HIGH-FREQUENCY ANALYSIS OF PARABOLIC STOCHASTIC PDES

By Carsten Chong

Institut de mathématiques, École Polytechnique Fédérale de Lausanne, carsten.chong@epfl.ch

We consider the problem of estimating stochastic volatility for a class of second-order parabolic stochastic PDEs. Assuming that the solution is observed at high temporal frequency, we use limit theorems for multipower variations and related functionals to construct consistent nonparametric estimators and asymptotic confidence bounds for the integrated volatility process. As a byproduct of our analysis, we also obtain feasible estimators for the regularity of the spatial covariance function of the noise.

1. Introduction. A central objective of stochastic modeling is to capture the fluctuations of a system evolving under the influence of random noise. Being able to quantify the degree of variability and uncertainty in such a system is inevitable for the control and prediction of its future behavior.

In the financial and econometrics literature, a key concept designed to measure and describe the amount of randomness present in the evolution of asset prices, interest rates or other financial indices is that of *stochastic volatility*. Over the past decades, a huge amount of work has been devoted to building stochastic volatility models that are able to reproduce stylized features found in empirical financial data. We only refer to [11] for a comprehensive overview.

Of course, the notion of stochastic volatility is not only limited to mathematical finance. For example, in the literature of turbulence, it is commonly referred to as *intermittency*; see [8, 44, 49] for various models of stochastic intermittency. In a related context, the phenomenon of intermittency has also been intensively studied in the theory of *stochastic partial differential equations* (stochastic PDEs). To be more precise, let us consider a parabolic stochastic PDE of the form

$$(1.1) \qquad \partial_t Y(t, x) = \frac{\kappa}{2} \Delta Y(t, x) - \lambda Y(t, x) + \sigma(t, x) \dot{W}(t, x),$$

where $\kappa > 0$ is a diffusion or viscosity constant, $\lambda > 0$ is a damping rate, σ is a predictable random field and $\dot{W}$ is a Gaussian noise. When we consider (1.1) for $t \geq 0$ and $x \in \mathbb{R}$ with an initial condition at $t = 0$ that is bounded away from 0, it is known from [26] that if σ is a linear function of the solution with sufficient growth, then the solution exhibits a strong mass concentration at large times by forming exponentially large peaks on exponentially small areas. On the other hand, when σ is a bounded function of the solution, this kind of intermittent behavior does not occur. Hence, the knowledge of the form of σ is essential for determining the behavior of the solution Y to (1.1).

Furthermore, in many applications of (1.1), or stochastic PDEs of a similar form, the random field σ models the level of noise that acts on a process described by an otherwise deterministic PDE. Examples include [17] on term structure models, [24, 25] on plankton distribution, [34] on the motion of particles in gravitational fields, [50, 51] on precipitation models and [53, 54] on neuron spikes. In these applications, the knowledge of σ is essential for assessing to which degree the solution to (1.1) deviates from the solution to the deterministic PDE.

1.1. *Objective and related literature.* Motivated from the applications mentioned above, the purpose of the present article is to establish consistent estimators and asymptotic confidence bounds for the random field σ in (1.1), which we henceforth call the *(stochastic) volatility process* (even outside the financial context). To this end, we assume that we are given observations of a single path of the solution $Y(t, x)$ at a finite number of spatial points $x_1, \ldots, x_N \in \mathbb{R}^d$ and at a high number of time points $t = \Delta_n, 2\Delta_n, \ldots, [T/\Delta_n]\Delta_n$ within a finite interval $[0, T]$ with $T < \infty$ ($[\cdot]$ stands for the integer part). Here, Δ_n is a small time step, and we seek estimators of σ with the properties mentioned above when $\Delta_n \to 0$. Hence, our observation scheme has *high frequency* in time and *low resolution* in space. This is a realistic framework for many of the applications mentioned above, where high-frequency recordings are only available at a small number of measuring sites.

The high-frequency analysis of Itô semimartingales has been fully accomplished in the past 10 years; see the treatises [3, 32] for a complete account. For example, given a continuous semimartingale $X(t) = \int_0^t \sigma(s) \, dB(s)$ where B is a Brownian motion and σ a predictable process, the basic idea to estimate σ is to consider *(normalized) power variations* of X, that is,

$$(1.2) \qquad V_p^n(X, t) = \Delta_n \sum_{i=1}^{[t/\Delta_n]} \left| \frac{\Delta_i^n X}{\sqrt{\Delta_n}} \right|^p, \qquad t \in [0, T], p > 0,$$

where $\Delta_i^n X = X(i\Delta_n) - X((i-1)\Delta_n)$ is an *increment* of X from $(i-1)\Delta_n$ to $i\Delta_n$. Under minimal assumptions on σ, one can show that

$$(1.3) \qquad V_p^n(X, t) \overset{\text{u.c.p.}}{\Longrightarrow} V_p(X, t) = \mathbb{E}[|Z|^p] \int_0^t |\sigma(s)|^p \, ds,$$

where $Z \sim N(0, 1)$ and $\overset{\text{u.c.p.}}{\Longrightarrow}$ denotes uniform convergence in probability on compacts; see Theorem 3.4.1 in [32]. Under further regularity assumptions on σ, we have an associated central limit theorem of the form

$$(1.4) \qquad \Delta_n^{-\frac{1}{2}} \left(V_p^n(X, t) - V_p(X, t) \right) \overset{\text{st}}{\Longrightarrow} \mathcal{Z},$$

where $\mathcal{Z}$ is a Gaussian process with independent increments and explicitly known variance, conditionally on σ; see Theorem 5.3.6 in [32]. In (1.4), $\overset{\text{st}}{\Longrightarrow}$ denotes functional stable convergence in law with respect to the uniform topology; see Section 3.2 in [3], Section 2.2 in [32] or [46]. The two results (1.3) and (1.4) can then be used to construct asymptotic confidence bounds for the *integrated volatility process* $\int_0^t |\sigma(s)|^p \, ds$, which is what we understand by "estimating σ" throughout this article.

When we leave the class of semimartingales and consider a moving average process of the form $X_t = \int_{-\infty}^t g(t - s)\sigma(s) \, dB_s$, where g is a kernel that is smooth except at the origin, then the functionals

$$(1.5) \qquad V_p^n(X, t) = \Delta_n \sum_{i=1}^{[t/\Delta_n]} \left| \frac{\Delta_i^n X}{\tau_n} \right|^p, \qquad t \in [0, T], p > 0,$$

with a normalizing factor τ_n depending on the singularity of g at the origin, still satisfy (1.3) and (1.4) (with a slightly larger variance for $\mathcal{Z}$) if g and σ are sufficiently regular; see [4, 5, 18]. In particular, this applies to fractional Brownian motion with Hurst parameter $H < \frac{3}{4}$; see [20, 29].

In the context of stochastic PDEs, estimation problems have been considered by many authors; see [15] for a recent survey. The majority of literature in this respect focuses on the estimation of κ, assuming that σ is constant and known, and that the solution to (1.1), or

certain transformations thereof, is observed continuously in time and/or space; see [15, 36] and the references therein. In practice, of course, measurements are *discrete*, and the amount of literature is much smaller when it comes to estimating σ based on discrete observation schemes.

When σ is a deterministic function of t only, and (1.1) is considered in one spatial dimension on an interval, [48] constructs an estimator for σ based on high-frequency observations in time of a fixed number of Fourier coefficients of Y. But these are solutions to certain stochastic differential equations, and hence semimartingales, so the estimation problem can be fully solved by the techniques of [3, 32]. By contrast, the solution Y itself at fixed spatial positions is *not* a semimartingale (if $\dot{W}$ is a Gaussian space–time white noise and $d = 1$, it has a nontrivial finite quartic variation in time; see [52, 54]). Assuming high-frequency observations of Y in time (as we do in this work), but still with deterministic σ that only depends on t (but not on x), the papers [9, 16] use power variations as in (1.5) with $p = 4$ and $p = 2$, respectively, to establish asymptotic confidence bounds for σ. We also refer to [35] for related results from a probabilistic point of view and to [10] for some extensions of [9].

With stochastic σ, [47] shows a variant of (1.3) with $p = 4$ for the solution Y to (1.1) if $d = 1$ and $\dot{W}$ is a space–time white noise; see also [16, 27]. The papers [6, 43] contain certain limit theorems for space–time moving averages when σ is independent of the noise W (so by conditioning on σ, this reduces to the case of deterministic σ). Apart from these particular cases, to our best knowledge, no further results are available for (1.1), and in particular, no central limit theorems as in (1.4) exist in the case of stochastic σ, and not even a law of large numbers as in (1.3) if $\dot{W}$ is a spatially colored noise.

It is therefore the main objective of this paper to fill this gap and to derive consistent estimators and confidence bounds for the integrated volatility process (with respect to time and for fixed values of x) if σ is a random field. In fact, we consider much more general functionals than (1.5). Given a sufficiently regular evaluation function $f: \mathbb{R}^{N \times L} \to \mathbb{R}^M$ with $L, M \in \mathbb{N}$, we consider *(normalized) variation functionals* of the form $V_f^n(Y_x, t) = (V_f^n(Y_x, t)_1, \ldots, V_f^n(Y_x, t)_M)'$ where

$$(1.6) \qquad V_f^n(Y_x, t)_m = \Delta_n \sum_{i=1}^{[t/\Delta_n]-L+1} f_m\left(\frac{\Delta_i^n Y_x}{\tau_n}, \ldots, \frac{\Delta_{i+L-1}^n Y_x}{\tau_n}\right)$$

for $t \in [0, T]$ and $m = 1, \ldots, M$. Here, Y_x is the N-dimensional process whose jth component is $Y(\cdot, x_j)$, $\Delta_i^n Y_x = Y_x(i\Delta_n) - Y_x((i - 1)\Delta_n) = (\Delta_i^n Y(\cdot, x_1), \ldots, \Delta_i^n Y(\cdot, x_N))'$, and τ_n is a normalizing factor to be introduced in Section 2. The main examples for f are multipowers, which we will study in detail in Section 2.2. The reader may consult [3, 32] and [4, 5, 18] for multipower variations of semimartingales and moving averages, respectively.

1.2. *Results and methodology.* After a short introduction to stochastic PDEs, the two main limit theorems are formulated in Section 2.1. Theorem 2.1 gives a law of large numbers for the functionals in (1.6), while Theorem 2.3 gives the associated central limit theorem at a rate of $\sqrt{\Delta_n}$. Section 2.2 shows how these limit theorems apply to the important example of multipower variations (see Corollary 2.9), which will then be used in Section 2.3 to construct feasible estimators for σ. It turns out that we can even estimate the spatial correlation structure of the noise $\dot{W}$ in (1.1), which we assume to be parametrized by an exponent α. Theorem 2.12 addresses the case of estimating σ when α is known, while Theorems 2.13 and 2.14 propose two estimation procedures for α and Theorem 2.16 one for σ when α is unknown. Our results indicate that this spatial correlation index α plays a very similar role to the kernel smoothness parameter in [4, 5, 18].

The proofs will be given in Section 3. As we already know from the semimartingale case considered in [3, 32], having a stochastic instead of a deterministic volatility process complicates the proofs considerably, in particular for the central limit theorem. For instance, the proofs of [9, 16] do not apply in our context as they make heavy use of the Gaussian distribution of the solution Y to (1.1) when σ is nonrandom. Conceptually, Y can be viewed as a moving average process in space and time; see formula (2.2) below, so it seems natural to transfer techniques from [4, 5, 18] to the stochastic PDE setting. A crucial step in their proofs is to factorize the volatility process out of the stochastic integral by discretizing σ along a subgrid of $\Delta_n, 2\Delta_n, \ldots, [T/\Delta_n]\Delta_n$ *before* showing the actual central limit theorem, and then to prove, using fractional calculus methods from [19], that this discretization only induces an asymptotically negligible error. If one wishes to apply this method to stochastic PDEs, one would have to discretize the volatility process σ both in time and space. Although the heat kernel in (2.2) is concentrated around the origin, in general, this localization is simply not strong enough *in space* on a $\sqrt{\Delta_n}$ rate, which would be needed for the central limit theorem. Thus, we see no way to apply the methods of [4, 5, 18] to (1.1); cf. part (4) of Remark 2.4.

Instead, we will show that a *combination* of the *martingale methods* of [3, 32] (for the discretization part and the identification of the limit law) with *analysis on the Wiener space* as in [4, 5, 18] (for tightness) will give the desired central limit theorem for (1.6). The advantage of this strategy is that we only need to make spatial approximations of σ *after* the actual central limit theorem, where we can use *symmetry* properties of certain measures related to the heat kernel to compensate for its bad spatial concentration properties. Since Y is not a semimartingale, for this method to work, we have to use a complex procedure to approximate $V_f^n(Y_x, t)$ by martingale-type sums in a first step.

Our "martingale proof" also provides an interesting alternative to proving limit theorems for moving average processes in the purely temporal case. For instance, with the new method, the results of [4, 5, 18] can be extended to allow for volatility processes that are semimartingales, which include the majority of stochastic volatility models available in the literature. Moreover, we believe that the martingale techniques we develop in this paper will pave the way for further statistical procedures to estimate spot volatility, to handle measurement errors, or to detect and estimate the density of jumps for stochastic PDEs (and moving average processes). We refer to Chapter III.8 in [3], [31] and [1, 2, 33], respectively, for the corresponding results in the semimartingale framework, which are all proved with martingale techniques.

This paper is accompanied by some Supplementary Material in [14]. All references and numberings starting with a letter, like (A.1) or Lemma B.1; refer to [14], except for Assumptions A and B, which are stated in Section 2.

In what follows, we often write $\iint_a^b = \int_a^b \int_{\mathbb{R}^d}$, $\iint = \iint_{-\infty}^\infty$, $\iiint_a^b = \int_a^b \int_{\mathbb{R}^d} \int_{\mathbb{R}^d}$, and $\iiint = \iiint_{-\infty}^\infty$. Moreover, $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$.

