

Common price and volatility jumps in noisy high-frequency data

Markus Bibinger

*Faculty of Mathematics and Computer Science, Philipps-Universität Marburg,
e-mail: bibinger@uni-marburg.de*

Lars Winkelmann*

*Department of Economics, Freie Universität Berlin,
e-mail: Lars.Winkelmann@fu-berlin.de*

Abstract: We introduce a statistical test for simultaneous jumps in the price of a financial asset and its volatility process. The proposed test is based on high-frequency data and is robust to market microstructure frictions. For the test, local estimators of volatility jumps at price jump arrival times are designed using a nonparametric spectral estimator of the spot volatility process. A simulation study and an empirical example with NASDAQ order book data demonstrate the practicability of the proposed methods and highlight the important role played by price volatility co-jumps.

MSC 2010 subject classifications: Primary 62G10; secondary 62M10.

Keywords and phrases: High-frequency data, microstructure noise, non-parametric volatility estimation, volatility jumps.

Received June 2017.

1. Introduction

In recent years the broad availability of high-frequency intra-day financial data has spurred a considerable collection of works dedicated to statistical modeling and inference for such data. Semimartingales are a general class of time-continuous stochastic processes to model dynamics of intra-day log-prices in accordance with standard no arbitrage conditions. We consider a general Itô semimartingale log-price model allowing for stochastic volatility, price and volatility jumps as well as leverage. Uncertainty and risk in these models are usually ascribed to two distinct sources: First, the volatility process of the continuous semimartingale part that permanently influences observed returns and, second, occasional jumps in prices. Modeling and inference on the two components constitutes a core research topic in statistics, finance and econometrics bringing forth the seminal contributions by [6], [7], [10], [4] and much more literature devoted to this aspect. For asset pricing ([20], [42]), macro and monetary economics ([43]) and risk management ([34]) information about jumps is of key

*Lars Winkelmann acknowledges financial support from the Deutsche Forschungsgemeinschaft via CRC 649 ‘Ökonomisches Risiko’, Humboldt-Universität zu Berlin.

importance. While the literature on price jumps is well developed from both a statistical and empirical point of view, methods and evidence about volatility jumps are lagging behind. Empirical evidence about volatility jumps is usually based on methods for price jumps applied to an observable volatility measure like the index of implied volatility of S&P 500 index options (VIX), see [17] and [41]. Such modeling strategies inevitably restrict the number of target variables and the overall scope of empirical insights. Since price jumps have often been associated with macro announcements or firm specific news, a natural empirical question arises, if prices and their volatilities jump at common times stimulated by the same events, or not. Such common jumps of price and volatility are often excluded in the statistics literature to avoid technical difficulties. Beyond the question if one should include simultaneous jump times in price and volatility in a model, testing locally for volatility jumps opens up new ways to study effects of information processing and volatility persistence. This is also reflected in an increasing interest to separate the leverage effect in a continuous and a jump part in the current literature, see [1] and [30]. The asset pricing model of [39] illustrates economic forces behind contemporaneous price and volatility jumps. In their model, agents learn about the profitability of a firm in a changing political environment. A change in government policy does not only affect the expected profitability of a firm (price jump) but also triggers a simultaneous volatility jump induced by the impact uncertainty of the new policy.

This article presents a statistical test to decide whether intra-day log-prices exhibit common price and volatility jumps. Our main contribution is to extend the pioneering works by Jacod and Todorov [29] and Bandi and Renò [8] and to provide an approach for an observation model that accounts for market microstructure. It is widely acknowledged that due to market microstructure of financial data recorded at high frequencies, as effects of transaction costs and bid-ask bounce, log-prices are not directly adequately modeled by semimartingales. Taking microstructure frictions into account substantially changes statistical properties and involved mathematical concepts of estimators. We introduce a spectral spot volatility estimator for noisy observations. The test generalizes the theory by [29] for non-noisy observations. We obtain a statistical test by a neat combination of a stable central limit theorem at (almost) optimal rate for the spectral spot volatility estimator and a suitable test function. In analogy to [29], the new test is self-scaling in the volatility and rate-optimal. Those two properties are crucial to obtain an efficient method. The development of a test that can cope with noise is of high relevance and importance as Jacod and Todorov [29] already remark in their empirical application: “presence of microstructure noise in the prices is nonnegligible”. We show in simulations that compared to an application of the method by [29] based on skip-sampled returns, we can significantly improve the power of the test.

Jumps in prices and the volatility are of very different nature. Large price jumps become visible through large returns. More precisely, in a high-frequency context truncation techniques as suggested by [35], [32] and [24] can be used to identify returns that involve jumps. Up to some subtle changes due to dilution by microstructure, this remains valid also in the noisy setup. However, the

localization of jump times becomes less precise and more difficult under noise. A first localization method for price jumps in the noisy semimartingale model has been introduced by [21] using wavelets. Other localization approaches are included in [33] and in [16]. We adopt the methods from [16] to estimate the spot volatility in presence of price jumps and also to locate price-jump times by thresholding. Contrarily to price jumps, volatility jumps are latent and not as obvious as price jumps due to the fact that we can not observe the volatility path. The key element to determine volatility jumps will be efficient estimates of the instantaneous volatility from observed prices.

Our spectral spot volatility estimator relies on the Fourier method promoted by [40] and [12] for estimating quadratic (co-)variation, combined with truncation techniques of [16] to deal with price jumps. These methods attain lower variance bounds for integrated volatility estimation from noisy observations and are, compared to simple smoothing methods and especially skip-sampling to lower observation frequencies, more efficient. While we are the first who address the testing problem under noise, consistent spot volatility estimators under noise are available. [45] and [36] present local two-scales estimators and prove stable central limit theorems. The construction of a rate-optimal pre-average estimator is sketched in Section 8.7 of [3]. An alternative approach considering deterministic volatility is presented in [38]. For our estimator, we establish rate-optimality and a stable central limit theorem with smaller asymptotic variance compared to the pre-average approach. The asymptotic theory allows for general heteroscedastic, serially correlated and endogenous noise. With this estimation approach at hand, we design a test, comparing estimated local volatilities and their left limits at the estimated price-jump times. As a special case, this includes a local test for volatility jumps at some fixed time or stopping time. A test with fast convergence rate based on second order asymptotics of the estimator is suggested. While the overarching strategy follows [29], the specific test function and construction in the noisy observation case are different and profit from the spectral estimation methodology. Compared to previous estimation techniques to smooth noise, the asymptotic variance structure of the spectral volatility estimates in Theorem 1 admits a simpler form. This facilitates a test statistic which is self-scaling in the local volatility and thus furnishes an asymptotic distribution free test with the best possible rate. The Monte Carlo study corroborates the high precision of the methods in finite samples. Our data study shows that price volatility co-jumps occur and are practically relevant.

The paper is organized as follows. Section 2 introduces the model and the statistical problem. We discuss the main ideas for the construction of the test including a short review of the approach for non-noisy data. Section 2.2 describes the spectral spot volatility estimation. We state and discuss the assumptions imposed on the model for the asymptotic theory in Section 3.1 before presenting the main results in Section 3.2. Practical guidance for the implementation and a Monte Carlo study are given in Section 4. In Section 5 the methods are used to analyze price and volatility jumps in NASDAQ high-frequency intra-day trading data, reconstructed from the order book. Section 6 concludes. All proofs are gathered in Section 7.

2. Model, testing problem and statistical approach

Let $(\Omega^X, \mathcal{F}^X, (\mathcal{F}_t^X), \mathbb{P}^X)$ be a filtered probability space satisfying the usual conditions. The latent log-price process X follows an Itô semimartingale

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{R}} \delta(s, x) \mathbb{1}_{\{|\delta(s, x)| \leq 1\}} (\mu - \nu)(ds, dx) \quad (1) \\ + \int_0^t \int_{\mathbb{R}} \delta(s, x) \mathbb{1}_{\{|\delta(s, x)| > 1\}} \mu(ds, dx),$$

with W an (\mathcal{F}_t^X) -adapted standard Brownian motion, μ a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ with $\mathbb{R}_+ = [0, \infty)$ and an intensity measure (predictable compensator of μ) $\nu(ds, dx) = \lambda(dx) \otimes ds$ for a given σ -finite measure λ . We consider discrete observation times $i/n, i = 0, \dots, n$, on the time span $[0, 1]$. The prevalent model, capturing market microstructure effects which interfere the evolution of an underlying semimartingale log-price process at high frequencies, is an indirect observation model with noise:

$$Y_i = X_{i/n} + \epsilon_i, i = 0, \dots, n, \quad (2)$$

where $(\epsilon_i)_{0 \leq i \leq n}$ is a discretization of the continuous-time noise process $(U_t)_{t \in [0, 1]}$. We consider X and U on a common probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ with $\mathcal{F} = \sigma(U_s, s \leq 1) \vee \mathcal{F}^X$ and $\mathcal{F}_t = \sigma(U_s, s \leq t) \vee \mathcal{F}_t^X$. Here, for two σ -algebras \mathcal{F} and \mathcal{H} , we denote $\mathcal{F} \vee \mathcal{H}$ the smallest σ -algebra which contains $\mathcal{F} \cup \mathcal{H}$. X has the same form (1) on this space, see Section 16.1 of [28] for a formal construction of embedding X and U in a joint probability space. Regularity conditions on the characteristics of the efficient price X and the noise, under which we establish asymptotic results, are given in Section 3.1. In particular, we work with a general smoothness assumption on the volatility $(\sigma_t)_{t \in [0, 1]}$. Similar to [29], resulting convergence rates of the spot volatility estimator and the asymptotic test hinge on this smoothness. First, readers may think of the typical case that $(\sigma_t)_{t \in [0, 1]}$ is an Itô semimartingale with a representation as X in (1) and with locally bounded characteristics.

2.1. Test for common price and volatility jumps

In the presence of price jumps, we design a statistical test to decide if contemporaneous price and volatility jumps occur on the considered time interval $[0, 1]$. Let $(S_p)_{p \geq 1}$ be a sequence of stopping times exhausting the jumps of X . We denote the process of left limits of the volatility $\sigma_{t-} = \lim_{u \rightarrow t, u < t} \sigma_u$. We address the null hypothesis of no common jump of volatility and price on $[0, 1]$:

$$\mathbb{H}_{[0, 1]} : \sum_{S_p \leq 1} |\sigma_{S_p}^2 - \sigma_{S_p-}^2| = 0, \quad (3)$$

against the alternative hypothesis that there is at least one jump in the volatility at a jump time of X .

Our test for (3) relies on two main ingredients. First, localization of price jumps using thresholding. Second, a *local* test for volatility jumps. Suppose we want to test $H_0^* : |\sigma_s^2 - \sigma_{s-}^2| = 0$ at a specific time $s \in (0, 1)$, against the alternative hypothesis that the volatility exhibits a jump $|\sigma_s^2 - \sigma_{s-}^2| > 0$. For such a test we require estimates of the squared volatility at time s , $\hat{\sigma}_s^2$, and before time s , $\hat{\sigma}_{s-}^2$. An intuitive test statistic is the difference $\hat{\sigma}_s^2 - \hat{\sigma}_{s-}^2$. It turns out that a more general class of statistics $T^*(s) = g(\hat{\sigma}_s^2, \hat{\sigma}_{s-}^2)$ with a test function g facilitates improved asymptotic properties.

If discrete observations of the efficient log-price $X_{i/n}$, $i = 0, \dots, n$, were directly available, and if we assume for this motivation that there are no jumps in X , σ_s^2 and σ_{s-}^2 could be estimated by local versions of realized volatility:

$$\hat{\sigma}_s^2 = \frac{n}{k_n} \sum_{j=\lfloor sn \rfloor + 1}^{\lfloor sn \rfloor + k_n} (X_{\frac{j+1}{n}} - X_{\frac{j}{n}})^2, \quad \hat{\sigma}_{s-}^2 = \frac{n}{k_n} \sum_{j=\lfloor sn \rfloor - k_n}^{\lfloor sn \rfloor - 1} (X_{\frac{j}{n}} - X_{\frac{j-1}{n}})^2. \quad (4)$$

For an Itô semimartingale $(\sigma_t)_{t \in [0, 1]}$, $k_n = c\sqrt{n}$ with some constant c , $\hat{\sigma}_s^2$ yields rate-optimal spot volatility estimators, that is, $(\hat{\sigma}_s^2 - \sigma_s^2) = \mathcal{O}_{\mathbb{P}}(n^{-1/4})$. Further, on the null hypothesis that $\sigma_{s-} = \sigma_s$, for $k_n = cn^b$ with $b = 1/2 - \delta$ and $\delta > 0$ arbitrarily small, a stable central limit theorem can be proved

$$n^{b/2}(\hat{\sigma}_s^2 - \hat{\sigma}_{s-}^2) \xrightarrow{(st)} MN(0, 4\sigma_s^4).$$

For stochastic volatility the limit is mixed normal and it is important that the convergence holds stably in law to allow for confidence intervals. This is a stronger mode of weak convergence which is equivalent to joint weak convergence with every \mathcal{F}^X -measurable bounded random variable, see [28] for an overview on stable limit theorems. This limit theorem readily supplies an asymptotic test for a volatility jump at time s with a rate of convergence $n^{b/2}$. However, the convergence rate is rather slow and not optimal for this testing problem. For the test statistic

$$\mathcal{T}(s) = 2 \log \left(\frac{1}{2}(\hat{\sigma}_s^2 + \hat{\sigma}_{s-}^2) \right) - \log(\hat{\sigma}_s^2) - \log(\hat{\sigma}_{s-}^2) \quad (5)$$

one derives instead $n^b \mathcal{T}(s) \xrightarrow{(st)} \chi_1^2$ with a χ_1^2 limit distribution and a much faster rate. This improves the (asymptotic) power significantly. A key property is that the test is pivotal, since $\mathcal{T}(s)$ is *self-scaling* in the volatility. This means that it does not require some estimated asymptotic variance, since the limit does not depend on any unknown parameter. Such a local test is not separately highlighted in [29], but is contained as one ingredient of their general method. The final test statistic of [29] is a sum of these local test statistics over all estimated jump times.

It is not obvious how to construct a generalization of the local test for a volatility jump to the noisy observations setup (2). Spot volatility estimators, which are local versions of integrated volatility estimators under noise, are avail-

able, see for instance [45] and [36]. For an Itô semimartingale $(\sigma_t)_{t \in [0,1]}$ and i.i.d. noise with some moment assumption, stable central limit theorems

$$n^{\beta/2}(\hat{\sigma}_s^2 - \sigma_s^2) \xrightarrow{(st)} MN(0, \mathbf{AVAR}_s)$$

with optimal $\beta = 1/4 - \delta$, $\delta > 0$, can be proved. Based on $\hat{\sigma}_s^2 - \hat{\sigma}_{s-}^2$, a test with rate $n^{\beta/2}$ could be constructed. Asymptotic variances \mathbf{AVAR}_s of such estimators are usually sums of at least three addends: one depending on the noise variance, one including the quarticity σ_s^4 and a cross term depending on both. This applies to the asymptotic variances of the spot volatility estimators in [45] and [36], which, however, have sub-optimal slower convergence rates localizing a sub-optimal two-scales integrated volatility estimator. The construction of a rate-optimal pre-average spot volatility estimator with an asymptotic variance of the type above is sketched in Section 8.7 of [3]. Due to this structure of the asymptotic variance, it appears difficult to find a suitable test function that facilitates an asymptotic distribution free test with improved convergence rate.

Apart from attaining asymptotic efficiency, our main motivation to construct a method based on spectral spot volatility estimation is that we will be able to prove a stable central limit theorem

$$n^{\beta/2}(\hat{\sigma}_s^2 - \sigma_s^2) \xrightarrow{(st)} MN(0, 8\sigma_s^3\eta^{1/2})$$

under mild assumptions for semimartingale volatility. Here, $\eta = \mathbb{E}[\epsilon_i^2]$ is the variance of i.i.d. noise, while we consider more general heteroscedastic and serially correlated noise in Section 3. This enables us to find a suitable test function $g(\hat{\sigma}_s^2, \hat{\sigma}_{s-}^2)$, such that

$$n^\beta T_0(s) \xrightarrow{(st)} \chi_1^2, \quad (6)$$

for a test statistic $T_0(s)$ which is self-scaling in the volatility. The self-scaling property and the much faster convergence rate are key features to derive a reliable testing procedure.

To test the null hypothesis (3), local tests are performed at the estimated price-jump times which can be detected by truncation methods. Our asymptotic analysis provides results for the local test at some time s as a special case.

The tests for common price and volatility jumps of [29] for direct observations and our generalization for noisy observations both restrict to finitely many large price adjustments at whose arrival times local tests are performed. Testing for volatility jumps over an interval instead would require a sequence of tests for volatility jumps at infinitely many points and is rather connected to a high-dimensional testing problem. A theory without noise recently has been presented in [14] and a generalization of the techniques, which are quite different to [29], to the model with noise is a challenging topic for future research. It is clear that detecting volatility jumps from noisy observations of the price is especially difficult if we do not specify where to look for potential volatility jumps and

the finite-sample performance of a global test is limited, see Section 6 of [14]. Restricting to local tests for volatility jumps as in this work facilitates a larger power in finite-sample applications.

2.2. Spectral spot volatility estimators

Consider a sequence of equispaced partitions of the considered time span $[0, 1]$ into bins $[kh_n, (k+1)h_n), k = 0, \dots, h_n^{-1} - 1$. For a simple notation suppose $nh_n \in \mathbb{N}$, such that on each bin we enclose nh_n noisy observations. A main idea of spectral volatility estimation, constructed in [15], is to perform optimal parametric estimation procedures localized on the bins. Based on these local estimates, one can build estimators for the spot and the integrated squared volatility. We utilize L^2 -orthogonal functions $(\Phi_{jk})_{1 \leq j \leq J_n}$ for spectral frequencies $1 \leq j \leq J_n$ in the Fourier domain up to a spectral cut-off $J_n \leq nh_n$. For $1 \leq j \leq J_n, 0 \leq k \leq h_n^{-1} - 1$ and $0 \leq t \leq 1$, we define

$$\begin{aligned} \Phi_{j0}(t) &= \left(\sqrt{2h_n n} \sin \left(\frac{j\pi}{2nh_n} \right) \right)^{-1} \sin(j\pi h_n^{-1} t) \mathbb{1}_{[0, h_n]}(t), \\ \Phi_{jk}(t) &= \Phi_{j0}(t - kh_n). \end{aligned} \quad (7)$$

The indicator functions localize the sine functions to the bins. For the spectral volatility estimation, local linear combinations of the noisy data are used with local weights obtained by evaluating the functions (7) on the discrete grid of observation times $i/n, i = 0, \dots, n$. We use the notion of empirical scalar products and norms for functions f, g as follows:

$$\langle f, g \rangle_n := \frac{1}{n} \sum_{l=1}^n f \left(\frac{l}{n} \right) g \left(\frac{l}{n} \right) \quad \text{and} \quad \|f\|_n^2 := \frac{1}{n} \sum_{l=1}^n f^2 \left(\frac{l}{n} \right) = \langle f, f \rangle_n. \quad (8)$$

The empirical norms of the sine functions above give for all bins $k = 0, \dots, h_n^{-1} - 1$:

$$\|\Phi_{jk}\|_n^2 = (4n^2 \sin^2(j\pi/(2nh_n)))^{-1}, \quad (9)$$

and we have the discrete orthogonality relations

$$\langle \Phi_{jk}, \Phi_{rk} \rangle_n = \|\Phi_{jk}\|_n^2 \delta_{jr}, \quad j, r \in \{1, \dots, J_n\}, k = 0, \dots, h_n^{-1} - 1, \quad (10)$$

where $\delta_{jr} = \mathbb{1}_{\{j=r\}}$ is Kronecker's delta. The latter rely on basic discrete Fourier analysis, a detailed proof is given in [15]. The central building blocks of spectral volatility estimation are the spectral statistics

$$S_{jk} = \|\Phi_{jk}\|_n^{-1} \sum_{i=1}^n \Delta_i^n Y \Phi_{jk} \left(\frac{i}{n} \right), \quad j = 1, \dots, J_n, k = 0, \dots, h_n^{-1} - 1, \quad (11)$$

in which observed returns $\Delta_i^n Y = Y_{i/n} - Y_{(i-1)/n}, i = 1, \dots, n$, are smoothed by bin-wise linear combinations. Since the weight functions $\Phi_{jk}(t)$ are non-zero only

on the k th bin, the spectral statistics (S_{jk}) include returns $(\Delta_i^n Y), i = knh_n + 1, \dots, (k + 1)nh_n$ only over the bin under consideration. In absence of price jumps, bin-wise estimates for the squared volatility $\sigma_{kh_n}^2, k = 0, \dots, h_n^{-1} - 1$, are provided by weighted sums of bias-corrected squared spectral statistics:

$$\zeta_k(Y) = \sum_{j=1}^{J_n} w_{jk} \left(S_{jk}^2 - \|\Phi_{jk}\|_n^{-2} \frac{\hat{\eta}_{kh_n}}{n} \right). \tag{12}$$

For the moment, readers can interpret $(\eta_t)_{t \in [0,1]}$ as time varying variance function of the observation errors in (2) and $\hat{\eta}_{kh_n}$ some consistent estimator. In Section 3.1, this is further generalized. The oracle optimal weights

$$w_{jk} = I_k^{-1} I_{jk} = \frac{\left(\sigma_{kh_n}^2 + \|\Phi_{jk}\|_n^{-2} \frac{\eta_{kh_n}}{n} \right)^{-2}}{\sum_{m=1}^{J_n} \left(\sigma_{kh_n}^2 + \|\Phi_{mk}\|_n^{-2} \frac{\eta_{kh_n}}{n} \right)^{-2}}, \tag{13}$$

with $I_k = \sum_{j=1}^{J_n} I_{jk}, I_{jk} = \frac{1}{2} (\sigma_{kh_n}^2 + \|\Phi_{jk}\|_n^{-2} \eta_{kh_n} / n)^{-2}$, follow from minimization of the variance under the constraint of unbiasedness. For a fully adaptive approach we apply a two-stage method and obtain adaptive local estimates $\zeta_k^{ad}(Y)$ by plugging in estimated optimal weights \hat{w}_{jk} in (12).

Remark 1. *Spectral statistics are related to pre-averages used by [26], but the two estimators can not be transformed into one another, see Remark 5.2 in [27] for a discussion of their connection. One difference is that for the spectral method we start with a histogram structure and not a rolling kernel and then smooth bin-wise noisy observations in the Fourier domain. The statistics (11) de-correlate the data for different frequencies and form their local principal components. This is key to the asymptotic efficiency attained by the spectral estimators as shown in [40] and [12]. The latter shows that the estimator’s asymptotic variance coincides with the minimum asymptotic variance among all asymptotically unbiased estimators. We refer to Remark 3.1 of [27] for a recent discussion about efficient volatility estimation under noise.*

The spectral volatility estimation provides local estimates (12) for the squared volatility $\sigma_{kh_n}^2, k = 0, \dots, h_n^{-1} - 1$. In order to derive an estimate $\hat{\sigma}_s^2$ at some time s , we average the statistics $\zeta_k(Y)$ over a local window around s of length $(r_n^{-1} h_n) \rightarrow 0$ as $n \rightarrow \infty, r_n^{-1} \in \mathbb{N}$, slowly enough to ensure $r_n^{-1} \rightarrow \infty$. In the presence of jumps in (1), truncation disentangles bin-wise statistics (12) which include jumps from all others. We use the methods from [16] to cope with price jumps for volatility estimation. If $h_n |\zeta_k(Y)| > u_n$ for a threshold sequence $u_n = c h_n^\tau, \tau \in (0, 1)$, with some constant c , the statistic is too large to be driven by the continuous part and is evoked by a jump of X . In order to estimate the volatility, we thus truncate $\zeta_k(Y)$ for these k . For estimating the squared volatility and its left limit at a certain time s , we use two disjoint windows after and before s , respectively.

