Annals of the Institute of Statistical Mathematics (2021) 73:649–669
https://doi.org/10.1007/s10463-020-00762-3

Check for updates

Estimation for high-frequency data under parametric market microstructure noise

Simon Clinet[1] **· Yoann Potiron**[2]

Received: 11 September 2019 / Revised: 4 August 2020 / Published online: 16 September 2020

Abstract
We develop a general class of noise-robust estimators based on the existing estimators in the non-noisy high-frequency data literature. The microstructure noise is a parametric function of the limit order book. The noise-robust estimators are constructed as plug-in versions of their counterparts, where we replace the efficient price, which is non-observable, by an estimator based on the raw price and limit order book data. We show that the technology can be applied to five leading examples where, depending on the problem, price possibly includes infinite jump activity and sampling times encompass asynchronicity and endogeneity.

Keywords Functionals of volatility · High-frequency covariance · High-frequency data · Limit order book · Parametric market microstructure noise

1 Introduction

It is now widely acknowledged that the availability of high-frequency data has led to a more accurate description of financial markets. Over the past decades, empirical studies have unveiled several aspects of the frictionless efficient price. Accordingly, the assumptions on the latter have been gradually weakened to the extent that it is common nowadays to represent it as a general Itô semi-martingale including

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10463-020-00762-3) contains supplementary material, which is available to authorized users.

✉ Simon Clinet
clinet@keio.jp
http://user.keio.ac.jp/~clinet/

Yoann Potiron
potiron@fbc.keio.ac.jp
http://www.fbc.keio.ac.jp/~potiron

[1] Faculty of Economics, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan

[2] Faculty of Business and Commerce, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan

jumps with infinite activity. Moreover, the sampling times are also often considered as asynchronous, random, and even sometimes endogenous, i.e. possibly correlated with the efficient price. The accessibility of high-frequency data has also shed light on the frictions, or so-called market microstructure noise (MMN), which get prominent as the sampling frequency increases. As a matter of fact, realized volatility (i.e. summing the square returns), which is efficient in the absence of frictions, becomes badly biased when the frequency increases. This was visible on the signature plot in Andersen et al. (2001a). A typical challenge that faces a theoretical statistician today is to incorporate jumps, asynchronicity, endogeneity and frictions into the model.

A frequently used set-up is

$$\underbrace{Z_{t_i}}_{\text{observed price}} = \underbrace{X_{t_i}}_{\text{efficient price}} + \underbrace{\epsilon_{t_i}}_{\text{MMN}}, \tag{1}$$

where ϵ_{t_i} is i.i.d. and latent. In two nice and independent papers, Li et al. (2016) and Chaker (2017), and subsequently Clinet and Potiron (2019a, b), consider the following parametric form for the noise to estimate volatility:

$$\underbrace{Z_{t_i}}_{\text{observed price}} = \underbrace{X_{t_i}}_{\text{efficient price}} + \underbrace{\underbrace{\phi\big(Q_{t_i},\theta_0\big)}_{\text{parametric noise}} + \underbrace{\epsilon_{t_i}}_{\text{residual noise}}}_{\text{MMN}}, \tag{2}$$

where Q_{t_i} is the information from the limit order book and ϕ is a function known to the statistician. A simple and familiar example was introduced in Roll (1984), and specified in e.g. Hasbrouck (2002), where

$$\phi\big(Q_{t_i},\theta_0\big) = I_{t_i}\theta_0, \tag{3}$$

with I_{t_i} corresponding to the trade direction, i.e. 1 if the transaction at time t_i is buyer initiated and -1 if seller initiated, and θ_0 standing for half of the effective spread. In Glosten and Harris (1988), the extension includes the trading volume V_{t_i} and takes on the form

$$\phi\big(Q_{t_i},\theta_0\big) = I_{t_i}\theta_0^{(1)} + I_{t_i}V_{t_i}\theta_0^{(2)}. \tag{4}$$

A different model features information about the quoted spread S_{t_i}, where

$$\phi\big(Q_{t_i},\theta_0\big) = I_{t_i}S_{t_i}\theta_0. \tag{5}$$

This model can be seen as an updated time-varying Roll model, as the quoted spread is nowadays available in the structure of current limit order book markets, whereas it was not observed at the time when Roll model was proposed.

There are two regimes related to the parametric model (2), i.e. the null residual noise and nonzero residual noise. To estimate volatility, the cited papers rely on a plug-in procedure. In a first step, they provide estimators of the parameter θ_0 and establish fast convergence rate which satisfies

$$N\left(\hat{\theta} - \theta_0\right) = O_{\mathbb{P}}(1), \tag{6}$$

where N stands for the number of observations and pre-estimate the efficient price via

$$\hat{X}_{t_i} = Z_{t_i} - \phi\left(Q_{t_i}, \hat{\theta}\right). \tag{7}$$

In a second step, one can apply a "usual" estimator of volatility, considering the observed price as in fact the pre-estimated efficient price. More specifically, in case of absence of residual noise, the cited papers implement realized volatility and retrieve efficiency of the method. In the presence of residual noise, they also provide residual noise robust estimators.

In this paper, we will assume the null residual noise regime, which we agree is quite a strong assumption (at first glance). Indeed, from a theoretical statistician standpoint, the nonzero residual noise regime, of which the common set-up (1) is a particular case, is obviously more challenging. Nonetheless, the original papers Li et al. (2016) and Chaker (2017) most likely wanted to select empirically variables from the limit order book that fully explain the MMN. Actually, in their empirical study on four stocks and one day, Li et al. (2016) find that the residual noise of models such as (3) and (4) accounts for 20–30% of the total MMN variance, which is quite low and yet not negligible. Chaker (2017) proposes and implements on a full year of one stock from the New York Stock Exchange tests for the absence of residual noise. She finds rejection rate around 15–25% for (3), and 10–30% in the case of (4), here again quite nice results but not indicating the absence of residual noise. More recently, implemented on a month with 31 constituents from the CAC 40, Clinet and Potiron (2019b) find that the "best" model among several competitors from the financial economics literature is (5), with related residual noise accounting for as low as 1% of the MMN variance, and results in line with previous findings for the other models. Finally, in an extensive study on 50 stocks randomly selected from the S&P 500 during the period 2009–2017, Clinet and Potiron (2019a) exhibit (5) as the model explaining the most variance of the MMN, with residual noise accounting for (almost) 0% of the total MMN variance. Those two empirical studies back up the null residual noise regime.

When implementing a non-noise-robust procedure with high frequency data, it is often the case that the applied statistician faces a dilemma in using tick-by-tick data on the statistical principle that one should not throw away data, or subsampling—say every five minutes—in respect to the limited theoretical assumptions. We argue that the plug-in approach is a cheap method that kills two birds with one stone. On the one hand, it provides the theoretical statistician with a simple and transparent method for adding MMN in his theory. On the other hand, this will be useful for the applied statistician as he/she will be able to use tick-by-tick data when implementing the related estimator. This strategy is actually successfully used in Andersen et al. (2019). In particular, our paper enlightens the theoretical aspect of the plug-in approach.

To do so, we describe the general framework as follows. If we define the horizon time as T, one typically seeks to estimate the random integrated parameter

$$\Xi = \int_0^T \xi_t \mathrm{d}t, \tag{8}$$

where the spot parameter ξ_t is a stochastic process which can correspond to the volatility, the high-frequency covariance, functionals of volatility and volatility of volatility, employing a given data-based estimator $\widetilde{\Xi}(X_{t_0}, \dots, X_{t_N})$. In the absence of noise, $\widetilde{\Xi}$ usually enjoys a stable central limit theorem of the form

$$N^\kappa \left(\widetilde{\Xi} - \Xi \right) \to \mathcal{MN}(AB, AVAR), \tag{9}$$

where $\kappa > 0$ corresponds to the rate of convergence, and $\mathcal{MN}(AB, AVAR)$ designates a mixed normal distribution of random bias AB and random variance $AVAR$ (due to the fact that the parameter itself is random). In addition, for the purpose of practical implementation, one typically provides a related studentized central limit theorem, i.e. data-based statistics $\widetilde{AB}(X_{t_0}, \dots, X_{t_N})$ and $\widetilde{AVAR}(X_{t_0}, \dots, X_{t_N})$ such that

$$N^\kappa \frac{\widetilde{\Xi} - N^{-\kappa}\widetilde{AB} - \Xi}{\sqrt{\widetilde{AVAR}}} \to \mathcal{N}(0, 1). \tag{10}$$

Accordingly, when observations are contaminated by the parametric noise, we propose to exploit the corresponding class of plug-in estimators to estimate the integrated parameter. They are constructed as $\widehat{\Xi} = \widetilde{\Xi}(\widehat{X}_{t_0}, \dots, \widehat{X}_{t_N})$, $\widehat{AB} = \widetilde{AB}(\widehat{X}_{t_0}, \dots, \widehat{X}_{t_N})$ and $\widehat{AVAR} = \widetilde{AVAR}(\widehat{X}_{t_0}, \dots, \widehat{X}_{t_N})$. This plug-in approach seems to be traced back to the framework of the model with uncertainty zones from Robert and Rosenbaum (2010, 2012).

