

# Parameter estimation for SPDEs based on discrete observations in time and space

Florian Hildebrandt<sup>1</sup> and Mathias Trabs<sup>2</sup>

<sup>1</sup>*Florian Hildebrandt,  
Fachbereich Mathematik,  
Universität Hamburg,  
Bundesstraße 55, 20146 Hamburg, Germany  
e-mail: [florian.hildebrandt@uni-hamburg.de](mailto:florian.hildebrandt@uni-hamburg.de)*

<sup>2</sup>*Mathias Trabs,  
Fachbereich Mathematik,  
Universität Hamburg,  
Bundesstraße 55, 20146 Hamburg, Germany  
e-mail: [mathias.trabs@uni-hamburg.de](mailto:mathias.trabs@uni-hamburg.de)*

**Abstract:** Parameter estimation for a parabolic linear stochastic partial differential equation in one space dimension is studied observing the solution field on a discrete grid in a fixed bounded domain. Considering an infill asymptotic regime in both coordinates, we prove central limit theorems for realized quadratic variations based on temporal and spatial increments as well as on double increments in time and space. Resulting method of moments estimators for the diffusivity and the volatility parameter inherit the asymptotic normality and can be constructed robustly with respect to the sampling frequencies in time and space. Upper and lower bounds reveal that in general the optimal convergence rate for joint estimation of the parameters is slower than the usual parametric rate. The theoretical results are illustrated in a numerical example.

**AMS 2000 subject classifications:** Primary 62F12, 62M09; secondary 60H15, 60F05.

**Keywords and phrases:** Central limit theorems, infill asymptotics, optimal rate of convergence, realized quadratic variation, stochastic partial differential equations.

Received July 2020.

## 1. Introduction

Stochastic partial differential equations (SPDEs) combine the ability of deterministic PDE models to describe complex mechanisms with the key feature of diffusion models, namely a stochastic signal which evolves within the system. While SPDEs have been intensively studied in stochastic analysis, their statistical theory is only at its beginnings. Since we first need to have a thorough statistical understanding for basic SPDEs before more complex models can be studied, let us consider the prototype for the large class of parabolic SPDEs

given by the stochastic heat equation on  $[0, 1]$ :

$$\begin{cases} dX_t(x) = \vartheta_2 \frac{\partial^2}{\partial x^2} X_t(x) dt + \sigma dW_t(x), \\ X_t(0) = X_t(1) = 0, \\ X_0 = \xi, \end{cases} \quad (1)$$

where  $dW$  denotes white noise in space and time,  $\xi$  is some independent initial condition and we impose Dirichlet boundary conditions. More general, we will later incorporate also a first and zero order term in the differential operator. The statistical aim is to infer on the diffusivity parameter  $\vartheta_2 > 0$  and the diffusion or volatility parameter  $\sigma^2 > 0$ .

In the seminal works by Huebner et al. [17] as well as Huebner and Rozovskii [18] a spectral approach has been considered where the processes  $t \mapsto u_\ell(t) := \langle X_t, e_\ell \rangle_{L^2}$  are observable for the eigenfunctions  $e_\ell$  of the underlying differential operator. These so called Fourier modes  $u_\ell$  are independent and satisfy Ornstein-Uhlenbeck dynamics. Consequently, classical results from statistics for stochastic processes can be applied directly. While the spectral approach is studied in numerous papers, see Lototsky [26] or Cialenco [6] for a review, this specific observation scheme is limiting and too restrictive in potential applications. Especially, for more general equations the eigenfunctions will depend on unknown parameters, which is already the case if we add a first order term  $\vartheta_1 \frac{\partial}{\partial x} X_t(x) dt$  with unknown  $\vartheta_1 \in \mathbb{R}$  in (1).

Complementarily to this spectral approach, the canonical problem of parameter estimation based on discrete observations of the solution field of the SPDE recently attracted an increased research activity. Assuming  $X$  is observed on a discrete grid

$$\{(t_i, y_k)\}_{i=0, \dots, N, k=0, \dots, M} \subset [0, T] \times [0, 1],$$

approximate maximum likelihood estimators have been first investigated by Markussen [28] for  $T \rightarrow \infty$ . For various linear SPDEs central limit theorems for method of moment type estimators based on realized quadratic variations have been studied by Torres et al. [36], Cialenco and Huang [7], Cialenco and Kim [8], Bibinger and Trabs [2, 3], Chong [4, 5], Shevchenko et al. [35], Liu and Tudor [25], as well as Kaino and Uchida [22]. However, all these works only give partial answers to the estimation problem. Even for the stochastic heat equation, there is neither a sharp analysis for joint estimation of  $\vartheta_2$  and  $\sigma^2$  nor the case where the number of spatial observations  $M$  dominates the number of temporal observations  $N$  has been explored in general.

Therefore, in this relatively young research field basic and elementary questions even for simple (linear, parabolic) SPDEs still need to be answered. This becomes most important with regard to an increasing number of SPDE models in applications, e.g., in neurobiology [38], for the description of oceans [13, 32], climate modelling [16] or the description of interest rates [9, 34].

In order to provide a complete statistical analysis of parametric estimation for linear parabolic SPDEs in dimension one based on discrete observations on a finite time horizon  $T > 0$ , our main contributions reveal that:

- (i)  $\vartheta_2$  and  $\sigma^2$  cannot be jointly estimated if  $N$  or  $M$  is fixed.
- (ii) The optimal convergence rate for estimating  $(\vartheta_2, \sigma^2)$  jointly is given by  $1/\sqrt{M^3 \wedge N^{3/2}}$  which is generally slower than the parametric rate  $1/\sqrt{MN}$ .
- (iii) Quadratic variations based on space-time increments, see (4) below, satisfy a central limit theorem with  $1/\sqrt{MN}$ -rate, regardless of the relation of  $N$  and  $M$ . Furthermore, they can be used to implement a joint estimator for  $(\sigma^2, \vartheta_2)$  that reaches the optimal rate  $1/\sqrt{M^3 \wedge N^{3/2}}$ .

In view of (i), we will consider the double asymptotic regime  $M, N \rightarrow \infty$  in our analysis which results in infill asymptotics in time and space. Since the vector of observations  $(X_{t_i}(y_k))_{i,k}$  is normally distributed with only two unknown parameters in equation (1), it might surprise that, in general, there is no joint estimator for  $(\sigma^2, \vartheta_2)$  with parametric rate. The lower bound which verifies this statement is at the heart of our analysis. It shows that the parametric rate can only be achieved if  $N$  and  $M^2$  are of the same order of magnitude. In view of the parabolic nature of the stochastic heat equation, this particular asymptotic regime  $N \approx M^2$  implies that we add the same amount of information in time and space as  $N$  and  $M$  increase. In this sense we have a balanced design. An unbalanced regime  $N = o(M^2)$  or  $M = o(\sqrt{N})$  causes a deterioration of the convergence rate. Note the resemblance of the well known consequence of the parabolicity of equation (1), that its solution admits a (nontrivial finite) quadratic variation in space versus a quartic variation in time.

Our statistical analysis also gives insights into the relation between the spectral and the discrete observation scheme. While both are heuristically comparable in view of the discrete Fourier transform, it turns out that there are important differences. In particular, the fully discrete observation scheme is not statistically equivalent (in the sense of Le Cam) to time discrete observations of the first  $M$  Fourier modes in general.

Our estimators rely on realized quadratic variations, taking into account time and space increments

$$(\Delta_i^N X)(y_k) := X_{t_{i+1}}(y_k) - X_{t_i}(y_k), \quad (2)$$

$$(\delta_k^M X)(t_i) := X_{t_i}(y_{k+1}) - X_{t_i}(y_k), \quad (3)$$

respectively, as well as the space-time increments or double increments

$$\begin{aligned} D_{ik} &:= (\delta_k^M \circ \Delta_i^N)X = (\Delta_i^N \circ \delta_k^M)X \\ &= X_{t_{i+1}}(y_{k+1}) - X_{t_{i+1}}(y_k) - X_{t_i}(y_{k+1}) + X_{t_i}(y_k). \end{aligned} \quad (4)$$

In contrast to the maximum likelihood approach which requires inversion of the large  $MN \times MN$  covariance matrix, method of moments type estimators based on (2)–(4) are easy to implement. For one-parameter processes, power variations are a standard tool in the statistical literature. Also, from a probabilistic point of view, there is a certain amount of literature devoted to variations based on double increments for some random field models, see e.g. [31, 33].

For the stochastic heat equation it is observed in [3] that a central limit theorem for realized temporal quadratic variations requires that the observation

frequency in time dominates the observation frequency in space, more precisely,  $M = o(\sqrt{N})$  is necessary. Complementarily, we show that the realized spatial quadratic variation satisfies a central limit theorem if  $N = o(M)$ . The remaining gap can be filled by double increments and the corresponding realized space-time quadratic variation turns out to be robust with respect to the sampling frequencies  $M$  and  $N$ . Based on these statistics, we construct method of moments estimators for  $\vartheta_2$  and  $\sigma^2$  (as well as  $\vartheta_1$  from a first order term). Our rate optimal joint estimator of all identifiable parameters is a least-squares estimator relying on double increments. Our proofs employ directly the Gaussian distribution of  $X$  which allows for an explicit covariance condition for asymptotic normality of quadratic forms of Gaussian triangular schemes. Also, our estimators could be directly generalized to a nonparametric model with time dependent coefficients, as indicated in [3, 5].

Note that the solution process  $X$  to the SPDE (1) admits continuous trajectories only in one spatial dimension. In the multi-dimensional case one could consider noise processes which are more regular in space as studied by Chong [5]. Alternatively, Kriz and Maslowski [24] as well as Altmeyer and Reiß [1] generalize the spectral approach to the observation of functionals  $\langle X_t, K \rangle$  for some (localizing) kernel  $K$ . In one space dimension, a second canonical problem is to analyze equation (1) on the whole real line instead of the unit interval, as considered in, e.g., [2, 7]. In this case, the Laplacian does not have a discrete spectrum and the solution together with its covariance structure changes. Nevertheless, in accordance with the previously named works, we expect a similar behavior of our estimators on both spatial domains  $[0, 1]$  and  $\mathbb{R}$ .

This work is organized as follows: In Section 2 we give a precise definition of the model and study probabilistic properties of the solution field. In Section 3 we present the central limit theorems for realized quadratic variations based on space and double increments. The resulting method of moments estimators are constructed in Section 4. Lower bounds are derived in Section 5. In Section 6 we illustrate our results with a numerical example. The proofs of the main results are collected in Section 7 while auxiliary results are postponed to the appendix.

## 2. Properties of the solution process

For parameters  $\sigma^2 > 0$  and  $\vartheta = (\vartheta_2, \vartheta_1, \vartheta_0) \in \mathbb{R}_+ \times \mathbb{R}^2$ , we consider the linear parabolic SPDE

$$\begin{cases} dX_t(x) = \left( \vartheta_2 \frac{\partial^2}{\partial x^2} X_t(x) + \vartheta_1 \frac{\partial}{\partial x} X_t(x) + \vartheta_0 X_t(x) \right) dt + \sigma dW_t(x), \\ X_t(0) = X_t(1) = 0, \\ X_0 = \xi, \end{cases} \quad (5)$$

$x \in (0, 1), t \geq 0$ , driven by a cylindrical Brownian motion  $W$  and where  $\xi$  is some independent initial condition, satisfying  $\xi \in L^2((0, 1))$  almost surely. More precisely, we study the weak solution  $X = (X_t(x), t \geq 0, x \in [0, 1])$  of the equation  $dX_t = A_\vartheta X_t dt + \sigma dW_t$  associated with the differential operator

$A_\vartheta = \vartheta_2 \frac{\partial^2}{\partial x^2} + \vartheta_1 \frac{\partial}{\partial x} + \vartheta_0$ . As usual, the Dirichlet boundary condition in (5) is implemented in the domain  $\mathcal{D}(A_\vartheta) = H^2((0, 1)) \cap H_0^1((0, 1))$  of  $A_\vartheta$  where  $H^k((0, 1))$  denotes the  $L^2$ -Sobolev spaces of order  $k \in \mathbb{N}$  and with  $H_0^1((0, 1))$  being the closure of  $C_c^\infty((0, 1))$  in  $H^1((0, 1))$ . The cylindrical Brownian motion  $W$  is defined as a linear mapping  $L^2((0, 1)) \ni u \mapsto W \cdot (u)$  such that  $t \mapsto W_t(u)$  is a one-dimensional standard Brownian motion for all normalized  $u \in L^2((0, 1))$  and such that the covariance structure is  $\text{Cov}(W_t(u), W_s(v)) = (s \wedge t) \langle u, v \rangle$ , for  $u, v \in L^2((0, 1))$ ,  $s, t \geq 0$ .  $W$  can thus be understood as the anti-derivative in time of space-time white noise.

The differential operator  $A_\vartheta$  has a complete orthonormal system of eigenvectors. Indeed, the eigenpairs  $(-\lambda_\ell, e_\ell)_{\ell \geq 1}$  associated with  $A_\vartheta$  are given by

$$e_\ell(y) = \sqrt{2} \sin(\pi \ell y) e^{-\kappa y/2}, \quad \lambda_\ell = \vartheta_2(\pi^2 \ell^2 + \Gamma), \quad y \in [0, 1], \ell \in \mathbb{N},$$

where

$$\kappa := \frac{\vartheta_1}{\vartheta_2} \quad \text{and} \quad \Gamma := \frac{\vartheta_1^2}{4\vartheta_2^2} - \frac{\vartheta_0}{\vartheta_2}.$$

The functions  $(e_\ell)_{\ell \geq 1}$  are orthonormal with respect to the weighted  $L^2$ -inner product

$$\langle u, v \rangle := \langle u, v \rangle_\vartheta := \int_0^1 u(x)v(x)e^{\kappa x} dx, \quad u, v \in L^2((0, 1)).$$

Note that in absence of the first derivative in  $A_\vartheta$ , i.e.  $\vartheta_1 = 0$ , the system  $(e_\ell)_{\ell \geq 1}$  reduces to the usual sine-base and  $\langle \cdot, \cdot \rangle$  to the standard inner product on  $L^2([0, 1])$ . In general, both the eigenpairs and the inner product depend on the model parameters. Hence, they are not accessible from a statistical point of view.

Throughout, we restrict the parameter space to

$$\Theta = \left\{ (\sigma^2, \vartheta_2, \vartheta_1, \vartheta_0) \in \mathbb{R}^4 : \sigma^2, \vartheta_2, \frac{\vartheta_1^2}{4\vartheta_2^2} - \frac{\vartheta_0}{\vartheta_2} + \pi^2 > 0 \right\}$$

such that all the eigenvalues are negative and  $A_\vartheta$  is a negative self-adjoint operator. Consequently, the weak solution to the SPDE (5) exists and is given by the variation of constants formula  $X_t = e^{tA_\vartheta} \xi + \sigma \int_0^t e^{(t-s)A_\vartheta} dW_s$ ,  $t \geq 0$ , where  $(e^{tA_\vartheta})_{t \geq 0}$  denotes the strongly continuous semigroup generated by  $A_\vartheta$ , see [10, Theorem 5.4].

Since  $(e_\ell)_{\ell \geq 1}$  is a complete orthonormal system, the cylindrical Brownian motion  $W$  can be realized via  $W_t = \sum_{\ell \geq 1} \beta_\ell(t) e_\ell$  in the sense that  $W_t(\cdot) = \sum_{\ell \geq 1} \beta_\ell(t) \langle \cdot, e_\ell \rangle$  for a sequence of independent standard Brownian motions  $(\beta_\ell)_{\ell \geq 1}$ . In terms of the projections or Fourier modes  $u_\ell(t) := \langle X_t, e_\ell \rangle$ ,  $t \geq 0$ ,  $\ell \in \mathbb{N}$ , we obtain the representation

$$X_t(x) = \sum_{\ell \geq 1} u_\ell(t) e_\ell(x), \quad t \geq 0, x \in [0, 1], \quad (6)$$

where  $(u_\ell)_{\ell \geq 1}$  are one-dimensional independent processes satisfying the Ornstein-Uhlenbeck dynamics  $du_\ell(t) = -\lambda_\ell u_\ell(t) dt + \sigma d\beta_\ell(t)$  or, equivalently,

$$u_\ell(t) = u_\ell(0)e^{-\lambda_\ell t} + \sigma \int_0^t e^{-\lambda_\ell(t-s)} d\beta_\ell(s), \quad u_\ell(0) = \langle \xi, e_\ell \rangle$$

in the sense of the usual finite-dimensional stochastic integral. We will assume throughout that  $\{\beta_\ell, u_\ell(0), \ell \in \mathbb{N}\}$  is an independent family and  $u_\ell(0) \sim \mathcal{N}(0, \sigma^2/(2\lambda_\ell))$  such that each coefficient process  $u_\ell$  is stationary with covariance  $\text{Cov}(u_\ell(s), u_\ell(t)) = \frac{\sigma^2}{2\lambda_\ell} e^{-\lambda_\ell|t-s|}$ ,  $s, t \geq 0$ . Note that this assumption is only imposed for the sake of simplicity. As follows from a perturbation argument as in [3, 15], our estimation methods apply to general sufficiently regular initial conditions.

From representation (6) it is evident that  $X$  is a two-parameter centered Gaussian field. Therefore, the model is completely specified by its covariance structure

$$\text{Cov}(X_s(x), X_t(y)) = \sigma^2 \sum_{\ell \geq 1} \frac{e^{-\lambda_\ell|t-s|}}{2\lambda_\ell} e_\ell(x)e_\ell(y), \quad s, t \geq 0, x, y \in [0, 1]. \quad (7)$$

While  $\sigma^2$  is only a multiplicative factor, the covariance structure depends on  $\vartheta$  through  $\lambda_\ell$  and  $e_\ell$ . By Kolmogorov's criterion there is a continuous version of the process  $(X_t(x), t \geq 0, x \in [0, 1])$ , cf. [10, Chapter 5.5]. In particular, point evaluations  $X_t(x)$  for fixed values of  $t$  and  $x$  are well defined.

For a fixed spatial location  $x$  the sample paths of the process  $X_t(x)$  are not semi-martingales. In fact,  $t \mapsto X_t(x)$  is only Hölder continuous of order almost  $1/4$  [10, Theorem 5.22] and thus has infinite quadratic variation over any time interval. On the other hand, regarding  $X$  as a function of space at a fixed point in time substantially simplifies the probabilistic structure of the process:

**Proposition 2.1.** *Fix  $t \geq 0$  and define  $\Gamma_0 = \sqrt{|\Gamma|}$ .*

(i) *For  $x \leq y$ ,*

$$\text{Cov}(X_t(x), X_t(y)) = \frac{\sigma^2}{2\vartheta_2} e^{-\frac{\kappa}{2}(x+y)} \cdot \begin{cases} \frac{\sin(\Gamma_0(1-y)) \sin(\Gamma_0 x)}{\Gamma_0 \sin(\Gamma_0)}, & \Gamma < 0, \\ x(1-y), & \Gamma = 0, \\ \frac{\sinh(\Gamma_0(1-y)) \sinh(\Gamma_0 x)}{\Gamma_0 \sinh(\Gamma_0)}, & \Gamma > 0. \end{cases}$$

(ii) *The process  $[0, 1] \ni x \mapsto Z(x) := X_t(x)$  is an Itô diffusion. In particular,*

$$dZ(x) = \sqrt{\frac{\sigma^2}{2\vartheta_2}} e^{-\frac{\kappa}{2}x} dB(x) - \begin{cases} \left( \frac{\Gamma_0 \cos(\Gamma_0(1-x))}{\sin(\Gamma_0(1-x))} + \frac{\kappa}{2} \right) Z(x) dx, & \Gamma < 0, \\ \left( \frac{1}{1-x} + \frac{\kappa}{2} \right) Z(x) dx, & \Gamma = 0, \\ \left( \frac{\Gamma_0 \cosh(\Gamma_0(1-x))}{\sinh(\Gamma_0(1-x))} + \frac{\kappa}{2} \right) Z(x) dx, & \Gamma > 0, \end{cases}$$

where  $B(\cdot) = B_t(\cdot)$  is a standard Brownian motion.

Note the similarity between the covariance structures of  $X_t(\cdot)$  and of the Brownian bridge, especially in the case  $\Gamma = 0$ . This resemblance is in line with the Dirichlet boundary conditions  $X_t(0) = X_t(1) = 0$  in our model.

*Remark 2.2.* For  $N \geq 2$  and fixed  $0 \leq t_1 < t_2 < \dots < t_N$  the multi-dimensional process  $x \mapsto (X_{t_1}(x), \dots, X_{t_N}(x))$  is not an Itô diffusion. Indeed, it is not even a Markov process: Take  $N = 2$  and let  $s < t$ . It is a well known fact that for Markov processes past and future are independent, given the present state. For  $x < y < z$  on the other hand, using the Gaussianity of  $X$ , the (Gaussian) conditional distribution of  $(X_s(x), X_t(z))$  given  $(X_s(y), X_t(y))$  can be computed explicitly. From here, independence is easily disproved by checking the non-diagonal entries of the conditional covariance matrix.

We conclude this section by studying absolute continuity properties for different parameter values  $(\sigma^2, \vartheta)$  which in particular has implications for their identifiability. To that aim we introduce the notations

$$\begin{aligned} (X_t(\cdot), t \in [0, T]) &\sim P_{(\sigma^2, \vartheta)} \text{ on } C([0, T], L^2[0, 1]), \\ (X_{t_0}(x), x \in [0, 1]) &\sim P_{(\sigma^2, \vartheta)}^{(t_0, \cdot)} \text{ on } L^2[0, 1], \\ (X_t(x_0), t \in [0, T]) &\sim P_{(\sigma^2, \vartheta)}^{(\cdot, x_0)} \text{ on } L^2[0, T] \end{aligned}$$

for fixed values  $t_0 \geq 0$ ,  $x_0 \in (0, 1)$  and a finite time horizon  $T > 0$ . Further, for probability measures  $Q$  and  $P$  we write  $Q \sim P$  if they are equivalent.

**Proposition 2.3.** *Let  $t_0 \geq 0$ ,  $x_0 \in (0, 1)$  be fixed and consider a finite time horizon  $T > 0$ . For any two sets of parameters  $(\sigma^2, \vartheta)$ ,  $(\tilde{\sigma}^2, \tilde{\vartheta}) \in \Theta$  we have*

$$\begin{aligned} (i) \quad &P_{(\sigma^2, \vartheta)} \sim P_{(\tilde{\sigma}^2, \tilde{\vartheta})} \text{ if and only if } (\sigma^2, \vartheta_2, \vartheta_1) = (\tilde{\sigma}^2, \tilde{\vartheta}_2, \tilde{\vartheta}_1), \\ (ii) \quad &P_{(\sigma^2, \vartheta)}^{(t_0, \cdot)} \sim P_{(\tilde{\sigma}^2, \tilde{\vartheta})}^{(t_0, \cdot)} \text{ if and only if } \begin{pmatrix} \sigma^2 \\ \vartheta_2 \end{pmatrix}, \kappa = \begin{pmatrix} \tilde{\sigma}^2 \\ \tilde{\vartheta}_2 \end{pmatrix}, \tilde{\kappa}, \\ (iii) \quad &P_{(\sigma^2, \vartheta)}^{(\cdot, x_0)} \sim P_{(\tilde{\sigma}^2, \tilde{\vartheta})}^{(\cdot, x_0)} \text{ if and only if } \frac{\sigma^2}{\sqrt{\vartheta_2}} e^{-\kappa x_0} = \frac{\tilde{\sigma}^2}{\sqrt{\tilde{\vartheta}_2}} e^{-\tilde{\kappa} x_0}, \end{aligned}$$

where  $\kappa = \vartheta_1/\vartheta_2$ ,  $\tilde{\kappa} = \tilde{\vartheta}_1/\tilde{\vartheta}_2$ .

Firstly, (i) shows that it is impossible to estimate  $\vartheta_0$  consistently on a finite time horizon. Secondly, (ii) and (iii) reveal that an estimator that only exploits the temporal or spatial covariance structure cannot consistently estimate any other parameters than  $(\sigma^2/\sqrt{\vartheta_2}, \kappa)$  or  $(\sigma^2/\vartheta_2, \kappa)$ , respectively. On the other hand, such estimators can be constructed by using squared time increments at least at two different spatial positions (cf. [3, Theorem 4.2]) or squared space increments (cf. Section 4), respectively.

### 3. Central limit theorems for realized quadratic variations

We will now study central limit theorems for realized quadratic variations based on the space and double increments from (3) and (4), respectively. To fix assumptions and notation, let  $X$  be given by (6) and suppose we have  $(M+1)(N+1)$

time and space discrete observations

$$X_{t_i}(y_k), \quad i = 0, \dots, N, k = 0, \dots, M,$$

on a regular grid  $(t_i, y_k) \subset [0, T] \times [0, 1]$  with a fixed time horizon  $T > 0$  and  $M, N \in \mathbb{N}_0$ . More precisely, assume that

$$y_k = b + k\delta \quad \text{and} \quad t_i = i\Delta \quad \text{where} \quad \delta = \frac{1 - 2b}{M}, \quad \Delta = \frac{T}{N}$$

for some fixed  $b \in [0, 1/2)$ . The spatial locations  $y_k$  are thus equidistant inside a (possibly proper) sub-interval  $[b, 1 - b] \subset [0, 1]$ . For most of our estimators, we will require a strictly positive value for  $b$  in order to exclude undesired boundary effects that lead to biased estimates. Note that whenever  $M \rightarrow \infty$  or/and  $N \rightarrow \infty$ , we obtain infill asymptotics in space  $\delta \rightarrow 0$  or/and time  $\Delta \rightarrow 0$ , respectively.

Throughout,  $M, N \rightarrow \infty$  should be understood in the sense of  $\min(M, N) \rightarrow \infty$ . For two sequences  $(a_n), (b_n)$ , we write  $a_n \lesssim b_n$  to indicate that there exists some  $c > 0$  such that  $|a_n| \leq c \cdot |b_n|$  for all  $n \in \mathbb{N}$  and we write  $a_n \asymp b_n$  if  $a_n \lesssim b_n \lesssim a_n$ . If  $a_n = a$  for some  $a \in \mathbb{R}$  and all  $n \in \mathbb{N}$ , we write  $(a_n) \equiv a$ . The Euclidian norm on  $\mathbb{R}^d$  is denoted by  $\|\cdot\|$ . Moreover,  $\|\cdot\|_2$  denotes the spectral norm and  $\|\cdot\|_F$  denotes the Frobenius norm for matrices.

The realized quadratic variations of  $X$  can be regarded as sums of squares of certain Gaussian random vectors. Hence, our central limit theorems embed into the literature on quadratic forms in random variables and their asymptotic properties, see e.g. [29]. Our key tool for proving asymptotic normality is the following proposition which is tailor made for the situation present in this work and which gives an explicit covariance condition that ensures convergence to the normal distribution.

**Proposition 3.1.** *Let  $(Z_{i,n}, 1 \leq i \leq d_n, n \in \mathbb{N})$  be a triangular array satisfying  $(Z_{1,n} \dots, Z_{d_n,n}) \sim \mathcal{N}(0, \Sigma_n)$  for a covariance matrix  $\Sigma_n \in \mathbb{R}^{d_n \times d_n}, n \in \mathbb{N}$ , and let  $(\alpha_{i,n}, 1 \leq i \leq d_n, n \in \mathbb{N})$  be a deterministic triangular array with values in  $\{-1, 1\}$ . Define  $S_n := \sum_{i=1}^{d_n} \alpha_{i,n} Z_{i,n}^2$  for  $n \geq 1$ . If  $\|\Sigma_n\|_2^2 / \text{Var}(S_n) \rightarrow 0$  as  $n \rightarrow \infty$ , then we have*

$$\frac{S_n - \mathbf{E}(S_n)}{\sqrt{\text{Var } S_n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1) \quad \text{for } n \rightarrow \infty.$$

The proof relies on the fact that  $S_n$  can be represented as a linear combination of independent  $\chi^2(1)$ -distributed random variables.  $\|\Sigma_n\|_2^2 / \text{Var}(S_n) \rightarrow 0$  then implies that the corresponding Lyapunov condition is fulfilled. In this section we only require  $\alpha_{i,n} = 1$  for all  $i$  and  $n$ , i.e.  $S_n = \|Z_{\bullet,n}\|^2$ . The general case will be necessary to verify asymptotic normality of the least-squares estimator in Section 4. It is worth noting that Proposition 3.1 reveals a quite elementary proof strategy to verify several central limit theorems in [3, 7, 35, 36] instead of advanced techniques from Malliavin calculus or mixing theory.



*Remark 3.2.*

1. If  $\alpha_{i,n} = 1$  for all  $i, n$ , it follows from Isserlis' theorem [21] that  $\text{Var}(S_n) = 2\|\Sigma_n\|_F^2$  and thus, the condition for asymptotic normality may be written as  $\|\Sigma_n\|_2/\|\Sigma_n\|_F \rightarrow 0$ . This condition is essentially optimal: In case of independent observations it is in fact equivalent to asymptotic negligibility of the individual normalized and centered summands and hence equivalent to Lindeberg's condition.
2. The spectral norm is bounded by the maximum absolute row sum. Writing  $\Sigma_n = (\sigma_{ij}^{(n)})_{i,j}$ , asymptotic normality thus holds under the sufficient condition

$$\frac{\left(\max_{i \leq d_n} \sum_{j=1}^{d_n} |\sigma_{ij}^{(n)}|\right)^2}{\text{Var } S_n} \rightarrow 0, \quad n \rightarrow \infty. \quad (8)$$

So far, the double asymptotic regime  $M, N \rightarrow \infty$  has only been studied for time increments  $(\Delta_i^N X)(y_k) = X_{t_{i+1}}(y_k) - X_{t_i}(y_k)$ : If  $b > 0$  and if there exists  $\rho \in (0, 1/2)$  such that  $M = \mathcal{O}(N^\rho)$ , then the *rescaled realized temporal quadratic variation*

$$V_t := \frac{1}{MN\sqrt{\Delta}} \sum_{i=0}^{N-1} \sum_{k=0}^{M-1} e^{\kappa y_k} (\Delta_i^N X)^2(y_k) \quad (9)$$

satisfies

$$\sqrt{MN} \left( V_t - \frac{\sigma^2}{\sqrt{\pi\vartheta_2}} \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left( 0, \frac{B\sigma^4}{\pi\vartheta_2} \right), \quad N, M \rightarrow \infty, \quad (10)$$

where

$$B = 2 + \sum_{J=1}^{\infty} \left( 2\sqrt{J} - \sqrt{J+1} - \sqrt{J-1} \right)^2, \quad (11)$$

cf. [3, Thm. 3.4]. Note that this result is only valid under the condition (roughly)  $M = o(\sqrt{N})$ , i.e., if the observation frequency in time is much higher than in space. This constraint is due to a non-negligible correlation of realized temporal quadratic variations at two neighboring points in space if the distance  $\delta$  of these points is small compared to  $\Delta$  or, equivalently, if  $M$  is large compared to  $N$ .

