Examen cours de finance statistique

Janvier 2025

No document allowed, 3 hours

Recommended timing: Questions 1-2: 30 min, Questions 3-4-5: 30 min, Article reading: 1 hour, Writing: 1 hour.

- (3 pts) We consider the Markowitz portfolio problem in the case of no risk free asset and N risky assets with expectation vector μ and covariance matrix Ω. Assuming we have a target expected return for the portfolio given by some positive constant μ_p, derive the optimal (relative) weights for each risky asset in the Markowitz problem in term of Ω, the $N \times 2$ matrix $\mathbf{U} = [\mu\ \mathbf{1}]$ and the vector $u = [\mu_p\ 1]'$.
- (2 pts) Show that in the case of no risk free asset, for optimal Markowitz portfolios, the variance is a quadratic function of the expected return μ_p.
- (1.5 pts) Ordinary least squares, ridge estimator, Lasso estimator: definitions, closed form solutions when possible, pros and cons.
- (1.5 pts) How can we set an optimal tick value for an asset ? (5-10 lines).
- (2 pts) How would you demonstrate that rough volatility models are superior to conventional stochastic volatility models? (1 page maximum, several answers are possible. Making connections with arbitrage and market impact is a good idea).
- (10 pts) Summarize (quickly) and comment the enclosed article in light of what has been seen in class. Discuss the obtained results, the strengths and weaknesses of the approach, the relevant points and the limitations. Suggest way to improve this work or interesting directions to extend it.

When Frictions are Fractional: Rough Noise in High-Frequency Data

Carsten H. Chong*
Department of Information Systems, Business Statistics and Operations Management,
The Hong Kong University of Science and Technology
and
Thomas Delerue
Institute of Epidemiology, Helmholtz Munich
and
Guoying Li
Department of Statistics, Columbia University

Abstract

The analysis of high-frequency financial data is often impeded by the presence of noise. This article is motivated by intraday return data in which market microstructure noise appears to be *rough*, that is, best captured by a continuous-time stochastic process that locally behaves as fractional Brownian motion. Assuming that the underlying efficient price process follows a continuous Itô semimartingale, we derive consistent estimators and asymptotic confidence intervals for the roughness parameter of the noise and the integrated price and noise volatilities, in all cases where these quantities are identifiable. In addition to desirable features such as serial dependence of increments, compatibility between different sampling frequencies and diurnal effects, the rough noise model can further explain divergence rates in volatility signature plots that vary considerably over time and between assets.

Keywords: Hurst parameter, market microstructure noise, mixed fractional Brownian motion, mixed semimartingales, volatility estimation, volatility signature plot.

*We thank Yacine Aït-Sahalia, Torben Andersen, Patrick Cheridito, Jean Jacod, Fabian Mies, Serena Ng, Mark Podolskij, Walter Schachermayer, Viktor Todorov and participants at various conferences and seminars for valuable comments and suggestions, which greatly improved the paper. We would also like to thank the Editor, an Associate Editor and a referee, whose detailed comments on earlier drafts led to a substantial improvement of the paper. The second author is partially supported by the Deutsche Forschungsgemeinschaft, project number KL 1041/7-2.

1 Introduction

A classical statistical inference problem consists in separating a signal X from a noise term Z when only their sum

$$Y = X + Z \tag{1.1}$$

is observed. This paper analyzes a particular instance of this problem in which the signal term X is a continuous Itô semimartingale of the form

$$X_t = X_0 + \int_0^t a_s \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}B_s, \tag{1.2}$$

the noise term Z is a rough fractional process to be specified in Section 2 below, and the observations $\{Y_{i\Delta_n} : i = 0, \dots, [T/\Delta_n]\}$ are recorded on a regularly spaced time grid where $\Delta_n \to 0$ and $T > 0$ is fixed.

The motivation for this problem comes from the statistical analysis of high-frequency financial data, where X models the efficient logarithmic price of an asset (i.e., its economic value in a frictionless setting) and Z denotes microstructure noise, which in practice arises due to bid–ask bounces, transaction costs and other market frictions. In this context, a key quantity of interest is the *integrated (price) volatility* $C_T = \int_0^T \sigma_s^2 \,\mathrm{d}s$. In the absence of noise, estimating C_T is a straightforward matter: given observations $\{X_{i\Delta_n} : i = 0, \dots, [T/\Delta_n]\}$, the *realized variance (RV)* defined by $\sum_{i=1}^{[T/\Delta_n]} (\Delta_i^n X)^2$, where $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$, is a consistent estimator of C_T as $\Delta_n \to 0$ (Jacod & Protter 2012*a*).

However, in practice, RV typically explodes as the sampling frequency increases, indicating the presence of noise. In order to deal with noisy observations, a common approach in the literature is to model Z at the observation times $i\Delta_n$ as

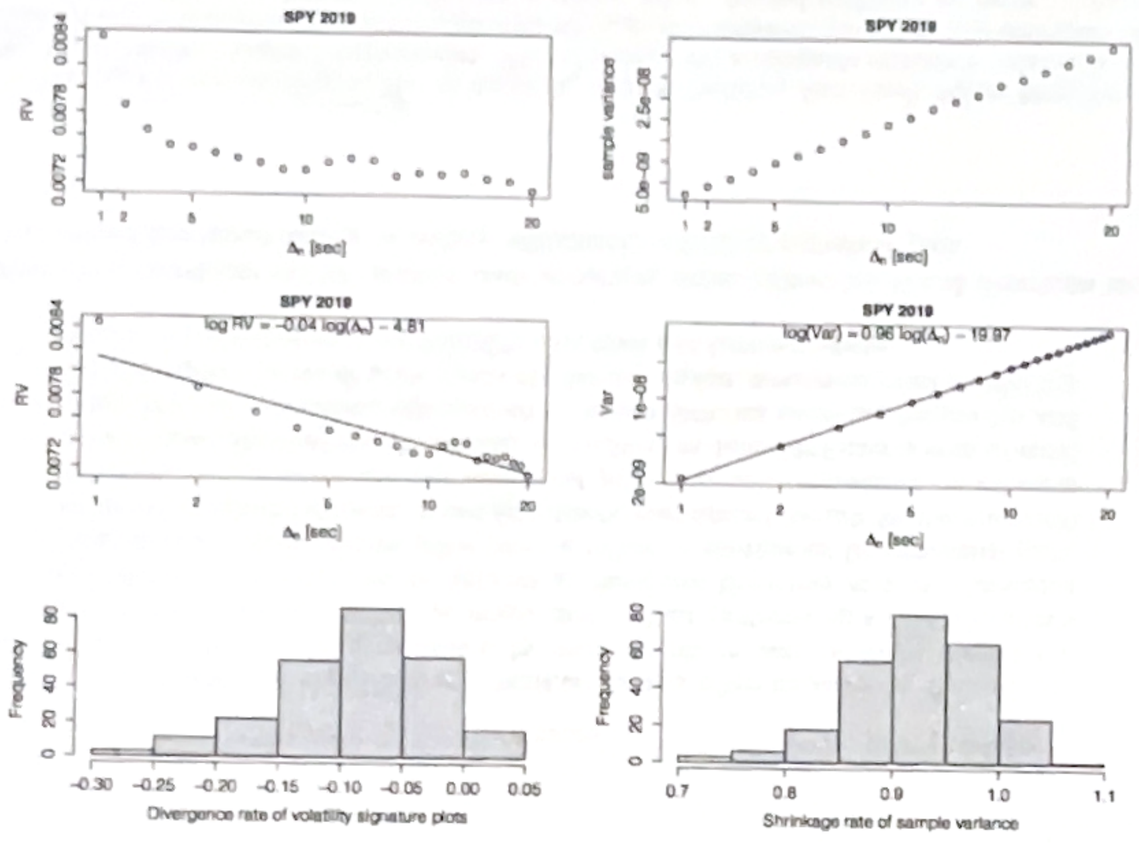
$$Z_{i\Delta_n} = \varepsilon_i^n, \tag{1.3}$$

where $(\varepsilon_i^n)_{i=1}^{[T/\Delta_n]}$ is a discrete time series for each n, and to construct noise-robust estimators of C_T based on that assumption.[1] The current paper is motivated by statistical properties found in certain samples of high-frequency financial data that cannot be explained by such discrete noise models. For instance, if the noise Z is independent of X and takes the form (1.3), where $\varepsilon = (\varepsilon_i^n)_{i=1}^{[T/\Delta_n]}$ is a stationary time series with a distribution that does not depend on n, it is a simple consequence of the law of large numbers (LLN) that

$$\Delta_n \sum_{i=1}^{[T/\Delta_n]} (\Delta_i^n Y)^2 \xrightarrow{\mathrm{P}} 2\,\mathrm{Var}(\varepsilon)(1 - r(1)),$$

where r is the autocorrelation function (ACF) of ε_i^n. In particular, the RV of the observed process Y should be of order Δ_n^{-1}. However, as we can see in the volatility signature plot of Figure 1 (a), the divergence rate of RV in real samples can be much slower. An almost equivalent way of illustrating this observation is to consider *variance plots*, in which the sample variance of increments of Y is computed as a function of the sampling frequency.

[1] Examples for ε_i^n include rounding noise (Delattre & Jacod 1997, Li & Mykland 2007, Robert & Rosenbaum 2010, Rosenbaum 2009), white noise (Bandi & Russell 2006, Barndorff-Nielsen et al. 2008, Podolskij & Vetter 2009, Zhang et al. 2005), AR- or MA-type noise (Aït-Sahalia et al. 2011, Da & Xiu 2021, Hansen & Lunde 2006), and certain non-parametric extensions thereof (Jacod et al. 2009, 2017, Li & Linton 2022).



(a) Volatility signature plots (b) Variance plots

Figure 1: (a) Volatility signature plot and (b) variance plot for 2019 SPY transaction data (top). The same plots on a log–log scale (middle) reveal a divergence rate of -0.04 for RV and a shrinkage rate of increments of 0.96 for the whole year. (The divergence rate of RV is α if $\mathrm{RV} \sim C_1 \Delta_n^\alpha$ for some $C_1 > 0$; the shrinkage rate of increments is β if $\mathrm{Var}(\Delta_i^n Y) \sim C_2 \Delta_n^\beta$ for some $C_2 > 0$.) The histograms (bottom) show the daily divergence rates in volatility signature plots and the daily shrinkage rates of price increments in 2019 SPY transaction data. Each data point corresponds to one trading day.

In our data sample, we observe shrinking price increments; see Figure 1 (b). By contrast, in discrete noise models,

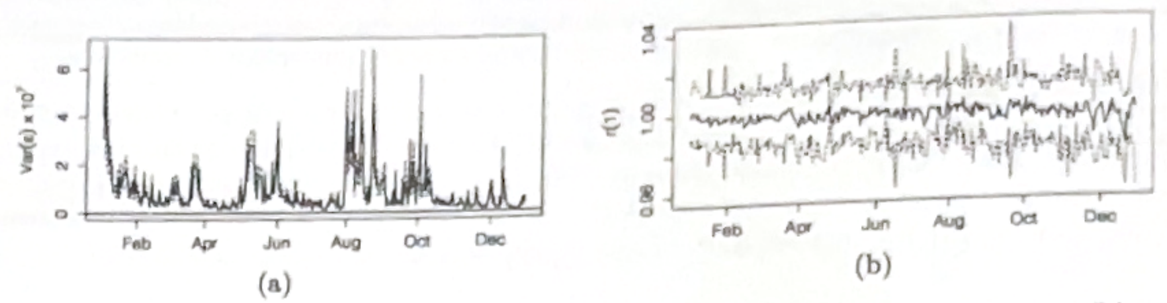
$$\mathrm{Var}(\Delta_i^n Y) \sim \mathrm{Var}(\varepsilon_i^n - \varepsilon_{i-1}^n) = 2\,\mathrm{Var}(\varepsilon)(1 - r(1)), \tag{1.4}$$

so asymptotically noise increments do not shrink.