The main contribution of this paper is presented in Sect. 4, where we state that under parametric noise the central limit theorems (9) and (10) still hold when we substitute the estimators by their related plug-in version in five leading examples of the literature. Depending on the problem at hand, price possibly features jumps with infinite activity and sampling times include asynchronicity and endogeneity. The first example considers the threshold realized volatility inspired by Andersen et al. (2001b), Barndorff-Nielsen and Shephard (2002b) and Mancini (2009). Technically, we extend the central limit theory of realized volatility under endogenous sampling in Li et al. (2014), which includes no jumps to allow for jumps with infinite activity. The second example deals with the threshold bipower variation, which was originally with no threshold in Barndorff-Nielsen and Shephard (2004), and from Corsi et al. (2010) and Vetter (2010). In the third example, we discuss the Hayashi and Yoshida (2005) estimator to estimate high-frequency covariance. The fourth example is devoted to the local estimator from Jacod and Rosenbaum (2013) which estimates functionals of volatility. Finally, we focus on the estimator of volatility of volatility introduced in Vetter (2015) in the last example.

In all those examples, the only required assumption on $\widehat{\theta}$ to obtain (9) and (10) is the fast convergence (6), which is already obtained in a general setting where price process features big jumps in Li et al. (2016), so that our contribution in that respect boils down to adding possible small jumps. Moreover, the asymptotic properties in both equations remain unchanged, whereas the rate of convergence is slower in the i.i.d latent noise case. It means that the parametric noise assumption induces faster rates of convergence than the i.i.d condition, but it is fair to say that we play a different game in this paper as plug-in estimators exploit supplementary data available from the limit order book.

The rest of this paper is structured as follows. Section 2 introduces the model. Section 3 is devoted to the estimation. The five examples are developed in Sect. 4. We conclude in Sect. 5. Proofs can be found in Sect. 6 in Electronic supplementary material.

2 Model

Almost all the quantities defined in what follows are multi-dimensional. Accordingly, the notation $x^{(k)}$ refers to the kth component of x. We define the horizon time as $T > 0$, and the (possibly random) number of observations[1] as N. The observation times, which satisfy $0 \leq t_0^{(k)} \leq \cdots \leq t_N^{(k)} \leq T$, are possibly asynchronous, i.e. they may differ from one price component to the next (see Sect. 4.3), and endogenous, i.e. correlated with X_t (as in Sects. 4.1 and 4.3). When observations are regular and synchronous, we have $\Delta_i t := t_i - t_{i-1} = T/n := \Delta$ (as in Sects. 4.2, 4.4 and 4.5), which implicitly means that $N = n$ and t_i are 1-dimensional, although the price process can be multi-dimensional.

In view of the empirical findings described in the introduction, it is natural to specify (2) as the "pure" parametric noise model via

$$\underbrace{Z_{t_i}}_{\text{observed price}} = \underbrace{X_{t_i}}_{\text{efficient price}} + \underbrace{\phi\left(Q_{t_i}, \theta_0\right)}_{\text{parametric noise}}, \tag{11}$$

where the parameter $\theta_0 \in \Theta \subset \mathbb{R}^l$ with Θ a compact set, the impact function ϕ is known of class C^3 in its second argument, and $Q_{t_i} \in \mathbb{R}^q$ includes *observable* information[2] at the observation time t_i from the limit order book such as the aforementioned trade type (Roll 1984), trading volume (Glosten and Harris 1988) and quoted

[1] All the defined quantities are implicitly or explicitly indexed by n (except for the integrated parameter which does not depend on n). For example N should be thought and considered as N_n. Consistency and convergence in law refer to the behavior as $n \to \infty$. A full specification of the model also involves the stochastic basis $\mathcal{B} = (\Omega, \mathbb{P}, \mathcal{F}, \mathbf{F})$, where $\mathcal{F}$ is a σ-field and $\mathbf{F} = (\mathcal{F}_t)_{t \in [0,T]}$ is a filtration, which will be example-specific. We assume that all the processes (including the integrated parameter ξ_t) are $\mathbf{F}$-adapted (either in a continuous or discrete meaning for Q_{t_i}) and that the observation times t_i are $\mathbf{F}$-stopping times. Also, when referring to Itô-semimartingale and stable convergence in law, we automatically mean that the statement is relative to $\mathbf{F}$. Finally, we assume in (13) that W is also a Brownian motion under the larger filtration $\mathcal{H}_t = \mathcal{F}_t \vee \sigma\{Q_{t_i}, 0 \leq i \leq N\}$.

[2] Note that we do not assume that Q_t exists for any $t \in [0,T] - \{t_0, \ldots, t_N\}$ as it is often the case in the i.i.d setting, see, e.g., the framework in Jacod et al. (2009).

bid-ask spread, but also possibly the duration time between two trades (Almgren and Chriss 2001), the quoted depth (Kavajecz 1999), the order flow imbalance (Cont et al. 2014), etc. In practice, ϕ could be always chosen as (5), although we do not specify this particular model in the paper for generality purposes. Further discussion is available in: Black (1986), Hasbrouck (1993), O'hara (1995), Madhavan et al. (1997), Madhavan (2000), Stoll (2000) and Hasbrouck (2007) among other prominent works. One can also look at the review from Diebold and Strasser (2013). Finally, on the grounds that the one-lag autocorrelation in mid price returns is often found positive empirically, Andersen et al. (2017) extend the usual martingale-plus-noise setting to allow for positivity in the one-lag serial autocorrelation. Note that the model of (11), without residual noise, is theoretically interesting because it allows to adapt the existing methods by plugging in the estimated price in place of the existing estimator.

Finally, we assume that

$$\max_{i,j,k}\left|Q^{(k,j)}_{t_i^{(k)}}\right| = O_{\mathbb{P}}(1), \tag{12}$$

where $Q^{(k)}_{t_i^{(k)}} = (Q^{(k,1)}_{t_i^{(k)}}, \dots, Q^{(k,j_k)}_{t_i^{(k)}})$ corresponds to the information related to $X^{(k)}$ at time $t_i^{(k)}$. The latent d-dimensional log-price X_t possibly including jumps and its related d^2-dimensional spot volatility $c_t = \sigma_t\sigma_t^T$ are Itô-semimartingales of the form

$$\begin{aligned} X_t =& X_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + \int_0^t \int_{\mathbb{R}} \delta(s,z)\mathbf{1}_{\{\|\delta(s,z)\|\le 1\}}(\mu-\nu)(\mathrm{d}s,\mathrm{d}z) \\ &+ \int_0^t \int_{\mathbb{R}} \delta(s,z)\mathbf{1}_{\{\|\delta(s,z)\|>1\}}\mu(\mathrm{d}s,\mathrm{d}z), \end{aligned} \tag{13}$$

$$\begin{aligned} c_t =& c_0 + \int_0^t \widetilde{b}_s \mathrm{d}s + \int_0^t \widetilde{\sigma}_s \mathrm{d}W'_s + \int_0^t \int_{\mathbb{R}} \widetilde{\delta}(s,z)\mathbf{1}_{\{\|\widetilde{\delta}(s,z)\|\le 1\}}(\mu-\nu)(\mathrm{d}s,\mathrm{d}z) \\ &+ \int_0^t \int_{\mathbb{R}} \widetilde{\delta}(s,z)\mathbf{1}_{\{\|\widetilde{\delta}(s,z)\|>1\}}\mu(\mathrm{d}s,\mathrm{d}z), \end{aligned} \tag{14}$$