In the situation where the number of spatial observations dominates the number of temporal observations the above result is not applicable. In this case, spatial increments  $(\delta_k^M X)(t_i) = X_{t_i}(y_{k+1}) - X_{t_i}(y_k)$  and the corresponding rescaled realized spatial quadratic variations

$$V_{\text{sp}}(t_i) := \frac{1}{M\delta} \sum_{k=0}^{M-1} e^{\kappa y_k} (\delta_k^M X)^2(t_i)$$

at time  $t_i$  turn out to be useful. In contrast to squared time increments, which have to be renormalized by  $\sqrt{\Delta}$  due to the roughness of  $t \mapsto X_t(y)$ , squared space increments have to be renormalized by  $\delta$  due to the semi-martingale nature of  $y \mapsto X_t(y)$ .

In the extreme case where observations are only available at one point  $t$  in time (and assuming  $\vartheta_1 = \vartheta_0 = 0$  as well as  $X_0 = 0$ ) Cialenco and Huang [7] showed that  $V_{\text{sp}}(t)$  is asymptotically normal with  $1/\sqrt{M}$ -rate of convergence. An analogous result has been proved by Shevchenko et al. [35] for the wave equation. Proposition 2.1 reveals that  $V_{\text{sp}}(t)$  is in fact a rescaled realized quadratic variation of the Itô diffusion  $y \mapsto X_t(y)$ . Hence,

$$\sqrt{M} \left( V_{\text{sp}}(t) - \frac{\sigma^2}{2\vartheta_2} \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left( 0, \frac{\sigma^4}{2\vartheta_2^2} \right), \quad M \rightarrow \infty,$$

follows from standard theory on quadratic variation for semi-martingales. In order to generalize this central limit theorem to the double asymptotic regime  $M, N \rightarrow \infty$ , we define the time average of the *rescaled realized spatial quadratic variations*:

$$V_{\text{sp}} := \frac{1}{N} \sum_{i=0}^{N-1} V_{\text{sp}}(t_i) = \frac{1}{NM\delta} \sum_{i=0}^{N-1} \sum_{k=0}^{M-1} e^{\kappa y_k} (\delta_k^M X)^2(t_i). \quad (12)$$

**Theorem 3.3.** *Let  $b \in [0, 1/2)$ . If  $N/M \rightarrow 0$  then*

$$\sqrt{MN} \left( V_{\text{sp}} - \frac{\sigma^2}{2\vartheta_2} \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left( 0, \frac{\sigma^4}{2\vartheta_2^2} \right), \quad M, N \rightarrow \infty.$$

*Remark 3.4.*

1. The condition  $N/M \rightarrow 0$  is necessary in order to neglect the bias: The proof of the theorem reveals that  $\delta^{-1} \mathbf{E} \left( e^{-\kappa y_k} (\delta_k^M X)^2(t_i) \right) - \frac{\sigma^2}{2\vartheta_2} \approx \delta$  and consequently, the overall bias is of the order

$$\mathbf{E} \left( \sqrt{MN} \left( V_{\text{sp}} - \frac{\sigma^2}{2\vartheta_2} \right) \right) \approx \sqrt{MN} \cdot \delta \approx \sqrt{\frac{N}{M}}.$$

2. A similar central limit theorem (with asymptotic mean and variance dependent on  $\kappa$ ) can also be expected to hold when neglecting the exponential factors  $e^{\kappa y_k}$  in the definition of  $V_{\text{sp}}$ . However, considering this rescaled version of the quadratic variation is useful with a view to parameter estimation (especially of  $\kappa$ ), as shall become clear in the following section.

We conclude that the central limit theorem for realized temporal quadratic variations  $V_t$  holds when (roughly)  $M = o(\sqrt{N})$ , whereas the central limit theorem for realized spatial quadratic variations  $V_{\text{sp}}$  is fulfilled if  $N = o(M)$ . To close the remaining gap, we finally study the space-time increments  $D_{ik}$  from (4). The corresponding rescaled realized quadratic variations are robust with respect to the sampling regime, as indicated by the representation

$$D_{ik} = \sum_{\ell \geq 1} (u_\ell(t_{i+1}) - u_\ell(t_i))(e_\ell(y_{k+1}) - e_\ell(y_k))$$

in terms of the series expansion (6).

In contrast to the case of space increments (and in line with the result for time increments), we impose  $b > 0$  for the remainder of this section. Inspection of the proofs suggests that this condition may be relaxed to  $b \rightarrow 0$  as long as the decay is sufficiently slow. As a first step, we calculate the expectation of the double increments

**Proposition 3.5.** *Let  $b \in (0, 1/2)$ . Then:*

(i) *It holds uniformly in  $0 \leq k \leq M - 1$  and  $1 \leq i \leq N - 1$  that*

$$\mathbf{E} (D_{ik}^2) = \sigma^2 e^{-\kappa y_k} \Phi_{\vartheta}(\delta, \Delta) + \mathcal{O} \left( \delta \sqrt{\Delta} \left( \delta \wedge \sqrt{\Delta} \right) \right), \quad \max(\delta, \Delta) \rightarrow 0,$$

where

$$\begin{aligned} \Phi_{\vartheta}(\delta, \Delta) &:= F_{\vartheta_2}(0, \Delta) \left( 1 + e^{-\kappa \delta} \right) - 2F_{\vartheta_2}(\delta, \Delta) e^{-\kappa \delta/2} \quad \text{and} \\ F_{\vartheta_2}(\delta, \Delta) &:= \sum_{\ell \geq 1} \frac{1 - e^{-\pi^2 \vartheta_2 \ell^2 \Delta}}{\pi^2 \vartheta_2 \ell^2} \cos(\pi \ell \delta). \end{aligned}$$

(ii) *Assuming that  $r = \lim \delta/\sqrt{\Delta} \in [0, \infty]$  exists,  $\Phi_{\vartheta}$  admits three different asymptotic regimes:*

$$\Phi_{\vartheta}(\delta, \Delta) = \begin{cases} \frac{1}{\vartheta_2} \cdot \delta + o(\delta), & r = 0, \\ \psi_{\vartheta_2}(r) \cdot \sqrt{\Delta} + o(\sqrt{\Delta}), & r \in (0, \infty), \\ \frac{2}{\sqrt{\vartheta_2 \pi}} \cdot \sqrt{\Delta} + o(\sqrt{\Delta}), & r = \infty, \end{cases}$$

where

$$\psi_{\vartheta_2}(r) := \frac{2}{\sqrt{\pi \vartheta_2}} \left( 1 - e^{-\frac{r^2}{4\vartheta_2}} + \frac{r}{\sqrt{\vartheta_2}} \int_{\frac{r}{2\sqrt{\vartheta_2}}}^{\infty} e^{-z^2} dz \right). \quad (13)$$

If, moreover,  $\delta/\sqrt{\Delta} \equiv r \in (0, \infty)$ , we have

$$\Phi_{\vartheta}(\delta, \Delta) = e^{-\kappa \delta/2} \psi_{\vartheta_2}(r) \cdot \sqrt{\Delta} + \mathcal{O}(\Delta^{3/2}). \quad (14)$$

*Remark 3.6.* The first order constants appearing in the asymptotic expressions in (ii) stem from a first derivative of  $F_{\vartheta_2}(\cdot, \Delta)$  in 0 in case  $r = 0$  and a Riemann sum approximation of  $F_{\vartheta_2}(\delta, \Delta)$  in case  $r \neq 0$ , respectively. Assuming for simplicity that  $\kappa = 0$ , the proof of Proposition 3.5 shows a more precise expression for the remainder terms in case  $r \in \{0, \infty\}$ :

$$\mathbf{E} (D_{ik}^2) = \begin{cases} \frac{1}{\vartheta_2} \cdot \delta + \mathcal{O}(\delta^2/\sqrt{\Delta}), & r = 0, \\ \frac{2}{\sqrt{\pi \vartheta_2}} \cdot \sqrt{\Delta} + \mathcal{O}(\Delta^{3/2}/\delta^2), & r = \infty. \end{cases}$$

Thus, if our analysis of the remainder terms is sharp (which we believe is the case), the first order approximations have a poor quality if  $\delta/\sqrt{\Delta}$  converges slowly.

Proposition 3.5 suggests to renormalize double increments with  $\delta$  if  $\delta/\sqrt{\Delta} \rightarrow 0$  and with  $\sqrt{\Delta}$  otherwise, which is in line with the renormalization of  $V_{\text{sp}}$  and  $V_{\text{sp}}$ , respectively. However, this approach might not be feasible: Firstly, it requires the knowledge which asymptotic regime is present, i.e., whether or not  $\delta/\sqrt{\Delta} \rightarrow 0$ . Especially for one given set of observations this information may be inaccessible. In this case renormalizing with  $\Phi_{\vartheta}(\delta, \Delta)$  automatically captures the correct asymptotic regime. Secondly, if  $r \in \{0, \infty\}$ , the previous remark shows that the asymptotic expressions for  $\Phi_{\vartheta}(\delta, \Delta)$  may lead to an undesirably large bias. In fact, in order to obtain a central limit theorem with  $1/\sqrt{MN}$ -rate of convergence, we would have to impose the assumptions  $N^2/M \rightarrow 0$  and  $M^5/N \rightarrow 0$ , respectively. These constraints are even more restrictive than the ones required for time or space increments.

Therefore, we renormalize with  $\Phi_{\vartheta}(\delta, \Delta)$  and introduce the *rescaled realized quadratic space-time variation*

$$\mathbb{V} := \frac{1}{MN\Phi_{\vartheta}(\delta, \Delta)} \sum_{k=0}^{M-1} \sum_{i=0}^{N-1} e^{\kappa y_k} D_{ik}^2.$$

**Theorem 3.7.** *Let  $b > 0$ . If either  $\delta/\sqrt{\Delta} \rightarrow r \in \{0, \infty\}$  or  $\delta/\sqrt{\Delta} \equiv r \in (0, \infty)$ , then*

$$\sqrt{MN}(\mathbb{V} - \sigma^2) \xrightarrow{\mathcal{D}} \mathcal{N}(0, C(r/\sqrt{\vartheta_2})\sigma^4), \quad N, M \rightarrow \infty,$$

where  $C(\cdot)$  is a bounded continuous function on  $[0, \infty]$ , given by (26), satisfying

$$C(0) = 3 \quad \text{and} \quad C(\infty) = 3 + \frac{3}{2} \sum_{J=1}^{\infty} \left( \sqrt{J-1} - \sqrt{J+1} - 2\sqrt{J} \right)^2.$$

The condition  $\delta/\sqrt{\Delta} \equiv r \in (0, \infty)$  can be relaxed to  $\delta/\sqrt{\Delta} \rightarrow r \in (0, \infty)$  as long as the convergence is fast enough which we omit for the sake of simplicity. If  $\delta/\sqrt{\Delta} \equiv r \in (0, \infty)$  holds, (14) shows that the renormalization  $\Phi_{\vartheta}(\delta, \Delta)$  and its first order approximation are close enough to be exchanged in the previous theorem. In this case we obtain a central limit theorem with a simpler renormalization which particularly does not depend on the model parameters:

**Corollary 3.8.** *If  $b > 0$  and  $\delta/\sqrt{\Delta} \equiv r \in (0, \infty)$ , then*

$$\mathbb{V}_r := \frac{1}{MN\sqrt{\Delta}} \sum_{k=0}^{M-1} \sum_{i=0}^{N-1} \exp\left(\frac{\kappa}{2}(y_k + y_{k+1})\right) D_{ik}^2 \quad (15)$$

satisfies with  $\psi_{\vartheta_2}(r)$  from (13) and  $C(\cdot)$  from (26):

$$\sqrt{MN}(\mathbb{V}_r - \psi_{\vartheta_2}(r)\sigma^2) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, C(r/\sqrt{\vartheta_2})\psi_{\vartheta_2}^2(r)\sigma^4\right), \quad N, M \rightarrow \infty.$$

Although the above corollary is only valid in the very particular regime  $\delta \approx \sqrt{\Delta}$ , it will also prove to be enormously useful for general sampling regimes when applied to appropriate subsets of the data. Indeed, its power stems from the fact that its asymptotic mean value  $\sigma^2\psi_{\vartheta_2}(r)$  enables separating the parameters  $\sigma^2$  and  $\vartheta_2$  by considering multiple values of  $r$ , see Section 4.

*Remark 3.9.*

1. The previous central limit results are satisfied for a possibly growing time horizon  $T_{N,\Delta} := N\Delta$ , too. Theorem 3.3 only requires that  $T_{N,\Delta} > \varepsilon$  for some  $\varepsilon > 0$ . Theorem 3.7 holds if  $T_{N,\Delta} = o(M)$  and, in particular, Corollary 3.8 is applicable if  $N\Delta^{3/2} \rightarrow 0$ .
2. Having obtained the asymptotic distribution of the different types of quadratic variations, future research should also be devoted to deriving corresponding Berry-Esseen type theorems. As in, e.g., [27], this usually requires advanced tools from Malliavin calculus.

To end this section, we compare the realized quadratic variations  $V_t, V_{\text{sp}}$  and  $\mathbb{V}$  and their asymptotic variances. For this purpose, we scale the statistics in such a way that they are asymptotically centered around the same mean, say  $\sigma^2$ :

$$\tilde{V}_t = \sqrt{\pi\vartheta_2}V_t, \quad \tilde{V}_{\text{sp}} = 2\vartheta_2V_{\text{sp}}, \quad \tilde{V} = \mathbb{V}. \quad (16)$$

For simplicity, let  $\kappa = 0$ . Plugging in the asymptotic expressions for  $\Phi_\vartheta(\delta, \Delta)$  from Proposition 3.5 shows that

$$\tilde{V} \approx \frac{1}{2} \sum_{k=0}^{M-1} \sum_{i=0}^{N-1} D_{ik}^2 \cdot \begin{cases} \frac{2\vartheta_2}{NM\delta}, & \delta/\sqrt{\Delta} \rightarrow 0, \\ \frac{\sqrt{\vartheta_2\pi}}{NM\sqrt{\Delta}}, & \delta/\sqrt{\Delta} \rightarrow \infty. \end{cases}$$

Therefore,  $\tilde{V}$  approximately coincides with  $\tilde{V}_{\text{sp}}$  and  $\tilde{V}_t$  for  $r \in \{0, \infty\}$ , respectively, except for the factor 1/2 and using double increments instead of time or space increments, respectively.

Further, denoting the asymptotic variances of  $\tilde{V}_t, \tilde{V}_{\text{sp}}$  and  $\tilde{V}$  by  $\mathfrak{S}_t, \mathfrak{S}_{\text{sp}}$  and  $\mathfrak{S}(r)$ , respectively, we observe the relations  $\mathfrak{S}(\infty) = \frac{3}{2}\mathfrak{S}_t$  and  $\mathfrak{S}(0) = \frac{3}{2}\mathfrak{S}_{\text{sp}}$ , where the factor 3/2 occurs since each double increment consists of two space or time increments, respectively.

#### 4. Parameter estimation

In view of the covariance structure of the observation vector and the fact that the value of  $\vartheta_0$  is irrelevant from a statistical point of view (cf. Proposition 2.3), we consider the parameter vector

$$\eta = (\sigma^2, \vartheta_2, \kappa).$$

It is straightforward to use the results from the previous section to construct method of moments estimators for the volatility parameter  $\sigma^2$  or the diffusivity parameter  $\vartheta_2$ , provided that the other two parameters in  $(\sigma^2, \vartheta_2, \kappa)$  are known, respectively. Doing so, we generalize the spatial increments based estimator from [7] to the double asymptotic regime and we complement the time increments based methods in [3, 5]. Our estimators do not hinge on  $\vartheta_0$  (or  $\Gamma$ ) such that the knowledge of its true value is not required.

Assuming firstly that  $\vartheta_2$  and  $\kappa$  are known, we obtain the following volatility estimators:

$$\hat{\sigma}_{\text{sp}}^2 := \tilde{V}_{\text{sp}}, \quad \hat{\sigma}_{\text{t}}^2 := \tilde{V}_{\text{t}} \quad \text{and} \quad \hat{\sigma}^2 := \mathbb{V}$$

where  $\tilde{V}_{\text{sp}}$  and  $\tilde{V}_{\text{t}}$  have been introduced in (16).

**Proposition 4.1.**

(i) If  $N = o(M)$ , then we have

$$\sqrt{MN} (\hat{\sigma}_{\text{sp}}^2 - \sigma^2) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2\sigma^4), \quad N, M \rightarrow \infty.$$

(ii) If  $M = \mathcal{O}(N^\rho)$  for some  $\rho \in (0, 1/2)$ , then we have with  $B$  defined in (11):

$$\sqrt{MN} (\hat{\sigma}_{\text{t}}^2 - \sigma^2) \xrightarrow{\mathcal{D}} \mathcal{N}(0, B\sigma^4), \quad N, M \rightarrow \infty.$$

(iii) If  $\sqrt{N} = o(M)$ ,  $M = o(\sqrt{N})$  or  $\sqrt{N}/M \approx \delta/\sqrt{\Delta} \equiv r > 0$ , then we have  $C(\cdot)$  from (26):

$$\sqrt{MN} (\hat{\sigma}^2 - \sigma^2) \xrightarrow{\mathcal{D}} \mathcal{N}(0, C(r/\sqrt{\vartheta_2})\sigma^4), \quad N, M \rightarrow \infty,$$

As discussed above, the double increments estimator has a larger variance than the single increments estimators. Hence, if one of the regimes  $N = o(M)$  or  $M = o(\sqrt{N})$  certainly applies, the single increments estimators are preferable. If none of the regimes is present or the situation is unclear, one can profit from the robustness of the double increments estimator with respect to the sampling regime.

If  $N = o(M)$ , the situation is close to that of  $N$  independent semi-martingales (cf. Proposition 2.1) and the asymptotic variance  $2\sigma^4$  of the spatial increments estimator equals the Cramér-Rao lower bound for estimating  $\sigma^2$ , as can be seen by a simple calculation. Consequently,  $\hat{\sigma}_{\text{sp}}^2$  is an asymptotically efficient estimator. The efficiency loss of the other estimators is due to the fact that for increasingly more temporal observations the infinite dimensional nature of the process  $X$  becomes apparent, leading to non-negligible covariances between increments.

If  $\sigma^2$  and  $\kappa$  are known, the diffusivity  $\vartheta_2$  can be estimated by

$$\hat{\vartheta}_{2,\text{sp}} := \frac{\sigma^2}{2V_{\text{sp}}} \quad \text{and} \quad \hat{\vartheta}_{2,\text{t}} := \frac{\sigma^4}{\pi V_{\text{t}}^2}$$

using  $V_{\text{sp}}$  and  $V_{\text{t}}$  from (12) and (9), respectively. Due to the non-trivial dependence of the renormalization  $\Phi_\vartheta(\delta, \Delta)$  on  $\vartheta$ , it is not apparent how to construct a method of moments estimator for  $\vartheta_2$  based on Theorem 3.7 in general. However, if  $\delta/\sqrt{\Delta} \equiv r > 0$ , the renormalization can be decoupled from the unknown parameter as exploited in Corollary 3.8. Since the function  $\vartheta_2 \mapsto \psi_{\vartheta_2}(r)$  has range  $(0, \infty)$  and is monotonic, there is an inverse  $Q_r(\cdot)$  and we can define the method of moments estimator

$$\hat{\vartheta}_{2,r} = Q_r(\mathbb{V}_r/\sigma^2)$$

with  $\mathbb{V}_r$  from (15). As a direct consequence of the delta method, the relation

$$Q'_r(\psi_{\vartheta_2}(r)) = \left( \frac{\partial}{\partial \vartheta_2} \psi_{\vartheta_2}(r) \right)^{-1} = -\vartheta_2^{3/2} \sqrt{\pi} \left( 1 - e^{-\frac{r^2}{4\vartheta_2}} + \frac{2r}{\sqrt{\vartheta_2}} \int_{\frac{r}{2\sqrt{\vartheta_2}}} e^{-z^2} dz \right)^{-1}$$

and the above central limit theorems, we obtain:

**Proposition 4.2.**

(i) If  $N = o(M)$ , then we have

$$\sqrt{MN} \left( \hat{\vartheta}_{2,\text{sp}} - \vartheta_2 \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left( 0, 2\vartheta_2^2 \right), \quad N, M \rightarrow \infty.$$

(ii) If  $M = \mathcal{O}(N^\rho)$  for some  $\rho \in (0, 1/2)$ , then we have with  $B$  from (11):

$$\sqrt{MN} \left( \hat{\vartheta}_{2,t} - \vartheta_2 \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left( 0, 4\vartheta_2^2 B \right), \quad N, M \rightarrow \infty.$$

(iii) If  $\sqrt{N}/M \approx \delta/\sqrt{\Delta} \equiv r > 0$ , then we have with  $C(\cdot)$  from (26):

$$\sqrt{MN} \left( \hat{\vartheta}_{2,r} - \vartheta_2 \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left( 0, C \left( \frac{r}{\sqrt{\vartheta_2}} \right) \left( \psi_{\vartheta_2}(r) / \frac{\partial}{\partial \vartheta_2} \psi_{\vartheta_2}(r) \right)^2 \right)$$

for  $N, M \rightarrow \infty$ .

We now consider parameter estimation when  $(\sigma^2, \vartheta)$  is unknown. Recall from Proposition 2.3 and its subsequent discussion that  $\vartheta_0$  cannot be estimated consistently on a finite time horizon. Moreover, it is not possible to estimate other parameters than  $(\sigma^2/\sqrt{\vartheta_2}, \kappa)$  or  $(\sigma^2/\vartheta_2, \kappa)$  only based on the temporal or the spatial covariance structure, respectively. Estimation of  $(\sigma^2/\sqrt{\vartheta_2}, \kappa)$  via a least squares procedure based on temporal increments is discussed in [3] in the  $M = o(\sqrt{N})$  regime. Analogously, it is possible to estimate  $(\rho^2, \kappa)$ , where  $\rho^2 = \sigma^2/\vartheta_2$ , using spatial increments and Theorem 3.3: Provided that  $N = o(M)$ , classical minimum-contrast estimation theory reveals that

$$(\hat{\rho}^2, \hat{\kappa}) := \arg \min_{(\tilde{\rho}^2, \tilde{\kappa})} \sum_{k=0}^{M-1} \left( \frac{2}{N\delta} \sum_{i=0}^{N-1} (\delta_k^M X)^2(t_i) - \tilde{\rho}^2 e^{-\tilde{\kappa} y_k} \right)^2$$

satisfies a central limit theorem with rate  $1/\sqrt{MN}$ . We omit a detailed analysis of this estimator.

To estimate all three identifiable parameters  $\eta = (\sigma^2, \vartheta_2, \kappa)$ , we employ a least squares approach based on double increments. Due to the highly nontrivial dependence of the normalization  $\Phi_\vartheta(\delta, \Delta)$  on  $\vartheta$ , a direct application of Theorem 3.7 is impossible. Assuming, however, a balanced design in the sense of  $\delta/\sqrt{\Delta} \equiv r \in (0, \infty)$ , we can use Corollary 3.8 where the normalization is decoupled from the unknown parameter  $\vartheta$ .

Let  $\delta/\sqrt{\Delta} \equiv r \in (0, \infty)$  and define  $\bar{D}_{ik} := D_{ik} + D_{(i+1)k}$  as well as  $z_k = (y_k + y_{k+1})/2$ . Corollary 3.8 suggests that

$$\frac{1}{N\sqrt{\Delta}} \sum_{i=0}^{N-1} D_{ik}^2 \approx e^{-\kappa z_k} \sigma^2 \psi_{\vartheta_2}(r) \quad \text{and} \quad \frac{1}{N\sqrt{2\Delta}} \sum_{i=0}^{N-2} \bar{D}_{ik}^2 \approx e^{-\kappa z_k} \sigma^2 \psi_{\vartheta_2}(r/\sqrt{2}).$$

By considering the two different sampling frequency ratios  $r$  and  $r/\sqrt{2}$ , we can distinguish  $\sigma^2$  and  $\vartheta_2$  instead of recovering only the product  $\sigma^2\psi_{\vartheta_2}(r)$ . To estimate  $\eta = (\sigma^2, \vartheta_2, \kappa)$ , we thus introduce the contrast process

$$\begin{aligned}
 K_{M,N}(\tilde{\eta}) &:= K_{M,N}^1(\tilde{\eta}) + K_{M,N}^2(\tilde{\eta}) \quad \text{where} \\
 K_{M,N}^1(\tilde{\eta}) &:= \frac{1}{M} \sum_{k=0}^{M-1} \left( \frac{1}{N\sqrt{\Delta}} \sum_{i=0}^{N-1} D_{ik}^2 - f_{\tilde{\eta}}^1(z_k) \right)^2, \\
 K_{M,N}^2(\tilde{\eta}) &:= \frac{1}{M} \sum_{k=0}^{M-1} \left( \frac{1}{N\sqrt{2\Delta}} \sum_{i=0}^{N-2} \bar{D}_{ik}^2 - f_{\tilde{\eta}}^2(z_k) \right)^2,
 \end{aligned}$$

and  $f_{\tilde{\eta}}^\nu(z) := \sigma^2 e^{-\kappa z} \psi_{\vartheta_2}(r/\sqrt{\nu})$ ,  $\nu = 1, 2$ . The corresponding minimum-contrast estimator is given by

$$\hat{\eta} = \arg \min_{\tilde{\eta} \in H} K_{M,N}(\tilde{\eta}), \tag{17}$$

where  $H$  is some subset of  $(0, \infty)^2 \times \mathbb{R}$  containing the true parameter  $\eta$ .

**Theorem 4.3.** *Assume  $b > 0$  and  $\delta/\sqrt{\Delta} \equiv r > 0$ . If  $\eta = (\sigma^2, \vartheta_2, \kappa)$  lies in the interior of  $H$  for some compact set  $H \subset (0, \infty)^2 \times \mathbb{R}$ , then the least squares estimator  $\hat{\eta}$  from (17) satisfies*

$$\sqrt{MN}(\hat{\eta} - \eta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Omega_\eta^r), \quad M, N \rightarrow \infty,$$

where  $\Omega_\eta^r \in \mathbb{R}^{3 \times 3}$  is a strictly positive definite covariance matrix, explicitly given by (30).

*Remark 4.4.*

1. Based on  $\hat{\eta}$ , we can define  $\hat{\vartheta}_1 := \hat{\eta}_2 \hat{\eta}_3 = \hat{\vartheta}_2 \hat{\kappa}$  to estimate  $\vartheta_1$ . The delta method then yields a central limit theorem for  $(\hat{\sigma}^2, \hat{\vartheta}_2, \hat{\vartheta}_1)$ .
2. If  $\vartheta_1$  is known and the sample size is sufficiently large, the estimator for  $(\sigma^2, \vartheta_2)$  can be computed without solving a minimization problem: For simplicity, assume  $\vartheta_1 = 0$  and let

$$V^1 := \frac{1}{MN\sqrt{\Delta}} \sum_{k=0}^{M-1} \sum_{i=0}^{N-1} D_{ik}^2, \quad V^2 := \frac{1}{MN\sqrt{2\Delta}} \sum_{k=0}^{M-1} \sum_{i=0}^{N-1} \bar{D}_{ik}^2.$$

Further, denote by  $G_r$  the inverse function of  $\vartheta_2 \mapsto \psi_{\vartheta_2}(r)/\psi_{\vartheta_2}(r/\sqrt{2})$ , whose existence is part of the proof of the above theorem. Then, we have

$$\hat{\vartheta}_2 = G_r(V^1/V^2), \quad \hat{\sigma}^2 = V^1/\psi_{\hat{\vartheta}_2}(r),$$

provided that  $V_1/V_2$  lies in the range of  $\vartheta_2 \mapsto \psi_{\vartheta_2}(r)/\psi_{\vartheta_2}(r/\sqrt{2})$ . Due to consistency of  $(V^1, V^2)$ , the latter is true with probability tending to one.

Even when  $\delta/\sqrt{\Delta} \equiv r > 0$  does not hold, there are always subsets of the data having the balanced sampling design. Hence, the estimation procedure treated



in Theorem 4.3 can be generalized to an arbitrary set  $\{X_{t_i}(y_k), i \leq N, k \leq M\}$  of discrete observations by considering an averaged version of the above contrast process. To that aim, choose  $v, w \in \mathbb{N}$  such that  $v \approx \max(1, N/M^2)$  and  $w \approx \max(1, M/\sqrt{N})$ . Then,  $\tilde{\Delta} := v\Delta$  and  $\tilde{\delta} := w\delta$  satisfy

$$r := \tilde{\delta}/\sqrt{\tilde{\Delta}} \approx 1.$$

Using double increments on the coarser grid

$$D_{v,w}(i, k) = X_{t_{i+v}}(y_{k+w}) - X_{t_i}(y_{k+w}) - X_{t_{i+v}}(y_k) + X_{t_i}(y_k),$$

we set

$$\begin{aligned} & \mathcal{K}_{N,M}^\nu(\tilde{\eta}) \\ & := \frac{1}{M-w+1} \sum_{k=0}^{M-w} \left( \frac{1}{N-\nu v+1} \sum_{i=0}^{N-\nu v} \frac{D_{\nu v,w}^2(i, k)}{\sqrt{\nu v \tilde{\Delta}}} - f_{\tilde{\eta}}^\nu\left(\frac{y_k + y_{k+w}}{2}\right) \right)^2 \end{aligned}$$

where  $f_\eta^\nu(z) := 2\sigma^2\psi_{\vartheta_2}(r/\sqrt{\nu})e^{-\kappa z}$  and  $\nu = 1, 2$ . The final estimator for  $\eta$  is then defined as

$$\hat{\eta}_{v,w} := \arg \min_{\tilde{\eta} \in H} (\mathcal{K}_{N,M}^1(\tilde{\eta}) + \mathcal{K}_{N,M}^2(\tilde{\eta})). \quad (18)$$

The rate of convergence of this estimation procedure is inherited from the observations on the coarser grids  $\{(t_{i+jv}, y_{k+lw}) : 0 \leq j \leq N/v - 1, 0 \leq l \leq M/w - 1\}$ ,  $i = 0, \dots, v - 1, k = 0, \dots, w - 1$ , on which we calculate the double increments. Each such subset consists of

$$\frac{M}{w} \cdot \frac{N}{v} \approx (M \wedge N^{1/2})(N \wedge M^2) = M^3 \wedge N^{3/2}$$

observations and has a balanced design by construction. Therefore, Theorem 4.3 implies the convergence rate  $1/\sqrt{M^3 \wedge N^{3/2}}$ .