To corroborate these findings, we perform additional analyses on our data sample, the results of which are shown in Figure 2. Panel (a) shows Jacod et al. (2017)'s point estimators and 95%-confidence bands for $\mathrm{Var}(\varepsilon)$, indicating that $\mathrm{Var}(\varepsilon)$ is significantly different from 0. This suggests that

$$\mathrm{Var}(Z_{i\Delta_n}) = \mathrm{Var}(\varepsilon_i^n) \text{ is bounded away from 0.} \tag{1.5}$$

Panel (b) shows Jacod et al. (2017)'s point estimators and 95%-confidence intervals for the first-order autocorrelation $r(1)$ of the noise. As we can see, there is a high correlation



(a) (b)

Figure 2: (a) Estimators of $\mathrm{Var}(\varepsilon)$ and (b) estimators of $r(1)$ including 95%-confidence intervals. The analysis is based on 2019 SPY transaction data sampled at $\Delta_n = 1$ sec.

between ε_i^n and ε_{i+1}^n, which is not significantly different from 1 on most days. By (1.4), it follows that $\mathrm{Var}(\varepsilon_i^n - \varepsilon_{i-1}^n) \approx 0$, which we interpret as

$$\mathrm{Var}(\Delta_i^n Z) = \mathrm{Var}(\varepsilon_i^n - \varepsilon_{i-1}^n) \to 0. \tag{1.6}$$

In conclusion, there is strong empirical evidence that market microstructure noise in our data sample is *non-shrinking* (because of (1.5)) but with *shrinking increments* (because of (1.6)).[2] To our best knowledge, all microstructure noise models that have been considered so far in the literature are either non-shrinking with non-shrinking increments (as in (1.3)) or shrinking with (necessarily) shrinking increments (e.g., Aït-Sahalia & Xiu (2019), Da & Xiu (2021), Kalnina & Linton (2008)). The goal of this work is to fill in this gap.

To this end, we introduce in Section 2 a non-shrinking microstructure noise model with shrinking increments that we call the *rough noise model*. We establish a CLT for variation functionals of the resulting semimartingale plus noise process in Section 3. A major challenge here is the subtle interplay between the semimartingale and the noise process, leading to (a potentially large number of) intermediate limits between the LLN and the CLT. While the LLN and the CLT limits only depend on the noise, these intermediate limits depend on coefficients of the semimartingale and of the noise process at the same time. In Section 4, we then combine the CLT of Section 3 with a generalized method of moments (GMM) approach to derive consistent and asymptotically mixed normal estimators of integrated price volatility, integrated noise volatility and the roughness of the noise. Section 5 contains a simulation and Section 6 shows an empirical study of SPY transaction data. Section 7 concludes. The supplement contains more details about modeling noise in continuous time (Appendix A), a multivariate extension of the CLT for mixed semimartingales (Appendix B) and its proof (Appendix C), the proof of the results in Section 4 (Appendix D) and further empirical results (Appendix E).

2 Model

Both the noise and the efficient price process are defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ satisfying the usual conditions. A natural way of implementing the

[2] We neither claim that this is universally true nor that noise has shrinking increments in tick time, which is the primary concern in Jacod et al. (2017). It is well known (see e.g., Hansen & Lunde (2006)) that the distributional properties of noise can be very different in calendar time compared to tick time.

shrinking increments property (1.6) observed in our data is to model the noise process Z in continuous time. Indeed, if $Z_{i\Delta_n}$ does not change much on average from i to $i+1$, this implies some form of continuity (e.g., in probability) between them. Therefore, $\{Z_{i\Delta_n} : i \in \mathbb{N}\}$, at least for large n, essentially determines a continuous-time process $(Z_t)_{t\geq 0}$. A continuous-time noise model further has the advantage that it is compatible between different sampling frequencies, a property that is typically hard to satisfy for colored noise models with non-shrinking increments (see Section 7.1.2 in Aït-Sahalia & Jacod (2014)). We give more econometric background on about modeling noise in continuous time in Appendix A.

Assumption (Z). *The process $(Z_t)_{t\geq 0}$ is given by*

$$Z_t = Z_0 + \int_0^t g(t-s)\rho_s \, \mathrm{d}W_s, \qquad t \geq 0, \tag{2.1}$$

where W is a standard $\mathbb{F}$-Brownian motion and $(\rho_t)_{t\geq 0}$ is an $\mathbb{F}$-adapted locally bounded process. The kernel $g\colon (0,\infty) \to \mathbb{R}$ is of the form

$$g(t) = K_H^{-1} t^{H-\frac{1}{2}} + g_0(t) \tag{2.2}$$

for some $H \in (0, \frac{1}{2})$, $K_H = \sqrt{2H\sin(\pi H)\Gamma(2H)}/\Gamma(H+\frac{1}{2})$ is a normalization constant and $g_0\colon [0,\infty) \to \mathbb{R}$ is a smooth function with $g_0(0) = 0$.[3]

Remark 2.1. By the Wold–Karhunen representation theorem (Doob 1953, Theorem XII.5.3), every second-order stationary process, up to deterministic or finite-variation components, has the form $\int_0^t G(t-s)\,\mathrm{d}M_s$ for some kernel $G \in L^2((0,\infty))$ and some process $(M_t)_{t\geq 0}$ with second-order stationary and orthogonal increments. Therefore, if Z is stationary, (2.1) is quite a natural assumption on the noise process. Due to the presence of ρ, the process Z in (2.1) does not need to be stationary in general.

In the special case where $g_0 \equiv 0$ and $\rho_s \equiv \rho$ is a constant, Z is—up to a term of finite variation—simply a multiple of *fractional Brownian motion (fBM)*. If further $X_t = \sigma B_t$ with constant volatility σ, then the resulting observed process $Y_t = \sigma B_t + \rho Z_t$ is a *mixed fractional Brownian motion (mfBM)* as introduced by Cheridito (2001). Our model for the observed price process, as the sum of X in (1.2) and Z in (2.1), can be viewed as a non-parametric generalization of mfBM that allows for stochastic volatility in both its Brownian and its fractional component. We do keep the parameter H, though, which we refer to as the *roughness parameter* of Z (or Y).[4] In analogy with mfBM, we call

$$Y_t = X_t + Z_t = Y_0 + \int_0^t a_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}B_s + \int_0^t g(t-s)\rho_s \, \mathrm{d}W_s, \qquad t \geq 0, \tag{2.3}$$

the observed price process in our model, a *mixed semimartingale*.

[3] The condition $H < \frac{1}{2}$ is not restrictive for the purpose of modeling microstructure noise: if $H = \frac{1}{2}$, then Z has the same smoothness as Brownian motion, so in general, there will be no way to discern Z from the efficient price process X; if $H > \frac{1}{2}$, then Z is smoother than X and RV remains a consistent estimator of C_T. The normalization K_H is chosen in such a way that $\mathbb{E}[(Z_{t+\delta} - Z_t)^2]/\delta^{2H} \to 1$ as $\delta \to 0$ if $\rho \equiv 1$.

[4] Fractional processes are also used in Mandelbrot (1997), Bayraktar et al. (2004), Bianchi & Pianese (2018) to model asset prices. In these works, the primary interest is short-/long-range dependence, which is determined by the behavior of g at $t = \infty$. Our interest, by contrast, is the behavior of g around $t = 0$, which governs the local regularity, or *roughness*, of the fractional process. Since our model does not specify the behavior of g at $t = \infty$ (due to the presence of g_0 in (2.2)), we refer to H as the roughness parameter of Z.

Remark 2.2. In recent years, there has been growing interest in *rough volatility models* (Gatheral et al. 2018), where σ is modeled by a rough process. In this paper, by contrast, we are concerned with roughness of observed prices, caused by microstructure noise. Roughness on the price level and roughness on the volatility level imply different features of asset returns and must be modeled and analyzed separately. For instance, if $Y_t = X_t = \int_0^t \sigma_s \, \mathrm{d}B_s$, without noise but with a rough σ, RV will *not* explode in volatility signature plots. In fact, in the absence of noise, the asymptotics of RV do *not* depend on the roughness of volatility (Jacod & Protter 2012*b*, Theorem 5.4.2). Therefore, the empirical findings discussed so far and below can neither be explained by nor do they indicate rough volatility.

On an abstract level, the statistical problem we are facing in this paper is a deconvolution problem: given a semimartingale process X and rough process Z, how can we recover the two (or certain components of the two, such as volatility) based on observing their sum $Y = X + Z$. The next result, which follows from (van Zanten 2007, Corollary 2.2), puts a constraint on the identifiability of the (smoother) semimartingale signal:

Proposition 2.3. *Assume that Y is an mfBM, that is, $Y = X + Z$ where $X = \sigma B$ and $Z = \rho B^H$ for some $\rho, \sigma \in (0,\infty)$, B is a Brownian motion and B^H is an independent fBM with Hurst parameter $H \in (0, \frac{1}{2})$. For any $T > 0$, the laws of $(Y_t)_{t\in[0,T]}$ and $(Z_t)_{t\in[0,T]}$ are mutually equivalent if $H \in (0, \frac{1}{4})$ and mutually singular if $H \in [\frac{1}{4}, \frac{1}{2})$.*

In other words, if $H \in (0, \frac{1}{4})$, due to the roughness of the noise, there is no way to consistently estimate σ on a finite time interval. This is conceptually similar to the fact that the finite-variation part of a semimartingale cannot be estimated consistently in finite time if there is a Brownian component. We will comment on possible pathways to estimate σ if $H < \frac{1}{4}$ in Section 7.

Remark 2.4. The case of white noise, which formally corresponds to $H = 0$ in terms of roughness, is special in this context: it is rougher than Z in (2.1), but $C_T = \int_0^T \sigma_s^2 \, \mathrm{d}s$ can still be recovered through subsampling (Zhang et al. 2005) or pre-averaging (Jacod et al. 2009). Indeed, if k_n is an increasing sequence and Z is a white noise, then $k_n^{-1}\sum_{j=0}^{k_n} Y_{(i+j)\Delta_n} \approx X_{i\Delta_n}$ by the law of large numbers. By contrast, if $H \in (0, \frac{1}{2})$, the process Z in (2.1) is *continuous* (and so is Y in (1.1)), which implies that $k_n^{-1}\sum_{j=0}^{k_n} Y_{(i+j)\Delta_n} \approx Y_{i\Delta_n}$, so pre-averaging does not remove the noise part at all! Therefore, while classical noise-robust volatility estimators work well if Z is a modulated white noise, they become inconsistent for C_T if $H \in (0, \frac{1}{2})$.

3 Central limit theorem for variation functionals

Our estimators are based on limit theorems for power variations and related functionals in an infill asymptotic setting. To keep notation simple, we only discuss the one-dimensional case here; a multivariate extension of Theorem 3.1 is stated and proved in Appendix B. Given a test function $f\colon \mathbb{R} \to \mathbb{R}$, our goal is to establish a CLT for *normalized variation functionals*

$$V_f^n(Y,t) = \Delta_n \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} f\Big(\frac{\Delta_i^n Y}{\Delta_n^H}\Big),$$

where $\Delta_i^n Y = Y_{i\Delta_n} - Y_{(i-1)\Delta_n}$. For semimartingales, this is a well studied topic; see Aït-Sahalia & Jacod (2014) and Jacod & Protter (2012*b*) for in-depth treatments of this

subject. For fractional Brownian motion or moving-average processes as in (2.1), the theory is similarly well understood; see Barndorff-Nielsen et al. (2011) and Brouste & Fukasawa (2018). Surprisingly, it turns out that the mixed case is more complicated than the "union" of the purely semimartingale and the purely fractional case. For instance, as we elaborate in Remark 3.4, already for power variations of even order, we may have a large number of higher-order bias terms. Our CLT will be proved under the following set of assumptions.

