Testing if the market microstructure noise is fully explained by the informational content of some variables from the limit order book[☆]

Simon Clinet [a,b], Yoann Potiron [c,*]

[a] Faculty of Economics, Keio University, 2-15-45 Mita, Minato-ku, Tokyo, 108-8345, Japan
[b] CREST, Japan Science and Technology Agency, Japan
[c] Faculty of Business and Commerce, Keio University, 2-15-45 Mita, Minato-ku, Tokyo, 108-8345, Japan

ARTICLE INFO

ABSTRACT

In this paper, we build tests for the presence of residual noise in a model where the market microstructure noise is a known parametric function of some variables from the limit order book. The tests compare two distinct quasi-maximum likelihood estimators of volatility, where the related model includes a residual noise in the market microstructure noise or not. The limit theory is investigated in a general nonparametric framework. In the presence of residual noise, we examine the central limit theory of the related quasi-maximum likelihood estimation approach.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

If one can sample directly from the efficient price, the estimation of volatility is a well-studied matter. The realized volatility (RV) estimator, i.e. summing the square of the log-returns, is both consistent and efficient. However, in practice the

observed price does not behave as expected. When sampling at high frequency, it can be quite different from the efficient price due to bid–ask bounce mechanism, the spread, the fact that transactions lie on a tick grid, etc. Market microstructure noise (MMN) typically degrades RV to the extent that it is highly biased when performed on tick by tick data.

One approach to overcome the problem consists in sub-sampling, say every 5 min, as in the pioneer work from Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002). On the contrary, Aït-Sahalia et al. (2005) recognize the MMN as an inherent part of the data and advice for the use of the Quasi-Maximum Likelihood Estimator (QMLE) which was later shown to be robust to time-varying volatility in Xiu (2010). Concurrent methods include and are not limited to: the Two-Scale Realized Volatility (TSRV) in Zhang et al. (2005), the Multi-Scale Realized Volatility in Zhang (2006), the Pre-averaging approach (PAE) in Jacod et al. (2009), realized kernels (RK) in Barndorff-Nielsen et al. (2008) and the spectral approach considered in Altmeyer and Bibinger (2015).

Those two approaches have cons in that the sub-sampling technique discards a large proportion of the data and the noise-robust estimators have a slower rate of convergence than RV in the absence of noise. The only exception is the QMLE proposed by Da and Xiu (2017), which is automatically adaptive to noise magnitude and always enjoys optimal rate. In this paper, we consider the MMN as a salient part of the financial market, but using the growing limit order book (LOB) big data available to the econometrician we ask a different question. Can we test whether the MMN is fully explained by the informational content of some variables from the limit order book and can we estimate the parameter of the model? If the MMN can be expressed as an observable function, then we can estimate the efficient price and use RV including all the data points. This idea is not new, and actually our work is heavily based on the two very nice papers Li et al. (2016) and Chaker (2017). We will explain the differences later in the introduction.

In fact it is rather natural to see the MMN as a function of some variables of the LOB as in the pioneer work from Roll (1984) where the trade type, i.e. whether the trade was buyer or seller initiated, is used to correct for the bid–ask bounce effect in the observed price. The related observed price Z_{t_i} is then defined as

$$\underbrace{Z_{t_i}}_{\text{observed price}} = \underbrace{X_{t_i}}_{\text{efficient price}} + \underbrace{I_i \theta_0}_{\text{MMN}}, \tag{1.1}$$

where θ_0 can be interpreted as one-half of the effective bid–ask spread, I_i is equal to 1 if the trade at time t_i is buyer-initiated and -1 if seller-initiated. A simple extension where the spread S_i is time-varying is given by

$$Z_{t_i} = X_{t_i} + \frac{1}{2} I_i S_i \theta_0. \tag{1.2}$$

Discussion and related leading models can be found in: Black (1986), Hasbrouck (1993), O'hara (1995), Madhavan et al. (1997), Madhavan (2000), Stoll (2000) and Hasbrouck (2007) among other prominent work.

The question we address is: can we trust models such as (1.1) or (1.2)? To investigate it, we introduce the general set-up as

$$\underbrace{Z_{t_i}}_{\text{observed price}} = \underbrace{X_{t_i}}_{\text{efficient price}} + \underbrace{\underbrace{\phi(Q_i, \theta_0)}_{\text{explicative part}} + \underbrace{\epsilon_{t_i}}_{\text{residual noise}}}_{\text{MMN}}, \tag{1.3}$$

where Q_i are observable variables included in the LOB while ϕ is known to the econometrician, and we develop tests for the presence of the residual noise ϵ_{t_i} at any given sampling frequency. The associated null hypothesis is such that $\{\epsilon_{t_i} = 0, \phi = \Phi\}$ and the alternative $\{\text{Var}[\epsilon_{t_i}] > 0, \phi = \Phi \text{ or } \phi = 0\}$, where $\Phi := \Phi(Q_i, \theta_0) \neq 0$ is known up to the parameter θ_0.

Our tests are based on Hausman (1978) tests[1] developed in Aït-Sahalia and Xiu (2016), which is restricted to the case $\Phi = 0$. The authors consider the difference $\widehat{\sigma}_{RV}^2 - \widehat{\sigma}_{QMLE}^2$, where $\widehat{\sigma}_{RV}^2 = T^{-1} \sum_i (Z_{t_i} - Z_{t_{i-1}})^2$ and $\widehat{\sigma}_{QMLE}^2$ corresponds to the QMLE in a model where $\phi = 0$. The Hausman test statistic is of the form $H = n(\widehat{\sigma}_{RV}^2 - \widehat{\sigma}_{QMLE}^2)^2 / \widehat{V}$, where $\widehat{V}$ is an estimator of $AVAR(\widehat{\sigma}_{RV}^2 - \widehat{\sigma}_{QMLE}^2)$ under the null hypothesis. Under the alternative, RV is not consistent whereas the QMLE stays consistent so that the authors show that H explodes in that case.

To test the presence of residual noise in (1.3) in the case where $\phi \neq 0$, we consider the Hausman tests comparing two distinct QMLE related to a model including the explicative part, i.e. $\widehat{\sigma}_{exp}^2$ which is restricted to a null residual noise and $\widehat{\sigma}_{err}^2$ including the residual noise. As far as the authors know, $\widehat{\sigma}_{err}^2$ is novel in the particular context of high frequency data. They respectively play the role of $\widehat{\sigma}_{RV}^2$ and $\widehat{\sigma}_{QMLE}^2$, but they are both distinct from the latter related to a model with $\phi = 0$. Note that Aït-Sahalia and Xiu (2016) consider other candidates for testing, including the PAE, that we set aside in this paper.

The estimator $\widehat{\sigma}_{exp}^2$ corresponds exactly to estimators considered in Li et al. (2016) (the so-called estimated-price RV in the latter paper) and Chaker (2017) (although the latter work is restricted to a linear ϕ). Moreover, the E-QMLE discussed in Li et al. (2016) (Section 2.2.1), i.e. first estimating the price and then applying the usual QMLE to the estimates, is asymptotically equivalent to $\widehat{\sigma}_{err}^2$. In addition, Chaker (2017) actually provides tests from a different nature for the presence of residual noise when ϕ is linear. Finally, some extensions are considered in Section 4.4 of Potiron and Mykland (2016).

Our first main theoretical contribution includes the investigation of the joint limit theory of $(\widehat{\sigma}_{exp}^2, \widehat{\sigma}_{err}^2, \widehat{a}_{err}^2, \widehat{\theta}_{exp}, \widehat{\theta}_{err})$, where $\widehat{a}_{err}^2$ is an estimator of the residual noise and $(\widehat{\theta}_{exp}, \widehat{\theta}_{err})$ are estimates of the parameter both obtained via QMLE under

---

[1] As far as we know, the use of Hausman tests in high-frequency data can be traced back to the TSRV and Huang and Tauchen (2005).

small residual noise, i.e. $\text{Var}[\epsilon_{t_i}] = O(1/n)$. The marginal limit theory for $\widehat{\sigma}^2_{exp}$ boils down to Theorem 2 in Li et al. (2016) as $\widehat{\sigma}^2_{exp}$ is equal to the estimated-price RV estimator, and to Theorem 4 (i) in Chaker (2017) when ϕ is linear. In addition, used in conjunction with the toolkit in Aït-Sahalia and Xiu (2016), one could easily obtain an asymptotically equivalent joint limit theory of $(\widehat{\sigma}^2_{exp}, \widehat{\sigma}^2_{err}, \widehat{a}^2_{err})$ as the E-QMLE is asymptotically equivalent to $\widehat{\sigma}^2_{err}$. However the main difference between our setting and that of Aït-Sahalia and Xiu (2016) is that we have stochastic observation times whereas the cited authors only consider regular sampling times. In particular, in case of regular observation times, our contribution boils down to the marginal and joint limits for $(\widehat{\theta}_{exp}, \widehat{\theta}_{err})$. We further demonstrate that only $\widehat{\sigma}^2_{err}$ is residual noise robust so that we can consider the corresponding Hausman statistic to test the presence of residual noise.