**Proposition 4.5.** *Assume  $b > 0$  and let  $\eta = (\sigma^2, \vartheta_2, \kappa)$  lie in the interior of  $H$  for some compact set  $H \subset (0, \infty)^2 \times \mathbb{R}$ . If there exist values  $v \approx \max(1, N/M^2) \in \mathbb{N}$  and  $w \approx \max(1, M/\sqrt{N}) \in \mathbb{N}$  such that  $w\delta/\sqrt{v\tilde{\Delta}}$  is constant, then the estimator given by (18) satisfies*

$$\|\hat{\eta}_{v,w} - \eta\| = \mathcal{O}_P\left(\frac{1}{\sqrt{M^3 \wedge N^{3/2}}}\right), \quad M, N \rightarrow \infty.$$

*Remark 4.6.*

1. The same rate of convergence is achieved if, instead of averaging, one computes the contrast process from a single balanced sub-sample and discards the remaining data. Thus, if  $M^2/N \rightarrow \{0, \infty\}$ , the optimal rate of convergence can be reached by using only a small portion of the available data. On the other hand, our simulation study in Section 6 suggests that using the whole data set is beneficial for the asymptotic variance of the estimator.

2. Integer values  $v$  and  $w$  such that  $w\delta/\sqrt{v\Delta}$  is constant exist, for instance, if the observations are recorded at a dyadic grid, i.e.  $M = 2^m$  and  $N = 4^n$  where  $m, n \rightarrow \infty$ .

Compared to the thinning method of [22], this rate is a considerable improvement. Indeed, it is (almost) optimal in the minimax sense, as shown in Section 5.

We conclude this section by remarking that estimation of all four parameters  $(\sigma^2, \vartheta_2, \vartheta_1, \vartheta_0)$ , which requires  $T \rightarrow \infty$ , is treated in Markussen [28] from a time series perspective and in Kaino and Uchida [22] as well as Hildebrandt [15], assuming high frequency observations in time and space.

### 5. Lower bounds

Our next theorem proves that the estimator  $\hat{\eta}$  from (18) for  $\eta = (\sigma^2, \vartheta_2, \kappa)$  is optimal in the minimax sense, up to a logarithmic factor. To obtain a lower bound, it suffices to consider the sub-problem where  $\vartheta_1 = \vartheta_0 = 0$  and only  $(\sigma^2, \vartheta_2)$  has to be estimated.

**Theorem 5.1.** *Let  $\vartheta_1 = \vartheta_0 = 0$ ,  $(\sigma^2, \vartheta_2) \in H$  for some open set  $H \subset (0, \infty)^2$  and consider observations at  $t_i = i/N$ ,  $i \leq N$ , and  $y_k = b + k\delta$ ,  $k \leq M$ , for some  $b \in [0, 1/2) \cap \mathbb{Q}$ . Then:*

- (i) *If  $\min(M, N)$  remains finite, there is no consistent estimator of  $(\sigma^2, \vartheta_2)$ .*
- (ii) *There is a constant  $c > 0$  such that*

$$\liminf_{M, N \rightarrow \infty} \inf_T \sup_{(\sigma^2, \vartheta_2) \in H} \mathbf{P}_{(\sigma^2, \vartheta_2)} \left( \left\| T - \begin{pmatrix} \sigma^2 \\ \vartheta_2 \end{pmatrix} \right\| > c \cdot r_{M, N} \right) > 0,$$

$$\text{where } r_{M, N} := \begin{cases} N^{-3/4}, & \frac{M}{\sqrt{N}} \gtrsim 1, \\ \left( M^3 \log \frac{N}{M^2} \right)^{-1/2}, & \frac{M}{\sqrt{N}} \rightarrow 0. \end{cases}$$

and  $\inf_T$  is taken over all estimators  $T$  of  $(\sigma^2, \vartheta_2)$  based on observations  $\{X_{t_{i+1}}(y_k) - X_{t_i}(y_k), i < N, k \leq M\}$ .

*Remark 5.2.* The lower bound for the case  $M/\sqrt{N} \gtrsim 1$  is also valid for estimators based on  $\{X_{t_i}(y_k), i \leq N, k \leq M\}$  instead of the increments. We conjecture that this is also true for the case  $M/\sqrt{N} \rightarrow 0$ .

This lower bound shows that, in general,  $(\sigma^2, \vartheta_2)$  cannot be estimated with the parametric rate  $1/\sqrt{MN}$ , in contrast to a conjecture in [7]. Instead, we observe a phase transition in the rate depending on the sampling frequency. The parametric rate can only be attained for a balanced design  $N \approx M^2$ .

The proof of Theorem 5.1 relies on the standard lower bound technique, cf. Tsybakov [37]. Using an inequality by Ibragimov and Has'minskii [20], we will bound the Hellinger distance of the laws of the observations in terms of the corresponding Fisher information for suitably chosen reparametrizations of  $(\sigma^2, \vartheta_2)$ .

For each sampling regime we choose a reparametrization  $(\gamma_1, \gamma_2)$  of  $(\sigma^2, \vartheta_2)$  in such a way that  $\gamma_1$  can be estimated with parametric rate, even without knowledge of  $\gamma_2$ . Bounding the Fisher information for  $\gamma_2$ , we then obtain a lower bound for the simpler problem of estimating the one dimensional parameter  $\gamma_2$ , assuming that  $\gamma_1$  is known. Clearly, the resulting lower bound for  $\gamma_2$  carries over to  $(\gamma_1, \gamma_2)$  and consequently to  $(\sigma^2, \vartheta_2)$ . The main effort, noting that the observations are significantly correlated, is to derive sharp upper bounds for the Fisher information in the different sampling regimes.

In the case  $M/\sqrt{N} \gtrsim 1$  we apply the following bound on the Fisher information for discrete observations of the first  $M$  coefficient processes. Thanks to the Markov property, the probability density function for discrete observations of an Ornstein-Uhlenbeck process is provided by the transition density and allows for explicit computations.

**Proposition 5.3.** *Let  $\vartheta_1 = \vartheta_0 = 0$  and consider a sample  $(u_\ell(i\Delta), \ell \leq M, i \leq N)$  where  $(u_\ell, \ell \in \mathbb{N})$  are independent Ornstein-Uhlenbeck processes given by*

$$du_\ell(t) = -\lambda_\ell u_\ell(t) dt + \sigma d\beta_\ell(t), \quad u_\ell(0) \sim \mathcal{N}\left(0, \frac{\sigma^2}{2\lambda_\ell}\right).$$

*Consider the reparametrization  $(\sigma^2, \rho^2)$  where  $\rho^2 = \sigma^2/\vartheta_2$  and the corresponding Fisher information  $J_{N,M}$ . For  $\max(M, N) \rightarrow \infty$ , the diagonal entries of  $J_{N,M}$  satisfy*

$$J_{N,M}(\sigma^2) = \mathcal{O}(N^{3/2} \wedge (MN)) \quad \text{and} \quad J_{N,M}(\rho^2) = \mathcal{O}(M^3 \wedge (MN)). \quad (19)$$

*In particular,  $\min(J_{N,M}(\sigma^2), J_{N,M}(\rho^2)) \lesssim N^{3/2} \wedge M^3$  for  $\max(N, M) \rightarrow \infty$ .*

*Remark 5.4.*

1. If  $M \lesssim \sqrt{N}$  and  $\sigma^2$  is known, Proposition 5.3 suggests a lower bound of  $M^{-3/2}$  for estimation of  $\vartheta_2$  in the spectral approach. Indeed, this rate is achieved by the maximum likelihood estimator for time continuous observations of the coefficient processes, cf. [26].
2. The reparametrization was chosen since  $\sigma^2$  can be computed from the quadratic variation of any coefficient process  $u_\ell$  when  $N \rightarrow \infty$ , while  $\rho^2$  can be computed from the empirical variance of  $\ell u_\ell(t_i)$ ,  $\ell \leq M$ , for a fixed  $t_i$  as  $M \rightarrow \infty$ , even without knowledge of the other parameter, respectively.

Letting  $M \rightarrow \infty$ , Proposition 5.3 suggests that based on observations of the coefficient processes it is not possible to estimate  $\sigma^2$  (and in particular  $(\sigma^2, \vartheta_2)$ ) at a rate faster than  $N^{-3/4}$ . Further, assuming  $\vartheta_1 = 0$ , the eigenfunctions  $e_\ell(\cdot)$  do not depend on unknown parameters and hence, the space-time discrete observations of the SPDE may be reconstructed from  $\{u_\ell(t_i), i \leq N, \ell \in \mathbb{N}\}$ . Consequently, the lower bound  $N^{-3/4}$  carries over to discrete observations of the SPDE.

Although the lower bounds resulting from Proposition 5.3 and Theorem 5.1 are almost the same, their proofs require a very different reasoning if  $M/\sqrt{N} \rightarrow$

0: In this case, if  $\sigma^2$  is known, Proposition 4.2 shows that it is possible to estimate  $\vartheta_2$  with parametric rate of convergence based on discrete observations of the SPDE whereas Proposition 5.3 suggests that  $\vartheta_2 = \sigma^2/\rho^2$  cannot be estimated at a faster rate than  $M^{-3/2}$  based on the coefficient processes. In particular, both observation schemes are not asymptotically equivalent in the sense of Le Cam.

To derive the lower bound in the case  $M/\sqrt{N} \rightarrow 0$ , we consider the situation where observations are recorded at rational positions  $y_k = \frac{k}{M}$ ,  $k = 1, \dots, M-1$ , where we work with  $M-1$  instead of  $M$  spatial observations for ease of notation. Thus, we potentially add spatial observations on the margin  $[0, b) \cup (1-b, 1]$  which can only increase the amount of information contained in the data. Since  $e_\ell(\cdot) = \sqrt{2} \sin(\pi \ell \cdot)$  is the sine basis, trigonometric identities imply that the vectors

$$\bar{e}_k := (e_k(y_1), \dots, e_k(y_{M-1})) \in \mathbb{R}^{M-1}, \quad k \in \mathbb{N},$$

satisfy  $\bar{e}_{k+2M} = \bar{e}_k$  for all  $k \in \mathbb{N}$  and  $\langle \bar{e}_k, \bar{e}_l \rangle = M \mathbf{1}_{\{k=l \neq M\}} - M \mathbf{1}_{\{k+l=2M\}}$  for  $k, l \leq 2M$ . Equivalently,  $(e_k)_{k=1, \dots, M-1}$  form an orthonormal basis with respect to the empirical scalar product and the relations for  $(\bar{e}_k)_{k \geq 1}$  follow from the symmetry of the sine. Therefore, observing  $\{X_{t_i}(y_k), i \leq N, k \leq M-1\}$  is equivalent to observing  $\{U_k(t_i), i \leq N, k \leq M-1\}$  with

$$U_k(t) := \frac{1}{M} \langle X_t(y), \bar{e}_k \rangle = \sum_{\ell \in \mathcal{I}_k^+} u_\ell(t) - \sum_{\ell \in \mathcal{I}_k^-} u_\ell(t) \quad (20)$$

and  $\mathcal{I}_k^+ := \{k + 2M\ell, \ell \geq 0\}$ ,  $\mathcal{I}_k^- := \{2M - k + 2M\ell, \ell \geq 0\}$ . Since the sets  $\mathcal{I}_k = \mathcal{I}_k^+ \cup \mathcal{I}_k^-$  are disjoint for different values of  $k$ , the processes  $\{U_1, \dots, U_{M-1}\}$  are independent which simplifies the calculation of the Fisher information considerably. Based on their spectral densities and Whittle's formula (35) for the asymptotic Fisher information of a stationary Gaussian time series, we obtain the following result for the increment processes  $\bar{U}_k$ ,  $k \leq M-1$ , defined by

$$\bar{U}_k(j) := U_k(t_{j+1}) - U_k(t_j), \quad j = 0, \dots, N-1. \quad (21)$$

**Proposition 5.5.** *Consider the parametrization  $(\sigma_0^2, \vartheta_2)$  where  $\sigma_0^2 := \sigma^2/\sqrt{\vartheta_2}$ . If  $M/\sqrt{N} \rightarrow 0$ , the Fisher information  $J_{M,N}$  with respect to  $\vartheta_2$  of a sample  $\{\bar{U}_k(j), j \leq N-1, k \leq M-1\}$  satisfies*

$$J_{M,N}(\vartheta_2) = \mathcal{O}\left(M^3 \log \frac{N}{M^2}\right).$$

Note that the reparametrization allows for estimation of  $\sigma_0^2 = \sigma^2/\sqrt{\vartheta_2}$  with parametric rate based on time increments in the regime  $M/\sqrt{N} \rightarrow 0$ , even when  $\vartheta_2$  is unknown. We have considered  $\bar{U}_k$  instead of  $U_k$  due to the technical reason that the  $N$ -th order Fourier approximation of the spectral density of the increment process is positive and hence, a spectral density as well. We conjecture that the same bound holds for the Fisher information of  $U_k$ .

## 6. Simulations

The following numerical example illustrates the asymptotic results for the estimators derived in Section 4. In order to simulate  $X$  on a grid in time and space, we use the *replacement method* developed in [14]: We choose  $L \gg 1/\sqrt{\Delta}$  and simulate for  $\ell \leq L$  the coefficient processes  $(u_\ell(t_i))_{i \leq N}$  from the representation  $X_{t_i}(y_k) = \sum_{\ell \geq 1} u_\ell(t_i) e_\ell(y_k)$  exploiting their AR(1)-structure, namely

$$u_\ell(0) = \frac{\sigma}{\sqrt{2\lambda_\ell}} N_0^\ell, \quad u_\ell(t_{i+1}) = e^{-\lambda_\ell \Delta} u_\ell(t_i) + \sigma \sqrt{\frac{1 - e^{-2\lambda_\ell \Delta}}{2\lambda_\ell}} N_i^\ell, \quad i \in \mathbb{N},$$

where  $(N_i^\ell)$  are independent standard normal random variables. Then, the remaining values  $u_\ell(t_i)$  for  $\ell > L$  are replaced by independent random variables with distribution  $\mathcal{N}(0, \frac{\sigma^2}{2\lambda_\ell})$ . Thanks to relation (20) and the stability of the normal distribution under summation, this can be done without truncation in terms of the number of coefficient processes. It is shown in [14] that — in contrast to naive truncation of the series representation of  $X$  — the distribution of the random vector obtained by the replacement method is close to the true one in total variation distance. Hence, the same holds for the limiting distribution of statistics computed from the simulated data.

Letting  $T = 1$ , we have considered a fixed number  $N = 2^{10}$  or  $N = 2^{14}$  of temporal observations, while  $M$  varies in the set  $\{15, 29, 57, 113, 225, 449, 897, 1793, 3585, 7169\}$ . The precise values for  $M$  stem from the procedure of lying a dyadic grid on  $[0, 1]$  and then removing the points on the margin  $[0, b) \cup (b-1, 1]$  where  $b = 2^{-4}$ . In fact, all observations are obtained as subsets of a simulation of  $X$  on the full grid  $((i/\bar{N}, k/\bar{M}), i \leq \bar{N}, k \leq \bar{M})$  with  $\bar{M} = 2^{13}$  and  $\bar{N} = 2^{14}$ . We have used the replacement method with  $L = \bar{M}$  which is justified in view of  $\bar{M} = 2^{13} \gg 2^7 = \sqrt{\bar{N}}$ . The parameters are chosen to be  $\sigma^2 = 0.1$ ,  $\vartheta_2 = 0.5$ ,  $\vartheta_1 = -0.4$  and  $\vartheta_0 = 0.3$ .

First, we consider the estimators for the volatility  $\sigma^2$  and the diffusivity  $\vartheta_2$  which have been analyzed in Propositions 4.1 and 4.2, respectively. Figure 1 shows the normalized (with respect to  $1/(MN)$  as well as the constants  $\sigma^4$  and  $\vartheta_2^2$ , respectively) mean squared errors based on 500 Monte Carlo iterations plotted against the logarithm of the sampling ratio  $\sqrt{\bar{N}}/M$ . The simplified double increments estimator  $\hat{\vartheta}_{2,r}$  is computed with  $r = (1 - 2b)\frac{\sqrt{\bar{N}}}{M}$ . Using the same value for  $r$ , the simplified double increments estimator for  $\sigma^2$  is computed by replacing the normalization  $\Phi_\vartheta(\delta, \Delta)$  with  $e^{-\kappa\delta/2}\psi_{\vartheta_2}(r)\sqrt{\Delta}$ .

As expected, the estimators based on temporal increments only achieve the parametric rate of convergence as long as  $M$  is not too large, whereas estimators based on space increments only work well when  $M$  is not too small. The estimators based on double increments perform very well throughout any regime depicted in the plot. Even the simplified versions work surprisingly well, although their applicability is only supported by our theory as long as  $M \approx \sqrt{\bar{N}}$ . In particular, the double increments estimator for  $\sigma^2$  can barely be distinguished from the simplified one. Furthermore, as suggested by the theory, the simulations show that the estimators based on space increments or time increments

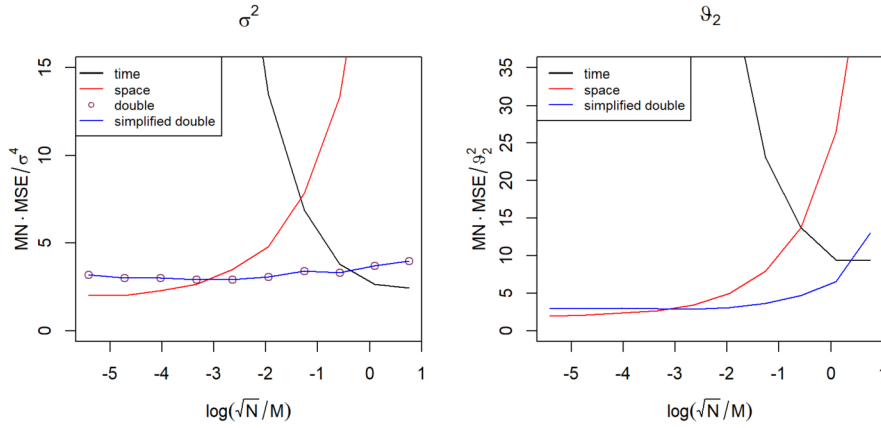


FIG 1. Normalized mean squared errors of estimators for  $\sigma^2$  (left) and  $\vartheta_2$  (right) with  $T = 1$ ,  $N = 2^{10}$  and  $M \in [15, 7169]$ .

have a smaller mean squared error than the double increments estimators in the regimes  $\sqrt{N}/M \rightarrow 0$  or  $\sqrt{N}/M \rightarrow \infty$ , respectively.

The above estimators require all but one of the parameters  $(\sigma^2, \vartheta_2, \kappa)$  to be known. Within the more difficult statistical problem where all parameters are unknown,  $\eta = (\sigma^2, \vartheta_2, \vartheta_1)$  can be estimated by  $\hat{\eta}$  from (17) and by  $\hat{\eta}_{v,w}$  from (18). Furthermore, we have implemented a data-thinning version of the estimator where the contrast process uses only a single balanced sub-sample and discards the remaining data instead of averaging. For the estimator  $\hat{\eta}_{v,w}$  and its thinning version, we set  $v = \lceil \max(1, \frac{N}{M^2}) \rceil$  and  $w = \lceil \max(1, M/\sqrt{N}) \rceil$  where  $\lceil \cdot \rceil$  indicates rounding to the next integer. The minimization problems were numerically solved using the nonlinear least squares function `nls` from R. Figure 2 shows the logarithm of the mean squared errors plotted against the logarithm of the sampling ratio  $\sqrt{N}/M$ , again based on 500 Monte Carlo iterations. Here, displaying the mean squared errors on the logarithmic scale helps in distinguishing the different curves and provides a close-up view at their behavior when they are very small.

For the fixed value  $N = 2^{14}$ , taking  $M = 113$  results in a balanced regime and, in particular, we have  $v = w = 1$ . Thus, the definitions of all estimators agree, leading to an intersection of the three curves at  $\log(\sqrt{N}/M) \approx 0.12$ . In contrast to the double increments estimators for single parameters,  $\hat{\eta}$  only produces good results as long as  $M \approx \sqrt{N}$ , which is covered by the theoretical foundation. In fact, with the smallest number of spatial observations,  $M = 15$ , the optimization algorithm was even unable to detect a minimum in almost 3/5 of the simulation runs and the mean squared error is computed based on the remaining data. Unsurprisingly, the other two estimators have, overall, a much better performance. On the contrary, when  $M = 57$  ( $\log(\sqrt{N}/M) \approx -0.56$ ) the estimator  $\hat{\eta}$  works slightly better. This can be explained by the fact that, here, the choice of  $v$  and  $w$  is too conservative in the sense that  $v \vee w > 1$  although the

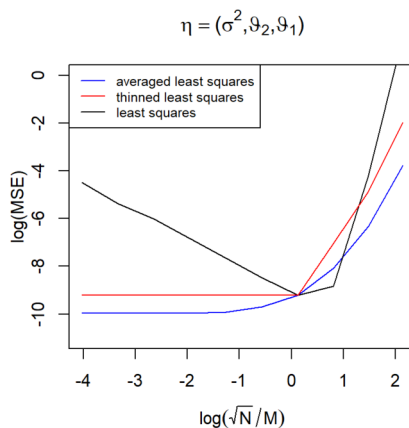


FIG 2. *Logarithm of the mean squared errors of the least squares estimator  $\hat{\eta}$ , its averaging version  $\hat{\eta}_{v,w}$  and a thinning version exploiting only one balanced sub-sample. The time horizon is  $T = 1$  and the sample sizes are  $N = 2^{14}$  and  $M \in [15, 7169]$ .*

regime is still reasonably balanced. Furthermore, we see that it is only possible to profit from an increasing number of spatial observations up to a certain degree: For  $M \leq \sqrt{N}$ , the optimal rate is  $M^{-3/2}$  and the empirical mean squared error of  $\hat{\eta}_{v,w}$  as well as its thinning version becomes increasingly smaller. For  $M \geq \sqrt{N}$ , the optimal rate is  $N^{-3/4}$  and, indeed, the empirical mean squared errors become stationary. Furthermore, while the latter two estimators have a similar qualitative behavior, the mean squared error of  $\hat{\eta}_{v,w}$  is consistently smaller. As announced in Remark 4.6, this indicates that using the whole data results in an improved asymptotic variance.

## 7. Proofs of the main results

### 7.1. Proofs for the central limit theorems for realized quadratic variations

First, we prove the generic central limit result in Proposition 3.1. Afterwards, we can verify the central limit theorems for realized quadratic variations based on spatial increments (Theorem 3.3) and double increments (Theorem 3.7).

*Proof of Proposition 3.1.* Since  $\Sigma_n = Q_n^\top \Lambda_n Q_n$  for an orthogonal matrix  $Q_n \in \mathbb{R}^{d_n \times d_n}$  and a diagonal matrix  $\Lambda_n$ , the vector  $Z_{\bullet,n}$  has the same distribution as  $B_n X^n$  for  $B_n := Q_n^\top \Lambda_n^{1/2}$  and  $X^n := (X_1, \dots, X_{d_n})$  with independent standard normal random variables  $(X_k)_{k \in \mathbb{N}}$ . Denoting  $A_n = \text{diag}(\alpha_{1,n}, \dots, \alpha_{d_n,n})$ , we obtain  $S_n = Z_{\bullet,n}^\top A_n Z_{\bullet,n} \stackrel{\mathcal{D}}{=} X^{n\top} B_n^\top A_n B_n X^n$ . Furthermore,  $B_n^\top A_n B_n$  is symmetric such that  $B_n^\top A_n B_n = P_n^\top \Gamma_n P_n$  where  $P_n$  is an orthogonal matrix and  $\Gamma_n$  is a diagonal matrix. Since  $P_n X^n \sim \mathcal{N}(0, E_{d_n})$ , we conclude as in

[29, p. 36]

$$S_n \stackrel{\mathcal{D}}{=} X^n \top B_n \top A_n B_n X^n = (P_n X^n) \top \Gamma (P_n X^n) \stackrel{\mathcal{D}}{=} X^n \top \Gamma_n X^n = \sum_{i=1}^{d_n} \gamma_{i,n} X_i^2,$$

where  $\gamma_{i,n}$ ,  $i \leq d_n$  are the eigenvalues of  $B_n \top A_n B_n$ . The statement now follows by Lyapunov's condition and  $\|B_n\|_2^2 = \|\Sigma_n\|_2$ :

$$\begin{aligned} \frac{\sum_{i=1}^{d_n} \gamma_{i,n}^4 \mathbf{E} \left( (X_k^2 - \mathbf{E} X_k^2)^4 \right)}{(\text{Var} S_n)^2} &\approx \frac{\sum_{i=1}^{d_n} \gamma_{i,n}^4}{\left( \sum_{i=1}^{d_n} \gamma_{i,n}^2 \right)^2} \lesssim \frac{\max_{i \leq d_n} \gamma_{i,n}^2}{\sum_{i=1}^{d_n} \gamma_{i,n}^2} = \frac{\|B_n \top A_n B_n\|_2^2}{\text{Var} S_n} \\ &\leq \frac{(\|B_n\|_2^2 \|A_n\|_2)^2}{\text{Var} S_n} = \frac{\|\Sigma\|_2^2}{\text{Var} S_n}. \quad \square \end{aligned}$$

Throughout, for a function  $f: \mathbb{R} \rightarrow \mathbb{R}$  we use the notation

$$D_\delta f(x) := f(x + \delta) - f(x) \quad \text{and} \quad D_\delta^2 f(x) := f(x + 2\delta) - 2f(x + \delta) + f(x).$$

*Proof of Theorem 3.3.* We abbreviate the (rescaled) space increments by

$$S_{ik} := (\delta_k^M X)(t_i) \quad \text{and} \quad \tilde{S}_{ik} := e^{\kappa y_k/2} (\delta_k^M X)(t_i).$$

*Step 1.* We calculate the asymptotic mean of  $V_{\text{sp}}$ . Application of the trigonometric identity  $\sin(\alpha) \sin(\beta) = \frac{1}{2} (\cos(\alpha - \beta) - \cos(\alpha + \beta))$  yields

$$\begin{aligned} &e^{\kappa x/2} (e_\ell(x + \delta) - e_\ell(x)) e^{\kappa y/2} (e_\ell(y + \delta) - e_\ell(y)) \\ &= g(\delta) (2 \cos(\pi \ell(y - x)) - \cos(\pi \ell(y - x - \delta)) - \cos(\pi \ell(y - x + \delta))) \quad (22) \\ &\quad + (g(2\delta) + g(0) - 2g(\delta)) (\cos(\pi \ell(y - x))) \\ &\quad + 2g(\delta) \cos(\pi \ell(y + x + \delta)) - g(0) \cos(\pi \ell(y + x)) - g(2\delta) \cos(\pi \ell(x + y + 2\delta)), \end{aligned}$$

where  $g(x) = \exp(-\kappa x/2)$ . Plugging in  $x = y$  gives

$$\begin{aligned} &e^{\kappa y} (e_\ell(y + \delta) - e_\ell(y))^2 \\ &= 2(1 - \cos(\pi \ell \delta)) + 2(1 - g(\delta)) (\cos(\pi \ell \delta) - 1) + (g(2\delta) + g(0) - 2g(\delta)) \quad (23) \\ &\quad + 2g(\delta) \cos(\pi \ell(2y + \delta)) - g(2\delta) \cos(2\pi \ell(y + \delta)) - g(0) \cos(2\pi \ell y). \end{aligned}$$

Thus, in terms of

$$f(y) := \sum_{\ell \geq 1} \frac{1}{2\lambda_\ell} \cos(\pi \ell y), \quad y \in [0, 1],$$

we have

$$\begin{aligned} \mathbf{E} \left( e^{\kappa y} (X_t(y + \delta) - X_t(y))^2 \right) &= \sigma^2 \sum_{\ell \geq 0} \frac{1}{2\lambda_\ell} e^{\kappa y} (e_\ell(y + \delta) - e_\ell(y))^2 \\ &= \sigma^2 \left( -2D_\delta f(0) - 2D_\delta g(0) D_\delta f(0) + f(0) D_\delta^2 g(0) - D_\delta^2 (g(\cdot) f(2y + \cdot)) (0) \right). \end{aligned}$$



Owing to its closed form expression in (41) below, we see that  $f \in C_b^\infty([0, 2])$  and  $f'(0) = -\frac{1}{4\vartheta_2}$ . Hence,

$$\mathbf{E} \left( e^{\kappa y} (X_t(y + \delta) - X_t(y))^2 \right) = -2\sigma^2 f'(0) \cdot \delta + \mathcal{O}(\delta^2) = \frac{\sigma^2}{2\vartheta_2} \cdot \delta + \mathcal{O}(\delta^2).$$

For  $y = y_k$  we obtain the asymptotic mean  $\mathbf{E}(V_{\text{sp}}) = \frac{\sigma^2}{2\vartheta_2} + \mathcal{O}(\delta)$  and in particular, under the condition  $N/M \rightarrow 0$ ,

$$\sqrt{MN} \left( V_{\text{sp}} - \frac{\sigma^2}{2\vartheta_2} \right) = \sqrt{MN} (V_{\text{sp}} - \mathbf{E}(V_{\text{sp}})) + o(1).$$

Step 2. We calculate the asymptotic variance. By Isserlis' Theorem [21] we have

$$\text{Cov}((\tilde{S}_{ik})^2, (\tilde{S}_{jl})^2) = 2 \text{Cov}(\tilde{S}_{ik}, \tilde{S}_{jl})^2.$$

Together with the symmetry  $\text{Cov}(\tilde{S}_{ik}, \tilde{S}_{jl}) = \text{Cov}(\tilde{S}_{jk}, \tilde{S}_{il})$ , this implies

$$\text{Var}(V_{\text{sp}}) = \frac{2}{N^2 M^2 \delta^2} (v_1 + v_2 + v_3 + v_4), \quad \text{where}$$

$$\begin{aligned} v_1 &:= \sum_{i=0}^{N-1} \sum_{k=0}^{M-1} \text{Var}(\tilde{S}_{ik})^2, \\ v_2 &:= 2 \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} \sum_{k=0}^{M-1} \text{Cov}(\tilde{S}_{ik}, \tilde{S}_{jk})^2, \\ v_3 &:= 2 \sum_{i=0}^{N-1} \sum_{k=0}^{M-2} \sum_{l=k+1}^{M-1} \text{Cov}(\tilde{S}_{ik}, \tilde{S}_{il})^2, \\ v_4 &:= 4 \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} \sum_{k=0}^{M-2} \sum_{l=k+1}^{M-1} \text{Cov}(\tilde{S}_{ik}, \tilde{S}_{jl})^2. \end{aligned}$$

We have already shown that  $\text{Var}(\tilde{S}_{ik}) = \mathbf{E}((\tilde{S}_{ik})^2) = \frac{\sigma^2}{2\vartheta_2} \cdot \delta + \mathcal{O}(\delta^2)$ . Therefore,

$$v_1 = NM\delta^2 \cdot \frac{\sigma^4}{4\vartheta_2^2} + \mathcal{O} \left( \frac{N}{M^2} \right) = NM\delta^2 \cdot \frac{\sigma^4}{4\vartheta_2^2} + o \left( \frac{N}{M} \right).$$

In the sequel, we show that the remaining covariances do not contribute to the asymptotic variance.