2. Model and main results. On a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}}, \mathbb{P})$ satisfying the usual conditions, we consider the stochastic PDE (1.1) for $t \in \mathbb{R}$ and $x \in \mathbb{R}^d$ driven by a zero-mean $(\mathcal{F}_t)_{t \in \mathbb{R}}$-Gaussian noise $\dot{W}$ which is white in time but possibly colored in space. More precisely, we have an L^2-valued centered Gaussian measure $W(A)$, indexed by bounded Borel sets $A \in \mathcal{B}_b(\mathbb{R}^{d+1})$, such that $W(A)$ is independent of $\mathcal{F}_t$ if $A \cap ((-\infty, t] \times \mathbb{R}^d) = \varnothing$, and such that for any $A_1, A_2 \in \mathcal{B}_b(\mathbb{R}^{d+1})$,

$$\mathbb{E}[W(A_1)W(A_2)] = \iiint \mathbb{1}_{A_1}(s, y)\mathbb{1}_{A_2}(s, z)\Lambda(\mathrm{d}y, \mathrm{d}z)\,\mathrm{d}s.$$

In this paper, we assume that $\Lambda(\mathrm{d}y, \mathrm{d}z) = F(z - y)\,\mathrm{d}y\,\mathrm{d}z$ where F is the *Riesz kernel* $F(x) = c_\alpha |x|^{-\alpha}$ for some $\alpha \in (0, d \wedge 2)$ and the normalizing constant is given by $c_\alpha =$

$\pi^{d/2-\alpha}\Gamma(\frac{\alpha}{2})/\Gamma(\frac{d-\alpha}{2})$. Here and throughout the paper, $|x|$ denotes the Euclidean norm. In dimension 1, we also allow for the case where $\dot{W}$ is a *Gaussian space–time white noise*, which corresponds to $\Lambda(\mathrm{d}y, \mathrm{d}z) = \delta_z(\mathrm{d}y)\,\mathrm{d}z$ and $F(x) = \delta_0(x)$. We set $\alpha = 1$ and $c_\alpha = 1$ in this case. In dimensions $d \geq 2$, it is well known that no function-valued solution to (1.1) exists if $\dot{W}$ is a space–time white noise, or if $2 \leq \alpha < d$; see [22].

By the classical integration theory of [55] (see also [22, 42] for extensions), an Itô integral against W can be constructed for integrands from the space $\mathcal{L}(W)$ of predictable random fields $\phi \colon \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ satisfying

$$\mathbb{E}\left[\iiint |\phi(s, y)\phi(s, z)|\Lambda(\mathrm{d}y, \mathrm{d}z)\,\mathrm{d}s\right] < \infty.$$

In particular, as soon as the predictable random field σ satisfies

$$(2.1) \qquad \sup_{(t,x)\in\mathbb{R}\times\mathbb{R}^d} \mathbb{E}[\sigma^2(t, x)] < \infty,$$

the stochastic PDE (1.1) for $(t, x) \in \mathbb{R} \times \mathbb{R}^d$ admits a *mild solution* given by

$$(2.2) \qquad Y(t, x) = \iint_{-\infty}^t G(t - s, x - y)\sigma(s, y)W(\mathrm{d}s, \mathrm{d}y), \quad (t, x) \in \mathbb{R} \times \mathbb{R}^d,$$

where

$$(2.3) \qquad G(t, x) = G_x(t) = (2\pi\kappa t)^{-\frac{d}{2}} e^{-\frac{|x|^2}{2\kappa t} - \lambda t} \mathbb{1}_{t>0}, \quad (t, x) \in \mathbb{R} \times \mathbb{R}^d,$$

is the heat kernel for (1.1). We remark that although the integration theory in [22, 42, 55] is developed for $t \geq 0$, their results extend without any change to the case $t \in \mathbb{R}$. Also, while all our results are formulated for (1.1) with $t \in \mathbb{R}$, they remain valid if (1.1) is considered for $t \geq 0$ as soon as the initial condition at $t = 0$ is sufficiently regular; see Remark 2.7 below.

As soon as σ is jointly stationary with the increments of W, the mild solution in (2.2) is stationary in space and time. In particular, if $\sigma \equiv 1$, all components of $\Delta_i^n Y_x$ are normally distributed with mean 0 and variance

$$(2.4) \qquad \begin{aligned} \tau_n^2 &= \mathbb{E}\left[\left|\iint_{-\infty}^{i\Delta_n} (G_{x_j-y}(i\Delta_n - s) - G_{x_j-y}((i-1)\Delta_n - s))W(\mathrm{d}s, \mathrm{d}y)\right|^2\right] \\ &= \iiint_0^\infty (G_y(s) - G_y(s - \Delta_n))(G_z(s) - G_z(s - \Delta_n))\Lambda(\mathrm{d}y, \mathrm{d}z)\,\mathrm{d}s, \end{aligned}$$

which depends neither on i nor on j. This will be the normalizing factor we choose in (1.6) so that $V_f^n(Y_x, t)$ is typically "of order 1" and we may hope for convergence as $n \to \infty$. An explicit formula for τ_n can be found in Lemma B.1 in the Supplementary Material [14].

2.1. *Limit theorems for normalized variation functionals.*

A first-order limit theorem for the normalized variation functionals (1.6) can be shown under mild assumptions on f and σ. In what follows, the Euclidean norm $|z|$ for some matrix $z \in \mathbb{R}^{N \times L}$ is defined by viewing z as an element of $\mathbb{R}^{NL}$.

ASSUMPTION A. There exists $p \geq 2$ with the following properties:

A1. The function $f \colon \mathbb{R}^{N \times L} \to \mathbb{R}^M$ is continuous and satisfies $f(z) = o(|z|^p)$ as $|z| \to \infty$.
A2. For some $\epsilon > 0$, we have

$$\sup_{(t,x)\in\mathbb{R}\times\mathbb{R}^d} \mathbb{E}[|\sigma(t, x)|^{p+\epsilon}] < \infty.$$

A3. The random field σ is uniformly L^2-continuous on $\mathbb{R} \times \mathbb{R}^d$: as $\epsilon \to 0$,

$$w(\epsilon) = \sup\{\mathbb{E}[|\sigma(t,x) - \sigma(s,y)|^2]: |t-s| + |x-y| < \epsilon\} \to 0.$$

The next theorem is our first main result. If $(X^n(t))_{t \geq 0}$ and $(X(t))_{t \geq 0}$ are stochastic processes, we write $X^n \overset{L^1}{\Longrightarrow} X$ or $X^n(t) \overset{L^1}{\Longrightarrow} X(t)$ if for every $T > 0$, we have $\mathbb{E}[\sup_{t \in [0,T]} |X^n(t) - X(t)|] \to 0$ as $n \to \infty$.

THEOREM 2.1 (Law of large numbers). *Under Assumption A, we have*

$$(2.5) \qquad V_f^n(Y_x, t) \overset{L^1}{\Longrightarrow} V_f(Y_x, t) = \int_0^t \mu_f(\sigma^2(s, x_1), \dots, \sigma^2(s, x_N)) \, ds.$$

Here, $\mu_f \colon \mathbb{R}^N \to \mathbb{R}^M$ is the function given by $\mu_f(v_1, \dots, v_N) = \mathbb{E}[f(Z)]$, where $Z = (Z_{jk})_{j,k=1}^{N,L}$ is multivariate normal with mean 0 and

$$(2.6) \qquad \mathrm{Cov}(Z_{j_1 k_1}, Z_{j_2 k_2}) = \Gamma_{|k_1 - k_2|} v_j \mathbb{1}_{j_1 = j_2 = j}$$

and

$$(2.7) \qquad \Gamma_0 = 1, \Gamma_r = \frac{1}{2}((r+1)^{1-\frac{\alpha}{2}} - 2r^{1-\frac{\alpha}{2}} + (r-1)^{1-\frac{\alpha}{2}}), \quad r \geq 1.$$

REMARK 2.2. Fix some $m = 1, \dots, M$. If f_m only depends on the variables $(z_{jk}: j \in J, k = 1, \dots, L)$, where $J \subseteq \{1, \dots, N\}$, that is, if $V_f^n(Y_x, t)_m$ only uses the increments observed at the points $(x_j: j \in J)$, then its limit in (2.5) will only depend on $(\sigma(\cdot, x_j): j \in J)$. In other words, by taking measurements at x_j, one can obtain isolated information about $\sigma(\cdot, x_j)$, independently from the values of σ at all other positions.

In order to obtain a central limit theorem for (2.5), we need to put stronger regularity assumptions on f and σ, which is already necessary for semimartingales (cf. [3, 32]) and moving averages (cf. [4, 5, 18]). In the first two references, σ itself has to be a semimartingale, while in the next three references, σ has to be (essentially) Hölder continuous with exponent $> \frac{1}{2}$. We will assume that σ has one of these two properties plus additional regularity in space.

ASSUMPTION B.

B1. The function $f \colon \mathbb{R}^{N \times L} \to \mathbb{R}^M$ is even [i.e., we have $f(z) = f(-z)$ for all $z \in \mathbb{R}^{N \times L}$] and four times continuously differentiable. Moreover, there are $p \geq 2$ and $C > 0$ such that

$$|f_m(z)| \leq C(1 + |z|^p), \qquad \left|\frac{\partial}{\partial z_\alpha} f_m(z)\right| \leq C(1 + |z|^{p-1}),$$

$$\left|\frac{\partial^2}{\partial z_\alpha \partial z_\beta} f_m(z)\right| + \left|\frac{\partial^3}{\partial z_\alpha \partial z_\beta \partial z_\gamma} f_m(z)\right| + \left|\frac{\partial^4}{\partial z_\alpha \partial z_\beta \partial z_\gamma \partial z_\delta} f_m(z)\right| \leq C(1 + |z|^{p-2})$$

for all $m \in \{1, \dots, M\}$ and $\alpha, \beta, \gamma, \delta \in \{1, \dots, N\} \times \{1, \dots, L\}$.

B2. If F is the Riesz kernel with $0 < \alpha < 1$, or $\alpha = 1$ and $\dot{W}$ is *not* a space–time white noise, we assume for each $m = 1, \dots, M$ that $f_m(z)$ only depends on $z_{j1}, \dots, z_{jL}$ for some $j = j(m) \in \{1, \dots, N\}$.

B3. The volatility process σ takes the form

$$(2.8) \qquad \sigma(t,x) = \sigma^{(0)}(t,x) + \iint_{-\infty}^t K(t-s, x-y)\rho(s,y)W'(ds, dy)$$

for $(t,x) \in \mathbb{R} \times \mathbb{R}^d$, with the following specifications:

- $\sigma^{(0)}$ is a predictable process satisfying

$$(2.9) \qquad \sup_{(t,x)\in\mathbb{R}\times\mathbb{R}^d} \mathbb{E}\big[|\sigma^{(0)}(t,x)|^{2p+\epsilon}\big] < \infty$$

for some $\epsilon > 0$, and

$$(2.10) \qquad \sup_{x\in\mathbb{R}^d} \mathbb{E}\big[|\sigma^{(0)}(t,x) - \sigma^{(0)}(s,x)|^{2p}\big]^{\frac{1}{2p}} \le C'|t-s|^{\gamma}$$

for some $\gamma \in (\frac{1}{2}, 1]$ and $C' > 0$. In addition, for every $t \in \mathbb{R}$, the mapping $x \mapsto \sigma^{(0)}(t,x)$ is almost surely twice differentiable such that for $j, k = 1, \ldots, d$,

$$(2.11) \qquad \sup_{(t,x)\in\mathbb{R}\times\mathbb{R}^d} \mathbb{E}\left[\left|\frac{\partial}{\partial x_j}\sigma^{(0)}(t,x)\right|^p + \left|\frac{\partial^2}{\partial x_j \partial x_k}\sigma^{(0)}(t,x)\right|^p\right] < \infty.$$

- W' is an $(\mathcal{F}_t)_{t\in\mathbb{R}}$-Gaussian noise that is white in time and possibly colored in space [such that (W, W') is bivariate Gaussian] with spatial covariance $\Lambda'(dy, dz) = F'(z - y)\,dy\,dz$, where F' is the Riesz kernel with some $\alpha' \in (0, 2) \cap (0, d]$ (including the possibility that $F' = \delta_0$ when $d = 1$).
- $K: [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ is a kernel such that the partial derivatives $\frac{\partial}{\partial t}K$, $\frac{\partial}{\partial x_j}K$, $\frac{\partial^2}{\partial x_j \partial x_k}K$, and $\frac{\partial^3}{\partial x_j \partial x_k \partial x_l}K$ exist and belong to $\mathcal{L}(W')$ for all $j, k, l = 1, \ldots, d$.
- ρ is a predictable process satisfying the same moment condition (2.9) as $\sigma^{(0)}$, and furthermore, for some $\epsilon' > 0$ and $C'' > 0$,

$$(2.12) \qquad \sup_{x\in\mathbb{R}^d} \mathbb{E}\big[|\rho(t,x) - \rho(s,x)|^{2p+\epsilon}\big]^{\frac{1}{2p+\epsilon}} \le C''|t-s|^{\epsilon'}.$$

For our second main result, we use $\overset{\text{st}}{\Longrightarrow}$ to denote functional stable convergence in law in the space of càdlàg functions $[0, \infty) \to \mathbb{R}^M$, equipped with the local uniform topology, while stable convergence in law between finite-dimensional random variables will be denoted by $\overset{\text{st}}{\longrightarrow}$. We refer the reader to [3, 32] for a definition of this mode of convergence and also for the definition of a *very good filtered extension* of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$. The only property of stable convergence in law we need is the following (see, e.g., Proposition 2(i) in [46]):

$$(2.13) \qquad X_n \overset{\text{st}}{\longrightarrow} X, \qquad Y_n \overset{\mathbb{P}}{\longrightarrow} Y \quad \Longrightarrow \quad (X_n, Y_n) \overset{\text{st}}{\longrightarrow} (X, Y).$$

Since the limiting objects in (2.5) are random, this will allow us to studentize (2.14) below and obtain feasible confidence bounds for σ. Just convergence in law, of course, will not suffice for this purpose.