When the optimal weights (13) are known, an *oracle spot volatility estimator*

$\hat{\sigma}_{s,or}^2$ for $s \in [r_n^{-1}h_n, 1 - r_n^{-1}h_n)$ is:

$$\hat{\sigma}_{s,or}^2 = \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \sum_{j=1}^{J_n} w_{jk} (S_{jk}^2 - \|\Phi_{jk}\|_n^{-2} \frac{\eta_{kh_n}}{n}) \mathbb{1}_{\{h_n |\zeta_k(Y)| \leq u_n\}}, \quad (14a)$$

and the estimator for $\hat{\sigma}_{s-,or}^2$:

$$\hat{\sigma}_{s-,or}^2 = \sum_{k=\lfloor sh_n^{-1} \rfloor - r_n^{-1}}^{\lfloor sh_n^{-1} \rfloor - 1} r_n \sum_{j=1}^{J_n} w_{jk} (S_{jk}^2 - \|\Phi_{jk}\|_n^{-2} \frac{\eta_{kh_n}}{n}) \mathbb{1}_{\{h_n |\zeta_k(Y)| \leq u_n\}}. \quad (14b)$$

Close to the boundaries, $s \in [0, r_n^{-1}h_n) \cup (1 - r_n^{-1}h_n, 1]$, we shrink one window length accordingly. Since the optimal weights (13) hinge on the unknown squared volatility and the noise level $(\eta_t)_{t \in [0,1]}$, we proceed with a two-step estimation approach. First, select a pilot spectral cut-off $J_n^{pi} \ll nh_n$, and build *pilot estimators* for the squared volatility

$$\hat{\sigma}_{s,pil}^2 = \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \sum_{j=1}^{J_n^{pi}} (J_n^{pi})^{-1} (S_{jk}^2 - \|\Phi_{jk}\|_n^{-2} \frac{\hat{\eta}_{kh_n}}{n}) \times \mathbb{1}_{\left\{h_n \left| \sum_{j=1}^{J_n^{pi}} (J_n^{pi})^{-1} (S_{jk}^2 - \|\Phi_{jk}\|_n^{-2} \frac{\hat{\eta}_{kh_n}}{n}) \right| \leq u_n \right\}}, \quad (15)$$

and $\hat{\sigma}_{s-,pil}^2$ analogously. The pilot estimators are hence averages of squared, bias-corrected spectral statistics over r_n^{-1} bins and J_n^{pi} spectral frequencies. In the second step, these pilot estimators are plugged into (13) to determine adaptive weights \hat{w}_{jk} for the final estimators. We write

$$\zeta_k^{ad}(Y) = \sum_{j=1}^{J_n} \hat{w}_{jk} \left(S_{jk}^2 - \|\Phi_{jk}\|_n^{-2} \frac{\hat{\eta}_{kh_n}}{n} \right). \quad (16)$$

The *spectral estimators of the squared spot volatility* at time s and its left limit are:

$$\hat{\sigma}_s^2 = \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \sum_{j=1}^{J_n} \hat{w}_{jk} (S_{jk}^2 - \|\Phi_{jk}\|_n^{-2} \frac{\hat{\eta}_{kh_n}}{n}) \mathbb{1}_{\{h_n |\zeta_k^{ad}(Y)| \leq u_n\}}, \quad (17a)$$

$$\hat{\sigma}_{s-}^2 = \sum_{k=\lfloor sh_n^{-1} \rfloor - r_n^{-1}}^{\lfloor sh_n^{-1} \rfloor - 1} r_n \sum_{j=1}^{J_n} \hat{w}_{jk} (S_{jk}^2 - \|\Phi_{jk}\|_n^{-2} \frac{\hat{\eta}_{kh_n}}{n}) \mathbb{1}_{\{h_n |\zeta_k^{ad}(Y)| \leq u_n\}}. \quad (17b)$$

Estimates (17a) and (17b) are truncated local averages of the statistics (16). Our approach entails several tuning parameters whose practical choice is discussed in Section 4.2.

3. Asymptotic theory

3.1. Assumptions with discussion

We start with the assumptions on the characteristics of X in (1) which are similar to the ones in [29].

Assumption 1. *For the adapted and locally bounded drift process $(b_s)_{s \geq 0}$, we require a minimal smoothness condition that for $0 \leq t < s \leq 1$, some constant C and some $\iota > 0$:*

$$\mathbb{E}[(b_s - b_t)^2 | \mathcal{F}_t] \leq C (s - t)^\iota. \tag{18}$$

The volatility process σ_t is càdlàg and neither σ_t nor $\sigma_{t-} = \lim_{u \rightarrow t, u < t} \sigma_s$ vanishes.

Assumption (H-r). *We assume that $\sup_{\omega, x} |\delta(t, x)|/\gamma(x)$ is locally bounded for a non-negative deterministic function γ satisfying $\int_{\mathbb{R}} (\gamma^r(x) \wedge 1) \lambda(dx) < \infty$.*

We index the assumption in $r \in [0, 2]$ to highlight the role of the jump activity index r . The larger r , the more general jump components are included in our model. In particular for $r = 0$ we consider jumps of finite activity. Imposing $r < 1$ instead allows for infinite activity jumps which are absolutely summable. We state the assumptions on characteristics of X with respect to $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, with the usual extension from $(\Omega^X, \mathcal{F}^X, (\mathcal{F}_t^X), \mathbb{P}^X)$. Especially, (W_t) in (1) is also a standard Brownian motion on this space. For the volatility process, our target of inference, we work with the following general smoothness condition determined by a smoothness parameter $\alpha \in (0, 1]$.

Assumption $(\sigma\text{-}\alpha)$. *The process σ_t satisfies $\sigma_t = f_\sigma(\sigma_t^{(A)}, \sigma_t^{(B)})$ with some function $f_\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}$, continuously differentiable in both coordinates, and two (\mathcal{F}_t) -adapted processes $\sigma_t^{(A)}, \sigma_t^{(B)}$, where*

- $\sigma^{(A)}$ is an Itô semimartingale

$$\begin{aligned} \sigma_t^{(A)} &= \sigma_0^{(A)} + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \tilde{\sigma}_s^* dW'_s \\ &\quad + \int_0^t \int_{\mathbb{R}} \tilde{\delta}(s, x) \mathbb{1}_{\{|\tilde{\delta}(s, x)| \leq 1\}} (\tilde{\mu} - \tilde{\nu})(ds, dx) \\ &\quad + \int_0^t \int_{\mathbb{R}} \tilde{\delta}(s, x) \mathbb{1}_{\{|\tilde{\delta}(s, x)| > 1\}} \tilde{\mu}(ds, dx), \end{aligned} \tag{19}$$

with an (\mathcal{F}_t) -Brownian motion W' independent of W , locally bounded characteristics $\tilde{\sigma}, \tilde{\sigma}^*, \tilde{b}, \tilde{\mu}$ and a random variable $\sigma_0^{(A)}$. $\sigma_t^{(A)}$ satisfies Assumptions 1 and (H-2) for $\alpha \leq 1/2$. For $\alpha > 1/2$, the continuous martingale part of $\sigma^{(A)}$ vanishes and $\sigma^{(A)}$ satisfies Assumptions 1 and (H- α^{-1}).

- $\sigma^{(B)}$ lies in a Hölder ball of order α almost surely, i.e. $|\sigma_t^{(B)} - \sigma_s^{(B)}| \leq L|t - s|^\alpha$, for all $t, s \in [0, 1]$ and a random variable L for which at least fourth moments exist.

The smaller α , the less restrictive is Assumption $(\sigma-\alpha)$. It is natural to develop results for general $\alpha \in (0, 1]$ to cover a broad framework and preserve some freedom in the model. This is particularly important, since the precision of nonparametrically estimating a process (or function) foremost hinges on its smoothness α . Therefore, convergence rates in Section 3.2 hinge on α . In the composition of the volatility in Assumption $(\sigma-\alpha)$, $\sigma_t^{(B)}$ can contain a non-Lipschitz seasonality component (Lipschitz continuous seasonalities can as well be modeled by the drift of $\sigma_t^{(A)}$). As pointed out by [29], $\sigma_t^{(B)}$ can also be a long-memory volatility component as the prominent exponential fractional Ornstein-Uhlenbeck model by [19].

While an i.i.d. assumption on the noise is standard in most works, empirical findings, for instance by [22], motivate to allow for serial correlation and endogeneity in the noise. We develop our theory under the following general assumption.

Assumption $(\eta-p)$. *The noise $(\epsilon_i)_{0 \leq i \leq n}$ process is centered, $\mathbb{E}[\epsilon_i] = 0$. For some $p \geq 4$, its \mathcal{F}^X -conditional law has finite p -th moments, $\mathbb{E}[\epsilon_i^p | \mathcal{F}^X] < \infty$ almost surely for all $i = 0, \dots, n$. The long-run variance process converges*

$$\sum_{l=-\lfloor tn \rfloor}^{n-\lfloor tn \rfloor} \text{Cov}(\epsilon_{\lfloor tn \rfloor}, \epsilon_{\lfloor tn \rfloor + l}) \rightarrow \eta_t, \quad (20)$$

for $t \in [0, 1]$ uniformly on compacts in probability and we have the mixing behavior

$$\sup_{i=0, \dots, n} |\text{Cov}(\epsilon_i, \epsilon_{i+l})| = \mathcal{O}(|l|^{-1-\varrho}), \quad (21)$$

for some $\varrho > 0$, which is specified in the discussion below Theorem 1. The process $(\eta_t)_{t \in [0, 1]}$ is locally bounded and satisfies for all $t, (t+s) \in [0, 1]$ the mild smoothness condition:

$$|\eta_{t+s} - \eta_t| \leq K s^{(1/2+\delta) \vee \alpha}, \quad (22)$$

with some $\delta > 0$. The noise ϵ_i is for all i uncorrelated to $(\Delta_l^n X)_{l=1, \dots, (i-\tilde{Q}-1 \vee 1)}$ for some $\tilde{Q} < \infty$ and

$$\sum_{l=\lfloor tn \rfloor - \tilde{Q}}^{\lfloor tn \rfloor} \mathbb{E}[\epsilon_{\lfloor tn \rfloor} \Delta_l^n X] \rightarrow \rho_t, \quad (23)$$

for some continuous bounded function $(\rho_t)_{t \in [0, 1]}$. Furthermore, the noise does not vanish, $\eta_t > 0$ for all $t \in [0, 1]$.

The case that $\text{Cov}(\epsilon_i, \epsilon_{i+l}) = 0$ for all $l \neq 0$ and $\eta = \text{Var}(\epsilon_i)$ constant for all i is tantamount to the classical setup with i.i.d. noise. In general the noise is serially correlated, endogenous and heteroscedastic. Different to Assumption (GN) in Section 7.2 of [3], we do not assume that the noise is conditionally centered

to include the correlation to the increments of X in (23). The endogeneity condition (23) includes linear models of the form $\epsilon_i = \sum_{l=i-Q}^i c_l \sqrt{n} \Delta_l^n X + U_i$, with U_i exogenous errors and constants c_l , similar as in Equation (6) of [31] or considered by [9]. If we knew the process $(\eta_t)_{t \in [0,1]}$, Assumption (η -p) with a mild lower bound for ϱ would be sufficient for our asymptotic results. For an adaptive method, however, we need to estimate the process $(\eta_t)_{t \in [0,1]}$. Consistent estimation of the noise long-run variance (20) requires stronger structural assumptions. For a Q -dependent noise process, that is, $\sup_{i=0, \dots, n} |\text{Cov}(\epsilon_i, \epsilon_{i+q})| = 0$ for $q > Q$ and some given $Q < \infty$, and if η in (20) is time-invariant, consistent estimation with \sqrt{n} -convergence rate of η has been established by [23]. In [13] it is shown how Q can be found adaptively if it is unknown. Consistent estimation of the noise variance process under heteroscedasticity, but without serial correlations, is discussed in [27]. For the fully adaptive method, we tighten the assumptions on the noise as follows.

Assumption 2. *Assumption (η -p) holds with $p \geq 8$. Moreover,*

$$\sup_{i=0, \dots, n} |\text{Cov}(\epsilon_i, \epsilon_{i+q})| = 0$$

for all $q > Q$ with some $Q < \infty$.

Assumption 2 is satisfied by a Q -dependent noise process. Then, a consistent estimation of the long-run noise variance (20) process is possible.

Proposition 3.1. *Under Assumption 2, for $h_n = \kappa_1 n^{-1/2} \log(n)$, for all $k = 0, \dots, h_n^{-1} - 1$, the locally constant approximated noise long-run variance process can be estimated with accuracy*

$$\hat{\eta}_{kh_n} = \eta_{kh_n} + \mathcal{O}_{\mathbb{P}}(n^{-\beta}). \quad (24)$$

Our estimator is given in (43) in the appendix. It is somewhat related to the methods from [23] and [13], but localized to bins.

The assumptions on the noise are more general than in other works on spectral volatility estimation as in [5] and in [13]. In particular, to the best of our knowledge, we consider for the first time heteroscedastic and serially correlated, endogenous noise.

Remark 2 (Non-equidistant observations). *For a coherent and simple exposition of the construction of the spectral estimator in (7)-(11), we discuss equidistant observations which allows us to rely on discrete-time Fourier identities in (10). Considering a heteroscedastic noise-level, our analysis and results are at the same time informative about non-equidistant observations. For general observation schemes $t_i^n, i = 0, \dots, n$, we impose the condition that a differentiable cdf F exists such that observation times $t_i^n = F^{-1}(i/n)$ are obtained by a quantile transformation from the equidistant setting. Moreover, we require that the derivative F' is strictly positive and satisfies the same smoothness as (η_t) in (22). These assumptions are the same as in Assumption (Obs-d) of [5]. Then, all our asymptotic results transfer from the equidistant to this general*

setting when we replace η_s by $\eta_s(F^{-1})'(s)$. This follows directly by the asymptotic equivalence of the respective experiments established in [12]. In particular, having locally less frequent observations is equivalent to having locally an increased noise level. Therefore, under the imposed conditions, $(\eta_t)_{t \in [0,1]}$ and $(F^{-1})'(t), t \in [0,1]$, may be pooled. Note that adding the factor $(F^{-1})'(s)$ to the noise level η_s is the same as generalizing the frequently occurring factor $\|\Phi_{jk}\|_n^{-2} \eta_{kh_n}/n$ to $\|\Phi_{jk}\|_n^{-2} \eta_{kh_n}/(nF'(kh_n))$, where $nh_n F'(kh_n)$ gives the local sample size. In the equidistant case this is nh_n and we have that $F'(s) = 1$ is constant.

3.2. Asymptotic results

Our first main result is on the spot squared volatility estimator and its asymptotic distribution.

Theorem 1. *Suppose Assumptions 1, 2 and (H-r) with some $r < 2$ and smoothness Assumption $(\sigma-\alpha)$, $\alpha \in (0, 1]$. Fix some time $s \in (0, 1)$, at which we want to estimate σ_s^2 and σ_{s-}^2 with (17a) and (17b), respectively. Set $h_n = \kappa_1 n^{-1/2} \log(n)$ and $r_n = \kappa_2 n^{-\beta} \log(n)$ with constants κ_1, κ_2 and $J_n \rightarrow \infty$, $J_n = \mathcal{O}(\log(n))$, as $n \rightarrow \infty$. Then, as $n \rightarrow \infty$ and if*

$$0 < \beta < \left(\frac{\alpha}{2\alpha + 1} \wedge \tau \left(1 - \frac{r}{2} \right) \right), \quad (25)$$

and $\tau < 1 - \beta/(p - 2)$ when $p < \infty$ moments of the noise exist, with τ the truncation exponent in the sequence u_n in (15), (17a) and (17b), the estimators satisfy the \mathcal{F} -stable central limit theorem:

$$n^{\beta/2} \begin{pmatrix} \hat{\sigma}_s^2 - \sigma_s^2 \\ \hat{\sigma}_{s-}^2 - \sigma_{s-}^2 \end{pmatrix} \xrightarrow{(st)} MN \left(0, \begin{pmatrix} 8\sigma_s^3 \eta_s^{1/2} & 0 \\ 0 & 8\sigma_{s-}^3 \eta_s^{1/2} \end{pmatrix} \right). \quad (26)$$

For the oracle estimators (14a) and (14b), the same limit theorem applies under the less restrictive Assumption $(\eta-p)$ with $p = 8$, $\varrho > \beta$, and if $\tau < 1 - \beta/(p - 2)$. In fact, we can get arbitrarily close to the optimal rate for estimation which is known to be $n^{\alpha/(4\alpha+2)}$ in this case, see [37]. Balancing the squared bias and the variance guarantees that the estimators (17a) and (17b) attain the optimal rate. For a central limit theorem we avoid an asymptotic bias by slightly undersmoothing. Most interesting is the case when $\alpha \approx 1/2$, e.g. when the volatility is a semimartingale. Then the convergence rate is $n^{1/8}$. In case that $\alpha > 1/2$, we obtain faster convergence rates. In case that $\alpha = 1/2$ and if all moments of the noise process exist, for any $r < 3/2$ in Assumption (H-r), we can choose $\beta = 1/4 - \varepsilon$ with any $\varepsilon > 0$. Under the standard assumption that we only have Assumption $(\eta-p)$ with $p = 8$, the condition $\tau < 23/24$ results in $r < 34/23 \approx 1.478$. Hence, restricting to the condition that up to 8th moments of the noise exist leads only to a slightly less general condition on the jump activity. We point out that the restriction $r < 3/2$ on the jump activity,

to come close to the optimal convergence rate, is less restrictive than the one obtained for integrated squared volatility estimation, $r < 1$, in [16]. The reason is that for spot volatility estimation we can only obtain slower convergence rates by local smoothing compared to integrated volatility estimation. This, however, works also under more active jumps.

The limit variable in (26) is *mixed normal* which we denote by MN and defined on a product space of the original probability space (on which Y is defined) and an orthogonal space independent of \mathcal{F} . The convergence is \mathcal{F} -stable in law, marked (*st*). Stability of weak convergence then allows for a so-called *feasible* version of the limit theorem (26) that facilitates confidence sets.

Corollary 3.2. *Under the conditions of Theorem 1, and also for any J_n fixed as $n \rightarrow \infty$:*

$$r_n^{-1/2} \begin{pmatrix} \hat{I}_{\lfloor sh_n^{-1} \rfloor + 1}^{1/2} (\hat{\sigma}_s^2 - \sigma_s^2) \\ \hat{I}_{\lfloor sh_n^{-1} \rfloor - 1}^{1/2} (\hat{\sigma}_{s-}^2 - \sigma_{s-}^2) \end{pmatrix} \xrightarrow{(st)} N \left(0, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right), \quad (27)$$

with $\hat{I}_{\lfloor sh_n^{-1} \rfloor + 1}^{1/2}$ and $\hat{I}_{\lfloor sh_n^{-1} \rfloor - 1}^{1/2}$ the estimates of $I_{\lfloor sh_n^{-1} \rfloor + 1}$ and $I_{\lfloor sh_n^{-1} \rfloor - 1}$, as defined in the weights (13), obtained by inserting the pilot estimates.

The results proved for the spot volatility estimator provide a main building block for our asymptotic test, but are moreover of interest in their own right. They show that the spectral method renders effective spot squared volatility estimators under general noise and in the presence of jumps.

In the sequel, let $(S_p)_{p \geq 1}$ be a sequence of stopping times exhausting the jumps of X . We address the null hypothesis (3) that no common jumps of volatility and price occur on $[0, 1]$. Under the alternative hypothesis, there is at least one contemporaneous jump in volatility and price.

Analogously to [29], we specify test hypotheses more precisely by focusing on jumps of X with absolute values $|\Delta X_{S_p}| > a$ for $a \geq 0$ and write $\mathbb{H}(a)_{[0,1]}$. The reason for this is that a suitable test statistic and associated limit theory for $\mathbb{H}(a)_{[0,1]}$ with $a > 0$ works under a much more general setup with jumps of infinite variation while testing $\mathbb{H}(0)_{[0,1]}$ requires Assumption (H-0) to hold. In both cases, we concentrate on a finite number of (large) price jumps under the null hypothesis. From an applied point of view this is reasonable, since we are interested in volatility movements at finitely many relevant price adjustments on a fixed time interval.

Denote by $g : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ a test function with $g(x, x) = 0$ for all x . Let us now state the general form of our test statistics:

$$T_0(h_n, r_n, g) = \sum_{k=r_n^{-1}}^{h_n^{-1} - r_n^{-1} - 1} \hat{\eta}_{kh_n}^{-1/2} g(\hat{\sigma}_{kh_n}^2, \hat{\sigma}_{kh_n-}^2) \mathbb{1}_{\{h_n |\zeta_k^{ad}(Y)| > (u_n \vee a^2)\}}. \quad (28)$$

Under mild regularity assumptions on g in terms of differentiability in both coordinates, limit theorems for (28) can be proved. For testing $\mathbb{H}(a)_{[0,1]}$, we

consider two specific test functions in the following. Adjustments of the test (3) for sub-intervals of $[0, 1]$ are readily obtained by ignoring all jumps elsewhere.

Theorem 2. *Let S_1, \dots, S_{N_1} be a finite collection of jump times of X on $[0, 1]$, with $|\Delta X_{S_i}| > a$ for all i . Consider $\mathbb{H}(a)_{[0,1]}$, if either $a > 0$ and we impose the condition that the Lévy measure of X does not have an atom in $\{a\}$, or assume $r = 0$. On all assumptions of Theorem 1 and if $\tau < 3/4$ for $a = 0$, when inserting estimates (17a) and (17b) with $h_n = \kappa_1 n^{-1/2} \log(n)$, $r_n = \kappa_2 n^{-\beta} \log(n)$, $J_n \rightarrow \infty$, $J_n = \mathcal{O}(\log(n))$ in (28) with the test function*

$$g(x_1, x_2) = 2\sqrt{\frac{x_1 + x_2}{2}} - \sqrt{x_1} - \sqrt{x_2}, \quad (29)$$

the following asymptotic distribution of the test statistic applies under $\mathbb{H}(a)_{[0,1]}$:

$$n^\beta T_0(h_n, r_n, g) \xrightarrow{(st)} \chi_{N_1}^2. \quad (30)$$

Under the alternative hypothesis, $n^\beta T_0(h_n, r_n, g) \rightarrow \infty$ in probability. Therefore, we obtain an asymptotic distribution free test by the asymptotic χ^2 -distribution with N_1 degrees of freedom. The test with critical regions

$$C_n = \{n^\beta T_0(h_n, r_n, g) > q_{1-\alpha}(\chi_{N_1}^2)\}, \quad (31)$$

where $q_\alpha(\chi_{N_1}^2)$ denotes the α -quantile of the $\chi_{N_1}^2$ -distribution, has asymptotic level α and asymptotic power 1.

In fact, (31) contains the estimated number of price jumps \hat{N}_1 . Since $\mathbb{P}(\hat{N}_1 - N_1 > 0) \rightarrow 0$, (30) applies with N_1 also. A naive approach based on the asymptotic normality result (27) with test function $\tilde{g}(x_1, x_2) = (x_1 - x_2)$ yields as well an asymptotic test. It holds that

$$r_n^{-1/2} \left(2 \sum_{i=1}^{\hat{N}_1} \hat{I}_{[\hat{h}_n^{-1} S_i] + 1}^{-1} \right)^{-1/2} T_0(h_n, r_n, \tilde{g}) \xrightarrow{d} N(0, 1), \quad (32)$$

on the null hypothesis $\mathbb{H}(a)_{[0,1]}$. Apparently, the rate $r_n^{-1/2} \asymp n^{\beta/2}$,¹ close to $n^{1/8}$ for $\alpha \leq 1/2$, is slower and thus the test in Theorem 2 is preferable.

Remark 3. *As mentioned by [29], their test based on (5) corresponds to a two-sample likelihood ratio test for equal variances in a Gaussian parametric model with observations $\sqrt{n}\Delta_j^n X \stackrel{iid}{\sim} N(0, \sigma_{s-}^2)$, $j = \lfloor sn \rfloor - k_n, \lfloor sn \rfloor - 1$ and $\sqrt{n}\Delta_j^n X \stackrel{iid}{\sim} N(0, \sigma_s^2)$, $j = \lfloor sn \rfloor + 1, \lfloor sn \rfloor + k_n$. In this simpler model – closely related to our model in case of no noise – the likelihood ratio is*

$$\Lambda = \frac{(\hat{\sigma}_s^2 \hat{\sigma}_{s-}^2)^{k_n/2}}{\left(\frac{\hat{\sigma}_s^2 + \hat{\sigma}_{s-}^2}{2}\right)^{k_n}}, \text{ and } -2 \log(\Lambda) = k_n \left(2 \log \frac{\hat{\sigma}_s^2 + \hat{\sigma}_{s-}^2}{2} - \log \hat{\sigma}_s^2 - \log \hat{\sigma}_{s-}^2 \right),$$

¹We write $a_n \asymp b_n$ for asymptotically equivalent real sequences which means $a_n/b_n \rightarrow c$ for some constant c .

where the estimators (4) are the maximum likelihood estimators for this model, and we derive the convergence of $-2\log(\Lambda)$ to a χ_1^2 -distribution from the standard asymptotic theory for likelihood ratio tests.