For  $v_2$  we define  $\omega := \vartheta_2(\pi^2 \wedge (\pi^2 + \Gamma)) > 0$  such that  $\lambda_\ell \geq \omega \ell^2$  for all  $\ell \in \mathbb{N}$ . Since  $(e_\ell(y_{k+1}) - e_\ell(y_k))^2 \lesssim \ell^2 \delta^2$ , we get for  $J = |i - j| \geq 1$

$$\text{Cov}(\tilde{S}_{ik}, \tilde{S}_{jk}) = \sigma^2 \sum_{\ell \geq 1} \frac{e^{-\lambda_\ell J \Delta}}{2\lambda_\ell} e^{\kappa y_k} (e_\ell(y_{k+1}) - e_\ell(y_k))^2$$

$$\lesssim \delta^2 \sum_{\ell \geq 1} e^{-\omega \ell^2 J \Delta} \lesssim \frac{\delta^2}{\sqrt{J \Delta}}$$

where the last step follows by Riemann summation with mesh size  $\sqrt{J \Delta}$ . Since  $\frac{\log N}{M^2 \Delta} \leq \frac{N}{M^2 \Delta} = \frac{N^2}{M^2 T} \rightarrow 0$ ,

$$v_2 \lesssim \frac{M \delta^4}{\Delta} \sum_{i=0}^{N-1} \sum_{j=i+1}^N \frac{1}{(j-i)} \leq \frac{NM \delta^4}{\Delta} \sum_{i=1}^N \frac{1}{i} = \mathcal{O}\left(\frac{N \log N}{M^3 \Delta}\right) = o\left(\frac{N}{M}\right).$$

To bound  $v_3$  we follow the same strategy as for the mean: Since (22) consists exclusively of second order differences we have  $\text{Cov}(\tilde{S}_{ik}, \tilde{S}_{il}) = \mathcal{O}(\delta^2)$  for  $k \neq l$ . Therefore,  $v_3 = \mathcal{O}(NM^2 \delta^4) = o(N/M)$ .

To estimate  $v_4$ , we deduce from (22) for  $k < l$  and  $J = |i - j| \geq 1$  that

$$\begin{aligned} \text{Cov}(\tilde{S}_{ik}, \tilde{S}_{jl}) &= -g(\delta) D_\delta^2 f_{J \Delta}(y_l - y_{k+1}) \\ &\quad + f_{J \Delta}(y_l - y_k) D_\delta^2 g(0) - D_\delta^2 (g(\cdot) f_{J \Delta}(y_l + y_k + \cdot))(0), \quad \text{where} \\ f_t(y) &:= \sigma^2 \sum_{\ell \geq 1} \frac{e^{-\lambda_\ell t}}{2\lambda_\ell} \cos(\pi \ell y). \end{aligned}$$

By Riemann summation we have  $f_t''(y) \lesssim \sum_{\ell \geq 1} e^{-\lambda_\ell t} \lesssim \frac{1}{\sqrt{t}}$ . On the other hand, by Lemma A.7,

$$f_t''(y) \lesssim \frac{1}{y \wedge (2-y)} \sup_k \left| \frac{k^2}{\lambda_k} e^{-\lambda_k t} \right| \lesssim \frac{1}{y \wedge (2-y)}.$$

Therefore,

$$f_t''(y) \lesssim B(t, y) := \frac{1}{y \wedge (2-y)} \wedge \frac{1}{\sqrt{t}}.$$

Similarly,  $f_t(y), f_t'(y) \lesssim B(t, y)$  can be shown. We conclude

$$\begin{aligned} v_4 &\lesssim NM \sum_{i=0}^{N-1} \sum_{k=0}^{2M-2} \delta^4 B\left(i\Delta, \frac{k}{M}\right)^2 \lesssim \frac{N}{M^3} \sum_{i=1}^N \sum_{k=0}^M \frac{M^2}{k^2} \wedge \frac{1}{i\Delta} \\ &= \frac{N}{M^3} \sum_{i=1}^N \left( \sum_{k < M\sqrt{i\Delta}} \frac{1}{i\Delta} + \sum_{M \geq k \geq M\sqrt{i\Delta}} \frac{M^2}{k^2} \right) \lesssim \frac{N}{M^3} \sum_{i=1}^N \frac{M}{\sqrt{i\Delta}} \\ &\lesssim \frac{N^{3/2}}{M^2 \sqrt{\Delta}} = o\left(\frac{N}{M}\right) \end{aligned}$$

where the last step follows from  $\frac{\sqrt{N}}{M\sqrt{\Delta}} = \frac{N}{M} \frac{1}{\sqrt{T}} \rightarrow 0$ . Summing up, we have proved that

$$\text{Var}(V_{\text{sp}}) = \frac{\sigma^4}{2\vartheta_2^2} \cdot \frac{1}{MN} + o\left(\frac{1}{NM}\right).$$

*Step 3.* To prove asymptotic normality, we interpret the number of temporal and spatial observations as sequences  $M = M_n, N = N_n$  indexed by  $n \in \mathbb{N}$  and consider the triangular array  $(Z_{ik,n}, n \in \mathbb{N}, k < M_n, i < N_n)$ , where  $Z_{ik,n} = \tilde{S}_{ik}/\sqrt{NM\delta}$ . Since  $\text{Var}(\sum_{i,k} Z_{ik}^2) \approx (MN)^{-1}$ , Proposition 3.1 applies if:

$$\frac{1}{MN\delta^2} \left( \sum_{i,k} |\text{Cov}(\tilde{S}_{ik}, \tilde{S}_{jl})| \right)^2 \rightarrow 0$$

uniformly in  $j < N, l < M$  in view of criterion (8). The covariance bounds in Step 2 yield uniformly in  $j < N, k < M$ :

$$\begin{aligned} \sum_{k < M} |\text{Cov}(\tilde{S}_{jk}, \tilde{S}_{jl})| &= \mathcal{O}(\delta), & \sum_{i < N} |\text{Cov}(\tilde{S}_{il}, \tilde{S}_{jl})| &= \mathcal{O}(\delta^2 \sqrt{N}/\sqrt{\Delta}), \\ \left( \sum_{i \neq j, k \neq l} |\text{Cov}(\tilde{S}_{ik}, \tilde{S}_{jl})| \right)^2 &\lesssim MN \sum_{i \neq j, k \neq l} |\text{Cov}(\tilde{S}_{ik}, \tilde{S}_{jl})|^2 = o(N/M), \end{aligned}$$

where we have used the Cauchy-Schwarz inequality to obtain the last bound. It remains to note  $N/M \rightarrow 0$  and  $N\Delta \gtrsim 1$ . □

The proof of Theorem 3.7 is similar to the previous one but the more complex covariance structure of the double increments has to be taken into account carefully, see Section A.1. The (asymptotic) mean of the realized quadratic space-time variation is provided by Proposition 3.5, which we prove first. In the following, we write

$$\tilde{D}_{ik} := e^{\kappa y_k/2} D_{ik}. \tag{24}$$

*Proof of Proposition 3.5. Step 1.* We show asymptotic independence of  $\Gamma$ , i.e.,

$$\mathbf{E}((D_{ik})^2) = \sigma^2 \sum_{\ell \geq 1} \frac{1 - e^{-\pi^2 \vartheta_2 \ell^2 \Delta}}{\pi^2 \vartheta_2 \ell^2} (e_\ell(y_{k+1}) - e_\ell(y_k))^2 + \mathcal{O}(\delta \sqrt{\Delta} (\delta \wedge \sqrt{\Delta})).$$

Define  $f(x) := \frac{1 - e^{-x}}{x}$ . A first order Taylor approximation of  $f$  yields

$$\mathbf{E}((D_{ik})^2) = \sigma^2 \Delta \sum_{\ell \geq 1} f(\pi^2 \vartheta_2 \ell^2 \Delta) (e_\ell(y_{k+1}) - e_\ell(y_k))^2 + R$$

where  $R \lesssim \Delta^2 \sum_{\ell \geq 1} f'(\vartheta_2(\pi^2 \ell^2 + \xi_\ell)\Delta) (e_\ell(x + \delta) - e_\ell(x))^2$  for some  $|\xi_k| \leq |\Gamma|$ . Since

$$\begin{aligned} &(e_\ell(y + \delta) - e_\ell(y))^2 \\ &\lesssim \left( e^{-\kappa \delta/2} (\sin(\pi \ell (y + \delta)) - \sin(\pi \ell y)) + \sin(\pi \ell y) (e^{-\kappa \delta/2} - 1) \right)^2 \lesssim 1 \wedge (\ell \delta)^2 \end{aligned}$$

and noting that  $f'(x^2)$  and  $x^2 f'(x^2)$  are integrable, we deduce

$$R \lesssim \Delta^2 \sum_{\ell \geq 1} (1 \wedge (\ell \delta)^2) f'(\vartheta_2(\pi^2 \ell^2 + \xi_\ell)\Delta)$$

$$= \mathcal{O}(\Delta^{3/2} \wedge (\delta^2 \sqrt{\Delta})) = \mathcal{O}((\delta \Delta) \wedge (\delta^2 \sqrt{\Delta})).$$

Step 2. We verify (i). Thanks to Step 1 we may assume  $\lambda_\ell = \pi^2 \vartheta_2 \ell^2$ . It follows from (23) that

$$\begin{aligned} \mathbf{E}(\tilde{D}_{ik}^2) &= \sigma^2 e^{-\kappa y} \left( F_{\vartheta_2}(0, \Delta) (1 + e^{-\kappa \delta}) - 2F_{\vartheta_2}(\delta, \Delta) e^{-\kappa \delta/2} \right) \\ &\quad - \sigma^2 e^{-\kappa y} D_\delta^2 \left( g(\cdot) F_{\vartheta_2}(2y_k + \cdot, \Delta) \right) (0). \end{aligned}$$

Consequently, it remains to show

$$D_\delta^2 \left( g(\cdot) F_{\vartheta_2}(2y + \cdot, \Delta) \right) (0) = \mathcal{O} \left( \delta \sqrt{\Delta} \left( \delta \wedge \sqrt{\Delta} \right) \right)$$

uniformly in  $y \in [b, 1 - b]$ . As before, this is done by showing

$$F_{\vartheta_2}(x, \Delta) \lesssim \Delta, \quad \frac{\partial F_{\vartheta_2}(x, \Delta)}{\partial x} \lesssim \Delta \quad \text{and} \quad \frac{\partial^2 F_{\vartheta_2}(x, \Delta)}{\partial x^2} \lesssim \sqrt{\Delta}$$

uniformly in  $x \in [2b, 2(1 - b)]$ . From Lemma A.8 it follows that  $F_{\vartheta_2}(x, \Delta) = \Delta \sum_{\ell \geq 1} f(\lambda_\ell \Delta) \cos(\pi \ell x) = \mathcal{O}(\Delta)$ . In order to access the first two derivatives of  $F_{\vartheta_2}(\cdot, \Delta)$ , we split it into two summands,

$$F_{\vartheta_2}(x, \Delta) = \underbrace{\Delta \sum_{\ell \geq 1} \frac{1}{1 + \lambda_\ell \Delta} \cos(\pi \ell x)}_{=: H_\Delta(x)} + \underbrace{\Delta \sum_{\ell \geq 1} \left( \frac{1 - e^{-\lambda_\ell \Delta}}{\lambda_\ell \Delta} - \frac{1}{1 + \lambda_\ell \Delta} \right) \cos(\pi \ell x)}_{=: G_\Delta(x)}.$$

Using the cosine series formula (41), we can compute

$$H_\Delta(x) = \frac{1}{\vartheta_2 \pi^2} \sum_{\ell \geq 1} \frac{1}{\ell^2 + \frac{1}{\pi^2 \vartheta_2 \Delta}} \cos(\pi \ell x) = \frac{\sqrt{\Delta}}{2\sqrt{\vartheta_2}} \frac{\cosh\left(\frac{1}{\sqrt{\vartheta_2} \Delta}(x - 1)\right)}{\sinh\left(\frac{1}{\sqrt{\vartheta_2} \Delta}\right)} - \frac{\Delta}{2},$$

from which it easily follows that  $H'_\Delta(x) \lesssim \Delta$  and  $H''_\Delta(x) \lesssim \sqrt{\Delta}$ . The derivatives of

$$G_\Delta(x) = \Delta \sum_{\ell \geq 1} h(\ell \sqrt{\Delta}) \cos(\pi \ell x), \quad \text{where} \quad h(z) := \frac{1 - e^{-z}(1 + z)}{z(1 + z)},$$

can be bounded summand-wisely,

$$\begin{aligned} G'_\Delta(x) &\approx \sqrt{\Delta} \sum_{\ell \geq 1} (\ell \sqrt{\Delta}) h(\ell \sqrt{\Delta}) \sin(\pi \ell x) \lesssim \Delta, \\ G''_\Delta(x) &\approx \sum_{\ell \geq 1} (\ell^2 \Delta) h(\ell \sqrt{\Delta}) \cos(\pi \ell x) \lesssim \sqrt{\Delta}, \end{aligned}$$

where the bounds follow from the Riemann sum approximations in Lemma A.8, owing to  $xh(x)|_{x=0} = x^2h(x)|_{x=0} = 0$ .

Step 3. We show the asymptotic expressions in (ii). Due to a Riemann sum argument, we have  $\|F_{\vartheta_2}(\cdot, \Delta)\|_\infty \lesssim \sqrt{\Delta}$  and consequently,

$$\begin{aligned} \Phi_\vartheta(\delta, \Delta) &= 2(F_{\vartheta_2}(0, \Delta) - F_{\vartheta_2}(\delta, \Delta)) + F_{\vartheta_2}(0, \Delta) \left[1 + e^{-\kappa\delta} - 2e^{-\kappa\delta/2}\right] \\ &\quad - 2(F_{\vartheta_2}(\delta, \Delta) - F_{\vartheta_2}(0, \Delta)) \left(e^{-\kappa\delta/2} - 1\right) \\ &= 2(F_{\vartheta_2}(0, \Delta) - F_{\vartheta_2}(\delta, \Delta)) + \mathcal{O}(\delta\sqrt{\Delta}). \end{aligned}$$

In the case  $\delta/\sqrt{\Delta} \rightarrow 0$  Taylor’s formula yields

$$F_{\vartheta_2}(0, \Delta) - F_{\vartheta_2}(\delta, \Delta) = -\delta \frac{\partial F_{\vartheta_2}(0, \Delta)}{\partial x} - \frac{\delta^2}{2} \frac{\partial^2 F_{\vartheta_2}(\eta, \Delta)}{\partial x^2}$$

for some  $\eta \in [0, \delta]$ . We employ the representation  $F_{\vartheta_2}(\cdot, \Delta) = H_\Delta + G_\Delta$  from Step 2: Since  $\sin(0) = 0$  we have  $\frac{\partial F_{\vartheta_2}(0, \Delta)}{\partial x} = H'_\Delta(0) = -\frac{1}{2\vartheta_2}$ . Further,  $H''_\Delta(\eta) = 1/\sqrt{\Delta}$  and the Riemann sum argument yields  $G''_\Delta(\eta) \lesssim \sum_{\ell \geq 1} (\ell^2 \Delta) h(\ell\sqrt{\Delta}) \lesssim 1/\sqrt{\Delta}$ . Therefore,  $F_{\vartheta_2}(0, \Delta) - F_{\vartheta_2}(\delta, \Delta) = \frac{1}{2\vartheta_2} \cdot \delta + \mathcal{O}\left(\frac{\delta^2}{\sqrt{\Delta}}\right)$ .

If  $\delta/\sqrt{\Delta} \rightarrow \infty$ , Lemma A.8 implies  $F_{\vartheta_2}(\delta, \Delta) = -\frac{\Delta}{2} + \mathcal{O}\left(\frac{\Delta^{3/2}}{\delta^2}\right)$  and Lemma A.9 yields

$$F_{\vartheta_2}(0, \Delta) = \sqrt{\Delta} \int_0^\infty \frac{1 - e^{-\pi^2 \vartheta_2 z^2}}{\pi^2 \vartheta_2 z^2} dz - \frac{\Delta}{2} + \mathcal{O}(\Delta^{3/2}). \tag{25}$$

Since  $\int_0^\infty \frac{1 - e^{-\pi^2 \vartheta_2 z^2}}{\pi^2 \vartheta_2 z^2} dz = \frac{1}{\sqrt{\vartheta_2 \pi}}$ , we obtain  $F_{\vartheta_2}(0, \Delta) - F_{\vartheta_2}(\delta, \Delta) = \frac{\sqrt{\Delta}}{\sqrt{\vartheta_2 \pi}} + \mathcal{O}\left(\frac{\Delta^{3/2}}{\delta^2}\right)$ .

Finally, we derive the asymptotic expression for the case  $\delta/\sqrt{\Delta} \equiv r$ , while  $\delta/\sqrt{\Delta} \rightarrow r$  can be handled similarly. We have

$$\begin{aligned} \Phi_\vartheta(\delta, \Delta) &= 2(F_{\vartheta_2}(0, \Delta) - F_{\vartheta_2}(\delta, \Delta))e^{-\kappa\delta/2} + F_{\vartheta_2}(0, \Delta)(1 + e^{-\kappa\delta} - 2e^{-\kappa\delta/2}) \\ &= 2(F_{\vartheta_2}(0, \Delta) - F_{\vartheta_2}(\delta, \Delta))e^{-\kappa\delta/2} + \mathcal{O}(\Delta^{3/2}) \end{aligned}$$

and since  $1 - \cos(0) = 0$ , Lemma A.9 yields

$$\begin{aligned} F_{\vartheta_2}(0, \Delta) - F_{\vartheta_2}(r\sqrt{\Delta}, \Delta) &= \sum_{\ell \geq 1} \frac{1 - e^{-\pi^2 \vartheta_2 \ell^2 \Delta}}{\pi^2 \vartheta_2 \ell^2} \left(1 - \cos(\pi \ell r \sqrt{\Delta})\right) \\ &= \sqrt{\Delta} \int_0^\infty \frac{1 - e^{-\pi^2 \vartheta_2 z^2}}{\pi^2 \vartheta_2 z^2} (1 - \cos(\pi r z)) dz + \mathcal{O}(\Delta^{3/2}). \end{aligned}$$

It remains to compute the integral. By substituting  $\tilde{r} = r/\sqrt{\vartheta_2}$  we can pass to

$$\int_0^\infty \frac{1 - e^{-\pi^2 \vartheta_2 z^2}}{\pi^2 \vartheta_2 z^2} (1 - \cos(\pi r z)) dz = \frac{1}{\pi\sqrt{\vartheta_2}} \left(h_1(\tilde{r}) - h_2(\tilde{r})\right)$$

where

$$h_1(\tilde{r}) = \int_0^\infty \frac{1 - \cos(\tilde{r}z)}{z^2} dz, \quad h_2(\tilde{r}) = \int_0^\infty e^{-z^2} \frac{1 - \cos(\tilde{r}z)}{z^2} dz.$$

To compute  $h_1$ , note that  $S(z) + \frac{\cos(z)-1}{z}$  is an antiderivative of  $\frac{1-\cos(z)}{z^2}$ , where  $S(z) = \int_0^z \frac{\sin(h)}{h} dh$  is the sine integral. Consequently, a substitution and the property  $\lim_{z \rightarrow \infty} S(z) = \pi/2$  yield

$$h_1(\tilde{r}) = \tilde{r} \int_0^\infty \frac{1 - \cos(z)}{z^2} dz = \frac{\pi \tilde{r}}{2}.$$

To treat  $h_2$ , note that  $h_2(0) = h_2'(0) = 0$  and hence,  $h_2(\tilde{r}) = \int_0^{\tilde{r}} \int_0^s h_2''(u) du ds$ . Now, plugging in  $h_2''(\tilde{r}) = \int_0^\infty e^{-z^2} \cos(\tilde{r}z) dz = \frac{\sqrt{\pi}}{2} e^{-\tilde{r}^2/4}$  and integrating by parts yields

$$h_2(\tilde{r}) = \frac{\sqrt{\pi}}{2} \int_0^{\tilde{r}} \int_0^s e^{-u^2/4} du = \sqrt{\pi} \tilde{r} \int_0^{\tilde{r}/2} e^{-u^2} du + \sqrt{\pi} (e^{-\tilde{r}^2/4} - 1).$$

The claim thus follows from

$$\begin{aligned} h_1(\tilde{r}) - h_2(\tilde{r}) &= \frac{\pi \tilde{r}}{2} \left( 1 - \frac{2}{\sqrt{\pi}} \int_0^{\tilde{r}/2} e^{-u^2} du \right) + \sqrt{\pi} (1 - e^{-\tilde{r}^2/4}) \\ &= \tilde{r} \sqrt{\pi} \int_{\tilde{r}/2}^\infty e^{-u^2} du + \sqrt{\pi} (1 - e^{-\tilde{r}^2/4}). \quad \square \end{aligned}$$

*Proof of Theorem 3.7.* Asymptotic normality follows just like in the proof of Theorem 3.3. Using the notation from the proof of the latter theorem (with space increments replaced by double increments) we have

$$\text{Var}(\mathbb{V}) = \frac{2}{M^2 N^2 \Phi_\theta^2(\delta, \Delta)} (v_1 + v_2 + v_3 + v_4).$$

To determine the asymptotic variances, we have to treat the three different sampling regimes separately.

Case  $\delta/\sqrt{\Delta} \rightarrow 0$ . By Lemmas A.1 and A.2 we have

$$\text{Var}(\tilde{D}_{ki})^2 = \frac{\sigma^4}{4\vartheta_2^2} e^{-\kappa\delta} \cdot \delta^2 + o(\delta^2), \quad \text{Cov}(\tilde{D}_{ki}, \tilde{D}_{k(i+1)})^2 = \frac{\sigma^4}{4\vartheta_2^2} e^{-\kappa\delta} \cdot \delta^2 + o(\delta^2)$$

as well as

$$\begin{aligned} \text{Cov}(\tilde{D}_{ki}, \tilde{D}_{kj})^2 &= o\left(\frac{\delta^2}{|i-j|^5}\right), \quad |i-j| \geq 2, \\ \text{Cov}(\tilde{D}_{ki}, \tilde{D}_{lj})^2 &= \mathcal{O}\left(\frac{\delta^4}{(|i-j|+1)^4} \left(\frac{M^2}{(k-l)^2} \wedge \frac{1}{\Delta}\right)\right), \quad k \neq l. \end{aligned}$$

Since  $\sum_{k \leq M} (\frac{M^2}{k^2} \wedge \frac{1}{\Delta}) \lesssim \frac{M}{\sqrt{\Delta}}$  (cf. the proof of Theorem 3.3), the latter covariances are negligible for the asymptotic variance. Inserting  $\Phi_{\vartheta}^2(\delta, \Delta) = \frac{e^{-\kappa\delta}}{\vartheta_2} \delta^2 + o(\delta^2)$  from Proposition 3.5 yields the claim.

Case  $\delta/\sqrt{\Delta} \rightarrow \infty$ . By Lemmas A.1 and A.3 we have

$$\text{Var}(\tilde{D}_{ki})^2 = \frac{4\sigma^4}{\pi\vartheta_2} e^{-\kappa\delta} \cdot \Delta + o(\Delta), \quad \text{Cov}(\tilde{D}_{ki}, \tilde{D}_{(k+1)i})^2 = \frac{\sigma^4}{\pi\vartheta_2} e^{-\kappa\delta} \cdot \Delta + o(\Delta).$$

From  $d(J) := \sqrt{J-1} + \sqrt{J+1} - 2\sqrt{J} = \mathcal{O}(J^{-3/2})$  and  $\sqrt{\Delta}/\delta \rightarrow 0$ , it follows for  $J = |i-j| \geq 1$  that

$$\begin{aligned} \text{Cov}(\tilde{D}_{ki}, \tilde{D}_{kj})^2 &= \frac{\sigma^4}{\pi\vartheta_2} d(J)^2 e^{-\kappa\delta} \cdot \Delta + o\left(\frac{\Delta}{J^{3/2}}\right) + \mathcal{O}(\Delta^3), \\ \text{Cov}(\tilde{D}_{ki}, \tilde{D}_{(k+1)j})^2 &= \frac{\sigma^4}{4\pi\vartheta_2} d(J)^2 e^{-\kappa\delta} \cdot \Delta + o\left(\frac{\Delta}{J^{3/2}}\right) + \mathcal{O}(\Delta^3). \end{aligned}$$

Note that the  $\mathcal{O}(\Delta^3)$ -term is negligible for the asymptotic variance since

$$N^2 M \Delta^3 = MN\Delta \cdot N\Delta^2 = MN\Delta \cdot \frac{T}{M} \cdot M\sqrt{\Delta} \cdot \sqrt{\Delta} = o(NM\Delta).$$

The remaining covariances do not contribute to the asymptotic variance since for  $|k-l| \geq 2$  we have

$$\text{Cov}(\tilde{D}_{ki}, \tilde{D}_{lj})^2 = \mathcal{O}\left(\frac{\Delta\delta^4}{(J+1)^3}\right) + \mathcal{O}\left(\frac{\Delta^2}{(J+1)^2} \frac{M^2}{(k-l)^2}\right).$$

The claim is now proved by inserting  $\Phi_{\vartheta}^2(\delta, \Delta) = \frac{4}{\pi\vartheta_2} e^{-\kappa\delta} \Delta + o(\Delta)$  and noting that we have

$$\frac{1}{N} \sum_{\substack{i,j=0 \\ i \neq j}}^{N-1} d(|i-j|)^2 = \frac{2}{N} \sum_{i=1}^{N-1} \sum_{j=1}^i d(j)^2 \longrightarrow 2 \sum_{j \geq 1} d(j)^2, \quad N \rightarrow \infty$$

by Cesàro summation.