When there is no residual noise in the model, i.e. $\epsilon_{t_i} = 0$, a byproduct of our contribution is that the parameter estimators are asymptotically equivalent. Subsequently, following the procedure considered in Chaker (2017) and Li et al. (2016), we can consistently estimate the efficient price directly from the data as

$$\widehat{X}_{t_i} = Z_{t_i} - \phi(Q_i, \widehat{\theta}_{exp}). \tag{1.4}$$

This procedure seems to be traced back to the model with uncertainty zones, which was introduced in Robert and Rosenbaum (2010, 2012). See also the pioneer work from Hansen and Lunde (2006) and the more recent work from Andersen et al. (2017) for efficient price estimation, although in a slightly different context.

When we assume a residual noise in the model, we examine the measure of goodness of fit introduced in Li et al. (2016), which corresponds to the proportion of MMN variance explained by the explicative part. Such measure can be estimated using the parameter and the residual noise variance estimates obtained with the QMLE related to the model including the residual noise. Our second main contribution establishes the corresponding central limit theorem in case when the variance of the residual noise stays constant. This goes one step further than Theorem 3 from Li et al. (2016) in that the noise variance does not shrink to 0 asymptotically, and that we can actually provide a more reliable residual noise variance estimator, along with the asymptotic theory. Also, the convergence rate is smaller than the pre-estimation (1.4)-TSRV approach considered in Chaker (2017) (see Theorem 4 (ii)). In particular, volatility estimation is naturally not as fast as when we assume small noise in the model.

We implement the tests over a one month period with tick by tick data, and find out that the linear signed spread model (1.2) consistently stands out from many other alternatives including Roll model (1.1). The tests further reveal that the large majority of stocks can be reasonably considered as free from residual noise with such model. Moreover, we implement the tests from Aït-Sahalia and Xiu (2016) regarding the estimated efficient price (1.4) as the given observed price. They largely corroborate the findings.

As far as we know, there are at least another paper in volatility estimation closely related to our work. The impact of ϕ on RV is thoroughly discussed in Diebold and Strasser (2013). In that paper, the authors study several leading models from the market microstructure literature. Unfortunately, their assumption of constant volatility is quite strong.

The remainder of the paper is structured as follows. The model is introduced in Section 2. The limit theory of the QMLE under small noise, the Hausman tests and the efficient price estimator are developed in Section 3. We discuss about measure of goodness of fit estimation, central limit theory under large noise and guidance for implementation of volatility estimation in Section 4. Section 5 performs a Monte Carlo experiment to assess finite sample performance of the tests and validation of the sequence to estimate volatility. Section 6 is devoted to an empirical study. We conclude in Section 7. Theoretical details and proofs can be found in the Appendix.

2. Model

For a given horizon time $T > 0$, we make observations[2] contaminated by the MMN at (possibly random) times $0 = t_0 \leq \cdots \leq t_N \leq T$ of the efficient log-price X_t, and we assume that we have the additive decomposition

$$\underbrace{Z_{t_i}}_{\text{observed price}} = \underbrace{X_{t_i}}_{\text{efficient price}} + \underbrace{\underbrace{\phi(Q_i, \theta_0)}_{\text{explicative part}} + \underbrace{\epsilon_{t_i}}_{\text{residual noise}}}_{\text{MMN}}.$$

Here the parameter $\theta_0 \in \Theta \subset \mathbb{R}^d$, where Θ is a compact set. The impact function ϕ is known, of class C^m in θ with $m > d/2 + 2$, and ϵ_{t_i} corresponds to the remaining noise. Finally, $Q_i \in \mathbb{R}^q$ includes *observable* information from the LOB such as the trade type I_i, the trading volume V_i (Glosten and Harris, 1988), the duration time between two trades D_i (Almgren and Chriss, 2001), the quoted depth[3] QD_i (Kavajecz, 1999), the bid–ask spread S_i, the order flow imbalance[4] OFI_i (Cont et al., 2014). The introduced MMN complies with the empirical evidence about autocorrelated noise[5] (see, e.g., Kalnina and Linton (2008) and Aït-Sahalia et al. (2011)). Some examples of ϕ can be consulted on Table 1.

---

[2] All the considered quantities are implicitly or explicitly indexed by n. Consistency and convergence in law refer to the behavior as $n \to \infty$. A full specification of the model actually 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. We assume that all the processes are $\mathbf{F}$-adapted (either in a continuous or discrete meaning) and that the observation times t_i are $\mathbf{F}$-stopping times. Also, when referring to Itô-semimartingale, we automatically mean that the statement is relative to $\mathbf{F}$.

[3] The ask (bid) depth specifies the volume available at the best ask (bid).

[4] It is defined as the imbalance between supply and demand at the best bid and ask prices (including both quotes and cancellations).

[5] Although not with endogenous and/or heteroskedastic noise.

Table 1
Overview of models and limit order book variables.

Limit order book variables		
Name	Symbol	Definition
trade type	I_i	1 if the trade at time t_i is buyer-initiated and -1 if seller-initiated
trading volume	V_i	number of shares traded at t_i
duration time	D_i	$D_i = t_i - t_{i-1}$
quoted depth	QD_i	The ask/bid depth specifies the volume available at the best ask/bid
bid–ask spread	S_i	$S_i = A_i - B_i$, with A_i best ask price and B_i best bid price
order flow imbalance	OFI_i	Imbalance between supply and demand at the best bid and ask prices (including both quotes and cancellations) between t_{i-1} and t_i

Models		
Name	$\phi(Q_i, \theta_0)$	Related literature
null	0	
Roll	$I_i\theta_0$	(Roll, 1984)
Glosten–Harris	$I_i(\theta_0^{(1)} + V_i\theta_0^{(2)})$	(Glosten and Harris, 1988)
signed timestamp	$I_i D_i^{-1}\theta_0$	(Almgren and Chriss, 2001)
signed spread	$\frac{1}{2}I_i S_i\theta_0$	
signed quoted depth	$I_i QD_i\theta_0$	(Kavajecz, 1999)
order flow imbalance	$OFI_i\theta_0$	(Cont et al., 2014)
NL signed spread	$I_i \frac{S_i\theta_0}{1+S_i\theta_0}$	
general	$I_i(\theta_0^{(1)} + V_i\theta_0^{(2)}$ $+ D_i^{-1}\theta_0^{(3)} + S_i\theta_0^{(4)}$ $+ QD_i\theta_0^{(5)})$ $+ OFI_i\theta_0^{(6)}$	

The efficient price

The latent log-price X_t is an Itô-semimartingale of the form

$$dX_t = b_t dt + \sigma_t dW_t + dJ_t, \tag{2.1}$$

$$d\sigma_t = \widetilde{b}_t dt + \widetilde{\sigma}_t^{(1)} dW_t + \widetilde{\sigma}_t^{(2)} d\widetilde{W}_t + d\widetilde{J}_t, \tag{2.2}$$

with $(W_t, \widetilde{W}_t)$ which is a 2 dimensional standard Brownian motion, the drift $(b_t, \widetilde{b}_t)$ which is componentwise locally bounded, $(\sigma_t, \widetilde{\sigma}_t^{(1)}, \widetilde{\sigma}_t^{(2)})$[6] which is componentwise locally bounded, itself an Itô process and $\inf_t(\min(\sigma_t, \widetilde{\sigma}_t^{(2)})) > 0$ a.s. We also assume that $(J_t, \widetilde{J}_t)$ is a 2 dimensional pure jump process[7] of finite activity.

The observation times

Crucial to the estimation is the robustness of the procedure when considering tick-time volatility instead of calendar-time volatility. For instance, Patton (2011, see, e.g., p. 299) compares empirically the accuracy of estimators and mentions that using tick-time sampling leads to more accurate volatility estimation, although the considered estimators are a priori not robust to such sampling procedure. Also, Xiu (2010) and Aït-Sahalia and Xiu (2016, Section 5, p. 17) compute the likelihood estimators estimating tick-time volatility even though the theory only covers the regular observation times framework.

We introduce the notation $\Delta := T/n$. We consider the random discretization scheme used in Clinet and Potiron (2018a, Section 4) and adapted from Jacod and Protter (2011) (see Section 14.1). We assume that there exists an Itô-semimartingale $\alpha_t > 0$ which satisfies Assumption 4.4.2 p. 115 in Jacod and Protter (2011) and is locally bounded and locally bounded away from 0, and i.i.d $U_i > 0$ that are independent with each other and from other quantities such that

$$t_0 = 0, \tag{2.3}$$

$$t_i = t_{i-1} + \Delta\alpha_{t_{i-1}}U_i. \tag{2.4}$$

We further assume that $\mathbb{E}U_i = 1$, and that for any $q > 0$, $m_q := \mathbb{E}U_i^q < \infty$, is independent of n. If we define $\pi_t := \sup_{i\geq 1} t_i - t_{i-1}$ and the number of observations before t as $N(t) = \sup\{i \in \mathbb{N}|0 < t_i \leq t\}$ we have that $\pi_t \to^{\mathbb{P}} 0$ and that[8]

$$\frac{N(t)}{n} \to^{\mathbb{P}} \frac{1}{T}\int_0^t \frac{1}{\alpha_s}ds. \tag{2.5}$$

When there is no room for confusion, we sometimes drop T in the expression, i.e we use $N := N(T)$.

---

[6] A nice review on the use of stochastic volatility in financial mathematics can be found in Ghysels et al. (1996).