The model with noise is more complicated. Our test from Theorem 2 does not directly correspond to a parametric likelihood ratio test and our estimators (17a) and (17b) do also not agree with the non-explicit maximum likelihood estimators in this model. The choice of g in (29) is motivated by studying which properties in (5) are important for the asymptotic pivotal distribution under the null. Any function of the form $g(x, y) = 2f(\frac{x+y}{2}) - f(x) - f(y)$, with some twice differentiable function f , is suitable for the construction of tests (in both models) with the fast convergence rate based on second order asymptotics of the estimators, since $\frac{d}{dx}g(\sigma_s^2, \sigma_s^2) = \frac{d}{dy}g(\sigma_s^2, \sigma_s^2) = 0$. On the other hand, that the statistic (5) is self-scaling in the volatility leading to the pivotal limit distribution is due to the identity

$$\frac{d^2}{dx^2}g(\sigma_s^2, \sigma_s^2) = \frac{d^2}{dy^2}g(\sigma_s^2, \sigma_s^2) = -\frac{1}{2}f''(\sigma_s^2) = (n^b \text{Var}(\hat{\sigma}_s^2))^{-1}$$

denoting f'' the second derivative of f . With $f(x) = \log(x)$, it holds that $-\frac{1}{2}f''(\sigma_s^2) = (2\sigma_s^4)^{-1}$, which guarantees the above identity in the model without noise. In light of the efficient asymptotic variance under noise in Theorem 1, it is natural to choose $f(x) = \sqrt{x}$, such that

$$-\frac{1}{2}f''(\sigma_s^2) = \frac{1}{8\sigma_s^3} = \frac{\eta_s^{1/2}}{n^b \text{Var}(\hat{\sigma}_s^2)}.$$

Since the noise level η_s can be estimated with a much faster rate of convergence than σ_s^2 – even under our general assumptions for the noise – this choice of (29) facilitates (30).

The particular choice of the spectral estimators (17a) and (17b) is not crucial for the construction of the test. Any rate-optimal spot volatility estimator may be used when it is possible to find a function f satisfying the above identities. However, with a more complex asymptotic variance structure, for instance for pre-average or realized kernel estimators, this appears to be difficult. Estimators attaining the same efficient variance as in (26) may be used with the same function g in (29), to derive a test with the same asymptotic properties. A localized QMLE as discussed by [18], for instance, could allow for analogous results.

4. Implementation and numerical study

4.1. Setup of Monte Carlo simulation study

The simulation study examines the finite-sample performance of the proposed methods. We implement a model where observed log-prices are given by

$$Y_{i/n} = \int_0^{\frac{i}{n}} \varphi_t \sigma_t dW_t + \int_0^{\frac{i}{n}} \int_{\mathbb{R}} x \mu(dt, dx, dy) + \epsilon_i,$$

with jump intensity measure $\nu(dt, dx, dy) = \lambda dt \Pi(dx)\Pi(dy)$ and with Gaussian jump sizes $\Pi \sim N(H, H/100)$ whose magnitude depend on a parameter H . The efficient semimartingale log-price process is recorded with additive microstructure noise

$$\epsilon_i = \theta \epsilon_{i-1} + u_i, \quad u_i \stackrel{iid}{\sim} N\left(0, \eta(1 - \theta)^{-2}\right), \quad i = 1, \dots, n, |\theta| < 1. \quad (33)$$

In line with empirical evidence, this model generates serially correlated noise. We further consider two different noise models (34) and (35) below. We set values of η according to realistic noise-to-signal ratios. We use the median value of the estimated measure $n\eta\left(\int_0^1 \varphi_t^4 \sigma_t^4\right)^{-1/2}$ found in a comprehensive data study in [13]. Sample sizes $n = 30,000$ and $n = 5,000$ in our simulations suggest $\eta^{1/2} \approx 0.005$ and $\eta^{1/2} \approx 0.015$, which we use in the following as two realistic noise levels. According to the data summary in Table 5, 30,000 is a sample size that matches (approximately) the average daily observation numbers of our empirical data. We additionally analyze the methods' performance for smaller samples sizes $n = 5,000$, which is realistic for less frequently traded assets. We set $\theta = 0.6$ equal to the empirically motivated value in [13].

$\varphi_t = 1 - \frac{3}{5}\sqrt{t} + \frac{1}{10}t^2$ mimics a deterministic volatility intra-day seasonality pattern and σ_t^2 a random stochastic volatility component with leverage:

$$d\sigma_t^2 = 6(1 - \sigma_t^2) dt + \sigma_t^2 dB_t + dJ_t.$$

B is a standard Brownian motion with $d[B, W]_t = \rho dt$, where we fix $\rho = 0.2$.

The jump measure above has a second real argument to incorporate instantaneous arrivals of volatility jumps. The volatility jump component is of the form

$$J_t = \gamma \int_0^t \int_{\mathbb{R}} y \mu(dt, dx, dy) + \int_0^t \int_{\mathbb{R}} z \tilde{\mu}(dt, dz)$$

with $\gamma \in \mathbb{R}$ and intensity measure $\tilde{\nu}(dt, dz) = dt \Pi(dz)$. Setting $\gamma = 0$ results in no common price and volatility jumps which means the null hypothesis is valid. To simulate the model under the alternative hypothesis, we set $\gamma = 1$ instead.

4.2. Choice of tuning parameters

In the sequel, we provide advice on how to specify the tuning parameters that are involved in the nonparametric procedures. We also conduct a sensitivity analysis for the Monte Carlo study to find suitable values.

First, the bin-width $h_n \asymp n^{-1/2} \log n$ balances the number of observations on bins nh_n , which should be large enough to smooth out noise, and the discretization error by approximating volatility bin-wise constant. The sensitivity analysis will show that the final test is very robust to modifications of h_n . We advise to select h_n such that the number of observations on bins is at least 50 within a range to 250 observations for typical high-frequency financial data. This results in a time resolution of 50-150 bins per trading day.

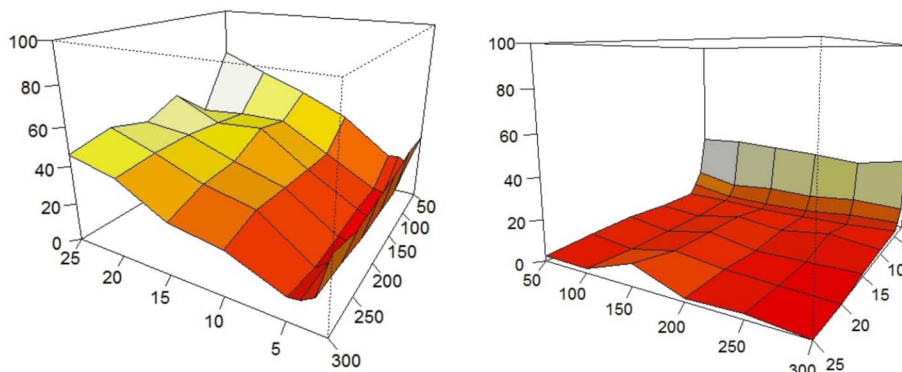


FIG 1. Empirical percentage type-II-error rate (right) and empirical percentage global testing error rate (left) for the test of size $\alpha = 0.05$, depending on tuning parameters h_n^{-1} and r_n^{-1} , with $50 \leq h_{30,000}^{-1} \leq 300$ and $2 \leq r_{30,000}^{-1} \leq 25$. The empirical type-II-error rate measures the empirical amount of realizations under the alternative hypothesis which are smaller or equal the .95-quantile of the $\chi_{N_1}^2$ -distribution. The global testing error rate is the sum of the type-II-error rate and the misspecification of the size, that is, the difference between $(1-\alpha)$ and the empirical amount of realizations smaller or equal the $(1-\alpha)$ -quantile of the $\chi_{N_1}^2$ -distribution, this time on the null hypothesis.

For the spot volatility estimators (17a) and (17b) and the pilot estimator (15), we fix spectral cut-offs J_n and J_n^{pi} , respectively. The values of the spectral cut-offs do not influence the methods when set sufficiently large. Since the weights (13) decay exponentially for $j \gtrsim \sqrt{n}h_n \asymp \log n$, the addends with j large become negligible, such that it suffices to choose $J_n \asymp \log n$. The proportionality constant should be larger than 1, we take values between 3 and 12. The pilot estimators (15) instead use averages over frequencies $j = 1, \dots, J_n^{pi}$, such that we fix J_n^{pi} to be smaller. We thus use $J_n^{pi} \asymp \log n$ with a proportionality factor smaller than for J_n . The threshold sequence u_n determines the bins on which large returns are ascribed to jumps. We use the practical selection presented in [16].

The most influential tuning parameter for our test is the size of the smoothing window $r_n \asymp n^{-\beta} \log n$. If we choose r_n larger, the spot volatility estimates have smaller variance but the bias for rapidly varying volatilities increases. For $\alpha = 1/2$, we know the exact order of r_n depending on n . There is, however, no simple rule of thumb to fix the constant κ_2 , and we conduct an extensive sensitivity analysis to find the best suitable values. The sensitivity analysis reveals that in order to detect volatility jumps and separate them from a rough continuous semimartingale volatility component, we should use rather small smoothing window sizes.

We investigate the performance of the test for common price and volatility jumps depending on the tuning parameters h_n and r_n in the Monte Carlo simulation. We implement the setup from paragraph 4.1 with $\lambda = 2$, $\eta^{1/2} = 0.005$ and $H = 0.25$ for both sample sizes $n = 30,000$ and $n = 5,000$. We set $J_n = 30$ in

TABLE 1
Empirical power of the $\alpha = 0.05$ -test for $n = 5,000$ depending on tuning parameters $h_{5,000}$ and $r_{5,000}$.

h_n^{-1} : r_n^{-1}	2	3	4	5	6	7	8	9	10
20	0.498	0.737	0.784	0.852	0.890	0.842	0.869	0.843	0.831
30	0.557	0.801	0.852	0.896	0.901	0.898	0.927	0.925	0.937
40	0.571	0.831	0.879	0.927	0.934	0.944	0.927	0.942	0.943
50	0.601	0.834	0.906	0.922	0.954	0.949	0.948	0.950	0.957
60	0.603	0.836	0.914	0.933	0.943	0.945	0.968	0.968	0.972
70	0.595	0.879	0.921	0.931	0.950	0.965	0.967	0.966	0.970
80	0.611	0.848	0.931	0.949	0.965	0.971	0.970	0.972	0.983
90	0.629	0.840	0.926	0.957	0.956	0.977	0.977	0.982	0.984
100	0.626	0.842	0.930	0.956	0.978	0.974	0.983	0.973	0.991

TABLE 2
Empirical size of the $\alpha = 0.05$ -test, that is, the empirical amount of realizations smaller or equal the 0.95-quantile of the asymptotic $\chi_{N_1}^2$ -distribution, for $n = 5,000$ depending on tuning parameters $h_{5,000}$ and $r_{5,000}$.

h_n^{-1} : r_n^{-1}	2	3	4	5	6	7	8	9	10
20	0.953	0.851	0.747	0.732	0.630	0.603	0.541	0.421	0.459
30	0.975	0.893	0.794	0.753	0.680	0.614	0.592	0.541	0.491
40	0.975	0.914	0.856	0.781	0.697	0.684	0.608	0.616	0.528
50	0.973	0.915	0.845	0.804	0.742	0.669	0.675	0.606	0.535
60	0.977	0.908	0.855	0.795	0.774	0.737	0.662	0.635	0.614
70	0.976	0.909	0.868	0.792	0.762	0.711	0.673	0.625	0.612
80	0.979	0.911	0.868	0.806	0.787	0.734	0.635	0.666	0.612
90	0.962	0.924	0.872	0.817	0.771	0.713	0.688	0.667	0.603
100	0.959	0.906	0.879	0.795	0.778	0.728	0.720	0.644	0.660

all configurations which is large enough to guarantee high efficiency but smaller than nh_n in any configuration. J_n^{pi} is set equal to 25.

Figure 1 shows the empirical power and a global testing error including misspecification of the size for a typical testing level $\alpha = 0.05$ and for $n = 30,000$. The power of all configurations is quite high. Starting with values $r_{30,000}^{-1} = 2$, that means the smoothing window is two bins in each direction, the power significantly increases by choosing larger values of $r_{30,000}^{-1}$. However, larger values of $r_{30,000}^{-1}$ lead to a misspecification of the size. The global testing error which adds the misspecification of size with equal weight to the power is minimal for $r_{30,000}^{-1} = 4$. On the other hand, the performance is remarkably robust across all considered values of $h_{30,000}$.

The precise values of empirical power and size for $n = 5,000$, depending on $r_{5,000}$ and $h_{5,000}$ are given in Table 1 and Table 2. Again, the global error measure becomes minimal when $r_{5,000}^{-1} = 4$, not changing much for $r_{5,000}^{-1} = 3$ or 5, and being very robust with respect to $h_{5,000}$.

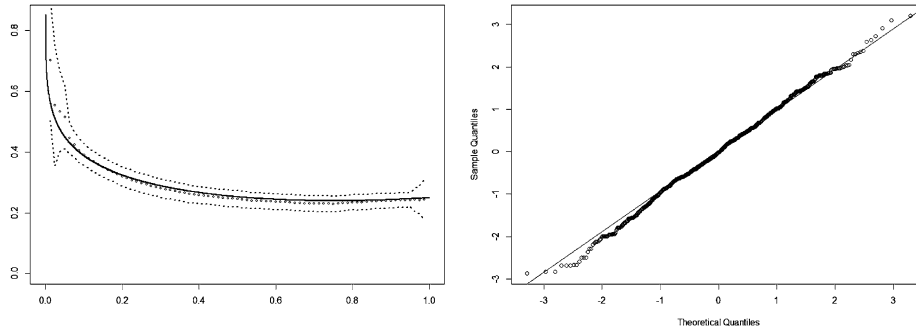


FIG 2. Bin-wise averages of spot squared volatility estimates (points) with bin-wise standard deviations (dashed lines) in comparison to the true spot squared volatility (solid line), for $n = 5,000$ left. Qq-normal plot for feasible versions of the estimates on bin 40 at $t = 1/2$ right.

4.3. Simulation results for spot volatility estimation with a comparison to a multi-scale approach

We analyze the accuracy of the spectral spot volatility estimator. First, we illustrate its performance in the model from Section 4.1, with only a non-random but time-varying volatility component $\varphi_t = 1 - \frac{3}{5}t^{1/5} + \frac{1}{10}t^2$ without volatility jumps. This allows a convenient visualization of the estimation uncertainty. We always use $h_n^{-1} = 150$ for $n = 30,000$, and $h_n^{-1} = 80$ for $n = 5,000$, $r_n^{-1} = 4$ and $J = 30$ and an average of estimators (17a) and (17b) for the spectral spot volatility estimation of σ_s^2 . For the noise specification (33) with pronounced serial correlations, we use the global version of (43) for the bias-correction terms. Figure 2 shows the theoretical squared volatility function in comparison to the bin-wise average estimates with standard deviations for $n = 5,000$ from 3,000 Monte Carlo runs. The empirical standard deviations on the bins (except the bins close to the boundaries) are quite close to their theoretical values $n^{-1/8} \sqrt{8\sigma_s^3 \eta_s^{1/2}}$. For instance, on bin 40 close to $t = 1/2$, we have a ratio of ca. 1.1 of empirical to theoretical standard deviation. Figure 2 also depicts the accuracy of the feasible central limit theorem (27) for bin 40.

Next, we compare the performance of our spectral spot volatility estimator to that of a noise-robust multi-scale spot volatility estimator. The multi-scale estimator for integrated volatility is adopted from [44]. Applied to all data it estimates $\int_0^1 \sigma_t^2 dt$ and we denote it by $\widehat{\langle X, X \rangle}_1$. In order to obtain an estimator of σ_t^2 at some $t \in (0, 1)$, we use a local difference $\widehat{\langle X, X \rangle}_t - \widehat{\langle X, X \rangle}_{t-\delta}$ with suitable small δ . This extends the methods by [36] and [45] from two-scale to multi-scale versions. Though no theoretical results are established for this estimator, it is clear that for optimal δ the approach renders a rate-optimal multi-scale spot volatility estimator. A tuning parameter, the multi-scale frequency, is chosen

TABLE 3
Accuracy of spectral and multi-scale spot volatility estimators.

n	noise model	$\eta^{1/2}$	MISE	
			spectral	multi-scale
30,000	(34)	0.01	0.0216	0.0713
30,000	(34)	0.005	0.0162	0.0421
30,000	(34)	0.0025	0.0146	0.0285
5,000	(34)	0.01	0.0328	0.0855
5,000	(34)	0.005	0.0246	0.0702
5,000	(34)	0.0025	0.0227	0.0698
30,000	(35)	0.01	0.0231	0.0792
30,000	(35)	0.005	0.0184	0.0555
30,000	(35)	0.0025	0.0170	0.0469
5,000	(35)	0.01	0.0597	0.1015
5,000	(35)	0.005	0.0540	0.0892
5,000	(35)	0.0025	0.0517	0.0875

data-driven in an optimal-way, for which a formula is provided in Section 6 of [11].

The multi-scale estimator gets biased under autocorrelated noise as in (33). Thus, we focus on noise models without serial correlation to draw a meaningful comparison. First, consider

$$\epsilon_i \stackrel{iid}{\sim} N\left(0, \eta \left(\int_0^1 \varphi_t^4 \sigma_t^4 dt \right)^{1/2}\right), \quad i = 0, \dots, n. \quad (34)$$

In this model, the bias-correction of the spectral estimator uses a standard noise variance estimator for i.i.d. noise. Further, we examine the estimators in the following noise model with time-varying and endogenous noise:

$$\epsilon_i \sim N\left(0, \eta \left(\int_0^1 \varphi_t^4 \sigma_t^4 dt \right)^{1/2} + \left(\sum_{l=1}^5 \frac{l}{15} |\Delta_{i-l}^n X| \right)^2\right), \quad i = 5, \dots, n, \quad (35)$$

and (34) for $i = 0, \dots, 5$. Here, we use locally bin-wise estimated noise levels for the bias-correction terms.

Since generated volatility paths in our simulation model are random and thus different in each run, we measure the discrepancy for each path. A suitable global quantity to assess the estimators' qualities from M Monte Carlo iterations is an average normalized mean integrated squared error

$$\text{MISE} = \frac{1}{M} \sum_{m=1}^M \int_0^1 \left(\frac{\hat{\sigma}_t^2}{\sigma_t^2} - 1 \right)^2 dt.$$

The integrals are approximated by sums. For the multi-scale estimator, we set $\delta = K_{MS}^{-1}$ and compute spot volatility estimates on a grid of K_{MS} equidistant time points. An optimization of the MISE led us to fix $K_{MS} = 30$ for $n = 30,000$,

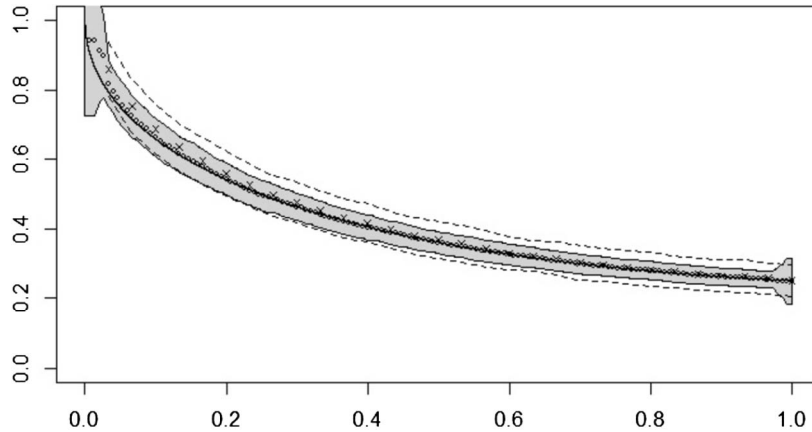


FIG 3. Bin-wise averages of spectral (points) and multi-scale (crosses) spot squared volatility estimates with bin-wise standard deviations (dashed lines) in comparison to the true spot squared volatility (solid line), for $n = 30,000$. The area around the spectral estimates determined by their standard deviations is gray colored such that the other dashed lines depict the standard deviations of the multi-scale estimates.

and $K_{MS} = 10$ for $n = 5,000$. For the spectral estimator the discretization is given by the h_n^{-1} bins of length h_n .

An overview of the results for different noise levels, each quantity based on $M = 3,000$ Monte Carlo runs, is given in Table 3. The spectral estimation outperforms the ad hoc multi-scale approach in each model specification. The efficiency gains are most relevant for larger noise and more frequent observations. Figure 3 visualizes spectral and multi-scale spot volatility estimates with their standard deviations when the true volatility is deterministic and given by $\varphi_t = 1 - \frac{3}{5}\sqrt{t} + \frac{1}{10}t^2$. The confidence regions sketched by the point-wise standard deviations are wider for the multi-scale than for the spectral estimator. We further see a small positive bias of the multi-scale estimates. The discretization, chosen to optimize MISE, is also coarser than the bins of the spectral method which we expect to be the main reason for this bias.

Overall, the estimation results for the spectral spot volatility estimator are promising. They confirm that it provides a useful statistical device which is of interest beyond its use as one ingredient for the statistical test for common price and volatility jumps.

4.4. Simulation results for the test with a comparison to a skip-sampling approach

In the sequel, we first study the empirical size and power of our test with respect to different calibrations of volatility jump sizes, noise level and number of observations. To evaluate the improved performance in comparison to the test by [29], we also implement the latter based on appropriately down sampled

TABLE 4
Parameter specification for Monte Carlo.

Scenario	n	λ	H	$\eta^{1/2}$	γ	h_n^{-1}	J	J^{pi}	r_n^{-1}
I-Hyp	30,000	2	0.25	0.005	0	150	30	25	4
I-Alt	30,000	2	0.25	0.005	1	150	30	25	4
II-Hyp	5,000	2	0.25	0.005	0	80	30	25	4
II-Alt	5,000	2	0.25	0.005	1	80	30	25	4
III-Alt	30,000	2	0.10	0.005	1	200	30	25	5
IV-Hyp	5,000	2	0.25	0.015	0	80	30	25	4
IV-Alt	5,000	2	0.25	0.015	1	80	30	25	4

discretized simulated paths.

The parameter configurations used in the Monte Carlo study for different scenarios are summarized in Table 4 together with the chosen tuning parameters according to the values found to be optimal in the sensitivity analysis. In scenario II (I) the average price jump is approx. 20 (60) times larger than the average absolute return. The identification of price jumps by truncation thus works with only very few errors. Hence, we can use the results from all Monte Carlo iterations to analyze our methods' performance. Examining the ability of thresholding to locate price jumps in different situations has been addressed in [16]. Here, the focus is on the test for common price and volatility jumps. The volatility jumps in scenarios I, II and IV are a bit smaller than half the size of the average range of the simulated continuous part of the intra-day volatility path. Figure 6 illustrates that in empirical applications much larger volatility jumps occur. In scenario III the jump in the volatility is less than 20% of the range of the continuous intra-day volatility motion. In scenarios I, II and IV we thus have a volatility jump size where the test should attain reasonable power, while scenario III investigates the behavior for rather small volatility jumps.

We compare the performance of our test based on the statistic (28) in scenario I for our simulated model with the method by [29]. We cannot apply the latter to the simulated $n = 30,000$ high-frequency observations, since the simulated data contains noise. If we apply the test for direct observations to noisy data, the statistics are heavily biased and the performance is very poor. Instead, we skip-sample simulations at a coarser frequency. A heuristic optimization leads us in scenario I of our simulation study to an optimal skip-sample frequency resulting in ca. 500 "de-noised" observations on $[0, 1]$. For intra-day NASDAQ data this translates in using one observation per 46.8 seconds. In [29] a one minute frequency for different – but also very liquid – data is used in the application part. Moderate changes of the skip-sampling frequency do not affect the results substantially. Figure 4 demonstrates a very good performance of our test in scenario I. The power is 97.7% for the $\alpha = 0.05$ -test and above 90% even for level $\alpha = 0.01$. Similar to our test, the performance of the Jacod-Todorov test applied to the 500 coarse returns is crucially influenced by the length of the smoothing window of local realized volatilities. We visualize two configu-

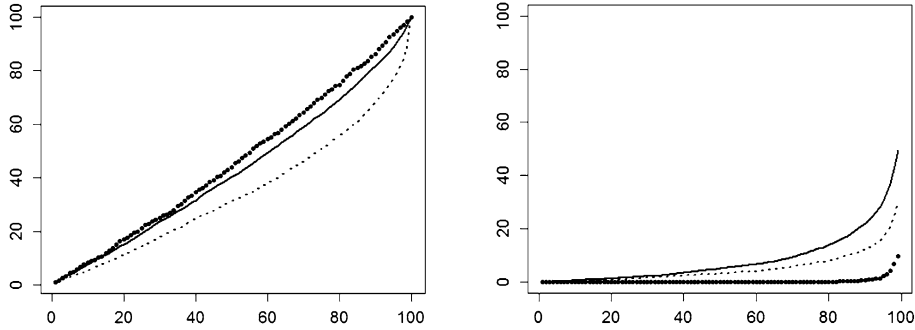


FIG 4. Comparison of the test and the Jacod-Todorov test based on skip-sampled data.

Empirical size and power of the tests in scenario I under the null hypothesis (left) and alternative hypothesis (right). Empirical amount of realizations smaller or equal percentiles of theoretical asymptotic distribution under the null (y-axis) against those percentiles (x-axis). The dotted line shows results for our test and the solid and dashed line two versions of the Jacod-Todorov test using two different tuning parameters. The skip-sampling frequency is optimized to allow for the highest power.