Assumption (CLT). *The observation process Y is given by the sum of X from (1.2) and Z from (2.1) with the following specifications:*

(i) The function $f\colon \mathbb{R} \to \mathbb{R}$ is even and infinitely differentiable. Moreover, all its derivatives (including f itself) have at most polynomial growth.

(ii) Both B and W are independent standard $\mathbb{F}$-Brownian motions, the drift a is locally bounded and $\mathbb{F}$-adapted, and σ is an $\mathbb{F}$-adapted locally bounded process such that for every $T > 0$, there is $K_1 \in (0, \infty)$ with

$$\mathbb{E}\big[1 \wedge |\sigma_t - \sigma_s|\big] \leq K_1 |t-s|^{\frac{1}{2}}, \qquad s,t \in [0,T]. \tag{3.1}$$

(iii) The noise volatility process ρ from (2.1) takes the form

$$\rho_t = \rho_t^{(0)} + \int_0^t \widetilde{b}_s \,\mathrm{d}s + \int_0^t \widetilde{\rho}_s \,\mathrm{d}\widetilde{W}_s, \qquad t \geq 0. \tag{3.2}$$

In (3.2), $\widetilde{W}$ is standard $\mathbb{F}$-Brownian motion that is jointly Gaussian with (B, W); $\widetilde{b}$ is locally bounded and $\mathbb{F}$-adapted; $\rho^{(0)}$ is an $\mathbb{F}$-adapted locally bounded process such that for all $T > 0$,

$$\mathbb{E}\big[1 \wedge |\rho_t^{(0)} - \rho_s^{(0)}|\big] \leq K_2 |t-s|^{\gamma}, \qquad s,t \in [0,T], \tag{3.3}$$

for some $\gamma \in (\frac{1}{2}, 1]$ and $K_2 \in (0, \infty)$; and $\widetilde{\rho}$ is an $\mathbb{F}$-adapted locally bounded process such that for all $T > 0$, there exist $\varepsilon > 0$ and $K_3 \in (0, \infty)$ with

$$\mathbb{E}\big[1 \wedge |\widetilde{\rho}_t - \widetilde{\rho}_s|\big] \leq K_3 |t-s|^{\varepsilon}, \qquad s,t \in [0,T]. \tag{3.4}$$

(iv) We have (2.2) with $H \in (0, \frac{1}{2})$ and some $g_0 \in C^\infty([0,\infty))$ with $g_0(0) = 0$.

The following CLT is our main technical result. We write $\mu_f(v) = \mathbb{E}[f(Z)]$ and $\gamma_f(v,q) = \mathrm{Cov}(f(Z), f(Z'))$, where (Z, Z') follows a centered bivariate normal distribution with $\mathrm{Var}(Z) = \mathrm{Var}(Z') = v$ and $\mathrm{Cov}(Z, Z') = q$. Moreover, we define

$$\Gamma_0^H = 1 \qquad \text{and} \qquad \Gamma_r^H = \frac{1}{2}\big((r+1)^{2H} - 2r^{2H} + (r-1)^{2H}\big), \quad r \geq 1, \tag{3.5}$$

and use $\overset{\mathrm{st}}{\Longrightarrow}$ (resp., $\overset{L^1}{\Longrightarrow}$) to denote functional stable convergence in law (resp., convergence in L^1) in the space of càdlàg functions equipped with the local uniform topology. In the parametric setup of an mfBM, the CLT for the test function $f(x) = x^2$ was obtained by Dozzi et al. (2015).

Theorem 3.1. *Grant Assumption (CLT) and let $N(H) = \lfloor 1/(2-4H) \rfloor$ for $H \in (0, \frac{1}{2})$. Then*

$$\Delta_n^{-\frac{1}{2}} \left\{ V_f^n(Y,t) - \int_0^t \mu_f(\rho_s^2)\,\mathrm{d}s - \sum_{j=1}^{N(H)} \frac{\Delta_n^{j(1-2H)}}{j!} \int_0^t \mu_f^{(j)}(\rho_s^2)\sigma_s^{2j}\,\mathrm{d}s \right\} \overset{\mathrm{st}}{\Longrightarrow} \mathcal{Z}, \tag{3.6}$$

where $\mu_f^{(j)}$ denotes the jth derivative of μ_f and $\mathcal{Z} = (\mathcal{Z}_t)_{t \geq 0}$ is a continuous process defined on a very good filtered extension $(\overline{\Omega}, \overline{\mathcal{F}}, (\overline{\mathcal{F}}_t)_{t\geq 0}, \overline{\mathbb{P}})$ of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ which, conditionally on $\mathcal{F}$, is a centered Gaussian process with independent increments and such that the conditional variance function $C_t = \overline{\mathbb{E}}[\mathcal{Z}_t^2 \mid \mathcal{F}]$ is given by

$$C_t = \int_0^t \left\{ \gamma_f(\rho_s^2, \rho_s^2) + 2\sum_{r=1}^{\infty} \gamma_f(\rho_s^2, \rho_s^2 \Gamma_r^H) \right\} \mathrm{d}s. \tag{3.7}$$

This result can be extended to a multivariate setting; see Theorem B.1 in Appendix B.

Remark 3.2. It suffices to require f be $2(N(H)+1)$-times continuously differentiable with derivatives of at most polynomial growth. An assumption as in (3.2) is standard for CLTs in high-frequency statistics. But here we need it for ρ (instead of σ), as the noise process dominates the efficient price process in the limit $\Delta_n \to 0$. Condition (3.1) on σ is satisfied if, for example, σ is itself a continuous Itô semimartingale. These assumptions do exclude the case of rough volatility. For quadratic functionals (as considered in Corollary 4.1 below), we conjecture that Assumption (CLT) can be relaxed to allow for rough (price and noise) volatility, if further structural assumptions are made concerning volatility of volatility (e.g., if both $\widetilde{\rho}$ and the volatility process of σ are again processes of fractional type); cf. Chong et al. (2023). To keep the exposition simple, we do not consider such an extension here.

Remark 3.3. Both the LLN limit $V_f(Y,t) = \int_0^t \mu_f(\rho_s^2)\,\mathrm{d}s$ and the conditional variance process $\mathcal{Z}$ are driven by the rough component Z. In other words, if $\sigma \equiv 0$ (i.e., in the pure fractional case), we would have (3.6) without the $\sum_{j=1}^{N(H)}$-expression; see Barndorff-Nielsen et al. (2011). Even if $\sigma \not\equiv 0$, in the case where $H < \frac{1}{4}$, no additional terms are present because $N(H) = 0$. This is in line with Proposition 2.3, which states that it is impossible to consistently estimate $C_t = \int_0^t \sigma_s^2\,\mathrm{d}s$ if $H < \frac{1}{4}$. If $H \in (\frac{1}{4}, \frac{1}{2})$, the "mixed" terms in the $\sum_{j=1}^{N(H)}$-expression will allow us to estimate C_t.

Remark 3.4. In the special case where $f(x) = x^{2p}$ for some $p \in \mathbb{N}$, (3.6) reads

$$\Delta_n^{-\frac{1}{2}} \left\{ V_f^n(Y,t) - \mu_{2p} \int_0^t \rho_s^{2p}\,\mathrm{d}s - \sum_{j=1}^{N(H)} \Delta_n^{j(1-2H)} \mu_{2p} \binom{p}{j} \int_0^t \rho_s^{2p-2j}\sigma_s^{2j}\,\mathrm{d}s \right\} \overset{\mathrm{st}}{\Longrightarrow} \mathcal{Z},$$

where μ_{2p} is the moment of order $2p$ of a standard normal variable. Typically, one is interested in estimating only one of the terms in the sum $\sum_{j=1}^{N(H)}$ at a time (e.g., $\int_0^t \sigma_s^{2p}\,\mathrm{d}s$ corresponding to $j = p$). All other terms (e.g., $j \neq p$) have to be considered as higher-order bias terms in this case. The appearance of (potentially many, if $N(H)$ is large) bias terms for test functions as simple as powers of even order neither happens in the pure semimartingale nor in the pure fractional setting.

The proof of Theorem 3.1 is deferred to Appendix C in the supplementary material. In addition to the usual steps that are common to CLTs in high-frequency statistics, there are two new challenges in the present setting:

(i) The observation process Y is *not* a semimartingale (and not even close to one). This is because the rough component Z *dominates* the efficient price process X in the limit as $\Delta_n \to 0$. In particular, the increments of Y remain conditionally *dependent* as $\Delta_n \to 0$.

(ii) If H is close to (but smaller than) $\frac{1}{2}$, the semimartingale part is only marginally smoother than the noise part. So for the CLT, there will be an intricate interplay between the efficient price process and the noise process.

To overcome the first challenge, we employ a multiscale analysis: by suitably truncating the increments of Y, we can restore, to some degree (not on the finest scale Δ_n but on some intermediate scale $\theta_n\Delta_n$ where $\theta_n \to \infty$), asymptotic conditional independence between increments of Y (see Lemma C.2). This in turn gives $V_f^n(Y,t)$, as a process in t, a semimartingale-like structure on this intermediate scale, which is sufficient for deriving the CLT when we center by appropriate conditional expectations (see (C.15)). However, because increments are still correlated on the finest scale, the limiting process is not the usual one for semimartingales but the one for fractional Brownian motion (see (B.9), in particular). Regarding the second challenge above, we find, to our surprise, that the semimartingale component *never* enters the CLT limit of $V_f^n(Y,t)$ when centered by conditional expectations (see Lemma C.3), no matter how close H is to $\frac{1}{2}$. By contrast, it does affect the limit behavior of these conditional expectations (Lemmas C.4–C.18), producing an H-dependent number of higher-order bias terms that neither appear in the pure semimartingale nor in the pure fractional setting.

4 Estimating the roughness parameter and integrated price and noise volatilities

In this section, we develop an estimation procedure for the roughness parameter of the noise and the integrated price (if $H > \frac{1}{4}$) and noise volatilities, that is, for H, $C_t = \int_0^t \sigma_s^2 \,\mathrm{d}s$ and $\Pi_t = \int_0^t \rho_s^2 \,\mathrm{d}s$. To avoid additional bias terms (cf. Remark 3.4), we use quadratic functionals only, that is, we consider $f_r(x) = x_1x_{r+1}$ for $x = (x_1,\dots,x_{r+1}) \in \mathbb{R}^{r+1}$ and $r \in \mathbb{N}_0$ and the associated variation functionals $V_{r,t}^n = V_{f_r}^n(Y,t) = \Delta_n^{1-2H}\sum_{k=1}^{\lfloor t/\Delta_n\rfloor - r} \Delta_k^n Y \Delta_{k+r}^n Y$. (This is a multivariate variation functional as considered in Appendix B.) Note that $V_{r,t}^n$ is not a statistic as it depends on the unknown parameter H. Therefore, we introduce $\widehat{V}_t^n = (\widehat{V}_{0,t}^n,\dots,\widehat{V}_{R,t}^n)^T$, a non-normalized version of $V_{r,t}^n$ that is a statistic:

$$\widehat{V}_{r,t}^n = \widehat{V}_{f_r}^n(Y,t) = \sum_{k=1}^{\lfloor t/\Delta_n\rfloor - r} \Delta_k^n Y \Delta_{k+r}^n Y, \qquad r \in \mathbb{N}_0.$$

Clearly, we have $\Delta_n^{1-2H}\widehat{V}_{r,t}^n = V_{r,t}^n$, so a multivariate extension of Theorem 3.1 (see Theorem B.1 in the appendix) immediately yields:

Corollary 4.1. *Let $\widehat{V}_t^n = (\widehat{V}_{0,t}^n,\dots,\widehat{V}_{R,t}^n)^T$ for a fixed but arbitrary $R \in \mathbb{N}_0$. For $H \in (0,\frac{1}{2})$,*

$$\Delta_n^{-\frac{1}{2}}\Big\{\Delta_n^{1-2H}\widehat{V}_t^n - \Gamma^H\int_0^t \rho_s^2\,\mathrm{d}s - e_1\int_0^t \sigma_s^2\,\mathrm{d}s\,\Delta_n^{1-2H}\mathbf{1}_{\{\frac{1}{4},\frac{1}{2}\}}(H)\Big\} \xRightarrow{\text{st}} \mathcal{Z}, \tag{4.1}$$

where $\Gamma^H = (\Gamma_0^H,\dots,\Gamma_R^H)^T$, $e_1 = (1,0,\dots,0)^T \in \mathbb{R}^{1+R}$ and $\mathcal{Z}$ is an $\mathbb{R}^{R+1}$-valued continuous process defined on $(\overline{\Omega},\overline{\mathcal{F}},(\overline{\mathcal{F}}_t)_{t\geq 0},\overline{\mathbb{P}})$ which, conditionally on $\mathcal{F}$, is a centered Gaussian process with independent increments such that for all $r,r' = 0,\dots,R$,

$$\mathcal{C}_{r,r'}^H(t) = \overline{\mathbb{E}}[\mathcal{Z}_t^r\mathcal{Z}_t^{r'} \mid \mathcal{F}] = \mathcal{C}_{r,r'}^H\int_0^t \rho_s^4\,\mathrm{d}s, \qquad \mathcal{C}_{r,r'}^H = v_{r,r'}^{H,0} + \sum_{k=1}^{\infty}(v_{r,r'}^{H,k} + v_{r',r}^{H,k}),$$
$$v_{r,r'}^{H,k} = \mathrm{Cov}(\Delta B_{1+k}^H\Delta B_{1+k+r}^H, \Delta B_1^H\Delta B_{1+r'}^H) = \Gamma_k^H\Gamma_{|r-r'+k|}^H + \Gamma_{|k-r'|}^H\Gamma_{k+r}^H, \tag{4.2}$$

where $\Delta B_t^H = B_t^H - B_{t-1}^H$ for a standard fractional Brownian motion B^H.

4.1 Asymptotically mixed normal estimators

The simplest estimator for H is obtained by calculating the rate of divergence in volatility signature plots, that is, by regressing $\log\Delta_n$ on $\log\widehat{V}_{0,t}^n$ (see also Rosenbaum (2011) for a more general but related concept). However, as noted by Dozzi et al. (2015) in their Remark 3.1, already in an mfBM model, this estimator only has a logarithmic rate of convergence. Indeed, as our simulation study in Section 5 shows, this estimator systematically overestimates H unless H is very close to $\frac{1}{2}$. In the pure fractional case, rate-optimal estimators are given by so-called change-of-frequency or autocorrelation estimators (Barndorff-Nielsen et al. 2011, Corcuera et al. 2013). Both extract information about H by considering the ratio of (different combinations of) $\widehat{V}_{r,t}^n$ for different values of r. For example, the simplest autocorrelation estimator is

$$\widehat{H}_{\text{acf}}^n = \frac{1}{2}\Big[1 + \log_2\Big(\frac{\widehat{V}_{1,t}^n}{\widehat{V}_{0,t}^n} + 1\Big)\Big], \tag{4.3}$$

which is based on the fact that $\widehat{V}_{1,t}^n/\widehat{V}_{0,t}^n = V_{1,t}^n/V_{0,t}^n \xrightarrow{\mathbb{P}} \Gamma_1^H = 2^{2H-1} - 1$. But due to the bias term that appears in (4.1) when $r = 0$, the convergence rate worsens and becomes suboptimal when (4.3) is applied to mixed semimartingales. A simple way to circumvent this problem is to consider ratios of $\widehat{V}_{r,t}^n$ for two different values of $r \neq 0$. This indeed leads to estimators of H with rate of convergence $\Delta_n^{-1/2}$, and the first rate-optimal estimator of H in the case of mfBM, which is constructed in Theorem 3.2 of Dozzi et al. (2015), is exactly of this type. However, these estimators suffer from a fundamental identification problem: if the observed price is simply $Y = \sigma B$ for some constant $\sigma > 0$ (i.e., there is no noise), then, by standard CLTs for Brownian motion, the ratio $\widehat{V}_{r_1,t}^n/\widehat{V}_{r_2,t}^n$ (for $r_1 \neq r_2$ with $r_1 \neq 0$ and $r_2 \neq 0$) converges stably in law to the ratio Z_1/Z_2 of two independent centered normal random variables. Because Z_1/Z_2 has a density supported on $\mathbb{R}$, estimators based on such ratios can generate estimates from any non-empty open interval with positive probability. Thus, with such estimators, it is impossible to tell whether there is evidence of rough noise or whether a small estimate of H is simply the result of chance. Even if H is less than but close to $\frac{1}{2}$, the finite-sample variance is so large that in their Remark 3.2, Dozzi et al. (2015) do not recommend using such estimators in practice.

Our strategy, by contrast, uses all lags $r = 0,\dots,R$ for some finite $R \in \mathbb{N}$ and is a two-step procedure. In a first step, we use the statistic

$$\widehat{T}^n = \frac{\widehat{V}_{1,t}^n}{\sqrt{\sum_{k=1}^{\lfloor t/\Delta_n\rfloor - 1}(\Delta_k^n Y\Delta_{k+1}^n Y)^2}} \tag{4.4}$$

to test for the presence of noise. If there is no noise (i.e., if $Y = X$ is a semimartingale), then $\widehat{T}^n \xrightarrow{\text{st}} N(0,1)$; cf. Theorem 8 of Andersen et al. (2023). If there is noise (i.e., $H < \frac{1}{2}$ and $\Pi_t \neq 0$), then it is easy to see that $\widehat{T}^n \to -\infty$ at a rate of $\Delta_n^{-1/2}$. Therefore, if $\widehat{T}^n > -q_n$ where $q_n = q \log \Delta_n^{-1}$ for some $q > 0$, we set

$$\widehat{H}^n = \tfrac{1}{2}, \quad \widehat{\Pi}_t^n = 0, \quad \widehat{C}_t^n = \widehat{V}_{0,t}^n. \tag{4.5}$$

In the absence of noise, this happens with probability converging to 1.

If $\widehat{T}^n \leq -q_n$, we construct an estimator $\widehat{\theta}_t^n = (\widehat{H}^n, \widehat{\Pi}_t^n, \widehat{C}_t^n)$ of $\theta_t = (H, \Pi_t, C_t)$ using a generalized method of moments (GMM) approach (Hansen 1982), by solving

$$\underset{\theta=(H,\Pi,C)}{\arg\min} \left\{ \|\widehat{\mathcal{W}}_n^{1/2}(\hat{V}_t^n - \Delta_n^{2H-1}\Pi\Gamma^H - Ce_1)\|_2^2 \quad \text{subject to } \Pi, C \geq 0,\ H \in (0, \tfrac{1}{2}] \right\}, \tag{4.6}$$

or rather

$$F_n(\theta) = \nabla_\theta \|\widehat{\mathcal{W}}_n^{1/2}(\hat{V}_t^n - \Delta_n^{2H-1}\Pi\Gamma^H - Ce_1)\|_2^2 = 0 \qquad \text{on } (0, \tfrac{1}{2}] \times [0, \infty)^2, \tag{4.7}$$

where $\widehat{\mathcal{W}}_n$ is a (possibly random) sequence of symmetric positive definite matrices in $\mathbb{R}^{(R+1)\times(R+1)}$ and $\|\cdot\|_2$ denotes the Euclidean norm. The main theorem of this paper is the following.

Theorem 4.2. *Suppose that (ii)–(iv) of Assumption (CLT) are satisfied. Further suppose that $R \geq 2$ and that $\widehat{\mathcal{W}}_n \xrightarrow{\text{P}} \mathcal{W}$, where $\mathcal{W} \in \mathbb{R}^{(R+1)\times(R+1)}$ is a deterministic symmetric positive definite matrix.*

(i) If $H \in (\frac{1}{4}, \frac{1}{2})$ and $\Pi_t, C_t > 0$ almost surely, then there exists a sequence of estimators $\widehat{\theta}_t^n = (\widehat{H}^n, \widehat{\Pi}_t^n, \widehat{C}_t^n)$ of $\theta_t = (H, \Pi_t, C_t)$ such that $\mathbb{P}(F_n(\widehat{\theta}_t^n) = 0) \to 1$ and

$$D_n(t)^{-1}(\widehat{\theta}_t^n - \theta_t) \xrightarrow{\text{st}} (E(t)E(t)^T)^{-1}E(t)\mathcal{W}^{1/2}\mathcal{Z} \tag{4.8}$$

where $\mathcal{Z}$ is the same process as in Corollary 4.1 and

$$D_n(t) = \begin{pmatrix} \Delta_n^{1/2} & 0 & 0 \\ 2\Delta_n^{1/2}|\log\Delta_n|\Pi_t & \Delta_n^{1/2} & 0 \\ 0 & 0 & \Delta_n^{2H-1/2} \end{pmatrix} \quad \text{and} \quad E(t) = (\Pi_t \partial_H \Gamma^H, \Gamma^H, e_1)^T \mathcal{W}^{1/2}.$$

In the last line, $\partial_H \Gamma^H$ is the entrywise derivative of Γ^H with respect to H.

(ii) If $H \in (0, \frac{1}{4})$ and $\Pi_t > 0$ almost surely, then there is a sequence of estimators $\widehat{\theta}_t^n = (\widehat{H}^n, \widehat{\Pi}_t^n)$ of $\theta_t' = (H, \Pi_t)$ such that $\mathbb{P}(F_n'(\widehat{\theta}_t^n) = 0) \to 1$ and

$$D_n'(t)^{-1}(\widehat{\theta}_t^n - \theta_t') \xrightarrow{\text{st}} (E'(t)E'(t)^T)^{-1}E'(t)\mathcal{W}^{1/2}\mathcal{Z}, \tag{4.9}$$

where $\mathcal{Z}$ is the same process as in Corollary 4.1,

$$F_n'(\theta') = F_n'(H, \Pi) = \nabla_{\theta'} \|\widehat{\mathcal{W}}_n^{1/2}(\hat{V}_t^n - \Delta_n^{2H-1}\Pi\Gamma^H)\|_2^2 \tag{4.10}$$

and

$$D_n'(t) = \begin{pmatrix} \Delta_n^{1/2} & 0 \\ 2\Delta_n^{1/2}|\log\Delta_n|\Pi_t & \Delta_n^{1/2} \end{pmatrix} \quad \text{and} \quad E'(t) = (\Pi_t \partial_H \Gamma^H, \Gamma^H)^T \mathcal{W}^{1/2}.$$

(iii) In the setup of (i) (resp., (ii)), the sequences $(\widehat{\theta}_t^n)_{n\in\mathbb{N}}$ (resp., $(\widehat{\theta}_t'^n)_{n\in\mathbb{N}}$) are locally unique in the sense that if $\widetilde{\theta}_t^n$ also satisfies $\mathbb{P}(F_n(\widetilde{\theta}_t^n) = 0) \to 1$ and $\mathbb{P}(\|\widetilde{\theta}_t^n - \theta_t\| \leq 1/(\log\Delta_n)^2) \to 1$ (resp., $\mathbb{P}(F_n'(\widetilde{\theta}_t'^n) = 0) \to 1$ and $\mathbb{P}(\|\widetilde{\theta}_t'^n - \theta_t'\| \leq 1/(\log\Delta_n)^2) \to 1$), then $\mathbb{P}(\widehat{\theta}_t^n = \widetilde{\theta}_t^n) \to 1$ (resp., $\mathbb{P}(\widehat{\theta}_t'^n = \widetilde{\theta}_t'^n) \to 1$). Moreover, in the situation considered in (ii), if $\widehat{\theta}_t^n = (\widehat{H}^n, \widehat{\Pi}_t^n, \widehat{C}_t^n)$ satisfies $\mathbb{P}(F_n(\widehat{\theta}_t^n) = 0) \to 1$, then (4.9) continues to hold with $\widehat{\theta}_t'^n = (\widehat{H}^n, \widehat{\Pi}_t^n)$ instead of $\widehat{\theta}_t'^n$.