[7] Jumps in volatility are a salient part of the data (see, e.g., Todorov and Tauchen (2011) for empirical evidence.)

[8] Actually the convergence is *u.c.p*, i.e. uniformly in probability on $[0, t]$ for any $t \in [0, T]$. Eq. (2.5) can be shown using Lemma 14.1.5 in Jacod and Protter (2011). The uniformity is obtained as a consequence of the fact that N_n and $\int_0^\cdot \frac{1}{\alpha_s}ds$ are increasing processes and Property (2.2.16) in Jacod and Protter (2011).

The information

Given the process X_t, the observed information Q_i is assumed to be conditionally stationary, i.e. for any $k, j, i_1, \ldots, i_k \in \mathbb{N}$ and for any continuous and bounded function f we have

$$\mathbb{E}\left[f(Q_{i_1+j}, \ldots, Q_{i_k+j}) \middle| X\right] = \mathbb{E}\left[f(Q_{i_1}, \ldots, Q_{i_k}) \middle| X\right] \text{ a.s.} \tag{2.6}$$

We introduce the difference between the explicative part taken in θ and in θ_0 as

$$W_i(\theta) := \phi(Q_i, \theta) - \phi(Q_i, \theta_0), \tag{2.7}$$

and for any $i, j, k, l \in \mathbb{N}$, and for any multi-indices $\boldsymbol{q} = (q_1, q_2)$, $\boldsymbol{r} = (r_1, r_2, r_3, r_4)$, where the subcomponents of $\boldsymbol{q}$ and $\boldsymbol{r}$ are themselves d dimensional multi-indices, the following quantities conditioned on the price process

$$\mathbb{E}\left[W_i(\theta) | X\right] = 0 \text{ a.s}, \tag{2.8}$$

$$\rho_j^{\boldsymbol{q}}(\theta) := \mathbb{E}\left[\frac{\partial^{q_1} W_i(\theta)}{\partial \theta^{q_1}} \frac{\partial^{q_2} W_{i+j}(\theta)}{\partial \theta^{q_2}} \middle| X\right] = \mathbb{E}\left[\frac{\partial^{q_1} W_i(\theta)}{\partial \theta^{q_1}} \frac{\partial^{q_2} W_{i+j}(\theta)}{\partial \theta^{q_2}}\right] \text{ a.s}, \tag{2.9}$$

$$\kappa_{j,k,l}^{\boldsymbol{r}}(\theta) := \operatorname{cum}\left[\frac{\partial^{r_1} W_i(\theta)}{\partial \theta^{r_1}}, \frac{\partial^{r_2} W_{i+j}(\theta)}{\partial \theta^{r_2}}, \frac{\partial^{r_3} W_{i+k}(\theta)}{\partial \theta^{r_3}}, \frac{\partial^{r_4} W_{i+l}(\theta)}{\partial \theta^{r_4}} \middle| X\right]$$

$$= \operatorname{cum}\left[\frac{\partial^{r_1} W_i(\theta)}{\partial \theta^{r_1}}, \frac{\partial^{r_2} W_{i+j}(\theta)}{\partial \theta^{r_2}}, \frac{\partial^{r_3} W_{i+k}(\theta)}{\partial \theta^{r_3}}, \frac{\partial^{r_4} W_{i+l}(\theta)}{\partial \theta^{r_4}}\right] \text{ a.s}, \tag{2.10}$$

where $\rho_j^{\boldsymbol{q}}(\theta)$ and $\kappa_{j,k,l}^{\boldsymbol{r}}(\theta)$ are assumed independent of n. Note that conditions (2.8)–(2.10) state that conditional moments of the information process (and its derivatives with respect to θ) up to the fourth order are independent of the efficient price. This is weaker than assuming the independence of Q and X (and thus it is weaker than the classical QMLE framework of Xiu (2010) where the MMN is assumed independent of X). When $\boldsymbol{q} = 0$ (respectively $\boldsymbol{r} = 0$), we refer directly to $\rho_j(\theta)$ (respectively $\kappa_{j,k,l}(\theta)$) in place of $\rho_j^{\boldsymbol{q}}(\theta)$ (respectively $\kappa_{j,k,l}^{\boldsymbol{q}}(\theta)$). To ensure the weak dependence of the information over time and the identifiability of θ_0, we also assume for any $i = 0, \ldots, m$ and $0 \le |\boldsymbol{q}|, |\boldsymbol{r}| \le m$ the following set of conditions:

$$\sup_{\theta \in \Theta} \sum_{j=0}^{+\infty} \left|\rho_j^{\boldsymbol{q}}(\theta)\right| < \infty \text{ a.s}, \tag{2.11}$$

$$\sup_{\theta \in \Theta} \sum_{j,k,l=0}^{+\infty} \left|\kappa_{j,k,l}^{\boldsymbol{r}}(\theta)\right| < \infty \text{ a.s}, \tag{2.12}$$

$$\mathbb{E}\left[\sup_{\theta \in \Theta} \left|\frac{\partial^j \mu_i(\theta)}{\partial \theta^j}\right|^p \middle| X\right] < \infty \text{ a.s, for any } p \ge 1, 0 \le j \le 2, \tag{2.13}$$

$$\frac{\partial \rho_0(\theta)}{\partial \theta} = 0 \Leftrightarrow \theta = \theta_0. \tag{2.14}$$

Remark 1. Conditions (2.11)–(2.12) ensure the weak dependence over time of the information process whereas Condition (2.14) implies the identifiability of θ_0 for the QMLE. They are needed in order to derive the limit theory of the QMLE estimators related to θ_0 that are defined in the next section. Note that we consider a setting where the information process is stationary when conditioned on the efficient price process, which was not assumed in Li et al. (2016). In particular conditions (2.11)–(2.12) are stronger forms for stationary sequences of Condition (A.xi), while (2.14) replaces the identifiability assumption (A.x) in their paper. The need for stronger assumptions is due to the fact that in this work, in addition to the consistency with rate of convergence N, we also prove the central limit theory for the QMLE related to θ_0. On the other hand, the moment condition (2.13) is weaker than the quite strong assumption (A.v) requiring that the information process is uniformly stochastically bounded.

The residual noise

The remaining noise is assumed independent of all the other processes, i.i.d with $\mathbb{E}[\epsilon_t] = 0$ and $\mathbb{E}[\epsilon_t^2] = a_0^2 > 0$, and with finite fourth moment.

Remark 2. Given the assumptions on the information process Q_i and on the residual noise ϵ_{t_i}, we have ruled out the case of an heteroskedastic MMN. Although empirical evidence indicates time dependence of the MMN (as pointed out in, e.g., Hansen and Lunde (2006)), incorporating heteroskedasticity in our model is beyond the scope of this paper. Note also that we only allow for a weak form of endogeneity for the explicative part $\phi(Q_i, \theta)$ (its conditional moments of order 4 or less should not depend on X_t). Again, we set aside stronger forms of endogeneity in this paper. Nevertheless, we have considered an endogenous and heteroskedastic residual noise in our simulation study and shown that the tests seem reasonably robust to such misspecification.

3. Tests for the presence of residual noise

3.1. Small noise alternative case

We first consider the simple semiparametric model where $X_t = \sigma_0 W_t$, the observations are regular $t_{i+1} - t_i = \Delta$ which implies that $N = n$, the residual noise ϵ_{t_i} is normally distributed with zero-mean and variance a_0^2. We further define $\Delta_N = T/N$ which in this simple model satisfies $\Delta_N = \Delta$. The null hypothesis is defined as $\mathcal{H}_0 : \{a_0^2 = 0, \phi = \Phi\}$ whereas the alternative is defined as $\mathcal{H}_1 : \{a_0^2 := \eta_0/n > 0, \phi = \Phi\}$, where $\Phi := \Phi(Q_i, \theta_0) \neq 0$ and η_0 is a constant which does not depend on n. The cases of large noise alternative and $\phi = 0$ alternative are respectively delayed to Sections 3.2 and 3.3. To ensure that our method is robust to general information, our strategy consists in considering two distinct likelihood functions conditioned on the information. We define the observed log returns $Y_i = Z_{t_i} - Z_{t_{i-1}}$, $Y = (Y_1, \ldots, Y_N)^T$. Moreover, the returns of information are denoted by $\mu_i(\theta) = \phi(Q_i, \theta) - \phi(Q_{i-1}, \theta)$, $\mu(\theta) = (\mu_1(\theta), \ldots, \mu_N(\theta))^T$ and we further define $\widetilde{Y}(\theta) = Y - \mu(\theta)$. Key to our analysis is that $\widetilde{Y}(\theta)$ is known to the econometrician.