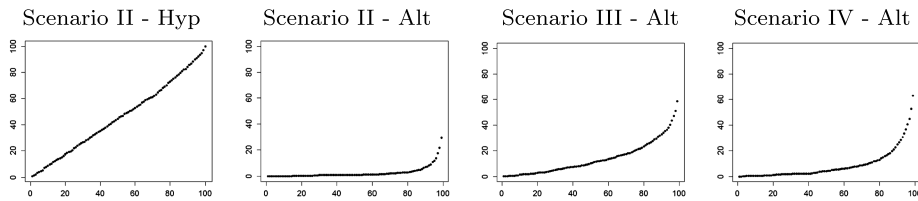


FIG 5. Empirical size and power of the tests in scenarios II, III and IV. Empirical amount of realizations smaller or equal percentiles of theoretical asymptotic distribution under the null (y-axis) against those percentiles (x-axis).

rations with $k_n = 50, 100$ in the spot volatility estimators given in (4). The choice $k_n = 100$ is in favor of higher power, but the accuracy of the asymptotic quantiles on the null hypothesis is not good. Setting $k_n = 50$, we obtain less power but the empirical quantiles on the null hypothesis track the asymptotic ones more closely. In all configurations, the performance of the Jacod-Todorov test applied to skip-sampled data is inferior to the power of our noise-robust approach. This is not surprising, since for our approach we rely on an efficient smoothing technique while skip-sampling can be seen as the simplest method to smooth out noise. The performance of the Jacod-Todorov test is reasonably well also, but in a situation with large available sample sizes and significant noise it is worth to apply the more efficient, noise-robust procedure. If sample sizes are smaller (and the noise not larger), the difference between the two methods becomes smaller. Figure 5 shows the performance in other scenarios II, III and IV. Decreasing the sample size to $n = 5,000$ observations in scenario II, while all parameters are equal as in scenario I, leads to a slightly smaller power and larger misspecification of the size. The power is still higher than for

the skip-sample approach, but the difference is less relevant. With the tuning parameters which minimize the global empirical testing error, the misspecification of the size is still acceptable. Larger noise levels result in smaller power as shown for scenario IV in Figure 5, while the fit of the size remains good. In this situation, the Jacod-Todorov method would only work for less frequent skip-sampling resulting as well in smaller power. For the alternative hypothesis with a small volatility jump in scenario III, a sensitivity analysis as in Section 4.2 led us to slightly different tuning parameters, $h_{30,000}^{-1} = 200$ and $r_{30,000}^{-1} = 5$. Since smaller bins give a higher time resolution, it is not surprising that detecting small volatility jumps in a rapidly time-varying spot volatility works better for a finer time resolution. On the other side, choosing r_n^{-1} slightly larger leads to almost the same window length $r_n^{-1}h_n$ for spot volatility estimation as before. The power for such small volatility jumps is less, but still ca. 60% for $\alpha = 0.05$.

5. Data study

To provide evidence about the practical relevance of price-volatility co-jumps and to study the usefulness of our estimators and test in a real-world data environment, we apply our methodology to stocks traded at the exchange platform NASDAQ. The data study is based on limit order book data taken from the online data tool LOBSTER². The example refers to stocks of the online and technology companies Amazon.com Inc. (AMZN), Apple Inc. (AAPL), Facebook Inc. (FB), Intel Corp. (INTC) and Microsoft Corp. (MSFT). We focus on transaction prices of 252 trading days in the year 2013. A trading day spans from 9:30 to 16:00 EDT and includes for a single stock a minimum of 4,267 (AMZN 2013-07-03) up to a maximum of 210,812 (FB 2013-10-31) transactions. One benefit of our estimator and test is that we can directly plug-in traded log-prices, reconstructed from the order book, without considering any skip-sampling or synchronization procedures. Since the method is robust against market microstructure noise, we efficiently take into account all information stored in the data.

Guided by our theoretical results and the simulations, estimates and tests are based on spectral statistics calculated for $k = 0, 1, \dots, h_n^{-1} - 1$ bins over a trading day, with $h_n^{-1} = \lfloor 3\sqrt{n}/\log(n) \rfloor$. We set $J = 30$ and $J^{pi} = 15$. Jumps in prices are detected with the locally adaptive threshold $\hat{u}_k = 2 \log(h_n^{-1})h_n\hat{\sigma}_{kh_n, pil}^2$, with $\hat{\sigma}_{kh_n, pil}^2$ the pilot estimator (15) of the spot squared volatility. We fix constant window lengths $r_n^{-1} = 4$. Surely, r_n^{-1} determines a crucial parameter which can be studied to learn about the persistence or live-time of a break in spot volatility. We apply the test to each day separately.

Table 5 reports the rejection rates for the 5% and 10% significance levels. Results indicate that on a 10% significance level 36% (INTC) up to 73% (AAPL) of jumps in prices are accompanied by jumps in volatility. It appears that the rejection rate decreases in the number of detected price jumps. This leads to

²LOBSTER academic data- lobsterdata.com, powered by NASDAQ OMX

TABLE 5
Testing for price and volatility jumps in NASDAQ order book data.

Stocks	# of days with price jumps	Rejection rate (price-volatility jumps)		Sample averages (whole year 2013)	
		$\alpha = 5\%$	$\alpha = 10\%$	n	$\Delta\hat{\sigma}_s^2$
Amazon	21	52.4%	61.9%	10,924	31.2%
Apple	22	63.6%	72.7%	36,947	36.5%
Facebook	37	46.0%	51.4%	41,354	27.8%
Intel	47	27.7%	36.2%	18,535	23.0%
Microsoft	22	50.0%	50.0%	28,052	31.2%

Notes: Estimation and test executed for each day in the year 2013 separately. n indicates the average number of observed trades per trading day, $\Delta\hat{\sigma}_s^2 = |\log(\hat{\sigma}_s^2) - \log(\hat{\sigma}_{s-}^2)|$ is the average estimated relative size of volatility jumps. Sample period: 01-02-2013 to 12-31-2013 (252 days).

relatively stable frequencies of price-volatility co-jumps over time across the considered stocks. Referring to the 5% significance level, the Amazon.com stock displays with around 4.4% of the trading days the lowest frequency of common price and volatility jumps. With around 6.7% of trading days, Facebook Inc. has the largest number of common jumps. Absolute jump sizes of the log squared volatility processes reported in Table 5 are considerably large.

Figure 6 illustrates the mechanisms behind the test for common price and volatility jumps. Left hand plots show an upward jump in prices on bin $k = 58$, whereas right hand plots show a downward jump in prices on bin $k = 39$. Both price jumps are associated with a significant contemporaneous upward jump in spot volatility. The p -value in both examples is 0.00. On the first example date, August, 13th 2013, the investor Carl Icahn has taken a large stake of AAPL stocks. On May 14th, the downward jump example date, figures of mobile phone sales have been reported.

We find evidence for frequent occurrences of simultaneous jumps in price and volatility and quite large volatility jump sizes. Yet, by far not all detected price jumps are accompanied by volatility jumps. Understanding the economic sources of different jump events and their consequences for price-volatility co-jumps is of interest for future research.

6. Conclusion

We present a new test for the presence of contemporaneous jumps of price and volatility based on high-frequency data. The test transfers the methodology of [29] to a setup accounting for microstructure noise by employing a spectral estimation of the spot volatility and an accurate test function. The nonparametric spot volatility estimator shows appealing asymptotic and finite-sample qualities and is of interest beyond the scope of this article. It opens up several new ways for inference in models for high-frequency financial data with noise.

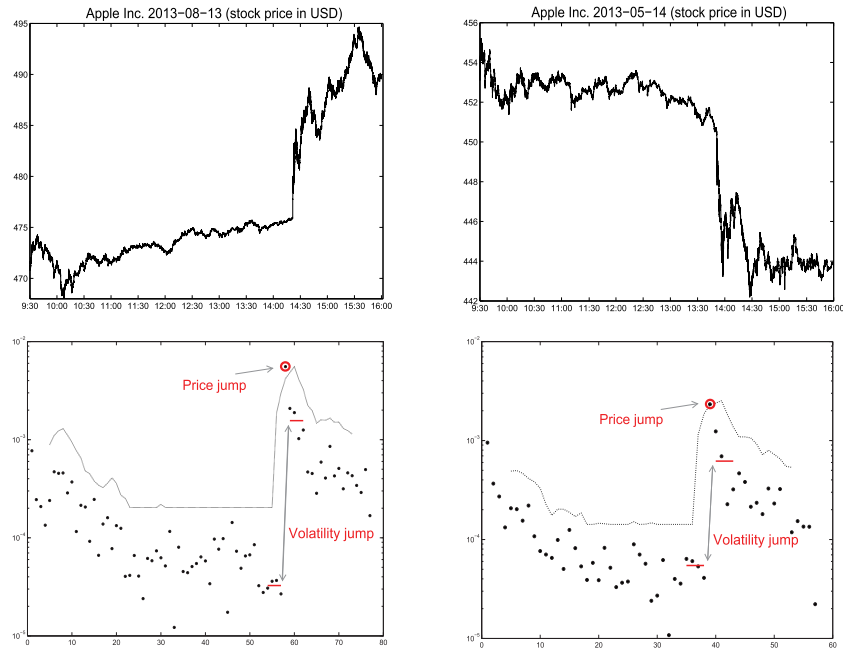


FIG 6. **Examples of common price and volatility jumps.** Upper figures indicate price processes of the Apple Inc. stock. Lower figures display the related spectral quadratic variation estimates for the bins $k = 0, \dots, h_n^{-1}$. Estimates exceeding the threshold (dotted line) are marked as price jumps. The difference between spot volatility estimates (bars) estimate the volatility jump. 2013-08-13: $n = 87,445$, 2013-05-14: $n = 40,707$.

Simulations demonstrate that the proposed noise-robust test increases the finite-sample performance considerably compared to an application of the test by [29] to skip-sampled data. Our data study reveals cogent significance of price and volatility co-jumps in NASDAQ high-frequency data. The presented methods can be generalized in various directions. For instance, our methods guide the way how a test for correlation of price and volatility jumps, as presented by [25] for a non-noisy observation design, can be constructed. A general global test for volatility jumps under noise generalizing the methods from [14] could be addressed with a related high-dimensional testing procedure.

7. Proofs

7.1. Preliminaries

On the finite time horizon $[0, 1]$, we may augment local boundedness to uniform boundedness in Assumption (H-r) and Assumption 1, such that we can assume that there exists a constant Λ with

$$\max \{ |b_s(\omega)|, \sigma_s^2(\omega), |X_s(\omega)|, |\delta_\omega(s, x)|/\gamma(x), \eta_s(\omega), \rho_s(\omega) \} \leq \Lambda,$$

for all $(\omega, s, x) \in (\Omega, \mathbb{R}_+, \mathbb{R})$. This standard procedure can be found in Section 4.4.1 of [28]. Throughout the proofs K is a generic constant and K_p a constant emphasizing dependence on p . We decompose the semimartingale X in its continuous part

$$C_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s,$$

and the jump component

$$J_t = \int_0^t \int_{\mathbb{R}} \delta(s, x) \mathbb{1}_{\{|\delta(s, x)| \leq 1\}} (\mu - \nu)(ds, dx) + \int_0^t \int_{\mathbb{R}} \delta(s, x) \mathbb{1}_{\{|\delta(s, x)| > 1\}} \mu(ds, dx).$$

The processes

$$\tilde{C}_t^n = \int_0^t \sigma_{\lfloor sh_n^{-1} \rfloor h_n} dW_s \tag{36}$$

serve as an approximation of C_t by simplified processes without drift and with locally constant volatility. We separate jumps with absolute value bounded from above by some $\varepsilon < 1$ and larger jumps:

$$\begin{aligned} J_t &= J(\varepsilon)_t + \int_0^t \int_{\mathbb{R} \setminus A_\varepsilon} \delta(s, x) \mathbb{1}_{\{|\delta(s, x)| \leq 1\}} (\mu - \nu)(ds, dx) \\ &\quad + \int_0^t \int_{\mathbb{R}} \delta(s, x) \mathbb{1}_{\{|\delta(s, x)| > 1\}} \mu(ds, dx), \end{aligned}$$

with $A_\varepsilon = \{z \in \mathbb{R} | \gamma(z) \leq \varepsilon\}$ and later let $\varepsilon \rightarrow 0$. Let us recall some usual estimates on Assumptions 1, (H-r) and $(\sigma-\alpha)$ which are crucial for the following proofs. For the continuous semimartingale part, we have

$$\forall p \in \mathbb{N}, s, t \geq 0 : \mathbb{E} [|C_{s+t} - C_s|^p | \mathcal{F}_s] \leq K_p t^{\frac{p}{2}}. \tag{37a}$$

For given $0 < \varepsilon < 1$, for $J(\varepsilon)$ the estimate

$$\begin{aligned} \mathbb{E} [|J(\varepsilon)_{s+t} - J(\varepsilon)_s|^p | \mathcal{F}_s] &\leq K_p \mathbb{E} \left[\left(\int_s^{s+t} \int_{A_\varepsilon} (\gamma^2(x) \wedge 1) \mu(d\tau, dx) \right)^{\frac{p}{2}} \right] \\ &\leq K_p t^{(\frac{p}{2} \wedge 1)} \gamma_\varepsilon^{(\frac{p}{2} \wedge 1)}, \quad \forall p \in \mathbb{N}, \forall s, t \geq 0, \end{aligned} \tag{37b}$$

holds with $\gamma_\varepsilon = \int_{A_\varepsilon} (\gamma^2(x) \wedge 1) \lambda(dx) \leq K\varepsilon^{(2-r)}$.

The continuous semimartingale increments satisfy local Gaussianity in the sense that

$$\begin{aligned} \mathbb{E} [|C_{s+t} - C_s - (\sigma_s(W_{s+t} - W_s))|^p | \mathcal{F}_s] &\leq K_p \mathbb{E} \left[\left(\int_s^{s+t} |\sigma_\tau - \sigma_s|^2 d\tau \right)^{\frac{p}{2}} | \mathcal{F}_s \right] \\ &\leq K_p t^{\frac{p}{2}} \mathbb{E} \left[\sup_{\tau \in [s, s+t]} (|\sigma_\tau - \sigma_s|^p) | \mathcal{F}_s \right] \\ &\leq K_p t^{\frac{p}{2}(1+2\alpha)}, \quad \forall p \in \mathbb{N}, s, t \geq 0, \end{aligned} \tag{37c}$$

on Assumption $(\sigma-\alpha)$. The probability of a frequent occurrence of large jumps is small. Precisely, the expectation of jumps with absolute value larger than ε is bounded:

$$\forall s, t \geq 0: \mathbb{E} [|J_{s+t} - J_s - (J(\varepsilon)_{s+t} - J(\varepsilon)_s)| | \mathcal{F}_s] \leq Kt\varepsilon^{-r}. \quad (37d)$$

Under Assumption $(H-r)$ with $r \geq 1$, the jumps moreover satisfy

$$\begin{aligned} \mathbb{E} [|J_t - J_s|^p | \mathcal{F}_s] &\leq K_p \mathbb{E} \left[\left(\int_s^t \int_{\mathbb{R}} (\gamma^r(x) \wedge 1) \lambda(dx) ds \right)^{p/r} \right] \\ &\leq K_p |t - s|^{((p/r) \wedge 1)}, \quad \forall s, t \geq 0. \end{aligned} \quad (37e)$$

Under Assumption $(\sigma-\alpha)$ for $0 \leq s < t \leq 1$, the squared volatility satisfies:

$$\mathbb{E} [|\sigma_t^2 - \sigma_s^2| | \mathcal{F}_s] \leq |t - s|^\alpha. \quad (37f)$$

Proofs of these bounds can be found, for instance, in Chapter 13 of [28]. (37b) follows from Equation (54) in [2].

In the sequel, we gather more properties of the basis functions (7). We define (Φ_{jk}) in (7) in the same way as [15] in their Equation (4b) to exploit discrete-time Fourier identities under equidistant sampling. The asymptotic properties of the estimator remain the same when we use

$$\tilde{\Phi}_{jk}(t) = \frac{\sqrt{2h_n}}{j\pi} \sin(j\pi h_n^{-1}(t - kh_n)) \mathbb{1}_{[kh_n, (k+1)h_n]}(t), \quad j \geq 1, \quad (38)$$

instead which equals the definition from Equation (2.2) in [12]. We heavily exploit the summation by parts identity for spectral statistics

$$S_{jk} = \|\Phi_{jk}\|_n^{-1} \left(\sum_{i=1}^n \Delta_i^n X \Phi_{jk}\left(\frac{i}{n}\right) - \sum_{i=1}^{n-1} \epsilon_i \varphi_{jk}\left(\frac{i+1/2}{n}\right) \frac{1}{n} \right), \quad (39)$$

with $\varphi_{jk}(t) = \sqrt{2}h_n^{-1/2} \cos(j\pi h_n^{-1}(t - kh_n)) \mathbb{1}_{[kh_n, (k+1)h_n]}(t)$, see Lemma 6.1 of [5]. For all (φ_{jk}) , it holds that

$$n^{-1} \sum_{i=1}^{n-1} \varphi_{jk}\left(\frac{i+1/2}{n}\right) \varphi_{j'k}\left(\frac{i+1/2}{n}\right) = \delta_{jj'}. \quad (40)$$

For the asymptotic theory, we shall further use the following identities

$$\int_0^1 \tilde{\Phi}_{jk}^2(t) dt = \frac{h_n^2}{\pi^2 j^2}, \quad \int_0^1 \varphi_{jk}^2(t) dt = 1, \quad (41a)$$

$$\int_0^1 \tilde{\Phi}_{jk}(t) \varphi_{uk}(t) dt = \frac{(1 - \cos(\pi j) \cos(\pi u)) 2h_n}{\pi^2(j^2 - u^2)}. \quad (41b)$$

The latter gives $4h_n/(\pi^2(j^2 - u^2))$ whenever j is odd and u even, or the other way round, and vanishes in all other cases. Recall the definition of the weights

(13). The magnitude of these weights is

$$\begin{aligned}
 w_{jk} &\leq I_{jk} = \frac{1}{2} \left(\sigma_{kh_n}^2 + \frac{\eta_{kh_n}}{n} \|\Phi_{jk}\|_n^{-2} \right)^{-2} = \mathcal{O} \left(\left(1 + \frac{j^2}{nh_n^2} \right)^{-2} \right) \\
 &= \begin{cases} \mathcal{O}(1) & \text{for } j \leq \sqrt{nh_n} \\ \mathcal{O}(j^{-4}n^2h_n^4) & \text{for } j > \sqrt{nh_n} \end{cases}, \tag{42}
 \end{aligned}$$

with $\|\Phi_{jk}\|_n^{-2} \approx \pi^2 j^2 h_n^{-2} = \left(\int_0^1 \tilde{\Phi}_{jk}^2(t) dt \right)^{-1}$.

In the proofs, we use the notation $\zeta_k^{ad}(Z)$ and $\zeta_k(Z)$ from (12) analogously also for different processes Z . This means that we insert in (12) spectral statistics $S_{jk}(Z)$, analogous to (11), computed from the sequence $Z_{i/n}, i = 0, \dots, n$, especially $\zeta_k(X)$ for the statistics based on the unobserved efficient price.

7.2. Estimation of the noise long-run variance

First, consider the standard case where $\alpha \leq 1/2$ in Assumption $(\sigma-\alpha)$, such that $\beta < 1/4$. To estimate (η_{kh_n}) under (22), we use nh_n observations on the bin $[kh_n, (k+1)h_n]$. For $k = 0, \dots, h_n^{-1} - 1$, and $u = 1, \dots, Q$, define the cumulative empirical autocorrelation statistics

$$\begin{aligned}
 Z_{kh_n}^{(u)} &= \frac{1}{2nh_n} \sum_{i=knh_n+1}^{(k+1)nh_n} (\Delta_i^n Y)^2 + \frac{1}{nh_n} \sum_{l=1}^u \sum_{i=knh_n+1}^{(k+1)nh_n-u} \Delta_i^n Y \Delta_{i+l}^n Y, \\
 \tilde{Z}_{kh_n}^{(u)} &= \frac{1}{2nh_n} \sum_{i=knh_n+1}^{(k+1)nh_n} (\Delta_i^n Y)^2 + \frac{1}{nh_n} \sum_{l=1}^u \sum_{i=knh_n+u+1}^{(k+1)nh_n} \Delta_i^n Y \Delta_{i-l}^n Y.
 \end{aligned}$$

For $u = 0$, the rescaled local realized volatilities in the first addend define $Z_{kh_n}^{(0)}$. We estimate η_{kh_n} by

$$\hat{\eta}_{kh_n} = \sum_{u=0}^Q (u+1) Z_{kh_n}^{(u)} + \sum_{u=1}^Q u \tilde{Z}_{kh_n}^{(u)}. \tag{43}$$

We assume that Q is known. However, the same result applies if the process is \tilde{Q} -dependent with $\tilde{Q} < Q$. It thus suffices to take Q sufficiently large. A statistical method to infer Q is provided by [13].

We consider separately the case $\alpha > 1/2$ with possible values $1/4 \leq \beta < 1/3$. Then, we exploit the increased smoothness of the noise by (22) to estimate (η_{kh_n}) with an improved convergence rate. We partition $[0, 1]$ in n/M_n windows of lengths M_n/n , each with M_n observations, where $M_n = c_M n^{1-(2\alpha+1)^{-1}}$. For a simple exposition we may suppose $M_n, n/M_n \in \mathbb{N}$ again. Completely analogously as before, we compute the cumulative empirical autocorrelation statistics $Z_{k \frac{M_n}{n}}^{(u)}, \tilde{Z}_{k \frac{M_n}{n}}^{(u)}$ for $k = 0, \dots, n/M_n - 1$. The estimates $(\hat{\eta}_{kh_n})$ are now obtained by

$$\hat{\eta}_{kh_n} = \hat{\eta}_{\tilde{k} \frac{M_n}{n}} \mathbb{1}_{\{kh_n \in [\tilde{k} \frac{M_n}{n}, (\tilde{k}+1) \frac{M_n}{n}]\}}, \tag{44}$$

with $\hat{\eta}_{\tilde{k} \frac{M_n}{n}}$ analogous to (43) over the coarser time windows.