THEOREM 2.3. *Under Assumption* B, *we have as $n \to \infty$,*

$$(2.14) \qquad \Delta_n^{-\frac{1}{2}}\big(V_f^n(Y_x, t) - V_f(Y, t)\big) \overset{\text{st}}{\Longrightarrow} \mathcal{Z},$$

where $(\mathcal{Z}(t) = (\mathcal{Z}_1(t), \ldots, \mathcal{Z}_M(t))')_{t\ge0}$ is a continuous process defined on a very good filtered extension $(\overline{\Omega}, \overline{\mathcal{F}}, (\overline{\mathcal{F}}_t)_{t\ge0}, \overline{\mathbb{P}})$ of the original probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$, which, conditionally on the σ-field $\mathcal{F}$, is a centered Gaussian process with independent increments such that the covariance function $\mathcal{C}_{m_1 m_2}(t) = \overline{\mathbb{E}}[\mathcal{Z}_{m_1}(t)\mathcal{Z}_{m_2}(t)|\mathcal{F}]$, for $m_1, m_2 =$

1, ..., M, is given by

$$\mathcal{C}_{m_1 m_2}(t) = \int_0^t \rho_{f_{m_1}, f_{m_2}}(0; \sigma^2(s, x_1), \ldots, \sigma^2(s, x_N)) \, ds$$

(2.15)
$$+ \sum_{r=1}^\infty \int_0^t \rho_{f_{m_1}, f_{m_2}}(r; \sigma^2(s, x_1), \ldots, \sigma^2(s, x_N)) \, ds$$

$$+ \sum_{r=1}^\infty \int_0^t \rho_{f_{m_2}, f_{m_1}}(r; \sigma^2(s, x_1), \ldots, \sigma^2(s, x_N)) \, ds.$$

In the last line, for $r \in \mathbb{N}_0$, we define

(2.16)
$$\rho_{f_{m_1}, f_{m_2}}(r; v_1, \ldots, v_N) = \text{Cov}(f_{m_1}(Z^{(1)}), f_{m_2}(Z^{(2)})),$$

where $Z^{(1)} = (Z_{jk}^{(1)})_{j,k=1}^{N,L}$ and $Z^{(2)} = (Z_{jk}^{(2)})_{j,k=1}^{N,L}$ are jointly Gaussian, both with the same law as the matrix Z in Theorem 2.1 *and cross-covariances*

(2.17)
$$\text{Cov}(Z_{j_1 k_1}^{(1)}, Z_{j_2 k_2}^{(2)}) = \Gamma_{|k_1 - k_2 + r|} v_j \mathbb{1}_{j_1 = j_2 = j}.$$

Part of the statement is that the series in (2.15) *converge in the L^1-sense.*

REMARK 2.4. Let us comment on the assumptions of Theorem 2.3.

(1) Assumption B1 can be relaxed by allowing, for example, f to be continuous but not differentiable at $z = 0$, very similar to [4] or Chapter 11.2 in [32]. Due to the technical proofs already needed under the stronger assumptions, we refrain from doing so in this paper.

(2) Increments at different measurements sites $x_j \neq x_{j'}$ contribute in the limit $n \to \infty$ independently to the right-hand side of (2.5); see Remark 2.2. However, in the cases specified in Assumption B2, the correlation between two such increments at different locations decays in general at a slower rate than $\sqrt{\Delta_n}$. So for Theorem 2.3 to hold, we must assume in these cases that each coordinate of f uses increments at no more than one measurement site. The symmetry assumption on f is standard and already needed in the semimartingale context in order to avoid an asymptotic bias; see Theorem 5.3.6 in [32].

(3) Assumption B3 on the temporal regularity of σ is the "union" of two typical cases considered in the literature. The part $\sigma^{(0)}$ is (essentially) Hölder continuous of order strictly larger than $\frac{1}{2}$, as considered, for instance, in [4, 5, 18]. If one wants to include volatility processes that are of the roughness of Brownian motion, one has to make further structural assumptions as in [3, 32], namely, that $\sigma^{(1)} = \sigma - \sigma^{(0)}$ is a semimartingale. As we will show in Lemma A.1, (2.8) is one possibility to obtain such a semimartingale structure, jointly in x.

(4) The volatility process must also have nice regularity in space. In fact, we assume that σ is pathwise twice differentiable in space. By Theorem 2.1, the limit of variation functionals taken at one measurement site only depends on the volatility at this site. However, this spatial concentration at the origin, which is due to the properties of the heat kernel, is very weak, so we must use the differentiability assumption *and* the symmetry of the heat kernel (notice, however, Remark 2.6) to obtain a localization at a faster rate than $\sqrt{\Delta_n}$. For more details, we refer the reader to Remark D.2 in the Supplementary Material. Examples of volatility models for $\sigma^{(1)}$ include Ornstein–Uhlenbeck processes in space and time (see [8, 38]) and their generalizations (see [12, 45]). If σ does not depend on x as in [9], then (2.11) is clearly satisfied.

REMARK 2.5. In principle, the results and techniques developed in this paper apply to more general equations than (1.1), or more general kernels G in (2.2) and spatial covariance functions F of the noise (for the existence of a Gaussian noise with $\Lambda(dy, dz) =$

$F(z - y) \, dy \, dz$, it is necessary and sufficient that the function F be the Fourier transform of a nonnegative tempered measure on $\mathbb{R}^d$; see Section 2 of [22] for more details).

As the proof shows, the law of large numbers (Theorem 2.1) continues to hold as long as $G \in \mathcal{L}(W)$ has a dominating singularity at the origin such that, for some $\Gamma_r \in \mathbb{R}$, the measures defined in (3.1) with the new G and F satisfy $\Pi_{r,0}^n \xrightarrow{\text{w}} \Gamma_r \delta_0$ and $|\Pi_{r,h}^n| \xrightarrow{\text{w}} 0$ for all $r = 0, 1, \ldots$ and $h \neq 0$.

For the central limit theorem (Theorem 2.3), we additionally need the following assumptions [expressed in terms of $\Pi_{r,h}^n$ and $|\Pi_{r,h}^n|$ from (3.1)]:

(1) G is symmetric in the sense that $G(t, x) = G(t, -x)$ for any $t > 0$ and $x \in \mathbb{R}^d$. In addition, for every $r \in \mathbb{N}_0$, one has, as $n \to \infty$,

$$\iiint_0^\infty \big(|y|^2 + |z|^2\big)|\Pi_{r,0}^n|(ds, dy, dz) = o(\sqrt{\Delta_n}).$$

(2) For every $r \in \mathbb{N}_0$, we have, as $n \to \infty$,

$$\big|\Pi_{r,0}^n([0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d) - \Gamma_r\big| = o(\sqrt{\Delta_n}).$$

Moreover, either f satisfies Assumption B2, or for all $r \in \mathbb{N}_0$ and $h \neq 0$,

$$\big|\Pi_{r,h}^n\big|([0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d) = o(\sqrt{\Delta_n}).$$

(3) There is some decreasing square-summable sequence $(\overline{\Gamma}_r : r \in \mathbb{N}_0)$ such that for all $n \in \mathbb{N}$, $r \in \mathbb{N}_0$, and $h \in \mathbb{R}^d$,

$$\big|\Pi_{r,h}^n\big|([0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d) \leq \overline{\Gamma}_r.$$

(4) There is $\nu > 1$ such that for all $\theta \in (0, 1)$,

$$\big|\Pi_{0,0}^n\big|\big((\Delta_n^{1-\theta}, \infty) \times \mathbb{R}^d \times \mathbb{R}^d\big) = O\big(\Delta_n^{\nu\theta}\big).$$

Condition (1) is needed for the reasons explained in part (4) of Remark 2.4 but can be relaxed; see Remark 2.6. Condition (2) is used for the terms $K_3^{n,i}$, $K_5^{n,i}$ and $K_6^{n,i}$ in the proof of Lemma 3.17. Condition (3) is crucial for the actual central limit theorem in Proposition 3.11 (if the asymptotic covariances of the increments fail to be square-summable, there is no hope to see a central limit theorem; cf. [39]). Finally, condition (4) is needed for nearly all approximations in the proof.

If the kernel has singularity fronts as, for example, in the case of the wave equation, the limits in (1.6) will have a different shape, which we shall discuss in a separate work.

REMARK 2.6. In the setting of Remark 2.5, the symmetry assumption on G can be weakened. Suppose that $G(t, x) = \widetilde{G}(t, x)H(x)$ where $\widetilde{G}$ satisfies condition (1) in Remark 2.5, and $H : \mathbb{R}^d \to \mathbb{R}$ is differentiable such that

$$(2.18) \qquad \sup_{i=1,\ldots,d} \sup_{u \in [0,1]} \bigg|\frac{\partial}{\partial x_i} H(ux)\bigg| + \sup_{u \in [0,1]} |H(ux)| \leq H_0(x)$$

for some function $H_0 : \mathbb{R}^d \to [0, \infty)$.

In the proof of Theorem 2.3, the symmetry of G is only used to show the identity (D.54). In the asymmetric case, we observe from (3.1) that $\Pi_{r,0}^n(ds, dy, dz) = H(y)H(z)\widetilde{\Pi}_{r,0}^n(ds, dy, dz)$ for all $r \in \mathbb{N}_0$, where $\widetilde{\Pi}_{r,h}^n$ is the measure that arises from the first equation in (3.1) when we replace G by $\widetilde{G}$. By the mean value theorem, applied to $(y, z) \mapsto H(y)H(z)$, and property (2.18), the left-hand side of (D.54) is bounded by a constant times

$$H(0)^2\bigg|\iiint_0^{(\lambda_n + (k \vee k'))\Delta_n} (y_l + z_l)\widetilde{\Pi}_{|k'-k|,0}^n(ds, dy, dz)\bigg|$$

$$+ \iiint_0^\infty |y_l + z_l|H_0(y)H_0(z)(|y| + |z|)\big|\widetilde{\Pi}_{|k'-k|,0}^n\big|(ds, dy, dz).$$

The first term is zero by symmetry. Thus, if we impose the condition

$$(2.19) \qquad \iiint_0^\infty (|y|^2 + |z|^2) H_0(y) H_0(z) |\widetilde{\Pi}_{r,0}^n| (\mathrm{d}s, \mathrm{d}y, \mathrm{d}z) = o(\sqrt{\Delta_n})$$

as $n \to \infty$ for all $r \in \mathbb{N}_0$, we no longer need G to be symmetric in Remark 2.5.

Let us apply the previous discussion to the important example where $\widetilde{G}$ is the heat kernel (2.3), and $H(x) = e^{\theta \cdot x}$ for some $\theta \in \mathbb{R}^d$, which corresponds to (1.1) with an additional gradient term (considered, in similar forms, by [9] and many of the applications mentioned in the Introduction). It is easily verified that the resulting kernel G belongs to $\mathcal{L}(W)$ (and hence, a stationary mild solution exists in the case of constant σ) if and only if $\lambda > \frac{\kappa}{2} |\theta|^2$. Under this additional constraint, one can then show that (2.19) holds with $H_0(x) = (1 + |\theta|)(e^{\theta \cdot x} \vee 1)$ in (2.18). Indeed, by symmetry considerations, the left-hand side of (2.19) is bounded by a constant times

$$\iiint_0^\infty (|y|^2 + |z|^2) e^{\theta \cdot y} e^{\theta \cdot z} |\widetilde{\Pi}_{r,0}^n| (\mathrm{d}s, \mathrm{d}y, \mathrm{d}z),$$

which equals the left-hand side of (B.19). Using the identities $G(t, x) = \widetilde{G}(t, x) e^{\theta \cdot x}$ and $G(t, x) = \widetilde{G}(t, x - \kappa \theta t) e^{\kappa |\theta| t / 2}$, it is not difficult to see that (B.10) and (B.21) remain valid, as well as (B.12), (B.17) and (B.18) if we replace λ by $\lambda_0 = \lambda - \frac{\kappa}{2} |\theta|$. One can now follow the arguments given in the proof of Lemma B.3 in order to show that (B.19), and hence (2.19) holds true.

REMARK 2.7. All results in this section remain valid if we consider (1.1) for $t > 0$ and $x \in \mathbb{R}^d$, subject to some bounded and sufficiently regular initial condition y_0 at time $t = 0$. Indeed, the mild solution is then given by

$$(2.20) \qquad Y(t, x) = \int_{\mathbb{R}^d} G(t, x - y) y_0(y) \, \mathrm{d}y + \iint_0^t G(t - s, x - y) \sigma(s, y) W(\mathrm{d}s, \mathrm{d}y)$$

for $(t, x) \in (0, \infty) \times \mathbb{R}^d$. Let us denote the first and the second term by $Y^{(0)}$ and $Y^{(1)}$, respectively, and fix $T > 0$. Under the hypothesis that y_0 is Hölder continuous with some exponent $> 1 - \frac{\alpha}{2}$ (resp., differentiable with a derivative that is Hölder continuous with some exponent $> 1 - \frac{\alpha}{2}$), we know from classical PDE theory (see Theorem 5.1.2(ii) in [37]) that $t \mapsto Y^{(0)}(t, x)$ is Hölder continuous on $[0, T]$ with some exponent $\eta > \frac{1}{2} - \frac{\alpha}{4}$ (resp., $\eta > 1 - \frac{\alpha}{4}$), uniformly for $x \in \mathbb{R}^d$. In particular, by (B.3), we have $|\Delta_i^n Y_x^{(0)} / \tau_n| \lesssim \Delta_n^\eta / \tau_n \lesssim \Delta_n^{\eta - 1/2 + \alpha/4}$, where the last exponent is strictly positive (resp., larger than $\frac{1}{2}$). From this, it is straightforward to deduce that the contribution of $Y^{(0)}$ to (1.6) is asymptotically negligible in Theorem 2.1 (resp., Theorem 2.3).