The proof can be found in Appendix D in the supplement and uses the theory of estimating equations (Jacod & Sørensen 2018, Mies & Podolskij 2023) to derive (4.8) and (4.9) from Corollary 4.1. The rates of convergence of $\widehat{H}^n$, $\widehat{\Pi}_t^n$ and $\widehat{C}_t^n$ (if $H > \frac{1}{4}$) are $\Delta_n^{-1/2}$, $\Delta_n^{-1/2}/|\log\Delta_n|$ and $\Delta_n^{1/2-2H}$, respectively. The additional logarithmic factor in estimating Π_t is due to the fact that H is unknown and already appears in the pure fractional setting (Brouste & Fukasawa 2018). The rate of convergence of $\widehat{C}_t^n$ decreases with H and C_t can no longer be consistently estimated if $H < \frac{1}{4}$. Therefore, the previous theorem yields a quantitative version of Proposition 2.3.

If $H < \frac{1}{4}$, there is no way to estimate C_t consistently on a finite time interval. This is why the case $H < \frac{1}{4}$ in (ii) has to be stated separately from part (i) where $H > \frac{1}{4}$. However, as a consequence of the last part of Theorem 4.2, there is no need in practice to know or distinguish whether $H < \frac{1}{4}$ or $H > \frac{1}{4}$, and we always recommend solving (4.6) to obtain estimates of H, Π_t and C_t. If $H > \frac{1}{4}$, we know that the resulting estimators are asymptotically mixed normal. If $H < \frac{1}{4}$, we still have asymptotic normality for the estimators of H and Π_t but the estimator of C_t is no longer consistent.

4.2 Feasible implementation

In order to obtain a consistent estimator of the asymptotic variance of $\widehat{H}^n$, $\widehat{\Pi}_t^n$ and $\widehat{C}_t^n$ (if $H \in (\frac{1}{4}, \frac{1}{2})$), we proceed analogously to Li & Xiu (2016) and define

$$\begin{aligned} \widehat{\Sigma}_n &= \widehat{\Sigma}_n^{(0)} + \sum_{\ell=1}^{\ell_n} K(\ell, \ell_n)(\widehat{\Sigma}_n^{(\ell)} + (\widehat{\Sigma}_n^{(\ell)})^T), \\ \widehat{\Sigma}_n^{(\ell)} &= \Delta_n \sum_{i=\ell+1}^{[t/\Delta_n]-R} \eta^{(i)}(\eta^{(i-\ell)})^T \in \mathbb{R}^{(R+1)\times(R+1)}, \qquad \eta^{(i)} = (\eta_0^{(i)}, \dots, \eta_R^{(i)})^T, \\ \eta_r^{(i)} &= \Delta_i^n Y \Delta_{i+r}^n Y - \widehat{m}_i^{n,r}, \qquad \widehat{m}_i^{n,r} = \frac{1}{k_n} \sum_{j=0}^{k_n-1} \Delta_{i+j}^n Y \Delta_{i+j+r}^n Y, \\ \widehat{\zeta}_n &= (\Delta_n^{2\widehat{H}^n}\widehat{\Pi}_t^n(\partial_H\Gamma^{\widehat{H}^n} - 2|\log\Delta_n|\Gamma^{\widehat{H}^n}), \Delta_n^{2\widehat{H}^n}\Gamma^{\widehat{H}^n}, \Delta_n e_1) \in \mathbb{R}^{(R+1)\times 3}, \end{aligned} \tag{4.11}$$

where K is a deterministic kernel function and k_n and ℓ_n are integer sequences.

Corollary 4.3. *Assume the conditions of Theorem 4.2 and that K is uniformly bounded with $K(\ell, \ell_n) \to 1$ for every fixed $\ell \geq 1$. Further suppose that k_n and ℓ_n increase to infinity such that $\ell_n/\sqrt{k_n} \to 0$ and $\ell_n\sqrt{k_n\Delta_n} \to 0$. If we denote the diagonal elements of the 3×3-matrix*

$$\mathbb{V}_n = \Delta_n(\widehat{\zeta}_n^T\widehat{\mathcal{W}}_n\widehat{\zeta}_n)^{-1}\widehat{\zeta}_n^T\widehat{\mathcal{W}}_n\widehat{\Sigma}_n\widehat{\mathcal{W}}_n\widehat{\zeta}_n(\widehat{\zeta}_n^T\widehat{\mathcal{W}}_n\widehat{\zeta}_n)^{-1} \tag{4.12}$$

by $\mathbb{V}_n^H$, $\mathbb{V}_n^\Pi$ and $\mathbb{V}_n^C$ and the distribution function of the standard normal law by Φ, then for any $\gamma \in (0,1)$,

$$[\widehat{H}^n \pm \Phi^{-1}((1-\gamma)/2)\sqrt{\mathbb{V}_n^H}], \quad [\widehat{\Pi}_t^n \pm \Phi^{-1}((1-\gamma)/2)\sqrt{\mathbb{V}_n^\Pi}], \quad [\widehat{C}_t^n \pm \Phi^{-1}((1-\gamma)/2)\sqrt{\mathbb{V}_n^C}]$$

are, respectively, asymptotic γ-confidence intervals for H, Π_t and C_t (if $H \in (\frac{1}{4}, \frac{1}{2})$).

4.3 Finite-sample considerations

As H approaches $\frac{1}{2}$, distinguishing volatility from a marginally rougher noise term becomes increasingly difficult. In this case, it can happen in finite samples that $\widehat{\Pi}_t^n$ yields a better approximation of C_t, while $\widehat{C}_t^n$ yields a better approximation of Π_t. As Π_t and C_t are not separable in the limit $H = \frac{1}{2}$, there is no way this can be prevented in general. However, if one is willing to incorporate a priori information such as the assumption that Π_t is smaller than C_t (which in our application is supported by previous empirical results of Aït-Sahalia & Yu (2009)), one can restrict the minimization problem (4.6) to solutions where $\Pi \leq C$, which by design eliminates the mix-ups mentioned before.[5] Thus, we implement estimators (H^n, Π_t^n, C_t^n) obtained as follows:

(i) If $\widehat{T}^n > -q_n$, we set $(H^n, \Pi_t^n, C_t^n) = (\widehat{H}^n, \widehat{\Pi}_t^n, \widehat{C}_t^n)$ from (4.5).

(ii) Otherwise, we compute (H^n, Π_t^n, C_t^n) by solving

$$\operatorname*{arg\,min}_{\theta=(H,\Pi,C)} \left\{ \|\widehat{\mathcal{W}}_n^{1/2}(\widehat{V}_t^n - \Delta_n^{2H-1}\Pi\Gamma^H - Ce_1)\|_2^2 \quad \text{subject to } C \geq \Pi \geq 0,\ H \in (0, \tfrac{1}{2}] \right\}.$$

As long as $\Pi_t \leq C_t$ if $H > \frac{1}{4}$, we have $\mathbb{P}(H^n = \widehat{H}^n,\ \Pi_t^n = \widehat{\Pi}_t^n,\ C_t^n = \widehat{C}_t^n \text{ if } H > \frac{1}{4}) \to 1$, which means that (H^n, Π_t^n, C_t^n) is only a finite-sample adjustment of $(\widehat{H}^n, \widehat{\Pi}_t^n, \widehat{C}_t^n)$ and the asymptotic results of Theorem 4.2 and Corollary 4.3 continue to hold for (H^n, Π_t^n, C_t^n).

5 Monte Carlo simulation

We evaluate the performance of H^n, Π^n and C^n when applied to the mfBM model

$$Y_t = X_t + Z_t = \sigma B_t + \rho B_t^H, \qquad t \in [0, T],$$

where $\sigma = 0.02$, B and B^H are independent and $T = 5$ trading days, each of which consists of 6.5 hours or $n = 23{,}400$ seconds. Accordingly, we choose $\Delta_n = 1/n = 1/23{,}400$. The values of H will be taken from the set

$$H \in \{0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45\}. \tag{5.1}$$

We also include "$H = 0.5$" (i.e., $\rho = 0$) and "$H = 0$," in which case $(B_t^0)_{t\in[0,T]}$ is a centered Gaussian white noise with variance $1/2$ (such that increments have variance 1). The value of ρ is chosen dependent on H such that noise accounts for $1/3$ of the log-return variance (i.e., such that $\rho^2\Delta_n^{2H-1}/(\rho^2\Delta_n^{2H-1} + \sigma^2) = \frac{1}{3}$). This choice roughly matches the empirical results of Aït-Sahalia & Yu (2009).

We choose $q_n = 1.645$ as the 95% standard normal quantile, which corresponds to an initial test for the presence of noise using (4.4) at a 5%-level. If "no noise" is rejected, we compute an estimate H^n following the procedure described in Section 4.3 using five days of simulated data. We choose $R = 10$, which corresponds to considering autocovariances up to a lag of ten seconds. Furthermore, we choose $k_n = 300 \approx 2\Delta_n^{-1/2}$, which corresponds to computing the local autocovariances $\widehat{m}_i^n$ in (4.11) over 5-minute intervals. For the computation of $\widehat{\Sigma}_n$, we take the Parzen kernel $K(\ell, \ell_n) = k(\ell/(\ell_n + 1))$, where $k(x) = (1 - 6x^2 + 6x^3)\mathbb{1}_{\{x\leq 1/2\}} + 2(1-x)^3\mathbb{1}_{\{x>1/2\}}$ and ℓ_n is selected according to the optimal procedure of Newey & West (1994). This guarantees that $\widehat{\Sigma}_n$ is positive semidefinite in finite samples and that the optimal ℓ_n, which is of order $\Delta_n^{-1/5}$, satisfies the rate conditions of Corollary 4.3 if k_n is of order $\Delta_n^{-1/2}$. The weight matrix $\widehat{\mathcal{W}}_n$ is chosen as $\widehat{W}_n = (\widehat{\Sigma}_n)^{-1}$ in order to obtain an optimal GMM procedure.

Table 1: Bias, SE and RMSE of H^n

H	0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
Bias	0.0091	-0.0005	-0.0007	-0.0007	-0.0008	-0.0010	-0.0012	-0.0016	-0.0042	-0.0255	-0.0184
SE	0.0122	0.0218	0.0222	0.0226	0.0235	0.0251	0.0280	0.0333	0.0420	0.0625	0.0890
RMSE	0.0153	0.0218	0.0222	0.0226	0.0235	0.0251	0.0280	0.0334	0.0423	0.0675	0.0912

As we can see from Table 1, the resulting estimator H^n is essentially unbiased, except when H is very close to 0 or 0.5 (but even in this case, the bias is very small). As a result, the main contribution to the root-mean-square error (RMSE) of H^n is the standard error (SE), which is increasing in H. This shows that given the same signal-to-noise ratio, estimating H becomes more difficult as H approaches 0.5. This is reasonable as it is impossible in the limit as $H \to 0.5$ to distinguish a fractional from a semimartingale process. At $H = 0.5$, the RMSE of H^n is completely due to the roughly 5% of cases where $\widehat{T}^n$ from (4.4) falsely detects the presence of noise.