Proof of Proposition 3.1

We begin with the case $\alpha \leq 1/2$ in Assumption $(\sigma\text{-}\alpha)$, such that $\beta < 1/4$. We prove that

$$\hat{\eta}_{kh_n} = \eta_{kh_n} + \mathcal{O}_{\mathbb{P}}(n^{-1/4}). \quad (45)$$

Considering the expectation of the cumulative empirical autocorrelation statistics, all terms involving increments $\Delta_i^n X$ are of order $\mathcal{O}_{\mathbb{P}}(n^{-1/2})$, and even smaller under exogenous noise. Thus, we have that

$$\begin{aligned} \mathbb{E}[Z_{kh_n}^{(u)} | \mathcal{F}_{kh_n}] &= \frac{1}{2nh_n} \sum_{i=knh_n+1}^{(k+1)nh_n} \mathbb{E}[\epsilon_i^2 + \epsilon_{i-1}^2 - 2\epsilon_i\epsilon_{i-1} | \mathcal{F}_{kh_n}] + \mathcal{O}_{\mathbb{P}}(n^{-1/2}) \\ &+ \frac{1}{nh_n} \sum_{l=1}^u \sum_{i=knh_n+1}^{(k+1)nh_n-u} \mathbb{E}[\epsilon_i\epsilon_{i+l} + \epsilon_{i-1}\epsilon_{i+l-1} - \epsilon_{i-1}\epsilon_{i+l} - \epsilon_i\epsilon_{i+l-1} | \mathcal{F}_{kh_n}] \\ &= \frac{1}{nh_n} \sum_{i=knh_n+1}^{(k+1)nh_n} \mathbb{E}[\epsilon_i^2 - \epsilon_i\epsilon_{i-1} | \mathcal{F}_{kh_n}] + \mathcal{O}_{\mathbb{P}}(n^{-1/2}) \\ &+ \frac{1}{nh_n} \sum_{i=knh_n+1}^{(k+1)nh_n-u} \mathbb{E}[\epsilon_i\epsilon_{i+u} + \epsilon_{i-1}\epsilon_i - \epsilon_{i-1}\epsilon_{i+u} - \epsilon_i^2 | \mathcal{F}_{kh_n}], \end{aligned}$$

where we use $(nh_n)^{-1} = \mathcal{O}(n^{-1/2})$ for the first and the telescoping sum for the second addend. We obtain that

$$\mathbb{E}[Z_{kh_n}^{(u)} | \mathcal{F}_{kh_n}] = \frac{1}{nh_n} \sum_{i=knh_n+1}^{(k+1)nh_n-u} \mathbb{E}[\epsilon_i\epsilon_{i+u} - \epsilon_{i-1}\epsilon_{i+u} | \mathcal{F}_{kh_n}] + \mathcal{O}_{\mathbb{P}}(n^{-1/2}),$$

for all $0 \leq u \leq Q$. Summing over $u \in \{0, \dots, Q\}$, we exploit another telescoping sum:

$$\begin{aligned} \sum_{u=0}^Q (u+1) \mathbb{E}[Z_{kh_n}^{(u)} | \mathcal{F}_{kh_n}] &= \frac{1}{nh_n} \sum_{i=knh_n+1}^{(k+1)nh_n-Q} \sum_{u=0}^Q (u+1) (\text{Cov}(\epsilon_i, \epsilon_{i+u} | \mathcal{F}_{kh_n}) \\ &\quad - \text{Cov}(\epsilon_{i-1}, \epsilon_{i+u} | \mathcal{F}_{kh_n})) + \mathcal{O}_{\mathbb{P}}(n^{-1/2}) \\ &= \frac{1}{nh_n} \sum_{i=knh_n+1}^{(k+1)nh_n-Q} \sum_{u=0}^Q \text{Cov}(\epsilon_i, \epsilon_{i+u} | \mathcal{F}_{kh_n}) + \mathcal{O}_{\mathbb{P}}(n^{-1/2}), \end{aligned}$$

since $\text{Cov}(\epsilon_{i-1}, \epsilon_{i+Q} | \mathcal{F}_{kh_n}) = 0$. There are at most $\tilde{Q} < \infty$ addends $i = knh_n + 1, \dots, knh_n + \tilde{Q}$, for that $\mathbb{E}[\epsilon_i | \mathcal{F}_{kh_n}] \neq 0$ is possible by endogeneity, which are asymptotically negligible in the above sum. A similar computation for $\tilde{Z}_{kh_n}^{(u)}$

gives:

$$\begin{aligned} \mathbb{E}[\tilde{Z}_{kh_n}^{(u)} | \mathcal{F}_{kh_n}] &= \frac{1}{2nh_n} \sum_{i=knh_n+1}^{(k+1)nh_n} \mathbb{E}[\epsilon_i^2 + \epsilon_{i-1}^2 - 2\epsilon_i\epsilon_{i-1} | \mathcal{F}_{kh_n}] + \mathcal{O}_{\mathbb{P}}(n^{-1/2}) \\ &+ \frac{1}{nh_n} \sum_{l=1}^u \sum_{i=knh_n+u+1}^{(k+1)nh_n} \mathbb{E}[\epsilon_i\epsilon_{i-l} + \epsilon_{i-1}\epsilon_{i-l-1} - \epsilon_{i-1}\epsilon_{i-l} - \epsilon_i\epsilon_{i-l-1} | \mathcal{F}_{kh_n}] \\ &= \frac{1}{nh_n} \sum_{i=knh_n+1}^{(k+1)nh_n} \mathbb{E}[\epsilon_{i-1}^2 - \epsilon_i\epsilon_{i-1} | \mathcal{F}_{kh_n}] + \mathcal{O}_{\mathbb{P}}(n^{-1/2}) \\ &+ \frac{1}{nh_n} \sum_{i=knh_n+u+1}^{(k+1)nh_n} \mathbb{E}[\epsilon_i\epsilon_{i-1} + \epsilon_{i-1}\epsilon_{i-u-1} - \epsilon_{i-1}^2 - \epsilon_i\epsilon_{i-u-1} | \mathcal{F}_{kh_n}], \end{aligned}$$

and thus that

$$\begin{aligned} \sum_{u=1}^Q u \mathbb{E}[\tilde{Z}_{kh_n}^{(u)} | \mathcal{F}_{kh_n}] &= \frac{1}{nh_n} \sum_{i=knh_n+Q+1}^{(k+1)nh_n} \sum_{u=1}^Q u (\text{Cov}(\epsilon_i, \epsilon_{i-u} | \mathcal{F}_{kh_n}) \\ &\quad - \text{Cov}(\epsilon_i, \epsilon_{i-u-1} | \mathcal{F}_{kh_n})) + \mathcal{O}_{\mathbb{P}}(n^{-1/2}) \\ &= \frac{1}{nh_n} \sum_{i=knh_n+Q+1}^{(k+1)nh_n} \sum_{u=1}^Q \text{Cov}(\epsilon_i, \epsilon_{i-u} | \mathcal{F}_{kh_n}) + \mathcal{O}_{\mathbb{P}}(n^{-1/2}), \end{aligned}$$

since $\text{Cov}(\epsilon_i, \epsilon_{i-Q-1} | \mathcal{F}_{kh_n}) = 0$ for all, except finitely many, i . This yields that for the estimator (43)

$$\mathbb{E}[\hat{\eta}_{kh_n} | \mathcal{F}_{kh_n}] = \frac{1}{nh_n} \sum_{i=knh_n+Q+1}^{(k+1)nh_n-Q} \sum_{u=-Q}^Q \text{Cov}(\epsilon_i, \epsilon_{i+u} | \mathcal{F}_{kh_n}) + \mathcal{O}_{\mathbb{P}}(n^{-1/2}), \tag{46}$$

such that $\sup_{t \in [kh_n, (k+1)h_n]} |\eta_{kh_n} - \eta_t| = \mathcal{O}(h_n^{(1/2+\delta)\vee\alpha}) = \mathcal{O}(n^{-1/4})$ and (20) give that

$$\mathbb{E}[\hat{\eta}_{kh_n} | \mathcal{F}_{kh_n}] = \eta_{kh_n} + \mathcal{O}_{\mathbb{P}}(n^{-1/4}). \tag{47}$$

The following bound for the conditional variance of the estimator (43) completes the proof of (45). It holds uniformly in k that

$$\begin{aligned} \text{Var}(Z_{kh_n}^{(u)} | \mathcal{F}_{kh_n}) &\leq \frac{1}{2n^2h_n^2} \sum_{i=knh_n+1}^{(k+1)nh_n} \sum_{u=-Q}^Q \text{Cov}((\Delta_i^n \epsilon)^2, (\Delta_{i+u}^n \epsilon)^2 | \mathcal{F}_{kh_n}) \\ &\quad + \frac{2}{n^2h_n^2} \sum_{i,j,l,u} \text{Cov}(\Delta_i^n \epsilon \Delta_{i+l}^n \epsilon, \Delta_j^n \epsilon \Delta_{j+u}^n \epsilon | \mathcal{F}_{kh_n}) + \mathcal{O}_{\mathbb{P}}(n^{-1/2}) \\ &= \mathcal{O}_{\mathbb{P}}((nh_n)^{-1}Q^3) = \mathcal{O}_{\mathbb{P}}(n^{-1/2}), \end{aligned}$$

since the covariances vanish whenever the difference of two indices exceeds $Q < \infty$. Analogously, we derive that $\text{Var}(\tilde{Z}_{kh_n}^{(u)} | \mathcal{F}_{kh_n}) = \mathcal{O}_{\mathbb{P}}(n^{-1/2})$ for all k . This readily implies that $\text{Var}(\hat{\eta}_{kh_n}) = \mathcal{O}(n^{-1/2})$, and with Chebyshev's inequality and (47) we conclude that $\hat{\eta}_{kh_n} = \eta_{kh_n} + \mathcal{O}_{\mathbb{P}}(n^{-1/4})$.

It remains to prove (43) for $\alpha \geq 1/2$. Then, $\sup_{t \in [\tilde{k}M_n/n, (\tilde{k}+1)M_n/n]} |\eta_{\tilde{k}M_n/n} - \eta_t| = \mathcal{O}((M_n/n)^\alpha) = \mathcal{O}(n^{-\frac{\alpha}{2\alpha+1}}) = \mathcal{O}(n^{-\beta})$ by (25). Repeating the steps for estimates $\hat{\eta}_{kh_n}$ from nh_n observations, we now obtain with M_n observations, for all $\tilde{k} = 0, \dots, n/M_n - 1$, that

$$\begin{aligned} \hat{\eta}_{\tilde{k} \frac{M_n}{n}} &= \eta_{\tilde{k} \frac{M_n}{n}} + \mathcal{O}_{\mathbb{P}}(M_n^{-1/2}) = \eta_{\tilde{k} \frac{M_n}{n}} + \mathcal{O}_{\mathbb{P}}\left(n^{-\frac{1}{2}\left(1 - \frac{1}{2\alpha+1}\right)}\right) \\ &= \eta_{\tilde{k} \frac{M_n}{n}} + \mathcal{O}_{\mathbb{P}}(n^{-\frac{\alpha}{2\alpha+1}}) = \eta_{\tilde{k} \frac{M_n}{n}} + \mathcal{O}_{\mathbb{P}}(n^{-\beta}). \end{aligned}$$

This proves Proposition 3.1. \square

7.3. Stable convergence of the spot squared volatility estimators

We first prove two lemmas, one on moments of the noise terms in the spectral statistics and one on moments of the statistics (12).

Lemma 1. *On Assumption (η -p) with $p \geq 4$:*

$$\mathbb{E}\left[\left(\sum_{i=1}^{n-1} \epsilon_i \varphi_{jk} \left(\frac{i+1/2}{n}\right) n^{-1}\right)^2 \middle| \mathcal{F}_{kh_n}\right] = \eta_{kh_n} n^{-1} + \mathcal{O}_{\mathbb{P}}(n^{-1}), \quad (48a)$$

$$\mathbb{E}\left[\left(\sum_{i=1}^{n-1} \epsilon_i \varphi_{jk} \left(\frac{i+1/2}{n}\right) n^{-1}\right)^4 \middle| \mathcal{F}_{kh_n}\right] = 3\eta_{kh_n}^2 n^{-2} + \mathcal{O}_{\mathbb{P}}(n^{-2}). \quad (48b)$$

Under Assumption (η -p) with $p = 2p', p' > 2$, it holds that

$$\mathbb{E}\left[\left(\sum_{i=1}^{n-1} \epsilon_i \varphi_{jk} \left(\frac{i+1/2}{n}\right) n^{-1}\right)^{2p'} \middle| \mathcal{F}_{kh_n}\right] \leq K_{p'} \eta_{kh_n}^{p'} n^{-p'} + \mathcal{O}_{\mathbb{P}}(n^{-p'}). \quad (48c)$$

Proof.

$$\begin{aligned} &\mathbb{E}\left[\left(\sum_{i=1}^{n-1} \epsilon_i \varphi_{jk} \left(\frac{i+1/2}{n}\right) n^{-1}\right)^2 \middle| \mathcal{F}_{kh_n}\right] \\ &= \sum_{i=1}^{n-1} \sum_{l=-i}^{n-i-1} \mathbb{E}[\epsilon_i \epsilon_{i+l} | \mathcal{F}_{kh_n}] \frac{\varphi_{jk} \left(\frac{i+1/2}{n}\right) \varphi_{jk} \left(\frac{i+l+1/2}{n}\right)}{n^2} \\ &= (\eta_{kh_n} + \mathcal{O}_{\mathbb{P}}(1)) n^{-1} \sum_{i=1}^{n-1} \varphi_{jk}^2 \left(\frac{i+1/2}{n}\right) n^{-1} + R_n \\ &= \eta_{kh_n} n^{-1} + \mathcal{O}_{\mathbb{P}}(n^{-1}) + R_n. \end{aligned}$$

To control the remainder R_n , we perform a Taylor expansion

$$\varphi_{jk}\left(\frac{i+l+1/2}{n}\right) - \varphi_{jk}\left(\frac{i+1/2}{n}\right) = \sum_{r \geq 1} \frac{\varphi_{jk}^{(r)}\left(\frac{i+1/2}{n}\right)}{r!} \frac{l^r}{n^r},$$

with $\varphi_{jk}^{(r)}$ the existing r th derivative of φ_{jk} . If $|\mathbb{E}[\epsilon_i \epsilon_{i+l}]| \leq |l|^{-1-\varrho}$ for all $i = 0, \dots, n-l$, it follows for any $i = knh_n + 1, \dots, (k+1)nh_n$ that

$$\sum_{l=knh_n+1-i}^{(k+1)nh_n-i} \mathbb{E}[\epsilon_i \epsilon_{i+l} | \mathcal{F}_{kh_n}] \frac{l^r}{n^r h_n^r} = \mathcal{O}\left(\sum_{l=1}^{nh_n} \frac{l^{r-1-\varrho}}{n^r h_n^r}\right) = \mathcal{O}((nh_n)^{-\varrho})$$

which tends to zero and is $\mathcal{O}(n^{-\beta/2})$ when $\varrho > \beta$. Since $\varphi_{jk}^{(r)} \lesssim h_n^{-r} \varphi_{jk}$ and φ_{jk} is zero outside the interval $[kh_n, (k+1)h_n]$ it follows that $R_n = \mathcal{O}_{\mathbb{P}}(n^{-1}(nh_n)^{-\varrho})$.

Considering fourth moments yields

$$\begin{aligned} & \mathbb{E}\left[\left(\sum_{i=1}^{n-1} \epsilon_i \varphi_{jk}\left(\frac{i+1/2}{n}\right) n^{-1}\right)^4 \middle| \mathcal{F}_{kh_n}\right] \\ &= \sum_{i,u,l,v} \frac{\mathbb{E}[\epsilon_i \epsilon_l \epsilon_u \epsilon_v | \mathcal{F}_{kh_n}]}{n^4} \varphi_{jk}\left(\frac{i+1/2}{n}\right) \varphi_{jk}\left(\frac{l+1/2}{n}\right) \varphi_{jk}\left(\frac{u+1/2}{n}\right) \varphi_{jk}\left(\frac{v+1/2}{n}\right) \\ &= 3n^{-4} \sum_{i=1}^{n-1} \sum_{l=-i}^{n-i+1} \mathbb{E}[\epsilon_i \epsilon_{i+l} | \mathcal{F}_{kh_n}] (1 + \mathcal{O}_{\mathbb{P}}(1)) \varphi_{jk}\left(\frac{i+1/2}{n}\right) \varphi_{jk}\left(\frac{i+l+1/2}{n}\right) \\ & \quad \times \sum_{u=1}^{n-1} \sum_{v=-u}^{n-u+1} \mathbb{E}[\epsilon_u \epsilon_{u+v} | \mathcal{F}_{kh_n}] \varphi_{jk}\left(\frac{u+1/2}{n}\right) \varphi_{jk}\left(\frac{u+v+1/2}{n}\right) \\ &= 3n^{-2} \eta_{kh_n}^2 \left(\sum_{u=1}^{n-1} \frac{\varphi_{jk}^2\left(\frac{u+1/2}{n}\right)}{n}\right)^2 + \tilde{R}_n. \end{aligned}$$

The conditional expectation $\mathbb{E}[\epsilon_i \epsilon_l \epsilon_u \epsilon_v | \mathcal{F}_{kh_n}]$ is negligible unless $|i-l|$ and $|u-v|$ are small, or $|i-u|$ and $|l-v|$ are small, or $|i-v|$ and $|u-l|$ are small. In the first identity, we have neglected the sum over terms where all four indices are close which is of the order

$$(1 + \mathcal{O}(1)) \sum_{i=1}^n \frac{\varphi_{jk}^4\left(\frac{i+1/2}{n}\right)}{n} n^{-3} \cdot \mathcal{O}_{\mathbb{P}}(1) = \mathcal{O}_{\mathbb{P}}(h_n^{-1} n^{-3}) = \mathcal{O}_{\mathbb{P}}(n^{-2})$$

given that $\mathbb{E}[\epsilon_i^4 | \mathcal{F}^X] < \infty$ almost surely for all i . That no fourth moments of the noise appear in the leading term is natural, as in standard proofs of central limit theorems using a moment method, since there are only n addends with $i = l = u = v$. That the remainder term \tilde{R}_n is asymptotically negligible follows with the Taylor expansion from above.

Analogously, given that $2p'$ th moments of the noise process exist for some $p' > 2$, an analogous computation yields that

$$\begin{aligned} & \mathbb{E}\left[\left(\sum_{i=1}^{n-1} \epsilon_i \varphi_{jk}\left(\frac{i+1/2}{n}\right) n^{-1}\right)^{2p'} \middle| \mathcal{F}_{kh_n}\right] \\ &= ((2p' - 1) \cdot (2p' - 3) \cdot \dots \cdot 1) \frac{\eta_{kh_n}^{p'}}{n^{p'}} \left(\sum_{u=1}^{n-1} \frac{\varphi_{jk}^2\left(\frac{u+1/2}{n}\right)}{n}\right)^{p'} + \mathcal{O}_{\mathbb{P}}(n^{-p'}). \quad \square \end{aligned}$$

Lemma 2. *On Assumptions 1, $(\sigma-\alpha)$, $(H-r)$ and $(\eta-p)$, we obtain the moment bounds*

$$\mathbb{E}[|\zeta_k(C + \epsilon)|^p | \mathcal{F}_{kh_n}] \leq K_p (\log(n) + \mathcal{O}_{\mathbb{P}}(1)). \tag{49}$$

Proof. First, (12) is a convex combination and applying Jensen’s inequality (for convex combinations) and Young’s inequality, we derive that

$$\begin{aligned} \mathbb{E}[|\zeta_k(C + \epsilon)|^p | \mathcal{F}_{kh_n}] &\leq \sum_{j=1}^{J_n} w_{jk} \mathbb{E}[|S_{jk}^2(C + \epsilon) - \|\Phi_{jk}\|_n^{-2} \frac{\hat{\eta}_{kh_n}}{n}|^p | \mathcal{F}_{kh_n}] \\ &\leq \sum_{j=1}^{J_n} w_{jk} 2^{p-1} \mathbb{E}[|S_{jk}(C + \epsilon)|^{2p} + \|\Phi_{jk}\|_n^{-2} \frac{\hat{\eta}_{kh_n}}{n}|^p | \mathcal{F}_{kh_n}]. \end{aligned}$$

For the second addends, we obtain with (42) and $J_n = \mathcal{O}(\log(n))$ that

$$\begin{aligned} \sum_{j=1}^{J_n} w_{jk} 2^{p-1} \|\Phi_{jk}\|_n^{-2} \frac{\eta_{kh_n}}{n}|^p &\leq K_p \sum_{j=1}^{J_n} w_{jk} \left(\frac{j^2}{nh_n^2}\right)^p \\ &\leq K_p \left(\sum_{j=1}^{J_n} j^{2p} (\log(n))^{-2p}\right) \leq K_p \log(n). \end{aligned}$$

With Proposition 3.1 this bound applies to the conditional expectation with $\hat{\eta}_{kh_n}$ also.

For the term with spectral statistics $S_{jk}(C + \epsilon)$, depending on the process $(C_t)_{t \in [0,1]}$ and the noise, we infer with Young’s inequality and since $\mathbb{E}[\Delta_i^n C \Delta_l^n C] = \mathcal{O}(n^{-2})$ for all $i \neq l$, that

$$\begin{aligned} \mathbb{E}[|S_{jk}(C + \epsilon)|^{2p} | \mathcal{F}_{kh_n}] &\leq 2^{2p-1} \left(\mathbb{E}\left[\left(\|\Phi_{jk}\|_n^{-2} \sum_{i=1}^n (\Delta_i^n C)^2 \Phi_{jk}^2\left(\frac{i}{n}\right)\right)^p \middle| \mathcal{F}_{kh_n}\right]\right. \\ &\quad \left. + \mathbb{E}\left[\left\|\Phi_{jk}\|_n^{-1} \sum_{i=1}^{n-1} \epsilon_i \varphi_{jk}\left(\frac{i+1/2}{n}\right) \frac{1}{n}\right\|^{2p} \middle| \mathcal{F}_{kh_n}\right]\right). \end{aligned}$$

Applying Jensen’s inequality again yields for the first addends

$$\begin{aligned} & \mathbb{E} \left[\left(\|\Phi_{jk}\|_n^{-2} \sum_{i=1}^n (\Delta_i^n C)^2 \Phi_{jk}^2 \left(\frac{i}{n} \right) \right)^p \mid \mathcal{F}_{kh_n} \right] \\ & \leq \|\Phi_{jk}\|_n^{-2} \sum_{i=1}^n \frac{\Phi_{jk}^2 \left(\frac{i}{n} \right)}{n} n^p \mathbb{E} [(\Delta_i^n C)^{2p} \mid \mathcal{F}_{kh_n}] \leq K_p \end{aligned}$$

by (37a). For the noise term, Lemma 1 implies that

$$\begin{aligned} \mathbb{E} \left[\|\Phi_{jk}\|_n^{-1} \sum_{i=1}^{n-1} \epsilon_i \varphi_{jk} \left(\frac{i+1/2}{n} \right) \frac{1}{n} \right]^{2p} \mid \mathcal{F}_{kh_n} & \leq K_p \|\Phi_{jk}\|_n^{-2p} \eta_{kh_n}^p (1 + \mathcal{O}_{\mathbb{P}}(1)) n^{-p} \\ & \leq K_p \left(\frac{j^2}{nh_n^2} \right)^p (1 + \mathcal{O}_{\mathbb{P}}(1)) \leq K_p (1 + \mathcal{O}_{\mathbb{P}}(1)) \end{aligned}$$

for all $j = 1, \dots, J_n = \mathcal{O}(\log(n))$. Inserting the bounds above yields (49). \square

Proof of Theorem 1

The proof is structured in five steps. We establish the marginal stable central limit theorem for the estimator (17a). Since we may consider the continuous martingale part of X time-reversed, the mathematical analysis for the second component follows the same arguments and we restrict ourselves to the right-limit case explicitly. Then, we address the joint convergence in the fifth step of the proof. The Steps 1-4 are structured according to the following decomposition:

$$\begin{aligned} n^{\beta/2} (\hat{\sigma}_s^2 - \sigma_s^2) & = n^{\beta/2} \left(\left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \zeta_k^{ad}(Y) \mathbb{1}_{\{h_n |\zeta_k^{ad}(Y)| \leq u_n\}} \right) - \sigma_s^2 \right) \\ & = n^{\beta/2} \left(\left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \zeta_k(\tilde{C}^n + \epsilon) \right) - \sigma_s^2 \right) \\ & \quad + n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n (\zeta_k(C + \epsilon) - \zeta_k(\tilde{C}^n + \epsilon)) \right) \\ & \quad + n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n (\zeta_k(Y) \mathbb{1}_{\{h_n |\zeta_k(Y)| \leq u_n\}} - \zeta_k(C + \epsilon)) \right) \\ & \quad + n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n (\zeta_k^{ad}(Y) \mathbb{1}_{\{h_n |\zeta_k^{ad}(Y)| \leq u_n\}} - \zeta_k(Y) \mathbb{1}_{\{h_n |\zeta_k(Y)| \leq u_n\}}) \right). \end{aligned}$$

In Step 1, we establish the stable limit theorem for the oracle spectral estimator (14a) built from observations of the process \tilde{C}^n in the simplified model with

noise. Working more generally than under Assumption 2 with Proposition 3.1, just suppose that we have some estimator

$$\mathbb{E}[\hat{\eta}_{kh_n} | \mathcal{F}_{kh_n}] = \frac{1}{nh_n} \sum_{i=knh_n}^{(k+1)nh_n-1} \sum_{u=knh_n-i}^{(k+1)nh_n-1-i} \text{Cov}(\epsilon_i, \epsilon_{i+u} | \mathcal{F}_{kh_n}) + o_{\mathbb{P}}(n^{-\beta/2}), \tag{50}$$

as well as

$$\text{Var}(\hat{\eta}_{kh_n} | \mathcal{F}_{kh_n}) = o_{\mathbb{P}}(n^{-\beta}). \tag{51}$$

Then, on Assumptions 1, (η -p) with $p = 8, \varrho > \beta$, (H-r) with $r < 2$ and (σ - α) and if $0 < \beta < \alpha/(2\alpha + 1)$, as $n \rightarrow \infty$:

Step 1 : $n^{\beta/2} \left(\left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \zeta_k(\tilde{C}^n + \epsilon) \right) - \sigma_s^2 \right) \xrightarrow{(st)} MN(0, 8\sigma_s^3 \eta_s^{1/2}). \tag{52}$

Proof of Step 1: In order to prove a point-wise central limit theorem we verify three conditions: one addressing the conditional bias, one the variance and one Lindeberg-type criterion. Additionally we have to show that the convergence holds *stably* in law.

