stochastic basis with usual assumptions, and let $\{w_j, \ j \geq 1\}$ be a collection of independent standard Brownian motions on this basis. Assume that G is a bounded and smooth domain in $\mathbb{R}^d$, and let us denote by $\boldsymbol{\Delta}$ the Laplace operator on G with zero boundary conditions. The corresponding scale of Sobolev spaces will be denoted by $H^s(G)$, or simply H^s, for $s \in \mathbb{R}$. It is well known (cf. [Shu01]) that: a) the set $\{h_k\}_{k\in\mathbb{N}}$ of eigenfunctions of $\boldsymbol{\Delta}$ forms a complete orthonormal system in $L^2(G)$; b) the corresponding eigenvalues $\nu_k, k \in \mathbb{N}$, can be arranged such that $0 < -\nu_1 \leq -\nu_2 \leq \ldots$, and there exists a positive constant ϖ so that

$$\lim_{k\to\infty} |\nu_k| k^{-2/d} = \varpi.$$

In what follows, we will use the notation $\lambda_k := \sqrt{-\nu_k}, \ k \in \mathbb{N}$, and $\Lambda = \sqrt{-\boldsymbol{\Delta}}$. Also, for two sequences of numbers $\{a_n\}$ and $\{b_n\}$, we will write $a_n \sim b_n$, if there exists a nonzero and finite number c such that $\lim_{n\to\infty} a_n/b_n = c$, and $a_n \simeq b_n$, if $\lim_{n\to\infty} a_n/b_n = 1$.

We consider the following stochastic PDE

$$\mathrm{d}U(t,x) + \theta(-\boldsymbol{\Delta})^\beta U(t,x)\,\mathrm{d}t = \sigma \sum_{k\in\mathbb{N}} \lambda_k^{-\gamma} h_k(x)\,\mathrm{d}w_k(t), \quad t \in [0,T], \ U(0,x) = U_0, \ x \in G, \quad (2.1)$$

where $\theta > 0$, $\beta > 0$, $\gamma \geq 0$, $\sigma > 0$, and $U_0 \in H^s(G)$ for some $s \in \mathbb{R}$.

Using standard arguments (cf. [Cho07, LR17, LR18]), it can be proved that if $2(\gamma - s)/d > 1$, then (2.1) has a unique solution U, weak in PDE sense and strong in probability sense, such that

$$U \in L^2(\Omega \times [0,T]; H^{s+\beta}) \cap L^2(\Omega; C((0,T); H^s)).$$

In what follows, we will assume that $s \geq 0$, and $2\gamma > d$. We denote by $u_k, k \in \mathbb{N}$, the Fourier coefficient of the solution U of (2.1) with respect to $h_k, k \in \mathbb{N}$, i.e. $u_k(t) = (U(t), h_k)_0, k \in \mathbb{N}$. Let H^N be the finite dimensional subspace of $L^2(G)$ generated by $\{h_k\}_{k=1}^N$, and denote by P_N the projection operator of $L^2(G)$ into H^N, and put $U^N = P_N U$, or equivalently $U^N := (u_1, \ldots, u_N)$. Clearly, the Fourier mode $u_k, k \in \mathbb{N}$, follows the dynamics of an Ornstein-Uhlenbeck process given by

$$\mathrm{d}u_k = -\theta\lambda_k^{2\beta} u_k\,\mathrm{d}t + \sigma\lambda_k^{-\gamma}\,\mathrm{d}w_k(t), \quad u_k(0) = (U_0, h_k), \ t \geq 0.$$

We denote by $\mathbb{P}_\theta^{N,T}$ the probability measure on $C([0,T]; H^N) \simeq C([0,T]; \mathbb{R}^N)$ generated by the U^N. The measures $\mathbb{P}_\theta^{N,T}$ are equivalent for different values of the parameter θ, and the Radon-Nikodym derivative, or likelihood ratio, has the form

$$\frac{\mathbb{P}_\theta^{N,T}}{\mathbb{P}_{\theta_0}^{N,T}}(U^N) = \exp\left(-\frac{(\theta - \theta_0)}{\sigma^2}\sum_{k=1}^N \lambda_k^{2\beta+2\gamma}\int_0^T u_k(t)\,\mathrm{d}u_k(t) - \frac{(\theta^2 - \theta_0^2)}{2\sigma^2}\sum_{k=1}^N \lambda_k^{4\beta+2\gamma}\int_0^T u_k^2(t)\,\mathrm{d}t\right).$$

By maximizing the log likelihood ratio with the respect to the parameter of interest θ, we obtain the Maximum Likelihood Estimator (MLE) for θ given by

$$\widehat{\theta}_{N,T} := -\frac{\sum_{k=1}^{N} \lambda_k^{2\beta+2\gamma} \int_0^T u_k(t) \, \mathrm{d}u_k(t)}{\sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t) \, \mathrm{d}t}, \quad N \in \mathbb{N}, \ T > 0. \tag{2.2}$$

Let us also compute the Fisher information related to $\mathbb{P}_\theta^{N,T}/\mathbb{P}_{\theta_0}^{N,T}$. For simplicity, set $U_0 = 0$. Namely,

$$\begin{aligned}
\mathcal{I}_{N,T} &:= \int \left| \frac{\partial}{\partial \theta} \log \frac{\mathrm{d}\mathbb{P}_\theta^{N,T}}{\mathrm{d}\mathbb{P}_{\theta_0}^{N,T}} \right|^2 \left(\frac{\mathrm{d}\mathbb{P}_{\theta_0}^{N,T}}{\mathrm{d}\mathbb{P}_\theta^{N,T}} \right)^{-1} \mathrm{d}\mathbb{P}_{\theta_0}^{N,T} \\
&= -\int \frac{\partial^2}{\partial \theta^2} \log \frac{\mathrm{d}\mathbb{P}_\theta^{N,T}}{\mathrm{d}\mathbb{P}_{\theta_0}^{N,T}} \left(\frac{\mathrm{d}\mathbb{P}_{\theta_0}^{N,T}}{\mathrm{d}\mathbb{P}_\theta^{N,T}} \right)^{-1} \mathrm{d}\mathbb{P}_{\theta_0}^{N,T} \\
&= \frac{1}{\sigma^2} \sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \mathbb{E}\left[\int_0^T u_k^2 \, \mathrm{d}t \right].
\end{aligned}$$

By direct evaluations, we have that

$$\mathbb{E}\left[\int_0^T u_k^2 \, \mathrm{d}t \right] = \frac{\sigma^2 \lambda_k^{-2\gamma-2\beta}}{2\theta_0} \left(T - \frac{1 - e^{-2\theta_0 \lambda_k^{2\beta} T}}{2\theta_0 \lambda_k^{2\beta}} \right),$$

which yields

$$\begin{aligned}
\mathcal{I}_{N,T} &= \frac{1}{2\theta_0} \sum_{k=1}^{N} \lambda_k^{2\beta} \left(T - \frac{1 - e^{-2\theta_0 \lambda_k^{2\beta} T}}{2\theta_0 \lambda_k^{2\beta}} \right) \simeq \frac{T}{2\theta_0} \sum_{k=1}^{N} \lambda_k^{2\beta}, \quad \text{as } T \to \infty \\
&\simeq \frac{\varpi^\beta d T N^{\frac{2\beta}{d}+1}}{(4\beta + 2d)\,\theta_0}, \quad \text{as } N, T \to \infty.
\end{aligned} \tag{2.3}$$

In particular, note that $\mathcal{I}_{N,T} \to \infty$, when $N, T \to \infty$.

3 Asymptotics in large time and large number of Fourier modes

It is known that the estimator $\widehat{\theta}_{N,T}$ is unbiased, strongly consistent and asymptotically normal in two asymptotic regimes: $N \to \infty$ and T fixed, and $T \to \infty$ and N fixed; see for instance [CX15, Cia18] and references therein. In particular, for every fixed $T > 0$, $\lim_{N \to \infty} \theta_{N,T} = \theta_0$, with probability one, and

$$\text{w-}\lim_{N \to \infty} N^{\beta/d + \frac{1}{2}} \left(\widehat{\theta}_{N,T} - \theta_0 \right) = \mathcal{N}\left(0, \frac{(4\beta/d + 2)\theta_0}{\varpi^\beta T} \right), \tag{3.1}$$

where w-lim denotes the limit in distribution[1], and $\mathcal{N}(0, \bar{\sigma}^2)$ is a Gaussian random variable[2] with mean zero and variance $\bar{\sigma}^2$. Similarly, for every fixed $N \in \mathbb{N}$, $\lim_{T \to \infty} \theta_{N,T} = \theta_0$, with probability one, and

$$\text{w-}\lim_{T \to \infty} \sqrt{T} \left(\widehat{\theta}_{N,T} - \theta_0 \right) = \mathcal{N}(0, 2\theta_0/J), \tag{3.2}$$

---

[1] Whenever convenient, we will also use the notation '$\xrightarrow{d}$' to denote the convergence in distribution of random variables.

[2] Throughout the text we will use the notation $\mathcal{N}(\mu_0, \sigma_0^2)$ to denote a Gaussian random variable with mean μ_0 and variance σ_0^2.

where $J = \sum_{k=1}^{N} \lambda_k^{2\beta}$.

To the best of our knowledge, the asymptotic properties of $\widehat{\theta}_{N,T}$ when (both) $T, N \to \infty$ is not studied in the current literature. Besides this being an important question alone, the obtained results in this section will also serve as theoretical basis for investigating the statistical properties of the discretized version of the MLE studied later in this paper. In view of the above, naturally one should expect that the joint time-space consistency $\lim_{N,T\to\infty} \theta_{N,T} = \theta_0$ is satisfied. On the other hand, by (3.1)

$$\text{w-}\lim_{T\to\infty} \lim_{N\to\infty} \sqrt{T} N^{\beta/d+\frac{1}{2}} (\widehat{\theta}_{N,T} - \theta_0) = \mathcal{N}\left(0, \frac{(4\beta/d + 2)\theta_0}{\varpi^\beta}\right), \tag{3.3}$$

and by (3.2) same result holds for the swaped limiting order $T \to \infty, N \to \infty$. While (3.3) does not have great statistical meaning, it leads to a reasonable ansatz that same identity (3.3) should be satisfied when both $N, T \to \infty$. Also note that, in view of (2.3), this estimator is also optimal in the mean-square sense, having the rate of convergence dictated by the Fisher information. Next we give a rigourous proof of these results.