where W_t is a d-dimensional Brownian motion and W'_t is a d^2-dimensional Brownian motion possibly correlated with W_t, the d-dimensional b_t and d^2-dimensional $\widetilde{b}_t$ drifts are locally bounded, σ_t and the d^2-dimensional $\widetilde{c}_t = \widetilde{\sigma}_t\widetilde{\sigma}_t^T$ are locally bounded, μ is a Poisson random measure on $\mathbb{R}^+ \times E$ where E is an auxiliary Polish space, with the related intensity measure, i.e. the nonrandom predictable compensator, $\nu(\mathrm{d}t,\mathrm{d}z) = \mathrm{d}t \otimes \lambda(\mathrm{d}z)$ for some σ-finite measure λ on $\mathbb{R}^+$. Finally, $\delta = \delta(\omega,t,z)$ (respectively $\widetilde{\delta}$) is a predictable $\mathbb{R}^d$-valued ($\mathbb{R}^{d\times d}$-valued) function on $\Omega \times \mathbb{R}^+ \times \mathbb{R}$ such that locally $\sup_{\omega,t} \| \delta(\omega,t,z) \|^r \le \gamma(z)$ $(\sup_{\omega,t} \| \widetilde{\delta}(\omega,t,z) \|^{\widetilde{r}} \le \gamma(z))$ for some nonnegative bounded λ-integrable function γ and some[3] $r \in [0,1)$ $(\widetilde{r} = 2)$.

[3] Here the restriction $r < 1$ follows from Jacod and Rosenbaum (2013). Indeed, even for the realized volatility problem, (16) may not happen in the case $r > 1$. Indeed, it yields a different optimal rate of convergence as shown in Jacod and Reiss (2014) (of the form $N^{\kappa}\log N$ for some $\kappa > 0$). Moreover, as explained in their Remark 3.4, a CLT is not even achievable in some cases. The case $r = 1$ is let aside. Such bordercase is examined in Vetter (2010) when considering the bipower variation.

Furthermore, we define the "genuine" drift as $b'_t = b_t - \int \delta(t,z)\mathbf{1}_{\{||\delta(t,z)||\leq 1\}}\lambda(\mathrm{d}z)$, the continuous part of X_t as

$$X'_t = X_0 + \int_0^t b'_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s,$$

and the jump part as $J_t = \sum_{s\leq t} \Delta X_s$. Key to our analysis is the decomposition

$$X_t = X'_t + J_t. \tag{15}$$

3 Estimation under parametric noise

3.1 Integrated parameter estimation

The object of interest can be the integrated volatility, etc. In the non-noisy version of the problem, the typical scenario is such that the high-frequency data user has a data-based estimator $\widetilde{\Xi}\left(X_{t_0}, \dots, X_{t_N}\right)$ of (8), such as the standard realized volatility (RV), i.e. $RV = \sum_{i=1}^N \Delta_i X^2$ where $\Delta_i A = A_{t_i} - A_{t_{i-1}}$, and possibly a related central limit theorem and a studentized version of it. In all generality, they, respectively, take the form of

$$N^\kappa\left(\widetilde{\Xi} - \Xi\right) \to \mathcal{MN}(AB, AVAR), \tag{16}$$

where $\kappa > 0$ corresponds to the rate of convergence, and

$$N^\kappa \frac{\widetilde{\Xi} - N^{-\kappa}\widetilde{AB} - \Xi}{\sqrt{\widetilde{AVAR}}} \to \mathcal{N}(0,1), \tag{17}$$

where $\widetilde{AB}\left(X_{t_0}, \dots, X_{t_N}\right)$ and $\widetilde{AVAR}\left(X_{t_0}, \dots, X_{t_N}\right)$ are also data-based statistics which, respectively, correspond to the asymptotic bias and the asymptotic variance estimator. The aim of this section is to equip the high-frequency data user with noise-robust estimators which are based on $\widetilde{\Xi}$.

To estimate the integrated parameter, we first need an estimator of the noise parameter θ_0 defined as $\widehat{\theta}$. We assume that $\widehat{\theta}$ satisfies

$$N\left(\widehat{\theta} - \theta_0\right) = O_{\mathbb{P}}(1). \tag{18}$$

The techniques of this paper are estimator independent and only require (18). In Sect. 3.2, we provide the form of the estimators from the literature which satisfy (18) (see Proposition 1 below). Based on $\widehat{\theta}$, the efficient price is naturally estimated as

$$\widehat{X}_{t_i} = Z_{t_i} - \phi\left(Q_{t_i}, \widehat{\theta}\right). \tag{19}$$

This estimator was already used in Li et al. (2016), Chaker (2017) and Clinet and Potiron (2019b). The related plug-in estimator is constructed as

$$\widehat{\Xi} = \widetilde{\Xi}\left(\widehat{X}_{t_0}, \dots, \widehat{X}_{t_N}\right). \tag{20}$$

For instance, in the case of RV, we obtain that $\widehat{RV} = \sum_{i=1}^{N} \Delta_i \widehat{X}^2$. Similarly, we introduce $\widehat{AB} = \widetilde{AB}(\widehat{X}_{t_0}, \dots, \widehat{X}_{t_N})$ and $\widehat{AVAR} = \widetilde{AVAR}(\widehat{X}_{t_0}, \dots, \widehat{X}_{t_N})$.

We end this section with a succinct remark on the theoretical implications of (18). At this point, the reader may notice that the fast rate N^{-1} in (18), which implies the approximation $\widehat{X}_{t_i} = X_{t_i} + \psi_i(\widehat{\theta})$ with $\psi_i(\widehat{\theta}) = \phi(Q_{t_i}, \theta_0) - \phi(Q_{t_i}, \widehat{\theta}) = O_{\mathbb{P}}(N^{-1})$ by (3.1), suggests that the perturbation $\psi_i(\widehat{\theta})$ acts as an additional drift component and therefore could be systematically treated as such in all derivations. There is, however, a fundamental difference between the two quantities, in that drift returns $\Delta_i B = \int_{t_{i-1}}^{t_i} b_s \mathrm{d}s$ are typically adapted, hence $\mathcal{F}_{t_i}$ measurable, whereas $\psi_i(\widehat{\theta})$, through $\widehat{\theta}$, depends not only on the additional observations $(Q_{t_j})_{j=0,\dots,N}$ but also on the whole trajectory of the price process X, that is $\mathcal{F}_T$. This may pose a problem when considering, for instance, terms of the form $A_{i-1}\Delta_i M$ where $\Delta_i M$ is a martingale increment (even for the augmented filtration $\mathcal{H}_t = \mathcal{F}_t \vee \sigma\{Q_{t_i}, 0 \leq i \leq N\}$). Indeed, when $A_{i-1} = \Delta_{i-1}B$, it naturally preserves the martingale structure of $A_{i-1}\Delta_i M$. On the other hand, if $A_{i-1} = \psi_{i-1}(\widehat{\theta})$, such a structure is broken, and additional arguments are necessary in order to retrieve the desired order of the increment $A_{i-1}\Delta_i M$. In this simple example, the problem can be circumvented with a Taylor expansion $\psi_{i-1}(\widehat{\theta}) \approx (\widehat{\theta} - \theta_0)^T \partial_\theta \psi_{i-1}(\theta_0) + r_{i-1}(\widehat{\theta})$, using that now $\partial_\theta \psi_{i-1}(\theta_0)$ is $\mathcal{H}_{t_{i-1}}$ measurable, and that $r_{i-1}(\widehat{\theta})$ is of order N^{-2}.

3.2 Noise parameter estimation

Several estimators have been proposed by Li et al. (2016), Chaker (2017), Clinet and Potiron (2019b) in different settings when we assume a null residual noise $\epsilon_t = 0$. The estimator from Chaker (2017) coincides with the minimum mean square error (MSE) estimator from Li et al. (2016) when ϕ is linear (which is the related assumption of the former paper). Moreover, the quasi maximum likelihood estimation (QMLE) from Clinet and Potiron (2019b) reduces to the MSE, due to the Gaussian form of the quasi likelihood function. Accordingly, we review the MSE procedure below and give the related limit theory for the noise parameter estimator.