First, we establish asymptotic unbiasedness of the local estimates (12):

$$\mathbb{E}[\zeta_k(\tilde{C}^n + \epsilon) | \mathcal{F}_{kh_n}] = \sigma_{kh_n}^2 + o_{\mathbb{P}}(n^{-\beta/2}) \text{ for all } k. \tag{53}$$

Using the summation by parts identity (39), we decompose

$$\begin{aligned} \mathbb{E}[\zeta_k(\tilde{C}^n + \epsilon) | \mathcal{F}_{kh_n}] &= \mathbb{E} \left[\sum_{j=1}^{J_n} w_{jk} \left(S_{jk}^2(\tilde{C}^n + \epsilon) - \|\Phi_{jk}\|_n^{-2} \frac{\hat{\eta}_{kh_n}}{n} \right) | \mathcal{F}_{kh_n} \right] \\ &= \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \left(\mathbb{E} \left[\left(\sum_{i=1}^n \Delta_i^n \tilde{C}^n \Phi_{jk} \left(\frac{i}{n} \right) \right)^2 | \mathcal{F}_{kh_n} \right] \right. \\ &\quad - 2 \mathbb{E} \left[\sum_{i=1}^n \Delta_i^n \tilde{C}^n \Phi_{jk} \left(\frac{i}{n} \right) \sum_{l=1}^{n-1} \epsilon_l \varphi_{jk} \left(\frac{l+1/2}{n} \right) \frac{1}{n} | \mathcal{F}_{kh_n} \right] \\ &\quad \left. + \mathbb{E} \left[\left(\sum_{i=1}^{n-1} \epsilon_i \varphi_{jk} \left(\frac{i+1/2}{n} \right) \frac{1}{n} \right)^2 - \frac{\hat{\eta}_{kh_n}}{n} | \mathcal{F}_{kh_n} \right] \right) \end{aligned}$$

and consider the three terms separately. For the first term we obtain with the martingale property that

$$\begin{aligned} &\sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \mathbb{E} \left[\left(\sum_{i=1}^n \Delta_i^n \tilde{C}^n \Phi_{jk} \left(\frac{i}{n} \right) \right)^2 | \mathcal{F}_{kh_n} \right] \\ &= \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \sum_{i=1}^n \frac{\sigma_{kh_n}^2}{n} \Phi_{jk}^2 \left(\frac{i}{n} \right) = \sigma_{kh_n}^2. \end{aligned}$$

For the noise and bias-correction term, we obtain with the bound for the remainder from Lemma 1 and with (50) that

$$\begin{aligned} & \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \mathbb{E} \left[\left(\sum_{i=1}^{n-1} \epsilon_i \varphi_{jk} \left(\frac{i+1/2}{n} \right) \frac{1}{n} \right)^2 - \frac{\hat{\eta}_{kh_n}}{n} | \mathcal{F}_{kh_n} \right] \\ &= \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \sum_{i=knh_n}^{(k+1)nh_n-1} \sum_{l=knh_n-i}^{(k+1)nh_n-i-1} \left(\mathbb{E}[\epsilon_i \epsilon_{i+l} | \mathcal{F}_{kh_n}] n^{-2} \varphi_{jk} \left(\frac{i+1/2}{n} \right) \right. \\ & \quad \left. \times \varphi_{jk} \left(\frac{i+l+1/2}{n} \right) - \frac{\mathbb{E}[\hat{\eta}_{kh_n} | \mathcal{F}_{kh_n}]}{n} \right) \\ &= \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \sum_{i=knh_n}^{(k+1)nh_n-1} \sum_{l=knh_n-i}^{(k+1)nh_n-i-1} \mathbb{E}[\epsilon_i \epsilon_{i+l} | \mathcal{F}_{kh_n}] \\ & \quad \times \left(\frac{\varphi_{jk} \left(\frac{i+1/2}{n} \right) \varphi_{jk} \left(\frac{i+l+1/2}{n} \right)}{n^2} - \frac{1}{n^2 h_n} \right) + \mathcal{O}_{\mathbb{P}}(n^{-\beta/2}) \\ &= \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} n^{-1} (\eta_{kh_n} + \mathcal{O}_{\mathbb{P}}(1)) \left(\sum_{i=knh_n}^{(k+1)nh_n-1} \frac{\varphi_{jk}^2 \left(\frac{i+1/2}{n} \right)}{n} - 1 + \mathcal{O}(n^{-\beta/2}) \right) \\ & \quad + \mathcal{O}_{\mathbb{P}}(n^{-\beta/2}) = \mathcal{O}_{\mathbb{P}}(n^{-\beta/2}), \end{aligned}$$

by (41a) since $\|\Phi_{jk}\|_n^{-2} n^{-1}$ is uniformly bounded. The expectation of cross terms clearly vanishes under independence of noise and (X_t) . Under (23), we derive that

$$\begin{aligned} & \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \mathbb{E} \left[\sum_{i=1}^n \Delta_i^n \tilde{C}^n \Phi_{jk} \left(\frac{i}{n} \right) \sum_{l=1}^{n-1} \epsilon_l \varphi_{jk} \left(\frac{l+1/2}{n} \right) \frac{1}{n} | \mathcal{F}_{kh_n} \right] \\ &= \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \sum_{i=knh_n}^{(k+1)nh_n-1} \sum_{l=knh_n}^i \mathbb{E}[\epsilon_i \Delta_l^n \tilde{C}^n | \mathcal{F}_{kh_n}] \Phi_{jk} \left(\frac{l}{n} \right) \varphi_{jk} \left(\frac{i+1/2}{n} \right) \frac{1}{n} \\ &= \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} (\rho_{kh_n} + \mathcal{O}_{\mathbb{P}}(1)) \sum_{i=knh_n}^{(k+1)nh_n-1} \frac{1}{n} \left(\Phi_{jk} \left(\frac{i}{n} \right) + \mathcal{O}((h_n n)^{-1}) \right) \\ & \quad \times \left(\varphi_{jk} \left(\frac{i}{n} \right) + \mathcal{O}(n^{-1}) \right) \\ &= \mathcal{O}_{\mathbb{P}}((nh_n)^{-1} + h_n^{3/2}) = \mathcal{O}_{\mathbb{P}}(n^{-\beta/2}), \end{aligned}$$

since $\sum_{i=knh_n}^{(k+1)nh_n-1} \Phi_{jk} \left(\frac{i}{n} \right) \varphi_{jk} \left(\frac{i}{n} \right) = 0$ and using the bound $|\int \Phi_{jk}(t) dt| \leq 2\sqrt{2}h_n^{3/2}j^{-2}$, whereas (φ_{jk}) integrate to zero. To put it simply, that the integrals in (41b) vanish for $j = u$ guarantees that the endogenous noise does not induce any non-negligible bias term. This completes the proof of (53).

For the expectation of the left-hand side in (52), we deduce that

$$\begin{aligned} n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \mathbb{E}[\zeta_k(\tilde{C}^n + \epsilon) - \sigma_s^2 | \mathcal{F}_{kh_n}] \right) &= n^{\beta/2} \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n (\sigma_{kh_n}^2 - \sigma_s^2) \\ &= \mathcal{O}_{\mathbb{P}} \left(n^{\beta/2} r_n \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} (kh_n)^\alpha \right) = \mathcal{O}_{\mathbb{P}} \left(n^{\beta/2} (h_n/r_n)^\alpha \right) \\ &= \mathcal{O}_{\mathbb{P}} \left(n^{\beta(\alpha+1/2)} n^{-\alpha/2} \log^\alpha(n) \right) = \mathcal{O}_{\mathbb{P}}(1), \end{aligned}$$

because $\alpha > 0$ and $\beta < \alpha(2\alpha + 1)^{-1}$. By (51) and using that $\|\Phi_{jk}\|_n^{-2} n^{-1}$ is uniformly bounded for all j , we obtain that

$$\begin{aligned} \text{Var} \left(\sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \frac{\hat{\eta}_{kh_n}}{n} | \mathcal{F}_{kh_n} \right) &= \left(\sum_{j=1}^{J_n} w_{jk} n^{-1} \|\Phi_{jk}\|_n^{-2} \right)^2 \text{Var}(\hat{\eta}_{kh_n} | \mathcal{F}_{kh_n}) \\ &= \mathcal{O}_{\mathbb{P}}(n^{-\beta}). \end{aligned}$$

Thus, the estimation of η_{kh_n} in the bias-correction is negligible in the variance of $\hat{\sigma}_s^2$. In case of exogenous noise, with Lemma 1, we can readily adopt the identity

$$\text{Var} \left(\zeta_k(\tilde{C}^n + \epsilon) | \mathcal{F}_{kh_n} \right) = \sum_{j=1}^{J_n} w_{jk}^2 I_{jk}^{-1} = I_k^{-1} \quad (54)$$

from Section 6.2.2 of [5] with I_k, I_{jk} from (13). We consider additionally the conditional variance terms due to endogenous noise under condition (23). With similar estimates for the remainders as in the bias term above, we obtain that

$$\begin{aligned} &\text{Cov} \left(\left(\sum_{i=1}^n \Delta_i^n \tilde{C}^n \Phi_{jk} \left(\frac{i}{n} \right) \right)^2, \left(\sum_{l=1}^{n-1} \epsilon_l \varphi_{uk} \left(\frac{l+1/2}{n} \right) \frac{1}{n} \right)^2 | \mathcal{F}_{kh_n} \right) \\ &= \sum_{i,p=1}^n \sum_{l,q=1}^{n-1} \left(\mathbb{E}[\Delta_i^n \tilde{C}^n \Delta_p^n \tilde{C}^n \epsilon_l \epsilon_q | \mathcal{F}_{kh_n}] - \mathbb{E}[\Delta_i^n \tilde{C}^n \Delta_p^n \tilde{C}^n | \mathcal{F}_{kh_n}] \mathbb{E}[\epsilon_l \epsilon_q | \mathcal{F}_{kh_n}] \right) \\ &\quad \times n^{-2} \Phi_{jk} \left(\frac{i}{n} \right) \Phi_{jk} \left(\frac{p}{n} \right) \varphi_{uk} \left(\frac{l+1/2}{n} \right) \varphi_{uk} \left(\frac{q+1/2}{n} \right) \\ &= 2 \sum_{l=knh_n}^{(k+1)nh_n-1} \sum_{i=knh_n}^l \mathbb{E}[\epsilon_l \Delta_i^n \tilde{C}^n | \mathcal{F}_{kh_n}] \sum_{q=knh_n}^{(k+1)nh_n-1} \sum_{p=knh_n}^q \mathbb{E}[\epsilon_q \Delta_p^n \tilde{C}^n | \mathcal{F}_{kh_n}] \\ &\quad \times n^{-2} \Phi_{jk} \left(\frac{i}{n} \right) \Phi_{jk} \left(\frac{p}{n} \right) \varphi_{uk} \left(\frac{l+1/2}{n} \right) \varphi_{uk} \left(\frac{q+1/2}{n} \right) (1 + \mathcal{O}_{\mathbb{P}}(1)) \\ &= 2 \sum_{l=knh_n}^{(k+1)nh_n-1} (\rho_{kh_n} + \mathcal{O}_{\mathbb{P}}(1)) n^{-1} \Phi_{jk} \left(\frac{l}{n} \right) \varphi_{uk} \left(\frac{l}{n} \right) \\ &\quad \times \sum_{q=knh_n}^{(k+1)nh_n-1} (\rho_{kh_n} + \mathcal{O}_{\mathbb{P}}(1)) n^{-1} \Phi_{jk} \left(\frac{q}{n} \right) \varphi_{uk} \left(\frac{q}{n} \right) + \mathcal{O}_{\mathbb{P}}(h_n^{3/2} n^{-1}) \end{aligned}$$

$$= 2\rho_{kh_n}^2 \left(\int_0^1 \tilde{\Phi}_{jk}(t)\varphi_{uk}(t) dt \right)^2 (1 + \mathcal{O}_{\mathbb{P}}(1)) + \mathcal{O}_{\mathbb{P}}(h_n^{3/2}n^{-1}).$$

In the first identity the terms for $i = p$ and i not close to l, q cancel. We used the smoothness of (Φ_{jk}) and (φ_{jk}) again. Analogously, we obtain that

$$\begin{aligned} & \text{Cov}\left(\left(\sum_{l=1}^{n-1} \epsilon_l \varphi_{jk}\left(\frac{l+1/2}{n}\right)\frac{1}{n}\right)^2, \left(\sum_{i=1}^n \Delta_i^n \tilde{C}^m \Phi_{uk}\left(\frac{i}{n}\right)\right)^2 \middle| \mathcal{F}_{kh_n}\right) \\ &= 2\rho_{kh_n}^2 \left(\int_0^1 \tilde{\Phi}_{uk}(t)\varphi_{jk}(t) dt \right)^2 (1 + \mathcal{O}_{\mathbb{P}}(1)) + \mathcal{O}_{\mathbb{P}}(h_n^{3/2}n^{-1}). \end{aligned}$$

With similar computations, we obtain that

$$\begin{aligned} & \text{Cov}\left(\sum_{i=1}^n \Delta_i^n \tilde{C}^m \Phi_{jk}\left(\frac{i}{n}\right) \sum_{l=1}^{n-1} \epsilon_l \varphi_{jk}\left(\frac{l+1/2}{n}\right)\frac{1}{n}, \left(\sum_{p=1}^n \Delta_p^n \tilde{C}^m \Phi_{uk}\left(\frac{p}{n}\right)\right)^2 \middle| \mathcal{F}_{kh_n}\right) \\ &= 2\rho_{kh_n} \sigma_{kh_n}^2 \left(\int_0^1 \tilde{\Phi}_{jk}(t)\tilde{\Phi}_{uk}(t) dt \int_0^1 \tilde{\Phi}_{uk}(t)\varphi_{jk}(t) dt \right) (1 + \mathcal{O}_{\mathbb{P}}(1)) + \mathcal{O}_{\mathbb{P}}(h_n^{3/2}n^{-1}) \\ &= \mathcal{O}_{\mathbb{P}}(h_n^{3/2}n^{-1}), \end{aligned}$$

since $\int_0^1 \tilde{\Phi}_{jk}(t)\tilde{\Phi}_{uk}(t) dt \int_0^1 \tilde{\Phi}_{uk}(t)\varphi_{jk}(t) dt = 0$ for all j, u . Analogously, since $\int_0^1 \varphi_{jk}(t)\varphi_{uk}(t) dt \int_0^1 \tilde{\Phi}_{jk}(t)\varphi_{uk}(t) dt = 0$ for all j, u , the conditional covariance of cross terms and noise terms is of the same order $h_n^{3/2}n^{-1}$ in probability. The only other (at first) non-negligible additional conditional variance term thus comes from

$$\begin{aligned} & \text{Cov}\left(\sum_{i=1}^n \Delta_i^n \tilde{C}^m \Phi_{jk}\left(\frac{i}{n}\right) \sum_{l=1}^{n-1} \epsilon_l \varphi_{jk}\left(\frac{l+1/2}{n}\right)\frac{1}{n}, \right. \\ & \left. \sum_{p=1}^n \Delta_p^n \tilde{C}^m \Phi_{uk}\left(\frac{p}{n}\right) \sum_{q=1}^{n-1} \epsilon_q \varphi_{uk}\left(\frac{q+1/2}{n}\right)\frac{1}{n} \middle| \mathcal{F}_{kh_n}\right). \end{aligned}$$

Using the same approximations as in the previous terms and subtracting the term already contained in I_{jk}^{-1} from the exogenous setup, we obtain the overall *additional* conditional variance

$$\begin{aligned} & \sum_{j,u=1}^{J_n} w_{jk} w_{uk} \|\Phi_{jk}\|_n^{-2} \|\Phi_{uk}\|_n^{-2} \left(\text{Cov}\left(\left(\sum_{i=1}^n \Delta_i^n \tilde{C}^m \Phi_{jk}\left(\frac{i}{n}\right)\right)^2, \right. \right. \\ & \left. \left. \left(\sum_{l=1}^{n-1} \epsilon_l \varphi_{uk}\left(\frac{l+1/2}{n}\right)\frac{1}{n}\right)^2 \middle| \mathcal{F}_{kh_n}\right) \right) \\ & + \text{Cov}\left(\left(\sum_{l=1}^{n-1} \epsilon_l \varphi_{jk}\left(\frac{l+1/2}{n}\right)\frac{1}{n}\right)^2, \left(\sum_{i=1}^n \Delta_i^n \tilde{C}^m \Phi_{uk}\left(\frac{i}{n}\right)\right)^2 \middle| \mathcal{F}_{kh_n}\right) \end{aligned}$$

$$\begin{aligned}
& + 4 \operatorname{Cov} \left(\sum_{i=1}^n \Delta_i^n \tilde{C}^n \Phi_{jk} \left(\frac{i}{n} \right) \sum_{l=1}^{n-1} \epsilon_l \varphi_{jk} \left(\frac{l+1/2}{n} \right) \frac{1}{n}, \right. \\
& \quad \left. \sum_{p=1}^n \Delta_p^n \tilde{C}^n \Phi_{uk} \left(\frac{p}{n} \right) \sum_{q=1}^{n-1} \epsilon_q \varphi_{uk} \left(\frac{q+1/2}{n} \right) \frac{1}{n} \middle| \mathcal{F}_{kh_n} \right) \\
& - \sum_{j=1}^{J_n} w_{jk}^2 4 \|\Phi_{jk}\|_n^{-2} \sigma_{kh_n}^2 \frac{\eta kh_n}{n} \\
& = \sum_{j,u=1}^{J_n} w_{jk} w_{uk} \|\Phi_{jk}\|_n^{-2} \|\Phi_{uk}\|_n^{-2} \rho_{kh_n}^2 (1 + o_{\mathbb{P}}(1)) \left(\mathcal{O}_{\mathbb{P}}(h_n^{3/2} n^{-1}) \right. \\
& \quad + 2 \left(\int \tilde{\Phi}_{jk}(t) \varphi_{uk}(t) dt \right)^2 + 2 \left(\int \tilde{\Phi}_{uk}(t) \varphi_{jk}(t) dt \right)^2 \\
& \quad \left. + 4 \int \tilde{\Phi}_{jk}(t) \varphi_{uk}(t) dt \int \tilde{\Phi}_{uk}(t) \varphi_{jk}(t) dt \right).
\end{aligned}$$

However, by (41b) the integrals sum up to zero. Since the remainder is $\mathcal{O}_{\mathbb{P}}((\log(n))^3 n^{-1/4})$, the effect of the endogenous noise becomes negligible at first asymptotic order. We conclude (54).

In the sequel we write w_{jk}, I_{jk}, I_k as functions of the squared volatility and η : $I_j(\sigma^2, \eta) = \frac{1}{2}(\sigma^2 + \|\Phi_{jk}\|_n^{-2} \frac{\eta}{n})^{-2}$, $I(\sigma^2, \eta) = \sum_{j=1}^{J_n} I_j(\sigma^2, \eta)$ and $w_j(\sigma^2, \eta) = (I(\sigma^2, \eta))^{-1} I_j(\sigma^2, \eta)$. Note that $\|\Phi_{jk}\|_n^{-2}$ is equal for all k such that the time-dependence of I, I_j, w_j is only in the squared volatility σ^2 and η . For the sum of conditional variances of the left-hand side of (52), we obtain that

$$\begin{aligned}
n^\beta \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n^2 \operatorname{Var}(\zeta_k(\tilde{C}^n + \epsilon) | \mathcal{F}_{kh_n}) &= n^\beta r_n \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} \sum_{j=1}^{J_n} r_n w_{jk}^2 I_{jk}^{-1} + o_{\mathbb{P}}(1) \right) \\
&= \log(n) I^{-1} \left(\sigma_{\lfloor sh_n^{-1} \rfloor h_n}^2, \eta_{\lfloor sh_n^{-1} \rfloor h_n} \right) + R_n.
\end{aligned}$$

We exploit bounds on the derivative of the weights with respect to σ^2 and η

$$\frac{\partial w_j(\sigma^2, \eta)}{\partial \sigma^2} = \mathcal{O}(w_j(\sigma^2, \eta) \log^2(n)), \quad (55)$$

here and several times below. The bound is proved as Equation (77) in [5]. $\partial w_j(\sigma^2, \eta) / (\partial \eta)$ can be bounded analogously. Observe that by the chain and product differentiation rule

$$\begin{aligned}
\frac{\partial}{\partial \sigma^2} (w_j^2(\sigma^2, \eta) (I_j(\sigma^2, \eta))^{-1}) &= 2w_j(\sigma^2, \eta) \frac{\partial w_j(\sigma^2, \eta)}{\partial \sigma^2} (I_j(\sigma^2, \eta))^{-1} \\
&\quad + w_j^2(\sigma^2, \eta) 4(\sigma^2 + \|\Phi_{jk}\|_n^{-2} \frac{\eta}{n}).
\end{aligned}$$

Thus, we can find an upper bound for the remainder R_n using

$$\begin{aligned} \sum_{j=1}^{J_n} (1 \vee \|\Phi_{jk}\|_n^{-2} n^{-1}) (1 \wedge \|\Phi_{jk}\|_n^8 n^4) &= \mathcal{O}\left(\sum_{j=1}^{\lfloor \sqrt{n}h_n \rfloor} 1 + \sum_{j=1}^{J_n} \|\Phi_{jk}\|_n^6 n^3\right) \\ &= \mathcal{O}(\log^6(n)) \\ \Rightarrow R_n &= \mathcal{O}_{\mathbb{P}}\left(n^\beta r_n^2 \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} \log^6(n) (\sigma_{kh_n}^2 - \sigma_{\lfloor sh_n^{-1} \rfloor h_n}^2)\right) \\ &= \mathcal{O}_{\mathbb{P}}\left(\log^7(n) (h_n/r_n)^\alpha\right) \end{aligned}$$

with (42), which tends to zero as $n \rightarrow \infty$ because $\alpha > 0$. By (22), the locally constant approximation of the long-run noise variance induces an error of smaller or at most equal order.

The Lindeberg condition is proved by the stronger Lyapunov criterion considering fourth moments:

$$n^{2\beta} \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n^4 \mathbb{E} \left[\zeta_k^4 (\tilde{C}^n + \epsilon) | \mathcal{F}_{kh_n} \right] = \mathcal{O}_{\mathbb{P}}\left(n^{-\beta} \log(n)\right) = \mathcal{O}_{\mathbb{P}}(1),$$

using Lemma 2 (replacing C by \tilde{C}^n , the proof of Lemma 2 applies in the same way). We obtain the variance in (52), since the bin-wise Fisher informations

$$I_k = \frac{1}{2} \sum_{j=1}^{J_n} \left(\sigma_{kh_n}^2 + \|\Phi_{jk}\|_n^{-2} \frac{\eta_{kh_n}}{n} \right)^{-2}$$

satisfy the following convergences (see Section 6.2.2 of [5]):

$$\frac{1}{\log(n)} I_k \longrightarrow \int_0^\infty \frac{1}{2} \left(\sigma_{kh_n}^2 + \eta_{kh_n} \pi^2 x^2 \right)^{-2} dx = \left(8 \sigma_s^3 \eta_s^{1/2} \right)^{-1}, \quad (56)$$

and the reciprocal of the right-hand side thus constitutes the asymptotic variance of $\hat{\sigma}_s^2$.

Finally, stability of the weak convergence is proved similarly as in Proposition 8.2 of [29]. For later use, let us directly consider a collection of times where we consider estimates of the spot volatilities instead of only one fixed time. In particular, for our test, we shall focus on finitely many jumps of X with absolute value larger than some constant. Consider a finite set $(S_p)_{1 \leq p \leq P}$ with fix $P < \infty$ of ordered stopping times exhausting those jump arrivals of X on $[0, 1]$. The restriction of Ω to

$$\Omega_n = \{ \omega \in \Omega | S_1 > r_n^{-1} h_n, S_P < 1 - r_n^{-1} h_n, \forall p : (S_p - S_{p-1}) > 2r_n^{-1} h_n \} \quad (57)$$

satisfies $\mathbb{P}(\Omega_n) \rightarrow 1$ as $n \rightarrow \infty$. Thus, we work on Ω_n . We aim at establishing for

$$\alpha_n = n^{\beta/2} \left(\sum_{k=\lfloor S_p h_n^{-1} \rfloor + 1}^{\lfloor S_p h_n^{-1} \rfloor + r_n^{-1}} r_n \zeta_k(\tilde{C}^n + \epsilon) - \sigma_{S_p}^2, \right. \tag{58}$$

$$\left. \sum_{k=\lfloor S_p h_n^{-1} \rfloor - r_n^{-1}}^{\lfloor S_p h_n^{-1} \rfloor - 1} r_n \zeta_k(\tilde{C}^n + \epsilon) - \sigma_{S_p^-}^2 \right)_{1 \leq p \leq P}$$

that $\mathbb{E}[Zg(\alpha_n)] \rightarrow \mathbb{E}[Zg(\alpha)]$ with $\alpha = (2\sqrt{2}\sigma_{S_p}^{3/2}\eta_{S_p}^{1/4}U_p, 2\sqrt{2}\sigma_{S_p^-}^{3/2}\eta_{S_p^-}^{1/4}U'_p)_{1 \leq p \leq P}$ for any \mathcal{F} -measurable bounded random variable Z and continuous bounded function g and for (U_p, U'_p) a sequence of standard normals defined on an exogenous space being independent of \mathcal{F} . This is the definition of the claimed \mathcal{F} -stable convergence.