Case  $\delta/\sqrt{\Delta} \equiv r \in (0, \infty)$ . For  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  define

$$\begin{aligned} D_x^2 f(x, y) &:= f(x+2, y) + f(x, y) - 2f(x+1, y), \\ D_y^2 f(x, y) &:= f(x, y+2) + f(x, y) - 2f(x, y+1). \end{aligned}$$

We show that the asymptotic variance is given by  $C(r/\sqrt{\vartheta_2})\sigma^4$  where

$$C(h) := \frac{2}{\Lambda_{0,0}^2(h)} \sum_{j,l \in \mathbb{Z}} \Lambda_{j,l}^2(h), \quad \Lambda_{j,l}(h) := (D_x^2 D_y^2 G_h)(|j|-1, |l|-1) \quad (26)$$

$$\text{with } G_h(j, l) := \sqrt{|j|} P\left(\frac{h|l|}{\sqrt{|j|}}\right) \mathbf{1}_{\{j \neq 0\}} \quad (27)$$

and  $P(x) := \frac{1}{2\sqrt{\pi}} \left( \exp\left(-\frac{x^2}{4}\right) - x \int_{x/2}^{\infty} e^{-z^2} dz \right)$ : Define

$$\xi_{i-j,k-l}^{\Delta} := \begin{cases} 2D_{\delta}F|_{i-j,\Delta}(0), & l = k, \\ D_{\delta}^2F|_{i-j,\Delta}(|k-l-1|\delta), & l \neq k, \end{cases}$$

with  $\delta = r\sqrt{\Delta}$  such that Lemma A.1 reads as

$$\text{Cov}(\tilde{D}_{ik}, \tilde{D}_{ik}) = -\sigma^2 e^{-\kappa\delta/2} \xi_{i-j,k-l}^{\Delta} + \mathcal{O}\left(\frac{\Delta^{3/2}}{(J+1)^{3/2}}\right). \tag{28}$$

Since each term  $\xi_{J,L}^{\Delta}$  is a Riemann sum multiplied by  $\sqrt{\Delta}$ , we have for  $J, L \geq 0$

$$\lim_{\Delta \rightarrow 0} \Delta^{-1/2} \xi_{J,L}^{\Delta} = - \begin{cases} 2(\Psi_r(J, 1) - \Psi_r(J, 0)), & L = 0, \\ \Psi_r(J, L-1) + \Psi_r(J, L+1) - 2\Psi_r(J, L), & L \geq 1, \end{cases}$$

where

$$\Psi_r(J, L) := \begin{cases} \int_0^{\infty} \frac{1 - e^{-\pi^2 \vartheta_2 z^2}}{\pi^2 \vartheta_2 z^2} \cos(\pi r L z) dz, & J = 0, \\ \int_0^{\infty} \frac{2e^{-J\pi^2 \vartheta_2 z^2} - e^{-(J+1)\pi^2 \vartheta_2 z^2} - e^{-(J-1)\pi^2 \vartheta_2 z^2}}{2\pi^2 \vartheta_2 z^2} \cos(\pi r L z) dz, & J \geq 1. \end{cases}$$

By symmetry of the cosine,

$$\lim_{M, N \rightarrow \infty} \Delta^{-1/2} \xi_{J,L} = - \left( \Psi_r(J, |L| - 1) + \Psi_r(J, |L| + 1) - 2\Psi_r(J, |L|) \right)$$

also holds for negative  $L$ . Hence, we can write for all  $L \in \mathbb{Z}$  and  $J \geq 0$  and with  $G$  from (27)

$$\begin{aligned} \Psi_r(J, L) &= \int_0^{\infty} \frac{2e^{-J\pi^2 \vartheta_2 z^2} - e^{-(J+1)\pi^2 \vartheta_2 z^2} - e^{-|J-1|\pi^2 \vartheta_2 z^2}}{2\pi^2 \vartheta_2 z^2} \cos(\pi r L z) dz \\ &= \left( G_{r/\sqrt{\vartheta_2}}(J+1, L) + G_{r/\sqrt{\vartheta_2}}(J-1, L) - 2G_{r/\sqrt{\vartheta_2}}(J, L) \right) / \sqrt{\vartheta_2}, \end{aligned}$$

where the last equality follows from

$$\frac{G_{r/\sqrt{\vartheta_2}}(j, l)}{\sqrt{\vartheta_2}} = \int_0^{\infty} \frac{1 - e^{-|j|\pi^2 \vartheta_2 z^2}}{2\pi^2 \vartheta_2 z^2} \cos(\pi r l z) dz, \quad j, l \in \mathbb{Z},$$

which may be shown analogously to the calculation of  $\psi_{\vartheta_2}(r)$ . Consequently, for all  $J \in \{1 - N, \dots, N - 1\}$  and  $L \in \{1 - M, \dots, M - 1\}$  we have

$$\lim_{M, N \rightarrow \infty} \Delta^{-1/2} \xi_{J,L} = -\Lambda_{J,L}(r/\sqrt{\vartheta_2})/\sqrt{\vartheta_2}.$$



The usual Riemann sum argument yields  $F_{J,\Delta}(0) \lesssim \frac{\sqrt{\Delta}}{(J+1)^{3/2}} \lesssim \frac{\sqrt{\Delta}}{(J+1)}$  for  $J \geq 0$  and Lemma A.3 (more precisely (45)) yields  $F_{J,\Delta}(L\delta) \lesssim \frac{\Delta}{(J+1)L\delta} \lesssim \frac{\sqrt{\Delta}}{(J+1)(L+1)}$  for  $J \in \mathbb{N}_0$  and  $L \geq 1$ . We obtain

$$\Delta^{-1/2} \xi_{J,L}^\Delta = \mathcal{O} \left( \frac{1}{(|J|+1)(|L|+1)} \right), \quad J, L \in \mathbb{Z}. \tag{29}$$

Therefore,

$$\text{Var} \left( \frac{1}{\sqrt{NM\Delta}} \sum_{i=0}^{N-1} \sum_{k=0}^{M-1} \tilde{D}_{ik}^2 \right) = \frac{2\sigma^4}{NM\Delta} \sum_{i,j=0}^{N-1} \sum_{k,l=0}^{M-1} (\xi_{i-j,k-l}^\Delta)^2 + o(1).$$

By dominated convergence and taking Cesàro limits twice, we conclude

$$\begin{aligned} & \lim_{M,N \rightarrow \infty} \text{Var} \left( \frac{1}{\sqrt{NM\Delta}} \sum_{i=0}^{N-1} \sum_{k=0}^{M-1} \tilde{D}_{ik}^2 \right) \\ &= \lim_{M,N \rightarrow \infty} \frac{2\sigma^4}{\vartheta_2 NM} \sum_{i,j=0}^{N-1} \sum_{k,l=0}^{M-1} \Lambda_{i-j,k-l}^2(r/\sqrt{\vartheta_2}) = \frac{2\sigma^4}{\vartheta_2} \sum_{i,k \in \mathbb{Z}} \Lambda_{i,k}^2(r/\sqrt{\vartheta_2}). \end{aligned}$$

Since  $\psi_{\vartheta_2}(r) = -\Lambda_{0,0}(r/\sqrt{\vartheta_2})/\sqrt{\vartheta_2}$ , we have  $\Phi_\vartheta^2(\delta, \Delta) = e^{-\kappa\delta} \Lambda_{0,0}^2(r/\sqrt{\vartheta_2})/\vartheta_2 \cdot \Delta + o(\Delta)$  and dividing by  $\lim_{M,N \rightarrow \infty} \Delta^{-1} \Phi_\vartheta^2(\delta, \Delta) = \Lambda_{0,0}^2(r/\sqrt{\vartheta_2})/\vartheta_2$  yields the claimed asymptotic variance.  $\square$

**7.2. Proofs for the estimators**

Propositions 4.1 and 4.2 follow immediately from the central limit theorems for the realized quadratic variations and the delta method. Before proving Theorem 4.3, we introduce some notation that will be used throughout the proof and we state the asymptotic covariance matrix explicitly. Recall the definition of  $\Lambda_{i,k}(\cdot)$  from (26) and for any  $i, k \in \mathbb{Z}$  let

$$\begin{aligned} A_{ik}^r &:= -\Lambda_{ik}(r/\sqrt{\vartheta_2})/\sqrt{\vartheta_2}, & A_r &:= \sum_{i,k \in \mathbb{Z}} (A_{ik}^r)^2, \\ B_{ik}^r &:= 2A_{ik}^r + A_{(i-1)k}^r + A_{(i+1)k}^r, & B_r &:= \sum_{i,k \in \mathbb{Z}} (B_{ik}^r)^2, \\ C_{ik}^r &:= A_{ik}^r + A_{(i-1)k}^r, & C_r &:= \sum_{i,k \in \mathbb{Z}} (C_{ik}^r)^2. \end{aligned}$$

In terms of

$$\begin{aligned} H(x) &:= \frac{4x}{\sqrt{\pi}} \left( 1 - e^{-x^2} + 2x \int_x^\infty e^{-z^2} dz \right), \\ H'(x) &= \frac{4}{\sqrt{\pi}} \left( 1 - e^{-x^2} + 4x \int_x^\infty e^{-z^2} dz \right), \quad x \geq 0, \end{aligned}$$

we have  $\psi_{\vartheta_2}(r) = \frac{1}{r}H\left(\frac{r}{2\sqrt{\vartheta_2}}\right)$  and  $\frac{\partial}{\partial \vartheta_2}\psi_{\vartheta_2}(r) = -H'\left(\frac{r}{2\sqrt{\vartheta_2}}\right)\frac{1}{4\vartheta_2^{3/2}}$ . Denoting  $r_i := r/\sqrt{i}$ , let  $h_\eta^i(z) := e^{-\kappa z}g_\eta^i(z)$  where

$$g_\eta^i(z) := e^{-\kappa z} \left( \frac{1}{r_i} H\left(\frac{r_i}{2\sqrt{\vartheta_2}}\right), -\frac{\sigma^2}{4\vartheta_2^{3/2}} H'\left(\frac{r_i}{2\sqrt{\vartheta_2}}\right), -z \frac{\sigma^2}{r_i} H\left(\frac{r_i}{2\sqrt{\vartheta_2}}\right) \right)^\top$$

is the gradient of  $\eta \mapsto f_\eta^i(z)$  with  $i = 1, 2$  and  $z \in [b, 1 - b]$ . Moreover, we write  $\langle f, g \rangle_b := \frac{1}{1-2b} \int_b^{1-b} f(x)g(x)dx$  for  $f, g \in L^2([b, 1 - b])$ . We will prove that the asymptotic covariance matrix equals

$$\Omega_\eta^r := V^{-1}UV^{-1}, \tag{30}$$

where  $U = U(\eta)$  and  $V = V(\eta)$  are defined via

$$\begin{aligned} U_{ij} &:= 4\sigma^4 \left( 2A_r \langle (h_\eta^1)_i, (h_\eta^1)_j \rangle_b + B_r \langle (h_\eta^2)_i, (h_\eta^2)_j \rangle_b \right. \\ &\quad \left. + \sqrt{2}C_r \left( \langle (h_\eta^1)_i, (h_\eta^2)_j \rangle_b + \langle (h_\eta^2)_i, (h_\eta^1)_j \rangle_b \right) \right), \\ V_{ij} &:= 2 \left( \langle (g_\eta^1)_i, (g_\eta^1)_j \rangle_b + \langle (g_\eta^2)_i, (g_\eta^2)_j \rangle_b \right), \quad i, j \in \{1, 2, 3\}. \end{aligned}$$

*Proof of Theorem 4.3.* The proof uses the classical theory on minimum contrast estimators, see e.g. [11]. In particular, the mean value theorem yields

$$-\dot{K}_{N,M}(\eta) = \dot{K}_{N,M}(\hat{\eta}) - \dot{K}_{N,M}(\eta) = \left( \int_0^1 \ddot{K}_{N,M}(\eta + \tau(\hat{\eta} - \eta)) d\tau \right) (\hat{\eta} - \eta)$$

as soon as  $[\hat{\eta}, \eta] \subset H$ , where  $\dot{K}_{N,M}$  and  $\ddot{K}_{N,M}$  denote gradient and Hessian with respect to  $\eta$ , respectively. In the sequel, we will verify that  $K_{N,M}$  is associated with the contrast function  $K(\eta, \tilde{\eta}) := K^1(\eta, \tilde{\eta}) + K^2(\eta, \tilde{\eta})$ , where

$$K^i(\eta, \tilde{\eta}) := \frac{1}{1-2b} \int_b^{1-b} (f_\eta^i(z) - f_{\tilde{\eta}}^i(z))^2 dz$$

(Steps 1-2), show consistency of  $\hat{\eta}$  (Step 3), prove asymptotic normality of  $\dot{K}_{N,M}(\eta)$  with covariance matrix  $U$  (Steps 4-7) and deduce stochastic convergence of  $\int_0^1 \ddot{K}_{N,M}(\eta + \tau(\hat{\eta} - \eta)) d\tau$  to the invertable matrix  $V$  (Steps 8-9). The result then follows from Slutsky's Lemma and  $-\sqrt{MN}V(\eta)^{-1}\dot{K}_{N,M}(\eta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Omega_\eta^r)$ .

*Step 1.* We show that  $K$  is a contrast function in the sense that for each  $\eta$  the function  $\tilde{\eta} \mapsto K(\eta, \tilde{\eta})$  attains its unique minimum in  $\tilde{\eta} = \eta$ . Since  $f_\eta^i(\cdot)$  is continuous it is sufficient to show that  $(f_\eta^1, f_\eta^2) = (f_{\tilde{\eta}}^1, f_{\tilde{\eta}}^2)$  if and only if  $\eta = \tilde{\eta}$ . Clearly,  $(f_\eta^1, f_\eta^2) = (f_{\tilde{\eta}}^1, f_{\tilde{\eta}}^2)$  holds if and only if  $\kappa = \tilde{\kappa}$  and  $\sigma^2\psi_{\vartheta_2}(r_i) = \tilde{\sigma}^2\psi_{\tilde{\vartheta}_2}(r_i)$  for  $i = 1, 2$ . Therefore, in order to prove identifiability, it is sufficient to show that  $\vartheta_2 \mapsto \psi_{\vartheta_2}(r_1)/\psi_{\vartheta_2}(r_2)$  is injective, which in turn is implied by strict monotonicity of  $H(r_1z)/H(r_2z)$  in  $z > 0$ . We show that the corresponding derivative or, equivalently, the function  $z \mapsto H'(r_1z)H(r_2z)r_1 - H'(r_2z)H(r_1z)r_2$ , is strictly

negative for all  $z > 0$ : For  $x > 0$  define  $p(x) = \int_x^\infty e^{-z^2} dz$  and  $q(x) = 1 - e^{-x^2}$ . A simple calculation shows that

$$\begin{aligned} &H'(r_1z)H(r_2z)r_1 - H'(r_2z)H(r_1z)r_2 \\ &= \frac{32}{\pi}r_1r_2z \left( p(r_1z)q(r_2z)r_1z - p(r_2z)q(r_1z)r_2z \right) \end{aligned}$$

which is strictly negative if we can show that  $p(b)q(a)b - p(a)q(b)a < 0$  for all  $0 < a < b$ . Now, a substitution yields  $p(x) = x \int_1^\infty e^{-x^2t^2} dt$  and  $q(x) = 2x^2 \int_0^1 se^{-x^2s^2} ds$  and therefore,

$$p(b)q(a)b - p(a)q(b)a = 2a^2b^2 \int_0^1 \int_1^\infty s \left( e^{-b^2t^2 - a^2s^2} - e^{-a^2t^2 - b^2s^2} \right) dt ds < 0$$

follows from negativity of the integrand.

*Step 2.*  $K$  is the contrast function associated with the process  $K_{N,M}$  in the sense that  $K_{N,M}(\tilde{\eta}) \xrightarrow{\mathbf{P}^{\eta}} K(\eta, \tilde{\eta})$ ,  $N, M \rightarrow \infty$ , for all  $\tilde{\eta} \in H$ : Recall from the proof of Theorem 3.7 that for  $i, j, k, l \in \mathbb{N}$  we have

$$\text{Cov}(D_{ik}, D_{jl}) = \sigma^2 e^{-\kappa \frac{z_k+z_l}{2}} \xi_{i-j, k-l}^\Delta + \mathcal{O} \left( \frac{\Delta^{3/2}}{(|i-j|+1)^{3/2}} \right), \tag{31}$$

$$\xi_{i,k}^\Delta = \mathcal{O} \left( \frac{\sqrt{\Delta}}{(|i|+1)(|k|+1)} \right) \tag{32}$$

and  $\lim_{N,M \rightarrow \infty} \Delta^{-1/2} \xi_{i-j, k-l}^\Delta = A_{ik}^r = -\Lambda_{ik}(r/\sqrt{\vartheta_2})/\sqrt{\vartheta_2}$ . Now, in terms of

$$r_{ik}(\eta) = D_{ik}^2/\sqrt{\Delta} - f_\eta^1(z_k), \quad R_k(\eta) = \frac{1}{N} \sum_{i=0}^{N-1} r_{ik}(\eta)$$

we can write

$$\begin{aligned} K_{N,M}^1(\tilde{\eta}) &= \frac{1}{M} \sum_{k=0}^{M-1} (f_\eta^1(z_k) - f_{\tilde{\eta}}^1(z_k))^2 \\ &\quad + \frac{2}{M} \sum_{k=0}^{M-1} R_k(\eta) (f_\eta^1(z_k) - f_{\tilde{\eta}}^1(z_k)) + \frac{1}{M} \sum_{k=0}^{M-1} R_k^2(\eta). \end{aligned} \tag{33}$$

Clearly, the first summand converges to  $K^1(\eta, \tilde{\eta})$ . To prove that the other two summands are negligible, note that

$$\begin{aligned} &\mathbf{E}(r_{ik}r_{jl}) \\ &= \mathbf{E} \left( (D_{ik}^2/\sqrt{\Delta} - \mathbf{E}(D_{ik}^2/\sqrt{\Delta}) + \mathcal{O}(\Delta))(D_{jl}^2/\sqrt{\Delta} - \mathbf{E}(D_{jl}^2/\sqrt{\Delta}) + \mathcal{O}(\Delta)) \right) \\ &= \frac{1}{\Delta} \text{Cov}(D_{ik}^2, D_{jl}^2) + \mathcal{O}(\Delta^2) = \frac{2}{\Delta} \text{Cov}(D_{ik}, D_{jl})^2 + \mathcal{O}(\Delta^2) \end{aligned}$$

$$= \mathcal{O}\left(\frac{1}{(|i-j|+1)^2(|k-l|+1)^2}\right) + \mathcal{O}(\Delta^2).$$

By Markov's inequality and boundedness of  $\phi(\cdot) = f_\eta^1(\cdot) - f_{\hat{\eta}}^1(\cdot)$ , we have for any  $\varepsilon > 0$ ,

$$\begin{aligned} \mathbf{P}\left(\left|\frac{1}{M}\sum_{k=0}^{M-1}R_k\phi(z_k)\right|\geq\varepsilon\right) &\leq\frac{1}{\varepsilon^2M^2}\sum_{k,l=0}^{M-1}|\mathbf{E}(R_kR_l)\phi(z_k)\phi(z_l)| \\ &\lesssim\frac{1}{M^2}\sum_{k,l=0}^{M-1}|\mathbf{E}(R_kR_l)|\leq\frac{1}{M^2N^2}\sum_{k,l=0}^{M-1}\sum_{i,j=0}^{N-1}|\mathbf{E}(r_{ik}r_{jl})|=o(1), \end{aligned}$$

hence, the second summand in (33) converges to zero in probability. For the third summand the same conclusion holds since

$$\mathbf{E}\left(\frac{1}{M}\sum_{k=0}^{M-1}R_k^2\right)=\frac{1}{M}\sum_{k=0}^{M-1}\mathbf{E}(R_k^2)=\frac{1}{MN^2}\sum_{k=0}^{M-1}\sum_{i,j=0}^{N-1}\mathbf{E}(r_{ik}r_{jk})=o(1)$$

and  $L^1$ -convergence implies convergence in probability.  $K_{N,M}^2$  can be handled similarly by considering a decomposition into two sums of non-overlapping increments:

$$\bar{R}_k(\eta)=2\left(\frac{1}{2N}\sum_{\substack{i\leq N-1 \\ i\text{ even}}}\bar{r}_{ik}(\eta)+\frac{1}{2N}\sum_{\substack{i\leq N-1 \\ i\text{ odd}}}\bar{r}_{ik}(\eta)\right)$$

where  $\bar{r}_{ik} = \bar{D}_{ik}^2/\sqrt{2\Delta} - f_\eta^2(z_k)$ .

Step 3. Consistency of  $\hat{\eta}$  follows from uniform convergence in probability of the contrast process. Since  $K_{N,M}$  and  $K$  are continuous, this in turn follows from

$$\forall\varepsilon>0:\lim_{h\rightarrow 0}\limsup_{M,N\rightarrow\infty}\mathbf{P}_\eta\left(\sup_{|\eta_1-\eta_2|<h}|K_{N,M}(\eta_1)-K_{N,M}(\eta_2)|\geq\varepsilon\right)=0:$$

By compactness of the parameter space, for each  $a > 0$  there exists  $h > 0$  such that  $\|f_{\eta_1}^i - f_{\eta_2}^i\|_\infty, \|(f_{\eta_1}^i)^2 - (f_{\eta_2}^i)^2\|_\infty \leq a$  for all  $|\eta_1 - \eta_2| < h$ . Therefore,

$$\begin{aligned} |K_{N,M}^1(\eta_1)-K_{N,M}^1(\eta_2)| &\leq\frac{2}{M}\sum_{k=0}^{M-1}\left(\frac{1}{N\sqrt{\Delta}}\sum_{i=0}^{N-1}D_{ik}^2\right)|f_{\eta_2}^1(z_k)-f_{\eta_1}^1(z_k)| \\ &\quad +\frac{1}{M}\sum_{k=0}^{M-1}|f_{\eta_1}^1(z_k)^2-f_{\eta_2}^1(z_k)^2| \\ &\leq a\left(\frac{2}{M}\sum_{k=0}^{M-1}\left(\frac{1}{N\sqrt{\Delta}}\sum_{i=0}^{N-1}D_{ik}^2\right)+1\right) \end{aligned}$$

and, hence,

$$\limsup_{M,N\rightarrow\infty}\mathbf{P}_\eta\left(\sup_{|\eta_1-\eta_2|<h}|K_{N,M}^1(\eta_1)-K_{N,M}^1(\eta_2)|\geq\varepsilon\right)$$

$$\begin{aligned} &\leq \limsup_{M,N \rightarrow \infty} \frac{1}{\varepsilon} \mathbf{E} \left( \sup_{|\eta_1 - \eta_2| < h} |K_{N,M}^1(\eta_1) - K_{N,M}^1(\eta_2)| \right) \\ &\leq \limsup_{M,N \rightarrow \infty} \frac{a}{\varepsilon} \mathbf{E} \left( \frac{2}{M} \sum_{k=0}^{M-1} \left( \frac{1}{N\sqrt{\Delta}} \sum_{i=0}^{N-1} D_{ik}^2 \right) + 1 \right) \lesssim \frac{a}{\varepsilon}. \end{aligned}$$

The same argument applies to  $K_{N,M}^2$  and the result follows.

*Step 4.* Let  $F_1, F_2 \in C^1([0, 1])$  and  $(a_k)_{k \in \mathbb{Z}}$  be absolutely summable. Then we can write

$$\begin{aligned} \frac{1}{n} \sum_{k,l=0}^{n-1} a_{k-l} F_1(z_k) F_2(z_l) &= \frac{a_0}{n} (F_1(z_0) F_2(z_0) + \dots + F_1(z_{n-1}) F_2(z_{n-1})) \\ &\quad + \frac{a_1}{n} (F_1(z_1) F_2(z_0) + \dots + F_1(z_{n-1}) F_2(z_{n-2})) \\ &\quad + \frac{a_{-1}}{n} (F_1(z_0) F_2(z_1) + \dots + F_1(z_{n-2}) F_2(z_{n-1})) \\ &\quad + \dots \end{aligned}$$

and, consequently, we have  $\frac{1}{n} \sum_{k,l=0}^{n-1} a_{k-l} F_1(z_k) F_2(z_l) \rightarrow \langle F_1, F_2 \rangle_b \cdot \sum_{k \in \mathbb{Z}} a_k$  for  $n \rightarrow \infty$  by dominated convergence.

*Step 5.* We show that the asymptotic covariance matrix of  $\sqrt{NM} \dot{K}_{N,M}(\eta)$  is given by  $U$ : We have  $\dot{K}_{N,M}(\eta) = \dot{K}_{N,M}^1(\eta) + \dot{K}_{N,M}^2(\eta)$  as well as

$$\dot{K}_{N,M}^1(\eta) = -\frac{2}{M} \sum_{k=0}^{M-1} \left( \frac{1}{N\sqrt{\Delta}} \sum_{i=0}^{N-1} D_{ik}^2 - f_\eta^1(z_k) \right) g_\eta^1(z_k)$$

and similarly for  $\dot{K}_{N,M}^2(\eta)$ . From Isserlis' theorem, (31) and  $\bar{D}_{ik} = D_{ik} + D_{(i+1)k}$  it follows that

$$\begin{aligned} \text{Cov}(D_{ik}^2, D_{jl}^2) &= 2 \left( \sigma^2 e^{-\frac{z_k+z_l}{2}} \xi_{i-j,k-l}^\Delta + \mathcal{O} \left( \frac{\Delta^{3/2}}{(|i-j|+1)^{3/2}} \right) \right)^2, \\ \text{Cov}(\bar{D}_{ik}^2, \bar{D}_{jl}^2) &= 2 \left( \sigma^2 e^{-\frac{z_k+z_l}{2}} (2\xi_{i-j,k-l}^\Delta + \xi_{i-j-1,k-l}^\Delta + \xi_{i-j+1,k-l}^\Delta) \right. \\ &\quad \left. + \mathcal{O} \left( \frac{\Delta^{3/2}}{(|i-j|+1)^{3/2}} \right) \right)^2, \\ \text{Cov}(D_{ik}^2, \bar{D}_{jl}^2) &= 2 \left( \sigma^2 e^{-\frac{z_k+z_l}{2}} (\xi_{i-j,k-l}^\Delta + \xi_{i-j-1,k-l}^\Delta) + \mathcal{O} \left( \frac{\Delta^{3/2}}{(|i-j|+1)^{3/2}} \right) \right)^2. \end{aligned}$$

Now, for any  $1 \leq e, f \leq 3$ , the first summand in the expansion

$$\begin{aligned} &\text{Cov}((\dot{K}_{N,M})_e, (\dot{K}_{N,M})_f) \\ &= \text{Cov}((\dot{K}_{N,M}^1)_e, (\dot{K}_{N,M}^1)_f) + \text{Cov}((\dot{K}_{N,M}^2)_e, (\dot{K}_{N,M}^2)_f) \\ &\quad + \text{Cov}((\dot{K}_{N,M}^1)_e, (\dot{K}_{N,M}^2)_f) + \text{Cov}((\dot{K}_{N,M}^2)_e, (\dot{K}_{N,M}^1)_f) \end{aligned} \tag{34}$$

is given by

$$\text{Cov}((\dot{K}_{N,M}^1)_e, (\dot{K}_{N,M}^1)_f) = \frac{4}{M^2 N^2 \Delta} \sum_{i,j=0}^{N-1} \sum_{k,l=0}^{M-1} \text{Cov}(D_{ik}^2, D_{jl}^2) (g_\eta^1)_e(z_k) (g_\eta^1)_f(z_l).$$

Like in the proof of Theorem 3.7, the covariances may be replaced by their asymptotic expressions due to dominated convergence. Further, using  $(h_\eta^i)_e(z) = e^{-\kappa z} (g_\eta^i)_e(z)$  and Step 4, we have

$$MN \cdot \text{Cov}((\dot{K}_{N,M}^1)_e, (\dot{K}_{N,M}^1)_f) \rightarrow 8\sigma^4 \sum_{i,k \in \mathbb{Z}} (A_{i,k}^r)^2 \cdot \langle (h_\eta^1)_e, (h_\eta^1)_f \rangle_b$$

for  $M, N \rightarrow \infty$ . Analogously,

$$MN \cdot \text{Cov}((\dot{K}_{N,M}^2)_e, (\dot{K}_{N,M}^2)_f) \rightarrow 4\sigma^4 \sum_{i,k \in \mathbb{Z}} (B_{i,k}^r)^2 \cdot \langle (h_\eta^2)_e, (h_\eta^2)_f \rangle_b,$$

$$MN \cdot \text{Cov}((\dot{K}_{N,M}^1)_e, (\dot{K}_{N,M}^2)_f) \rightarrow 4\sqrt{2}\sigma^4 \sum_{i,k \in \mathbb{Z}} (C_{i,k}^r)^2 \cdot \langle (h_\eta^1)_e, (h_\eta^2)_f \rangle_b$$

and insertion into (34) yields the claimed asymptotic covariance matrix.

*Step 6.*  $U$  is strictly positive definite: It is sufficient to show that  $C_r < \sqrt{A_r B_r}$ , then it follows for any  $\alpha \in \mathbb{R}^3 \setminus \{0\}$  and  $H_\alpha^i = \sum_{j=1}^3 \alpha_j (h_\eta^i)_j$ ,  $i = 1, 2$  that

$$\begin{aligned} \alpha^\top U \alpha &= 4\sigma^4 \left( 2A_r \|H_\alpha^1\|_b^2 + B_r \|H_\alpha^2\|_b^2 + 2\sqrt{2}C_r \langle H_\alpha^1, H_\alpha^2 \rangle_b \right) \\ &> 4\sigma^4 \left( 2A_r \|H_\alpha^1\|_b^2 + B_r \|H_\alpha^2\|_b^2 + 2\sqrt{2A_r B_r} \langle H_\alpha^1, H_\alpha^2 \rangle_b \right) \\ &= 8\sigma^4 \left\| \sqrt{2A_r} H_\alpha^1 + \sqrt{B_r} H_\alpha^2 \right\|_b^2 \geq 0, \end{aligned}$$

where we may assume  $\langle H_\alpha^1, H_\alpha^2 \rangle_b < 0$  since otherwise  $\alpha^\top U \alpha > 0$  follows immediately from the first equality. Now, consider  $(A_{i,k}^r)$  and  $(B_{i,k}^r)$  as elements in the Hilbert space  $\ell^2$  of square summable sequences indexed by  $\mathbb{Z} \times \mathbb{Z}$ . Clearly,  $A_r = \|(A_{i,k}^r)\|_{\ell^2}^2$ ,  $B_r = \|(B_{i,k}^r)\|_{\ell^2}^2$  and a direct calculation shows that  $C_r = \langle (A_{i,k}^r), (B_{i,k}^r) \rangle_{\ell^2}$ . Thus, by the Cauchy-Schwarz inequality we have  $C_r \leq \sqrt{A_r B_r}$  and equality is ruled out by the fact that  $(A_{i,k}^r)$  and  $(B_{i,k}^r)$  are not linearly dependent.

*Step 7.* We show  $\sqrt{NM} \dot{K}_{N,M}^1(\eta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, U)$  under  $\mathbf{P}_\eta$ . In view of the Cramér-Wold device, we have to prove  $\sqrt{NM} \alpha^\top \dot{K}_{N,M} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \alpha^\top U \alpha)$  for any  $\alpha \in \mathbb{R}^3$ . Let  $s_{ik}$  and  $Z_{ik}$  be given by the relation  $s_{ik} Z_{ik}^2 = -\frac{2\alpha^\top f_\eta^1(z_k)}{\sqrt{NM\Delta}} D_{ik}^2$  where  $s_{ik} \in \{-1, 1\}$  is deterministic. Analogously, define  $\bar{s}_{ik}$  and  $\bar{Z}_{ik}^2$ . Then,  $\mathcal{Z}_{N,M} = (Z_{ik}, \bar{Z}_{j,l})_{i,j,k,l}$  is a Gaussian vector and from Proposition 3.5 it follows that

$$\sqrt{NM} \alpha^\top \dot{K}_{N,M}(\eta) = S_{N,M} - \mathbf{E}(S_{N,M}) + o(1)$$

where  $S_{N,M} = \sum_{i=0}^{N-1} \sum_{k=0}^{M-1} s_{ik} Z_{ik}^2 + \sum_{i=0}^{N-1} \sum_{k=0}^{M-2} \bar{s}_{ik} \bar{Z}_{ik}^2$ . From Steps 5 and 6 we can deduce that  $\text{Var}(S_{N,M}) \rightarrow \alpha^\top U \alpha > 0$ ,  $N, M \rightarrow \infty$  and thus, in

view of criterion (8), asymptotic normality follows if the absolute row sums of the covariance matrix of  $\mathcal{Z}_{N,M}$  vanish uniformly. This in turn is a simple consequence of (31) and (32).

*Step 8.* In order to prove  $\int_0^1 \ddot{K}_{N,M}(\eta + \tau(\hat{\eta} - \eta)) d\tau \xrightarrow{\mathbf{P}_\eta} V(\eta)$ , it suffices to show  $\ddot{K}_{N,M}(\eta_{N,M}) \xrightarrow{\mathbf{P}_\eta} V(\eta)$  for any consistent estimator  $\eta_{N,M}$  of  $\eta$ . We have

$$\ddot{K}_{N,M}^1(\eta) = \frac{2}{M} \sum_{k=0}^{M-1} g_\eta^1(z_k) g_\eta^1(z_k)^\top - \frac{2}{M} \sum_{k=0}^{M-1} \left( \frac{1}{N\sqrt{\Delta}} \sum_{i=0}^{N-1} D_{ik}^2 - f_\eta^1(z_k) \right) \ddot{f}_\eta^1(z_k)$$

and analogously for  $\ddot{K}_{N,M}^2$ . By using  $\mathbf{P}_\eta(\eta_{N,M} \in H) \rightarrow 1$  and the uniform continuity of  $f_\eta^i(z)$  and its derivatives in the parameter  $(z, \eta) \in [0, 1] \times H$ , it is straightforward to show  $\ddot{K}_{N,M}(\eta_{N,M}) - \ddot{K}_{N,M}(\eta) \xrightarrow{\mathbf{P}_\eta} 0$ . Now, write  $V = 2(V^1 + V^2)$  where  $V^i$  is the Gram matrix of the functions  $\{(g_\eta^i)_1, (g_\eta^i)_2, (g_\eta^i)_3\}$  with respect to the inner product  $\langle \cdot, \cdot \rangle_b$ , i.e.  $V_{ef}^i = \langle (g_\eta^i)_e, (g_\eta^i)_f \rangle_b$ ,  $1 \leq e, f \leq 3$ . Clearly, first summand of  $\ddot{K}_{N,M}^1(\eta)$  converges to  $2V^1$  while the calculations of Step 2 show that the second summand converges to 0 in probability. The same reasoning holds for  $\ddot{K}_{N,M}^2(\eta)$  and the result follows.