Next, we study the distribution of the pivotal quantity $(H^n - H)/\sqrt{\mathbb{V}_n^H}$. As Figure 3 shows, for all considered values of H except $H \in \{0.05, 0.40, 0.45\}$, the sample quantiles of $(H^n - H)/\sqrt{\mathbb{V}_n^H}$ match standard normal quantiles quite well, confirming the finite-sample reliability of the distributional approximations in Theorem 4.2 and Corollary 4.3 for inferential purposes. If $H = 0.05$ (resp., $H \in \{0.40, 0.45\}$), the low (resp., high) quantiles of $(H^n - H)/\sqrt{\mathbb{V}_n^H}$ are essentially flat, while higher (resp., lower) quantiles are approximately normal. This, of course, is due to the fact that $H^n \in [0, 0.5]$ by construction.

Finally, in Figure 4, we compare our estimator H^n with three alternatives: the estimator $\widetilde{H}_{\text{DMS}}^n = \frac{1}{2}(1 + \log_{2+}[(\widehat{V}_{0,t}^{n/4} - \widehat{V}_{0,t}^{n/2})/(\widehat{V}_{0,t}^{n/2} - \widehat{V}_{0,t}^n)])$ of Dozzi et al. (2015), where $\log_{2+} x = \log_2 x$ if $x > 0$ and $\log_{2+} x = 0$ otherwise; the estimator $\widetilde{H}_{\text{VS}}^n = \frac{1}{2}(\widetilde{\beta}_{\text{VS}}^n + 1)$ based on volatility signature plots, where $\widetilde{\beta}_{\text{VS}}^n$ is the slope estimate in a linear regression of $\log \widehat{V}_{0,t}^{n/i}$ on $\log i$ for $i = 1, \ldots, 20$; and the autocorrelation estimator $\widetilde{H}_{\text{acf}}^n$ from (4.3).

As the first plot of Figure 4 shows, the estimators based on volatility signature plots and first-order autocorrelation have large upward biases except when H is very close to $\frac{1}{2}$. The estimator by Dozzi et al. (2015) is essentially bias-free except for $H = 0.5$ where it shows a large upward bias. On top of that, as the second plot shows, the SE of this estimator explodes as H approaches 0.5, confirming Dozzi et al. (2015)'s observation that

[5] The assumption $\Pi_t \leq C_t$ does *not* imply that noise only accounts for a small proportion of the log-return variance. Indeed, this proportion is given by $\Delta_n^{2H-1}\Pi_t/(\Delta_n^{2H-1}\Pi_t + C_t)$, which can be large because of Δ_n^{2H-1} even if $\Pi_t \leq C_t$.

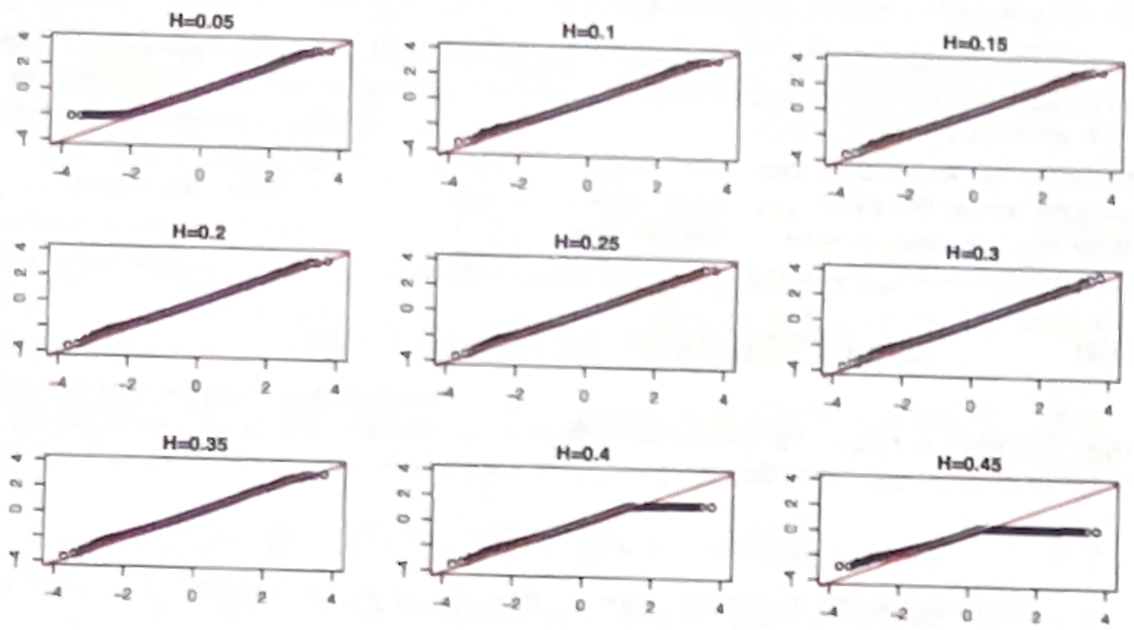


Figure 3: Sample quantiles of $(H^n - H)/\sqrt{V_n^H}$ against standard normal quantiles.

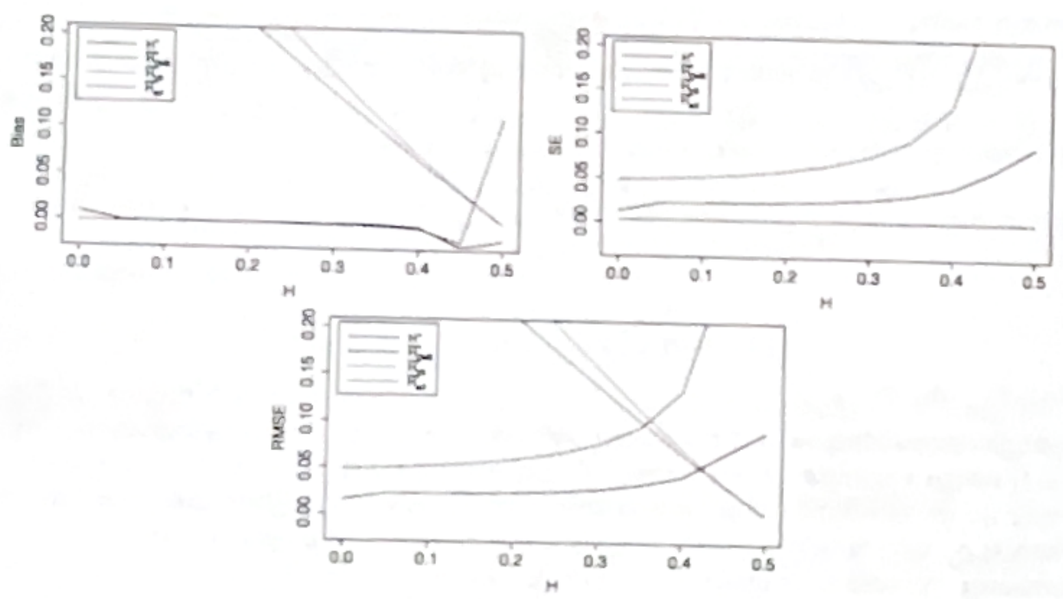


Figure 4: Bias, SE and RMSE of H^n, $\widetilde{H}^n_{\mathrm{DMS}}$, $\widetilde{H}^n_{\mathrm{VS}}$ and $\widetilde{H}^n_{\mathrm{acf}}$ in absolute numbers.

this estimator is highly unstable. For most values of H, our estimator H^n achieves the best RMSE results, which confirms the benefit of initially testing for the presence of noise as a bias–variance trade-off.

Next, we turn to volatility estimation. Having obtained an estimate of H, we estimate $C_T - C_{T-1}$ and $\Pi_T - \Pi_{T-1}$, that is, price and noise volatility on the last trading day by repeating steps (ii) and (iii) in Section 4.3 but with H fixed at the previously obtained

estimate.

Table 2: Bias, SE and RMSE of C^n/σ^2 and Π^n/ρ^2

H	Quantiles of C^n/σ^2					Quantiles of Π^n/ρ^2				
	2.5%	25%	50%	75%	97.5%	2.5%	25%	50%	75%	97.5%
0	0.9545	0.9809	0.9943	1.0076	1.0321	0.9513	1.0159	1.0696	1.3740	2.4227
0.05	0.9517	0.9827	0.9993	1.0166	1.0460	0.3849	0.7162	0.9946	1.3753	2.4168
0.10	0.9460	0.9810	0.9997	1.0184	1.0512	0.3752	0.7144	0.9898	1.3751	2.4606
0.15	0.9371	0.9782	0.9997	1.0212	1.0578	0.3600	0.7091	0.9902	1.3895	2.5175
0.20	0.9234	0.9743	1.0003	1.0249	1.0670	0.3394	0.6899	0.9891	1.4025	2.6643
0.25	0.8997	0.9667	1.0006	1.0314	1.0816	0.3034	0.6625	0.9831	1.4453	2.9399
0.30	0.8488	0.9534	1.0013	1.0411	1.1057	0.2553	0.6195	0.9831	1.5328	3.5292
0.35	0.7008	0.9245	1.0027	1.0613	1.1495	0.1810	0.5461	0.9797	1.7186	5.2889
0.40	0.4371	0.8262	1.0043	1.1116	1.2342	0.0830	0.4092	0.9685	2.3267	6.6144
0.45	0.5378	0.6161	1.0215	1.2578	1.3914	0.0054	0.1464	0.9661	2.8184	3.1781
0.50	0.9758	0.9930	0.9996	1.0060	1.0174	-	-	-	-	-

Table 2 summarizes the performance of Π^n and C^n as respective estimators of $\Pi_T - \Pi_{T-1}$ and $C_T - C_{T-1}$. For all values of $H \leq 0.30$, the performance of C^n is quite good, with a relative error of less than 5% (resp., 16%) in 50% (resp., 95%) of the cases. For $H = 0.5$, the performance is very good, too, with less than 3% error in 95% of the cases. The most difficult case is when H is relatively close to but not equal to 0.5. While C^n remains essentially unbiased in this case, the relative error increases with H. This is expected as it becomes increasingly more difficult as $H \to 0.5$ to statistically distinguish a semimartingale process from a fractional one (the two being indistinguishable in the limit $H = 0.5$). The results for the noise volatility estimator Π^n are qualitatively similar, except that the dispersion of estimates is generally higher in comparison with C^n. One explanation is that noise is smaller than volatility in our simulation (and typically in practice as well).

An interesting observation is that even for $H \leq 0.25$, our estimator C^n yields very precise estimates in the simulation study, although C_T cannot be consistently estimated for $H < \frac{1}{4}$ according to Proposition 2.3. This is because we have fixed the same noise-to-signal ratio for all values of H, which implies that ρ is smaller for smaller values of H. If we had fixed the same ρ for all values of H, then as $\Delta_n \to 0$, the percentage of log-return variance explained by noise increases very fast to 100% for small values to H, which is quite different from what has been observed in practice (Aït-Sahalia & Yu 2009).

6 Empirical analysis

We apply our estimators H^n, C^n and Π^n to SPY transaction data from 2013–2022. In Appendix E of the supplement, we carry out a similar analysis for transaction data of single-name stocks. For each trading day in the ten-year period, we collect all trades from 9:30am to 4:00pm Eastern Time from the TAQ database. We apply mild data cleaning

procedures and sample in calendar time every second using the previous-tick method.[6] All tuning parameters are chosen exactly as in the Monte Carlo. We compute estimates of H on a moving window of five business days and use these estimates to compute daily estimates of integrated price and noise volatility.