Theorem 3.1. *Assume that $\beta > 1/2$ and $\gamma > d/2$. Then, $\widehat{\theta}_{N,T}$ is strongly consistent, i.e.*

$$\lim_{N,T\to\infty} \widehat{\theta}_{N,T} = \theta_0, \quad \text{with probability one,} \tag{3.4}$$

and asymptotically normal, i.e.

$$\text{w-}\lim_{N,T\to\infty} \sqrt{T} N^{\frac{\beta}{d}+\frac{1}{2}} \left(\widehat{\theta}_{N,T} - \theta_0\right) = \mathcal{N}\left(0, \frac{(4\beta/d + 2)\theta_0}{\varpi^\beta}\right). \tag{3.5}$$

Proof. For simplicity, we set $U_0 = 0$, and hence $u_k(0) = 0$ for all $k \geq 1$. Since

$$u_k(t) = \sigma \lambda_k^{-\gamma} \int_0^t e^{-\theta_0 \lambda_k^{2\beta}(t-s)} \, dw_k(s), \ k \geq 1, \tag{3.6}$$

it is straightforward to show that

$$\mathbb{E}u_k^2(t) = \sigma^2 \lambda_k^{-2\beta-2\gamma} \frac{\left(1 - e^{-2\theta_0 \lambda_k^{2\beta} t}\right)}{2\theta_0}, \tag{3.7}$$

$$\mathbb{E}u_k^4(t) = 3\sigma^4 \lambda_k^{-4\beta-4\gamma} \frac{\left(1 - e^{-2\theta_0 \lambda_k^{2\beta} t}\right)^2}{(2\theta_0)^2}. \tag{3.8}$$

We note that

$$\begin{aligned}
\widehat{\theta}_{N,T} - \theta_0 &= -\frac{\sigma \sum_{k=1}^{N} \lambda_k^{2\beta+\gamma} \int_0^T u_k(t) \, dw_k(t)}{\sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t) \, dt} \\
&= -\frac{\sigma \sum_{k=1}^{N} \xi_{k,T}}{\sum_{k=1}^{N} \mathbb{V}\mathrm{ar}\,(\xi_{k,T})} \cdot \frac{\sum_{k=1}^{N} \mathbb{V}\mathrm{ar}\,(\xi_{k,T})}{\sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t) \, dt},
\end{aligned} \tag{3.9}$$

where

$$\xi_{k,T} := \lambda_k^{2\beta+\gamma} \int_0^T u_k(t) \, dw_k(t).$$

To show consistency (3.4), we will use the strong law of large numbers [Shi96, Theorem IV.3.2]. From (3.7), we have

$$\mathbb{V}\mathrm{ar}(\xi_{k,T}) = \lambda_k^{4\beta+2\gamma} \int_0^T \mathbb{E}u_k^2(t)\,\mathrm{d}t = \sigma^2 \lambda_k^{2\beta} \int_0^T \frac{1 - e^{-2\theta_0 \lambda_k^{2\beta} t}}{2\theta_0}\,\mathrm{d}t \simeq \frac{\sigma^2 \lambda_k^{2\beta} T}{2\theta_0}, \quad \text{as } T \to \infty,$$

and thus,

$$\sum_{k=1}^N \mathbb{V}\mathrm{ar}(\xi_{k,T}) \simeq \frac{\sigma^2 T}{2\theta_0} \sum_{k=1}^N \lambda_k^{2\beta}, \quad \text{as } T \to \infty$$

$$\simeq \frac{\sigma^2 \varpi^\beta dT N^{\frac{2\beta}{d}+1}}{(4\beta + 2d)\theta_0}, \quad \text{as } N, T \to \infty. \tag{3.10}$$

Moreover, using (3.8), we get that

$$\mathbb{V}\mathrm{ar}\left(\lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t)\,\mathrm{d}t\right) \le \mathbb{E}\left(\lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t)\,\mathrm{d}t\right)^2 \le \lambda_k^{8\beta+4\gamma} T \int_0^T \mathbb{E}u_k^4(t)\,\mathrm{d}t$$

$$\sim \lambda_k^{4\beta} T^2, \quad \text{as } T \to \infty. \tag{3.11}$$

Hence, there exists $T_0 > 0$ such that for all $T \ge T_0$,

$$\sum_{N=1}^\infty \frac{\mathbb{V}\mathrm{ar}\left(\xi_{N,T}\right)}{\left(\sum_{k=1}^N \mathbb{V}\mathrm{ar}\left(\xi_{k,T}\right)\right)^2} \le \frac{C_1}{T} \sum_{N=1}^\infty \frac{\lambda_N^{2\beta}}{\left(\sum_{k=1}^N \lambda_k^{2\beta}\right)^2} \le \frac{C_2}{T} \sum_{N=1}^\infty \frac{1}{N^{2+\frac{2\beta}{d}}} \le C_3 < \infty,$$

$$\sum_{N=1}^\infty \frac{\mathbb{V}\mathrm{ar}\left(\lambda_N^{4\beta+2\gamma} \int_0^T u_N^2(t)\,\mathrm{d}t\right)}{\left(\sum_{k=1}^N \mathbb{V}\mathrm{ar}\left(\xi_{k,T}\right)\right)^2} \le C_4 \sum_{N=1}^\infty \frac{\lambda_N^{4\beta}}{\left(\sum_{k=1}^N \lambda_k^{2\beta}\right)^2} \le C_5 \sum_{N=1}^\infty \frac{1}{N^2} < \infty,$$

where $C_1, C_2, C_3, C_4, C_5 > 0$ are some constants[3] independent of T. Using the uniform boundedness of the above series, and employing the strong law of large numbers, we deduce that for every $\varepsilon > 0$ and $T \ge T_0$, there exists $N_0 > 0$ independent of T such that for $N \ge N_0$,

$$\left| \frac{\sigma \sum_{k=1}^N \xi_{k,T}}{\sum_{k=1}^N \mathbb{V}\mathrm{ar}\left(\xi_{k,T}\right)} \right| < \varepsilon, \quad \text{and} \quad \left| \frac{\sum_{k=1}^N \mathbb{V}\mathrm{ar}\left(\xi_{k,T}\right)}{\sum_{k=1}^N \lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t)\,\mathrm{d}t} - 1 \right| < \varepsilon$$

with probability one. Therefore,

$$\lim_{N,T\to\infty} \frac{\sigma \sum_{k=1}^N \xi_{k,T}}{\sum_{k=1}^N \mathbb{V}\mathrm{ar}\left(\xi_{k,T}\right)} = 0 \quad \text{and} \quad \lim_{N,T\to\infty} \frac{\sum_{k=1}^N \mathbb{V}\mathrm{ar}\left(\xi_{k,T}\right)}{\sum_{k=1}^N \lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t)\,\mathrm{d}t} = 1 \tag{3.12}$$

with probability one. From here, and using (3.9), the proof of (3.4) is complete.

Next, we will prove asymptotic normality property (3.5), starting with representation

$$\widehat{\theta}_{N,T} - \theta_0 = -\frac{\sigma \sum_{k=1}^N \xi_{k,T}}{\left(\sum_{k=1}^N \mathbb{V}\mathrm{ar}(\xi_{k,T})\right)^{1/2}} \cdot \frac{1}{\left(\sum_{k=1}^N \mathbb{V}\mathrm{ar}(\xi_{k,T})\right)^{1/2}} \cdot \frac{\sum_{k=1}^N \mathbb{V}\mathrm{ar}(\xi_{k,T})}{\sum_{k=1}^N \lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t)\,\mathrm{d}t}. \tag{3.13}$$

---

[3]Notoriously, we will denote by C with subindexes generic constants that may change from line to line.

Let us consider the first term in (3.13). We will show that

$$\text{w-}\lim_{N,T\to\infty} \frac{\sigma \sum_{k=1}^{N} \xi_{k,T}}{\left(\sum_{k=1}^{N} \mathbb{V}\text{ar}(\xi_{k,T})\right)^{1/2}} = \mathcal{N}(0,\sigma^2).$$

By Burkholder–Davis–Gundy inequality and Cauchy–Schwartz inequality, we have

$$\mathbb{E}\xi_{k,T}^4 = \mathbb{E}\left(\lambda_k^{2\beta+\gamma}\int_0^T u_k(t)\,\mathrm{d}w_k(t)\right)^4 \leq C_1\lambda_k^{8\beta+4\gamma}\mathbb{E}\left(\int_0^T u_k^2(t)\,\mathrm{d}t\right)^2$$

$$\leq C_1\lambda_k^{8\beta+4\gamma}T\int_0^T \mathbb{E}u_k^4(t)\,\mathrm{d}t,$$

for some $C_1 > 0$. By (3.8) and (3.11), there exists $T_1 > 0$ such that for all $T \geq T_1$, $\mathbb{E}\xi_{k,T}^4 \leq C_2\lambda_k^{4\beta}T^2$, for some $C_2 > 0$, and hence there exists $N_1 > 0$ independent of T and T_1 such that for all $N \geq N_1$ and for all $T \geq T_1$,

$$\sum_{k=1}^{N} \mathbb{E}\xi_{k,T}^4 \leq C_3 N^{\frac{4\beta}{d}+1}T^2,$$

for some $C_3 > 0$, independent of N, and T. We will verify the classical Lindeberg condition [Shi96, Theorem III.5.1], namely that for every $\varepsilon > 0$,

$$\lim_{N,T\to\infty} \frac{\sum_{k=1}^{N} \mathbb{E}\left(\xi_{k,T}^2 \mathbb{1}_{\left\{|\xi_{k,T}|>\varepsilon\sqrt{\sum_{k=1}^{N} \mathbb{V}\text{ar}(\xi_{k,T})}\right\}}\right)}{\sum_{k=1}^{N} \mathbb{V}\text{ar}(\xi_{k,T})} = 0.$$