We assume that $\theta := (\theta_0^{(1)}, \dots, \theta_0^{(d)})$, where for each component $k = 1, \dots, d$ we have $\theta_0^{(k)} := (\theta_0^{(k,1)}, \dots, \theta_0^{(k,l_k)})$, which corresponds to the parameter related to the kth component of the observed price. More specifically, we assume the componentwise form

$$Z_{t_i}^{(k)} = X_{t_i}^{(k)} + \phi\left(Q_{t_i}^{(k)}, \theta_0^{(k)}\right). \tag{21}$$

Accordingly, we consider the estimation of $\theta_0^{(k)}$ separately and thus we can assume that $d = 1$ in what follows without loss of generality. The estimator $\widehat{\theta}^{(MSE)}$ is given by

$$\widehat{\theta}^{(MSE)} = \underset{\theta \in \Theta}{\text{argmin}}\ Q_N(Z,\theta), \quad \text{where}$$
$$Q_N(Z,\theta) = \frac{1}{2}\sum_{i=1}^{N}\left(\Delta_i Z - \mu_i(\theta)\right)^2,$$

where $\mu_i(\theta) = \phi(Q_{t_i},\theta) - \phi(Q_{t_{i-1}},\theta)$.

When ϕ is linear, the problem boils down to a linear regression. As a result the estimator admits the explicit form

$$\widehat{\theta}^{(MSE)} = \left(\mathbb{M}^{\mathrm{T}}\mathbb{M}\right)^{-1}\mathbb{M}^{T}\Delta Z, \tag{22}$$

where $\Delta Z := \left(\Delta_1 Z, \dots, \Delta_N Z\right)$, and as soon as the matrix

$$\mathbb{M} := \left(\Delta_i Q^{(j)}\right)_{1\leq i\leq N, 1\leq j\leq l}$$

is such that $\mathbb{M}^T\mathbb{M}$ is invertible.

We now recall the limit theory associated with $\widehat{\theta}^{(MSE)}$ under the framework of Li et al. (2016) which in particular includes jumps with infinite activity. In the next proposition, Condition **A** assumes the local boundedness of b and σ, the summability of the jump process, and several standard identifiability assumptions of most functions which depend on the parameter θ and the sequence $(Q_{t_i})_{i\in\mathbb{N}}$. Details can be found in Li et al. (2016, p. 35).

Proposition 1 [Theorem 1 from Li et al. (2016)] *Assume Condition* **A** *from* Li et al. (2016). *Then*

$$N\left(\widehat{\theta}^{(MSE)} - \theta_0\right) = O_{\mathbb{P}}(1).$$

4 Applications of the method

In what follows, we state that the plug-in estimators are noise-robust for five leading examples taken from the literature, and that the central limit theorems (9) and (10) hold under parametric noise. In Example 4.1, we study the threshold realized volatility in the case of infinite activity jumps in price and endogeneity in arrival times. We go one step further the central limit theory of realized volatility with in Li et al. (2014), which includes no jumps when there is endogeneity in observation times, to allow for jumps with infinite activity. We first state the central limit theorems related to threshold realized volatility, and then the theory associated with the plug-in estimators. In Example 4.2, we consider the threshold bipower variation under infinite activity jumps and regular observations. In Example 4.3, we develop the Hayashi–Yoshida estimator of high-frequency covariance in a no-jump setup, and asynchronous and endogenous observation times. In Example 4.4, we consider the estimation of functionals of volatility when the price can exhibit jumps with infinite activity and observations are regular. Finally, we address the case of volatility of volatility for continuous price and volatility processes and regular observation times in Example 4.5.

4.1 Threshold realized volatility

The parameter is $\xi_t = \sigma_t^2$, and the rate of convergence $\kappa = 1/2$ if observations are not contaminated by the noise. When the price is continuous and observations are regular, a popular estimator of $\Xi = \int_0^T \sigma_s^2 \mathrm{d}s$ is RV considered in Andersen et al. (2001a, b), but also in Barndorff-Nielsen and Shephard (2002a, b), Meddahi (2002), etc. Jacod and Protter (1998) showed that

$$n^{1/2}\left(RV - \int_0^T \sigma_s^2 \mathrm{d}s\right) \to \mathcal{MN}\left(0, 2T\int_0^T \sigma_s^4 \mathrm{d}s\right).$$

When observations are not regular, the AVAR is equal to $2T\int_0^T \sigma_t^4 \mathrm{d}H_t$, where $H_t = \lim T^{-1} N \sum_{t_i \leq t}(t_i - t_{i-1})^2$ is the so-called quadratic variation of time (see Zhang 2001; Mykland and Zhang 2006), provided that such a quantity exists. When observations are endogenous, Li et al. (2014) show that the limit distribution of $n^{1/2}(RV - \Xi)$ includes an asymptotic bias and that the related AVAR is altered. In addition, they prove that the informational content of arrival times can be useful to estimate the asymptotic bias and the AVAR.

Our aim is to allow for parametric noise in this endogenous setting, while also including jumps in the price process. As far as the authors know, no general theory[4] includes general endogeneity and jumps, even when observations are not noisy. Accordingly, we first extend the results of Li et al. (2014) when adding jumps. Then, we show that the technology of this paper applies in such a general setting, and this part essentially boils down to applying the arguments of Li et al. (2016).

Although no theory exists under endogeneity, Theorem 13.2.4 (p. 383) in Jacod and Protter (2011) can be used when observations are regular. We consider a similar threshold RV, originally in the spirit of Mancini (2009, 2011), and defined as $\widetilde{\Xi} = \sum_{i=1}^N \left(\Delta_i X\right)^2 \mathbf{1}_{\{|\Delta_i X| \leq w_i\}}$, where $w_i = \alpha \Delta_i t^{\bar{\varpi}}$, $\bar{\varpi} \in (1/(2(2-r)), 1/2)$ and $\alpha > 0$ is a tuning parameter. In the next theorem, we provide the related central limit theorem and show that the condition of our paper holds.

Theorem 2 *We assume that* $\inf_{t\in(0,T]} \sigma_t > 0$. *We further suppose that there exists non-random* $\widetilde{u}_t$ *and* $\widetilde{v}_t$ *such that*

$$n \sum_{0<t_i\leq t} \left(\Delta_i X'\right)^4 \to^{\mathbb{P}} \int_0^t \widetilde{u}_s \sigma_s^4 \mathrm{d}s, \tag{23}$$

[4] Remark 6 (p. 36) in Li et al. (2016) suggests that the threshold RV estimator can be used under endogeneity, but there is no formal proof and this is limited to the case of jumps with finite activity.

$$n^{1/2} \sum_{0<t_i \le t} \left(\Delta_i X'\right)^3 \to^{\mathbb{P}} \int_0^t \widetilde{v}_s \sigma_s^3 \mathrm{d}s, \tag{24}$$

where $\widetilde{u}_t\sigma_t^4$, $\widetilde{v}_t\sigma_t^3$ *and* $\widetilde{v}_t^2\sigma_t^4$ *are integrable, and* $\widetilde{v}_t$ *locally bounded and bounded away from 0. Furthermore, we assume that* t_i, b_t, σ_t and δ *are generated by finitely many Brownian motions.*[5] *Finally we assume that* $N/n \to^{\mathbb{P}} F$ *for some random variable* F, *and that* $n\Delta_i t$ *are locally bounded and locally bounded away from 0. Then, stably in law as* $n \to \infty$, *we have*

$$N^{1/2}\left(\widetilde{\Xi} - \Xi\right) \to \frac{2}{3}\int_0^T v_s\sigma_s \mathrm{d}X'_s + \int_0^T \sqrt{\frac{2}{3}u_s - \frac{4}{9}v_s^2}\sigma_s^2 \mathrm{d}B_s, \tag{25}$$

where $v_s = \sqrt{F}\widetilde{v}_s$, $u_s = F\widetilde{u}_s$ *and* B_t *is a standard Brownian motion independent of the other quantities.*[6] *Moreover, we have*

$$N^{1/2}\left(\widehat{\Xi} - \Xi\right) \to \frac{2}{3}\int_0^T v_s\sigma_s \mathrm{d}X'_s + \int_0^T \sqrt{\frac{2}{3}u_s - \frac{4}{9}v_s^2}\sigma_s^2 \mathrm{d}B_s. \tag{26}$$