We are left to show that $Y^{(1)}$ has the same asymptotic behavior as the expression in (2.2). For the law of large numbers, this is straightforward, while for the central limit theorem, it can be proved analogously to Step 1 in Section 3.2. The details are omitted at this point. We further remark that the assumption $\lambda > 0$ is superfluous when (1.1) is considered for $t \geq 0$, and it is sufficient to formulate Assumptions A2 and A3 as well as Assumption B3 with $t \in [0, T]$ for any $T > 0$ instead of $t \in \mathbb{R}$, and we may replace $-\infty$ in (2.8) by 0.

REMARK 2.8. In the literature of stochastic PDEs, one often considers equations where the random field σ is an explicit functional of the solution Y, that is, equation (2.20) where $\sigma = B(Y)$ and B is an operator satisfying certain regularity and growth conditions such that

(2.20) admits a mild solution; see [21]. While Assumption A is relatively weak in this situation (e.g., it is satisfied if

$$B(Y)(t, x) = b(Y(t, x)),$$

(2.21)
$$B(Y)(t, x) = \iint_0^t H(t - s, x - y) Y(s, y) \, dy \, ds \quad \text{or}$$

$$B(Y)(t, x) = \iint_0^t K(t - s, x - y) Y(s, y) W(ds, dy),$$

and b is Lipschitz continuous, $H \in L^1([0, T] \times \mathbb{R}^d)$, and $K \mathbb{1}_{[0,T]} \in \mathcal{L}(W)$ for all $T > 0$), Assumption B is more restrictive. In fact, for the functionals in (2.21), it is only satisfied if b is constant, or if $B(Y)$ is the second or third expression with functions H and K that are sufficiently smooth.

2.2. *Multipower variations.* We apply Theorems 2.1 and 2.3 to an important class of functionals, namely to so-called *multipower variations* $V_\Phi^n(Y_x, t)$ or *signed multipower variations* $V_\Psi^n(Y_x, t)$, where $\Phi, \Psi \colon \mathbb{R}^{N \times L} \to \mathbb{R}^N$ (note that $N = M$) are given by

(2.22)
$$\Phi_m(z) = \Phi_m\big((z_{jk})_{j,k=1}^{N,L}\big) = \prod_{k=1}^L |z_{mk}|^{w_{mk}},$$

(2.23)
$$\Psi_m(z) = \Psi_m\big((z_{jk})_{j,k=1}^{N,L}\big) = \prod_{k=1}^L (z_{mk})^{w_{mk}}, \quad m = 1, \dots, N,$$

with $w_{mk} \geq 0$ in (2.22) and $w_{mk} \in \mathbb{N}_0$ in (2.23). We shall write $w = (w_{mk})_{m,k=1}^{N,L}$, $w_m = w_{m1} + \cdots + w_{mL}$ and $\overline{w} = \max\{w_1, \dots, w_N\}$. If we want to emphasize the dependence on w, we write $\Phi(z) = \Phi(w; z)$ and $\Psi(z) = \Psi(w; z)$. If $w_{mk} = p_k$ for all m and k, we write

(2.24)
$$\Phi(p_1, \dots, p_L; z) = \Phi(w; z) \quad \text{and} \quad \Psi(p_1, \dots, p_L; x) = \Psi(w; z).$$

For multipowers, Theorems 2.1 and 2.3 take the following form.

COROLLARY 2.9. *Assume that $\alpha \in (0, 2) \cap (0, d]$.*

(1) *If Assumptions A2 and A3 hold with $p = \overline{w} \vee 2$, then we have*

(2.25)
$$V_{\Phi_m|\Psi_m}^n(Y_x, t) \overset{L^1}{\Longrightarrow} \mu_{\Phi_m|\Psi_m} \int_0^t |\sigma(s, x_m)|^{w_m} \, ds$$

for all $m = 1, \dots, N$, where $\mu_f = \mu_f(1, \dots, 1)$ (as defined in Theorem 2.1) and $\Phi_m|\Psi_m$ means that we can either take Φ_m or Ψ_m in (2.25). Note that $\mu_{\Psi_m} = 0$ if w_m is odd.

(2) *Suppose that $w_{mk} \in \{0, 2\}$ or $w_{mk} \geq 4$ in the case of (2.22), and that all w_m are even in the case of (2.23). Further assume that Assumption B3 holds with $p = \overline{w}$. Then (2.14) holds for $f = \Phi|\Psi$, and the $\mathcal{F}$-conditional covariance processes in (2.15) are given by*

(2.26)
$$C_{m_1 m_2}(t) = \begin{cases} \rho_{\Phi_m|\Psi_m} \int_0^t |\sigma(s, x_m)|^{2w_m} \, ds, & m_1 = m_2 = m, \\ 0, & m_1 \neq m_2, \end{cases}$$

where $\rho_f = \rho_{f,f}(0; 1, \dots, 1) + 2 \sum_{r=1}^\infty \rho_{f,f}(r; 1, \dots, 1)$ (as defined in Theorem 2.3).

Because of their particular importance in high-frequency statistics, we further specialize Corollary 2.9 to the normalized power variations

(2.27)
$$V_p^n(Y_x, t) = \big(V_p^n(Y_x, t)_m\big)_{m=1}^N = \left(\Delta_n \sum_{i=1}^{[t/\Delta_n]} \left| \frac{\Delta_i^n Y(\cdot, x_m)}{\tau_n} \right|^p \right)_{m=1}^N,$$

where $p > 0$ and which corresponds to the special case $L = 1$ and the function $\Phi(p, \cdot)$ in (2.24).

COROLLARY 2.10.   *Assume that* $\alpha \in (0, 2) \cap (0, d]$.

(1) *Let* $p > 0$ *and* σ *be a predictable random field satisfying Assumptions* A2 (*with* $p \vee 2$ *instead of* p) *and* A3. *Then for every* $m = 1, \ldots, N$,

$$(2.28) \qquad V_p^n(Y_x, t)_m \overset{L^1}{\Longrightarrow} \mu_p \int_0^t |\sigma(s, x_m)|^p \, ds,$$

where $\mu_p = \mathbb{E}[|Z|^p]$ *with* $Z \sim N(0, 1)$.

(2) *Let* $p = 2$ *or* $p \geq 4$ *and suppose that Assumption* B3 *holds. Then*

$$(2.29) \qquad \left(\Delta_n^{-\frac{1}{2}} \left(V_p^n(Y_x, t)_m - \mu_p \int_0^t |\sigma(s, x_m)|^p \, ds \right) \right)_{m=1}^N \overset{\text{st}}{\Longrightarrow} \mathcal{Z},$$

where $\mathcal{Z}$ *is a process as described after* (2.14) *and*

$$(2.30) \qquad \mathcal{C}_{m_1 m_2}(t) = \begin{cases} R_p \int_0^t |\sigma(s, x_m)|^{2p} \, ds, & m_1 = m_2 = m, \\ 0, & m_1 \neq m_2. \end{cases}$$

In the previous line, $R_p = \rho_p(1) + 2 \sum_{r=1}^\infty \rho_p(\Gamma_r)$, *where* Γ_r *is defined in* (2.7), *and* $\rho_p(r) = \text{Cov}(|X|^p, |Y|^p)$ *for* $(X, Y) \sim N\left(0, \left(\begin{smallmatrix} 1 & r \\ r & 1 \end{smallmatrix}\right)\right)$.

EXAMPLE 2.11.   If $\dot{W}$ is a space–time white noise and $p = 2$ or $p = 4$, by expressing x^2 and x^4 in terms of Hermite polynomials and then using Lemma 1.1.1 in [40], we obtain

$$R_2 = 2 + 4 \sum_{r=1}^\infty \left(\frac{1}{2}\sqrt{r+1} - \sqrt{r} + \frac{1}{2}\sqrt{r-1} \right)^2 = 2.357487\ldots,$$

$$(2.31) \qquad R_4 = 96 + 144 \sum_{r=1}^\infty \left(\frac{1}{2}\sqrt{r+1} - \sqrt{r} + \frac{1}{2}\sqrt{r-1} \right)^2$$

$$+ 48 \sum_{r=1}^\infty \left(\frac{1}{2}\sqrt{r+1} - \sqrt{r} + \frac{1}{2}\sqrt{r-1} \right)^4 = 109.223069\ldots,$$

which are larger than the corresponding constants 2 and 96 in the semimartingale framework (cf. Theorem 6.1 and Example 6.5 in [3]) and are the same as in the setting of moving average processes (cf. Theorem 4 in [4]). The reason for larger constants compared to the semimartingale case is the nonvanishing asymptotic correlation between increments of $Y(\cdot, x_m)$. Let us also remark that $R_2 = \pi \Gamma$ for the constant Γ in Theorem 4.2 of [9], and that $R_4 = \check{\sigma}^2$ for the constant $\check{\sigma}^2$ in equation (A.2) of [16].

2.3. *Estimation of volatility and spatial noise correlation index.*   In this section, we will explain how Theorems 2.1 and 2.3 can be applied to estimate the volatility process σ and the spatial correlation index α of the noise.

For both problems, the knowledge of the parameter λ is irrelevant as we shall see. This is important because there is no way to estimate λ consistently under our observation scheme. Indeed, a Girsanov argument (see Proposition 1.6 in [41]) shows that for constant σ, the laws of the solution Y on a compact space–time set are equivalent for different values of λ.

Furthermore, for the estimation of σ, we will assume that the parameter κ is known. In fact, if $N = 1$ and σ is a constant, then $Y(\cdot, x_1)$ is a stationary Gaussian process whose distribution

is completely determined by its covariance function. By (B.6) and the scaling properties of the normal distribution, this only depends on the ratio $\sigma^2/\kappa^{\alpha/2}$, so there is no way to identify the pair (σ, κ) based on observations of $Y(\cdot, x_1)$. If measurements are recorded at $N \geq 2$ spatial positions, then by the second statement of Lemma B.2(1), the normalized increments at different locations are asymptotically uncorrelated, and hence independent. Thus, it is impossible to consistently estimate both κ and σ based on observations at finitely many space points.

Despite this restriction, even if κ is unknown, the subsequent results can be easily modified to yield consistent and asymptotically mixed normal estimators for the *viscosity-adjusted volatility* $\kappa^{-\alpha w_m/4} \int_0^t |\sigma(s, x_m)|^{w_m} \, ds$ or the *relative volatility* $\int_0^t |\sigma(s, x_m)|^{w_m} \, ds / \int_0^T |\sigma(s, x_m)|^{w_m} \, ds$ with $m = 1, \ldots, N$, w_m as below and $t \in [0, T]$. Both quantities are constant multiples of the integrated volatility, and thus completely describe the shape of the temporal fluctuations of σ, which is sufficient for many applications; see, for example, [7], where the concept of relative volatility was introduced and further applied to turbulence data.

We first consider the situation when the spatial correlation index α is *known*. Then Corollary 2.9 immediately yields consistent estimators and asymptotic confidence bounds for the integrated volatility process at the measurement sites $x_1, \ldots, x_N$. In the theorems of this section, we will often divide by asymptotic $\mathcal{F}$-conditional variances during studentization procedures which may be zero in some degenerate situations. In these cases, convergence in probability and stable convergence in law should be understood in restriction to the set where all involved realized variation functionals are strictly positive. For the theoretical background of this concept for stable convergence in law, we refer the reader to Chapter 3.2 in [3] and to [46].

THEOREM 2.12. *Assume that $\alpha \in (0, 2) \cap (0, d]$ and $\kappa > 0$ are known. Define $\widetilde{V}^n_{\Phi|\Psi}(Y_x, t)$ in the same way as $V^n_{\Phi|\Psi}(Y_x, t)$ but with τ_n replaced by*

$$(2.32) \qquad \widetilde{\tau}_n^2 = \frac{\pi^{\frac{d}{2} - \alpha} \Gamma(\frac{\alpha}{2})}{(2\kappa)^{\frac{\alpha}{2}} (1 - \frac{\alpha}{2}) \Gamma(\frac{d}{2})} \Delta_n^{1 - \frac{\alpha}{2}}.$$

Then, under the hypotheses of Corollary 2.9 (1) and (2), we have

$$(2.33) \qquad \widetilde{V}^n_{\Phi_m|\Psi_m}(Y_x, t) \xrightarrow{L^1} \mu_{\Phi_m|\Psi_m} \int_0^t |\sigma(s, x_m)|^{w_m} \, ds, \quad m = 1, \ldots, N,$$

and, for every $T > 0$,

$$
(2.34) \qquad
\begin{aligned}
&\left\{ \Delta_n^{-\frac{1}{2}} \frac{\mu_{\Phi_m|\Psi_m}}{\sqrt{\rho_{\Phi_m|\Psi_m}}} \sqrt{\frac{\mu_{\Phi_m|\Psi_m(2w;\cdot)}}{\widetilde{V}^n_{\Phi_m|\Psi_m(2w;\cdot)}(Y_x, T)}} \right. \\
&\left. \times \left(\frac{\widetilde{V}^n_{\Phi_m|\Psi_m}(Y_x, T)}{\mu_{\Phi_m|\Psi_m}} - \int_0^T |\sigma(s, x_m)|^{w_m} \, ds \right) \right\}_{m=1}^N \xrightarrow{st} N(0, \mathrm{Id}_N)
\end{aligned}
$$

as $n \to \infty$. The left-hand sides of (2.33) and (2.34) are independent of λ.