In the absence of residual noise, the observed returns can be expressed as

$$Y_i = \sigma_0(W_{t_i} - W_{t_{i-1}}) + \mu_i(\theta_0). \tag{3.1}$$

It is then clear that $\widetilde{Y}(\theta_0)$ is i.i.d normally distributed centered with variance $\sigma_0^2 \Delta_N$ and the log-likelihood can be expressed as

$$l_{exp}(\sigma^2, \theta) = -\frac{N}{2}\log(\sigma^2 \Delta_N) - \frac{N}{2}\log(2\pi) - \frac{1}{2\sigma^2 \Delta_N}\widetilde{Y}(\theta)^T\widetilde{Y}(\theta). \tag{3.2}$$

When the residual noise is present, Aït-Sahalia et al. (2005) show that in the case where there is no information, i.e.

$$Y_i = \sigma_0(W_{t_i} - W_{t_{i-1}}) + (\epsilon_{t_i} - \epsilon_{t_{i-1}}), \tag{3.3}$$

Y features a MA(1) process so that the log-likelihood process of the model is

$$l(\sigma^2, a^2) = -\frac{1}{2}\log \det(\Omega) - \frac{N}{2}\log(2\pi) - \frac{1}{2}Y^T \Omega^{-1} Y, \tag{3.4}$$

where Ω is the matrix

$$\Omega = \begin{pmatrix} \sigma^2 \Delta_N + 2a^2 & -a^2 & 0 & \cdots & 0 \\ -a^2 & \sigma^2 \Delta_N + 2a^2 & -a^2 & \ddots & \vdots \\ 0 & -a^2 & \sigma^2 \Delta_N + 2a^2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -a^2 \\ 0 & \cdots & 0 & -a^2 & \sigma^2 \Delta_N + 2a^2 \end{pmatrix}. \tag{3.5}$$

When incorporating non-null information, the model for the returns can be written as

$$Y_i = \sigma_0(W_{t_i} - W_{t_{i-1}}) + \mu_i(\theta_0) + (\epsilon_{t_i} - \epsilon_{t_{i-1}}). \tag{3.6}$$

It is then immediate to see that $\widetilde{Y}(\theta_0)$ follows a MA(1) dynamic so that we can substitute the log-likelihood function by

$$l_{err}(\sigma^2, \theta, a^2) = -\frac{1}{2}\log \det(\Omega) - \frac{N}{2}\log(2\pi) - \frac{1}{2}\widetilde{Y}(\theta)^T \Omega^{-1}\widetilde{Y}(\theta). \tag{3.7}$$

To assess the central limit theory, we consider the general framework specified in Section 2 and define the quadratic variation as

$$T\overline{\sigma}_0^2 := \int_0^T \sigma_s^2 ds + \sum_{0 < s \leq T} \Delta J_s^2,$$

where $\Delta J_s = J_s - J_{s-}$, and we assume that $\overline{\sigma}_0^2 \in \left[\underline{\sigma}^2, \overline{\sigma}^2\right]$ almost surely, where $\underline{\sigma}^2 > 0$. This assumption is necessary to maximize the quasi likelihood function on a well-defined bounded space. This may seem to be a somewhat restrictive condition on the volatility process, but since $\overline{\sigma}^2$ can be taken arbitrarily large, it does not affect the implementation of the estimation procedure in practice. Under $\mathcal{H}_0$ and assuming null information, Aït-Sahalia and Xiu (2016) show that the QMLE associated to (3.4) is optimal with rate of convergence $n^{1/2}$. When incorporating information into the model, both QMLE related to (3.2) and (3.7) also turn out to converge with rate $n^{1/2}$. Formally, we assume that $\upsilon_0 := (\overline{\sigma}_0^2, \theta_0) \in \Upsilon$, where $\Upsilon = \left[\underline{\sigma}^2, \overline{\sigma}^2\right] \times \Theta$. We define $\widehat{\upsilon}_{exp} := (\widehat{\sigma}_{exp}^2, \widehat{\theta}_{exp})$ and $\widehat{\xi}_{err} := (\widehat{\sigma}_{err}^2, \widehat{\theta}_{err}, \widehat{a}_{err}^2)$ as respectively one solution to the equation $\partial_\upsilon l_{exp}(\upsilon) = 0$ on the interior of Υ and one solution to the equation $\partial_\xi l_{err}(\xi) = 0$ on $\Upsilon \times \left[-\underline{\eta}/n, \overline{\eta}/n\right]$, where $\overline{\eta} > 0$ and $0 < \underline{\eta} < \underline{\sigma}^4/4$. This corresponds to an extension of parameter space as a^2 can take negative values, as in Aït-Sahalia and Xiu (2016, see the discussion at the bottom of p. 8). Such extension is needed because under $\mathcal{H}_0$ and with the non-extended space $[0, \overline{\eta}/n]$, the parameter $a_0^2 = 0$ would lie on the boundary of the parameter space, making the above procedure inconsistent. In the following theorem, we give the joint limit distribution of $(\widehat{\upsilon}_{exp}, \widehat{\xi}_{err})$ assuming that the noise process is of order $1/\sqrt{n}$. We also specify the limit under $\mathcal{H}_0$, i.e when there is no residual noise.

Theorem 3.1 (*Joint Central Limit Theorem for* $(\widehat{\upsilon}_{exp}, \widehat{\xi}_{err})$ *Under the Small Residual Noise Framework*)**.** *Assume that* $a_0^2 = \eta_0 T/n$ *and that* $\mathrm{cum}_4[\epsilon] = \mathcal{K}T^2/n^2$ *for some fixed* $\eta_0 \geq 0$, $\mathcal{K} \geq 0$, *where* $\mathrm{cum}_4[\epsilon]$ *is the fourth order cumulant of* ϵ_t. *Then, we have* $\mathcal{G}_T$-*stably*[9] *in law that*

$$
\begin{pmatrix}
N^{1/2}\big(\widehat{\sigma}_{exp}^2 - \overline{\sigma}_0^2 - 2\widetilde{\eta}_0\big) \\
N\big(\widehat{\theta}_{exp} - \theta_0 - N^{-1}B_{\theta_0,exp}\big) \\
N^{1/2}\big(\widehat{\sigma}_{err}^2 - \overline{\sigma}_0^2\big) \\
N\big(\widehat{\theta}_{err} - \theta_0 - N^{-1}B_{\theta_0,err}\big) \\
N^{3/2}\big(\widehat{a}_{err}^2 - a_0^2\big)
\end{pmatrix}
\to \mathcal{MN}\left(0, \frac{\mathcal{Q}}{T} \times
\begin{pmatrix}
2 & 0 & 2 & 0 & 0 \\
0 & \dfrac{T\int_0^T \sigma_s^2 ds}{\mathcal{Q}} U_{\theta_0}^{-1} & 0 & \dfrac{T\int_0^T \sigma_s^2 ds}{\mathcal{Q}} U_{\theta_0}^{-1} & 0 \\
2 & 0 & 6 & 0 & -2T \\
0 & \dfrac{T\int_0^T \sigma_s^2 ds}{\mathcal{Q}} U_{\theta_0}^{-1} & 0 & \dfrac{T\int_0^T \sigma_s^2 ds}{\mathcal{Q}} U_{\theta_0}^{-1} & 0 \\
0 & 0 & -2T & 0 & T^2
\end{pmatrix} + \mathbf{V}\right),
$$

where $\mathcal{Q} = T^{-1}\int_0^T \alpha_s^{-1} ds \left\{ \int_0^T \sigma_s^4 \alpha_s ds + \sum_{0<s\leq T} \Delta J_s^2(\sigma_s^2 \alpha_s + \sigma_{s-}^2 \alpha_{s-}) \right\}$, $\widetilde{\eta}_0 = (T^{-1}\int_0^T \alpha_s^{-1} ds)\eta_0$, $\widetilde{\mathcal{K}} = (T^{-1}\int_0^T \alpha_s^{-1} ds)^2 \mathcal{K}$,

$$
U_{\theta_0} = \mathbb{E}\left[\frac{\partial \mu_1(\theta_0)}{\partial \theta} \cdot \frac{\partial \mu_1(\theta_0)}{\partial \theta}^T \right],
$$

$\mathbf{V} := \mathbf{V}(\mathcal{Q}, \overline{\sigma}_0^2, \widetilde{\eta}_0, \widetilde{\mathcal{K}}, \theta_0)$ *is an additional matrix due to the presence of residual noise of the form*

$$
\mathbf{V} = \begin{pmatrix}
\mathbf{V}_{11} & 0 & \mathbf{V}_{13} & 0 & \mathbf{V}_{15} \\
0 & \mathbf{V}_{22} & 0 & \mathbf{V}_{24} & 0 \\
\mathbf{V}_{13} & 0 & \mathbf{V}_{33} & 0 & \mathbf{V}_{35} \\
0 & \mathbf{V}_{24} & 0 & \mathbf{V}_{44} & 0 \\
\mathbf{V}_{15} & 0 & \mathbf{V}_{35} & 0 & \mathbf{V}_{55}
\end{pmatrix},
$$

and $B_{\theta_0,exp}, B_{\theta_0,err}$ *are two bias terms due to the presence of jumps in the price process. The exact expression of* $\mathbf{V}$, $B_{\theta_0,exp}$, *and* $B_{\theta_0,err}$ *can be found in* Appendix B.