Remark 3 If observations are regular, then $F = 1$, $u_s = 3T$ and $v_s = 0$ for all $s \in [0, T]$. Therefore, (25) and (26) can be specified as

$$n^{1/2}\left(\widetilde{\Xi} - \Xi\right) \to \mathcal{MN}\left(0, 2T\int_0^T \sigma_s^4 \mathrm{d}s\right), \tag{27}$$

$$n^{1/2}\left(\widehat{\Xi} - \Xi\right) \to \mathcal{MN}\left(0, 2T\int_0^T \sigma_s^4 \mathrm{d}s\right). \tag{28}$$

We provide now jump-robust estimators of $AB = (2/3)\int_0^T v_s\sigma_s \mathrm{d}X'_s$ and $AVAR = \int_0^T (\frac{2}{3}u_s - \frac{4}{9}v_s^2)\sigma_s^4 \mathrm{d}s$ based on the non-jump-robust estimators provided in Li et al. (2014). Accordingly, we chop the data into B blocks of h observations (except for the last block which might include less observations). We set $h = \lfloor n^\beta \rfloor$, where $1/2 < \beta < 1$. We can estimate $v_{t_{hi}}\sigma_{t_{hi}}$ as

$$\widetilde{v\sigma}_i = \frac{N^{1/2} \sum_{j=h(i-1)+1}^{hi} \left(\Delta_j X\right)^3 \mathbf{1}_{\{|\Delta_j X| \le w_j\}}}{\sum_{j=h(i-1)+1}^{hi} \left(\Delta_j X\right)^2 \mathbf{1}_{\{|\Delta_j X| \le w_j\}}},$$

and AB and AVAR as

[5] i.e. we assume that t_i are $\mathbf{G}$-stopping times, where $\mathbf{G} = (\mathcal{G}_t)_{t\in[0,T]}$ is a sub-filtration of $\mathbf{F}$ generated by finitely many Brownian motions, and that b_t, σ_t and δ are adapted to $\mathbf{G}$.

[6] Here and in the other theorems, we mean that B_t is independent of the underlying σ-field $\mathbf{F}$.

$$\widetilde{AB} = \sum_{i=1}^{B} \underbrace{\frac{2}{3}\widetilde{v\sigma}_i \left\{ \sum_{j=h(i-1)+1}^{hi} \Delta_j X \mathbf{1}_{\{|\Delta_j X| \leq w_j\}} \right\}}_{\widetilde{AB}_i},$$

$$\widetilde{AVAR} = \frac{2N}{3} \sum_{i=1}^{N} \left(\Delta_i X\right)^4 \mathbf{1}_{\{|\Delta_i X| \leq w_i\}} - \sum_{i=1}^{B} \widetilde{AB}_i^2.$$

Recalling that $\widehat{AB}$ and $\widehat{AVAR}$ are constructed, respectively, as $\widetilde{AB}$ and $\widetilde{AVAR}$ when replacing X by $\widehat{X}$, we provide now the studentized version of the previous central limit theorems.

Corollary 4 *We have*

$$N^{1/2} \frac{\widetilde{\Xi} - N^{-1/2}\widetilde{AB} - \Xi}{\sqrt{\widetilde{AVAR}}} \rightarrow \mathcal{N}(0, 1), \tag{29}$$

$$N^{1/2} \frac{\widehat{\Xi} - N^{-1/2}\widehat{AB} - \Xi}{\sqrt{\widehat{AVAR}}} \rightarrow \mathcal{N}(0, 1). \tag{30}$$

Remark 5 If observations are regular, there is no asymptotic bias and $AVAR$ can be estimated using the plug-in estimator of quarticity obtained in Sect. 4.4. In view of Theorem 11 which implies the consistency of the plug-in estimator, we obtain directly by the stable convergence obtained in Theorem 2 that (30) holds.

Remark 6 (Estimating volatility under i.i.d noise) Alternative approaches to estimate integrated volatility under latent i.i.d noise include and are not limited to: the Quasi-Maximum Likelihood Estimator (QMLE) from Aït-Sahalia et al. (2005) which was later shown to be robust to time-varying volatility in Xiu (2010), the two-scale realized volatility in Zhang et al. (2005), the multi-scale realized volatility in Zhang (2006), the pre-averaging approach in Jacod et al. (2009), realized kernels in Barndorff-Nielsen et al. (2008) and the spectral approach considered in Altmeyer and Bibinger (2015) based on Reiss (2011). Clinet and Potiron (2018) discussed AVAR reduction when considering local estimators. In addition, Li et al. (2013) consider endogenous arrival times.

4.2 Threshold bipower variation

Here again $\xi_t = \sigma_t^2$. The bipower variation $BV = \frac{\pi}{2} \sum_{i=2}^{N} | \Delta_i X || \Delta_{i-1} X |$ [more generally multipower variation from Barndorff-Nielsen and Shephard (2004, 2006)] was originally introduced as an alternative measure robust to finite-activity jumps. In case of regular observations and no jump, Barndorff-Nielsen et al. (2006a, b) established the central limit theory. See also Kinnebrock and Podolskij (2008) for related development. In case of finite-activity jumps, see also Barndorff-Nielsen et al. (2006c).

If jumps exhibit infinite activity, Vetter (2010) shows that BV is no longer consistent, but the jump-robust threshold estimator

$$\widetilde{\Xi} = \frac{\pi}{2} \sum_{i=2}^{N} \mid \Delta_i X \mid \mathbf{1}_{\{|\Delta_i X| \leq w\}} \mid \Delta_{i-1} X \mid \mathbf{1}_{\{|\Delta_{i-1} X| \leq w\}}$$

is consistent, where $w = \alpha \Delta^{\bar{\varpi}}$, $\bar{\varpi} \in (0, 1/2)$. Moreover, he also shows the related central limit theory. See also Corsi et al. (2010) for related work. Finally, the general theory (Theorem 13.2.1 (p. 380)) from Jacod and Protter (2011) can be applied too. All those papers have in common that they assume regular observations, and we follow the same setting to show that the techniques of this paper can be used in this example too. We provide the formal result in what follows.

Theorem 7 *We have that*

$$n^{1/2}\left(\widehat{\Xi} - \widetilde{\Xi}\right) \to^{\mathbb{P}} 0. \tag{31}$$

In particular, stably in law as $n \to \infty$,

$$n^{1/2}\left(\widehat{\Xi} - \Xi\right) \to \frac{\pi}{2}\sqrt{\left(1 + \frac{4}{\pi} - \frac{12}{\pi^2}\right)T} \int_0^T \sigma_s^2 \mathrm{d}B_s, \tag{32}$$

where B_t is a Brownian motion independent of the other quantities.

In this example, we have that $AVAR = \frac{\pi^2}{4}(1 + \frac{4}{\pi} - \frac{12}{\pi^2})T \int_0^T \sigma_s^4 \mathrm{d}s$, which can be estimated by $\widehat{AVAR} = \frac{\pi^2}{4}(1 + \frac{4}{\pi} - \frac{12}{\pi^2})T \widehat{\int_0^T \sigma_s^4 \mathrm{d}s}$, where the plug-in estimator of quarticity $\widehat{\int_0^T \sigma_s^4 \mathrm{d}s}$ is defined as a particular case of Sect. 4.4 (i.e. $\widehat{\int_0^T \sigma_s^4 \mathrm{d}s}$ corresponds to the estimator given in (39) below with $g(x) = x^2$). We also provide the related studentized central limit theorem.

Corollary 8 *We have*

$$n^{1/2} \frac{\widehat{\Xi} - \Xi}{\sqrt{\widehat{AVAR}}} \to \mathcal{N}(0, 1). \tag{33}$$

4.3 Hayashi–Yoshida estimator of high-frequency covariance

We assume here that X_t is 2-dimensional and that $\xi_t = \rho_t \sigma_t^{(1)} \sigma_t^{(2)}$, where the high-frequency correlation ρ_t satisfies $d\langle W^{(1)}, W^{(2)} \rangle_t = \rho_t \mathrm{d}t$. The rate of convergence is $\kappa = 1/2$ in this problem too. We consider that observations are non-synchronous. In this framework and assuming that the price is continuous, Hayashi and Yoshida

(2005) bring forward the so-called Hayashi–Yoshida estimator and establish the consistency in case sampling times are independent from the price process. This is extended in an endogenous setting in Hayashi and Kusuoka (2008). The related central limit theory can be found in Hayashi and Yoshida (2008, 2011) and Potiron and Mykland (2017), where the latter work considers general endogenous arrival times. See also the remarkable work from Bibinger and Vetter (2015) and Martin and Vetter (2020) in a jumpy setting, and Koike (2014a, b, 2016) which incorporates jumps, noise and some kind of endogeneity into the model.