By Cauchy-Schwartz inequality and Chebyshev inequality,

$$\sum_{k=1}^{N} \mathbb{E}\left(\xi_{k,T}^2 \mathbb{1}_{\left\{|\xi_{k,T}|>\varepsilon\sqrt{\sum_{k=1}^{N} \mathbb{V}\text{ar}(\xi_{k,T})}\right\}}\right) \leq \sum_{k=1}^{N} \left(\mathbb{E}\xi_{k,T}^4\right)^{1/2}\left(\mathbb{P}\left(|\xi_{k,T}|>\varepsilon\sqrt{\sum_{k=1}^{N} \mathbb{V}\text{ar}(\xi_{k,T})}\right)\right)^{1/2}$$

$$\leq \frac{\sum_{k=1}^{N} \mathbb{E}\xi_{k,T}^4}{\varepsilon^2 \sum_{k=1}^{N} \mathbb{V}\text{ar}(\xi_{k,T})}.$$

Consequently,

$$\frac{\sum_{k=1}^{N} \mathbb{E}\left(\xi_{k,T}^2 \mathbb{1}_{\left\{|\xi_{k,T}|>\varepsilon\sqrt{\sum_{k=1}^{N} \mathbb{V}\text{ar}(\xi_{k,T})}\right\}}\right)}{\sum_{k=1}^{N} \mathbb{V}\text{ar}(\xi_{k,T})} \leq \frac{\sum_{k=1}^{N} \mathbb{E}\xi_{k,T}^4}{\varepsilon^2 \left(\sum_{k=1}^{N} \mathbb{V}\text{ar}(\xi_{k,T})\right)^2} \sim \frac{1}{\varepsilon^2 N}, \quad \text{as } N, T \to \infty.$$

Thus,

$$\frac{\sum_{k=1}^{N} \xi_{k,T}}{\left(\sum_{k=1}^{N} \mathbb{V}\text{ar}(\xi_{k,T})\right)^{1/2}} \xrightarrow[N,T\to\infty]{d} \mathcal{N}(0,1). \tag{3.14}$$

We also note that

$$\left(\sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma}\int_0^T \mathbb{E}u_k^2(t)\,\mathrm{d}t\right)^{1/2} \simeq \frac{\sigma\sqrt{\varpi^\beta d}\sqrt{T}N^{\frac{\beta}{d}+\frac{1}{2}}}{\sqrt{(4\beta+2d)\theta_0}} \quad \text{as } N, T \to \infty.$$

In view of the strong law of large numbers,

$$\lim_{N,T\to\infty} \frac{\sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t)\,\mathrm{d}t}{\sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \int_0^T \mathbb{E}u_k^2(t)\,\mathrm{d}t} = 1$$

with probability one. Finally, combining all the above and using Slutzky's theorem, (3.5) follows at once. This completes the proof. $\square$

3.1 Asymptotic normality of the MLE by Malliavin-Stein's approach

In this section, we give an alternative proof of (3.5) using tools and results from Malliavin calculus. While this method of proof has its stand along value, we also believe that it will serve as a theoretical base for future studies related to other discretized estimators and different asymptotic regimes.

As before, let $U_0 = 0$, and for convenience, in this section we will use the following notations:

$$F_{N,T} := \widehat{\theta}_{N,T} - \theta_0 = -\frac{\sigma \sum_{k=1}^{N} \lambda_k^{2\beta+\gamma} \int_0^T u_k(t)\,\mathrm{d}w_k(t)}{\sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t)\,\mathrm{d}t} =: -\frac{\boldsymbol{F}_1(N,T)}{\boldsymbol{F}_2(N,T)},$$

$$\widehat{F}_{N,T} := -\frac{\sigma \sum_{k=1}^{N} \lambda_k^{2\beta+\gamma} \int_0^T u_k(t)\,\mathrm{d}w_k(t)}{\mathbb{E}\left[\sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t)\,\mathrm{d}t\right]} = -\frac{\boldsymbol{F}_1(N,T)}{C_{N,T}^2},$$

where $C_{N,T}^2 := \mathbb{E}\left(\boldsymbol{F}_2(N,T)\right)$. We note that $\widehat{F}_N$ can be written as a double stochastic integral and in view of [NP12, Theorem 2.7.7], $\widehat{F}_N$ belongs to the second-order chaos; see also Appendix B. Next we present a key technical result.

Lemma 3.2. *Let $\mathcal{H}$ be the space endowed with the inner product defined in* (B.1). *Let D be the Malliavin derivative defined in* (B.3). *Then, we have*

$$\sqrt{\mathbb{V}\mathrm{ar}\left(\frac{1}{2}\|C_{N,T}D\widehat{F}_{N,T}\|_{\mathcal{H}}^2\right)} \longrightarrow 0, \ \ as \ N,T \to \infty.$$

The proof of Lemma 3.2 is deferred to Appendix A.

To prove asymptotic normality of $\widehat{\theta}_{N,T}$, we will show that $C_{N,T}F_{N,T} \xrightarrow{d} \mathcal{N}(0,\sigma^2)$, as $N,T \to \infty$. The variance σ^2 comes from the fact $\mathbb{E}\left(C_{N,T}^2\widehat{F}_{N,T}^2\right) = \sigma^2$. We split $C_{N,T}F_{N,T}$ into

$$C_{N,T}F_{N,T} = C_{N,T}\left(F_{N,T} - \widehat{F}_{N,T}\right) + C_{N,T}\widehat{F}_{N,T}. \tag{3.15}$$

We note that

$$C_{N,T}(F_{N,T} - \hat{F}_{N,T}) = \frac{C_{N,T}^2}{\boldsymbol{F}_2(N,T)}\frac{\boldsymbol{F}_1(N,T)}{C_{N,T}}\left(1 - \frac{\boldsymbol{F}_2(N,T)}{C_{N,T}^2}\right).$$

From (3.12),

$$\frac{C_{N,T}^2}{\boldsymbol{F}_2(N,T)} \longrightarrow 1, \quad 1 - \frac{\boldsymbol{F}_2(N,T)}{C_{N,T}^2} \longrightarrow 0, \quad as \ N,T \to \infty$$

with probability 1. On the other hand, by Lemma 3.2, we have

$$\frac{\boldsymbol{F}_1(N,T)}{C_{N,T}} = C_{N,T}\widehat{F}_{N,T} \xrightarrow{d} \mathcal{N}(0,\sigma^2), \quad as \ N,T \to \infty.$$

Hence, by Slutzky's theorem, we deduce that $C_{N,T}(F_{N,T} - \widehat{F}_{N,T}) \xrightarrow{d} 0$, as $N,T \to \infty$, which consequently implies that

$$C_{N,T}(F_{N,T} - \widehat{F}_{N,T}) \longrightarrow 0, \quad \text{as } N,T \to \infty, \quad \text{in probability.} \tag{3.16}$$

To deal with the second term in (3.15), we note that by Lemma 3.2 and Proposition B.3, we get that

$$\lim_{N,T \to \infty} d_{TV}\left(C_{N,T}\widehat{F}_{N,T}, \mathcal{N}(0, \sigma^2)\right) = 0.$$

Consequently, Theorem B.2 implies that

$$\text{w-}\lim_{N,T \to \infty} C_{N,T}\widehat{F}_{N,T} = \mathcal{N}\left(0, \sigma^2\right).$$

This, combined with (3.15) and (3.16), implies that

$$\text{w-}\lim_{N,T \to \infty} C_{N,T}F_{N,T} = \mathcal{N}(0, \sigma^2).$$

Finally, note that, by (3.10)

$$\frac{C_{N,T}}{\sigma} \simeq \frac{\sqrt{\varpi^\beta d}\sqrt{T}N^{\frac{\beta}{d}+\frac{1}{2}}}{(4\beta + 2d)\theta_0}, \quad \text{as } N,T \to \infty,$$

which implies (3.5), and the proof is complete.

4 Asymptotic properties of the discretized MLE

In this section, we investigate statistical properties of the discretized version of MLE (2.2). Towards this end, we assume that the Fourier modes $u_k(t)$, $k \geq 1$, are observed at a uniform time grid

$$0 = t_0 < t_1 < \cdots < t_M = T, \quad \text{with } \Delta t := t_i - t_{i-1} = \frac{T}{M}, \ i = 1, \ldots, M.$$

We consider the discretized MLE $\widetilde{\theta}_{N,M,T}$ defined by

$$\widetilde{\theta}_{N,M,T} := -\frac{\sum_{k=1}^{N} \lambda_k^{2\beta+2\gamma} \sum_{i=1}^{M} u_k(t_{i-1})\left[u_k(t_i) - u_k(t_{i-1})\right]}{\sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \sum_{i=1}^{M} u_k^2(t_{i-1})\Delta t}.$$

We are interested in studying the asymptotic properties of $\widetilde{\theta}_{N,M,T}$, as $N, M, T \to \infty$.

For simplicity of writing, we also introduce the following notations:

$$Y_{N,M,T} := \sum_{k=1}^{N} \lambda_k^{2\beta+\gamma} \sum_{i=1}^{M} u_k(t_{i-1})\left(w_k(t_i) - w_k(t_{i-1})\right), \qquad Y_{N,T} := \sum_{k=1}^{N} \lambda_k^{2\beta+\gamma} \int_0^T u_k(t)\, dw_k(t),$$

$$I_{N,M,T} := \sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \sum_{i=1}^{M} u_k^2(t_{i-1})\Delta t, \qquad I_{N,T} := \sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \int_0^T u_k^2(t)\, dt,$$

$$V_{N,M,T} := \sum_{k=1}^{N} \lambda_k^{4\beta+2\gamma} \sum_{i=1}^{M} u_k(t_{i-1}) \int_{t_{i-1}}^{t_i} \left(u_k(t) - u_k(t_{i-1})\right) dt, \qquad \Upsilon := \left(\frac{\varpi^\beta}{(4\beta/d + 2)\theta_0}\right)^{1/2}.$$

A key step in the proofs of the main results is to write $\widetilde{\theta}_{N,M,T}$ as

$$\widetilde{\theta}_{N,M,T} - \theta_0 = \frac{\theta_0 V_{N,M,T}}{I_{N,M,T}} - \frac{\sigma Y_{N,M,T}}{I_{N,M,T}}. \tag{4.1}$$

Below we present several important technical results, and to streamline the presentation their proofs are deferred to Appendix A.