If α is unknown, we first have to find a consistent estimator for α, for which we propose two solutions. The first estimator is a regression-type estimator similar to the change-of-frequency estimator in [5, 18] for the kernel singularity of a moving average process and similar to the estimator for the Hölder index of a Gaussian process proposed in [29]. Define

the function $\Phi^{(2)}(p; \cdot)\colon \mathbb{R}^{N \times 2} \to \mathbb{R}^N$ by

$$\Phi_m^{(2)}(p; x) = \Phi_m^{(2)}(p; (x_{j1}, x_{j2})_{j=1}^N) = |x_{m1} + x_{m2}|^p, \quad m = 1, \ldots, N.$$

Furthermore, recalling $\rho_p(r)$ from Corollary 2.10, define

$$(2.35) \qquad \mathcal{C}_0(\alpha) = \left(\frac{4}{p \log 2}\right)^2 \left(\mathcal{C}_{11} - 2\frac{\mathcal{C}_{12}}{(2 + 2\Gamma_1)^{\frac{p}{2}}} + \frac{\mathcal{C}_{22}}{(2 + 2\Gamma_1)^p}\right),$$

where

$$\mathcal{C}_{11} = \rho_p(1) + 2\sum_{r=1}^{\infty} \rho_p(\Gamma_r),$$

$$\mathcal{C}_{22} = (2 + 2\Gamma_1)^p \left(\rho_p(1) + 2\sum_{r=1}^{\infty} \rho_p\left(\frac{2\Gamma_r + \Gamma_{r-1} + \Gamma_{r+1}}{2 + 2\Gamma_1}\right)\right),$$

$$(2.36) \qquad \mathcal{C}_{12} = (2 + 2\Gamma_1)^{\frac{p}{2}} \left(\rho_p\left(\sqrt{\frac{1 + \Gamma_1}{2}}\right) + \sum_{r=1}^{\infty} \rho_p\left(\frac{\Gamma_r + \Gamma_{r-1}}{\sqrt{2 + 2\Gamma_1}}\right)\right.$$

$$\left. + \sum_{r=1}^{\infty} \rho_p\left(\frac{\Gamma_r + \Gamma_{r+1}}{\sqrt{2 + 2\Gamma_1}}\right)\right).$$

Note that $\mathcal{C}_0$ depends on α via Γ_r; see (2.7).

THEOREM 2.13.   *Let $\alpha \in (0, 2) \cap (0, d]$.*

(1) *If $p > 0$ and Assumptions A2 and A3 hold with exponent $p \vee 2$, then*

$$(2.37) \qquad \begin{aligned} \widehat{\alpha}_n^{(p)} &= \frac{1}{N} \sum_{m=1}^N \widehat{\alpha}_n^{(p),m} = \frac{1}{N} \sum_{m=1}^N \left(2 - \frac{4}{p} \log_2\left(\frac{V_{\Phi_m^{(2)}(p;\cdot)}^n(Y_x, T)}{V_{\Phi_m(p;\cdot)}^n(Y_x, T)}\right)\right) \\ &= 2 - \frac{4}{pN} \sum_{m=1}^N \log_2\left(\frac{\sum_{i=1}^{[T/\Delta_n]-1} |\Delta_i^n Y_{x_m} + \Delta_{i+1}^n Y_{x_m}|^p}{\sum_{i=1}^{[T/\Delta_n]} |\Delta_i^n Y_{x_m}|^p}\right) \xrightarrow{\mathbb{P}} \alpha. \end{aligned}$$

(2) *If $p = 2$ or $p \geq 4$ and Assumption B3 holds for this value of p, then*

$$(2.38) \qquad \frac{N}{\Delta_n^{\frac{1}{2}}} \sqrt{\frac{2^p \Gamma(\frac{2p+1}{2})}{\pi^{\frac{1}{2}} \mathcal{C}_0(\widehat{\alpha}_n^{(p)})}} \left(\sum_{m=1}^N \frac{V_{\Phi_m(2p;\cdot)}^n(Y_x, T)}{(V_{\Phi_m(p;\cdot)}^n(Y_x, T))^2}\right)^{-\frac{1}{2}} (\widehat{\alpha}_n^{(p)} - \alpha) \xrightarrow{\text{st}} N(0, 1).$$

The left-hand sides of (2.37) *and* (2.38) *do not depend on the parameters κ and λ.*

The second estimator is a correlation estimator (compare with the modified realized variation ratio of [5]). To this end, we define

$$(2.39) \qquad \widetilde{\mathcal{C}}_0(\alpha) = \left(\frac{2}{\log 2}\right)^2 (\widetilde{C}_{11} - 2\widetilde{C}_{12}\Gamma_1 + \widetilde{C}_{22}\Gamma_1^2),$$

where

$$\widetilde{C}_{11} = 1 + \Gamma_1^2 + 2\sum_{r=1}^{\infty} (\Gamma_r^2 + \Gamma_{r+1}\Gamma_{r-1}), \qquad \widetilde{C}_{22} = 2 + 4\sum_{r=1}^{\infty} \Gamma_r^2,$$

$$\widetilde{C}_{12} = 2\Gamma_1 + 2\sum_{r=1}^{\infty} \Gamma_r(\Gamma_{r+1} + \Gamma_{r-1}).$$

THEOREM 2.14. *Let $\alpha \in (0, 2) \cap (0, d]$ and $F(x) = -2\log_2(1 + x)$.*

(1) *Under Assumptions* A2 *and* A3 *with $p = 2$, we have as $n \to \infty$,*

$$
(2.40) \qquad \widetilde{\alpha}_n = \frac{1}{N}\sum_{m=1}^{N}\widetilde{\alpha}_n^m = \frac{1}{N}\sum_{m=1}^{N} F\left(\frac{V_{\Psi_m(1,1;\cdot)}^n(Y_x, T)}{V_{\Phi_m(2;\cdot)}^n(Y_x, T)}\right) \xrightarrow{\mathbb{P}} \alpha.
$$

(2) *Under Assumption* B3 *with $p = 2$, we have as $n \to \infty$,*

$$
(2.41) \qquad \frac{N}{\Delta_n^{\frac{1}{2}}}\sqrt{\frac{3}{\widetilde{\mathcal{C}}_0(\widetilde{\alpha}_n)}}\left(\sum_{m=1}^{N}\frac{V_{\Phi_m(4;\cdot)}^n(Y_x, T)}{(V_{\Psi_m(1,1;\cdot)+\Phi_m(2;\cdot)}^n(Y_x, T))^2}\right)^{-\frac{1}{2}}(\widetilde{\alpha}_n - \alpha)
$$
$$
\xrightarrow{\text{st}} N(0, 1).
$$

Both quantities on the left-hand side of (2.40) *and* (2.41) *do not depend on κ and λ.*

REMARK 2.15. Let us compare the asymptotic variances of the two estimators $\widehat{\alpha}_n^{(p)}$ and $\widetilde{\alpha}_n$ in the case where $\sigma(t, x) \equiv \sigma$ is constant (but nonzero). If $p = 2$, then under the assumptions of Theorems 2.13 and 2.14,

$$
(2.42) \qquad \begin{aligned} \lim_{n\to\infty}\Delta_n^{-\frac{1}{2}}\text{Var}[\widehat{\alpha}_n^{(2)} - \alpha] &= \frac{1}{NT}\mathcal{C}_0(\alpha), \\ \lim_{n\to\infty}\Delta_n^{-\frac{1}{2}}\text{Var}[\widetilde{\alpha}_n - \alpha] &= \frac{1}{NT}\frac{\widetilde{\mathcal{C}}_0(\alpha)}{(1 + \Gamma_1)^2}. \end{aligned}
$$

A straightforward calculation shows that $\mathcal{C}_0(\alpha) = \widetilde{\mathcal{C}}_0(\alpha)/(1 + \Gamma_1)^2$ for all $\alpha \in (0, 2)$. In other words, from the viewpoint of asymptotic variance, the two estimators $\widehat{\alpha}_n^{(2)}$ and $\widetilde{\alpha}_n$ are equivalent. By varying the value of p for the estimator $\widehat{\alpha}_n^{(p)}$, we can further check how reliable the estimates for α are.

With (either of) the two estimators for α at hand, we can now proceed to the estimation of σ. The rate of convergence is slower by a logarithmic factor compared to the case where α is known; see Theorem 2.12. This is the same phenomenon that occurs when the smoothness and the variance of a Gaussian process are to be estimated at the same time; see [13, 28, 29].

THEOREM 2.16. *Assume that $\alpha \in (0, 2) \cap (0, d]$ and that κ is known. Let $\alpha_n = \widehat{\alpha}_n^{(p_0)}$ for some $p_0 > 0$ or $\alpha_n = \widetilde{\alpha}_n$, in which case we set $p_0 = 2$. Define*

$$
(2.43) \qquad \widehat{\tau}_n^2 = \frac{\pi^{\frac{d}{2}-\alpha_n}\Gamma(\frac{\alpha_n}{2})}{(2\kappa)^{\frac{\alpha_n}{2}}(1 - \frac{\alpha_n}{2})\Gamma(\frac{d}{2})}\Delta_n^{1-\frac{\alpha_n}{2}}, \quad n \in \mathbb{N},
$$

and the functionals $\widehat{V}_{\Phi|\Psi}^n(Y_x, t)$ in the same way as $V_{\Phi|\Psi}^n(Y_x, t)$ but with $\widehat{\tau}_n$ instead of τ_n.

(1) *If Assumptions* A2 *and* A3 *hold with $p_0 \vee \overline{w} \vee 2$, then for every $T \geq 0$ and $m = 1, \ldots, N$,*

$$
\widehat{V}_{\Phi_m|\Psi_m}^n(Y_x, T) \xrightarrow{\mathbb{P}} \mu_{\Phi_m|\Psi_m}\int_0^T |\sigma(s, x_m)|^{w_m}\, ds \quad as\ n \to \infty.
$$

(2) *Assume that all $w_{mk} \in \{0, 2\} \cup [4, \infty)$ in the case of Φ, and that all w_m are even in the case of Ψ. Also assume that $p_0 = 2$ or ≥ 4, and that Assumption* B3 *holds with exponent*

$p_0 \vee \overline{w}$. *Then, if* $Z \sim N(0, 1)$, *we have in the case* $\alpha_n = \widehat{\alpha}_n^{(p_0)}$,

$$
\left\{ \frac{\Delta_n^{-\frac{1}{2}}}{|\log \Delta_n|} \frac{4N}{w_m \widehat{V}_{\Phi_m | \Psi_m}^n(Y_x, T)} \sqrt{\frac{2^{p_0} \Gamma(\frac{2p_0+1}{2})}{\pi^{\frac{1}{2}} \mathcal{C}_0(\widehat{\alpha}_n^{(p_0)})}} \left(\sum_{j=1}^N \frac{V_{\Phi_j(2p_0; \cdot)}^n(Y_x, T)}{(V_{\Phi_j(p_0; \cdot)}^n(Y_x, T))^2} \right)^{-\frac{1}{2}} \right.
$$

$$
\left. \times \left(\widehat{V}_{\Phi_m | \Psi_m}^n(Y_x, T) - \mu_{\Phi_m | \Psi_m} \int_0^T |\sigma(s, x_m)|^{w_m} \, ds \right) \right\}_{m=1}^N \xrightarrow{\text{st}} \begin{pmatrix} Z \\ \vdots \\ Z \end{pmatrix},
$$

while in the case $\alpha_n = \widetilde{\alpha}_n$,

$$
\left\{ \frac{\Delta_n^{-\frac{1}{2}}}{|\log \Delta_n|} \frac{4N}{w_m \widehat{V}_{\Phi_m | \Psi_m}^n(Y_x, T)} \sqrt{\frac{3}{\widetilde{\mathcal{C}}_0(\widetilde{\alpha}_n)}} \left(\sum_{j=1}^N \frac{V_{\Phi_j(4; \cdot)}^n(Y_x, T)}{(V_{\Psi_j(1,1; \cdot) + \Phi_j(2, \cdot)}^n(Y_x, T))^2} \right)^{-\frac{1}{2}} \right.
$$

$$
\left. \times \left(\widehat{V}_{\Phi_m | \Psi_m}^n(Y_x, T) - \mu_{\Phi_m | \Psi_m} \int_0^T |\sigma(s, x_m)|^{w_m} \, ds \right) \right\}_{m=1}^N \xrightarrow{\text{st}} \begin{pmatrix} Z \\ \vdots \\ Z \end{pmatrix}.
$$

REMARK 2.17. Whereas different coordinates are asymptotically independent in (2.34), they are identical in the limit in Theorem 2.16 (2). The former is a consequence of (2.26), while the latter is due the fact that the dominating term in the case of unknown α comes from the difference $\alpha_n - \alpha$ (which is independent of m); see the proof of Theorem 2.16.

3. Overview of proofs. The main ideas for the proof of Theorems 2.1 and 2.3 are sketched in this section, while the details will be given in Sections C and D. The results of Sections 2.2 and 2.3 are shown in Section E.

Without risk of confusion, we shall use the notation $x = (x_1, \ldots, x_N)' \in (\mathbb{R}^d)^N$, $\underline{\Delta}_i^n Y_x = (\Delta_i^n Y_x, \ldots, \Delta_{i+L-1}^n Y_x) \in \mathbb{R}^{N \times L}$, and

$$
\Delta_i^n G_y(s) = G(i \Delta_n - s, y) - G((i-1)\Delta_n - s, y) \in \mathbb{R},
$$

$$
\Delta_i^n G_{x,y}(s) = (\Delta_i^n G_{x_1-y}(s), \ldots, \Delta_i^n G_{x_N-y}(s))' \in \mathbb{R}^N,
$$

$$
\underline{\Delta}_i^n G_{x,y}(s) = (\Delta_i^n G_{x,y}(s), \ldots, \Delta_{i+L-1}^n G_{x,y}(s)) \in \mathbb{R}^{N \times L}
$$

for $s \in \mathbb{R}$ and $y \in \mathbb{R}^d$. Similarly, we let $\underline{\sigma}(s, x) = (\sigma(s, x_1), \ldots, \sigma(s, x_N))'$ and $\underline{\sigma}^2(s, x) = (\sigma^2(s, x_1), \ldots, \sigma^2(s, x_N))'$. Moreover, we write $t_n^* = [t/\Delta_n] - L + 1$ for $t \in [0, \infty)$ and $A \lesssim B$ if $A \leq CB$ for some finite constant $C > 0$ that does not depend on any important parameter.