*Step 9.*  $V$  is strictly positive definite: Being Gram matrices,  $V^1$  and  $V^2$  are positive semi-definite and consequently, the same holds for  $V$ . Clearly, the only way  $V$  can be singular is if there exists  $\alpha \in \mathbb{R}^3$  such that  $0 = \alpha^\top V^i \alpha = \|\sum_{e=1}^3 \alpha_e (g_\eta^i)_e\|_b^2$  holds for both  $i \in \{1, 2\}$ . From the particular form of the functions  $(g_\eta^i)_e$  it is apparent that this would imply that  $\alpha_1 \psi_{\vartheta_2}(r_i) + \alpha_2 \sigma^2 \frac{\partial \psi_{\vartheta_2}(r_i)}{\partial \vartheta_2} = \alpha_3 = 0$  for both  $i \in \{1, 2\}$ , which is impossible.  $\square$

*Proof of Proposition 4.5.* We have to prove

$$\forall \varepsilon > 0 \exists C > 0 : \limsup_{N,M \rightarrow \infty} \mathbf{P}_\eta \left( \sqrt{M^3 \wedge N^{3/2}} \|\hat{\eta}_{v,w} - \eta\| \geq C \right) \leq \varepsilon.$$

Similar calculations as in Theorem 4.3 show that Steps 1-3 and 8-9 of the corresponding proof remain valid. Consequently, we have the representation  $-\dot{\mathcal{K}}_{N,M}(\eta) = V_{N,M}(\hat{\eta}_{v,w}, \eta)(\hat{\eta}_{v,w} - \eta)$ , where  $V_{N,M}(\tilde{\eta}, \eta) = \int_0^1 \ddot{K}_{N,M}(\eta + \tau(\tilde{\eta} - \eta)) d\tau$  as well as  $V_{N,M}(\hat{\eta}_{v,w}, \eta) \xrightarrow{\mathbf{P}_\eta} V(\eta)$  where  $V(\eta)$  is an invertible deterministic matrix. In particular, the set

$$A_{N,M} = \{V_{N,M}(\hat{\eta}_{v,w}, \eta) \text{ is invertible and } \|V_{N,M}(\hat{\eta}_{v,w}, \eta)^{-1}\|_2 \leq \|V(\eta)^{-1}\|_2 + 1\}$$

satisfies  $\mathbf{P}_\eta(A_{N,M}) \rightarrow 1$ . Further,  $\dot{\mathcal{K}}_{N,M}(\eta)$  can be written as an average of expressions of the type  $\dot{K}_{N,M}$  from Theorem 4.3 so that the calculations of Step 5 show together with the Cauchy-Schwarz inequality that  $\mathbf{E}_\eta \left( \|\dot{\mathcal{K}}_{N,M}(\eta)\|^2 \right) = \mathcal{O}((M^3 \wedge N^{3/2})^{-1})$ . Now,

$$\mathbf{P}_\eta \left( \sqrt{M^3 \wedge N^{3/2}} \|\hat{\eta}_{v,w} - \eta\| \geq C \right)$$

$$\leq \mathbf{P}_\eta \left( \left\{ \sqrt{M^3 \wedge N^{3/2}} \|\hat{\eta}_{v,w} - \eta\| \geq C \right\} \cap A_{N,M} \right) + \mathbf{P}_\eta(A_{N,M}^c).$$

The second summand becomes arbitrarily small as  $M, N \rightarrow \infty$ . For the first summand, let  $\gamma(\eta) = \|V(\eta)^{-1}\|_2 + 1$ , then it follows from Markov's inequality that

$$\begin{aligned} & \mathbf{P}_\eta \left( \left\{ \sqrt{M^3 \wedge N^{3/2}} \|\hat{\eta}_{v,w} - \eta\| \geq C \right\} \cap A_{N,M} \right) \\ &= \mathbf{P}_\eta \left( \left\{ \sqrt{M^3 \wedge N^{3/2}} \|V_{N,M}(\hat{\eta}_{v,w}, \eta)^{-1} \dot{\mathcal{K}}_{N,M}(\eta)\| \geq C \right\} \cap A_{N,M} \right) \\ &\leq \mathbf{P}_\eta \left( \left\{ \sqrt{M^3 \wedge N^{3/2}} \|\dot{\mathcal{K}}_{N,M}(\eta)\| \geq \frac{C}{\gamma(\eta)} \right\} \cap A_{N,M} \right) \\ &\leq \mathbf{P}_\eta \left( \sqrt{M^3 \wedge N^{3/2}} \|\dot{\mathcal{K}}_{N,M}(\eta)\| \geq \frac{C}{\gamma(\eta)} \right) \\ &\leq (M^3 \wedge N^{3/2}) \mathbf{E}_\eta(\|\dot{\mathcal{K}}_{N,M}(\eta)\|^2) \frac{\gamma(\eta)^2}{C^2} \lesssim \frac{1}{C^2}. \end{aligned} \quad \square$$

### 7.3. Proofs of the lower bounds

Before we prove Theorem 5.1, we verify its ingredients Proposition 5.3 and Proposition 5.5.

*Proof of Proposition 5.3.* By setting  $a = k^2$ ,  $\mu = \pi^2 \vartheta_2$  and  $\nu^2 = \frac{\sigma^2}{\pi^2 \vartheta_2}$  in Lemma A.4 and using independence of  $(u_\ell, \ell \in \mathbb{N})$ , we get the Fisher information matrix  $I$  for the parameters  $(\mu, \nu^2)$ , namely

$$\begin{aligned} I_{11} &= N \sum_{\ell=1}^M \frac{\ell^4 \Delta^2 (e^{-4\mu\ell^2\Delta} + e^{-2\mu\ell^2\Delta})}{(1 - e^{-2\mu\ell^2\Delta})^2} = N \sum_{\ell=1}^M g_{11}(\ell\sqrt{\Delta}), \\ I_{12} &= N \sum_{\ell=1}^M \frac{\ell^2 \Delta e^{-2\mu\ell^2\Delta}}{\nu^2 (1 - e^{-2\mu\ell^2\Delta})} = N \sum_{\ell=1}^M g_{12}(\ell\sqrt{\Delta}), \\ I_{22} &= \frac{(N+1)M}{2\nu^4} \end{aligned}$$

$$\text{with } g_{11}(x) := \frac{x^4(e^{-4\mu x^2} + e^{-2\mu x^2})}{(1 - e^{-2\mu x^2})^2}, \quad g_{12}(x) := \frac{x^2 e^{-2\mu x^2}}{\nu^2 (1 - e^{-2\mu x^2})}.$$

The Fisher information matrix  $J = J_{M,N}$  for the parameters  $(\sigma^2, \rho^2)$  can be computed via the change of variables formula  $J = A^\top I A$  where

$$A = \begin{pmatrix} \pi^2/\rho^2 & -\pi^2\sigma^2/\rho^4 \\ 0 & 1/\pi^2 \end{pmatrix}$$

is the Jacobian of the function transforming  $(\sigma^2, \rho^2)$  to  $(\mu, \nu^2)$ . Hence, the diagonal entries of  $J$  are given by

$$J_{11} = \frac{\pi^4}{\rho^4} I_{11}, \quad J_{22} = \frac{\pi^4 \sigma^4}{\rho^8} I_{11} - \frac{2\sigma^2}{\rho^4} I_{12} + \frac{1}{\pi^4} I_{22}.$$



If  $M\sqrt{\Delta}$  is bounded away from 0, then  $I_{11}$  can be interpreted as a Riemann sum. We obtain

$$J_{11} \approx I_{11} \approx N^{3/2} \int_0^{M\sqrt{\Delta}} g_{11}(x) dx \approx N^{3/2}.$$

On the other hand, if  $M\sqrt{\Delta} \rightarrow 0$ , it follows from Lemma A.10 and  $g_{11}(0) = \frac{1}{2\mu^2} = \frac{\rho^4}{2\pi^4\sigma^4}$ ,  $g_{12}(0) = \frac{1}{2\mu\nu^2} = \frac{1}{2\sigma^2}$  as well as  $g'_{11}(0) = g'_{12}(0) = 0$  that

$$\begin{aligned} I_{11} &= N^{3/2} (M\sqrt{\Delta}g_{11}(0) + \frac{M^2\Delta}{2}g'_{11}(0) + \mathcal{O}(M^3\Delta^{3/2})) = \frac{\rho^4}{2\pi^4\sigma^4}NM + \mathcal{O}(M^3), \\ I_{12} &= N^{3/2} (M\sqrt{\Delta}g_{12}(0) + \frac{M^2\Delta}{2}g'_{12}(0) + \mathcal{O}(M^3\Delta^{3/2})) = \frac{NM}{2\sigma^2} + \mathcal{O}(M^3), \\ I_{22} &= \frac{\pi^4}{2\rho^4}MN + \mathcal{O}(M). \end{aligned}$$

Therefore, the leading terms in  $J_{22}$  cancel and consequently,  $J_{22} = \mathcal{O}(M^3)$ .  $\square$

*Proof of Proposition 5.5.* For a discrete time, centered, stationary Gaussian process  $(Z_j)_{j \in \mathbb{Z}}$  whose covariance function depends on an unknown parameter  $\theta \in \mathbb{R}$  we denote the Fisher information of a sample  $(Z_0, \dots, Z_{n-1})$  with respect to  $\theta$  by  $I_n(Z)$ . A particularly useful result to calculate  $I_n(Z)$  for the above class of Gaussian processes is given by Whittle [39]:

$$\lim_{n \rightarrow \infty} \frac{1}{n} I_n(Z) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left( \frac{\frac{\partial}{\partial \theta} \phi_\theta(\omega)}{\phi_\theta(\omega)} \right)^2 d\omega, \quad n \rightarrow \infty, \quad (35)$$

where  $\phi(\omega) = \sum_{j \in \mathbb{Z}} \mathbf{E}[Z_0 Z_j] e^{-ij\omega}$ ,  $\omega \in [-\pi, \pi]$ , is the spectral density of  $Z$ .

Setting  $\theta = \pi^2 \vartheta_2$ , (35) cannot be directly applied to the process  $Z = \bar{U}_k$ , for  $1 \leq k \leq M-1$ , since  $\bar{U}_k$  arises from high-frequency increments of the continuous time process  $U_k$ . In this case, the spectral density  $\Phi_k^\Delta$  of  $\bar{U}_k$  hinges on  $\Delta = 1/N$  and therefore, even for large  $N$ ,  $I_N(\bar{U}_k)/N$  is not necessarily close to the asymptotic Fisher information defined in (35).

To circumvent this difficulty, consider the  $N$ -th order Fourier approximation to  $\Phi_k^\Delta$ :

$$\Phi_k^{N,\Delta}(\omega) = \sum_{j=1-N}^{N-1} \mathbf{E}[\bar{U}_k(0)\bar{U}_k(j)] e^{-ij\omega} \geq 0, \quad \omega \in [-\pi, \pi]. \quad (36)$$

Lemma A.6(i) verifies that  $\Phi_k^{N,\Delta}$  is positive. Therefore, there exists a stationary Gaussian process  $Y_k = (Y_k(j))_{j \in \mathbb{Z}}$  with spectral density  $\Phi_k^{N,\Delta}$ . Clearly,

$$(Y_k(j), \dots, Y_k(j+N-1)) \stackrel{\mathcal{D}}{=} (\bar{U}_k(0), \dots, \bar{U}_k(N-1)), \quad j \in \mathbb{N}_0,$$

and  $(Y_k(j), \dots, Y_k(j+N-1))$  is independent of  $(Y_k(h), \dots, Y_k(h+N-1))$  whenever  $|j-h| > 2N$ . Consequently, it is possible to extract  $L$  independent

copies of  $(\bar{U}_k(0), \dots, \bar{U}_k(N-1))$  from a sample  $(Y_k(0), \dots, Y_k(2NL-1))$  for any  $L \in \mathbb{N}$ . Now, using the fact that a statistic never has larger information than the data from which it is constructed (cf. [20, Theorem I.7.2]) yields

$$L \cdot I_N(\bar{U}_k) \leq I_{2NL}(Y_k). \tag{37}$$

For fixed  $\Delta = 1/N$  we can now apply Whittle’s formula (35) for  $L \rightarrow \infty$ : For each  $\varepsilon > 0$  we can choose  $L \in \mathbb{N}$  such that

$$I_{2NL}(Y_k) \leq 2NL(1 + \varepsilon)\mathcal{I}_k, \tag{38}$$

where

$$\mathcal{I}_k^{N,\Delta} := \frac{1}{4\pi} \int_{-\pi}^{\pi} S^2(\omega) d\omega, \quad S := \frac{\partial}{\partial \theta_2} \log \Phi_k^{N,\Delta}.$$

By combining (37) and (38) we get  $I_N(\bar{U}_k) \leq 2N\mathcal{I}_k$ . Proving below that uniformly in  $k = 0, \dots, M-1$

$$\mathcal{I}_k^{N,\Delta} \lesssim M^2 \Delta \log \frac{1}{M^2 \Delta}, \tag{39}$$

we obtain  $I_N(\bar{U}_k) \lesssim M^2 \log \frac{1}{M^2 \Delta}$  and the results follows by independence of the processes  $\bar{U}_1, \dots, \bar{U}_{M-1}$ .

In order to verify (39), we only have to consider the integral over  $[0, \pi]$  by symmetry. From Lemma A.6 we can deduce for  $\omega \geq k^2 \Delta$  that

$$S(\omega) \lesssim \begin{cases} \frac{M\sqrt{\Delta}}{\sqrt{\omega}}, & \omega \geq M^2 \Delta, \\ 1, & \omega \in [k^2 \Delta, M^2 \Delta], \end{cases}$$

implying  $\int_{k^2 \Delta}^{\pi} S^2(\omega) d\omega \lesssim M^2 \Delta \log \frac{1}{M^2 \Delta}$ . For  $\omega \leq k^2 \Delta$ , Lemma A.6 gives  $S(\omega) \lesssim (\frac{\omega^2}{k^4 \Delta^2} + k^2 e^{-\theta k^2}) / (\frac{\omega^2}{k^4 \Delta^2} + e^{-\theta k^2})$ . Since

$$\int_0^1 \frac{d\omega}{(\omega^2 + e^{-\theta k^2})^2} \leq \int_0^{e^{-\theta k^2/2}} \frac{1}{e^{-2\theta k^2}} d\omega + \int_{e^{-\theta k^2/2}}^1 \frac{1}{\omega^4} d\omega \lesssim \exp\left(\frac{3}{2}\theta k^2\right),$$

a substitution yields

$$\int_0^{k^2 \Delta} S^2(\omega) d\omega \lesssim k^2 \Delta \int_0^1 \left( \frac{\omega^2 + k^2 e^{-\theta k^2}}{\omega^2 + e^{-\theta k^2}} \right)^2 d\omega \lesssim M^2 \Delta. \quad \square$$

We can now conclude the main lower bound.

*Proof of Theorem 5.1.* The proof of the lower bound relies on the fact that if  $(P_\gamma)_{\gamma \in G}$  is a dominated family of distributions with a convex parameter space  $G \subset \mathbb{R}$ , then the Hellinger distance  $\mathcal{H}$  can be bounded in terms of the Fisher Information  $J$ : Let  $\nu$  be a dominating measure,  $p(\cdot, \gamma) = dP_\gamma/d\nu$  and  $g = \sqrt{p}$ . Then, as shown in [20, Theorem I.7.6], Jensen’s inequality yields

$$\mathcal{H}^2(P_\gamma, P_{\gamma+h}) = \int (g(x, \gamma) - g(x, \gamma+h))^2 \nu(dx)$$

$$\begin{aligned} &\leq h^2 \int_0^1 \int_0^1 \frac{\partial g}{\partial \gamma}(x, \gamma + sh)^2 ds \nu(dx) \\ &= \frac{h^2}{4} \int_0^1 \int \left( \frac{\partial}{\partial \theta} \log p(x, \gamma + sh) \right)^2 P_{\gamma+sh}(dx) ds = \frac{h^2}{4} \int_0^1 J(\gamma + sh) ds. \end{aligned}$$

Combining this bound of the Hellinger distance (in the setting of Theorem 5.1) with Theorem 2.2 by Tsybakov [37], it suffices that for each sampling regime there is a reparametrization  $(\gamma_1, \gamma_2)$  of  $(\sigma^2, \vartheta_2)$  such that the corresponding Fisher information satisfies  $J_{M,N}(\gamma_2) \lesssim r_{M,N}^{-2}$  locally uniformly. Inspection of the proofs of Propositions 5.3 and 5.5 shows that the bounds on the Fisher information are indeed locally uniform.

(ii) *Case  $M/\sqrt{N} \gtrsim 1$ .* For  $L \in \mathbb{N}$  define the process  $X^L$  via  $X_t^L(y) = \sum_{\ell=1}^L u_\ell(t) e_\ell(y)$ ,  $t \geq 0$ ,  $y \in [0, 1]$ , and let  $\mathcal{X}_{N,M}^L = \{X_{t_i}^L(y_k), i = 0, \dots, N-1, k = 0, \dots, M\}$  as well as  $\mathcal{X}_{N,M} = \mathcal{X}_{N,M}^\infty$ . Denoting the corresponding covariance matrices by  $\Sigma_{N,M}^L$  and  $\Sigma_{N,M}$  and using the result of [12], we can bound the total variation distance of the Gaussian distributions by  $\text{TV}(\Sigma_{N,M}, \Sigma_{N,M}^L) \leq \frac{3}{2} \|\Sigma_{N,M}^{-1/2}(\Sigma_{N,M}^L - \Sigma_{N,M})\Sigma_{N,M}^{-1/2}\|_F \leq \frac{3}{2} \|\Sigma_{N,M}^{-1/2}\|_F^2 \|\Sigma_{N,M}^L - \Sigma_{N,M}\|_F$ . Consequently, we can pick a sequence  $L_{N,M} \rightarrow \infty$  such that  $\mathcal{X}_{N,M}^{L_{N,M}}$  and  $\mathcal{X}_{N,M}$  are statistically equivalent in the sense of Le Cam and it is sufficient to derive a lower bound for  $\mathcal{X}_{N,M}^{L_{N,M}}$ , or even  $\{u_\ell(t_i), i \leq N, \ell \leq L_{N,M}\}$ . Assuming  $L_{N,M} \geq M$  without loss of generality, for this observation scheme Proposition 5.3 yields under the parametrization  $(\sigma^2/\vartheta_2, \sigma^2)$ :

$$J_{M,N}(\sigma^2) \lesssim N^{3/2} \wedge L_{N,M}^3 = N^{3/2} = r_{N,M}^{-2}.$$

*Case  $M/\sqrt{N} \rightarrow 0$ .* For  $b \in \mathbb{Q} \cap (0, 1/2)$  write  $b = p/q$  where  $p \in \mathbb{Z}$  and  $q \in \mathbb{N}$  such that  $y_k = \frac{pM+k(q-2p)}{qM}$ ,  $k \leq M$ , and consequently  $\{y_k, k = 0, \dots, M\}$  is a subset of  $\{z_k, k = 1, \dots, qM-1\}$  where  $z_k = \frac{k}{qM}$ . Now,  $qM\sqrt{\Delta} \rightarrow 0$  and since  $q^3 M^3 \log\left(\frac{1}{q^2 M^2 \Delta}\right) \lesssim M^3 \log\left(\frac{1}{M^2 \Delta}\right)$  Proposition 5.5 implies under the parametrization  $(\sigma^2/\sqrt{\vartheta_2}, \vartheta_2)$ :

$$J_{M,N}(\vartheta_2) \lesssim M^3 \log\left(\frac{1}{M^2 \Delta}\right) = r_{N,M}^{-2}.$$

(i) If  $\min(M, N)$  remains finite and  $M/\sqrt{N} \gtrsim 1$ , then  $N$  necessarily remains finite and the result follows from (ii). On the other hand, if  $M/\sqrt{N} \rightarrow 0$ , then  $M$  must remain finite. Like in the proof of (ii), extend the set of spatial locations to  $\{z_k, k < qM\}$  and consider the corresponding processes  $U_k, k = 1, \dots, qM-1$  from (20). A similar calculation as in the proof of Proposition 2.3 shows that for any  $k < qM$ , the laws of the independent continuous processes  $\{U_k(t), t \leq 1\}$  are absolutely continuous for different parameter values  $(\sigma^2, \vartheta_2)$  and  $(\tilde{\sigma}^2, \tilde{\vartheta}_2)$  as long as  $\sigma^2/\sqrt{\vartheta_2} = \tilde{\sigma}^2/\sqrt{\tilde{\vartheta}_2}$  and hence, consistent estimation of  $(\sigma^2, \vartheta_2)$  based on continuous or discrete observations is impossible: Note that the continuous spectral density of  $U_k$  is  $f_k(u) = \frac{1}{2u^2} \sum_{\ell \in \mathcal{I}_k} h_{(\sigma^2, \vartheta_2)}\left(\frac{\ell}{\sqrt{|u|}}\right)$ ,  $u \in \mathbb{R}$ , where

$h_{(\sigma^2, \vartheta_2)}$  is defined in the proof of Proposition 2.3. Now, a Riemann sum midpoint approximation, cf. Lemma A.9, shows that

$$\begin{aligned} f_k^+(u) &:= \frac{1}{2u^2} \sum_{\ell \geq 0} h_{(\sigma^2, \vartheta_2)}\left(\frac{k + 2M\ell}{\sqrt{u}}\right) \\ &= \frac{1}{2u^2} \left( \frac{\sqrt{u}}{2M} \int_{(k-M)/\sqrt{u}}^{\infty} h_{(\sigma^2, \vartheta_2)}(z) dz + \mathcal{O}\left(\frac{1}{\sqrt{u}}\right) \right), \\ f_k^-(u) &:= \frac{1}{2u^2} \sum_{\ell \geq 0} h_{(\sigma^2, \vartheta_2)}\left(\frac{2M - k + 2M\ell}{\sqrt{u}}\right) \\ &= \frac{1}{u^2} \left( \frac{\sqrt{u}}{4M} \int_{(M-k)/\sqrt{u}}^{\infty} h_{(\sigma^2, \vartheta_2)}(z) dz + \mathcal{O}\left(\frac{1}{\sqrt{u}}\right) \right). \end{aligned}$$

as  $u \rightarrow \infty$ . Since  $h_{(\sigma^2, \vartheta_2)}$  is symmetric around 0 we obtain

$$f_k(u) = f_k^+(u) + f_k^-(u) = \frac{1}{u^2} \left( \frac{\sqrt{u}}{2M} \int_0^{\infty} h_{(\sigma^2, \vartheta_2)}(z) dz + \mathcal{O}\left(\frac{1}{\sqrt{u}}\right) \right)$$

from which equivalence follows as in Proposition 2.3. □

#### 7.4. Proofs for Section 2

*Proof of Proposition 2.1.* Due to (7) and the trigonometric identity

$$\sin(\alpha) \sin(\beta) = \frac{1}{2} (\cos(\alpha - \beta) - \cos(\alpha + \beta)), \tag{40}$$

we have

$$\begin{aligned} &\text{Cov}(X_t(x), X_t(y)) \\ &= \frac{\sigma^2}{2\pi^2\vartheta_2} e^{-\frac{\kappa}{2}(x+y)} \sum_{\ell \geq 1} \frac{1}{\ell^2 + \Gamma/\pi^2} (\cos(\pi\ell(y-x)) - \cos(\pi\ell(x+y))). \end{aligned}$$

The claimed formulas now follow by inserting the closed expressions

$$\sum_{\ell \geq 1} \frac{1}{\ell^2 + \beta} \cos(\pi\ell x) = \begin{cases} -\frac{\pi \cos(\pi\sqrt{|\beta|}(x-1))}{2\sqrt{|\beta|} \sin(\pi\sqrt{|\beta|})} + \frac{1}{2|\beta|}, & -1 < \beta < 0 \\ \frac{\pi^2(x-1)^2}{4} - \frac{\pi^2}{12}, & \beta = 0 \\ \frac{\pi \cosh(\pi\sqrt{\beta}(x-1))}{2\sqrt{\beta} \sinh(\pi\sqrt{\beta})} - \frac{1}{2\beta}, & \beta > 0 \end{cases} \tag{41}$$

for  $x \in [0, 1]$  and again applying (40) and  $\sinh(\alpha) \sinh(\beta) = \frac{1}{2}(\cosh(\alpha + \beta) - \cosh(\alpha - \beta))$ , respectively. To prove the second statement we use the ansatz  $Z(x) = u(x)B(v(x))$ ,  $u, v$  positive and  $v$  non-decreasing, which is the general form of a Gaussian Markov process, cf. [30]. Comparison of covariance functions yields explicit expressions for  $u$  and  $v$ . Further, one has  $u(x)B(v(x)) \stackrel{(d)}{=} u(x) \int_0^x \sqrt{v'(z)} dB(z)$  for  $v(0) = 0$  and, thus, the claimed semi-martingale representation follows from Itô's formula. □

*Proof of Proposition 2.3.* The necessity of the conditions on the parameters follow from the fact that (i) the parameter  $\sigma^2/\sqrt{\vartheta_2}e^{-\kappa x_0}$  may be consistently estimated using time increments, see [3], and (ii) the parameters  $\frac{\sigma^2}{\vartheta_2}$  and  $\kappa$  may be consistently estimated by computing the quadratic variation of the process  $x \mapsto X_t(x)$  on two different sub-intervals of  $[0, 1]$  in view of Proposition 2.1.

It remains to prove sufficiency of the conditions on the parameters:

(i) is a simple consequence of [23, Proposition 1]: Set  $\lambda_\ell = \vartheta_2(\pi^2\ell^2 + \Gamma)$  and  $\tilde{\lambda}_\ell = \vartheta_2(\pi^2\ell^2 + \tilde{\Gamma})$  where  $\Gamma = \frac{\vartheta_1^2}{4\vartheta_2^2} - \frac{\vartheta_0}{\vartheta_2}$  and  $\tilde{\Gamma} = \frac{\vartheta_1^2}{4\vartheta_2^2} - \frac{\tilde{\vartheta}_0}{\vartheta_2}$ . Then, absolute continuity follows from  $\sum_{\ell \geq 1} \frac{(\lambda_\ell - \tilde{\lambda}_\ell)^2}{\lambda_\ell} < \infty$ . Thanks to (i) and due to the one to one correspondence between  $\Gamma$  and  $\vartheta_0$  we may assume  $\Gamma = \tilde{\Gamma} = 0$  for the remainder of the proof.

(ii) follows from the fact that  $\text{Cov}(X_{t_0}(x), X_{t_0}(y))$  only depends on  $(\frac{\sigma^2}{\vartheta_2}, \kappa)$  in view of the Gaussianity of  $X$ .

For (iii) note that  $t \mapsto X_t(x_0)$  is a stationary Gaussian process with covariance function

$$\rho(t) = \sigma^2 \sum_{k \geq 1} \frac{e^{-\lambda_k t}}{2\lambda_k} e_k^2(x_0).$$

Let

$$f_{(\sigma^2, \vartheta)}(u) = \frac{1}{2\pi} \int e^{-iut} \rho(|t|) dt = \frac{1}{\pi} \int_0^\infty \cos(ut) \rho(t) dt = \frac{\sigma^2}{2\pi} \sum_{\ell \geq 1} \frac{e_\ell^2(x_0)}{\lambda_\ell^2 + u^2}$$

be the spectral density of  $t \mapsto X_t(x_0)$ . By Theorem 17 and its preceding discussion in [19] it suffices to show

$$\exists r > 1 : \quad \lim_{u \rightarrow \infty} u^r f_{(\sigma^2, \vartheta)}(u) \in (0, \infty) \quad \text{and} \quad \frac{f_{(\sigma^2, \vartheta)} - f_{(\tilde{\sigma}^2, \tilde{\vartheta})}}{f_{(\sigma^2, \vartheta)}} \in L^2(\mathbb{R}).$$

To prove these statements, we may assume  $\kappa = 0$  without loss of generality. Set  $h_{(\sigma^2, \vartheta_2)}(z) = \frac{\sigma^2}{\pi(\pi^4 \vartheta_2^2 z^4 + 1)}$ ,  $z \in \mathbb{R}$ . By Lemma A.9 (ii) we have for  $u \rightarrow \infty$

$$\begin{aligned} f_{(\sigma^2, \vartheta)}(u) &= \frac{1}{u^2} \sum_{\ell \geq 1} h_{(\sigma^2, \vartheta_2)}\left(\frac{\ell}{\sqrt{u}}\right) \sin^2(\pi \ell x_0) \\ &= \frac{1}{u^2} \left( \frac{\sqrt{u}}{2} \int_0^\infty h_{(\sigma^2, \vartheta_2)}(z) dz + \mathcal{O}\left(\frac{1}{\sqrt{u}}\right) \right), \end{aligned}$$

which proves the first condition. Now, if  $\sigma^2/\sqrt{\vartheta_2} = \tilde{\sigma}^2/\sqrt{\tilde{\vartheta}_2}$ , a substitution yields  $\int_0^\infty h_{(\sigma^2, \vartheta_2)}(z) dz = \int_0^\infty h_{(\tilde{\sigma}^2, \tilde{\vartheta}_2)}(z) dz$  and, therefore, the second condition follows due to

$$\frac{f_{(\sigma^2, \vartheta)}(u) - f_{(\tilde{\sigma}^2, \tilde{\vartheta})}(u)}{f_{(\sigma^2, \vartheta)}(u)} = \mathcal{O}\left(\frac{1}{u}\right), \quad u \rightarrow \infty. \quad \square$$

**Appendix A: Remaining proofs and auxiliary results**

**A.1. Covariances of double increments**

The following three lemmas are used to calculate the asymptotic variance of  $\mathbb{V}$ . Recall the definition of  $\tilde{D}_{ik}$  from (24).