Table 3: Quantiles, mean and standard deviation of daily estimates of H and of the NSR

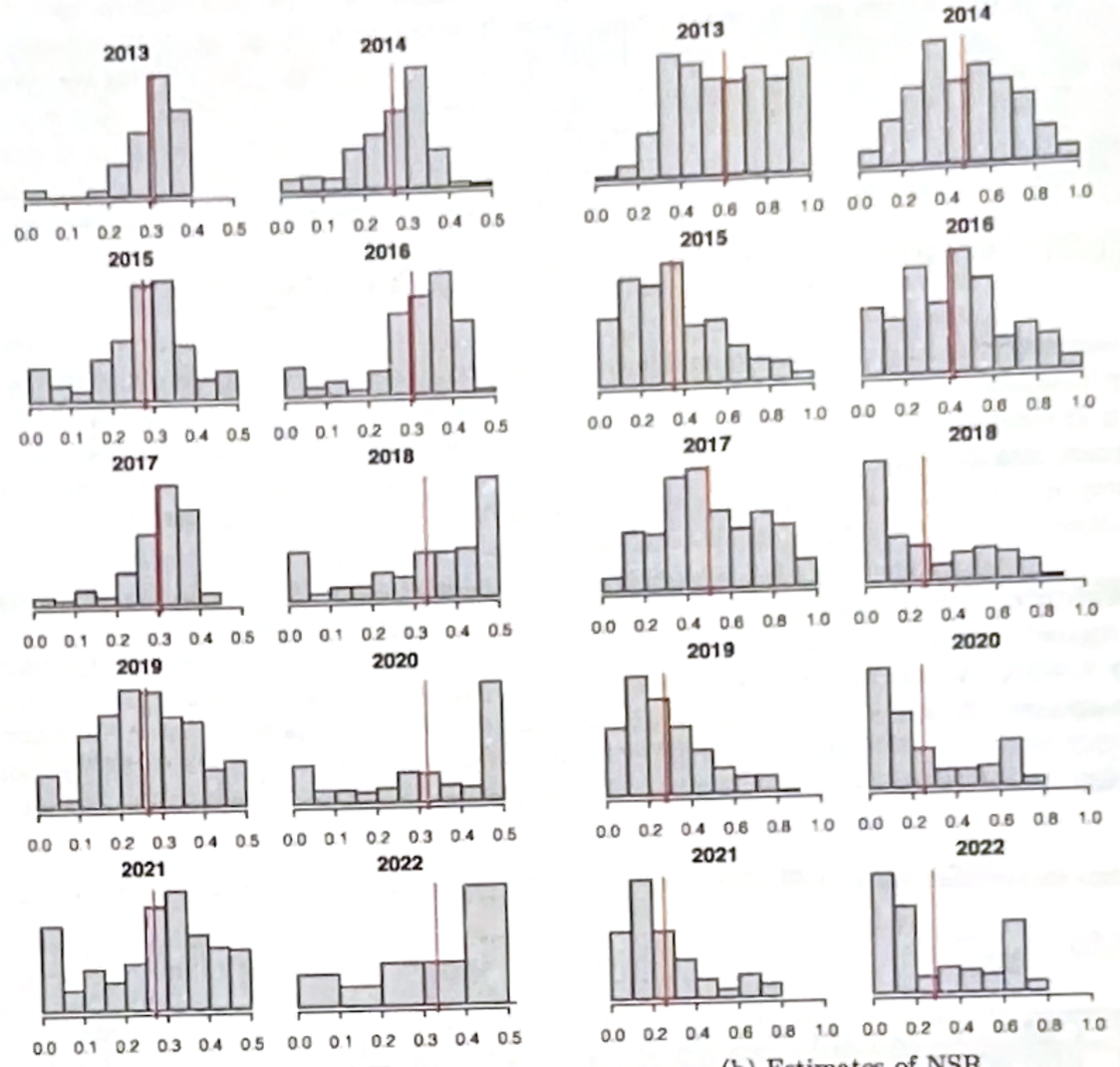
	2.5% Qu.	25% Qu.	50% Qu.	75% Qu.	97.5% Qu.	Mean	SD
Estimates of H	0.0010	0.2344	0.3136	0.3780	0.5000	0.2963	0.1285
Estimates of NSR	0.0000	0.1478	0.3386	0.5722	0.8973	0.3677	0.2537

Table 3 shows summary statistics of the daily estimators of H and of the noise-to-signal ratio (NSR) $\Pi^n \Delta_n^{2H^n-1}/(\Pi^n \Delta_n^{2H^n-1} + C^n)$ for the whole ten-year period. Next, we show in Figure 5 the empirical distributions of the estimates of H and the NSR separately for each year. While the average estimate of the noise roughness parameter remains in the region $[0.25, 0.35]$ for all years, without a prominent trend, the shape of the distribution does change over the years, from a more concentrated distribution in earlier years towards a more spread out one in recent years. Also, the number of days with no or almost no noise (i.e., H close to 0.5) tends to be much higher at the end than at the beginning of the considered period. This is in line with the histograms of the NSR estimates, which show a concentration around smaller values in recent years. In summary, while the average roughness of noise appears to be relatively stable, the average magnitude of noise (relative to volatility) seems to decrease over time. This is in agreement with other research (see e.g., Aït-Sahalia & Xiu (2019)) showing that the level of noise in high-frequency return data has been decreasing in recent years.

To further understand the time-dependence of our estimates, we plot as a function of time the daily estimates of H (including 95%-confidence intervals) in Figure 6[7] and of volatility and the NSR in Figure 7. While the estimates of H exhibit time-dependence in all considered years, the time variation is stronger in later years, confirming our earlier observation that the distribution of H becomes less concentrated around its mean recently. At the same time, the confidence intervals for H are typically wider in the second half of the considered data. This is line with our previous observation that the NSR decreases over time, which makes inference of H harder. We also note that for most of the time, H is significantly different from 0 (white noise case) and from $\frac{1}{2}$ (noise-free case), indicating

[6] We exclude trades with exchange codes `D` and `S`. Furthermore, we only keep trades with trade condition indicator equal to `0`, `00`, `1` or `01` and a trade condition that is either empty or equal to `@`, `C`, `E`, `F`, `I`, `M`, `O`, `Q`, `6` or any combination thereof. To account for outliers due to, for instance, big jumps, we further remove log-returns exceeding in absolute value three times the standard deviation of log-returns of the same day.

[7] To reduce oscillations, we show moving averages of H-estimates obtained as follows: for each day i, we compute an estimate H_i^n using data of the immediate past five days. If $H_i^n = 0.5$ (i.e., no noise is detected), we plot this estimate of H without confidence intervals. If $H_i^n < 0.5$ (i.e., there is noise), we plot $\sum_{j=0}^{3} a_{ij}^n H_{i-5j}^n$ as a point estimate of H together with corresponding confidence intervals. The weights a_{ij}^n are chosen to sum up to one and inversely proportional to the estimated asymptotic variance of H_{i-5j}^n. As each H_{i-5j}^n is an asymptotically unbiased estimator of H and H_i^n, H_{i-5}^n, H_{i-10}^n and H_{i-15}^n are asymptotically independent, this choice minimizes the asymptotic mean-squared error among all convex combinations of H_i^n, H_{i-5}^n, H_{i-10}^n and H_{i-15}^n. If any of the estimates H_{i-5j}^n with $j = 1, 2, 3$ equals 0.5 (i.e., the test in (i) of Section 4.3 fails to detect noise), we exclude it by setting its weight to 0.



(a) Estimates of H (b) Estimates of NSR

Figure 5: Histogram of daily estimates of H (a) and of the NSR (b) based on 1 second SPY transaction data over a period of ten years, with the mean indicated by a red line.

the presence of rough noise in the data. A notable exceptionis the period around the onset of the COVID-19 pandemic in spring 2020, where the data almost appears as noise-free.

Let us mention that the empirical evidence of rough noise reported in this section is not of universal nature and depends strongly on the considered asset and the time period. For instance, when analyzing single-name stocks in Appendix E, we find substantial variation in the roughness of noise, with some stocks such as Intel (INTC) or Coca-Cola (KO) exhibiting rough noise throughout most days in 2019, while other stocks such as Travelers (TRV) or United Technologies (UTX) showed little roughness during the same period. Finally, let us stress that our empirical findings also depend on the data cleaning method applied to the raw data. For instance, if we only kept trades of SPY on NYSE and NASDAQ (which is also a common cleaning procedure), then the resulting price series is almost noise-free in the most recent years of the sample.

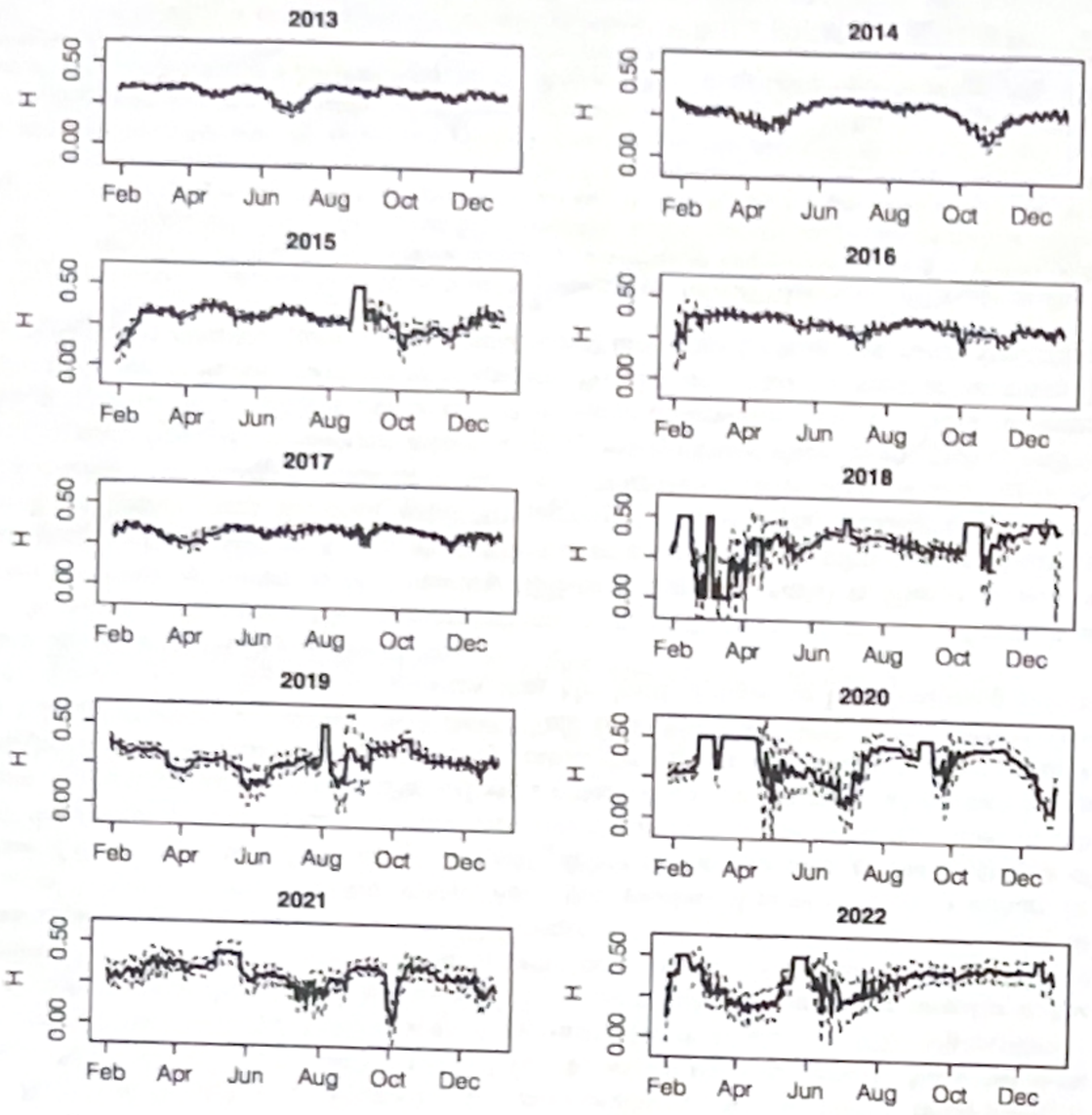


Figure 6: Daily estimates of H with asymptotic 95%-confidence intervals.