The following measures will play an important role in identifying the limit behavior of (1.6):

$$
\Pi_{r,h}^n(A) = \iiint_A \frac{G_y(s) - G_y(s - \Delta_n)}{\tau_n}
$$

$$
\times \frac{G_{z+h}(s + r\Delta_n) - G_{z+h}(s + (r-1)\Delta_n)}{\tau_n} \, ds \, \Lambda(dy, dz),
$$

(3.1)

$$
|\Pi_{r,h}^n|(A) = \iiint_A \frac{|G_y(s) - G_y(s - \Delta_n)|}{\tau_n}
$$

$$
\times \frac{|G_{z+h}(s + r\Delta_n) - G_{z+h}(s + (r-1)\Delta_n)|}{\tau_n} \, ds \, \Lambda(dy, dz),
$$

where $r \in \mathbb{N}_0$, $h \in \mathbb{R}^d$, and $A \in \mathcal{B}([0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d)$. By (2.4), we have $\Pi_{0,0}^n([0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d) = 1$, so $\Pi_{0,0}^n$ is a probability measure. In fact, if we consider an arbitrary, say, the first

component of the increment $\Delta_i^n Y_x$ with $\sigma \equiv 1$, then for $A_1 \in \mathcal{B}([0, \infty))$ and $A_2 \in \mathcal{B}(\mathbb{R}^d)$, $\Pi_{0,0}^n(A_1 \times A_2 \times A_2)$ is the proportion of the variance of $Y(i\Delta_n, x_1) - Y((i-1)\Delta_n, x_1)$ that is explained by the integral in (3.2) on the set $\{(s, y): (i\Delta_n - s, x_1 - y) \in A_1 \times A_2\}$.

In general, $\Pi_{r,h}^n([0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d) \in [0, 1]$ is the correlation between two increments $\Delta_i^n Y(\cdot, x_1)$ and $\Delta_{i+r}^n Y(\cdot, x_1 + h)$, taken at a temporal distance of $r\Delta_n$ and a spatial distance of h. The value $\Pi_{r,h}^n(A_1 \times A_2 \times A_3)$, with A_1 and A_2 as above and $A_3 \in \mathcal{B}(\mathbb{R}^d)$, then quantifies how much this correlation is caused by the restrictions of the corresponding integrals in (3.2) to $\{(s, y): (i\Delta_n - s, x_1 - y) \in A_1 \times A_2\}$ and $\{(s, z): ((i+r)\Delta_n - s, x_1 + h - z) \in A_1 \times A_3\}$, respectively. Some important properties of these measures are proved in Section B.

3.1. *Overview of the proof of Theorem* 2.1. By arguing componentwise, we may assume without loss of generality that $M = 1$. As a first step, we show that we may further assume that σ is a bounded random field.

LEMMA 3.1. *In order to prove Theorem* 2.1, *one may additionally assume that σ is uniformly bounded in* $(\omega, t, x) \in \Omega \times \mathbb{R} \times \mathbb{R}^d$.

For the remaining analysis, by writing f as the difference of its positive and negative part, which still satisfy Assumption A1, we may assume that f is nonnegative. Then both $V_f^n(Y_x, t)$ and $V_f(Y_x, t)$ are increasing processes in t, so it suffices to prove $\mathbb{E}[|V_f^n(Y_x, t) - V_f(Y_x, t)|] \to 0$ for every $t \geq 0$. By definition, we have

$$(3.2) \qquad \Delta_i^n Y_x = \iint \Delta_i^n G_{x,y}(s)\sigma(s, y)W(ds, dy).$$

As we shall see, asymptotically as $\Delta_n \to 0$, only the portion of the integral where s is close to $i\Delta_n$ contributes to the size of $\Delta_i^n Y_x$. More precisely, we have the following result.

LEMMA 3.2. *For $\epsilon > 0$, if we define*

$$(3.3) \qquad \alpha_x^{n,i,\epsilon} = \iint \Delta_i^n G_{x,y}(s)\sigma(s, y)\mathbb{1}_{s > i\Delta_n - \epsilon}W(ds, dy),$$

then, as $n \to \infty$,

$$V_f^n(Y_x, t) - \Delta_n \sum_{i=1}^{t_n^*} f\left(\frac{\alpha_x^{n,i,\epsilon}}{\tau_n}\right) \xrightarrow{L^1} 0.$$

As a next step, we discretize the volatility process in (3.3) along the points $i\Delta_n - \epsilon$. By the following lemma, this only introduces an asymptotically negligible error.

LEMMA 3.3. *If*

$$\widehat{\alpha}_x^{n,i,\epsilon} = \iint \Delta_i^n G_{x,y}(s)\sigma(i\Delta_n - \epsilon, y)\mathbb{1}_{s > i\Delta_n - \epsilon}W(ds, dy),$$

then

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \mathbb{E}\left[\sup_{t \in [0,T]}\left|\Delta_n \sum_{i=1}^{t_n^*}\left\{f\left(\frac{\alpha_x^{n,i,\epsilon}}{\tau_n}\right) - f\left(\frac{\widehat{\alpha}_x^{n,i,\epsilon}}{\tau_n}\right)\right\}\right|\right] = 0.$$

For small ϵ, many of the integrals in (3.3) are taken over disjoint intervals. By exploiting this kind of conditional independence, we shall prove the following result.

LEMMA 3.4.   *If σ is bounded, we have for every $t \geq 0$, as $\epsilon \to 0$,*

(3.4) $$\sup_{n \in \mathbb{N}} \mathbb{E}\left[\left|\Delta_n \sum_{i=1}^{t_n^*} \left\{ f\left(\frac{\widehat{\alpha}_x^{n,i,\epsilon}}{\tau_n}\right) - \mathbb{E}\left[f\left(\frac{\widehat{\alpha}_x^{n,i,\epsilon}}{\tau_n}\right)\Big|\mathcal{F}_{i\Delta_n - \epsilon}\right]\right\}\right|^2\right] \to 0.$$

Finally, we show that the conditional expectations in (3.4) converge to the correct limit.

LEMMA 3.5.   *If σ is bounded, we have for every $t \geq 0$, as $\epsilon \to 0$,*

(3.5) $$\limsup_{n \to \infty} \mathbb{E}\left[\left|\Delta_n \sum_{i=1}^{t_n^*} \mathbb{E}\left[f\left(\frac{\widehat{\alpha}_x^{n,i,\epsilon}}{\tau_n}\right)\Big|\mathcal{F}_{i\Delta_n - \epsilon}\right] - \int_0^t \mu_f(\underline{\sigma}^2(r,x))\,\mathrm{d}r\right|\right] \to 0.$$

3.2. *Overview of the proof of Theorem* 2.3.   At the heart of our proof, we use a martingale central limit theorem for triangular arrays (see Theorem 2.2.15 in [32]) to obtain the stable convergence in law in (2.14). However, since the solution process (2.2) to (1.1) at fixed spatial points lacks the semimartingale property, many approximations are needed—before and after—to turn the left-hand side of (2.14) into a term with a martingale structure.

Step 1: *Martingalization.*   We want to truncate the increments $\underline{\Delta}_i^n Y_x$ in a similar way as in (3.3) to make sure that a large portion of the truncated increments are stochastic integrals over disjoint intervals (and hence have some sort of conditional independence) for different values of i. However, if we take a fixed level of ϵ as in (3.3), the number of overlapping increments will still be of order ϵ/Δ_n. Consequently, the total approximation error for $V_f^n(Y_x, t)$ will be of order $\Delta_n(\epsilon/\Delta_n) = \epsilon$, which is not sufficient due to the $1/\sqrt{\Delta_n}$ prefactor. Of course, the best truncation that one can hope for is to only keep the integral in (3.3) on the set $\{s > (i - 1)\Delta_n\}$, so that a given increment will only overlap with a finite number of other increments. But since $|\Pi_{0,0}^n|((\Delta_n, \infty) \times \mathbb{R}^d \times \mathbb{R}^d) \not\to 0$, this approximation is simply not valid (even without dividing by $\sqrt{\Delta_n}$). The idea is therefore to consider a truncation in between, that is, to take the integral in (3.3) only on the set $\{s > i\Delta_n - \lambda_n \Delta_n\}$ where λ_n is a sequence increasing to ∞ with $\lambda_n \Delta_n \to 0$. As it turns out, the best (i.e., smallest) choice for λ_n is achieved when we carry out the truncation iteratively. Hence, in a first step, we consider truncations of the form

(3.6) $$\gamma_x^{n,i,0} = \iint \underline{\Delta}_i^n G_{x,y}(s)\sigma(s,y)\mathbb{1}_{s > i\Delta_n - \lambda_n^0 \Delta_n} W(\mathrm{d}s, \mathrm{d}y),$$

for which we have the following result.

LEMMA 3.6.   *If $\lambda_n^0 = [\Delta_n^{-a_0}]$ for some $a_0 > \frac{1}{\nu}$, where $\nu = 1 + \frac{\alpha}{2}$ as in Lemma B.4, then*

(3.7) $$\Delta_n^{\frac{1}{2}} \sum_{i=1}^{t_n^*} \left\{ f\left(\frac{\underline{\Delta}_i^n Y_x}{\tau_n}\right) - f\left(\frac{\gamma_x^{n,i,0}}{\tau_n}\right)\right\} \overset{L^1}{\Longrightarrow} 0.$$

Since two truncated increments $\gamma_x^{n,i,0}$ and $\gamma_x^{n,j,0}$ are defined on disjoint intervals as soon as $|i - j| > \lambda_n^0 + L - 1$, we can employ martingale techniques to improve (i.e., decrease) the order of λ_n^0. As mentioned above, we use an iterative truncation procedure and consider numbers $a_1 > \cdots > a_R > \frac{1}{2\nu}$ satisfying $a_r > \frac{a_{r-1}}{\nu}$ for all $r = 1, \ldots, R$. We define $\lambda_n^r = [\Delta_n^{-a_r}]$ and $\gamma_x^{n,i,r}$ as in (3.6) but with λ_n^0 replaced by λ_n^r. Furthermore, with the notation

(3.8) $$\mathcal{F}_i^n = \mathcal{F}_{i\Delta_n},$$

we introduce the variables

$$\delta_i^{n,r} = \Delta_n^{\frac{1}{2}} \left(f\left(\frac{\gamma_x^{n,i,r-1}}{\tau_n} \right) - f\left(\frac{\gamma_x^{n,i,r}}{\tau_n} \right) \right),$$

(3.9)

$$\bar{\delta}_i^{n,r} = \delta_i^{n,r} - \mathbb{E}\big[\delta_i^{n,r} | \mathcal{F}_{i-\lambda_n^{r-1}}^n \big]$$

for $r = 1, \ldots, R$. As $n \to \infty$, we can now shrink the domain of integration to the set $\{s > i\Delta_n - \lambda_n^R \Delta_n\}$ as the following two lemmas show.

LEMMA 3.7. *For every* $r = 1, \ldots, R$, $\sum_{i=1}^{t_n^*} \bar{\delta}_i^{n,r} \overset{L^1}{\Longrightarrow} 0$.

LEMMA 3.8. *For every* $r = 1, \ldots, R$, $\sum_{i=1}^{t_n^*} \mathbb{E}[\delta_i^{n,r} | \mathcal{F}_{i-\lambda_n^{r-1}}^n] \overset{L^1}{\Longrightarrow} 0$.

In what follows, we define $a = a_R$, $\lambda_n = \lambda_n^R$, and $\gamma_x^{n,i} = \gamma_x^{n,i,R}$. Since $\nu > 1$, after possibly increasing R, we may assume that a is larger but arbitrarily close to $\frac{1}{2\nu}$. Although the iterative truncation procedure above has greatly reduced the number of overlapping increments (one increment now overlaps with roughly λ_n instead of Δ_n^{-1} increments), this number λ_n is still increasing in n, and hence, the increments are still far from having a martingale structure. A classical block splitting technique, similar to [31] (see also Chapter 12.2.4 in [32]), will now help us to finally obtain a martingale array. To this end, define

$$V^n(t) = \sum_{i=1}^{t_n^*} \psi_i^n, \qquad \psi_i^n = \Delta_n^{\frac{1}{2}} \left(f\left(\frac{\gamma_x^{n,i}}{\tau_n} \right) - \mathbb{E}\left[f\left(\frac{\gamma_x^{n,i}}{\tau_n} \right) \Big| \mathcal{F}_{i-\lambda_n}^n \right] \right).$$

We now arrange the summands ψ_i^n into blocks of length $m\lambda_n$ (where $m \in \mathbb{N}$), leaving out $\lambda_n + L - 1$ terms between two consecutive blocks. More precisely, we decompose $V^n(t)$ into $V^n(t) = V^{n,m,1}(t) + V^{n,m,2}(t) + V^{n,m,3}(t)$ with

$$V^{n,m,1}(t) = \sum_{j=1}^{J^{n,m}(t)} V_j^{n,m},$$

$$V^{n,m,2}(t) = \sum_{j=1}^{J^{n,m}(t)} \sum_{k=1}^{\lambda_n+L-1} \psi_{(j-1)((m+1)\lambda_n+L-1)+m\lambda_n+k}^n,$$

and

$$V^{n,m,3}(t) = \sum_{i=J^{n,m}(t)((m+1)\lambda_n+L-1)+1}^{t_n^*} \psi_i^n.$$

Here,

$$V_j^{n,m} = \sum_{k=1}^{m\lambda_n} \psi_{(j-1)((m+1)\lambda_n+L-1)+k}^n, \quad j = 1, \ldots, J^{n,m}(t),$$

are the blocks that we have up to time t. There are $J^{n,m}(t) = \lfloor \frac{t_n^*}{(m+1)\lambda_n+L-1} \rfloor$ complete blocks and possibly a boundary term $V^{n,m,3}(t)$, while $V^{n,m,2}(t)$ contains all summands that have been left out between blocks.