Lemma 4.1. *For $0 < t < s \leq T$ and $k, l \in \mathbb{N}$, we have that*

$$\mathbb{E}(u_k(t)u_k(s)) = \frac{\sigma^2 \lambda_k^{-2\gamma-2\beta}}{2\theta_0} \left[e^{-\theta_0 \lambda_k^{2\beta}(s-t)} - e^{-\theta_0 \lambda_k^{2\beta}(s+t)} \right], \tag{4.2}$$

$$\mathbb{E}|u_k(t) - u_k(s)|^{2l} \leq C(l) \left(\sigma^2 \lambda_k^{-2\gamma} \right)^l |t-s|^l, \tag{4.3}$$

$$\mathbb{E}|u_k(t) + u_k(s)|^{2l} \leq \bar{C}(l) \left(\sigma^2 \lambda_k^{-2\gamma-2\beta} T \right)^l, \tag{4.4}$$

for some $C(l), \bar{C}(l) > 0$.

Lemma 4.2. *For each $T > 0$, $N, M \in \mathbb{N}$, there exist $C > 0$ independent of N, M, T such that*

$$\mathbb{E}|Y_{N,M,T} - Y_{N,T}|^2 \leq C \frac{T^2 N^{\frac{4\beta}{d}+1}}{M}, \tag{4.5}$$

$$\mathbb{E}|I_{N,M,T} - I_{N,T}|^2 \leq C \frac{T^4 N^{\frac{8\beta}{d}+1}}{M^2}, \tag{4.6}$$

$$\mathbb{E}|V_{N,M,T}|^2 \leq C \frac{T^4 N^{\frac{8\beta}{d}+1}}{M^2}. \tag{4.7}$$

Remark 4.3. As a direct consequence of (3.10) and (3.12), we have that

$$\lim_{T,N\to\infty} \frac{Y_{N,T}}{TN^{\frac{2\beta}{d}+1}} = 0, \quad \text{and} \quad \lim_{T,N\to\infty} \frac{I_{N,T}}{\sigma^2 \Upsilon^2 T N^{\frac{2\beta}{d}+1}} = 1, \tag{4.8}$$

with probability one. Moreover, following the lines of the proof of (3.12), one can show that for every fixed $T > 0$,

$$\lim_{N\to\infty} \frac{Y_{N,T}}{TN^{\frac{2\beta}{d}+1}} = 0, \quad \text{and} \quad \lim_{N\to\infty} \frac{I_{N,T}}{\sigma^2 \Upsilon^2 T N^{\frac{2\beta}{d}+1}} = 1, \tag{4.9}$$

with probability one.

With these at hand, we are ready to show that $\widetilde{\theta}_{N,M,T}$ is a weakly consistent estimator of θ_0.

Theorem 4.4. *Assume that $\beta > 1/2$ and $\gamma > d/2$. Then,*

$$\widetilde{\theta}_{N,M,T} \to \theta_0, \quad \text{in probability}, \tag{4.10}$$

as $N, M, T \to \infty$, or as $N, M \to \infty$ while T is fixed, and assuming that (in both cases)

$$\frac{T^2 N^{\frac{4\beta}{d}-1}}{M^2} \to 0. \tag{4.11}$$

Moreover, if $4\beta < d$, then (4.10) holds true with both T, M being fixed.

Proof. Let $\bar{L} := \mathbb{P}\left(\left|\widetilde{\theta}_{N,M,T} - \theta_0\right| > \varepsilon\right)$. In view of (4.1), we note that

$$\bar{L} \leq \mathbb{P}\left(\left|\frac{\sigma Y_{N,M,T}}{I_{N,M,T}}\right| > \varepsilon/2\right) + \mathbb{P}\left(\left|\frac{\theta_0 V_{N,M,T}}{I_{N,M,T}}\right| > \varepsilon/2\right).$$

Consequently, for an arbitrary fixed $\delta \in (0, \varepsilon/2)$, we have, using (A.7),

$$\bar{L} \leq \mathbb{P}\left(\frac{\theta_0|V_{N,M,T}|}{\sigma^2\Upsilon^2 TN^{\frac{2\beta}{d}+1}} > \delta\right) + \mathbb{P}\left(\frac{|Y_{N,M,T}|}{\sigma\Upsilon^2 TN^{\frac{2\beta}{d}+1}} > \delta\right)$$
$$+ 2\mathbb{P}\left(\left|\frac{I_{N,M,T}}{\sigma^2\Upsilon^2 TN^{\frac{2\beta}{d}+1}} - 1\right| > \frac{\varepsilon - 2\delta}{\varepsilon}\right)$$
$$=: L_1 + L_2 + 2L_3.$$

By Chebyshev inequality[4], and (4.7), we have that, for some constant (that may depend on δ) $C_1(\delta) > 0$,

$$L_1 \leq C_1(\delta)\frac{T^2 N^{\frac{4\beta}{d}-1}}{M^2}.$$

As far as L_2, we write

$$L_2 \leq \mathbb{P}\left(\frac{|Y_{N,M,T} - Y_{N,T}|}{\sigma\Upsilon^2 TN^{\frac{2\beta}{d}+1}} > \delta/2\right) + \mathbb{P}\left(\frac{|Y_{N,T}|}{\sigma\Upsilon^2 TN^{\frac{2\beta}{d}+1}} > \delta/2\right) =: L_{21} + L_{22}.$$

Again by Chebyshev inequality, and using (4.5), we get $L_{21} \leq C_2(\delta)/(NM)$, for some $C_2(\delta) > 0$. On the other hand, by (4.8), $L_{22} \to 0$, as $N, T \to \infty$. Moreover, by (4.9) $L_{22} \to 0$, as $N \to \infty$ with T being fixed. We treat L_3 similarly:

$$L_3 \leq \mathbb{P}\left(\frac{|I_{N,M,T} - I_{N,T}|}{\sigma^2\Upsilon^2 TN^{\frac{2\beta}{d}+1}} > \frac{\varepsilon - 2\delta}{2\varepsilon}\right) + \mathbb{P}\left(\left|\frac{I_{N,T}}{\sigma^2\Upsilon^2 TN^{\frac{2\beta}{d}+1}} - 1\right| > \frac{\varepsilon - 2\delta}{2\varepsilon}\right) =: L_{31} + L_{32}.$$

In view of (4.6), and Chebyshev inequality, we have the bound

$$L_{31} \leq C_3(\varepsilon)\frac{T^2 N^{\frac{4\beta}{d}-1}}{M^2}.$$

By (4.8), and respectively (4.9), we get that $L_{32} \to 0$, as $N, T \to \infty$, and, respectively, as $N \to \infty$ while T fixed. Hence, combining all the above bounds, we conclude that

$$\bar{L} \leq C(\varepsilon)\left(\frac{T^2 N^{\frac{4\beta}{d}-1}}{M^2} + \frac{1}{NM}\right).$$

Clearly, $\bar{L} \to 0$ for every $\varepsilon > 0$, as $N, T \to \infty$, or when T fixed and as $N \to \infty$, given that (4.11) is satisfied. This concludes the proof. □

Next we prove an asymptotic normality result for discretized MLE $\widetilde{\theta}_{N,M,T}$. As one may expect, the rate of convergence of $\widetilde{\theta}_{N,M,T}$ agrees with those from continuous time setup, and thus asymptotically is optimal in the mean-square sense. As usual, we will denote by Φ the cumulative probability function of a standard Gaussian random variable.

---

[4]Throughout we will use the following version of the Chebyshev inequality: $\mathbb{P}(|X| > a) \leq \mathbb{E}(X^2)/a^2$, for $a > 0$.

Theorem 4.5. *Assume that $\beta > 1/2$ and $\gamma > d/2$. Then,*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\Upsilon \sqrt{T} N^{\frac{\beta}{d} + \frac{1}{2}} \left(\theta_0 - \widetilde{\theta}_{N,M,T} \right) \leq x \right) - \Phi(x) \right| \to 0, \tag{4.12}$$

as $N, M, T \to \infty$, or as $N, M \to \infty$ while T is fixed, and such that (in both cases)

$$\frac{T^3 N^{6\beta/d}}{M^2} \to 0, \qquad or \qquad \frac{T N^{2\beta/d}}{M} \to 0. \tag{4.13}$$

Proof. We denote the left hand side of (4.12) by $\bar{K}$, and using (4.1), we write it as

$$\bar{K} = \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\sigma \Upsilon \sqrt{T} N^{\frac{\beta}{d} + \frac{1}{2}} \frac{\sigma Y_{N,M,T} - \theta_0 V_{N,M,T}}{I_{N,M,T}} \leq x \right) - \Phi(x) \right|.$$

Using (A.9), we continue

$$\bar{K} \leq \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{\sigma Y_{N,M,T} - \theta_0 V_{N,M,T}}{\sigma \Upsilon \sqrt{T} N^{\frac{\beta}{d} + \frac{1}{2}}} \leq x \right) - \Phi(x) \right| + \mathbb{P} \left(\left| \frac{I_{N,M,T}}{\sigma^2 \Upsilon^2 T N^{\frac{2\beta}{d} + 1}} - 1 \right| > \varepsilon \right) + \varepsilon$$

$$=: K_1 + K_2 + \varepsilon.$$

Consequently, by applying (A.8), we obtain

$$K_1 \leq \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{Y_{N,T}}{\sigma \Upsilon \sqrt{T} N^{\frac{\beta}{d} + \frac{1}{2}}} \leq x \right) - \Phi(x) \right| + \mathbb{P} \left(\frac{|Y_{N,M,T} - Y_{N,T}|}{\sigma \Upsilon \sqrt{T} N^{\frac{\beta}{d} + \frac{1}{2}}} > \varepsilon \right)$$

$$+ \mathbb{P} \left(\frac{\theta_0 |V_{N,M,T}|}{\sigma \Upsilon \sqrt{T} N^{\frac{\beta}{d} + \frac{1}{2}}} > \varepsilon \right) + 2\varepsilon =: K_{1,1} + K_{1,2} + K_{1,3} + 2\varepsilon.$$

Note that (3.14) implies that

$$\text{w-} \lim_{N,T \to \infty} \frac{Y_{N,T}}{\sigma \Upsilon \sqrt{T} N^{\frac{\beta}{d} + \frac{1}{2}}} = \mathcal{N}(0,1).$$

Moreover, using (3.1), one can easily derive that the above limit holds true for every fixed $T > 0$ and $N \to \infty$. Thus, $K_{1,1} \to 0$, as $N, T \to \infty$ (or as $N \to \infty$, when T is fixed). By Chebyshev inequality and by (4.5) and (4.7), we deduce

$$K_{1,2} \leq C_1(\varepsilon) \frac{T N^{\frac{2\beta}{d}}}{M}, \qquad K_{1,3} \leq C_2(\varepsilon) \frac{T^3 N^{\frac{6\beta}{d}}}{M^2},$$

for some $C_1(\varepsilon), C_2(\varepsilon) > 0$.