The strategy is to exclude intervals on which the spot estimators are built and conditioning. Thereto, define

$$B_n = \bigcup_{p=1}^P [(S_p - (r_n^{-1} + 1)h_n) \vee 0, (S_p + (r_n^{-1} + 1)h_n) \wedge 1]$$

and $\tilde{\mathcal{G}}_t^n$ as the smallest filtration to which \tilde{C}^n and U are adapted and such that the σ -field generated by the Poisson measure which determines S_1, \dots, S_P lies in $\tilde{\mathcal{G}}_0^n$. Then each α_n is $\tilde{\mathcal{G}}_1^n$ -measurable. The following decomposition of \tilde{C}^n is well-defined:

$$\tilde{X}(n)_t = \int_0^t \mathbb{1}_{B_n}(s) \sigma_{\lfloor sh_n^{-1} \rfloor h_n} dW_s, \quad \bar{X}(n)_t = \tilde{C}_t^n - \tilde{X}(n)_t,$$

and analogously (\tilde{U}_t) and (\bar{U}_t) . It is enough to consider Z being $\tilde{\mathcal{G}}_1^n$ -measurable, as we can simply substitute with $\mathbb{E}[Z|\tilde{\mathcal{G}}_1^n]$ otherwise. When \mathcal{H}_n is the σ -field generated by $\tilde{\mathcal{G}}_0^n, \bar{X}(n)_t$ and \bar{U}_t , $(\mathcal{H}_n)_n$ is an isotonic sequence and $\bigvee_n \mathcal{H}_n = \tilde{\mathcal{G}}_1^n$. Since $\mathbb{E}[Z|\mathcal{H}_n] \rightarrow Z$ in $L^1(\mathbb{P})$, it is enough to show

$$\mathbb{E}[Z\mathbb{1}_{\Omega_n}g(\alpha_n)] \rightarrow \mathbb{E}[Zg(\alpha)] = \mathbb{E}[Z]\mathbb{E}[g(\alpha)] \tag{59}$$

for Z \mathcal{H}_q -measurable for some q . We can use the approximation with constant \mathcal{H}_q -measurable squared volatilities $\sigma_{S_p}^2, \sigma_{S_p^-}^2$ and with η_{S_p} locally constant on the single intervals of B_n , where the errors have been bounded above. Restricted to Ω_n the vector α_n then includes only Brownian increments $\Delta_i^n W$ independent of the Brownian increments of $\bar{X}(n)_t$. Further, the noise is under Assumption $(\eta-p)$ only short-term dependent on the past and in particular covariances of any such Z and α_n tend to zero. Then for all $n \geq q$, conditional on \mathcal{H}_q , the vector α_n has a law asymptotically independent of $\bar{X}(n)_t$ and \bar{U}_t , such that the ordinary central limit theorem implies the claimed convergence. The above proof includes the stable convergence of the spot volatility estimator at one fixed time

$s \in (0, 1)$ as a special case. Thus, we have verified all conditions and infer the stable limit theorem (52). \square

To prove that the same limit theorem as (52) is valid for $n^{\beta/2}(\hat{\sigma}_s^2 - \sigma_s^2)$, we show for the other addends above that they converge to zero in probability for all $s \in (0, 1)$. We proceed with

$$\text{Step 2: } n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \left(\zeta_k(C + \epsilon) - \zeta_k(\tilde{C}^n + \epsilon) \right) \right) = o_{\mathbb{P}}(1), \quad (60)$$

under the same conditions as in Step 1. This remainder due to approximating C by the simplified processes \tilde{C}^n has exactly the same structure as the one for integrated squared volatility estimation examined in paragraph 6.3 of [5]. We just incorporate the additional jump component in the volatility using an estimate as (37e). Then, repeating the proof along the same lines, only changing the mean over all bins to the mean over local windows of size $r_n^{-1}h_n$, renders with $\beta < 1/2$ the order:

$$\zeta_k(C + \epsilon) - \zeta_k(\tilde{C}^n + \epsilon) = \mathcal{O}_{\mathbb{P}}(h_n^\alpha) = o_{\mathbb{P}}(n^{-\beta/2}),$$

uniformly for all k . Analogously to [5], we require here the mild condition (18).

$$\text{Step 3: } n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n (\zeta_k(Y) \mathbb{1}_{\{h_n|\zeta_k(Y)| \leq u_n\}} - \zeta_k(C + \epsilon)) \right) = o_{\mathbb{P}}(1), \quad (61)$$

when, additional to the assumptions for Steps 1 and 2, we have $\beta < \tau(1 - r/2)$ and $\tau < 1 - \beta/(p - 2)$ when $p < \infty$ moments of the noise exist in Assumption (η -p).

Proof of Step 3: This part of the proof is related to Chapter 13 of [28] and the proofs in [16]. Our strategy here is related, but slightly different. We differentiate three cases. For some fixed $\rho \in (0, 1)$, for instance $\rho = 1/2$, consider

$$\begin{aligned} & n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n (\zeta_k(Y) \mathbb{1}_{\{h_n|\zeta_k(Y)| \leq u_n\}} - \zeta_k(C + \epsilon)) \right) = \\ & n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \mathbb{1}_{\{h_n|\zeta_k(C + \epsilon)| > \rho u_n\}} (\zeta_k(Y) \mathbb{1}_{\{h_n|\zeta_k(Y)| \leq u_n\}} - \zeta_k(C + \epsilon)) \right) \\ & - n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \mathbb{1}_{\{h_n|\zeta_k(C + \epsilon)| \leq \rho u_n\}} \mathbb{1}_{\{h_n|\zeta_k(Y)| > u_n\}} \zeta_k(C + \epsilon) \right) \end{aligned}$$

$$+ n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \mathbb{1}_{\{h_n|\zeta_k(C+\epsilon)| \leq \rho u_n\}} \mathbb{1}_{\{h_n|\zeta_k(Y)| \leq u_n\}} (\zeta_k(Y) - \zeta_k(C + \epsilon)) \right).$$

We prove that all three sums tend to zero in probability. For the first term, when $h_n|\zeta_k(C + \epsilon)| > \rho u_n = c\rho h_n^\tau$, it suffices to prove that uniformly for all k :

$$|\zeta_k(Y)\mathbb{1}_{\{h_n|\zeta_k(Y)| \leq u_n\}} - \zeta_k(C + \epsilon)| = \mathcal{O}_{\mathbb{P}}(n^{-\beta/2}).$$

We can choose $N_0 \in \mathbb{N}$, such that $h_n^{N_0(1-\tau)} = \mathcal{O}(n^{-\beta/2-\varepsilon})$ for some $\varepsilon > 0$. Given that $h_n|\zeta_k(C + \epsilon)| > \rho u_n$, when we have enough moments of the noise such that $\tau < 1 - \beta/(p - 2)$, we conclude with Lemma 2 that

$$\begin{aligned} |\zeta_k(Y)\mathbb{1}_{\{h_n|\zeta_k(Y)| \leq u_n\}} - \zeta_k(C + \epsilon)| &\leq (h_n^{-1}u_n + |\zeta_k(C + \epsilon)|) \left| \frac{h_n\zeta_k(C + \epsilon)}{\rho u_n} \right|^{N_0+1} \\ &\leq (|\zeta_k(C + \epsilon)|^{N_0+1} + |\zeta_k(C + \epsilon)|^{N_0+2}h_n^{1-\tau}(c\rho)^{-1})(c\rho)^{-N_0}h_n^{N_0(1-\tau)} \\ &= \mathcal{O}_{\mathbb{P}}\left(\log(n)h_n^{N_0(1-\tau)}\right) = \mathcal{O}_{\mathbb{P}}(n^{-\beta/2}). \end{aligned}$$

This shows that the first sum above tends to zero in probability. Next, we prove that

$$\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \mathbb{1}_{\{h_n|\zeta_k(C+\epsilon)| \leq \rho u_n\}} \mathbb{1}_{\{h_n|\zeta_k(Y)| > u_n\}} \zeta_k(C + \epsilon) = \mathcal{O}_{\mathbb{P}}(n^{-\beta/2}). \tag{62}$$

We have the decomposition

$$\begin{aligned} \zeta_k(Y) &= \zeta_k(C + \epsilon) + \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \left(\left(\sum_{i=1}^n \Delta_i^n J\Phi_{jk}\left(\frac{i}{n}\right) \right)^2 \right. \\ &\quad \left. + 2 \sum_{i=1}^n \Delta_i^n J\Phi_{jk}\left(\frac{i}{n}\right) \sum_{v=1}^n \Delta_v^n C\Phi_{jk}\left(\frac{v}{n}\right) \right), \end{aligned}$$

neglecting cross terms of jumps and noise. All cross terms can be bounded using the Cauchy-Schwarz inequality. Observe that

$$\begin{aligned} \mathbb{1}_{\{h_n|\zeta_k(C+\epsilon)| \leq \rho u_n\}} \cap \{h_n|\zeta_k(Y)| > u_n\} &\leq \\ &\mathbb{1}_{\left\{h_n \left| \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \left(\sum_{i=1}^n \Delta_i^n J\Phi_{jk}(i/n) \right)^2 \right| > \tilde{\rho} u_n \right\}}, \end{aligned}$$

for some fix $\tilde{\rho} \in (0, 1)$ depending on ρ . This means that if the terms from the continuous part are not exceptionally large, the jumps need to be sufficiently large such that $h_n|\zeta_k(Y)| > u_n$. The simple uniform bound $\Phi_{jk}(t) \leq \sqrt{2}h_n^{-1/2}\|\Phi_{jk}\|_n$ yields that

$$h_n \left| \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \left(\sum_{i=1}^n \Delta_i^n J\Phi_{jk}\left(\frac{i}{n}\right) \right)^2 \right| \leq 2 \left(\sum_{i=n\lfloor kh_n \rfloor + 1}^{n\lfloor (k+1)h_n \rfloor} \Delta_i^n J \right)^2$$

and we obtain that

$$\mathbb{1}_{\{h_n|\zeta_k(C+\epsilon)|\leq\rho u_n\}}\mathbb{1}_{\{h_n|\zeta_k(Y)|>u_n\}}\leq\mathbb{1}_{\{|J_{(k+1)h_n}-J_{kh_n}|>\rho^*\sqrt{u_n}\}},$$

with $\rho^* = \bar{\rho}/\sqrt{2}$. Therefore, it is sufficient to prove that

$$\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} \mathbb{1}_{\{|J_{(k+1)h_n}-J_{kh_n}|>\rho^*\sqrt{u_n}\}} = \mathcal{O}_{\mathbb{P}}(n^{\beta/2}). \tag{63}$$

Similar terms have been addressed several times in the literature, see, for instance, (13.1.14) in [28]. Applying (37d) with $\epsilon = \rho^* u_n^{1/2}$, we derive the condition

$$r_n^{-1}h_n u_n^{-r/2} = \mathcal{O}(n^{\beta/2}) \Leftrightarrow 1 - \frac{\tau r}{2} > \beta, \tag{64}$$

to ensure (63). When $h_n|\zeta_k(C + \epsilon)| \leq \rho u_n$ and $h_n|\zeta_k(Y)| \leq u_n$, it follows that

$$h_n \left| \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \left(\sum_{i=1}^n \Delta_i^n J \Phi_{jk}(i/n) \right)^2 \right| \leq c u_n$$

with some constant $c < 4$. In this case, we obtain by (37b):

$$\left(|J_{(k+1)h_n} - J_{kh_n}| \wedge \sqrt{c u_n} \right) = \mathcal{O}_{\mathbb{P}}(h_n^{1/2} u_n^{1-r/2}),$$

and hence, if we can ensure that $h_n^{\tau(1-r/2)} = \mathcal{O}(n^{-\beta/2})$, using again $\Phi_{jk}(t) \leq \sqrt{2} h_n^{-1/2} \|\Phi_{jk}\|_n$,

$$\begin{aligned} |\zeta_k(Y) - \zeta_k(C + \epsilon)| &\leq c \left(\left| \sum_{j=1}^{J_n} w_{jk} \|\Phi_{jk}\|_n^{-2} \sum_{i=1}^n (\Delta_i^n J \wedge \sqrt{u_n})^2 \Phi_{jk}^2\left(\frac{i}{n}\right) \right| \wedge u_n \right) \\ &\leq 2c h_n^{-1} \left(\left(\sum_{i=n\lfloor kh_n \rfloor + 1}^{n\lfloor (k+1)h_n \rfloor} \Delta_i^n J \right)^2 \wedge u_n \right) \\ &= \mathcal{O}_{\mathbb{P}}(u_n^{1-r/2}) = \mathcal{O}_{\mathbb{P}}(n^{-\beta/2}), \end{aligned}$$

on the set where $\{h_n|\zeta_k(C + \epsilon)| \leq \rho u_n, h_n|\zeta_k(Y)| \leq u_n\}$. The condition $\beta < \tau(1 - r/2)$ implies (64) and is exactly what we need to complete the proof of (61).

Step 4 : $n^{\beta/2} \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n(\zeta_k^{ad}(Y)\mathbb{1}_{\{h_n|\zeta_k^{ad}(Y)|\leq u_n\}} - \zeta_k(Y)\mathbb{1}_{\{h_n|\zeta_p(Y)|\leq u_n\}}) = \mathcal{O}_{\mathbb{P}}(1).$

Proof of Step 4: In Step 3 we have not used the specific form of the oracle weights (13) and the proof analogously extends to

$$n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n(\zeta_k^{ad}(Y)\mathbb{1}_{\{h_n|\zeta_k^{ad}(Y)|\leq u_n\}} - \zeta_k^{ad}(C + \epsilon)) \right) = \mathcal{O}_{\mathbb{P}}(1). \tag{65}$$

Thus it suffices to prove that

$$n^{\beta/2} \left(\sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n (\zeta_k^{ad}(C + \epsilon) - \zeta_k(C + \epsilon)) \right) = \mathcal{O}_{\mathbb{P}}(1). \quad (66)$$

We decompose this remainder as follows. Since both, oracle weights $w_j(\sigma_{kh_n}^2, \eta_{kh_n})$ and estimated weights $w_j(\hat{\sigma}_{kh_n}^2, \hat{\eta}_{kh_n})$ sum up to one, we can replace $(S_{jk}^2 - \|\Phi_{jk}\|_n^{-2} \hat{\eta}_{kh_n}/n)$ by $(S_{jk}^2 - \mathbb{E}[S_{jk}^2])$. First, consider the difference of pre-estimated and oracle weights, when the pilot estimator is the same for the whole window. When $\max(\hat{\eta}_{\lfloor sh_n^{-1} \rfloor h_n} - \eta_{\lfloor sh_n^{-1} \rfloor h_n}, \hat{\sigma}_{\lfloor sh_n^{-1} \rfloor h_n}^{2,pilot} - \sigma_{\lfloor sh_n^{-1} \rfloor h_n}^2) = \mathcal{O}_{\mathbb{P}}(\delta_n)$ with $\delta_n \rightarrow 0$ as $n \rightarrow \infty$, we derive that

$$\begin{aligned} & \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \sum_{j=1}^{J_n} \left(w_j(\hat{\sigma}_{\lfloor sh_n^{-1} \rfloor h_n}^{2,pilot}, \hat{\eta}_{\lfloor sh_n^{-1} \rfloor h_n}) - w_j(\sigma_{\lfloor sh_n^{-1} \rfloor h_n}^2, \eta_{\lfloor sh_n^{-1} \rfloor h_n}) \right) \\ & \quad \times (S_{jk}^2 - \mathbb{E}[S_{jk}^2]) \\ &= r_n \sum_{j=1}^{J_n} \left(w_j(\hat{\sigma}_{\lfloor sh_n^{-1} \rfloor h_n}^{2,pilot}, \hat{\eta}_{\lfloor sh_n^{-1} \rfloor h_n}) - w_j(\sigma_{\lfloor sh_n^{-1} \rfloor h_n}^2, \eta_{\lfloor sh_n^{-1} \rfloor h_n}) \right) \\ & \quad \times \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} (S_{jk}^2 - \mathbb{E}[S_{jk}^2]) \\ &= \mathcal{O}_{\mathbb{P}} \left(r_n^{1/2} \sum_{j=1}^{J_n} (1 + \|\Phi_{jk}\|_n^{-2} n^{-1}) w_j(\sigma_{\lfloor sh_n^{-1} \rfloor h_n}^2, \eta_{\lfloor sh_n^{-1} \rfloor h_n}) \log(n) \delta_n \right) \\ &= \mathcal{O}_{\mathbb{P}}(n^{-\beta/2}). \end{aligned}$$

We have used that the expectation is zero and that the weights do not hinge on k . Then, we can bound the variance using the derivative bound (55). Covariances of the S_{jk}^2 over different bins for $k \neq k'$ are negligible what is shown in Step 5 of the proof. Finally, since $r_n^{1/2} = n^{-\beta/2} \sqrt{\log(n)}$ some $\delta_n < n^{-\epsilon}$ for any $\epsilon > 0$ is enough here, while we actually attain $\delta_n = n^{-\beta/2}$. It remains to bound

$$\begin{aligned} & \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n^2 \text{Var} \left(\sum_{j=1}^{J_n} \left(w_j(\hat{\sigma}_{kh_n}^{2,pilot}, \hat{\eta}_{kh_n}) - w_j(\hat{\sigma}_{\lfloor sh_n^{-1} \rfloor h_n}^{2,pilot}, \hat{\eta}_{\lfloor sh_n^{-1} \rfloor h_n}) \right) \right. \\ & \quad \left. \times (S_{jk}^2 - \mathbb{E}[S_{jk}^2]) \right) \\ &+ \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n^2 \text{Var} \left(\sum_{j=1}^{J_n} \left(w_j(\sigma_{kh_n}^2, \eta_{kh_n}) - w_j(\sigma_{\lfloor sh_n^{-1} \rfloor h_n}^2, \eta_{\lfloor sh_n^{-1} \rfloor h_n}) \right) \right. \\ & \quad \left. \times (S_{jk}^2 - \mathbb{E}[S_{jk}^2]) \right) \end{aligned}$$

$$= \mathcal{O}\left(r_n \log^5(n)(n^{-\beta} \vee (r_n^{-1}h_n)^{2\alpha})\right) = \mathcal{O}(n^{-\beta}).$$

This proves (66).

Step 5 :
$$n^\beta \sum_{\substack{k,k'=\lfloor sh_n^{-1} \rfloor + 1, \\ k \neq k'}}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n^2 \text{Cov}(\zeta_k(\tilde{C}^n + \epsilon), \zeta_{k'}(\tilde{C}^n + \epsilon)) = \mathcal{O}(1).$$

Moreover, it holds that $\text{Cov}(\hat{\sigma}_s^2, \hat{\sigma}_{s-}^2) = \mathcal{O}(n^{-\beta})$.

Proof of Step 5: Covariances of S_{jk}^2 and $S_{uk'}^2$ for different bins $k \neq k'$ are only due to the noise parts and the endogeneity between noise and signal terms. All covariances of the signal parts vanish by the martingale property of \tilde{C}_t^n . Under (23), covariances of S_{jk}^2 and $S_{uk'}^2$ due to correlations between $(\epsilon_i)_{0 \leq i \leq n}$ and $(\Delta_i^n X)_{1 \leq i \leq n}$ are only non-zero when $|k - k'| = 1$. Since there are only a finite number of indices with $|i - l| < \tilde{Q}$ on two neighboring bins, we obtain the bound

$$\begin{aligned} & \text{Cov}\left(\left(\sum_{i=1}^n \Delta_i^n \tilde{C}^n \Phi_{j(k-1)}\left(\frac{i}{n}\right)\right)^2, \left(\sum_{l=1}^{n-1} \epsilon_l \varphi_{uk}\left(\frac{l+1/2}{n}\right)\frac{1}{n}\right)^2\right) \\ &= \left(\sum_{i=knh_n}^{knh_n+\tilde{Q}} \sum_{l=i-\tilde{Q}}^i \mathbb{E}[\epsilon_i \Delta_l^n \tilde{C}^n] \Phi_{j(k-1)}\left(\frac{l}{n}\right) \varphi_{uk}\left(\frac{i+1/2}{n}\right)\right)^2 n^{-2}(1 + \mathcal{O}(1)) \\ &= \mathcal{O}(n^{-2}), \end{aligned}$$

uniformly for all k, u, j . We used the same approximations as for the variance terms in Step 1. For the two other covariance terms due to endogeneity, analogous estimates yield bounds of the same order. Under Assumption 2, a similar bound can be proved for the covariances due to serial correlation of the noise. Here, we provide a proof that does not use Assumption 2, but only Assumption (η -p) with (21). We derive that

$$\begin{aligned} & \text{Cov}(S_{jk}^2, S_{uk'}^2) \\ &= \|\Phi_{jk}\|_n^{-2} \|\Phi_{uk'}\|_n^{-2} \text{Cov}\left(\left(\sum_{i=1}^{n-1} \epsilon_i \varphi_{jk}\left(\frac{i+1/2}{n}\right)\frac{1}{n}\right)^2, \right. \\ & \quad \left. \left(\sum_{l=1}^{n-1} \epsilon_l \varphi_{uk'}\left(\frac{l+1/2}{n}\right)\frac{1}{n}\right)^2\right) \\ &= \sum_{i=knh_n}^{(k+1)nh_n-1} \sum_{p=knh_n-i}^{(k+1)nh_n-i-1} \sum_{l=k'nh_n}^{(k'+1)nh_n-1} \sum_{q=k'nh_n-p}^{(k'+1)nh_n-p-1} (\mathbb{E}[\epsilon_i \epsilon_{i+p} \epsilon_l \epsilon_{l+q}] \\ & \quad - \mathbb{E}[\epsilon_i \epsilon_{i+p}] \mathbb{E}[\epsilon_l \epsilon_{l+q}]) \\ & \times \varphi_{jk}\left(\frac{i+1/2}{n}\right) \varphi_{jk}\left(\frac{i+p+1/2}{n}\right) \varphi_{uk'}\left(\frac{l+1/2}{n}\right) \varphi_{uk'}\left(\frac{l+q+1/2}{n}\right) \\ & \times \frac{\|\Phi_{jk}\|_n^{-2} \|\Phi_{uk'}\|_n^{-2}}{n^4} \end{aligned}$$

$$\begin{aligned}
&\leq K \sum_{i=knh_n}^{(k+1)nh_n-1} \sum_{l=k'nh_n}^{(k'+1)nh_n-1} ((\mathbb{E}[\epsilon_i \epsilon_l])^2 + \mathfrak{o}(1)) \frac{\|\Phi_{jk}\|_n^{-2} \|\Phi_{uk'}\|_n^{-2}}{n^4} \varphi_{jk}^2\left(\frac{i+1/2}{n}\right) \\
&\times \varphi_{uk'}^2\left(\frac{l+1/2}{n}\right) \\
&\leq K \sum_{i=knh_n}^{(k+1)nh_n-1} \sum_{l=k'nh_n}^{(k'+1)nh_n-1} ((\mathbb{E}[\epsilon_i \epsilon_l])^2 + \mathfrak{o}(1)) n^{-2} \varphi_{jk}^2\left(\frac{i+1/2}{n}\right) \varphi_{uk'}^2\left(\frac{l+1/2}{n}\right) \\
&= \mathcal{O}(((k-k')nh_n)^{-2-\varrho}), \tag{67}
\end{aligned}$$

where we use similar approximations as in the proof of Lemma 1 and that $\int \varphi_{jk}^2(t) dt = 1$ for all (j, k) . Thereby, we obtain that

$$\begin{aligned}
&n^\beta \sum_{\substack{k, k' = \lfloor sh_n^{-1} \rfloor + 1, \\ k \neq k'}}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n^2 \text{Cov}(\zeta_k(\tilde{C}^n + \epsilon), \zeta_{k'}(\tilde{C}^n + \epsilon)) \\
&= n^\beta r_n \sum_{\substack{k, k' = \lfloor sh_n^{-1} \rfloor + 1, \\ k \neq k'}}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} r_n \sum_{u, j=1}^{J_n} w_{jk} w_{uk'} \text{Cov}(S_{jk}^2, S_{uk'}^2) \\
&= \mathfrak{o}(r_n^{-1} (\log(n))^2 (nh_n)^{-2}) = \mathfrak{o}(r_n^{-1} n^{-1}) = \mathfrak{o}(1).
\end{aligned}$$

This completes the proof of the marginal central limit theorem. At the same time, we obtain analogously

$$n^\beta \sum_{k=\lfloor sh_n^{-1} \rfloor + 1}^{\lfloor sh_n^{-1} \rfloor + r_n^{-1}} \sum_{k'=\lfloor sh_n^{-1} \rfloor - r_n^{-1}}^{\lfloor sh_n^{-1} \rfloor - 1} r_n^2 \text{Cov}(\zeta_k(\tilde{C}^n + \epsilon), \zeta_{k'}(\tilde{C}^n + \epsilon)) = \mathfrak{o}(1).$$

This yields that the covariances of $\hat{\sigma}_s^2$ and $\hat{\sigma}_{s-}^2$ are asymptotically negligible. We conclude the joint stable central limit theorem (26). \square

7.4. Asymptotic theory for the test

Denote by $\{S_1, \dots, S_{N_1}\}$ the finite sequence of stopping times exhausting the jumps of X on $[0, 1]$ with $|\Delta X_{S_i}| > a$ for all i and some $a \in \mathbb{R}_+$ and the Lévy measure of X does not have an atom in $\{a\}$. In the case of finite activity jumps, $r = 0$ in Assumption (H-r), we can set $a = 0$.