As we want to allow for quite exotic endogenous models, we follow Potiron and Mykland (2017). In particular, we assume no jumps in the setup. We describe the hitting boundary with time process (HBT) model introduced in the subsequent paper in what follows. In that model, eight stochastic processes (four of which are actually families of stochastic processes) are of interest, four for each asset. For the index $k = 1, 2$, we have the price process—$X_t^{(k)}$—and three other stochastic processes (two of which are actually families of processes)—$Y_t^{(k)}$, $d_t^{(k)}(s)$ and $u_t^{(k)}(s)$—related to the observation times of that process. Those four stochastic processes can be correlated, and we further assume that (X_t, Y_t) is a 4-dimensional Itô-process. For the process $k = 1, 2$, $Y_t^{(k)}$ stands for the continuous observation time process which drives the observation times related to $X_t^{(k)}$. The others four processes are the down processes $d_t^{(k)}(s)$ and the up processes $u_t^{(k)}(s)$. We assume that the down process takes only negative values and that the up process takes only positive values. A new observation time will be generated whenever one of those two processes is hit by the increment of the observation time process. Then, the increment of the observation time process will be reset to 0, and the next observation time will be produced whenever the up or the down process is hit again. Formally, if we let $\alpha > 0$ stand for the tick size, we define the first observation time as $t_0^{(k)} := 0$ and recursively $t_i^{(k)}$ as

$$t_i^{(k)} := \inf\left\{ t > t_{i-1}^{(k)} \ : \ \Delta Y^{(k)}_{\left[t_{i-1}^{(k)}, t\right]} \notin \left[\alpha d_t^{(k)}\left(t - t_{i-1}^{(k)}\right), \alpha u_t^{(k)}\left(t - t_{i-1}^{(k)}\right)\right] \right\}, \tag{34}$$

where $\Delta Y^{(k)}_{[a,b]} := Y_b^{(k)} - Y_a^{(k)}$. We define the Hayashi–Yoshida estimator as

$$\widetilde{\Xi} := \sum_{0 < t_i^{(1)},\, t_j^{(2)} < T} \Delta_i X^{(1)} \Delta_j X^{(2)} \mathbf{1}_{\left\{ \left[t_{i-1}^{(1)}, t_i^{(1)}\right) \cap \left[t_{j-1}^{(2)}, t_j^{(2)}\right) \neq \emptyset \right\}}. \tag{35}$$

In the asymptotic theory, we let $\alpha \to 0$. For the sake of Remark 5 (p. 25) in Potiron and Mykland (2017), α^{-1} is of the same order as $n^{1/2}$. We can now show that the techniques of this paper hold in this case too.

Theorem 9 *As the tick size $\alpha \to 0$, we have that*

$$\alpha^{-1}\left(\widehat{\Xi} - \widetilde{\Xi}\right) \to^{\mathbb{P}} 0. \tag{36}$$

In particular, under the assumptions of Potiron and Mykland (2017), *there exist AB and a process AV_t such that stably in law as the tick size $\alpha \to 0$,*

$$\alpha^{-1}\left(\widehat{\Xi} - \Xi\right) \to AB + \int_0^T \left(AV_s\right)^{1/2} \mathrm{d}B_s, \tag{37}$$

where B_t is a Brownian motion independent of the other quantities, AB and AV_t are defined in Section 4.3 of Potiron and Mykland (2017).

We define $\widetilde{AB}$ and $\widetilde{AVAR}$ following, respectively, (46) and (47) in Potiron and Mykland (2017) (Section 5, p. 28). Note that $\widetilde{AB}$ and $\widetilde{AVAR}$ are already of the right asymptotic order in the sense that $\alpha^{-1}\widetilde{AB} \to^{\mathbb{P}} AB$ and $\alpha^{-2}\widetilde{AVAR} \to^{\mathbb{P}} \int_0^T AV_s \mathrm{d}s$ (see (48) and (49) in Corollary 4 of the cited paper). We provide now the studentized version of (37).

Corollary 10 *We have*

$$\frac{\widehat{\Xi} - \widehat{AB} - \Xi}{\sqrt{\widehat{AVAR}}} \to \mathcal{N}(0, 1). \tag{38}$$

4.4 Functionals of volatility local estimator

The spot parameter is $\xi_t = g(c_t)$ for a given smooth function g on $\mathcal{M}_d^+$, the set of all non-negative symmetric $d \times d$ matrices. The problem was initiated by Barndorff-Nielsen and Shephard (2002a). See also Barndorff-Nielsen et al. (2006a), Mykland and Zhang (2012) (Proposition 2.17, p. 138) and Renault et al. (2017) for related developments. Here, the rate of convergence is $\kappa = 1/2$ again.

Local estimation (Mykland and Zhang 2009, Section 4.1, pp. 1421–1426) can make the mentioned estimators efficient. Jacod and Rosenbaum (2013) extended the method in several ways. To do that, they first propose an estimator of the spot volatility $\widetilde{c}_i$, and then take a Riemann sum of $g(\widetilde{c}_i)$.

For any matrix $a \in \mathcal{M}_d^+$, the related a^{ij} stands for the (i, j)-component of a. Moreover, for $b \in \mathbb{R}$, $[b]$ stands for the floor of b. Several results are of interest in Jacod and Rosenbaum (2013). In its most useful form (from our point of view), the estimator takes on the form

$$\widetilde{\Xi} = \Delta \sum_{i=1}^{[T/\Delta]-k+1} \left\{ g(\widetilde{c}_i) - \frac{1}{2k} \sum_{j,q,l,m=1}^{d} \partial^2_{jq,lm} g(\widetilde{c}_i) \left(\widetilde{c}_i^{jl} \widetilde{c}_i^{qm} + \widetilde{c}_i^{jm} \widetilde{c}_i^{ql} \right) \right\}, \tag{39}$$

with

$$\widetilde{c}_i^{lm} = \frac{1}{k\Delta} \sum_{j=0}^{k-1} \Delta_{i+j} X^l \Delta_{i+j} X^m \mathbf{1}_{\{\|\Delta_{i+j} X\| \le w\}},$$

for two sequences of integers k and $w = \alpha \Delta^{\bar{\omega}}$ for some $\alpha > 0$, and

$$\frac{2p-1}{2(2p-r)} \leq \bar{\varpi} < \frac{1}{2},$$

where we suppose that

$$\left\| \partial^j g(x) \right\| \leq K\big(1 + \|x\|^{p-j}\big), \;\; j = 0, 1, 2, 3 \tag{40}$$

for some constants $p \geq 3$, $K > 0$. In Eq. (39), $\widetilde{c}_i$ corresponds to an estimator of the spot volatility matrix, the first term is part of the Riemann sum, while the second term is required to remove the asymptotic bias of the first term in $\widetilde{\Xi}$, which explodes asymptotically. We show that the associated plug-in estimator $\widehat{\Xi}$ enjoys the same limit theory as $\widetilde{\Xi}$. More precisely, we have the following result.

Theorem 11 *Assume that $k^2\Delta \to 0$, $k^3\Delta \to \infty$. Let $\widetilde{\Xi}'$ be the estimator defined as in (39) where X_t is replaced by its continuous part X_t'. Then, we have the convergence*

$$n^{1/2}\Big(\widehat{\Xi} - \widetilde{\Xi}'\Big) \to^{\mathbb{P}} 0. \tag{41}$$

Moreover, stably in law, we have the convergence

$$n^{1/2}\Big(\widehat{\Xi} - \Xi\Big) \to \int_0^T \sqrt{T\overline{h}(c_s)}\mathrm{d}B_s, \tag{42}$$

where for $x \in \mathcal{M}_d^+$,

$$\overline{h}(x) = \sum_{j,q,l,m=1}^{d} \partial_{jq} g(x) \partial_{lm} g(x) \big(x^{jl}x^{qm} + x^{jm}x^{ql}\big),$$

and where B is a standard Brownian motion independent of the other quantities.