Similarly,

$$K_2 \leq \mathbb{P} \left(\left| \frac{I_{N,T}}{\sigma^2 \Upsilon^2 T N^{\frac{2\beta}{d} + 1}} - 1 \right| > \frac{\varepsilon}{2} \right) + \mathbb{P} \left(\frac{|I_{N,M,T} - I_{N,T}|}{\sigma^2 \Upsilon^2 T N^{\frac{2\beta}{d} + 1}} > \frac{\varepsilon}{2} \right) =: K_{2,1} + K_{2,2}.$$

By (4.8) or (4.9), $K_{2,1} \to 0$, as $N, T \to \infty$, or as $N \to \infty$ and T fixed. On the other hand, by (4.6),

$$K_{2,1} \leq C_3(\varepsilon) \frac{T^2 N^{\frac{4\beta}{d} - 1}}{M^2}.$$

Combining all the above, we conclude

$$\bar{K} \leq C_4(\varepsilon) \left(\frac{T N^{\frac{2\beta}{d}}}{M} + \frac{T^3 N^{\frac{6\beta}{d}}}{M^2} + \frac{T^2 N^{\frac{4\beta}{d} - 1}}{M^2} \right) \leq C_5(\varepsilon) \left(\frac{T N^{\frac{2\beta}{d}}}{M} + \frac{T^3 N^{\frac{6\beta}{d}}}{M^2} \right) + 3\varepsilon.$$

Since $\varepsilon > 0$ was chosen arbitrarily, and since $C_5(\varepsilon)$ is independent of N, M, T, and given that (4.13) is satisfied, we conclude that $\bar{K} \to 0$, as $N, M, T \to \infty$, or $N, M \to \infty$ and T being fixed. The proof is complete. $\qquad \square$

A   Proofs of technical lemmas

Proof of Lemma 3.2. We start by computing the Malliavin derivative of $\widehat{F}_{N,T}$. If $r \le t$ and for $1 \le k \le N$, then

$$
D_{r,k} u_k(t) = \sigma \lambda_k^{-\gamma} D_{r,k} \int_0^t e^{-\theta_0 \lambda_k^{2\beta}(t-s)} \, dw_k(s)
$$
$$
= \sigma \lambda_k^{-\gamma} e^{-\theta_0 \lambda_k^{2\beta}(t-r)}.
$$

Moreover, one has that $D_{r,j} u_k(t) = 0$ if $j \ne k$ or $r > t$. Therefore, for $r \le T$ and $1 \le j \le N$, we have by (B.4),

$$
D_{r,j} \widehat{F}_{N,T} = -\frac{\sigma}{C_{N,T}^2} \lambda_j^{2\beta+\gamma} u_j(r) - \frac{\sigma}{C_{N,T}^2} \sum_{k=1}^N \lambda_k^{2\beta+\gamma} \int_r^T D_{r,j} u_k(t) \, dw_k(t)
$$
$$
= -\frac{\sigma}{C_{N,T}^2} \lambda_j^{2\beta+\gamma} u_j(r) - \frac{\sigma}{C_{N,T}^2} \lambda_j^{2\beta+\gamma} \int_r^T D_{r,j} u_j(t) \, dw_j(t)
$$
$$
= -\frac{\sigma}{C_{N,T}^2} \lambda_j^{2\beta+\gamma} u_j(r) - \frac{\sigma^2}{C_{N,T}^2} \lambda_j^{2\beta} \int_r^T e^{-\theta_0 \lambda_j^{2\beta}(t-r)} \, dw_j(t). \tag{A.1}
$$

We continue by setting

$$
A := \left\| C_{N,T} D\widehat{F}_{N,T} \right\|_{\mathcal{H}}^2 = C_{N,T}^2 \| D\widehat{F}_{N,T} \|_{\mathcal{H}}^2,
$$

and in view of (A.1), we obtain

$$
A = C_{N,T}^2 \int_0^T \sum_{k=1}^N \left[\frac{\sigma}{C_{N,T}^2} \lambda_k^{2\beta+\gamma} u_k(r) + \frac{\sigma^2}{C_{N,T}^2} \lambda_k^{2\beta} \int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} dw_k(t) \right]^2 dr
$$
$$
= \int_0^T \sum_{k=1}^N \left[\frac{\sigma^2}{C_{N,T}^2} \lambda_k^{4\beta+2\gamma} u_k^2(r) + 2\frac{\sigma^3}{C_{N,T}^2} \lambda_k^{4\beta+\gamma} u_k(r) \int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \, dw_k(t) \right.
$$
$$
\left. + \frac{\sigma^4}{C_{N,T}^2} \lambda_k^{4\beta} \left(\int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \, dw_k(t) \right)^2 \right] dr
$$
$$
=: A_1 + A_2 + A_3.
$$

It is easy to see that for any process $\Phi = \{\Phi(s), s \in [0,t]\}$ such that $\sqrt{\mathbb{V}\mathrm{ar}(\Phi(s))}$ is integrable on $[0,t]$, it holds that $\sqrt{\mathbb{V}\mathrm{ar}\left(\int_0^t \Phi_s \, ds\right)} \le \int_0^t \sqrt{\mathbb{V}\mathrm{ar}(\Phi_s)} \, ds$. Therefore, we have

$$
\sqrt{\mathbb{V}\mathrm{ar}\left(\frac{1}{2} \| C_{N,T} D\widehat{F}_{N,T} \|_{\mathcal{H}}^2\right)} \le \frac{\sqrt{3}}{2} \left(\sqrt{\mathbb{V}\mathrm{ar}(A_1)} + \sqrt{\mathbb{V}\mathrm{ar}(A_2)} + \sqrt{\mathbb{V}\mathrm{ar}(A_3)} \right)
$$
$$
\le \frac{\sqrt{3}}{2} (B_1 + B_2 + B_3),
$$

where

$$B_1 := \frac{\sigma^2}{C_{N,T}^2} \int_0^T \left[\mathbb{V}\mathrm{ar} \left(\sum_{k=1}^N \lambda_k^{4\beta+2\gamma} u_k^2(r) \right) \right]^{1/2} \mathrm{d}r$$

$$B_2 := 2\frac{\sigma^3}{C_{N,T}^2} \int_0^T \left[\mathbb{V}\mathrm{ar} \left(\sum_{k=1}^N \lambda_k^{4\beta+\gamma} u_k(r) \int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} dw_k(t) \right) \right]^{1/2} \mathrm{d}r$$

$$B_3 := \frac{\sigma^4}{C_{N,T}^2} \int_0^T \left[\mathbb{V}\mathrm{ar} \left(\sum_{k=1}^N \lambda_k^{4\beta} \left(\int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} dw_k(t) \right)^2 \right) \right]^{1/2} \mathrm{d}r.$$

Note that by (3.7) and (3.8),

$$\mathbb{V}\mathrm{ar}\left(u_k^2(r) \right) = \mathbb{E}\left(u_k^4(r) \right) - \mathbb{E}^2\left(u_k^2(r) \right) = \frac{\sigma^4 \lambda_k^{-4\beta-4\gamma}}{2\theta_0^2} \left(1 - e^{-2\theta_0 \lambda_k^{2\beta} r} \right)^2. \tag{A.2}$$

For B_1, by the independence of $\{u_k\}_{k\geq 1}$ and (A.2),

$$B_1 = \frac{\sigma^2}{C_{N,T}^2} \int_0^T \left(\sum_{k=1}^N \lambda_k^{8\beta+4\gamma} \mathbb{V}\mathrm{ar}\left(u_k^2(r) \right) \right)^{1/2} \mathrm{d}r$$

$$= \frac{\sigma^4}{\sqrt{2}\theta_0 C_{N,T}^2} \int_0^T \left(\sum_{k=1}^N \lambda_k^{4\beta} \left(1 - e^{-2\theta_0 \lambda_k^{2\beta} r} \right)^2 \right)^{1/2} \mathrm{d}r$$

$$\simeq \frac{\sigma^4 T}{\sqrt{2}\theta_0 C_{N,T}^2} \left(\sum_{k=1}^N \lambda_k^{4\beta} \right)^{1/2} \quad \text{as } T \to \infty$$

$$\sim \frac{1}{N^{1/2}} \to 0, \quad \text{as } N, T \to \infty. \tag{A.3}$$

For B_2, we note that u_k and W_l are independent if $k \neq l$. Therefore, we rewrite B_2 as

$$B_2 = \frac{2\sigma^3}{C_{N,T}^2} \int_0^T \left[\sum_{k=1}^n \lambda_k^{8\beta+2\gamma} \mathbb{V}\mathrm{ar}\left(u_k(r) \int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \mathrm{d}W_k(t) \right) \right]^{1/2} \mathrm{d}r.$$

By straightforward calculations, we have that

$$\mathbb{V}\mathrm{ar}\left(u_k(r) \int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \mathrm{d}w_k(t) \right) \leq \mathbb{E}\left[u_k^2(r) \left[\int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \mathrm{d}w_k(t) \right]^2 \right]$$

$$= \mathbb{E}\left[u_k^2(r) \mathbb{E}\left[\left(\int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \mathrm{d}w_k(t) \right)^2 \Big| \mathcal{F}_r \right] \right]$$

$$= \frac{\sigma^2 \lambda_k^{-2\beta-2\gamma}}{2\theta_0} \left(1 - e^{-2\theta_0 \lambda_k^{2\beta} r} \right) \int_r^T e^{-2\theta_0 \lambda_k^{2\beta}(t-r)} \mathrm{d}t$$

$$\leq \frac{\sigma^2 \lambda_k^{-4\beta-2\gamma}}{4\theta_0^2}.$$

Therefore, we get

$$B_2 \leq \frac{\sigma^4}{\theta_0 C_{N,T}^2} \int_0^T \left(\sum_{k=1}^N \lambda_k^{4\beta} \right)^{1/2} \mathrm{d}r \sim \frac{1}{N^{1/2}} \to 0, \quad \text{as } N, T \to \infty. \tag{A.4}$$