In particular, $\mathbf{V}(\mathcal{Q}, \overline{\sigma}_0^2, 0, 0, \theta_0) = 0$, *and thus, under* $\mathcal{H}_0$, *we have* $\mathcal{G}_T$-*stably in law that*

$$
\begin{pmatrix}
N^{1/2}\big(\widehat{\sigma}_{exp}^2 - \overline{\sigma}_0^2\big) \\
N\big(\widehat{\theta}_{exp} - \theta_0 - N^{-1}B_{\theta_0,exp}\big) \\
N^{1/2}\big(\widehat{\sigma}_{err}^2 - \overline{\sigma}_0^2\big) \\
N\big(\widehat{\theta}_{err} - \theta_0 - N^{-1}B_{\theta_0,err}\big) \\
N^{3/2}\big(\widehat{a}_{err}^2 - 0\big)
\end{pmatrix}
\to \mathcal{MN}\left(0, \frac{\mathcal{Q}}{T} \times
\begin{pmatrix}
2 & 0 & 2 & 0 & 0 \\
0 & \dfrac{T\int_0^T \sigma_s^2 ds}{\mathcal{Q}} U_{\theta_0}^{-1} & 0 & \dfrac{T\int_0^T \sigma_s^2 ds}{\mathcal{Q}} U_{\theta_0}^{-1} & 0 \\
2 & 0 & 6 & 0 & -2T \\
0 & \dfrac{T\int_0^T \sigma_s^2 ds}{\mathcal{Q}} U_{\theta_0}^{-1} & 0 & \dfrac{T\int_0^T \sigma_s^2 ds}{\mathcal{Q}} U_{\theta_0}^{-1} & 0 \\
0 & 0 & -2T & 0 & T^2
\end{pmatrix}\right).
$$

Remark 3.2 (*Regular Sampling*)**.** If observations are regular, $\mathcal{Q}$, $\widetilde{\eta}_0$ and $\widetilde{\mathcal{K}}$ can be specified as

$$
\mathcal{Q} = \int_0^T \sigma_s^4 ds + \sum_{0<s\leq T} \Delta J_s^2(\sigma_s^2 + \sigma_{s-}^2), \quad \widetilde{\eta}_0 = \eta_0, \quad \widetilde{\mathcal{K}} = \mathcal{K}. \tag{3.8}
$$

We consider now the problem of testing $\mathcal{H}_0$ against $\mathcal{H}_1$. To do that we consider the Hausman statistics of the form

$$
S = N(\widehat{\sigma}_{exp}^2 - \widehat{\sigma}_{err}^2)^2 / \widehat{V}, \tag{3.9}
$$

where $\widehat{V}$ is a consistent estimator of $AVAR(\widehat{\sigma}_{exp}^2 - \widehat{\sigma}_{err}^2)$ that will be defined in what follows. We aim to show that S satisfies the key asymptotic properties

$$
S \to \chi^2 \quad \text{under } \mathcal{H}_0, \tag{3.10}
$$
$$
S \to \infty \quad \text{under } \mathcal{H}_1, \tag{3.11}
$$

where χ^2 is a standard chi-squared distribution. Actually, we can deduce (3.10) from Theorem 3.1 along with the consistency of $\widehat{V}$ and (3.11) is relatively easy to obtain. As in Aït-Sahalia and Xiu (2016), we consider three distinct scenarios, i.e.

  (i)   constant volatility
 (ii)   time-varying volatility and no price jump
(iii)   time-varying volatility and price jump

---

[9] The filtration $\mathbf{G} = (\mathcal{G}_t)_{0\leq t\leq T}$ is defined as $\mathcal{G}_t := \sigma\left\{ U_i^n, \alpha_s, X_s | (i,n) \in \mathbb{N}^2, 0 \leq s \leq t \right\}$.

This leads us to define two (one estimator is robust to two scenarios) distinct variance estimators $\widehat{V}_i$ and their affiliated statistics $S_i = N(\widehat{\sigma}^2_{exp} - \widehat{\sigma}^2_{err})^2/\widehat{V}_i$ in what follows. Due to the non regularity of arrival times, fourth power returns based estimators such as $\widehat{V}_5$ (defined in Appendix A) are inconsistent in general. We therefore consider bipower statistics, inspired by Barndorff-Nielsen and Shephard (2004b,a). If we assume (i) we have that $AVAR(\widehat{\sigma}^2_{exp} - \widehat{\sigma}^2_{err}) = 4T^{-2}\sigma_0^4 \int_0^T \alpha_s^{-1} ds \int_0^T \alpha_s ds$, which can be estimated by

$$\widehat{V}_1 = \frac{4N}{T^2} \sum_{i=2}^{N} \Delta\widehat{X}_i^2 \Delta\widehat{X}_{i-1}^2. \tag{3.12}$$

The estimator $\widehat{V}_1$ is also robust to (ii), where $AVAR(\widehat{\sigma}^2_{exp} - \widehat{\sigma}^2_{err}) = 4T^{-2} \int_0^T \alpha_s^{-1} ds \int_0^T \sigma_s^4 \alpha_s ds$. Under (iii), $AVAR(\widehat{\sigma}^2_{exp} - \widehat{\sigma}^2_{err}) = 4T^{-2} \int_0^T \alpha_s^{-1} ds \{\int_0^T \sigma_s^4 \alpha_s ds + \sum_{0<s\leq T} \Delta J_s^2(\sigma_s^2\alpha_s + \sigma_{s-}^2\alpha_{s-})\}$. If we introduce $\widetilde{k}$ which is random and satisfies $\widetilde{k}\Delta_N \to^{\mathbb{P}} 0$ and $\widetilde{u}_i = \widetilde{\alpha}(t_i - t_{i-1})^\omega$, we can estimate $AVAR(\widehat{\sigma}^2_{exp} - \widehat{\sigma}^2_{err})$ with

$$\widehat{V}_2 = \frac{4}{T} \left\{ \frac{1}{\Delta_N} \sum_{i=2}^{N} \Delta\widehat{X}_i^2 \Delta\widehat{X}_{i-1}^2 \mathbf{1}_{\{|\Delta\widehat{x}_i|\leq\widetilde{u}_i\}} \mathbf{1}_{\{|\Delta\widehat{x}_{i-1}|\leq\widetilde{u}_{i-1}\}} \right. \tag{3.13}$$
$$\left. + \sum_{i=\widetilde{k}+1}^{N-\widetilde{k}} \Delta\widehat{X}_i^2 \mathbf{1}_{\{|\Delta\widehat{x}_i|>\widetilde{u}_i\}} \left(\widehat{\sigma_{t_i}^2\alpha_{t_i}} + \widehat{\sigma_{t_{i-}}^2\alpha_{t_{i-}}} \right) \right\}, \text{ where}$$

$$\widehat{\sigma_{t_i}^2\alpha_{t_i}} = \frac{1}{\widetilde{k}\Delta_N} \sum_{j=i+1}^{i+\widetilde{k}} \Delta\widehat{X}_j^2 \mathbf{1}_{\{|\Delta\widehat{x}_j|\leq\widetilde{u}_j\}} , \widehat{\sigma_{t_{i-}}^2\alpha_{t_{i-}}} = \widehat{\sigma_{t_{i-\widetilde{k}-1}}^2\alpha_{t_{i-\widetilde{k}-1}}}.$$

We first show the consistency of the proposed estimators.

Proposition 3.3. *For any $i = 1, 2$ we have, as $n \to +\infty$,*

under $\mathcal{H}_0$: $\widehat{V}_i \to^{\mathbb{P}} AVAR(\widehat{\sigma}^2_{exp} - \widehat{\sigma}^2_{err})$,

under $\mathcal{H}_1$: $\widehat{V}_i = O_{\mathbb{P}}(1)$,

if we assume the related framework.

We then deduce asymptotic properties of the statistics.

Corollary 3.4. *Let $0 < \beta < 1$ and c_β the associated β-quantile of the standard chi-squared distribution. Under the related framework, the test statistics S_i satisfy*

$$\mathbb{P}(S_i > c_{1-\beta} \mid \mathcal{H}_0) \to \beta \text{ and } \mathbb{P}(S_i > c_{1-\beta} \mid \mathcal{H}_1) \to 1. \tag{3.14}$$

When there is no residual noise in the model, following the procedure considered in Li et al. (2016) and Chaker (2017), we can estimate the efficient price as

$$\widehat{X}_t = Z_{t_i} - \phi(Q_i, \widehat{\theta}_{exp}), \text{ for } t \in (t_{i-1}, t_i]. \tag{3.15}$$

By virtue of Theorem 3.1, we have that $\widehat{\theta}_{exp}$ is consistent and thus we can show the consistency of $\widehat{X}_t$. It is also immediate to see that

$$\widehat{\sigma}^2_{exp} = T^{-1} \sum_{i=1}^{N} (\widehat{X}_{t_i} - \widehat{X}_{t_{i-1}})^2. \tag{3.16}$$

Formally, the volatility estimator (3.16) expressed as a function of the estimated parameter is equal to the volatility estimator (also viewed as a function of the estimated parameter) considered in Li et al. (2016). Moreover, given the shape of l_{exp} in (3.2), the QMLE $\widehat{\theta}_{exp}$ and the least square estimator (9) in the cited paper (p. 35) coincide, implying that both volatility estimators are equal. The need to correct the price prior to using RV in (3.16) can be understood looking at Table 3 (p. 1324) from Diebold and Strasser (2013). In that table, the first column reports the limit of the naive RV. Accordingly, one can see that depending on the serial autocorrelation of ϕ, there will be one or more extra autocorrelation terms in the limit. Subsequently, the use of the price estimation in (3.16) permits to get rid of those additive terms.

The following corollary formally states the consistency of $\widehat{X}_t$ and the efficiency of RV when used on $\widehat{X}_t$. This corresponds exactly to Theorem 2 in Li et al. (2016). This also corresponds to Theorem 4 (i) in Chaker (2017) when ϕ is linear.