LEMMA 3.9.   *If $a > \frac{1}{2v}$ is sufficiently small, we have for $i = 2, 3$,*

$$\lim_{m \to \infty} \limsup_{n \to \infty} \mathbb{E}\Big[\sup_{t \in [0,T]} |V^{n,m,i}(t)|\Big] = 0.$$

For different values of j, the terms $V_j^{n,m}$ in the definition of $V^{n,m,1}$ comprise stochastic integrals over disjoint time domains. But the volatility process is evaluated continuously in time, so in order to finally obtain a martingale structure, we fix σ at the beginning of each block $V_j^{n,m}$. To this end, we define

(3.10)
$$\widehat{\xi}_{i,k}^n = \iint \frac{\Delta_i^n G_{x,y}(s)}{\tau_n} \sigma\big((i - \lambda_n - k)\Delta_n, y\big)\mathbb{1}_{s > (i - \lambda_n)\Delta_n} W(ds, dy),$$

$$\widehat{\psi}_{i,k}^n = \Delta_n^{\frac{1}{2}}\big(f(\widehat{\xi}_{i,k}^n) - \mathbb{E}[f(\widehat{\xi}_{i,k}^n)|\mathcal{F}_{i-\lambda_n-k}^n]\big),$$

and

(3.11)
$$\widehat{V}^{n,m,1}(t) = \sum_{j=1}^{J^{n,m}(t)} \widehat{V}_j^{n,m},$$

$$\widehat{V}_j^{n,m} = \sum_{k=1}^{m\lambda_n} \widehat{\psi}_{(j-1)((m+1)\lambda_n+L-1)+k,k}^n, \quad j = 1, \ldots, J^{n,m}(t).$$

LEMMA 3.10.   *If $a > \frac{1}{2v}$ is sufficiently small, we have for every $m \in \mathbb{N}$ that $V^{n,m,1}(t) - \widehat{V}^{n,m,1}(t) \overset{L^1}{\Longrightarrow} 0$ as $n \to \infty$.*

Step 2: *The martingale central limit theorem.* From (3.10) and (3.11), it is easy to see that for every $m \in \mathbb{N}$, the variables $((\widehat{V}_j^{n,m})_{j=1,\ldots,J^{n,m}(t)}: n \in \mathbb{N})$ form a *triangular array*, in the sense of Chapter 2.2.4 in [32], with respect to the filtrations $((\mathcal{F}_{j((m+1)\lambda_n+L-1)-\lambda_n}^n)_{j=0,\ldots,J^{n,m}(t)}: n \in \mathbb{N})$. We use Proposition 2.2.4 and Theorem 2.2.15 in [32] to establish the asymptotic distribution of $\widehat{V}^{n,m,1}(t)$.

PROPOSITION 3.11.   *For small $a > \frac{1}{2v}$, we have $\widehat{V}^{n,m,1} \overset{\text{st}}{\Longrightarrow} \mathcal{Z}^{(m)}$, where $\mathcal{Z}^{(m)}$ is a process characterized by the same properties as $\mathcal{Z}$ in Theorem 2.3 but with $\mathcal{C}(t)$ replaced by $\mathcal{C}^m(t) = \frac{m}{m+1}\mathcal{C}(t)$.*

We need two approximation results, whose proofs are given after that of Proposition 3.11.

LEMMA 3.12.   *In order to prove Proposition 3.11, we may assume that σ is bounded.*

LEMMA 3.13.   *In order to prove Proposition 3.11, we may assume that f is an even polynomial.*

Because $\mathbb{E}[\widehat{V}_j^{n,m}|\mathcal{F}_{(j-1)((m+1)\lambda_n+L-1)-\lambda_n}^n] = 0$ by construction, it remains to prove the following properties. In point (3) below, we consider a sequence $((W^\iota(t))_{t \in \mathbb{R}}: \iota \in \mathbb{N})$ of independent two-sided standard Brownian motions such that the stochastic integral of any $H \in \mathcal{L}(W)$ against W can be expressed as an L^2-series of stochastic integrals against W^ι. The existence of such a sequence follows as in Section 2.3 of [23]; see their Proposition 2.6(b) in particular.

(1) For all $t > 0$ and $m_1, m_2 = 1, \ldots, M$, we have, as $n \to \infty$,

$$
\begin{aligned}
(3.12) \qquad \mathcal{C}^{n,m}_{m_1 m_2}(t) &= \sum_{j=1}^{J^{n,m}(t)} \mathbb{E}\big[(\widehat{V}^{n,m}_j)_{m_1}(\widehat{V}^{n,m}_j)_{m_2}|\mathcal{F}^n_{(j-1)((m+1)\lambda_n+L-1)-\lambda_n}\big] \\
&\xrightarrow{\mathbb{P}} \mathcal{C}^m_{m_1 m_2}(t).
\end{aligned}
$$

(2) There is $q > 2$ such that for all $t > 0$ and $m \in \mathbb{N}$, we have as $n \to \infty$,

$$
(3.13) \qquad \sum_{j=1}^{J^{n,m}(t)} \mathbb{E}\big[|\widehat{V}^{n,m}_j|^q|\mathcal{F}^n_{(j-1)((m+1)\lambda_n+L-1)-\lambda_n}\big] \xrightarrow{\mathbb{P}} 0.
$$

(3) Let $(M(t))_{t \geq 0}$ be either the restriction to $[0, \infty)$ of W^ι for some $\iota \in \mathbb{N}$ or a bounded $(\mathcal{F}_t)_{t \geq 0}$-martingale that is orthogonal (in the martingale sense) to W^ι for all $\iota \in \mathbb{N}$. Then, given $m \in \mathbb{N}$ and $t > 0$, we have, as $n \to \infty$,

$$
(3.14) \qquad \sum_{j=1}^{J^{n,m}(t)} \mathbb{E}\big[\widehat{V}^{n,m}_j(M(\tau^n_j) - M(\tau^n_{j-1}))|\mathcal{F}^n_{(j-1)((m+1)\lambda_n+L-1)-\lambda_n}\big] \xrightarrow{\mathbb{P}} 0,
$$

where $\tau^n_j = \inf\{t \geq 0 \colon J^{n,m}(t) \geq j\} = (j((m+1)\lambda_n + L - 1) + L - 1)\Delta_n$.

Let us remark that in contrast to Theorem 2.2.15 in [32], we fix a countable, and not just a finite number of Brownian motions and then consider martingales orthogonal to all these Brownian motions. The proof of the mentioned theorem, which is based on [30], extends to this situation with no change.

Step 3: *Computing the conditional expectation.* Since $\mathcal{C}^m \xrightarrow{L^1} \mathcal{C}$, the results up to now and Proposition 2.2.4 in [32] imply that in the limit $n \to \infty$,

$$
(3.15) \qquad \Delta_n^{-\frac{1}{2}}\left(V^n_f(Y_x, t) - \Delta_n \sum_{i=1}^{t^*_n} \mathbb{E}\left[f\left(\frac{\gamma^{n,i}_x}{\tau_n}\right)\Big|\mathcal{F}^n_{i-\lambda_n}\right]\right) \xRightarrow{\text{st}} \mathcal{Z}.
$$

The conditional expectation cannot be computed explicitly as it involves the volatility process sampled continuously. The purpose of the next lemma is therefore to discretize σ at a fixed number of intermediate points (in the same spirit as Lemma 3 in [4]), where we use the following notation:

$$
(3.16) \qquad t^{n,i}_k = i\Delta_n - k\Delta_n, \qquad I^{n,i}_{k,l} = (t^{n,i}_k, t^{n,i}_l), \qquad \mathbb{1}^{n,i}_{k,l}(s) = \mathbb{1}_{I^{n,i}_{k,l}}(s).
$$

LEMMA 3.14. *If $a > \frac{1}{2\nu}$ is small enough, there exist numbers $a > a^{(1)} > \cdots > a^{(Q-1)}$ such that*

$$
(3.17) \qquad \Delta_n^{\frac{1}{2}} \sum_{i=1}^{t^*_n} \left\{\mathbb{E}\left[f\left(\frac{\gamma^{n,i}_x}{\tau_n}\right)\Big|\mathcal{F}^n_{i-\lambda_n}\right] - \mathbb{E}[f(\theta^n_i)|\mathcal{F}^n_{i-\lambda_n}]\right\} \xRightarrow{L^1} 0,
$$

where

$$
(3.18) \qquad \theta^n_i = \iint \frac{\Delta^n_i G_{x,y}(s)}{\tau_n} \sum_{q=1}^{Q} \sigma\big(t^{n,i}_{\lambda_n^{(q-1)}}, y\big)\mathbb{1}^{n,i}_{\lambda_n^{(q-1)}, \lambda_n^{(q)}}(s)W(\mathrm{d}s, \mathrm{d}y),
$$

and $\lambda_n^{(0)} = \lambda_n = [\Delta_n^{-a}]$, $\lambda_n^{(q)} = [\Delta_n^{-a^{(q)}}]$ for $q = 1, \ldots, Q - 1$, and $\lambda_n^{(Q)} = 0$.

Notice the difference between λ_n^r and a_r as defined before Lemma 3.7 and $\lambda_n^{(q)}$ and $a^{(q)}$ as defined in the previous lemma. In fact, we have

$$0 = \lambda_n^{(Q)} < \lambda_n^{(Q-1)} < \cdots < \lambda_n^{(0)} = \lambda_n = \lambda_n^R < \cdots < \lambda_n^0,$$

$$a^{(Q-1)} < \cdots < a^{(0)} = a = a_R < \cdots < a_0.$$

After suitable approximations, the conditional expectation of $f(\theta_i^n)$ can now be evaluated along these intermediate time points. To this end, define $m_r^{n,i} \in \mathbb{R}^{N \times L}$ and $v_r^{n,i} \in (\mathbb{R}^{N \times L})^2$ by

$$
\begin{aligned}
(m_r^{n,i})_{jk} &= \iint \frac{\Delta_{i+k-1}^n G_{x_j - y}(s)}{\tau_n} \\
&\quad \times \sum_{q=1}^r \sigma(t_{\lambda_n^{(q-1)}}^{n,i}, y) \mathbb{1}_{\lambda_n^{(q-1)}, \lambda_n^{(q)}}^{n,i}(s) W(ds, dy),
\end{aligned}
$$

(3.19)

$$
\begin{aligned}
(v_r^{n,i})_{jk, j'k'} &= \iiint \frac{\Delta_{i+k-1}^n G_{x_j - y}(s) \Delta_{i+k'-1}^n G_{x_{j'} - z}(s)}{\tau_n^2} \\
&\quad \times \sum_{q=r+1}^Q \sigma(t_{\lambda_n^{(q-1)}}^{n,i}, y) \sigma(t_{\lambda_n^{(q-1)}}^{n,i}, z) \mathbb{1}_{\lambda_n^{(q-1)}, \lambda_n^{(q)}}^{n,i}(s) \, ds \, \Lambda(dy, dz)
\end{aligned}
$$

for $r = 0, \ldots, Q$ and $j, j' = 1, \ldots, N$ and $k, k' = 1, \ldots, L$. In particular, we have $m_0^{n,i} = 0$, $m_Q^{n,i} = \theta_i^n$, and $v_Q^{n,i} = 0$. Then, recalling the definition of $\underline{\mu}_f$ after (C.8), the following results hold if a is sufficiently close to $\frac{1}{2v}$.

LEMMA 3.15. $\Delta_n^{\frac{1}{2}} \sum_{i=1}^{t_n^*} \{\mathbb{E}[f(\theta_i^n)|\mathcal{F}_{i-\lambda_n}^n] - \underline{\mu}_f(\mathbb{E}[v_0^{n,i}|\mathcal{F}_{i-\lambda_n}^n])\} \xrightarrow{L^1} 0$.

LEMMA 3.16. $\Delta_n^{\frac{1}{2}} \sum_{i=1}^{t_n^*} \{\underline{\mu}_f(\mathbb{E}[v_0^{n,i}|\mathcal{F}_{i-\lambda_n}^n]) - \underline{\mu}_f(v_0^{n,i})\} \xrightarrow{L^1} 0$.

Step 4: *Approximation of the Lebesgue integral.* We complete the proof of Theorem 2.3 by showing in two steps that $\Delta_n \sum_{i=1}^{[t/\Delta_n]} \underline{\mu}_f(v_0^{n,i})$ approximates the integral $V_f(Y_x, t) = \int_0^t \mu_f(\underline{\sigma}^2(s, x)) \, ds$ at a rate faster than $\sqrt{\Delta_n}$.

LEMMA 3.17. $\Delta_n^{\frac{1}{2}} \sum_{i=1}^{t_n^*} \{\underline{\mu}_f(v_0^{n,i}) - \mu_f(\underline{\sigma}^2((i-1)\Delta_n, x))\} \xrightarrow{L^1} 0$.

LEMMA 3.18. $\Delta_n^{-\frac{1}{2}} (\Delta_n \sum_{i=1}^{t_n^*} \mu_f(\underline{\sigma}^2((i-1)\Delta_n, x)) - V_f(Y_x, t)) \xrightarrow{L^1} 0$.

As expected from the semimartingale literature (cf. [32]), the two previous lemmas only hold true under strong regularity assumptions on σ. In fact, Assumption B2 and the spatial differentiability assumptions on σ in Assumption B3 are only needed for this step.