Let us now consider B_3. Since w_k and w_j are independents for $k \neq j$, we have

$$
\begin{aligned}
B_3 &= \frac{\sigma^4}{C_{N,T}^2} \int_0^T \left[\mathbb{V}\mathrm{ar} \left(\sum_{k=1}^N \lambda_k^{4\beta} \left(\int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \, \mathrm{d}w_k(t) \right)^2 \right) \right]^{1/2} \mathrm{d}r \\
&= \frac{\sigma^4}{C_{N,T}^2} \int_0^T \left[\sum_{k=1}^N \lambda_k^{8\beta} \mathbb{V}\mathrm{ar} \left(\int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \, \mathrm{d}w_k(t) \right)^2 \right]^{1/2} \mathrm{d}r \\
&\leq \frac{\sigma^4}{C_{N,T}^2} \int_0^T \left[\sum_{k=1}^N \lambda_k^{8\beta} \mathbb{E} \left(\int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \, \mathrm{d}w_k(t) \right)^4 \right]^{1/2} \mathrm{d}r \\
&= \frac{\sigma^4}{C_{N,T}^2} \int_0^T \left[3 \sum_{k=1}^N \lambda_k^{8\beta} \left[\mathbb{E} \left(\int_r^T e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \, \mathrm{d}w_k(t) \right)^2 \right]^2 \right]^{1/2} \mathrm{d}r \\
&\leq \frac{\sqrt{3}\sigma^4}{2\theta_0 C_{N,T}^2} T \left(\sum_{k=1}^N \lambda_k^{4\beta} \right)^{1/2} \sim \frac{1}{N^{1/2}} \to 0, \text{ as } N, T \to \infty. \quad \text{(A.5)}
\end{aligned}
$$

Finally, combining (A.3), (A.4) and (A.5), we have that for every $\varepsilon > 0$, there exist two independent constants $N_0, T_0 > 0$ such that for all $N \geq N_0$ and $T \geq T_0$,

$$
B_1 + B_2 + B_3 < \varepsilon.
$$

This completes the proof. □

Proof of Lemma 4.1. Using (3.6), (4.2) follows by direct evaluations. As far as (4.3) and (4.4), since $u_k(t) - u_k(s)$ is a Gaussian random variable, it is enough to prove (4.3) and (4.4) for $l = 1$. We note that for $t < s$,

$$
\begin{aligned}
\mathbb{E}|u_k(t) - u_k(s)|^2 &= \sigma^2 \lambda_k^{-2\gamma} \mathbb{E} \left[\int_0^t e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \, \mathrm{d}w_k(r) - \int_0^s e^{-\theta_0 \lambda_k^{2\beta}(s-r)} \, \mathrm{d}w_k(r) \right]^2 \\
&= \sigma^2 \lambda_k^{-2\gamma} \left[\int_0^t e^{-2\theta_0 \lambda_k^{2\beta}(t-r)} \, \mathrm{d}r - 2 \int_0^t e^{-\theta_0 \lambda_k^{2\beta}(t+s-2r)} \, \mathrm{d}r + \int_0^s e^{-2\theta_0 \lambda_k^{2\beta}(s-r)} \, \mathrm{d}r \right] \\
&= \frac{\sigma^2 \lambda_k^{-2\gamma-2\beta}}{2\theta_0} \left[2(1 - e^{-\theta_0 \lambda_k^{2\beta}(s-t)}) + (e^{-\theta_0 \lambda_k^{2\beta}t} + e^{-\theta_0 \lambda_k^{2\beta}s})(e^{-\theta_0 \lambda_k^{2\beta}s} - e^{-\theta_0 \lambda_k^{2\beta}t}) \right] \\
&\leq C\sigma^2 \lambda_k^{-2\gamma}|t-s|,
\end{aligned}
$$

for some $C > 0$, and where in the last inequality we used the fact that e^{-x} is Lipschitz continuous on $[0, \infty)$. Thus part (4.3) is proved. Finally, we have the estimates

$$
\begin{aligned}
\mathbb{E}|u_k(t) + u_k(s)|^2 &= \sigma^2 \lambda_k^{-2\gamma} \mathbb{E} \left[\int_0^t e^{-\theta_0 \lambda_k^{2\beta}(t-r)} \, \mathrm{d}w_k(r) + \int_0^s e^{-\theta_0 \lambda_k^{2\beta}(s-r)} \, \mathrm{d}w_k(r) \right]^2 \\
&= \sigma^2 \lambda_k^{-2\gamma} \left[\int_0^t e^{-2\theta_0 \lambda_k^{2\beta}(t-r)} \, \mathrm{d}r + 2 \int_0^t e^{-\theta_0 \lambda_k^{2\beta}(t+s-2r)} \, \mathrm{d}r + \int_0^s e^{-2\theta_0 \lambda_k^{2\beta}(s-r)} \, \mathrm{d}r \right] \\
&\leq C\sigma^2 \lambda_k^{-2\gamma-2\beta} T,
\end{aligned}
$$

for some $C > 0$, which implies (4.4). The proof is complete. □

Proof of Lemma 4.2. Since $u_k,\ k \geq 1$, are independent, and taking into account that

$$\int_0^T u_k(t)\, \mathrm{d}w_k(t) = \sum_{i=1}^M \int_{t_{i-1}}^{t_i} u_k(t)\, \mathrm{d}w_k(t),$$

we have that

$$
\begin{aligned}
\mathbb{E}|Y_{N,M,T} - Y_{N,T}|^2 &= \mathbb{E}\left| \sum_{k=1}^N \lambda_k^{2\beta+\gamma} \left[\sum_{i=1}^M u_k(t_{i-1})\,(w_k(t_i) - w_k(t_{i-1})) - \int_0^T u_k(t)\, \mathrm{d}w_k(t) \right] \right|^2 \\
&= \sum_{k=1}^N \lambda_k^{4\beta+2\gamma} \mathbb{E}\left| \sum_{i=1}^M u_k(t_{i-1})\,(w_k(t_i) - w_k(t_{i-1})) - \int_0^T u_k(t)\, \mathrm{d}w_k(t) \right|^2 \\
&= \sum_{k=1}^N \lambda_k^{4\beta+2\gamma} \mathbb{E}\left| \sum_{i=1}^M \int_{t_{i-1}}^{t_i} (u_k(t_{i-1}) - u_k(t))\, \mathrm{d}w_k(t) \right|^2 \\
&= \sum_{k=1}^N \lambda_k^{4\beta+2\gamma} \sum_{i=1}^M \int_{t_{i-1}}^{t_i} \mathbb{E}\left(u_k(t_{i-1}) - u_k(t)\right)^2 \mathrm{d}t
\end{aligned}
$$

and hence, by (4.3), there exist constants $C_1, C_2 > 0$, such that

$$
\begin{aligned}
\mathbb{E}|Y_{N,M,T} - Y_{N,T}|^2 &\leq C_1 \sum_{k=1}^N \lambda_k^{4\beta} \sum_{i=1}^M \int_{t_{i-1}}^{t_i} |t_{i-1} - t|\, \mathrm{d}t \\
&= C_1 \left(\sum_{k=1}^N \lambda_k^{4\beta} \right) \frac{T^2}{2M} \leq C_2 \frac{T^2 N^{\frac{4\beta}{d}+1}}{M}.
\end{aligned}
$$

Hence, (4.5) follows at once.

Next we will prove (4.6). We note that

$$
\begin{aligned}
\mathbb{E}|I_{N,M,T} - I_{N,T}|^2 &= \mathbb{E}\left| \sum_{k=1}^N \lambda_k^{4\beta+2\gamma} \left(\sum_{i=1}^M u_k^2(t_{i-1})(t_i - t_{i-1}) - \int_0^T u_k^2(t)\, \mathrm{d}t \right) \right|^2 \\
&= \sum_{k=1}^N \lambda_k^{8\beta+4\gamma} \mathbb{E}\left| \sum_{i=1}^M \int_{t_{i-1}}^{t_i} \left(u_k^2(t_{i-1}) - u_k^2(t) \right) \mathrm{d}t \right|^2.
\end{aligned}
$$

Consequently, letting $U_i(t) := u_k^2(t_{i-1}) - u_k^2(t),\ k \geq 1$, we continue

$$\mathbb{E}|I_{N,M,T} - I_{N,T}|^2 = \sum_{k=1}^N \lambda_k^{8\beta+4\gamma} \sum_{i=1}^M \mathbb{E}\left| \int_{t_{i-1}}^{t_i} U_i(t)\, \mathrm{d}t \right|^2 + 2 \sum_{k=1}^N \lambda_k^{8\beta+4\gamma} \sum_{i<j}^N \mathbb{E} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} U_i(t)U_j(s)\, \mathrm{d}s\, \mathrm{d}t$$
$$=: I_1 + I_2.$$

Note that by Cauchy-Schwartz inequality,

$$
\begin{aligned}
\mathbb{E}|U_i^2(t)| = \mathbb{E}|u_k^2(t_{i-1}) - u_k^2(t)|^2 &= \mathbb{E}|u_k(t_{i-1}) - u_k(t)|^2 |u_k(t_{i-1}) + u_k(t)|^2 \\
&\leq \left(\mathbb{E}|u_k(t_{i-1}) - u_k(t)|^4 \right)^{1/2} \left(\mathbb{E}|u_k(t_{i-1}) + u_k(t)|^4 \right)^{1/2}.
\end{aligned}
$$

Moreover, by (4.3) and (4.4),

$$\mathbb{E}|U_i^2(t)| \leq c_1 \lambda_k^{-4\gamma-2\beta} T |t - t_{i-1}|, \quad \text{for some } c_1 > 0. \tag{A.6}$$