Corollary 3.5. *Under $\mathcal{H}_0$, the estimator $\widehat{X}_t$ is consistent, i.e. for any $t \in [0, T]$,*

$$\widehat{X}_t \to^{\mathbb{P}} X_t. \tag{3.17}$$

Furthermore, we have $\mathcal{G}_T$-stably in law that

$$N^{1/2}\left(\sum_{i=1}^{N}(\widehat{X}_{t_i} - \widehat{X}_{t_{i-1}})^2 - \int_0^T \sigma_s^2 ds - \sum_{0 < s \le T} \Delta J_s^2\right) \to \mathcal{MN}(0, 2T\mathcal{Q}). \tag{3.18}$$

In particular, when observations are regular and the efficient price is continuous, this can be written as

$$N^{1/2}\left(\sum_{i=1}^{N}(\widehat{X}_{t_i} - \widehat{X}_{t_{i-1}})^2 - \int_0^T \sigma_s^2 ds\right) \to \mathcal{MN}\left(0, 2T\int_0^T \sigma_s^4 ds\right). \tag{3.19}$$

It is interesting to remark that when $J = 0$, convergence (3.18) shows that RV on the estimated price is efficient in the sense that its AVAR attains the nonparametric efficiency bound derived in Renault et al. (2017). Indeed, taking $T = 1$, note that our model of observation times falls under the setting of Renault et al. (2017) (see Assumption 2 and the short discussion below), where, in view of (2.10) on p. 447 in the aforementioned paper, we easily derive that $\alpha_s = T_s'^{-1}$. Thus, (3.18) can be rewritten as

$$n^{1/2}\left(\sum_{i=1}^{N}(\widehat{X}_{t_i} - \widehat{X}_{t_{i-1}})^2 - \int_0^1 \sigma_s^2 ds\right) \to \mathcal{MN}\left(0, 2\int_0^1 \sigma_s^4 T_s'^{-1} ds\right), \tag{3.20}$$

which corresponds precisely to the efficiency bound (3.18) on p. 454 in Renault et al. (2017) in the case $g(u, \sigma^2) = \sigma^2$.

3.2. Large noise alternative case

If one can detect small noise, one can a-priori detect large noise. In this section, we consider the large noise alternative $\widetilde{\mathcal{H}}_1 : \{a_0^2 := \eta_0, \phi = \Phi\}$, where we recall that $\eta_0 > 0$ and $\Phi := \Phi(Q_i, \theta_0) \ne 0$. We show that Proposition 3.3 and Corollary 3.4 remain valid in what follows. We have removed the statements related to $\mathcal{H}_0$ which obviously stay true.

Proposition 3.6. *Under the related framework, for any $i = 1, 2$ we have, as $n \to +\infty$,*

under $\widetilde{\mathcal{H}}_1$: $\widehat{V}_i = O_{\mathbb{P}}(N^2)$.

Corollary 3.7. *Under the related framework, the test statistics S_i satisfy*

$$\mathbb{P}(S_i > c_{1-\beta} \mid \widetilde{\mathcal{H}}_1) \to 1. \tag{3.21}$$

3.3. The $\phi = 0$ alternative case

So far we have assumed that the tests were conditional on a specific parametric model where ϕ is non-null, so that the null hypothesis and the alternative were considered under the constraint $\phi \ne 0$. In this part, we consider the pure i.i.d MMN alternative $\overline{\mathcal{H}}_1 : \{\phi = 0, a_0^2 > 0\}$, where the noise may be small ($a_0^2 = \eta_0 T/n$) or large ($a_0^2 = \eta_0 T$) with $\eta_0 > 0$. Note that this is precisely the same alternative as that of Aït-Sahalia and Xiu (2016). We prove in what follows that Proposition 3.3 and Corollary 3.4 remain true up to an innocuous assumption on the fitted model $\{\phi(., \theta), \theta \in \Theta\}$, which is satisfied on all the models considered in this paper. Here again we have removed the statements related to $\mathcal{H}_0$.

Proposition 3.8. *Assume that there exists $\widetilde{\theta}$ in the interior of Θ such that $\phi(., \widetilde{\theta}) = 0$. Then, under the related framework, for any $i = 1, 2$ we have, as $n \to +\infty$,*

under $\overline{\mathcal{H}}_1$ with $a_0^2 = \eta_0 T/n$: $\widehat{V}_i = O_{\mathbb{P}}(1)$,
under $\overline{\mathcal{H}}_1$ with $a_0^2 = \eta_0 T$: $\widehat{V}_i = O_{\mathbb{P}}(N_n^2)$.

Corollary 3.9. *Assume that there exists $\widetilde{\theta}$ in the interior of Θ such that $\phi(., \widetilde{\theta}) = 0$. Then, under the related framework, the test statistics S_i satisfy*

$$\mathbb{P}(S_i > c_{1-\beta} \mid \overline{\mathcal{H}}_1) \to 1. \tag{3.22}$$

4. Goodness of fit

The goal of this section is threefold. First, we introduce a measure of goodness of fit which can be used by the high frequency data user prior to testing to compare several models and assess if one or several candidates are worth testing. Second, we provide the central limit theory of the QMLE related to the model including the residual noise when assuming

that it is present, and we deduce an estimator of the measure. Finally, we give a practical guidance to estimate volatility — this sequence is illustrated in the finite sample analysis that follows.

4.1. Definition

Prior to looking at the Hausman tests, it is safer to assume a model where the variance of the residual noise $a_0^2 > 0$ is non negligible. We introduce the proportion of variance explained as

$$\pi_V := \frac{\mathbb{E}\big[\phi(Q_0, \theta_0)^2\big]}{\mathbb{E}\big[\phi(Q_0, \theta_0)^2\big] + a_0^2}, \tag{4.1}$$

which is a measure of goodness of fit of the model. This measure is almost identical to π_{exp} from Remark 8 (p. 37) in Li et al. (2016). The estimation of (4.1) is based on the QMLE related to the model including the residual noise and given in (4.5).

4.2. Central limit theory

Throughout the rest of this section we assume that the residual noise variance $a_0^2 > 0$ does not depend on n. When $\phi = 0$, this corresponds to a widespread assumption on the residual noise (which in this case corresponds exactly to the MMN). In this setting and further assuming that the volatility is constant, Aït-Sahalia et al. (2005) show that the MLE related to (3.4) is efficient with convergence rate $n^{1/4}$ and obtain the robustness of the MLE in case of departure from the normality of the noise. Xiu (2010) shows that the procedure is also robust to time-varying volatility. We further investigate in Clinet and Potiron (2018a) the behavior of the estimator when adding jumps to the price process and considering non regular stochastic arrival times. In what follows we show in particular that $\widehat{\sigma}_{err}^2$ converges at the same rate $n^{1/4}$. We assume that $\xi_0 := (\overline{\sigma}_0^2, a_0^2, \theta_0) \in \Xi$, where $\Xi = \big[\underline{\sigma}^2, \overline{\sigma}^2\big] \times \big[\underline{a}^2, \overline{a}^2\big] \times \Theta$ with $\underline{a}^2 > 0$. Finally, $\widehat{\xi}_{err}$ is defined as one solution to the equation $\partial_\xi l_{err}(\xi) = 0$ on the interior of Ξ.

Theorem 4.1. *We have $\mathcal{G}_T$-stably in law that*

$$\begin{pmatrix} N^{1/4}\big(\widehat{\sigma}_{err}^2 - \overline{\sigma}_0^2\big) \\ N^{1/2}\big(\widehat{a}_{err}^2 - a_0^2\big) \\ N^{1/2}\big(\widehat{\theta}_{err} - \theta_0\big) \end{pmatrix} \to \mathcal{MN}\left(0, \begin{pmatrix} \dfrac{5a_0 \mathcal{Q}}{T^{3/2}\overline{\sigma}_0} + \dfrac{3a_0\overline{\sigma}_0^3}{T^{1/2}} & 0 & 0 \\ 0 & 2a_0^4 + \mathrm{cum}_4[\epsilon] & 0 \\ 0 & 0 & a_0^2 V_{\theta_0}^{-1} \end{pmatrix}\right), \tag{4.2}$$

where the term $\mathrm{cum}_4[\epsilon]$ stands for the fourth order cumulant of ϵ, and V_{θ_0} is the Fisher information matrix related to θ_0 defined as

$$V_{\theta_0} = \mathbb{E}\left[\frac{\partial \phi\,(Q_0, \theta_0)}{\partial \theta} \cdot \frac{\partial \phi\,(Q_0, \theta_0)}{\partial \theta}^T\right].$$