7 Conclusion and future directions

Volatility estimation based on high-frequency return data is often impeded by the presence of market microstructure noise. In this paper, we propose to model microstructure noise as a continuous-time rough stochastic process. A distinctive feature of these mixed semimartingale models is a non-shrinking noise component with shrinking increments, which can explain a rich variety of scaling exponents in volatility signature plots.

Using CLTs for variation functionals and a GMM approach, we construct consistent and asymptotically mixed normal estimators for the roughness parameter H of the noise and the integrated price and noise volatilities, whenever these quantities are identifiable. In an empirical application, we find evidence of rough noise in high-frequency return data.

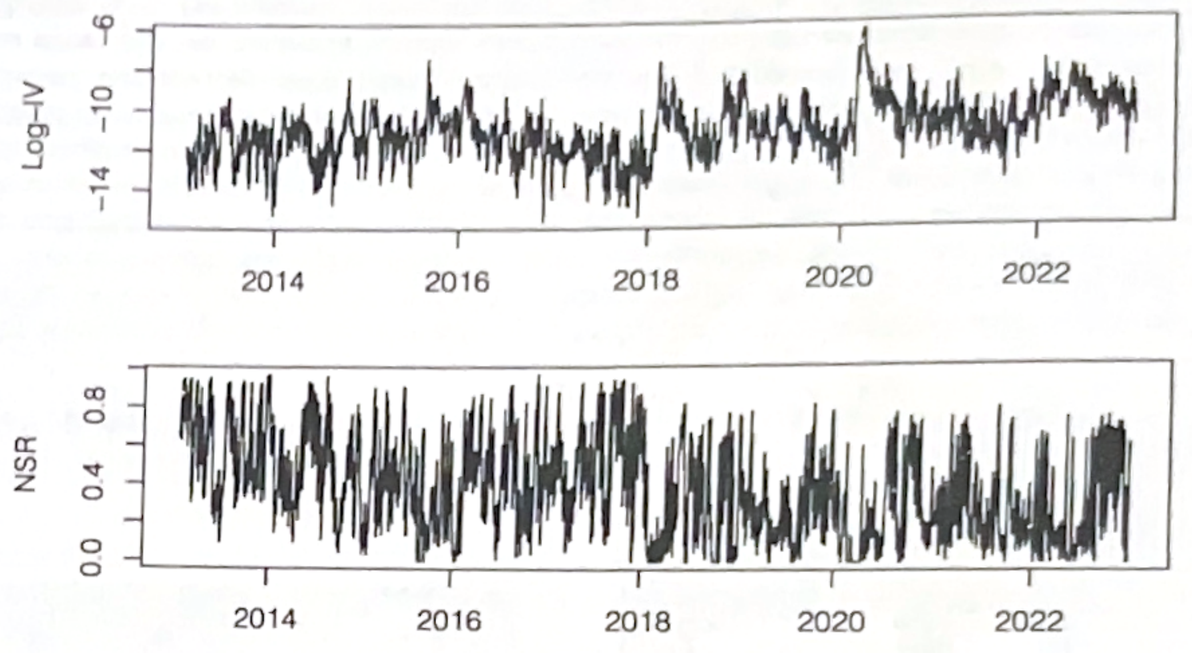


Figure 7: Daily estimates of log-integrated variance and of the NSR.

In this first paper, we do not examine the effect of jumps (Aït-Sahalia & Jacod 2009, Jacod & Todorov 2014) or irregular observation times (Barndorff-Nielsen & Shephard 2005, Chen et al. 2020, Jacod et al. 2017, 2019) on our estimators. Similarly, the current mixed semimartingale model does not capture rounding effects in observed prices (Aït-Sahalia & Jacod 2014, Delattre & Jacod 1997, Robert & Rosenbaum 2010, 2012), which are particularly relevant at the highest sampling frequencies. We leave it to future research to develop estimators that are robust to the aforementioned features of high-frequency data.

A current shortcoming of the mixed semimartingale model is that price volatility cannot be consistently estimated for $H < \frac{1}{4}$. At the same time, our simulation study shows that the estimators of volatility perform very well, even for $H < \frac{1}{4}$, at practically relevant levels of the noise-to-signal ratio. Therefore, an interesting future direction of research is to examine whether, and how, price volatility can be consistently estimated for all values of H if the noise volatility coefficient is assumed to be shrinking.

A Does microstructure noise exist in continuous time?

In the classical Roll (1984) model of transaction prices, deviations of the observed from the efficient price are due to bid–ask bounces associated to each single trade. This raises the question whether Assumption (Z), which postulates the existence of noise in continuous time, is appropriate. Moreover, another important source of noise is the discreteness of prices (see Harris (1990, 1991) and Delattre & Jacod (1997), Li & Mykland (2007), Robert & Rosenbaum (2010, 2012), Rosenbaum (2009)), which is clearly not satisfied by (2.3).

These seeming contradictions between classical market microstructure theory and our mixed semimartingale model can be resolved by taking into account the time scale at

which prices are observed. At low to medium frequency (e.g., if $\Delta_n \geq 5\,\text{min}$) and for liquid assets, it is a well established practice to consider noise as negligible and observed prices as essentially following semimartingale processes.[8] As Δ_n enters a high-frequency regime, noise becomes noticeable and even dominates when Δ_n approaches a few seconds. Finally, at ultra-high frequency, eventually all trades are recorded tick by tick and both transaction times and observed prices become discrete.

Without doubt, estimating volatility using tick-by-tick data (see, for example, Jacod et al. (2019), Li et al. (2014), Robert & Rosenbaum (2010, 2012)) necessitates a careful modeling of rounding effects and bid–ask bounces in prices. However, as we can see from Figure 8, prices sampled at 1 second in our 2019 SPY data do not show much discreteness or flat periods as opposed to, for example, a typical price path in 1999, which was before the decimalization on US stock exchanges. This is in agreement with our previous observation from Figure 1 (b) that price increments are still shrinking[9] at the frequencies we consider (rounding errors would induce a flattening in variance plots). As a result, rounding effects and bid–ask bounces do not seem to be the dominant source of noise in the data and at the frequency we consider.

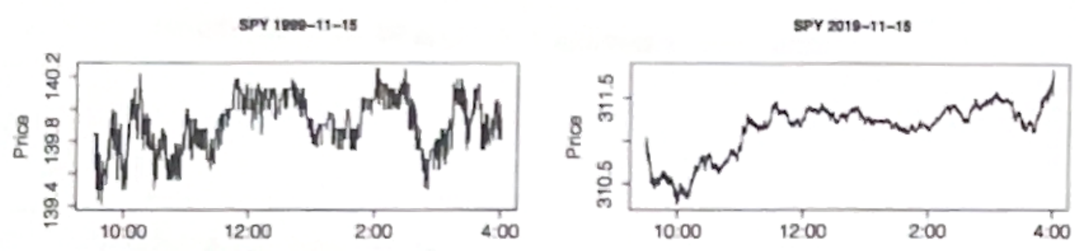


Figure 8: Two paths of SPY transaction prices, one from 1999 and one from 2019.

Next, we give two possible explanations for the existence of microstructure noise in continuous time. Both are related to the very reason why the efficient price X is typically assumed to be a semimartingale. First, according to the fundamental theorem of asset pricing, the absence of arbitrage in an idealized frictionless market implies that prices must be semimartingales (Delbaen & Schachermayer 1994). Real markets, of course, have transaction costs (e.g., bid–ask spreads and commissions). Transaction costs do not only generate trade-specific noise in the form of bid–ask bounces (as in the Roll (1984) model), but have the effect that the absence of arbitrage no longer implies the semimartingale property for prices. For example, both fBM and mfBM (which are special cases of our model) are known to not admit arbitrage in the presence of transaction costs (Cherny 2008, Guasoni et al. 2008, Jarrow et al. 2009). In other words, even if noise due to trading mechanisms is taken away, transaction costs may lead to an additional continuous noise component.

[8] This property can be realized in our model: The size of increments of Z over large time intervals is determined by the behavior of the kernel g_0 in (2.2) for large t, which is not further specified in our model. For instance, if Z is a standard fBM with $H \in (0, \frac{1}{2})$, $Z_{s+t} - Z_s$ is of lower order than $X_{s+t} - X_s$ for large t, so the effect of noise is negligible.

[9] An important detail: to calculate the variance of increments, we exclude periods of no observations (as they would artificially lower the variance) but include zero returns between identical observed prices.

Second, as shown by Aït-Sahalia & Jacod (2020), many microscopic models of tick-by-tick data are compatible (i.e., functionally converge in law to) macroscopic semimartingale models as time is stretched out. In this framework, microstructure noise can be viewed as the difference between the limiting semimartingale process X and the microscopic tick-by-tick observed price process Y (which evolves as a continuous-time but piecewise constant process). In this approach, the microstructure noise process $Z = Y - X$ is, by definition, a continuous-time process. Moreover, since it bridges a microscopic model with a classical white or colored noise as in (1.3) ("$H = 0$") and a noise-free macroscopic model ("$H = \frac{1}{2}$"), it seems reasonable to assume a locally fractional nature for Z with some $H \in (0, \frac{1}{2})$.

Finally, let us remark that including both a discrete and a continuous noise component would probably yield the most satisfying solution; but this is beyond the scope of the current paper. Also, a theoretical substantiation of the arguments in the previous paragraph (e.g., by exhibiting a tick-by-tick price model that converges to a mixed semimartingale on an intermediate time scale) remains open and is left to future research.

B A multivariate central limit theorem for variation functionals

Theorem B.1 can be extended to a multivariate setting that covers variation functionals of the form

$$V_f^n(Y,t) = \Delta_n \sum_{i=1}^{[t/\Delta_n]-L+1} f\Big(\frac{\underline{\Delta}_i^n Y}{\Delta_n^H}\Big),$$

where $f\colon \mathbb{R}^{d\times L} \to \mathbb{R}^M$ is some test function ($L, M \in \mathbb{N}$), Y is a d-dimensional process and

$$\Delta_i^n Y = Y_{i\Delta_n} - Y_{(i-1)\Delta_n} \in \mathbb{R}^d, \quad \underline{\Delta}_i^n Y = (\Delta_i^n Y, \Delta_{i+1}^n Y, \dots, \Delta_{i+L-1}^n Y) \in \mathbb{R}^{d\times L}. \tag{B.1}$$

In the following set of hypotheses, which is a direct multivariate extension of Assumption (CLT), $\|\cdot\|$ denotes the Euclidean norm (in $\mathbb{R}^n$ if applied to vectors and in $\mathbb{R}^{nm}$ if applied to a matrix in $\mathbb{R}^{n\times m}$).

Assumption (CLT_d). *The observation process Y is given by the sum of X from (1.2) and Z from (2.1) with the following specifications:*

(i) The function $f\colon \mathbb{R}^{d\times L} \to \mathbb{R}^M$ is even and infinitely differentiable. Moreover, all its derivatives (including f itself) have at most polynomial growth.

(ii) Both B and W are independent standard $\mathbb{F}$-Brownian motions in $\mathbb{R}^d$, the drift a is d-dimensional, locally bounded and $\mathbb{F}$-adapted, and σ is an $\mathbb{F}$-adapted locally bounded $\mathbb{R}^{d\times d}$-valued process such that for every $T > 0$, there is $K_1 \in (0,\infty)$ with

$$\mathbb{E}\big[1 \wedge \|\sigma_t - \sigma_s\|\big] \leq K_1|t-s|^{\frac{1}{2}}, \qquad s,t \in [0,T]. \tag{B.2}$$

(iii) The noise volatility process ρ takes the form

$$\rho_t = \rho_t^{(0)} + \int_0^t \widetilde{b}_s\,\mathrm{d}s + \int_0^t \widetilde{\rho}_s\,\mathrm{d}\widetilde{W}_s, \qquad t \geq 0. \tag{B.3}$$