Again by Cauchy-Schwartz inequality and (A.6), we have that

$$I_1 = \sum_{k=1}^{N} \lambda_k^{8\beta+4\gamma} \sum_{i=1}^{M} \mathbb{E} \left| \int_{t_{i-1}}^{t_i} U_i(t) \, \mathrm{d}t \right|^2 \leq \sum_{k=1}^{N} \lambda_k^{8\beta+4\gamma} \sum_{i=1}^{M} (t_i - t_{i-1}) \int_{t_{i-1}}^{t_i} \mathbb{E}|U_i^2(t)| \, \mathrm{d}t$$

$$\leq c_1 T \sum_{k=1}^{N} \lambda_k^{6\beta} \sum_{i=1}^{M} (t_i - t_{i-1}) \int_{t_{i-1}}^{t_i} (t - t_{i-1}) \, \mathrm{d}t = c_1 \sum_{k=1}^{N} \lambda_k^{6\beta} \frac{T^4}{2M^2}.$$

Turning to I_2, we first notice that

$$\mathbb{E}|U_i(t)U_j(s)| = \mathbb{E}\left[\left(u_k^2(t_{i-1}) - u_k^2(t) \right) \left(u_k^2(t_{j-1}) - u_k^2(s) \right) \right]$$

$$= \mathbb{E}\left[(u_k(t_{i-1}) - u_k(t)) (u_k(t_{i-1}) + u_k(t)) (u_k(t_{j-1}) - u_k(s)) (u_k(t_{j-1}) + u_k(s)) \right]$$

$$= \mathbb{E}\left[(u_k(t_{i-1}) - u_k(t)) (u_k(t_{j-1}) - u_k(s)) u_k(t_{i-1})u_k(t_{j-1}) \right]$$

$$+ \mathbb{E}\left[(u_k(t_{i-1}) - u_k(t)) (u_k(t_{j-1}) - u_k(s)) (u_k(t_{i-1})u_k(s)) \right]$$

$$+ \mathbb{E}\left[(u_k(t_{i-1}) - u_k(t)) (u_k(t_{j-1}) - u_k(s)) u_k(t)u_k(t_{j-1}) \right]$$

$$+ \mathbb{E}\left[(u_k(t_{i-1}) - u_k(t)) (u_k(t_{j-1}) - u_k(s)) u_k(t)u_k(s) \right].$$

By the Wick's Lemma [Bis08, Lemma 3.1], we continue

$$\mathbb{E}|U_i(t)U_j(s)| = \mathbb{E}\left[(u_k(t_{i-1}) - u_k(t)) (u_k(t_{i-1}) + u_k(t)) \right] \mathbb{E}\left[(u_k(t_{j-1}) - u_k(s)) (u_k(t_{j-1}) + u_k(s)) \right]$$

$$+ \mathbb{E}\left[(u_k(t_{i-1}) - u_k(t)) (u_k(t_{j-1}) - u_k(s)) \right] \mathbb{E}\left[(u_k(t_{i-1}) + u_k(t)) (u_k(t_{j-1}) + u_k(s)) \right]$$

$$+ \mathbb{E}\left[(u_k(t_{i-1}) - u_k(t)) (u_k(t_{j-1}) + u_k(s)) \right] \mathbb{E}\left[(u_k(t_{i-1}) + u_k(t)) (u_k(t_{j-1}) - u_k(s)) \right]$$

$$=: J_1 + J_2 + J_3.$$

For J_2, we have

$$\mathbb{E} (u_k(t_{i-1}) - u_k(t)) (u_k(t_{j-1}) - u_k(s)) = \mathbb{E} (u_k(t_{i-1})u_k(t_{j-1})) - \mathbb{E} (u_k(t_{i-1})u_k(s))$$

$$- \mathbb{E} (u_k(t)u_k(t_{j-1})) + \mathbb{E} (u_k(t)u_k(s)).$$

By (4.2), for $i < j$ and $t < s$,

$$\mathbb{E} (u_k(t_{i-1}) - u_k(t)) (u_k(t_{j-1}) - u_k(s)) = \frac{\sigma^2 \lambda_k^{-2\gamma-2\beta}}{2\theta_0} \left[e^{-\theta_0 \lambda_k^{2\beta}(t_{j-1}-t_{i-1})} - e^{-\theta_0 \lambda_k^{2\beta}(t_{j-1}+t_{i-1})} \right.$$

$$\left. - e^{-\theta_0 \lambda_k^{2\beta}(s-t_{i-1})} + e^{-\theta_0 \lambda_k^{2\beta}(t_{i-1}+s)} - e^{-\theta_0 \lambda_k^{2\beta}(t_{j-1}-t)} + e^{-\theta_0 \lambda_k^{2\beta}(t_{j-1}+t)} + e^{-\theta_0 \lambda_k^{2\beta}(s-t)} - e^{-\theta_0 \lambda_k^{2\beta}(s+t)} \right]$$

$$= \frac{\sigma^2 \lambda_k^{-2\gamma-2\beta}}{2\theta_0} \left[\left(e^{-\theta_0 \lambda_k^{2\beta}(t_{j-1}-t_{i-1})} - e^{-\theta_0 \lambda_k^{2\beta}(t_{j-1}-t)} \right) + \left(e^{-\theta_0 \lambda_k^{2\beta}(t_{j-1}+t)} - e^{-\theta_0 \lambda_k^{2\beta}(t_{j-1}+t_{i-1})} \right) \right.$$

$$\left. + \left(e^{-\theta_0 \lambda_k^{2\beta}(s-t)} - e^{-\theta_0 \lambda_k^{2\beta}(s-t_{i-1})} \right) + \left(e^{-\theta_0 \lambda_k^{2\beta}(t_{i-1}+s)} - e^{-\theta_0 \lambda_k^{2\beta}(s+t)} \right) \right]$$

$$\leq c_2 \lambda_k^{-2\gamma}(t - t_{i-1}),$$

for some $c_2 > 0$. By similar arguments, we also obtain

$$\mathbb{E}\left(u_k(t_{i-1}) + u_k(t)\right)\left(u_k(t_{j-1}) + u_k(s)\right) \leq c_3 \lambda_k^{-4\gamma}(s - t_{j-1}),$$

for some $c_3 > 0$. Thus,

$$J_2 \leq c_4 \lambda_k^{-4\gamma}(t - t_{i-1})(s - t_{j-1}),$$

for some $c_4 > 0$. By analogy, one can treat J_1 and J_3, and derive the following upper bounds:

$$J_1 \leq c_5 \lambda_k^{-4\gamma}(t - t_{i-1})(s - t_{j-1}), \qquad J_3 \leq c_6 \lambda_k^{-4\gamma}(t - t_{i-1})(s - t_{j-1}),$$

for some $c_5, c_6 > 0$. Finally, combining the above, we have

$$I_2 \leq c_7 \sum_{k=1}^{N} \lambda_k^{8\beta} \sum_{i<j} \int_{t_{j-1}}^{t_j} \int_{t_{i-1}}^{t_i} (t - t_{i-1})(s - t_{j-1}) \, \mathrm{d}t \, \mathrm{d}s$$

$$\leq c_8 \sum_{k=1}^{N} \lambda_k^{8\beta} \frac{T^4}{M^2}, \quad \text{for some } c_8, c_9 > 0.$$

Thus, using the estimates for I_1, I_2, and the fact that $\lambda_k \sim k^{1/d}$, we conclude that

$$I_1 + I_2 \leq c_9 \sum_{k=1}^{N} \lambda_k^{8\beta} \frac{T^4}{M^2} \leq C_2 \frac{T^4 N^{\frac{8\beta}{d}+1}}{M^2},$$

and hence (4.6) is proved. The estimate (4.7) is proved by similar arguments, and we omit the details here. This completes the proof. $\qquad\square$

A.1   Auxiliary results

For reader's convenience, we present here some simple, or well-known, results from probability. Let X, Y, Z be random variables, and assume that $Z > 0$ a.s.. For any $\varepsilon > 0$ and $\delta \in (0, \varepsilon/2)$, the following inequalities hold true.

$$\mathbb{P}(|Y/Z| > \varepsilon) \leq \mathbb{P}(|Y| > \delta) + \mathbb{P}(|Z - 1| > (\varepsilon - \delta)/\varepsilon), \tag{A.7}$$

$$\sup_{x \in \mathbb{R}} \left|\mathbb{P}(X + Y \leq x) - \Phi(x)\right| \leq \sup_{x \in \mathbb{R}} \left|\mathbb{P}(X + Y \leq x) - \Phi(x)\right| + \mathbb{P}(|Y| > \varepsilon) + \varepsilon, \tag{A.8}$$

$$\sup_{x \in \mathbb{R}} \left|\mathbb{P}(Y/Z \leq x) - \Phi(x)\right| \leq \sup_{x \in \mathbb{R}} \left|\mathbb{P}(Y \leq x) - \Phi(x)\right| + \mathbb{P}(|Z - 1| > \varepsilon) + \varepsilon, \tag{A.9}$$

where Φ denotes the probability function of a standard Gaussian random variable.