In particular, note that the asymptotic variance in the stable convergence can be expressed as

$$AVAR = T \int_0^T \overline{h}(c_s)\mathrm{d}s,$$

so that we naturally define the asymptotic variance estimator as

$$\widehat{AVAR} = T\Delta \sum_{i=1}^{[t/\Delta]-k+1} \overline{h}\big(\widehat{c}_i\big).$$

We easily deduce from Corollary 3.7 p. 1471 in Jacod and Rosenbaum (2013) the following studentized version of the above central limit theorem.

Corollary 12 *Under the assumptions of the previous theorem, we have the stable convergence in law*

$$\frac{n^{1/2}\left(\widehat{\Xi}-\Xi\right)}{\sqrt{\widehat{AVAR}}} \to \mathcal{N}(0,1).$$

Remark 13 (Estimation of functionals of volatility under i.i.d noise) Under i.i.d noise, no result with a general function $g(c_t)$ is available. Alternative approaches include: Jacod et al. (2010) for even power, Mancino and Sanfelici (2012) and also Andersen et al. (2014) in the special case of quarticity, and also Altmeyer and Bibinger (2015) when considering the tricity. See also the work from Potiron and Mykland (2016) (Section 4.2) for a local maximum-likelihood estimation with noise variance vanishing asymptotically.

4.5 Volatility of volatility

In this section we assume that X_t is 1-dimensional and we are interested in the spot parameter $\xi_t = \widetilde{\sigma}_t^2$ which corresponds to the so-called volatility of volatility process defined in (14). As far as we know, there is no result in the literature including noise into the model, but in the non-noisy scenario one can consult Vetter (2015) (Theorems 2.5 and 2.6) and Mykland et al. (2012) (Theorem 7 and Corollary 2). We follow here the former author, and aim to show the robustness of Theorem 2.6 when using plug-in estimators. Accordingly, we hereafter assume that both X_t and c_t are continuous processes, i.e. $\delta = \widetilde{\delta} = 0$ in (13)–(14). To our knowledge, the case with jumps in X_t and/or c_t remains an open question. The rate of convergence is $\kappa = 1/4$. Introducing the spot volatility estimator[7] for $i \in \{0, \dots, n-k\}$,

$$\widetilde{c}_i := \frac{n}{k}\sum_{j=1}^{k}\left(\Delta_{i+j}X\right)^2,$$

and the spot quarticity estimator

$$\widetilde{q}_i := \frac{n^2}{3k}\sum_{j=1}^{k}\left(\Delta_{i+j}X\right)^4,$$

the author defines the volatility of volatility estimator (see (2.5) on p. 2399 in the cited work) as

$$\widetilde{\Xi} := \sum_{i=0}^{[t/\Delta]-2k}\left\{\frac{3}{2k}\left(\widetilde{c}_{i+k}-\widetilde{c}_i\right)^2 - \frac{6}{k^2}\widetilde{q}_i\right\}.$$

Letting $\widehat{c}_i$, $\widehat{q}_i$, and $\widehat{\Xi}$ be the corresponding plug-in estimators, we obtain the following results.

[7] Note that the definition of $\widetilde{c}_i$ slightly diverges from the previous section.

Theorem 14 *Assume that $k = cn^{1/2} + o(n^{1/4})$ for some $c > 0$. Then stably in law,*

$$\sqrt{\frac{n}{k}}\left(\widehat{\Xi} - \Xi\right) \to \sqrt{T}\int_0^T \alpha_s \mathrm{d}B_s,$$

where B_t is a Brownian motion independent from the other quantities and

$$\alpha_s^2 = \frac{48}{c^4}\sigma_s^8 + \frac{12}{c^2}\sigma_s^4\widetilde{\sigma}_s^2 + \frac{151}{70}\widetilde{\sigma}_s^4.$$

Moreover, if we define

$$\begin{aligned} G^{(1)} &= \frac{T}{n}\sum_{i=0}^{[t/\Delta]-k} \widehat{q}_i^2, \\ G^{(2)} &= T\sum_{i=0}^{[t/\Delta]-2k} \left\{\frac{3}{2k}\left(\widehat{c}_{i+k} - \widehat{c}_i\right)^2 - \frac{6}{k^2}\widehat{q}_i\right\}\widehat{q}_i, \\ G^{(3)} &= \frac{Tn}{k^2}\sum_{i=0}^{[t/\Delta]-2k} \left(\widehat{c}_{i+k} - \widehat{c}_i\right)^4, \end{aligned}$$

and finally

$$\widehat{AVAR} = \frac{453}{280}G^{(3)} - \frac{n}{k^2}\frac{486}{35}G^{(2)} - \frac{n^2}{k^4}\frac{1038}{35}G^{(1)},$$

we can derive the following studentized version of the previous central limit theorem.

Corollary 15 *Under the assumptions of the previous theorem, we have the stable convergence in law, when k has the optimal rate $c\sqrt{n}$ for $c > 0$*

$$n^{1/4}\frac{\widehat{\Xi} - \Xi}{\sqrt{c\widehat{AVAR}}} \to \mathcal{N}(0, 1).$$

5 Conclusion

This paper develops plug-in estimators to estimate high-frequency quantities under parametric noise on five different examples. We do not find any particular difficulty when working out the theory of those examples. Another example of application can be found in Andersen et al. (2019).

Acknowledgements We would like to thank Selma Chaker, Yingying Li, Xinghua Zheng, Manh Cuong Pham, Mathias Vetter, two anonymous referees, the participants of the 2nd International Conference on Econometrics and Statistics in Hong Kong, and the Econometric Society Australasian Meeting 2018 in Auckland for helpful discussions and advice. The research of Yoann Potiron is supported by a special private grant from Keio University and Japanese Society for the Promotion of Science Grant-in-Aid for

Young Scientists No. 60781119. The research of Simon Clinet is supported by Japanese Society for the Promotion of Science Grant-in-Aid for Young Scientists No. 19K13671.

References

Aït-Sahalia, Y., Mykland, P. A., Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies*, *18*(2), 351–416.

Almgren, R., Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, *3*, 5–40.

Altmeyer, R., Bibinger, M. (2015). Functional stable limit theorems for quasi-efficient spectral covolatility estimators. *Stochastic Processes and Their Applications*, *125*(12), 4556–4600.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics*, *61*(1), 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, *96*(453), 42–55.

Andersen, T. G., Dobrev, D., Schaumburg, E. (2014). A robust neighborhood truncation approach to estimation of integrated quarticity. *Econometric Theory*, *30*(1), 3–59.

Andersen, T. G., Cebiroglu, G., Hautsch, N. (2017). *Volatility, information feedback and market microstructure noise*: *A tale of two regimes*. Retrieved 9 May 2017. Working paper available at https://www.paperssrncom/sol3/paperscfm?abstract_id=2921097.

Andersen, T. G., Thyrsgaard, M., Todorov, V. (2019). Time-varying periodicity in intraday volatility. *Journal of the American Statistical Association*, *114*(528), 1695–1707.

Barndorff-Nielsen, O. E., Shephard, N. (2002a). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society Series B (Methodological)*, *64*(2), 253–280.

Barndorff-Nielsen, O. E., Shephard, N. (2002b). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, *17*(5), 457–477.

Barndorff-Nielsen, O. E., Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, *2*(1), 1–37.

Barndorff-Nielsen, O. E., Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, *4*(1), 1–30.

Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J., Podolskij, M., Shephard, N. (2006a). A central limit theorem for realised power and bipower variations of continuous semimartingales. *From stochastic calculus to mathematical finance*, pp. 33–68. Berlin: Springer.

Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J., Shephard, N. (2006b). Limit theorems for bipower variation in financial econometrics. *Econometric Theory*, *22*(4), 677–719.

Barndorff-Nielsen, O. E., Shephard, N., Winkel, M. (2006c). Limit theorems for multipower variation in the presence of jumps. *Stochastic Processes and Their Applications*, *116*(5), 796–806.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, *76*(6), 1481–1536.

Bibinger, M., Vetter, M. (2015). Estimating the quadratic covariation of an asynchronously observed semimartingale with jumps. *Annals of the Institute of Statistical Mathematics*, *67*(4), 707–743.