**Lemma A.1.** *Let  $b \in (0, 1/2)$ . For  $J \geq 1$  define*

$$F_{J,\Delta}(z) = \sum_{\ell \geq 1} \frac{2e^{-\pi^2 \vartheta_2 J \ell^2 \Delta} - e^{-\pi^2 \vartheta_2 (J+1) \ell^2 \Delta} - e^{-\pi^2 \vartheta_2 (J-1) \ell^2 \Delta}}{2\pi^2 \vartheta_2 \ell^2} \cos(\pi \ell z)$$

and  $F_{0,\Delta} = F_{\vartheta_2}(\cdot, \Delta)$ . Then, for  $J = |i - j|$ ,

$$\text{Cov}(\tilde{D}_{ik}, \tilde{D}_{jl}) = -\sigma^2 e^{-\kappa \delta / 2} \cdot \begin{cases} 2D_\delta F_{J,\Delta}(0), & l = k, \\ D_\delta^2 F_{J,\Delta}(y_l - y_{k+1}), & l > k \end{cases} + \mathcal{O}\left(\frac{\sqrt{\Delta} \delta^2}{(J+1)^{3/2}}\right).$$

*Proof.* It immediately follows from the covariance structure  $\text{Cov}(u_\ell(s), u_\ell(t)) = \frac{\sigma^2}{2\lambda_\ell} e^{-\lambda_\ell |t-s|}$ ,  $s, t \geq 0$ , of the coefficient processes that

$$\begin{aligned} \text{Cov}(D_{ik}, D_{jl}) &= \sigma^2 \sum_{\ell \geq 1} (e_\ell(y_{k+1}) - e_\ell(y_k))(e_\ell(y_{l+1}) - e_\ell(y_l)) \\ &\quad \cdot \begin{cases} \frac{1 - e^{-\lambda_\ell \Delta}}{\lambda_\ell}, & J = 0, \\ \frac{2e^{-\lambda_\ell J \Delta} - e^{-\lambda_\ell (J+1) \Delta} - e^{-\lambda_\ell (J-1) \Delta}}{2\lambda_\ell}, & J \geq 1. \end{cases} \end{aligned}$$

*Step 1.* We show negligibility of  $\Gamma$ . From the first step of the last proof we already know that

$$\begin{aligned} &\text{Cov}(D_{ik}, D_{il}) \\ &= \sigma^2 \sum_{\ell \geq 1} \frac{1 - e^{-\pi^2 \vartheta_2 \ell^2 \Delta}}{\pi^2 \vartheta_2 \ell^2} (e_\ell(y_{k+1}) - e_\ell(y_k))(e_\ell(y_{l+1}) - e_\ell(y_l)) + \mathcal{O}\left(\sqrt{\Delta} \delta^2\right). \end{aligned}$$

For  $J \geq 1$  we will show now that

$$\begin{aligned} \text{Cov}(D_{ik}, D_{jl}) &= \sigma^2 \sum_{\ell \geq 1} \frac{2e^{-\pi^2 \vartheta_2 \ell^2 J \Delta} - e^{-\pi^2 \vartheta_2 \ell^2 (J+1) \Delta} - e^{-\pi^2 \vartheta_2 \ell^2 (J-1) \Delta}}{2\pi^2 \vartheta_2 \ell^2} \\ &\quad \cdot (e_\ell(y_{k+1}) - e_\ell(y_k))(e_\ell(y_{l+1}) - e_\ell(y_l)) + \mathcal{O}\left(\frac{\sqrt{\Delta} \delta^2}{(J+1)^{3/2}}\right). \end{aligned}$$

If  $J = 1$  this directly follows from the case  $J = 0$  since

$$\frac{2e^{-\lambda_\ell \Delta} - e^{-2\lambda_\ell \Delta} - 1}{2\lambda_\ell} = \frac{1 - e^{-2\lambda_\ell \Delta}}{2\lambda_\ell} - \frac{1 - e^{-\lambda_\ell \Delta}}{\lambda_\ell}. \tag{42}$$

For  $J \geq 2$  define  $g_J(x) = \frac{2e^{-Jx} - e^{-(J+1)x} - e^{-(J-1)x}}{2x}$ . A first order Taylor approximation of  $g_J$  gives

$$\begin{aligned} \text{Cov}(D_{ik}, D_{jl}) &= \Delta \sum_{\ell \geq 1} g_J(\lambda_\ell \Delta) (e_\ell(y_{k+1}) - e_\ell(y_k))(e_\ell(y_{l+1}) - e_\ell(y_l)) \\ &= \Delta \sum_{\ell \geq 1} g_J(\pi^2 \vartheta_2 \ell^2 \Delta) (e_\ell(y_{k+1}) - e_\ell(y_k))(e_\ell(y_{l+1}) - e_\ell(y_l)) + R, \end{aligned}$$

where  $R \lesssim \Delta^2 \sum_{\ell \geq 1} g'_J(\vartheta_2(\pi^2 \ell^2 + \xi_\ell) \Delta) \ell^2 \delta^2$  for some  $|\xi_\ell| \leq |\Gamma|$ . It can be shown easily that  $g'_J(x) \lesssim e^{-(J-1)x/2}$ . Therefore, for some  $\omega > 0$  and by regarding  $R$  as a Riemann sum with lag  $\sqrt{(J-1)\Delta}$ ,

$$R \lesssim \Delta^2 \sum_{\ell \geq 1} e^{-\omega(J-1)\ell^2 \Delta} \ell^2 \delta^2 \lesssim \frac{\sqrt{\Delta} \delta^2}{(J-1)^{3/2}} \lesssim \frac{\sqrt{\Delta} \delta^2}{(J+1)^{3/2}}.$$

Step 2. By Step 1 we may assume  $\lambda_\ell = \pi^2 \vartheta_2 \ell^2$ . By (23), we have

$$\begin{aligned} \text{Cov}(\tilde{D}_{ik}, \tilde{D}_{jk}) &= -2\sigma^2 g(\delta) D_\delta F_{J,\Delta}(0) + \sigma^2 F_{J,\Delta}(0) D_\delta^2 g(0) \\ &\quad - \sigma^2 D_\delta^2 (g(\cdot) F_{J,\Delta}(2y_k + \cdot)) (0) \end{aligned}$$

and by (22) for  $l > k$

$$\begin{aligned} \text{Cov}(\tilde{D}_{ik}, \tilde{D}_{jl}) &= -\sigma^2 g(\delta) D_\delta^2 F_{J,\Delta}(y_l - y_{k+1}) + \sigma^2 F_{J,\Delta}(y_l - y_k) D_\delta^2 g(0) \\ &\quad - \sigma^2 D_\delta^2 (g(\cdot) F_{J,\Delta}(y_l + y_k + \cdot)) (0). \end{aligned}$$

Hence, as in previous Lemmas it is sufficient to establish

$$F_{J,\Delta}(0), F_{J,\Delta}(z), F'_{J,\Delta}(z) F''_{J,\Delta}(z) \lesssim \frac{\sqrt{\Delta}}{J^{3/2}}, \quad z \in [2b, 2(1-b)].$$

For  $J = 0$  this was already proven in Proposition 3.5. The case  $J = 1$  follows from the case  $J = 0$  since (42) shows

$$F_{1,\Delta}(z) = \frac{1}{2} F_{2,\Delta}(z) - F_\Delta(z). \tag{43}$$

For  $J \geq 2$  we have

$$2e^{-\lambda_\ell J \Delta} - e^{-\lambda_\ell (J+1) \Delta} - e^{-\lambda_\ell (J-1) \Delta} \lesssim e^{-\lambda_\ell (J-1) \Delta} (\lambda_\ell \Delta)^2, \tag{44}$$

and therefore, again using a Riemann sum approximation with lag  $\sqrt{(J-1)\Delta}$ ,

$$F_{J,\Delta}(z) \lesssim F_{J,\Delta}(0) \lesssim \sum_{\ell \geq 1} \lambda_\ell \Delta^2 e^{-\lambda_\ell (J-1) \Delta} = \mathcal{O} \left( \frac{\sqrt{\Delta}}{(J-1)^{3/2}} \right).$$

The bound on the first derivative is provided by Lemma A.7,

$$F'_{J,\Delta}(z) \lesssim \sum_{\ell \geq 1} \frac{2e^{-\lambda_\ell J \Delta} - e^{-\lambda_\ell (J+1) \Delta} - e^{-\lambda_\ell (J-1) \Delta}}{2\lambda_\ell} \ell \sin(\pi \ell z)$$

$$\begin{aligned} &\lesssim \sup_{\ell} \left| \frac{2e^{-\lambda_{\ell}J\Delta} - e^{-\lambda_{\ell}(J+1)\Delta} - e^{-\lambda_{\ell}(J-1)\Delta}}{2\lambda_{\ell}} \ell \right| \frac{1}{z \wedge (2-z)} \\ &\lesssim \sup_{\ell} \left| \lambda_{\ell} \Delta^2 e^{-\lambda_{\ell}J\Delta} \ell \right| \lesssim \frac{\sqrt{\Delta}}{J^{3/2}}. \end{aligned}$$

Finally, to bound  $F''_{J,\Delta}$  we define  $h_J(z) = 2e^{-Jz^2} - e^{-(J+1)z^2} - e^{-(J-1)z^2}$ . Clearly,  $h_J(0) = 0$  and

$$h'_J(z) = -2(J-1)ze^{-(J-1)z^2} \underbrace{(2e^{-z^2} - e^{-2z^2} - 1)}_{\lesssim z^4} - 4e^{-(J-1)z^2} \underbrace{(ze^{-z^2} - ze^{-2z^2})}_{\lesssim z^3},$$

implying  $\|h'_J\|_{\infty} \lesssim J^{-3/2}$ . In view of Lemma A.8 this shows

$$\begin{aligned} F''_{J,\Delta}(z) &\lesssim \sum_{\ell \geq 1} \frac{2e^{-\lambda_{\ell}J\Delta} - e^{-\lambda_{\ell}(J+1)\Delta} - e^{-\lambda_{\ell}(J-1)\Delta}}{2\lambda_{\ell}} \ell^2 \cos(\pi \ell z) \\ &\lesssim \sum_{\ell \geq 1} h_J(\sqrt{\lambda_{\ell}\Delta}) \cos(\pi \ell z) = \mathcal{O} \left( \frac{1}{(z \wedge (2-z))^2} \frac{\sqrt{\Delta}}{J^{3/2}} \right). \quad \square \end{aligned}$$

**Lemma A.2.** For  $J \in \mathbb{N}_0$  and  $z \in (0, 2)$  it holds that

- (i)  $F_{J,\Delta}(0) - F_{J,\Delta}(\delta) = \delta \frac{1}{2\vartheta_2} \mathbf{1}_{\{J=0\}} - \delta \frac{1}{4\vartheta_2} \mathbf{1}_{\{J=1\}} + \mathcal{O} \left( \frac{\delta^2}{(J+1)^{5/2}\sqrt{\Delta}} \right)$
- (ii)  $2F_{J,\Delta}(z) - F_{J,\Delta}(z + \delta) - F_{J,\Delta}(z - \delta) = \mathcal{O} \left( \frac{\delta^2}{(J+1)^2} \left( \frac{1}{\sqrt{\Delta}} \wedge \frac{1}{z \wedge (2-z)} \right) \right)$ .

*Proof.* (i) The validity for the case  $J = 0$  follows from the proof of Proposition 3.5 (ii), the case  $J = 1$  follows from (43). For  $J \geq 2$  we have by Taylor's theorem

$$F_{J,\Delta}(0) - F_{J,\Delta}(\delta) = -\delta F'_{J,\Delta}(0) - \frac{\delta^2}{2} F''_{J,\Delta}(\xi)$$

for some  $\xi \in [0, \delta]$ . Now, the claim is proved by inserting  $F'_{J,\Delta}(0) = 0$  and noting due to (44):

$$\begin{aligned} \|F''_{J,\Delta}\|_{\infty} &\lesssim \sum_{\ell \geq 1} (2e^{-\lambda_{\ell}J\Delta} - e^{-\lambda_{\ell}(J+1)\Delta} - e^{-\lambda_{\ell}(J-1)\Delta}) \\ &\lesssim \sum_{\ell \geq 1} \lambda_{\ell}^2 \Delta^2 e^{-\lambda_{\ell}(J-1)\Delta} \lesssim \frac{1}{J^{5/2}\sqrt{\Delta}}. \end{aligned}$$

(ii) As in previous Lemmas it suffices to establish

$$F''_{J,\Delta}(z) \lesssim \frac{1}{(J+1)^2} \left( \frac{1}{\sqrt{\Delta}} \wedge \frac{1}{z \wedge (2-z)} \right).$$

For the case  $J = 0$  we employ the representation  $F_{\Delta} = H_{\Delta} + G_{\Delta}$  from Proposition 3.5. The validity of the bound on  $H''_{\Delta}$  follows from  $H''_{\Delta}(z) \lesssim \frac{1}{\sqrt{\Delta}} \wedge \frac{1}{z \wedge (2-z)}$ . The bound on  $G''_{\Delta}(z)$  follows from  $\|G''_{\Delta}\|_{\infty} \lesssim 1/\sqrt{\Delta}$  and  $G''_{\Delta}(z) \lesssim$



$\sup_{\ell} \left| \frac{1 - e^{-\lambda_{\ell} \Delta (1 + \lambda_{\ell} \Delta)}}{1 + \lambda_{\ell} \Delta} \right| \frac{1}{z \wedge (2 - z)} \lesssim \frac{1}{z \wedge (2 - z)}$ , see Lemma A.7. The case  $J = 1$  follows from the case  $J = 0$ , see (43). For  $J \geq 2$  we proceed in the same way: In the proof of (i) it was shown that  $\|F''_{\Delta, J}\|_{\infty} \lesssim \frac{1}{J^{5/2} \sqrt{\Delta}} \lesssim \frac{1}{J^2 \sqrt{\Delta}}$ . Finally, by Lemma A.7,

$$\begin{aligned} F''_{J, \Delta}(z) &\lesssim \sup_{\ell} \left| 2e^{-\lambda_{\ell} J \Delta} - e^{-\lambda_{\ell} (J+1) \Delta} - e^{-\lambda_{\ell} (J-1) \Delta} \right| \frac{1}{z \wedge (2 - z)} \\ &\lesssim \sup_{\ell} \left| (\lambda_{\ell} \Delta)^2 e^{-\lambda_{\ell} (J-1) \Delta} \right| \frac{1}{z \wedge (2 - z)} \lesssim \frac{1}{(J+1)^2} \frac{1}{z \wedge (2 - z)}. \quad \square \end{aligned}$$

**Lemma A.3.** For  $J \in \mathbb{N}_0$  and  $z \in (0, 2)$  we have

$$\begin{aligned} (i) \quad &F_{J, \Delta}(0) - F_{J, \Delta}(\delta) \\ &= \begin{cases} \frac{\sqrt{\Delta}}{\sqrt{\vartheta_2 \pi}} + \mathcal{O}\left(\frac{\Delta^{3/2}}{\delta^2}\right), & J = 0, \\ \frac{\sqrt{\Delta}}{2\sqrt{\pi \vartheta_2}} \left(\sqrt{J-1} + \sqrt{J+1} - 2\sqrt{J}\right) + \mathcal{O}\left(\Delta^{3/2} + \frac{\Delta}{(J+1)\delta}\right), & J \geq 1, \end{cases} \\ (ii) \quad &2F_{J, \Delta}(\delta) - F_{J, \Delta}(0) - F_{J, \Delta}(2\delta) \\ &= \begin{cases} -\frac{\sqrt{\Delta}}{\sqrt{\vartheta_2 \pi}} + \mathcal{O}\left(\frac{\Delta^{3/2}}{\delta^2}\right), & J = 0, \\ -\frac{\sqrt{\Delta}}{2\sqrt{\pi \vartheta_2}} \left(\sqrt{J-1} + \sqrt{J+1} - 2\sqrt{J}\right) + \mathcal{O}\left(\Delta^{3/2} + \frac{\Delta}{(J+1)\delta}\right), & J \geq 1, \end{cases} \\ (iii) \quad &2F_{J, \Delta}(z) - F_{J, \Delta}(z - \delta) - F_{J, \Delta}(z + \delta) = \mathcal{O}\left(\frac{\Delta}{J+1} \frac{1}{z \wedge (2 - z)}\right). \end{aligned}$$

*Proof.* (iii) It is sufficient to show

$$F_{J, \Delta}(z) = \mathcal{O}\left(\frac{\Delta}{J+1} \frac{1}{z \wedge (2 - z)}\right) \tag{45}$$

for  $J \in \mathbb{N}_0$  and  $z \in (0, 2)$ : If  $J = 0$ , Lemma A.7 gives

$$F_{\Delta}(z) \lesssim \sup_{\ell \geq 1} \left| \frac{1 - e^{-\lambda_{\ell} \Delta}}{\lambda_{\ell}} \right| \frac{1}{z \wedge (2 - z)} \lesssim \frac{\Delta}{z \wedge (2 - z)}.$$

By (43) this bound is also valid for  $F_{1, \Delta}(z)$ . For  $J \geq 2$  the same method gives

$$\begin{aligned} F_{J, \Delta}(z) &\lesssim \sup_{\ell \geq 1} \left| \frac{2e^{-\lambda_{\ell} J \Delta} - e^{-\lambda_{\ell} (J+1) \Delta} - e^{-\lambda_{\ell} (J-1) \Delta}}{\lambda_{\ell}} \right| \frac{1}{z \wedge (2 - z)} \\ &\lesssim \sup_{\ell \geq 1} \left| \lambda_{\ell} \Delta^2 e^{-\lambda_{\ell} J \Delta} \right| \frac{1}{z \wedge (2 - z)} \lesssim \frac{\Delta}{J} \frac{1}{z \wedge (2 - z)}, \end{aligned}$$

where we have used (44).

(i) The case  $J = 0$  was already shown in the proof of Proposition 3.5. For  $J \geq 1$  we prove

$$F_{J, \Delta}(0) = \frac{\sqrt{\Delta}}{2\sqrt{\pi \vartheta_2}} \left(\sqrt{J-1} + \sqrt{J+1} - 2\sqrt{J}\right) + \mathcal{O}(\Delta^{3/2}),$$

then (ii) follows in view of (45): If  $J = 1$  we use (25) to calculate

$$\begin{aligned} F_{1,\Delta}(0) &= \frac{1}{2}F_{2\Delta}(0) - F_{\Delta}(0) = \frac{1}{2} \left( \frac{\sqrt{2\Delta}}{\sqrt{\pi\vartheta_2}} - \Delta \right) - \left( \frac{\sqrt{\Delta}}{\sqrt{\pi\vartheta_2}} - \frac{\Delta}{2} \right) + \mathcal{O}(\Delta^{3/2}) \\ &= \frac{\sqrt{\Delta}}{2\sqrt{\pi\vartheta_2}} (\sqrt{2} - 2) + \mathcal{O}(\Delta^{3/2}). \end{aligned}$$

For  $J \geq 2$  define  $g_J(z) = \frac{2e^{-J\pi^2\vartheta_2z^2} - e^{-(J+1)\pi^2\vartheta_2z^2} - e^{-(J-1)\pi^2\vartheta_2z^2}}{2\pi^2\vartheta_2z^2}$ . Then,

$$\int_0^\infty g_J(z) dz = \frac{1}{2\sqrt{\pi\vartheta_2}} (\sqrt{J-1} + \sqrt{J+1} - 2\sqrt{J})$$

and since  $g_J(0) = 0$  we have by Lemma A.9

$$\begin{aligned} F_{J,\Delta}(0) &= \Delta \sum_{\ell \geq 1} g_J(\ell\sqrt{\Delta}) = \sqrt{\Delta} \int_0^\infty g_J(z) dz + \mathcal{O}(\Delta^{3/2}) \\ &= \frac{\sqrt{\Delta}}{2\sqrt{\pi\vartheta_2}} (\sqrt{J-1} + \sqrt{J+1} - 2\sqrt{J}) + \mathcal{O}(\Delta^{3/2}). \end{aligned}$$

Finally, (ii) is a direct consequence of (i). □

### A.2. Auxiliary results for the lower bounds

For the proofs of Propositions 5.3 and 5.5 we require the following auxiliary lemmas.

**Lemma A.4.** Consider a discrete sample  $(u(i\Delta), i = 0, \dots, N)$  of an Ornstein-Uhlenbeck process given by

$$du(t) = -a\mu u(t) dt + \nu\sqrt{\mu} dB_t, \quad u(0) \sim \mathcal{N}\left(0, \frac{\nu^2}{2a}\right)$$

and assume  $\Delta = 1/N$ . Then, the Fisher information  $I = I_N$  for the parameter  $(\mu, \nu^2)$  is given by

$$I_{11} = \frac{a^2\Delta(e^{-4\mu a\Delta} + e^{-2\mu a\Delta})}{(1 - e^{-2\mu a\Delta})^2}, \quad I_{12} = \frac{ae^{-2\mu a\Delta}}{\nu^2(1 - e^{-2\mu a\Delta})}, \quad I_{22} = \frac{N + 1}{2\nu^4}.$$

*Proof.* By the Markov property of  $u$ , the log-likelihood function of  $(\mu, \nu^2)$  is given by

$$\ell(\mu, \nu^2) = \log \pi_0(u(0)) + \sum_{i=0}^{N-1} \log p_\Delta(u(i\Delta), u((i+1)\Delta)),$$

where  $p_t(x, y) = \frac{1}{\sqrt{\pi\nu^2(1-e^{-2\mu at})/a}} \exp\left(-\frac{(y-xe^{-\mu at})^2}{\nu^2(1-e^{-2\mu at})/a}\right)$  is the transition density of  $u$  and  $\pi_0$  is the density of the initial distribution  $\mathcal{N}\left(0, \frac{\nu^2}{2a}\right)$ . By stationarity of  $u$ , the Fisher information simplifies to

$$I = -\mathbf{E} (D^2\ell(\mu, \nu^2)) = -\mathbf{E} (D^2 \log \pi_0(u(0))) - N\mathbf{E} (D^2 \log p_\Delta(u(0), u(\Delta))),$$

where we write  $D^2g$  for the Hessian of a function  $g$ . This expression can be computed explicitly, yielding the claimed formulas.  $\square$

**Lemma A.5.** *The function  $g: [0, \infty) \times [-\pi, \pi] \rightarrow \mathbb{R}$  defined by*

$$g(x, \omega) = \frac{2x^2 - \sinh(x^2) \cosh(x^2) + \cos(\omega)(\sinh(x^2) - 2x^2 \cosh(x^2))}{x^2(\cosh(x^2) - \cos(\omega))^2} (1 - \cos(\omega))$$

satisfies

- (i)  $\int_0^\infty g(x, \omega) dx = 0$ , for all  $\omega \in [-\pi, \pi]$ ,
- (ii)  $\sup_{|\omega| \leq \pi} \|\frac{\partial}{\partial x} g(\cdot, \omega)\|_{L^1} < \infty$ .
- (iii)  $|g(x, \omega)| \lesssim \frac{1+x^2}{x^4} \omega^2$  uniformly in  $\omega \in [-\pi, \pi]$ ,  $x > 0$ .

*Proof.* (i) follows from the fact that

$$G(x, \omega) := \frac{\sinh(x^2)(1 - \cos(\omega))}{x(\cosh(x^2) - \cos(\omega))}, \quad x > 0, \omega \in [-\pi, \pi],$$

is a primitive of  $x \mapsto g(x, \omega)$  and since  $\lim_{x \rightarrow \infty} G(x, \omega) = \lim_{x \rightarrow 0} G(x, \omega) = 0$  for all  $\omega \in [-\pi, \pi]$ .

(ii) can be shown by writing  $G(\cdot, \omega)$  as a sum of monotonic functions and noting that for a monotonic function  $g : \mathbb{R}_+ \rightarrow \mathbb{R}$  it holds that  $\|g'\|_{L^1} = |\lim_{x \rightarrow \infty} g(x) - \lim_{x \rightarrow 0} g(x)|$ .

Finally, (iii) follows by direct calculations.  $\square$

**Lemma A.6.** *Consider the parametrization of Proposition 5.5 and the function  $\Phi_k^{N, \Delta}$  from (36). If  $M\sqrt{\Delta} \rightarrow 0$ , then*

- (i)  $\Phi_k^{N, \Delta}(\omega) > 0$  for all  $\omega \in [-\pi, \pi]$ ,
- (ii)

$$\Phi_k^{N, \Delta}(\omega) \gtrsim \begin{cases} \frac{\sqrt{\Delta}}{M} \sqrt{|\omega|}, & |\omega| \geq M^2\Delta, & (46a) \\ \Delta, & k^2\Delta \leq |\omega| \leq M^2\Delta, & (46b) \\ \frac{\omega^2}{k^4\Delta} + \Delta e^{-\vartheta_2 k^2}, & |\omega| \leq k^2\Delta, & (46c) \end{cases}$$

(iii)

$$\frac{\partial}{\partial \vartheta_2} \Phi_k^{N, \Delta}(\omega) \lesssim \begin{cases} \Delta, & \omega \in [-\pi, \pi], & (47a) \\ \frac{\omega^2}{k^4\Delta} + \Delta k^2 e^{-\vartheta_2 k^2}, & |\omega| \leq k^2\Delta. & (47b) \end{cases}$$

*Proof.* Without loss of generality let  $\theta = \pi^2 \vartheta_2$  and  $\sigma_0^2 = \pi^2$ . We denote the covariance function of  $\bar{U}_k$  by  $\rho_k : \mathbb{Z} \rightarrow \mathbb{R}$  and write  $\Phi_k^N$  instead of  $\Phi_k^{N,\Delta}$ , i.e.  $\Phi_k^N(\omega) = \sum_{j=1-N}^{N-1} \rho_k(j) e^{-ij\omega}$ ,  $\omega \in [-\pi, \pi]$ . (i) Let  $r_k$  be the covariance function of the process  $(U_k(t_0), U_k(t_1), \dots)$ , i.e.

$$r_k(j) = \sum_{\ell \in \mathcal{I}_k} \frac{e^{-\theta \ell^2 |j| \Delta}}{2\sqrt{\theta} \ell^2}, \quad j \in \mathbb{Z},$$

where  $\mathcal{I}_k = \mathcal{I}_k^+ \cup \mathcal{I}_k^-$ . Note that  $r_k$  and  $\rho_k$  are related by  $\rho_k(j) = 2r_k(j) - r_k(j-1) - r_k(j+1)$ ,  $j \in \mathbb{Z}$ , which is a second order difference. Since  $x \mapsto e^{-x}$  has a positive second derivative, it follows that  $\rho_k(j) < 0$  if  $j \neq 0$ . On the other hand, for  $j = 0$  we have  $\rho_k(0) = \text{Var}(\bar{U}_k(t_0)) > 0$  and therefore,

$$\begin{aligned} \Phi_k^N(\omega) &= \rho_k(0) + 2 \sum_{j=1}^{N-1} \rho_k(j) \cos(j\omega) \\ &\geq \rho_k(0) + 2 \sum_{j=1}^{N-1} \rho_k(j) = 2(r_k(N-1) - r_k(N)) > 0. \end{aligned}$$

To treat (ii) and (iii) we calculate

$$\begin{aligned} \Phi_k^N(\omega) &= \sum_{j=1-N}^{N-1} \rho_k(j) e^{-ij\omega} \\ &= 2(1 - \cos(\omega)) \sum_{j=2-N}^{N-2} r_k(j) e^{-ij\omega} + 4r_k(N-1) \cos((N-1)\omega) \\ &\quad - 2r_k(N) \cos((N-1)\omega) - 2r_k(N-1) \cos((N-2)\omega). \end{aligned}$$

From  $\sum_{j=0}^{J-1} z^j = \frac{1-z^J}{1-z}$  for  $z \in \mathbb{C} \setminus \{1\}$  it follows that

$$\begin{aligned} &\sum_{j=1-J}^{J-1} e^{-\theta \ell^2 |j| \Delta} e^{-ij\omega} \\ &= \frac{1 - e^{-2\theta \ell^2 \Delta} + 2e^{-(J+1)\theta \ell^2 \Delta} \cos((J-1)\omega) - 2e^{-J\theta \ell^2 \Delta} \cos(J\omega)}{1 + e^{-2\theta \ell^2 \Delta} - 2e^{-\theta \ell^2 \Delta} \cos(\omega)} \\ &= \frac{\sinh(\theta \ell^2 \Delta) + e^{-J\theta \ell^2 \Delta} \cos((J-1)\omega) - e^{-(J-1)\theta \ell^2 \Delta} \cos(J\omega)}{\cosh(\theta \ell^2 \Delta) - \cos(\omega)} \end{aligned}$$

for  $J \geq 1$  and by elementary manipulations we can pass to the representation  $\Phi_k^N = \Phi + R_N$ , where

$$\Phi(\omega) = (1 - \cos(\omega)) \sum_{\ell \in \mathcal{I}_k} \frac{1}{\sqrt{\theta} \ell^2} \frac{\sinh(\theta \ell^2 \Delta)}{\cosh(\theta \ell^2 \Delta) - \cos(\omega)},$$

$$R_N(\omega) = \sum_{\ell \in \mathcal{I}_k} (1 - \cosh(\theta \ell^2 \Delta)) \frac{e^{-\theta \ell^2 (N-1)\Delta}}{\sqrt{\theta} \ell^2} \frac{e^{-\theta \ell^2 \Delta} \cos((N-1)\omega) - \cos(N\omega)}{\cosh(\theta \ell^2 \Delta) - \cos(\omega)}.$$

Note that we have suppressed the dependence on  $k$  for ease of notation. We remark that  $\Phi(\omega) = \sum_{j \in \mathbb{Z}} \rho_k(j) e^{-ij\omega}$ ,  $\omega \in [-\pi, \pi]$ , is the spectral density of the process  $(\tilde{U}_k(j))_{j \geq 0}$ .

(ii) To prove (46a) we note that for  $\omega \geq M^2 \Delta$  we have

$$\begin{aligned} & \left| e^{-\theta \ell^2 \Delta} \cos((N-1)\omega) - \cos(N\omega) \right| \\ &= \left| (e^{-\theta \ell^2 \Delta} - 1) \cos((N-1)\omega) + \cos((N-1)\omega) - \cos(N\omega) \right| \lesssim \ell^2 \Delta + \omega \lesssim \ell^2 \omega. \end{aligned}$$

Consequently,

$$\begin{aligned} R_N(\omega) &\lesssim \sum_{\ell \in \mathcal{I}_k} \ell^2 \Delta \sinh(\theta \ell^2 \Delta) \frac{e^{-\theta \ell^2 (N-1)\Delta}}{\sqrt{\theta} \ell^2} \frac{\ell^2 \omega}{\cosh(\theta \ell^2 \Delta) - \cos(\omega)} \\ &\lesssim \frac{\Delta}{\omega} \sum_{\ell \in \mathcal{I}_k} \frac{\sinh(\theta \ell^2 \Delta)}{\ell^2 (\cosh(\theta \ell^2 \Delta) - \cos(\omega))} (1 - \cos(\omega)) \lesssim \frac{1}{M^2} \Phi(\omega) \end{aligned}$$

and, hence,  $R_N$  is negligible compared to  $\Phi$ . In order to compute an asymptotic expression for  $\Phi$ , set

$$h(x, \omega) = \frac{\sinh(\theta x^2)(1 - \cos(\omega))}{x^2 (\cosh(\theta x^2) - \cos(\omega))}, \quad x > 0, \omega \in [-\pi, \pi].$$

We have  $\frac{\partial h}{\partial x} \leq 0$  and therefore,  $\left\| \frac{\partial}{\partial x} h(\cdot, \omega) \right\|_{L^1} = h(0, \omega) - \lim_{x \rightarrow \infty} h(x, \omega) = \theta$  is uniformly bounded in  $\omega$ . Thus, using the mean value theorem and a Riemann sum approximation with mesh size  $M\sqrt{\Delta}$  for  $\frac{\partial}{\partial x} h(\cdot, \omega)$ , we obtain

$$\Phi(\omega) \approx \Delta \sum_{\ell \in \mathcal{I}_k} h(\ell\sqrt{\Delta}, \omega) = \Delta \sum_{\ell=1}^{\infty} h(2\ell M\sqrt{\Delta}, \omega) + \mathcal{O}(\Delta).$$

Further, since

$$\left| \varepsilon \sum_{\ell \geq 1} f(\ell\varepsilon) - \int_0^\infty f(x) dx \right| \leq \varepsilon \|f'\|_{L^1} \tag{48}$$

for any function  $f \in C^1[0, \infty)$ , we get  $\Phi(\omega) \approx \frac{\sqrt{\Delta}}{M} \int_0^\infty h(x, \omega) dx + \mathcal{O}(\Delta)$ . Finally, due to

$$a + b \approx \max(a, b), \quad a, b > 0, \tag{49}$$

we have  $(\cosh(\theta \omega x^2) - \cos(\omega)) \approx \max(\cosh(\theta \omega x^2) - 1, 1 - \cos(\omega))$  and, consequently,

$$h(\sqrt{\omega}x, \omega) = \frac{\sinh(\theta \omega x^2)(1 - \cos(\omega))}{\omega x^2 (\cosh(\theta \omega x^2) - \cos(\omega))} \gtrsim \frac{\sinh(\theta \omega x^2)}{\omega x^2} \gtrsim 1, \quad x \leq \theta^{-1/2}.$$

Therefore,

$$\int_0^\infty h(x, \omega) dx = \sqrt{\omega} \int_0^\infty h(\sqrt{\omega}x, \omega) dx \gtrsim \sqrt{\omega},$$

finishing the proof of (46a).