Proposition 7.1. *On the null hypothesis $\mathbb{H}(a)_{[0,1]}$, when Assumptions 1, 2, (H-r) and $(\sigma-\alpha)$ are satisfied, the test statistic is asymptotically χ^2 -distributed,*

$$n^\beta T_0(h_n, r_n, g) \xrightarrow{(st)} \chi_{N_1}^2, \tag{68}$$

with N_1 degrees of freedom .

Corollary 7.2. *Under the alternative hypothesis $(\Omega \setminus \mathbb{H}(a)_{[0,1]})$, when there exists at least one $s \in [0, 1]$ with $|\Delta X_s| > a$ and $|\Delta \sigma_s^2| > 0$, it holds as $n \rightarrow \infty$ that:*

$$\mathbb{P}\left(n^\beta T_0(h_n, r_n, g) > q_{1-\alpha}(\chi_{N_1}^2)\right) \rightarrow 1. \tag{69}$$

Proof of Proposition 7.1.:

1. Detection of (large) price jump arrivals

Consider the set

$$\begin{aligned} \tilde{\Omega}_n = & \{\omega \in \Omega | S_i = k \cdot h_n, i = 1, \dots, N_1, k = 0, \dots, h_n^{-1}\}^{\mathbb{G}} \cup \\ & \{\omega \in \Omega | S_1 > r_n^{-1}h_n, S_{N_1} < 1 - r_n^{-1}h_n, S_i - S_{i-1} > 2r_n^{-1}h_n, i = 1, \dots, N_1 - 1\} \\ & \cup \{\omega \in \Omega | \exists(s, i) \text{ s.t. } |\Delta \sigma_s^2| > 0 \text{ and } s \in [S_i - r_n^{-1}h_n, S_i + r_n^{-1}h_n] \setminus \{S_i\}\}^{\mathbb{G}}. \end{aligned}$$

Since $\mathbb{P}(\tilde{\Omega}_n) \rightarrow 1$ as $n \rightarrow \infty$ with (37d), we work conditionally on $\tilde{\Omega}_n$. The jump times $\{S_i, i = 1, \dots, N_1\}$ are estimated with thresholding by $\{\hat{S}_i, i = 1, \dots, \hat{N}_1\}$, where we set $\hat{S}_i = kh_n$ when $h_n |\zeta_k^{ad}(Y)| > u_n \vee a^2$. We prove that

$$\begin{aligned} & \sum_{k=r_n^{-1}}^{h_n^{-1}-r_n^{-1}-1} g(\hat{\sigma}_{kh_n}^2, \hat{\sigma}_{kh_n-}^2) \mathbb{1}_{\{h_n |\zeta_k^{ad}(Y)| > u_n \vee a^2\}} - \sum_{s \leq 1} g(\hat{\sigma}_s^2, \hat{\sigma}_{s-}^2) \mathbb{1}_{\{|\Delta X_s| > a\}} \\ & = \mathcal{O}_{\mathbb{P}}(n^{-\beta}). \end{aligned} \tag{70}$$

Denote $\mathcal{K} = \{0 \leq k \leq h_n^{-1} - 1 | S_i \in (kh_n, (k+1)h_n)\}$ and $\mathcal{K}^{\mathbb{G}} = \{0, \dots, h_n^{-1} - 1\} \setminus \mathcal{K}$. First, we show that

$$\sum_{k \in \mathcal{K}^{\mathbb{G}}} g(\hat{\sigma}_{kh_n}^2, \hat{\sigma}_{kh_n-}^2) \mathbb{1}_{\{h_n |\zeta_k^{ad}(Y)| > u_n \vee a^2\}} = \mathcal{O}_{\mathbb{P}}(n^{-\beta}). \tag{71}$$

With the Markov inequality, Lemma 2 and using that at least $p = 8$ moments of the noise exist, we obtain that

$$\mathbb{P}\left(\sup_{k \in \mathcal{K}^{\mathbb{G}}} |\zeta_k^{ad}(Y)| > h_n^{-1}(u_n \vee a^2)\right) \leq Kh_n^{-1} \frac{\log(n)}{(u_n \vee a^2)^8} h_n^8,$$

for some constant K , and the same order without the factor h_n^{-1} for $\mathbb{P}(|\zeta_k^{ad}(Y)| > h_n^{-1}(u_n \vee a^2))$ and some $k \in \mathcal{K}^{\mathbb{G}}$. Indicator functions $\mathbb{1}_{A_n}$, with $p_n = \mathbb{P}(A_n) \rightarrow 0$, satisfy $\mathbb{1}_{A_n} = \mathcal{O}_{\mathbb{P}}(p_n^{1/2})$, using that $\mathbb{E}[\mathbb{1}_{A_n}] = p_n$ and $\text{Var}(\mathbb{1}_{A_n}) \leq p_n$. Most factors $g(\hat{\sigma}_{kh_n}^2, \hat{\sigma}_{kh_n-}^2)$ in (71) tend to zero in probability. When $|\Delta \sigma_s^2| = 0$ for all $s \in [(k-1)h_n, kh_n]$, we have that $g(\hat{\sigma}_{kh_n}^2, \hat{\sigma}_{kh_n-}^2) = \mathcal{O}_{\mathbb{P}}(n^{-\beta})$. However, for $k \in \mathcal{K}^{\mathbb{G}}$, jumps in $(\sigma_s^2)_{0 \leq s \leq 1}$ can occur. From the summability of $\sum_{s \leq 1} (\Delta \sigma_s^2)^2 < \infty$, it follows that at most n^{2v} volatility jumps of sizes bounded by n^{-v} , $v \in [0, 1/2]$, can occur. Since $g(x, y) = \mathcal{O}((x-y)^2)$ for (29) as $(x-y) \rightarrow 0$, for $a = 0$ we obtain that

$$\sum_{k \in \mathcal{K}^{\mathbb{G}}} g(\hat{\sigma}_{kh_n}^2, \hat{\sigma}_{kh_n-}^2) \mathbb{1}_{\{h_n |\zeta_k^{ad}(Y)| > u_n \vee a^2\}} = \mathcal{O}_{\mathbb{P}}(n^{-\beta} \log(n) h_n^{3-4\tau} + \log(n) h_n^{4(1-\tau)}),$$

respectively $n^{-\beta} \log(n)h_n^3 + \log(n)h_n^4$ for $a > 0$. For $a > 0$, (71) is clearly satisfied, while for $a = 0$ the condition

$$3 - 4\tau > 0 \quad \Rightarrow \quad \tau < 3/4$$

ensures (71). We have proven that the error due to false jump detections is asymptotically negligible.

It remains to prove that the error due to non-detection of one of the finitely many jump times S_1, \dots, S_{N_1} is also asymptotically negligible. This is ensured by

$$- \sum_{k \in \mathcal{K}} g(\hat{\sigma}_{kh_n}^2, \hat{\sigma}_{kh_n-}^2) \mathbb{1}_{\{h_n |\zeta_k^{ad}(Y)| \leq u_n \vee a^2\}} = \mathcal{O}_{\mathbb{P}}(n^{-\beta}). \quad (72)$$

By the results from Section 3.1.3 of [16],³ for $S_i \in ((k-1)h_n, kh_n)$, it holds that

$$h_n |\zeta_k^{ad}(Y)| = (\Delta X_{S_i})^2 + \xi_i = a^2 + \epsilon + \xi_i \quad \text{with } \xi_i = \mathcal{O}_{\mathbb{P}}(1) \text{ and } \epsilon > 0.^4$$

On the hypothesis, there are no simultaneous jumps in the volatility, i.e. $\sigma_{S_i}^2 - \sigma_{S_i-}^2 = 0$ for all $i = 1, \dots, N_1$. On the finitely many bins with $k \in \mathcal{K}$, we thus have that

$$\sup_{k \in \mathcal{K}} g(\hat{\sigma}_{kh_n}^2, \hat{\sigma}_{kh_n-}^2) = \mathcal{O}_{\mathbb{P}}(n^{-\beta}).$$

Hence, $\xi_i = \mathcal{O}_{\mathbb{P}}(1)$ suffices to ensure (72). (71) and (72) imply (70).

2. Stable convergence of spot volatility estimates around detected (large) price-jump times

The asymptotic distribution of the test statistic is derived with (70) and the stable convergences of the spot volatility estimates:

$$n^{\beta/2} \begin{pmatrix} \hat{\sigma}_s^2 - \sigma_s^2 \\ \hat{\sigma}_{s-}^2 - \sigma_{s-}^2 \end{pmatrix} \xrightarrow{(st)} MN \left(0, \begin{pmatrix} 8\sigma_s^3 \eta_s^{1/2} & 0 \\ 0 & 8\sigma_{s-}^3 \eta_s^{1/2} \end{pmatrix} \right),$$

which hold jointly for all $i = 1, \dots, N_1$. The stable limit theorems of the spot volatility estimators are given in Theorem 1. Concerning the convergence of the spot estimates at stopping times, observe that

- Thresholding and identification of a jump is based on $\zeta_k^{ad}(Y)$.
- Given that $h_n |\zeta_k^{ad}(Y)| > u_n \vee a^2$, $\hat{S}_i = kh_n$ for some $i \in \{1, \dots, \hat{N}_1\}$, $\hat{\sigma}_{\hat{S}_i}^2$ is computed from $\zeta_l^{ad}(Y)$, $l = (k+1), \dots, (k+r_n^{-1})$.
- Given $\hat{S}_i = kh_n$, $\hat{\sigma}_{\hat{S}_i-}^2$ is computed from $\zeta_l^{ad}(Y)$, $l = (k-r_n^{-1}), \dots, (k-1)$.

We restrict to $\tilde{\Omega}_n$ again. For the stability of weak convergences, we have already considered a sequence of stopping times in Step 1 of the proof of Theorem 1. Recall the definition of $\tilde{\mathcal{G}}_t^n$ from this paragraph. The $S_p, p = 1, \dots, N_1$, are

³See Proposition 3.2. of [16].

⁴Since the Lévy measure of X does not have an atom in $\{a\}$.

$\tilde{\mathcal{G}}_0^n$ -measurable random variables and denote i_p integer-valued $\tilde{\mathcal{G}}_0^n$ -measurable random variables such that $i_p h_n < S_p < (i_p + 1)h_n$. The stable limit theorem in Theorem 1 is valid when replacing the fixed time s by stopping times $S_p, p = 1, \dots, N_1$. Analogously as in Lemma 8.1 of [29], this readily follows with the points above by the asymptotic independence of the statistics in Step 1 of the proof of Theorem 1 with $s = S_p$ for $\hat{\sigma}_{S_p}^2$, or $s = i_p h_n$ for $\hat{\sigma}_{i_p h_n}^2$ respectively, from \mathcal{F}_{S_p} . Here, we exploit that the noise is under Assumption 2 only weakly serially dependent over asymptotically decreasing intervals and only dependent on finitely many preceding increments of X , and the strong Markov property of Brownian motion.

On assumption $(\sigma-\alpha)$, $\max_p |\sigma_{i_p h_n}^2 - \sigma_{S_p}^2| = \mathcal{O}_{\mathbb{P}}(h_n^\alpha) = \mathcal{O}_{\mathbb{P}}(n^{-\alpha/2})$, the latter being much smaller than $n^{-\beta/2}$. Therefore, a discretization of estimated jump arrivals is asymptotically negligible.

Moreover, on $\tilde{\Omega}_n$ all spot squared volatility estimates are computed from *disjoint* data subsets. Therefore, by (67), covariations between all estimates converge to zero in probability what implies joint weak convergence.⁵ Stability of the convergence of the vector has been established above in Step 1 of the proof of Theorem 1.

3. Convergence of the test statistic

For test functions which are twice continuously differentiable with bounded second derivatives, Taylor’s formula yields

$$\begin{aligned} g(x_1, x_2) - g(a_1, a_2) &= \frac{\partial g}{\partial x_1}(a_1, a_2)(x_1 - a_1) + \frac{\partial g}{\partial x_2}(a_1, a_2)(x_2 - a_2) \\ &+ \frac{\partial^2 g}{2 \partial x_1^2}(a_1, a_2)(x_1 - a_1)^2 + \frac{\partial^2 g}{2 \partial x_2^2}(a_1, a_2)(x_2 - a_2)^2 \\ &+ \frac{\partial^2 g}{\partial x_1 \partial x_2}(a_1, a_2)(x_1 - a_1)(x_2 - a_2) \\ &+ \mathcal{O}(\max((x_1 - a_1)^2, (x_2 - a_2)^2)). \end{aligned}$$

We apply the generalized Δ -method for stable convergence and set $(a_1, a_2) = (\sigma_{S_i}^2, \sigma_{S_{i-}}^2)$ and the random vector $(x_1, x_2) = (\hat{\sigma}_{S_i}^2, \hat{\sigma}_{S_{i-}}^2)$ with estimators (17a) and (17b) at the times $S_i, S_{i-}, i = 1, \dots, N_1$. When we focus on the test function (29) in Theorem 2, it holds that

$$\frac{\partial g}{\partial x_1}(\sigma_{S_i}^2, \sigma_{S_i}^2) = \frac{\partial g}{\partial x_2}(\sigma_{S_i}^2, \sigma_{S_i}^2) = g(\sigma_{S_i}^2, \sigma_{S_i}^2) = 0.$$

The second order term comes into play and the equalities

$$\frac{\partial^2 g}{\partial x_1^2}(\sigma_{S_i}^2, \sigma_{S_i}^2) = \frac{\partial^2 g}{\partial x_2^2}(\sigma_{S_i}^2, \sigma_{S_i}^2) = -\frac{\partial^2 g}{\partial x_1 \partial x_2}(\sigma_{S_i}^2, \sigma_{S_i}^2) = \frac{1}{8} \sigma_{S_i}^{-3}. \quad (73)$$

⁵Note that by Step 4 in the proof of Theorem 1, this is still true if the pre-estimated noise long-run variance was computed from all observations.

Under Assumption 2 we have by Proposition 3.1 estimators $\hat{\eta}_{kh_n}^{1/2} = \eta_{kh_n}^{1/2} + \mathcal{O}_{\mathbb{P}}(n^{-\beta})$ for all k . This renders the estimation errors of $\hat{\eta}_{kh_n}^{-1/2}$ in (28) asymptotically negligible in (30).

Cramér-Wold's theorem gives equivalence of the weak convergence of the vector $(\hat{\sigma}_{S_i}^2, \hat{\sigma}_{S_i-}^2)_{1 \leq i \leq N_1}$ to weak convergence of linear combinations. Under $\mathbb{H}(a)_{[0,1]}$, when $\sigma_{S_i}^2 = \sigma_{S_i-}^2$ for all i , the limit of $n^\beta T_0(h_n, r_n, g)$ can thus be described by a random variable

$$\sum_{i=1}^{N_1} \left(\frac{\partial^2 g}{2 \partial x_1^2}(\sigma_{S_i}^2, \sigma_{S_i}^2) Z_i^2 + \frac{\partial^2 g}{2 \partial x_2^2}(\sigma_{S_i}^2, \sigma_{S_i}^2) \tilde{Z}_i^2 + \frac{\partial^2 g}{\partial x_1 \partial x_2}(\sigma_{S_i}^2, \sigma_{S_i}^2) Z_i \tilde{Z}_i \right) 8\sigma_{S_i}^3,$$

where Z_i and \tilde{Z}_i , $i = 1, \dots, N_1$, are two independent collections of i.i.d. standard normals defined on the orthogonal extension of $(\Omega, \mathcal{F}, \mathbb{P})$ in the product space that accommodates all random variables. Since $(1/\sqrt{2})(Z_i - \tilde{Z}_i)$ are i.i.d. standard normals, the χ^2 -distribution with N_1 degrees of freedom appears as limiting distribution. Proposition 7.1 follows with the binomial formula and by the second derivatives of the test function (29) from (73). Even though the limit above could depend on the particular choice of stopping times its \mathcal{F} -conditional law does not. \square

Proof of Corollary 7.2.

Under the alternative hypothesis, $\sigma_{S_i}^2 \neq \sigma_{S_i-}^2$, for at least one $i \in \{1, \dots, N_1\}$. In this case, we have that

$$\begin{aligned} n^\beta T_0(h_n, r_n, g) &= \mathcal{O}_{\mathbb{P}}(1) + n^\beta \hat{\eta}_{\lfloor S_i h_n^{-1} \rfloor h_n}^{-1/2} g(\hat{\sigma}_{\lfloor S_i h_n^{-1} \rfloor h_n}^2, \hat{\sigma}_{\lfloor S_i h_n^{-1} \rfloor h_n-}^2) \\ &\quad \times \mathbb{1}_{\{h_n |\zeta_{\lfloor S_i h_n^{-1} \rfloor h_n}^{ad}(Y)| > (u_n \vee a^2)\}} \end{aligned}$$

with Proposition 7.1. Since

$$g(\hat{\sigma}_{\lfloor S_i h_n^{-1} \rfloor h_n}^2, \hat{\sigma}_{\lfloor S_i h_n^{-1} \rfloor h_n-}^2) \geq c \Delta \sigma_{S_i}^2 - \mathcal{O}_{\mathbb{P}}(1)$$

for some constant c and since $h_n |\zeta_k^{ad}(Y)| = (\Delta X_{S_i})^2 + \xi_i = a^2 + \epsilon + \xi_i$ with $\xi_i = \mathcal{O}_{\mathbb{P}}(1)$ and some $\epsilon > 0$, we conclude with the reverse triangle inequality that

$$\mathbb{P}\left(n^\beta T_0(h_n, r_n, g) > q_{1-\alpha}(\chi_{N_1}^2)\right) \rightarrow 1$$

for any arbitrarily small $\alpha > 0$. This proves Corollary 7.2. \square

Acknowledgement

The authors are grateful to two anonymous referees whose valuable comments and suggestions helped improving the paper.

References

- [1] Aït-Sahalia, Y., J. Fan, R. J. A. Laeven, C. D. Wang, and X. Yang (2017). Estimation of the continuous and discontinuous leverage effects. *Journal of the American Statistical Association*, 112(520), 1744–1758. [MR3750896](#)
- [2] Aït-Sahalia, Y. and J. Jacod (2010). Is Brownian motion necessary to model high-frequency data? *The Annals of Statistics* 38(5), 3093–3128. [MR2722465](#)
- [3] Aït-Sahalia, Y. and J. Jacod (2014). *High-frequency financial econometrics*. Princeton, NJ: Princeton University Press.
- [4] Aït-Sahalia, Y., L. Zhang, and P. A. Mykland (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies* 18, 351–416.
- [5] Altmeyer, R. and M. Bibinger (2015). Functional stable limit theorems for quasi-efficient spectral covolatility estimators. *Stochastic Processes and their Applications* 125(12), 4556–4600. [MR3406596](#)
- [6] Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39(4), 885–905.
- [7] Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* 96(453), 42–55. [MR1952727](#)
- [8] Bandi, F. and R. Renò (2016). Price and volatility co-jumps. *Journal of Financial Economics* 119(1), 107–146.
- [9] Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica* 76(6), 1481–1536. [MR2468558](#)
- [10] Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society* 64(2), 253–280. [MR1904704](#)
- [11] Bibinger, M. (2011). Efficient covariance estimation for asynchronous noisy high-frequency data. *Scandinavian Journal of Statistics* 38, 23–45. [MR2760138](#)
- [12] Bibinger, M., N. Hautsch, P. Malec, and M. Reiß (2014). Estimating the quadratic covariation matrix from noisy observations: Local method of moments and efficiency. *The Annals of Statistics* 42(4), 1312–1346. [MR3226158](#)
- [13] Bibinger, M., N. Hautsch, P. Malec, and M. Reiß (2017). Estimating the spot covariation of asset prices – statistical theory and empirical evidence. *Journal of Business & Economic Statistics*, forthcoming. [MR3055330](#)
- [14] Bibinger, M., M. Jirak, and M. Vetter (2017). Nonparametric change-point analysis of volatility. *The Annals of Statistics* 45(4), 1542–1578. [MR3670188](#)
- [15] Bibinger, M. and M. Reiß (2014). Spectral estimation of covolatility from noisy observations using local weights. *Scandinavian Journal of Statistics* 41(1), 23–50. [MR3181131](#)

- [16] Bibinger, M. and L. Winkelmann (2015). Econometrics of co-jumps in high-frequency data with noise. *Journal of Econometrics* 184(2), 361 – 378. [MR3291008](#)
- [17] Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica* 77(3), 623–685. [MR2531358](#)
- [18] Clinet, S. and Y. Potiron (2017). Efficient asymptotic variance reduction when estimating volatility in high frequency data. *arXiv:1701.01185*.
- [19] Comte, F. and E. Renault (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance* 8(4), 291–323. [MR1645101](#)
- [20] Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68(6), 1343–1376. [MR1793362](#)
- [21] Fan, J. and Y. Wang (2007). Multi-scale jump and volatility analysis for high-frequency data. *Journal of the American Statistical Association* 102(480), 1349–1362. [MR2372538](#)
- [22] Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics* 24(2), 127–161. [MR2234447](#)
- [23] Hautsch, N. and M. Podolskij (2013). Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: Theory, implementation, and empirical evidence. *Journal of Business & Economic Statistics* 31(2), 165–183. [MR3055330](#)
- [24] Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Processes and their Applications* 118(4), 517–559. [MR2394762](#)
- [25] Jacod, J., C. Klüppelberg, and G. Müller (2017). Testing for non-correlation between price and volatility jumps. *Journal of Econometrics* 197(2), 284–297. [MR3614202](#)
- [26] Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and their Applications* 119, 2803–2831. [MR2531091](#)
- [27] Jacod, J. and P. A. Mykland (2015). Microstructure noise in the continuous case: Approximate efficiency of the adaptive pre-averaging method. *Stochastic Processes and their Applications* 125, 2910 – 2936. [MR3343282](#)
- [28] Jacod, J. and P. Protter (2012). *Discretization of processes*. Springer. [MR2859096](#)
- [29] Jacod, J. and V. Todorov (2010). Do price and volatility jump together? *The Annals of Applied Probability* 20(4), 1425–1469. [MR2676944](#)
- [30] Kalnina, I. and D. Xiu (2017). Nonparametric Estimation of the Leverage Effect: A Trade-Off Between Robustness and Efficiency. *Journal of the American Statistical Association* 112(517), 384–396. [MR3646579](#)
- [31] Koike, Y. (2016). Estimation of integrated covariances in the simultaneous presence of nonsynchronicity, microstructure noise and jumps. *Econometric Theory* 32(3), 533–611. [MR3506433](#)
- [32] Lee, S. and P. A. Mykland (2008). Jumps in financial markets: A new

- nonparametric test and jump dynamics. *Review of Financial Studies* 21(6), 2535–2563.
- [33] Lee, S. and P. A. Mykland (2012). Jumps in equilibrium prices and market microstructure noise. *Journal of Econometrics* 168(2), 396–406. [MR2923776](#)
- [34] Liu, J., F. Longstaff, and J. Pan (2003). Dynamic asset allocation with event risk. *Journal of Finance* 58(1), 231–259.
- [35] Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics* 36(4), 270–296. [MR2528985](#)
- [36] Mancini, C., V. Mattiussi, and R. Reno (2015). Spot volatility estimation using delta sequences. *Finance and Stochastics* 19(2), 261–293. [MR3320322](#)
- [37] Munk, A. and J. Schmidt-Hieber (2010a). Lower bounds for volatility estimation in microstructure noise models. In J. O. Berger, T. T. Cai, and I. M. Johnstone (Eds.), *Borrowing Strength: Theory Powering Applications – A Festschrift for Lawrence D. Brown*, Volume 6 of *Collections*, pp. 43–55. Beachwood, Ohio, USA: Institute of Mathematical Statistics. [MR2798510](#)
- [38] Munk, A. and J. Schmidt-Hieber (2010b). Nonparametric estimation of the volatility function in a high-frequency model corrupted by noise. *Electronic Journal of Statistics* 4, 781–821. [MR2684388](#)
- [39] Pastor, L. and P. Veronesi (2012). Uncertainty about government policy and stock prices. *Journal of Finance* 67(4), 1219–1264.
- [40] Reiß, M. (2011). Asymptotic equivalence for inference on the volatility from noisy observations. *The Annals of Statistics* 39(2), 772–802. [MR2816338](#)
- [41] Tauchen, G. and V. Todorov (2011). Volatility jumps. *Journal of Business and Economic Statistics* 29(3), 356–371. [MR2848503](#)
- [42] Todorov, V. (2010). Variance risk-premium dynamics: The role of jumps. *Review of Financial Studies* 23(1), 345–383.
- [43] Winkelmann, L., M. Bibinger, and T. Linzert (2016). Ecb monetary policy surprises: Identification through cojumps in interest rates. *Journal of Applied Econometrics* 31(4), 613–629. [MR3512316](#)
- [44] Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli* 12(6), 1019–1043. [MR2274854](#)
- [45] Zu, Y. and H. P. Boswijk (2014). Estimating spot volatility with high-frequency financial data. *Journal of Econometrics* 181(2), 117 – 135. [MR3209859](#)