Remark 4.2 (*Variance Gain when Estimating the Quadratic Variation*). If we assume that the information process Q_i is i.i.d, it is also possible to directly estimate the quadratic variation and the global noise variance using the original QMLE of Xiu (2010) and generalized to our setting with jumps and stochastic observation times in Clinet and Potiron (2018a). Denoting such estimator by $(\widehat{\sigma}_{ori}^2, \widehat{a}_{ori}^2)$, we have:

$$\begin{pmatrix} N^{1/4}\big(\widehat{\sigma}_{ori}^2 - \overline{\sigma}_0^2\big) \\ N^{1/2}\big(\widehat{a}_{ori}^2 - \widetilde{a}_0^2\big) \end{pmatrix} \to \mathcal{MN}\left(0, \begin{pmatrix} \dfrac{5\widetilde{a}_0 \mathcal{Q}}{T^{3/2}\overline{\sigma}_0} + \dfrac{3\widetilde{a}_0\overline{\sigma}_0^3}{T^{1/2}} & 0 \\ 0 & 2\widetilde{a}_0^4 + \mathrm{cum}_4[\epsilon + \phi(Q_0, \theta_0)] \end{pmatrix}\right), \tag{4.3}$$

where $\widetilde{a}_0^2 = a_0^2 + \mathbb{E}[\phi(Q_0, \theta_0)^2]$. Therefore, in view of (4.2) and (4.3), accounting for the explicative part of the noise in the estimation process results in an asymptotic variance reduction for the volatility estimation of a factor

$$\frac{\widetilde{a}_0}{a_0} = \sqrt{1 + \frac{\mathbb{E}[\phi(Q_0, \theta_0)^2]}{a_0^2}} = \frac{1}{\sqrt{1 - \pi_V}}.$$

In particular, when the residual noise is negligible, i.e. in the limit $\pi_V \to 1$, we see that the asymptotic gain is infinite which is coherent with the fact that in such case the rate of convergence of $\widehat{\sigma}_{err}^2$ switches from $N^{1/4}$ to $N^{1/2}$ as in Theorem 3.1.

Remark 4.3 (*Connection to the Literature*). Li et al. (2016) consider a shrinking noise in their Theorem 3, and thus they obtain a faster rate of convergence $n^{1/2}$ for the volatility estimator. On the other hand, Chaker (2017) considers the same setting as ours in their Theorem 4, but as they are using a modification of the TSRV, they obtain the slower rate of convergence $n^{1/6}$.

Remark 4.4 (*Regular Sampling and Continuous Price Case*)**.** When the observation times are regularly spaced and the efficient price is continuous, (4.2) can be specified as

$$
\begin{pmatrix} n^{1/4}\big(\widehat{\sigma}_{err}^2 - \overline{\sigma}_0^2\big) \\ n^{1/2}\big(\widehat{a}_{err}^2 - a_0^2\big) \\ n^{1/2}\big(\widehat{\theta}_{err} - \theta_0\big) \end{pmatrix} \to \mathcal{MN}\left(0, \left(\begin{pmatrix} \dfrac{5a_0 \int_0^T \sigma_s^4 ds}{T\left(\int_0^T \sigma_s^2 ds\right)^{1/2}} + \dfrac{3a_0 \left(\int_0^T \sigma_s^2 ds\right)^{3/2}}{T^2} & 0 & 0 \\ 0 & 2a_0^4 + \mathrm{cum}_4[\epsilon] & 0 \\ 0 & 0 & a_0^2 V_{\theta_0}^{-1} \end{pmatrix}\right)\right). \tag{4.4}
$$

Remark 4.5 (*Local QMLE*)**.** Using the local QMLE, we could further reduce the AVAR of the volatility obtained in (4.2). In the case of regular sampling and continuous price process, we could be as close as possible from the lower efficiency bound defined in Reiss (2011). The proofs of this paper would straightforwardly adapt. The case $\phi = 0$ is treated in Clinet and Potiron (2018a).

Based on Theorem 4.1, we can consistently estimate π_V as

$$
\widehat{\pi}_V := \frac{(N+1)^{-1} \sum_{i=0}^{N} \phi(Q_i, \widehat{\theta}_{err})^2}{(N+1)^{-1} \sum_{i=0}^{N} \phi(Q_i, \widehat{\theta}_{err})^2 + \widehat{a}_{err}^2}. \tag{4.5}
$$

In accordance with our empirical findings (see Section 6), we also investigate what happens in the case where π_V approaches 1, which corresponds to the small noise framework of Section 3, where $a_0^2 = \eta_0 T/n$ for some fixed $\eta_0 \geq 0$, and where $\mathbb{E}[\phi(Q_0, \theta_0)^2] > 0$. It turns out that under this framework too $\widehat{\pi}_V$ converges to π_V. In particular, this also proves the consistency of $\widehat{\pi}_V$ in the case $a_0^2 = 0$, that is $\pi_V = 1$. More precisely, we have the following result.

Lemma 4.6 (*Consistency of $\widehat{\pi}_V$*)**.** *In the large noise case $a_0^2 > 0$, we have*

$$
\widehat{\pi}_V = \pi_V + o_{\mathbb{P}}(1).
$$

In the small noise case $a_0^2 = \eta_0 T/n$, $\eta_0 \geq 0$, with $\mathbb{E}[\phi(Q_0, \theta_0)^2] > 0$, we have

$$
\widehat{\pi}_V = \pi_V + o_{\mathbb{P}}\left(N^{-1}\right).
$$

4.3. Practical guidance to estimate volatility

In this section, we provide a sequence – not theoretically validated but which behaves correctly numerically in next section – to estimate volatility based on the introduced tests. As a matter of fact, the sequence may suffer from the so-called post-model selection problem (as in, e.g Leeb and Pötscher (2005)). This is due to a possible lack of uniformity when pre-testing the presence of residual noise, and may affect the finite sample performance of the volatility estimator constructed from the sequence hereafter. A solution to that issue would be to prove that the proposed inference is uniformly valid with respect to the residual noise magnitude, in a similar way as in Da and Xiu (2017), Section 4.3–4.4. In practice, Section 5 of the present paper suggests that in finite sample, the post-model selection problem does not seem to impact much volatility estimation for the considered models. Moreover, we provide steps – here again completely ad hoc, but implemented in our empirical study – to the empirical researcher to investigate if it is worth considering a specific ϕ when implementing the tests. A rigorous statistical approach to choose ϕ among a class of competitive models in practice based on Bayesian Information Criterion is beyond the scope of this paper and can be found in Clinet and Potiron (2018b).

We suggest the following sequence for volatility estimation:

- If the original Hausman tests from Aït-Sahalia and Xiu (2016) are not rejected, it is reasonably safe to use RV on the raw data, even though it does not necessarily mean that there is no MMN – in our simulation study and empirical study, we find that the tests are rejected (almost) all the time when used at the highest frequency on fairly liquid stocks –. We emphasize that although not reported in our numerical study results, the original tests from Aït-Sahalia and Xiu (2016) when implemented with estimators not robust to autocorrelated MMN (such as RK, PAE, QMLE) are distorted by the presence of MMN of the form $\phi(Q_i, \theta_0) + \epsilon_{t_i}$ with $\phi \neq 0$. Accordingly, in line with our numerical study, we strongly advise the user to implement $\widehat{\sigma}_{err}^2$ as volatility estimator to be compared with RV. This requires a priori to know ϕ and Q_i. Another alternative, which does not need any preestimation, consists in using an estimator robust to autocorrelated MMN such as in Da and Xiu (2017). We did not implement this type of estimator in our numerical study. Finally, we insist on the fact that the theory related to such Hausman tests has not been investigated.
- If the original tests are rejected and the tests considered in this paper are rejected, one should stick to $\widehat{\sigma}_{err}^2$.
- If the original tests are rejected and the tests of this paper are not rejected, then one should use $\widehat{\sigma}_{exp}^2$.

Finally, we recommend the empirical researcher the following steps to choose a specific ϕ prior to implementing the tests on several models:

- The user may implement the original Hausman tests from Aït-Sahalia and Xiu (2016) on the raw data. In agreement with the sequence of volatility estimation, we advise the user to choose a volatility estimator robust to autocorrelated MMN.
- If the results seem to indicate the presence of MMN, the user should estimate the ratio (4.1).
- If the ratio turns out to be close to 100%, then a proper investigation using the tests should be carried out.

To illustrate the method, we follow this procedure in our empirical study.

5. Finite sample performance

We now conduct a Monte Carlo experiment to assess finite sample performance of the tests, and validity of the sequence to estimate volatility described in Section 4.3 – which a priori is subject to multiple testing, model selection and post model selection issues – by comparing it to some leading estimators from the literature. We simulate M = 1000 Monte Carlo days of high-frequency returns where the related horizon time $T = 1/252$ is annualized. One working day corresponds to 6.5 h of trading activity, i.e. 23,400 s.

The efficient price

We introduce the Heston model with U-shape intraday seasonality component and jumps in both price and volatility as

$$dX_t = bdt + \sigma_t dW_t + dJ_t,$$
$$\sigma_t = \sigma_{t-,U}\sigma_{t,SV},$$

where

$$\sigma_{t,U} = C + Ae^{-at/T} + De^{-c(1-t/T)} - \beta\sigma_{\tau-,U}\mathbb{1}_{\{t \geq \tau\}},$$
$$d\sigma_{t,SV}^2 = \alpha(\bar{\sigma}^2 - \sigma_{t,SV}^2)dt + \delta\sigma_{t,SV}d\bar{W}_t,$$

with $b = 0.03$, $dJ_t = \nabla S_t dN_t$, $\nabla = T\bar{\sigma}^2$, the signs of the jumps $S_t = \pm 1$ are i.i.d symmetric, N_t is a homogeneous Poisson process with parameter $\bar{\lambda} = T$ so that the contribution of jumps to the total quadratic variation of the price process is around 50%, $C = 0.75$, $A = 0.25$, $D = 0.89$, $a = 10$, $c = 10$, the volatility jump size parameter $\beta = 0.5$, the volatility jump time τ follows a uniform distribution on $[0, T]$, $\alpha = 5$, $\bar{\sigma}^2 = 0.1$, $\delta = 0.4$, $\bar{W}_t$ is a standard Brownian motion such that $d\langle W, \bar{W}\rangle_t = \bar{\phi}dt$, $\bar{\phi} = -0.75$, $\sigma_{0,SV}^2$ is sampled from a Gamma distribution of parameters $(2\alpha\bar{\sigma}^2/\delta^2, \delta^2/2\alpha)$, which corresponds to the stationary distribution of the CIR process. To obtain more information about the model and values, see Clinet and Potiron (2018a). The model is inspired directly from Andersen et al. (2012) and Aït-Sahalia and Xiu (2016).