SUPPLEMENTARY MATERIAL

Supplement to "High-frequency analysis of parabolic stochastic PDEs" (DOI: [10.1214/19-AOS1841SUPP](https://doi.org/10.1214/19-AOS1841SUPP); .pdf). This paper is accompanied by supplementary material in [14]. Section A in [14] gives some auxiliary results needed for the proofs in this paper. In Section B, some important estimates related to the heat kernel are derived. Sections C and D provide the details for the proof of Theorems 2.1 and 2.3, respectively. Finally, Section E contains the proofs for Sections 2.2 and 2.3.

REFERENCES

[1] AÏT-SAHALIA, Y. and JACOD, J. (2009). Testing for jumps in a discretely observed process. *Ann. Statist.* **37** 184–222. MR2488349 https://doi.org/10.1214/07-AOS568

[2] AÏT-SAHALIA, Y. and JACOD, J. (2009). Estimating the degree of activity of jumps in high frequency data. *Ann. Statist.* **37** 2202–2244. MR2543690 https://doi.org/10.1214/08-AOS640

[3] AÏT-SAHALIA, Y. and JACOD, J. (2014). *High-Frequency Financial Econometrics*. Princeton Univ. Press, Princeton, NJ.

[4] BARNDORFF-NIELSEN, O. E., CORCUERA, J. M. and PODOLSKIJ, M. (2011). Multipower variation for Brownian semistationary processes. *Bernoulli* **17** 1159–1194. MR2854768 https://doi.org/10.3150/10-BEJ316

[5] BARNDORFF-NIELSEN, O. E., CORCUERA, J. M. and PODOLSKIJ, M. (2013). Limit theorems for functionals of higher order differences of Brownian semi-stationary processes. In *Prokhorov and Contemporary Probability Theory* (A. N. Shiryaev, S. R. S. Varadhan and E. L. Presman, eds.). *Springer Proc. Math. Stat.* **33** 69–96. Springer, Heidelberg. MR3070467 https://doi.org/10.1007/978-3-642-33549-5_4

[6] BARNDORFF-NIELSEN, O. E. and GRAVERSEN, S. E. (2011). Volatility determination in an ambit process setting. *J. Appl. Probab.* **48A** 263–275. MR2865631 https://doi.org/10.1239/jap/1318940470

[7] BARNDORFF-NIELSEN, O. E., PAKKANEN, M. S. and SCHMIEGEL, J. (2014). Assessing relative volatility/intermittency/energy dissipation. *Electron. J. Stat.* **8** 1996–2021. MR3273617 https://doi.org/10.1214/14-EJS942

[8] BARNDORFF-NIELSEN, O. E. and SCHMIEGEL, J. (2004). Lévy-based spatial-temporal modelling, with applications to turbulence. *Russian Math. Surveys* **59** 65–90.

[9] BIBINGER, M. and TRABS, M. (2018). Volatility estimation for stochastic PDEs using high-frequency observations. Available at arXiv:1710.03519 [math.ST].

[10] BIBINGER, M. and TRABS, M. (2019). On central limit theorems for power variations of the solution to the stochastic heat equation. Available at arXiv:1901.01026 [math.ST].

[11] BROCKWELL, P. J. (2009). An overview of asset-price models. In *Handbook on Financial Time Series* (T. Mikosch, J.-P. Kreiß, R. A. Davis and T. G. Andersen, eds.) 403–419. Springer, Berlin.

[12] BROCKWELL, P. J. and MATSUDA, Y. (2017). Continuous auto-regressive moving average random fields on $\mathbb{R}^n$. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 833–857. MR3641410 https://doi.org/10.1111/rssb.12197

[13] BROUSTE, A. and FUKASAWA, M. (2018). Local asymptotic normality property for fractional Gaussian noise under high-frequency observations. *Ann. Statist.* **46** 2045–2061. MR3845010 https://doi.org/10.1214/17-AOS1611

[14] CHONG, C. (2020). Supplement to "High-frequency analysis of parabolic stochastic PDEs." https://doi.org/10.1214/19-AOS1841SUPP.

[15] CIALENCO, I. (2018). Statistical inference for SPDEs: An overview. *Stat. Inference Stoch. Process.* **21** 309–329. MR3824970 https://doi.org/10.1007/s11203-018-9177-9

[16] CIALENCO, I. and HUANG, Y. (2017). A note on parameter estimation for discretely sampled SPDEs. Available at arXiv:1710.01649 [math.PR].

[17] CONT, R. (2005). Modeling term structure dynamics: An infinite dimensional approach. *Int. J. Theor. Appl. Finance* **8** 357–380. MR2144706 https://doi.org/10.1142/S0219024905003049

[18] CORCUERA, J. M., HEDEVANG, E., PAKKANEN, M. S. and PODOLSKIJ, M. (2013). Asymptotic theory for Brownian semi-stationary processes with application to turbulence. *Stochastic Process. Appl.* **123** 2552–2574. MR3054536 https://doi.org/10.1016/j.spa.2013.03.011

[19] CORCUERA, J. M., NUALART, D. and PODOLSKIJ, M. (2014). Asymptotics of weighted random sums. *Commun. Appl. Ind. Math.* **6** e–486, 11. MR3277312 https://doi.org/10.1685/journal.caim.486

[20] CORCUERA, J. M., NUALART, D. and WOERNER, J. H. C. (2006). Power variation of some integral fractional processes. *Bernoulli* **12** 713–735. MR2248234 https://doi.org/10.3150/bj/1155735933

[21] DA PRATO, G. and ZABCZYK, J. (1992). *Stochastic Equations in Infinite Dimensions*. *Encyclopedia of Mathematics and Its Applications* **44**. Cambridge Univ. Press, Cambridge. MR1207136 https://doi.org/10.1017/CBO9780511666223

[22] DALANG, R. C. (1999). Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s. *Electron. J. Probab.* **4** no. 6, 29 pp. MR1684157 https://doi.org/10.1214/EJP.v4-43

[23] DALANG, R. C. and QUER-SARDANYONS, L. (2011). Stochastic integrals for spde's: A comparison. *Expo. Math.* **29** 67–109. MR2785545 https://doi.org/10.1016/j.exmath.2010.09.005

[24] DENARO, G., VALENTI, D., LA COGNATA, A., SPAGNOLO, B., BONANNO, A., BASILONE, G., MAZZOLA, S., ZGOZI, S. W., ARONICA, S. et al. (2013). Spatio-temporal behaviour of the deep chlorophyll maximum in Mediterranean Sea: Development of a stochastic model for picophytoplankton dynamics. *Ecol. Complex.* **13** 21–34.

[25] EL SAADI, N. and BAH, A. (2007). An individual-based model for studying the aggregation behavior in phytoplankton. *Ecol. Model.* **204** 193–212.

[26] FOONDUN, M. and KHOSHNEVISAN, D. (2009). Intermittence and nonlinear parabolic stochastic partial differential equations. *Electron. J. Probab.* **14** 548–568. MR2480553 https://doi.org/10.1214/EJP.v14-614

[27] FOONDUN, M., KHOSHNEVISAN, D. and MAHBOUBI, P. (2015). Analysis of the gradient of the solution to a stochastic heat equation via fractional Brownian motion. *Stoch. Partial Differ. Equ. Anal. Comput.* **3** 133–158. MR3350450 https://doi.org/10.1007/s40072-015-0045-y

[28] FUKASAWA, M. and TAKABATAKE, T. (2019). Asymptotically efficient estimators for self-similar stationary Gaussian noises under high frequency observations. *Bernoulli* **25** 1870–1900. MR3961234 https://doi.org/10.3150/18-BEJ1039

[29] ISTAS, J. and LANG, G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. Henri Poincaré Probab. Stat.* **33** 407–436. MR1465796 https://doi.org/10.1016/S0246-0203(97)80099-4

[30] JACOD, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law. In *Séminaire de Probabilités, XXXI* (J. Azéma, M. Emery and M. Yor, eds.). *Lecture Notes in Math.* **1655** 232–246. Springer, Berlin. MR1478732 https://doi.org/10.1007/BFb0119308

[31] JACOD, J., PODOLSKIJ, M. and VETTER, M. (2010). Limit theorems for moving averages of discretized processes plus noise. *Ann. Statist.* **38** 1478–1545. MR2662350 https://doi.org/10.1214/09-AOS756

[32] JACOD, J. and PROTTER, P. (2012). *Discretization of Processes*. *Stochastic Modelling and Applied Probability* **67**. Springer, Heidelberg. MR2859096 https://doi.org/10.1007/978-3-642-24127-7

[33] JACOD, J. and TODOROV, V. (2014). Efficient estimation of integrated volatility in presence of infinite variation jumps. *Ann. Statist.* **42** 1029–1069. MR3210995 https://doi.org/10.1214/14-AOS1213

[34] JONES, B. J. T. (1999). The origin of scaling in the galaxy distribution. *Mon. Not. R. Astron. Soc.* **307** 376–386.

[35] LIU, J. and TUDOR, C. A. (2016). Central limit theorem for the solution to the heat equation with moving time. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **19** 1650005, 17. MR3474517 https://doi.org/10.1142/S0219025716500053

[36] LOTOTSKY, S. V. and ROZOVSKY, B. L. (2017). *Stochastic Partial Differential Equations*. *Universitext*. Springer, Cham. MR3674586 https://doi.org/10.1007/978-3-319-58647-2

[37] LUNARDI, A. (1995). *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. Birkhäuser, Basel. MR3012216

[38] NGUYEN, M. and VERAART, A. E. D. (2017). Spatio-temporal Ornstein–Uhlenbeck processes: Theory, simulation and statistical inference. *Scand. J. Stat.* **44** 46–80. MR3619694 https://doi.org/10.1111/sjos.12241

[39] NOURDIN, I., NUALART, D. and TUDOR, C. A. (2010). Central and non-central limit theorems for weighted power variations of fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 1055–1079. MR2744886 https://doi.org/10.1214/09-AIHP342

[40] NUALART, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd ed. *Probability and Its Applications (New York)*. Springer, Berlin. MR2200233

[41] NUALART, D. and PARDOUX, E. (1994). Markov field properties of solutions of white noise driven quasi-linear parabolic PDEs. *Stoch. Stoch. Rep.* **48** 17–44. MR1786190 https://doi.org/10.1080/17442509408833896

[42] NUALART, D. and QUER-SARDANYONS, L. (2007). Existence and smoothness of the density for spatially homogeneous SPDEs. *Potential Anal.* **27** 281–299. MR2336301 https://doi.org/10.1007/s11118-007-9055-3

[43] PAKKANEN, M. S. (2014). Limit theorems for power variations of ambit fields driven by white noise. *Stochastic Process. Appl.* **124** 1942–1973. MR3170230 https://doi.org/10.1016/j.spa.2014.01.005

[44] PEREIRA, R. M., GARBAN, C. and CHEVILLARD, L. (2016). A dissipative random velocity field for fully developed fluid turbulence. *J. Fluid Mech.* **794** 369–408. MR3482258 https://doi.org/10.1017/jfm.2016.166

[45] PHAM, V. S. and CHONG, C. (2018). Volterra-type Ornstein–Uhlenbeck processes in space and time. *Stochastic Process. Appl.* **128** 3082–3117. MR3834854 https://doi.org/10.1016/j.spa.2017.10.012

[46] PODOLSKIJ, M. and VETTER, M. (2010). Understanding limit theorems for semimartingales: A short survey. *Stat. Neerl.* **64** 329–351. MR2683464 https://doi.org/10.1111/j.1467-9574.2010.00460.x

[47] POSPÍŠIL, J. and TRIBE, R. (2007). Parameter estimates and exact variations for stochastic heat equations driven by space–time white noise. *Stoch. Anal. Appl.* **25** 593–611. MR2321899 https://doi.org/10.1080/07362990701282849

[48] PRAKASA RAO, B. L. S. (2002). Nonparametric inference for a class of stochastic partial differential equations based on discrete observations. *Sankhyā Ser. A* **64** 1–15. MR1968372

[49] ROBERT, R. and VARGAS, V. (2008). Hydrodynamic turbulence and intermittent random fields. *Comm. Math. Phys.* **284** 649–673. MR2452591 https://doi.org/10.1007/s00220-008-0642-y

[50] SIGRIST, F., KÜNSCH, H. R. and STAHEL, W. A. (2012). A dynamic nonstationary spatio-temporal model for short term prediction of precipitation. *Ann. Appl. Stat.* **6** 1452–1477. MR3058671 https://doi.org/10.1214/12-AOAS564

[51] SIGRIST, F., KÜNSCH, H. R. and STAHEL, W. A. (2015). Stochastic partial differential equation based modelling of large space–time data sets. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 3–33. MR3299397 https://doi.org/10.1111/rssb.12061

[52] SWANSON, J. (2007). Variations of the solution to a stochastic heat equation. *Ann. Probab.* **35** 2122–2159.

[53] TUCKWELL, H. C. (2013). Stochastic partial differential equations in neurobiology: Linear and nonlinear models for spiking neurons. In *Stochastic Biomathematical Models* (M. Bachar, J. Batzel and S. Ditlevsen, eds.). *Lecture Notes in Math.* **2058** 149–173. Springer, Heidelberg. MR3051032 https://doi.org/10.1007/978-3-642-32157-3_6

[54] WALSH, J. B. (1981). A stochastic model of neural response. *Adv. in Appl. Probab.* **13** 231–281. MR0612203 https://doi.org/10.2307/1426683

[55] WALSH, J. B. (1986). An introduction to stochastic partial differential equations. In *École D'été de Probabilités de Saint-Flour, XIV—1984* (P. L. Hennequin, ed.). *Lecture Notes in Math.* **1180** 265–439. Springer, Berlin. MR0876085 https://doi.org/10.1007/BFb0074920