Black, F. (1986). Noise. *The Journal of Finance*, *41*(3), 528–543.

Chaker, S. (2017). On high frequency estimation of the frictionless price: The use of observed liquidity variables. *Journal of Econometrics*, *201*(1), 127–143.

Clinet, S., Potiron, Y. (2018). Efficient asymptotic variance reduction when estimating volatility in high frequency data. *Journal of Econometrics*, *206*(1), 103–142.

Clinet, S., Potiron, Y. (2019a). Disentangling sources of high frequency market microstructure noise. *Journal of Business & Economic Statistics*, https://doi.org/10.1080/07350015.2019.1617158.

Clinet, S., Potiron, Y. (2019b). Testing if the market microstructure noise is fully explained by the informational content of some variables from the limit order book. *Journal of Econometrics*, *209*(2), 289–337.

Cont, R., Kukanov, A., Stoikov, S. (2014). The price impact of order book events. *Journal of Financial Econometrics*, *12*(1), 47–88.

Corsi, F., Pirino, D., Reno, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, *159*(2), 276–288.

Diebold, F. X., Strasser, G. (2013). On the correlation structure of microstructure noise: A financial economic approach. *The Review of Economic Studies*, *80*(4), 1304–1337.

Glosten, L. R., Harris, L. E. (1988). Estimating the components of the bid/ask spread. *Journal of Financial Economics*, *21*(1), 123–142.

Hasbrouck, J. (1993). Assessing the quality of a security market: A new approach to transaction-cost measurement. *The Review of Financial Studies*, *6*(1), 191–212.

Hasbrouck, J. (2002). Stalking the "efficient price" in market microstructure specifications: An overview. *Journal of Financial Markets*, *5*(3), 329–339.

Hasbrouck, J. (2007). *Empirical market microstructure*: *The institutions, economics, and econometrics of securities trading*. New York: Oxford University Press.

Hayashi, T., Kusuoka, S. (2008). Consistent estimation of covariation under nonsynchronicity. *Statistical Inference for Stochastic Processes*, *11*(1), 93–106.

Hayashi, T., Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli*, *11*(2), 359–379.

Hayashi, T., Yoshida, N. (2008). Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. *Annals of the Institute of Statistical Mathematics*, *60*(2), 367–406.

Hayashi, T., Yoshida, N. (2011). Nonsynchronous covariation process and limit theorems. *Stochastic Processes and Their Applications*, *121*(10), 2416–2454.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., Vetter, M. (2009). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and Their Applications*, *119*(7), 2249–2276.

Jacod, J., Podolskij, M., Vetter, M. (2010). Limit theorems for moving averages of discretized processes plus noise. *Annals of Statistics*, *38*(3), 1478–1545.

Jacod, J., Protter, P. (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability*, *26*, 267–307.

Jacod, J., Protter, P. E. (2011). *Discretization of processes*. Berlin: Springer.

Jacod, J., Reiss, M. (2014). A remark on the rates of convergence for integrated volatility estimation in the presence of jumps. *Annals of Statistics*, *42*(3), 1131–1144.

Jacod, J., Rosenbaum, M. (2013). Quarticity and other functionals of volatility: Efficient estimation. *Annals of Statistics*, *41*(3), 1462–1484.

Kavajecz, K. A. (1999). A specialist's quoted depth and the limit order book. *The Journal of Finance*, *54*(2), 747–771.

Kinnebrock, S., Podolskij, M. (2008). A note on the central limit theorem for bipower variation of general functions. *Stochastic Processes and Their Applications*, *118*(6), 1056–1070.

Koike, Y. (2014a). An estimator for the cumulative co-volatility of asynchronously observed semimartingales with jumps. *Scandinavian Journal of Statistics*, *41*(2), 460–481.

Koike, Y. (2014b). Limit theorems for the pre-averaged Hayashi–Yoshida estimator with random sampling. *Stochastic Processes and Their Applications*, *124*(8), 2699–2753.

Koike, Y. (2016). Estimation of integrated covariances in the simultaneous presence of nonsynchronicity, microstructure noise and jumps. *Econometric Theory*, *32*(3), 533–611.

Li, Y., Zhang, Z., Zheng, X. (2013). Volatility inference in the presence of both endogenous time and microstructure noise. *Stochastic Processes and Their Applications*, *123*(7), 2696–2727.

Li, Y., Mykland, P. A., Renault, E., Zhang, L., Zheng, X. (2014). Realized volatility when sampling times are possibly endogenous. *Econometric Theory*, *30*(03), 580–605.

Li, Y., Xie, S., Zheng, X. (2016). Efficient estimation of integrated volatility incorporating trading information. *Journal of Econometrics*, *195*(1), 33–50.

Madhavan, A. (2000). Market microstructure: A survey. *Journal of Financial Markets*, *3*(3), 205–258.

Madhavan, A., Richardson, M., Roomans, M. (1997). Why do security prices change? A transaction-level analysis of NYSE stocks. *The Review of Financial Studies*, *10*(4), 1035–1064.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, *36*(2), 270–296.

Mancini, C. (2011). The speed of convergence of the threshold estimator of integrated variance. *Stochastic Processes and Their Applications*, *121*(4), 845–855.

Mancino, M. E., Sanfelici, S. (2012). Estimation of quarticity with high-frequency data. *Quantitative Finance*, *12*(4), 607–622.

Martin, O., Vetter, M. (2020). The null hypothesis of (common) jumps in case of irregular and asynchronous observations. *Scandinavian Journal of Statistics*, *47*(3), 711–756.

Meddahi, N. (2002). A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics*, *17*(5), 479–508.
Mykland, P. A., Zhang, L. (2006). ANOVA for diffusions and Ito processes. *Annals of Statistics*, *34*(4), 1931–1963.
Mykland, P. A., Zhang, L. (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica*, *77*(5), 1403–1445.
Mykland, P. A., Zhang, L. (2012). The econometrics of high frequency data. In M. Kessler, A. Lindner, M. Sørensen (Eds.), *Statistical methods for stochastic differential equations*, pp. 109–190. Boca Raton, FL: Chapman and Hall/CRC Press.
Mykland, P. A., Shephard, N., Sheppard, K. (2012). Efficient and feasible inference for the components of financial variation using blocked multipower variation. Technical Report, University of Oxford, available at https://doi.org/10.2139/ssrn2008690.
O'hara, M. (1995). *Market microstructure theory*, Vol. 108. Cambridge, MA: Blackwell.
Potiron, Y., Mykland, P. A. (2016). *Local parametric estimation in high frequency data*. Working paper available at arXiv:160305700.
Potiron, Y., Mykland, P. A. (2017). Estimation of integrated quadratic covariation with endogenous sampling times. *Journal of Econometrics*, *197*(1), 20–41.
Reiss, M. (2011). Asymptotic equivalence for inference on the volatility from noisy observations. *Annals of Statistics*, *39*(2), 772–802.
Renault, E., Sarisoy, C., Werker, B. J. (2017). Efficient estimation of integrated volatility and related processes. *Econometric Theory*, *33*(2), 439–478.
Robert, C. Y., Rosenbaum, M. (2010). A new approach for the dynamics of ultra-high-frequency data: The model with uncertainty zones. *Journal of Financial Econometrics*, *9*(2), 344–366.
Robert, C. Y., Rosenbaum, M. (2012). Volatility and covariation estimation when microstructure noise and trading times are endogenous. *Mathematical Finance*, *22*(1), 133–164.
Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, *39*(4), 1127–1139.
Stoll, H. R. (2000). Presidential address: Friction. *The Journal of Finance*, *55*(4), 1479–1514.
Vetter, M. (2010). Limit theorems for bipower variation of semimartingales. *Stochastic Processes and Their Applications*, *120*(1), 22–38.
Vetter, M. (2015). Estimation of integrated volatility of volatility with applications to goodness-of-fit testing. *Bernoulli*, *21*(4), 2393–2418.
Xiu, D. (2010). Quasi-maximum likelihood estimation of volatility with high frequency data. *Journal of Econometrics*, *159*(1), 235–250.
Zhang, L. (2001). *From martingales to ANOVA*: *Implied and realized volatility*. PhD thesis, University of Chicago, Department of Statistics, Chicago.
Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, *12*(6), 1019–1043.
Zhang, L., Mykland, P. A., Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, *100*(472), 1394–1411.