To prove (46b) and (46c), let us write  $\Phi = \sum_{\ell \in \mathcal{I}_k} \varphi_\ell$  and  $R_N = \sum_{\ell \in \mathcal{I}_k} \varrho_\ell^N$ . Since the argument in the proof of (i) was on a summand-wise level, also each of the functions  $\varphi_\ell + \varrho_\ell^N$  is positive,  $\ell \in \mathbb{N}$ . Therefore, we can bound  $\Phi_k^N$  from below with the first summand,

$$\Phi_k^N \geq \varphi_k + \varrho_k^N = \varrho_k^N(0) + \varphi_k + (\varrho_k^N - \varrho_k^N(0)).$$

We show that there exists an environment  $U$  around zero and some  $\delta \in (0, 1)$  such that

$$|\varrho_k^N(\omega) - \varrho_k^N(0)| \leq (1 - \delta)\varphi_k(\omega), \quad \omega \in U : \tag{50}$$

A simple calculation yields

$$\begin{aligned} & \varrho_k^N(\omega) - \varrho_k^N(0) \\ &= e^{-(N-1)\theta k^2 \Delta} \frac{(\cos((N-1)\omega) - \cos(N\omega))(1 - \cosh(\theta k^2 \Delta))}{\sqrt{\theta} k^2 (\cosh(\theta k^2 \Delta) - \cos(\omega))} \\ &+ e^{-(N-1)\theta k^2 \Delta} \frac{(1 - e^{-\theta k^2 \Delta})(1 - \cos((N-1)\omega))(1 - \cosh(\theta k^2 \Delta))}{\sqrt{\theta} k^2 (\cosh(\theta k^2 \Delta) - \cos(\omega))} \\ &+ e^{-(N-1)\theta k^2 \Delta} \frac{(e^{-\theta k^2 \Delta} - 1)(1 - \cos(\omega))}{\sqrt{\theta} k^2 (\cosh(\theta k^2 \Delta) - \cos(\omega))}. \end{aligned}$$

Since  $\cos(x) - \cos(y) = -2 \sin \frac{x+y}{2} \sin \frac{x-y}{2}$ ,  $x, y \in \mathbb{R}$ , we have

$$\left| \cos((N-1)\omega) - \cos(N\omega) \right| = \left| 2 \sin \left( \frac{(2N-1)\omega}{2} \right) \sin \left( \frac{\omega}{2} \right) \right| \leq N\omega^2. \tag{51}$$

Therefore, for any  $\alpha > 0$  there exists an environment  $U$  of 0 such that

$$\begin{aligned} |\cos((N-1)\omega) - \cos(N\omega)| &\leq N\omega^2 \leq N(1 - \cos(\omega))(2 + \alpha) \\ 1 - \cos((N-1)\omega) &\leq \frac{N^2\omega^2}{2} \leq \frac{N^2}{2}(1 - \cos(\omega))(2 + \alpha) \end{aligned}$$

holds for all  $\omega \in U$ . Further, for all  $x \geq 0$  we have  $\cosh(x) - 1 \leq \frac{\sinh(x)x}{2}$ ,  $1 - e^{-x} \leq \sinh(x)$ , and consequently,

$$\begin{aligned} \frac{|\varrho_k^N(\omega) - \varrho_k^N(0)|}{\varphi_k(\omega)} &\leq e^{-(N-1)\theta k^2 \Delta} \left( 1 + \frac{2 + \alpha}{2} \theta k^2 + \frac{2 + \alpha}{4} \theta^2 k^4 \right) \\ &\leq \frac{2 + \alpha}{2} e^{\Delta \theta k^2} e^{-\theta k^2} \left( 1 + \theta k^2 + \frac{\theta^2 k^4}{2} \right) < \frac{2 + \alpha}{2} e^{\Delta \theta k^2}. \end{aligned}$$

Clearly, for  $\Delta$  sufficiently small one can choose  $\alpha$  in such a way that this bound is strictly less than 1 for all  $k \leq M - 1$ , yielding (50). Consequently, it is sufficient to prove (46b) and (46c) with  $\Phi_k^N$  replaced by  $\varphi_k + \varrho_k^N(0)$ : Now,

$$\varphi_k(0) + \varrho_k^N(0) = \varrho_k^N(0) = e^{-\theta k^2(N-1)\Delta} \frac{1 - e^{-\theta k^2\Delta}}{k^2} \approx \Delta e^{-\theta k^2}$$

and again by using (49), we get

$$\begin{aligned} \varphi_k(\omega) &\gtrsim \frac{\sinh(\theta k^2\Delta)}{k^2} \gtrsim \Delta, & \omega &\geq k^2\Delta, \\ \varphi_k(\omega) &\gtrsim (1 - \cos(\omega)) \frac{1}{\sqrt{\theta}k^2} \frac{\sinh(\theta k^2\Delta)}{\cosh(\theta k^2\Delta) - 1} \gtrsim \frac{\omega^2}{k^4\Delta}, & \omega &\leq k^2\Delta. \end{aligned}$$

(iii) We show (47a): We have  $\frac{\partial}{\partial\theta}\Phi(\omega) = \frac{\Delta}{2\sqrt{\theta}} \sum_{\ell \in \mathcal{I}_k} g(\ell\sqrt{\theta\Delta}, \omega)$  with  $g$  defined in Lemma A.5. Using the properties of  $g$  derived in Lemma A.5 and the Riemann sum approximation (48) with mesh size  $M\sqrt{\Delta}$ , we obtain

$$\frac{\partial}{\partial\theta}\Phi(\omega) \approx \Delta \sum_{\ell \geq 1} g(\ell M\sqrt{\Delta}, \omega) + \mathcal{O}(\Delta) = \frac{\sqrt{\Delta}}{M} \int_0^\infty g(x, \omega) dx + \mathcal{O}(\Delta) = \mathcal{O}(\Delta).$$

To show that also  $\frac{\partial}{\partial\theta}R_N$  is of the claimed order, we write

$$\begin{aligned} \varrho_\ell^N &= \alpha_\ell \beta_\ell \quad \text{where} \quad \alpha_\ell(\omega) = \frac{1 - \cosh(\theta\ell^2\Delta)}{\sqrt{\theta}\ell^2 (\cosh(\theta\ell^2\Delta) - \cos(\omega))}, \\ \beta_\ell(\omega) &= e^{-\theta\ell^2(N-1)\Delta} \left( e^{-\theta\ell^2\Delta} \cos((N-1)\omega) - \cos(N\omega) \right). \end{aligned}$$

The corresponding derivatives are given by

$$\begin{aligned} \frac{\partial}{\partial\theta}\alpha_\ell(\omega) &= \underbrace{\frac{\cosh(\theta\ell^2\Delta) - 1}{2\theta^{3/2}\ell^2 (\cosh(\theta\ell^2\Delta) - \cos(\omega))}}_{=: a_\ell^1(\omega)} \underbrace{- \frac{\Delta \sinh(\theta\ell^2\Delta) (1 - \cos(\omega))}{\sqrt{\theta} (\cosh(\theta\ell^2\Delta) - \cos(\omega))^2}}_{=: a_\ell^2(\omega)} \quad \text{and} \\ \frac{\partial}{\partial\theta}\beta_\ell(\omega) &= e^{-\theta\ell^2(N-1)\Delta} \ell^2\Delta \left( -N e^{-\theta\ell^2\Delta} \cos((N-1)\omega) + (N-1) \cos(N\omega) \right). \end{aligned}$$

Using the estimates

$$\frac{\cosh(x) - 1}{\cosh(x) - \cos(\omega)} \lesssim \frac{x^2}{x^2 \vee \omega^2}, \quad \frac{x \sinh(x)(1 - \cos(\omega))}{(\cosh(x) - \cos(\omega))^2} \lesssim \frac{x^2}{x^2 \vee \omega^2}$$

in combination with  $\beta_\ell(\omega) \lesssim e^{-\theta\ell^2(N-1)\Delta} ((\ell^2\Delta) \vee \omega)$  and  $b_\ell(\omega) := \frac{\partial}{\partial\theta}\beta_\ell(\omega) \lesssim e^{-\theta\ell^2(N-1)\Delta} \ell^2 ((\ell^2\Delta) \vee \omega)$  shows that any of the three products in

$$\frac{\partial}{\partial\theta}R_N = \sum_{\ell \in \mathcal{I}_k} a_\ell^1 \beta_\ell + a_\ell^2 \beta_\ell + \alpha_\ell b_\ell \tag{52}$$

can be bounded by

$$\sum_{\ell \in \mathcal{I}_k} e^{-\theta \ell^2 (N-1)\Delta} \frac{\ell^4 \Delta^2}{(\ell^4 \Delta^2) \vee \omega^2} ((\ell^2 \Delta) \vee \omega) \leq \Delta \sum_{\ell \in \mathcal{I}_k} e^{-\theta \ell^2 (N-1)\Delta} \ell^2 \lesssim \Delta.$$

Consequently, we have  $\frac{\partial}{\partial \theta} R_N = \mathcal{O}(\Delta)$ , which finishes the proof of (47a).

To prove (47b), we use property (iii) of Lemma A.5 to deduce

$$\frac{\partial}{\partial \theta} \Phi(\omega) \lesssim \omega^2 \Delta \sum_{\ell \in \mathcal{I}_k} \frac{1 + \theta \ell^2 \Delta}{\theta^2 \ell^4 \Delta^2} \lesssim \frac{\omega^2}{\Delta} \left( \frac{1 + \theta k^2 \Delta}{\theta^2 k^4} + \sum_{\ell \geq 1} \frac{1 + \theta (2\ell M)^2 \Delta}{\theta^2 (2\ell M)^4} \right) \lesssim \frac{\omega^2}{k^4 \Delta},$$

where the last step follows from  $k^2 \Delta \leq M^2 \Delta \rightarrow 0$ . Further, using (52),

$$\begin{aligned} & \frac{\partial}{\partial \theta} (R_N(\omega) - R_N(0)) \\ &= \sum_{\ell \in \mathcal{I}_k} a_\ell^1(\omega) (\beta_\ell(\omega) - \beta_\ell(0)) + \sum_{\ell \in \mathcal{I}_k} (a_\ell^1(\omega) - a_\ell^1(0)) \beta_\ell(0) + \sum_{\ell \in \mathcal{I}_k} a_\ell^2(\omega) \beta_\ell(\omega) \\ & \quad + \sum_{\ell \in \mathcal{I}_k} \alpha_\ell(\omega) (b_\ell(\omega) - b_\ell(0)) + \sum_{\ell \in \mathcal{I}_k} (\alpha_\ell(\omega) - \alpha_\ell(0)) b_\ell(0). \end{aligned} \tag{53}$$

Now, by (51), we have

$$\begin{aligned} & \beta_\ell(\omega) - \beta_\ell(0) \\ &= e^{-\theta \ell^2 (N-1)\Delta} ((e^{-\theta \ell^2 \Delta} - 1)(\cos((N-1)\omega) - 1) + \cos((N-1)\omega) - \cos(N\omega)) \\ & \lesssim e^{-\theta \ell^2 (N-1)\Delta} \ell^2 N \omega^2. \end{aligned}$$

In a similar way, using  $\omega \leq k^2 \Delta \leq \ell^2 \Delta$  for  $\ell \in \mathcal{I}_k$ , we can bound

$$\beta_\ell(\omega) \lesssim e^{-\theta \ell^2 (N-1)\Delta} ((\ell^2 \Delta) \vee \omega) \lesssim e^{-\theta \ell^2 (N-1)\Delta} \ell^2 \Delta$$

as well as

$$\begin{aligned} & \beta_\ell(0) \lesssim e^{-\theta \ell^2 (N-1)\Delta} \ell^2 \Delta, \\ & b_\ell(\omega) - b_\ell(0) \lesssim e^{-\theta \ell^2 (N-1)\Delta} \ell^4 N \omega^2, \quad b_\ell(0) \lesssim e^{-\theta \ell^2 (N-1)\Delta} \ell^4 \Delta. \end{aligned}$$

Also,

$$a_\ell^1(\omega) - a_\ell^1(0) \lesssim \frac{1 - \cos(\omega)}{\cosh(\theta \ell^2 \Delta) - \cos(\omega)} \lesssim \frac{1 - \cos(\omega)}{(\cosh(\theta \ell^2 \Delta) - 1)} \lesssim \frac{\omega^2}{k^4 \Delta^2}$$

and similarly,  $\alpha_\ell(\omega) - \alpha_\ell(0) \lesssim \frac{\omega^2}{k^4 \Delta^2}$ ,  $a_\ell^2(\omega) \lesssim \frac{\omega^2}{k^4 \Delta^2}$ ,  $a_\ell^1(\omega) \lesssim 1$  and  $\alpha_\ell(\omega) \lesssim 1$ .

Using the bounds just developed in combination with  $e^{-\theta \ell^2 (N-1)\Delta} \lesssim \frac{1}{k^4 \ell^m}$ ,  $m \in \mathbb{N}$ , shows that any of the five terms in (53) is of order  $\mathcal{O}(\frac{\omega^2}{k^4 \Delta})$  and hence,  $\frac{\partial}{\partial \theta} (R_N(\omega) - R_N(0)) \lesssim \frac{\omega^2}{k^4 \Delta}$ . Now, the proof of (47b) is finalized by

$$\frac{\partial}{\partial \theta} R_N(0)$$



$$\begin{aligned}
 &= \sum_{\ell \in \mathcal{I}_k} e^{-\theta \ell^2 (N-1)\Delta} \frac{2\theta \ell^2 (N-1)\Delta (e^{-\theta \ell^2 \Delta} - 1) + 2\theta \ell^2 \Delta e^{-\theta \ell^2 \Delta} + e^{-\theta \ell^2 \Delta} - 1}{2\theta^{3/2} \ell^2} \\
 &\lesssim \Delta \sum_{\ell \in \mathcal{I}_k} e^{-\theta \ell^2 (N-1)\Delta} \ell^2 \lesssim \Delta k^2 e^{-\theta k^2}. \quad \square
 \end{aligned}$$

**A.3. Bounds on Fourier series and Riemann summation**

The Lemmas in this section provide bounds for Fourier series and Taylor expansions for Riemann sums. Similar results are stated in Lemma 7.2 of [3].

**Lemma A.7.** *Let  $(a_n)$  be a real sequence and  $\tau \in \{\sin, \cos\}$ . Then,*

$$\left| \sum_{k=1}^N a_k \tau(ky) \right| \leq \frac{1 + 2K_N}{y \wedge (2\pi - y)} \sup_{n \leq N} |a_n|$$

holds for any  $y \in (0, 2\pi)$  where  $K_N$  is the number of monotone sections of  $(a_n)_{1 \leq n \leq N}$ .

*Proof.* By Lagrange’s trigonometric identities,

$$\begin{aligned}
 \sum_{k=1}^N \cos(ky) &= \frac{\sin((N + 1/2)y) - \sin(y/2)}{2 \sin(y/2)}, \\
 \sum_{k=1}^N \sin(ky) &= \frac{\cos(y/2) - \cos((N + 1/2)y)}{2 \sin(y/2)},
 \end{aligned}$$

we have  $\left| \sum_{k=M}^N \tau(ky) \right| \leq \frac{1}{\sin(y/2)} \leq \frac{1}{y \wedge (2\pi - y)}$  uniformly in  $M \leq N$ . Therefore,  $\left| \sum_{k=1}^N a_k \tau(ky) \right|$  can be decomposed into

$$\begin{aligned}
 &\left| a_1 \sum_{k=1}^N \tau(ky) + (a_2 - a_1) \sum_{k=2}^N \tau(ky) + (a_3 - a_2) \sum_{k=3}^N \tau(ky) \right. \\
 &\quad \left. + \dots + (a_N - a_{N-1}) \tau(Ny) \right| \\
 &\leq |a_1| \left| \sum_{k=1}^N \tau(ky) \right| + |a_2 - a_1| \left| \sum_{k=2}^N \tau(ky) \right| + |a_3 - a_2| \left| \sum_{k=3}^N \tau(ky) \right| \\
 &\quad + \dots + |a_N - a_{N-1}| |\tau(Ny)| \\
 &\leq \frac{1}{y \wedge (2\pi - y)} \left( |a_1| + \sum_{k=1}^{N-1} |a_{k+1} - a_k| \right) \leq \frac{1 + 2K_N}{y \wedge (2\pi - y)} \sup_{n \leq N} |a_n|,
 \end{aligned}$$

where the last inequality follows from the fact that if  $(a_k)_{N_0 \leq k \leq N_1}$  is monotone for some  $N_0 \leq N_1 \leq N$ , then  $\sum_{k=N_0}^{N_1-1} |a_{k+1} - a_k| = |a_{N_1} - a_{N_0}| \leq 2 \sup_{n \leq N} |a_n|$ .  $\square$

**Lemma A.8.** *Let  $g \in C^1(\mathbb{R}_+)$  be such that  $g'$  is bounded and has a finite number  $K$  of monotone sections. Then, for  $y \in (0, 2\pi)$ , as  $\varepsilon \rightarrow 0$ ,*

$$\begin{aligned} \sum_{k=1}^{\infty} g(k\varepsilon) \cos(ky) &= -\frac{g(0)}{2} + \mathcal{O}\left(\frac{\varepsilon \|g'\|_{\infty}}{(y \wedge (2\pi - y))^2}\right) \\ \sum_{k=1}^{\infty} g(k\varepsilon) \sin(ky) &= \frac{g(0)}{2} \cot\left(\frac{y}{2}\right) + \mathcal{O}\left(\frac{\varepsilon \|g'\|_{\infty}}{(y \wedge (2\pi - y))^2}\right). \end{aligned}$$

*Proof.* We use the formula  $\sin(\alpha) - \sin(\beta) = 2 \cos \frac{\alpha+\beta}{2} \sin \frac{\alpha-\beta}{2}$ ,  $\alpha, \beta \in \mathbb{R}$ , to calculate

$$\begin{aligned} &\frac{g(0)}{2} + \sum_{k=1}^{\infty} g(k\varepsilon) \cos(ky) \\ &= \frac{g(0)}{2} + \frac{1}{2 \sin \frac{y}{2}} \sum_{k=1}^{\infty} g(k\varepsilon) (\sin((k+1/2)y) - \sin((k-1/2)y)) \\ &= \frac{g(0)}{2} - \frac{g(\varepsilon)}{2} + \frac{1}{2 \sin \frac{y}{2}} \sum_{k=1}^{\infty} \sin((k+1/2)y) (g(k\varepsilon) - g((k+1)\varepsilon)) \\ &= -\frac{1}{2} \left( g'(\xi_0^\varepsilon) + \frac{1}{\sin \frac{y}{2}} \sum_{k=1}^{\infty} \sin((k+1/2)y) g'(\xi_k^\varepsilon) \right) \varepsilon \\ &\leq \frac{1 + 2K}{(y \wedge (2\pi - y))^2} \|g'\|_{\infty} \varepsilon, \end{aligned}$$

where  $\xi_k^\varepsilon \in [k\varepsilon, (k+1)\varepsilon]$ . Here, the last step follows from  $\sin((k+1/2)y) = \sin(ky) \cos(y/2) + \cos(ky) \sin(y/2)$  and then applying Lemma A.7. The second statement can be proved analogously, using the relation  $\cos(\alpha) - \cos(\beta) = -2 \sin \left(\frac{\alpha+\beta}{2}\right) \sin \left(\frac{\alpha-\beta}{2}\right)$ ,  $\alpha, \beta \in \mathbb{R}$ .  $\square$

**Lemma A.9.** *Let  $g \in C^2(\mathbb{R}_+) \cap L^1(\mathbb{R}_+)$ ,  $g' \in L^\infty(\mathbb{R}_+)$  and  $g'' \in L^1(\mathbb{R}_+)$ . Then,*

$$\begin{aligned} (i) \quad \varepsilon \sum_{k \geq 1} g(k\varepsilon) &= \int_0^\infty g(z) dz - \frac{g(0)}{2} \varepsilon + \mathcal{O}(\varepsilon^2 \|g''\|_{L^1}), \\ (ii) \quad \varepsilon \sum_{k \geq 1} g(k\varepsilon) \sin^2(ky) &= \frac{1}{2} \int_0^\infty g(z) dz + \mathcal{O}\left(\varepsilon^2 \left(\frac{\|g'\|_{\infty}}{(y \wedge (\pi - y))^2} \wedge \|g''\|_{L^1}\right)\right). \end{aligned}$$

*Proof.* For a detailed proof of (i) we refer to [3, Lemma 7.2]. The main idea is to regard each term  $\varepsilon g(k\varepsilon)$  as a midpoint integral approximation. Since  $\sin^2(y) = (1 - \cos(2y))/2$ , statement (ii) is a direct consequence of (i) and the previous lemma.  $\square$

**Lemma A.10.** *Let  $g \in C^2(\mathbb{R}_+)$  and  $M \rightarrow \infty$ ,  $M\varepsilon \rightarrow 0$ . Then,*

$$\varepsilon \sum_{k=1}^M g(k\varepsilon) = M\varepsilon g(0) + \frac{(M^2 + M)\varepsilon^2}{2} g'(0) + \mathcal{O}((M\varepsilon)^3).$$

*Proof.* First of all, by the midpoint rule there exist  $\eta_k \in [(k-1/2)\epsilon, (k+1/2)\epsilon]$  such that

$$\begin{aligned} \left| \epsilon \sum_{k=1}^M g(k\epsilon) - \int_{\epsilon/2}^{(M+1/2)\epsilon} g(x) dx \right| &= \left| \sum_{k=1}^M \int_{(k-1/2)\epsilon}^{(k+1/2)\epsilon} (g(k\epsilon) - g(x)) dx \right| \\ &\leq \epsilon^3 \sum_{k=1}^M |g''(\eta_k)| \lesssim M^3 \epsilon^3 \end{aligned}$$

and secondly, a Taylor approximation shows that

$$\int_{\epsilon/2}^{(M+1/2)\epsilon} g(x) dx = M\epsilon g(0) + \frac{(M^2 + M)\epsilon^2}{2} g'(0) + \mathcal{O}((M\epsilon)^3). \quad \square$$

## References

- [1] Altmeyer, R. and Reiß, M. (2021). Nonparametric estimation for linear SPDEs from local measurements. *Ann. Appl. Probab.*, 31(1):1–38.
- [2] Bibinger, M. and Trabs, M. (2019). On central limit theorems for power variations of the solution to the stochastic heat equation. In *Stochastic Models, Statistics and Their Applications. Springer Proceedings in Mathematics & Statistics*, volume 294, pages 69–84. [MR4043170](#)
- [3] Bibinger, M. and Trabs, M. (2020). Volatility estimation for stochastic PDEs using high-frequency observations. *Stochastic Process. Appl.*, 130(5):3005–3052. [MR4080737](#)
- [4] Chong, C. (2019). High-frequency analysis of parabolic stochastic PDEs with multiplicative noise: Part I. *arXiv preprint arXiv:1908.04145*. [MR4102691](#)
- [5] Chong, C. (2020). High-frequency analysis of parabolic stochastic PDEs. *Ann. Statist.*, 48(2):1143–1167. [MR4102691](#)
- [6] Cialenco, I. (2018). Statistical inference for SPDEs: an overview. *Stat. Inference Stoch. Process.*, 21(2):309–329. [MR3824970](#)
- [7] Cialenco, I. and Huang, Y. (2020). A note on parameter estimation for discretely sampled SPDEs. *Stoch. Dyn.*, 20(3). [MR4101083](#)
- [8] Cialenco, I. and Kim, H.-J. (2020). Parameter estimation for discretely sampled stochastic heat equation driven by space-only noise revised. *arXiv preprint arXiv:2003.08920*.
- [9] Cont, R. (2005). Modeling term structure dynamics: an infinite dimensional approach. *Int. J. Theor. Appl. Finance*, 8(3):357–380. [MR2144706](#)
- [10] Da Prato, G. and Zabczyk, J. (2014). *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, Cambridge. [MR3236753](#)
- [11] Dacunha-Castelle, D. and Duflo, M. (1986). *Probability and statistics. Vol. II*. Springer-Verlag, Berlin Heidelberg New York. [MR0815650](#)
- [12] Devroye, L., Mehrabian, A., and Reddad, T. (2020). The total variation distance between high-dimensional Gaussians. *arXiv preprint arXiv:1810.08693v5*.

- [13] Dostal, L. (2019). The effect of random wind forcing in the nonlinear Schrödinger equation. *Fluids*, 4(3):121.
- [14] Hildebrandt, F. (2020). On generating fully discrete samples of the stochastic heat equation on an interval. *Statist. Probab. Lett.*, 162. Article 108750. [MR4080720](#)
- [15] Hildebrandt, F. (2021). *Parameter estimation for SPDEs based on discrete observations in time and space*. PhD thesis, Staats-und Universitätsbibliothek Hamburg Carl von Ossietzky. Forthcoming.
- [16] Hottovy, S. and Stechmann, S. N. (2015). A spatiotemporal stochastic model for tropical precipitation and water vapor dynamics. *J. Atmospheric Sci.*, 72(12):4721–4738.
- [17] Huebner, M., Khasminskii, R., and Rozovskii, B. (1993). Two examples of parameter estimation for stochastic partial differential equations. In *Stochastic processes*, pages 149–160. Springer. [MR1427311](#)
- [18] Huebner, M. and Rozovskii, B. L. (1995). On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDE's. *Probab. Theory Related Fields*, 103(2):143–163. [MR1355054](#)
- [19] Ibragimov, I. and Rozanov, Y. (1978). *Gaussian random processes*. Springer-Verlag, Berlin Heidelberg New York. [MR0543837](#)
- [20] Ibragimov, I. A. and Has'minskii, R. Z. (1981). *Statistical estimation*, volume 16 of *Applications of Mathematics*. Springer-Verlag, New York Berlin. Asymptotic theory, Translated from the Russian by Samuel Kotz. [MR0620321](#)
- [21] Isserlis, L. (1918). On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12:134–139.
- [22] Kaino, Y. and Uchida, M. (2021). Parametric estimation for a parabolic linear SPDE model based on discrete observations. *J. Statist. Plann. Inference*, 211:190–220. [MR4153870](#)
- [23] Koski, T. and Loges, W. (1985). Asymptotic statistical inference for a stochastic heat flow problem. *Statist. Probab. Lett.*, 3:185–189. [MR0801688](#)
- [24] Kriz, P. and Maslowski, B. (2019). Central limit theorems and minimum-contrast estimators for linear stochastic evolution equations. *Stochastics*, 0(0):1–32. [MR4023510](#)
- [25] Liu, J. and Tudor, C. A. (2016). Central limit theorem for the solution to the heat equation with moving time. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 19(01):1650005. [MR3474517](#)
- [26] Lototsky, S. V. (2009). Statistical inference for stochastic parabolic equations: a spectral approach. *Publ. Mat.*, 53(1):3–45. [MR2474113](#)
- [27] Mahdi Khalil, Z. and Tudor, C. (2019). Estimation of the drift parameter for the fractional stochastic heat equation via power variation. *Mod. Stoch. Theory Appl.*, 6(4):397–417. [MR4047392](#)
- [28] Markussen, B. (2013). Likelihood inference for a discretely observed stochastic partial differential equation. *Bernoulli*, 9(5):745 – 762. [MR2047684](#)
- [29] Mathai, A. M. and Provost, S. B. (1992). *Quadratic Forms in Random*

- Variables*. Marcel Dekker, inc., New York. [MR1192786](#)
- [30] Neveu, J. (1968). Processus aléatoires gaussiens. Séminaire de mathématiques supérieures. Les Presses de l'Université de Montréal. [MR0272042](#)
- [31] Pakkanen, M. S. (2014). Limit theorems for power variations of ambit fields driven by white noise. *Stochastic Process. Appl.*, 124(5):1942–1973. [MR3170230](#)
- [32] Piterbarg, L. I. and Ostrovskii, A. G. (1997). *Advection and Diffusion in Random Media: Implications for Sea Surface Temperature Anomalies*. Springer Science & Business Media.
- [33] Réveillac, A., Stauch, M., and Tudor, C. A. (2012). Hermite variations of the fractional brownian sheet. *Stoch. Dyn.*, 12(3):1150021. [MR2926578](#)
- [34] Santa Clara, P. and Sornette, D. (2000). The dynamics of the forward interest rate curve with stochastic string shocks. *Rev. Financial Stud.*, 14(1):149–185.
- [35] Shevchenko, R., Slaoui, M., and Tudor, C. A. (2020). Generalized  $k$ -variations and Hurst parameter estimation for the fractional wave equation via Malliavin calculus. *J. Statist. Plann. Inference*, 207:155–180. [MR4066122](#)
- [36] Torres, S., Tudor, C., and Viens, F. (2014). Quadratic variations for the fractional-colored stochastic heat equation. *Electron. J. Probab.*, 19. [MR3256876](#)
- [37] Tsybakov, A. B. (2010). *Introduction to Nonparametric Estimation*. Springer, New York. [MR2724359](#)
- [38] Tuckwell, H. C. (2013). Stochastic partial differential equations in neurobiology: Linear and nonlinear models for spiking neurons. In *Stochastic Biomathematical Models*, pages 149–173. Springer. [MR3051032](#)
- [39] Whittle, P. (1953). The analysis of multiple stationary time series. *J. R. Stat. Soc. Ser. B*, 15:125–139. [MR0056902](#)