B   Elements of Malliavin calculus

In this section, we recall some facts from Malliavin calculus associated with a Gaussian process, that we use in Section 3.1. For more details, we refer to [Nua06]. Toward this end, let $T > 0$ be given. We consider the space $\mathcal{H} = L^2\left([0, T] \times \mathcal{M}\right)$, where $\mathcal{M}$ is the counting measure on $\mathbb{N}$, namely, for $v \in \mathcal{H}$,

$$v(t) = \sum_{k=1}^{\infty} v_k(t).$$

We endow $\mathcal{H}$ with the inner product and the norm

$$\langle u, v \rangle_{\mathcal{H}} := \sum_{k=1}^{\infty} \int_0^T u_k(t) v_k(t) \, \mathrm{d}t, \quad \text{and} \quad \|v\|_{\mathcal{H}} := \sqrt{\langle v, v \rangle_{\mathcal{H}}}, \quad u, v \in \mathcal{H}. \tag{B.1}$$

We fix an isonormal Gaussian process $W = \{W(h)\}_{h \in \mathcal{H}}$ on $\mathcal{H}$, defined on a suitable probability space $(\Omega, \mathscr{F}, \mathbb{P})$, such that $\mathscr{F} = \sigma(W)$ is the σ-algebra generated by W. Denote by $C_p^{\infty}(\mathbb{R}^n)$, the space of all smooth functions on $\mathbb{R}^n$ with at most polynomial growth partial derivatives. Let $\mathcal{S}$ be the space of simple functionals of the form

$$F = f(W(h_1), \ldots, W(h_n)), \quad f \in C_p^{\infty}(\mathbb{R}^n), \ h_i \in \mathcal{H}, \ 1 \leq i \leq n.$$

As usual, we define the Malliavin derivative D on $\mathcal{S}$ by

$$DF = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(W(h_1), \ldots, W(h_n)) h_i, \quad F \in \mathcal{S}. \tag{B.2}$$

We note that the derivative operator D is a closable operator from $L^p(\Omega)$ into $L^p(\Omega; \mathcal{H})$, for any $p \geq 1$. Let $\mathbb{D}^{1,p}$, $p \geq 1$, be the completion of $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,p} = \left(\mathbb{E}\big[|F|^p\big] + \mathbb{E}\big[\|DF\|_{\mathcal{H}}^p\big] \right)^{1/p}.$$

Also, for F of the form

$$F = f\left(W\left(\mathbb{1}_{[0,t_1]}\right), \ldots, W\left(\mathbb{1}_{[0,t_n]}\right)\right), \quad t_1, \ldots, t_n \in [0, T],$$

we define the Malliavin derivative of F at the point t as

$$D_t F = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\left(W\left(\mathbb{1}_{[0,t_1]}\right), \ldots, W\left(\mathbb{1}_{[0,t_n]}\right)\right) \mathbb{1}_{[0,t_i]}(t), \quad t \in [0, T],$$

where $\mathbb{1}_A$ denotes the indicator function of set A. For simplicity, from now on, we define $W(t) := W\left(\mathbb{1}_{[0,t]}\right)$, $t \in [0, T]$, to represent a standard Brownian motion on $[0, T]$. If $\mathscr{F}$ is generated by a collection of independent standard Brownian motions $\{W_k, \ k \geq 1\}$ on $[0, T]$, we define the Malliavin derivative of F at the point t by

$$D_t F := \sum_{k=1}^{\infty} D_{t,k} F := \sum_{k=1}^{\infty} \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(W_k(t_1), \ldots, W_k(t_n)) \mathbb{1}_{[0,t_i]}(t), \quad t \in [0, T]. \tag{B.3}$$

Next, we denote by $\boldsymbol{\delta}$, the adjoint of the Malliavin derivative D (as defined in (B.2)) given by the duality formula

$$\mathbb{E}\left(\boldsymbol{\delta}(v) F\right) = \mathbb{E}\left(\langle v, DF \rangle_{\mathcal{H}}\right),$$

for $F \in \mathbb{D}^{1,2}$ and $v \in \mathcal{D}(\boldsymbol{\delta})$, where $\mathcal{D}(\boldsymbol{\delta})$ is the domain of $\boldsymbol{\delta}$. If $v \in L^2(\Omega; \mathcal{H}) \cap \mathcal{D}(\boldsymbol{\delta})$ is a square integrable process, then the adjoint $\boldsymbol{\delta}(v)$ is called the Skorokhod integral of the process v (cf. [Nua06]), and it can be written as

$$\boldsymbol{\delta}(v) = \int_0^T v(t) \, \mathrm{d}W(t).$$

Proposition B.1. *[Nua06, Theorem 1.3.8] Suppose that $v \in L^2(\Omega; \mathcal{H})$ is a square integrable process such that $v(t) \in \mathbb{D}^{1,2}$ for almost all $t \in [0, T]$. Assume that the two parameter process $\{D_t v(s)\}$ is square integrable in $L^2([0, T] \times \Omega; \mathcal{H})$. Then, $\boldsymbol{\delta}(v) \in \mathbb{D}^{1,2}$ and*

$$D_t\left(\boldsymbol{\delta}(v)\right) = v(t) + \int_0^T D_t v(s) \, \mathrm{d}W(s), \quad t \in [0, T]. \tag{B.4}$$

Next, we present a connection between Malliavin calculus and Stein's method. For symmetric functions $f \in L^2\left([0, T]^q\right)$, $q \geq 1$, let us define the following multiple integral of order q

$$\mathbb{I}_q(f) = q! \int_0^T \mathrm{d}W(t_1) \int_0^{t_1} \mathrm{d}W(t_2) \cdots \int_0^{t_{q-1}} \mathrm{d}W(t_q) f(t_1, \ldots, t_q),$$

with $0 < t_1 < t_2 < \cdots < t_q < T$. Note that $\mathbb{I}_q(f)$ is also called the q-th Wiener chaos [NP12, Theorem 2.7.7]. Denote by $d_{TV}(F, G)$, the total variation of two random variables F and G.

Theorem B.2. *[NP12, Corollary 5.2.8] Let $F_N = \mathbb{I}_q(f_N)$, $N \geq 1$, be a sequence of random variables for some fixed integer $q \geq 2$. Assume that $\mathbb{E}\left(F_N^2\right) \to \sigma^2 > 0$, as $N \to \infty$. Then, as $N \to \infty$, the following assertions are equivalent:*

1. $F_N \xrightarrow{d} \mathcal{N} := \mathcal{N}(0, \sigma^2)$;

2. $d_{TV}(F_N, \mathcal{N}) \longrightarrow 0$.

We conclude this section with a result about an upper bound for the total variation of a q-th multiple integral and a Gaussian random variable.

Proposition B.3. *[NP12, Theorem 5.2.6] Let $q \geq 2$ be an integer, and let $F = \mathbb{I}_q(f)$ be a multiple integral of order q such that $\mathbb{E}(F^2) = \sigma^2 > 0$. Then, for $\mathcal{N} = \mathcal{N}(0, \sigma^2)$,*

$$d_{TV}(F, \mathcal{N}) \leq \frac{2}{\sigma^2} \sqrt{\mathbb{Var}\left(\frac{1}{q} \|DF\|_{\mathcal{H}}^2\right)}.$$

References

[Bis08]   J. P. N. Bishwal. *Parameter estimation in stochastic differential equations*, volume 1923 of *Lecture Notes in Mathematics*. Springer, Berlin, 2008.

[BT17]   M. Bibinger and M. Trabs. Volatility estimation for stochastic PDEs using high-frequency observations. *Preprint, arXiv:1710.03519*, 2017.

[BT19]   M. Bibinger and M. Trabs. On central limit theorems for power variations of the solution to the stochastic heat equation. *Preprint*, 2019.

[CCG19] Z. Cheng, I. Cialenco, and R. Gong. Bayesian estimations for diagonalizable bilinear SPDEs. *Forthcoming in Stochastic Processes and their Applications, DOI: 10.1016/j.spa.2019.03.020*, 2019.

[CGH11] I. Cialenco and N. Glatt-Holtz. Parameter estimation for the stochastically perturbed Navier-Stokes equations. *Stochastic Process. Appl.*, 121(4):701–724, 2011.

[CGH18] I. Cialenco, R. Gong, and Y. Huang. Trajectory fitting estimators for SPDEs driven by additive noise. *Statistical Inference for Stochastic Processes*, 21(1):1–19, 2018.

[CH17]   I. Cialenco and Y. Huang. A note on parameter estimation for discretely sampled SPDEs. *Preprint arXiv:1710.01649*, 2017.

[Cho07]   P. Chow. *Stochastic partial differential equations*. Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, Boca Raton, FL, 2007.

[Cho19]   C. Chong. High-frequency analysis of parabolic stochastic PDEs. *Forthcoming in Ann. Statist.*, 2019.

[Cia18]   I. Cialenco. Statistical inference for SPDEs: an overview. *Statistical Inference for Stochastic Processes*, 21(2):309–329, 2018.

[CLP09]   I. Cialenco, S. V. Lototsky, and J. Pospíšil. Asymptotic properties of the maximum likelihood estimator for stochastic parabolic equations with additive fractional Brownian motion. *Stoch. Dyn.*, 9(2):169–185, 2009.

[CX15]   I. Cialenco and L. Xu. Hypothesis testing for stochastic PDEs driven by additive noise. *Stochastic Process. Appl.*, 125(3):819–866, March 2015.

[Lot09]   S. V. Lototsky. Statistical inference for stochastic parabolic equations: a spectral approach. *Publ. Mat.*, 53(1):3–45, 2009.

[LR17]   S. V. Lototsky and B. L. Rozovsky. *Stochastic partial differential equations*. Universitext. Springer International Publishing, 2017.

[LR18]   S. V. Lototsky and B. L. Rozovsky. *Stochastic Evolution Systems. Linear theory and applications to non-linear filtering*, volume 89 of *Probability Theory and Stochastic Modelling*. Springer International Publishing, second edition edition, 2018.

[Mar03]   B. Markussen. Likelihood inference for a discretely observed stochastic partial differential equation. *Bernoulli*, 9(5):745–762, 2003.

[NP12]   I. Nourdin and G. Peccati. *Normal approximations with Malliavin calculus, from Stein's method to universality*, volume 192 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2012.

[Nua06]   D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[PR97]   L. I. Piterbarg and B. L. Rozovskii. On asymptotic problems of parameter estimation in stochastic PDE's: discrete time sampling. *Math. Methods Statist.*, 6(2):200–223, 1997.

[PR02]   B. L. S. Prakasa Rao. Nonparametric inference for a class of stochastic partial differential equations based on discrete observations. *Sankhyā Ser. A*, 64(1):1–15, 2002.

[PR03]   B. L. S. Prakasa Rao. Estimation for some stochastic partial differential equations based on discrete observations. II. *Calcutta Statist. Assoc. Bull.*, 54(215-216):129–141, 2003.

[PS19]   G. Pasemann and W. Stannat. Drift estimation for stochastic reaction-diffusion systems. *Preprint: arXiv:1904.04774v1*, 2019.

[PT07]   J. Pospíšil and R. Tribe. Parameter estimates and exact variations for stochastic heat equations driven by space-time white noise. *Stoch. Anal. Appl.*, 25(3):593–611, 2007.

[Shi96]   A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1996.

[Shu01]   M. A. Shubin. *Pseudodifferential operators and spectral theory*. Springer-Verlag, Berlin, second edition, 2001.