The observation times

We consider three levels of sampling: tick by tick, 15 s, 30 s. The observation times are generated regularly except for the tick by tick case. For the latter, we assume that $\alpha_t = 1/(e^{\beta_1} + \{e^{\beta_2} + e^{\beta_3}\}^2(t/T - e^{\beta_2}/(e^{\beta_2} + e^{\beta_3}))^2)$, and that U_i are following an exponential distribution with parameter $2T/23,400$. We have that the rate of arrival times α_t^{-1} exhibits a usual U-shape intraday pattern, as pointed out in Engle and Russell (1998) (see discussions in Section 5–6 and Figure 2) and Chen and Hall (2013) (see Section 5, pp. 1011–1017). We fix $\beta_1 = -0.84$, $\beta_2 = -0.26$ and $\beta_3 = -0.39$ following the empirical values exhibited in Clinet and Potiron (2018c), which implies that the sampling frequency is on average faster than one second.

The information

We implement two models: Roll model and the signed spread model. As in the simulation study from Li et al. (2016), the trade indicator I_i is simulated featuring a Bernoulli process with parameter $p = 1/2$ and with an autocorrelation chosen equal to 0.3. We fix the parameter $\theta = 0.0001$ in the case of Roll model. For the signed spread model, we further simulate the spread $\bar{S}_i$ as an AR(1) process with mean 0.000125, variance 10^{-10} and correlation parameter which amounts to 0.6. The parameter is chosen equal to $\theta = 0.80$. The values of the parameters correspond roughly to the fitted values.[10]

The residual noise

To assess finite sample performance of the tests, we consider two types of (finite sample) alternative. In $\mathcal{H}_1$, we assume that the residual noise is i.i.d normally distributed with zero-mean and variance $a_0^2 = 10^{-9}, 10^{-8}, 10^{-7}$. In $\mathcal{H}_2$, which in particular does not accommodate with the assumptions of this paper, we assume that the residual noise

$$\epsilon_{t_i} = \left(\frac{a_0}{\sqrt{3}} + \nu N_1\right)(\text{sign}(\Delta X_i)|N_2| + I_i|N_3| + N_4),$$

where N_1, N_2, N_3 and N_4 are standard independent normally distributed variables, and $\nu = a_0^2$. With that specification, the residual noise has also variance approximately equal to a_0^2 (with $a_0^4 \ll a_0^2$), is serially correlated (since I_i is serially correlated), heteroskedastic, endogenous as both correlated with the efficient returns and the explicative part of the MMN.

To validate the sequence to estimate volatility, we consider $a_0^2 = 0, 10^{-9}$, mix, where the latter corresponds to a setup where there is no residual noise for half of the days in the sample and a residual noise with variance $a_0^2 = 10^{-9}$ for the remaining half in the sample.

---

[10] Although not fully reported in the empirical study.

Remaining tuning parameters

Although the likelihood-based estimators do not require any tuning parameter, we need to select some parameters for the truncation method used when computing S_3 and S_5. We choose $k = \lfloor n^{1/2} \rfloor$, $\omega = 0.48$, $\widetilde{\alpha} = \alpha_0 \widehat{\sigma}_{exp}$, $\alpha_0 = 4$, and $\widetilde{k} = \lfloor N^{1/2} \rfloor$, consistently with Aït-Sahalia and Xiu (2016) (except for $\alpha_0 = 4$ which was set equal to 3, because this was yielding too many jumps detection in our case).

Concurrent volatility estimators and simulated model considered for comparison

We consider a group of eight concurrent volatility estimators which is a mix of estimators considered in this paper and leading estimators from the literature. S corresponds to the sequence introduced in Section 4.3. The QMLEexp is $\widehat{\sigma}_{exp}^2$, and is actually equal to estimated-price RV defined in Li et al. (2016) since both ϕ considered are linear. The QMLEerr is defined as $\widehat{\sigma}_{err}^2$. The E-QMLE is the two-step estimator with price estimation first and (regular) QMLE on the price estimates. We also have some popular estimators QMLE, PAE, RK, and RV.[11] In particular, excluding S, no tests are applied prior or post to the estimation methods.

The simulated model considered features time-varying volatility but does not incorporate jumps in the price process as most methods considered are not robust to such environment. Moreover, the sampling times are regular (with high frequency sampling every second) since some methods may be badly affected when they are not.

Results

We first discuss the results to assess finite sample performance of the tests. We compute S_1 and S_2 when dealing with the tick by tick simulated returns, and S_3, S_4 and S_5, which are introduced in the appendix, when looking at sparser observations. We report in Table 2 and in Table 3 the fraction of rejections of $\mathcal{H}_0$ at the 0.05 level for different scenarios. The statistics have desired fraction of rejections, and the power is reasonable (it is actually slightly better when the residual noise from the alternative has a general form which actually breaks the theoretical assumptions of this paper). There are two important lessons to take from this part. First, the power is much more satisfactory in the tick by tick case, and thus we insist that the high frequency data user should make inference using all the data available. Second, depending on the simulated scenario, i.e. constant volatility, time-varying volatility including jumps or not, the related statistic behave slightly better than the other statistics. One should confirm accordingly the type of data at hand prior to choosing the related statistic to use. This will heavily depend on data pre-processing, such as controlling for diurnal pattern in volatility (see Christensen et al. (2018)) and/or removing jumps.

We now discuss about the validity of the sequence given in Section 4.3. Table 4 reports the bias, standard deviation and RMSE of the eight concurrent volatility estimators for the three scenarios, i.e. no residual noise, residual noise and a mix of both aforementioned scenarios. As expected from Theorem 3.1, QMLEexp leads the cohort when there is no residual noise. QMLEerr has approximately a standard deviation $\sqrt{3}$ times as big as that of QMLEexp, which is in line with the theorem. The sequence using the tests' RMSE is very close to that of QMLEexp, although slightly bigger, which is due to the fact that the test to assess if the MMN is fully explained by some variables of the limit order book are (falsely) rejected one time over twenty. In case of non-zero residual noise, QMLEerr performs the best which is not surprising since the estimation procedure is residual noise robust. The sequence using the tests' RMSE is almost the same as that of QMLEerr, which is explained by the fact that the tests from this paper are (rightly) rejected 499 days over 500. On the other hand, QMLEexp suffers since it is not residual noise robust. Overall, the sequence using the tests is leading the group (in terms of RMSE) in the mix scenario (which is the most realistic case). In particular, note that the first step in the sequence using the original tests of Aït-Sahalia and Xiu (2016) implemented with QMLEerr as volatility estimator compared with RV has no distortion on the finite sample results since tests are rejected all the time. QMLEerr comes second while being relatively close from the sequence using the tests. E-QMLE is virtually tied with QMLEerr, which can be explained by the fact that they are asymptotically equivalent. The other estimators perform more poorly.

6. Empirical study

Our main dataset consists of one calendar month (April 2011) of trades and quotes for thirty one CAC 40 constituents traded on the Euronext NV. The data for individual stocks were obtained from the TAQ data and the Order book data.[12] We keep quotes corresponding to best bid/ask price, which are often referred to as Level 1 data. To obtain the information of trade type, we implement Section 3.4 in Muni Toke (2016).[13] The timestamp is rounded to the nearest millisecond.

To prevent from opening and closing effects, we restrict our dataset starting each day at 9:30am and ending at 4pm. We consider the data in tick time, for an average of 3000 daily trades and a quote/trade ratio bigger than 20. The most active days include more than 10,000 trades, whereas the less liquid days are around 500 trades. Descriptive statistics on the individual stocks are detailed on Table 5. Using the regular QMLE restricted to $\phi = 0$, we find that the MMN variance lies within 1.30×10^{-9} and 5.60×10^{-8}, taking the value 1.63×10^{-8} on average.

---

[11] Details on the choice of tuning parameters for the PAE and the RK can be obtained upon request to the authors.

[12] The data were obtained through Reuters and provided by the Chair of Quantitative Finance of Ecole Centrale Paris.

[13] The code is available on our websites. A comparison with the simple and popular Lee–Ready procedure introduced in Lee and Ready (1991) can be consulted in Section 5 of the cited paper.

Table 2

Simulation study results: fraction of rejections of the null hypothesis when the true scenario is the null hypothesis $\mathcal{H}_0$ or the alternative with i.i.d residual noise $\mathcal{H}_1$ at the 0.05 level. Note that $\xi^2 = a_0^2/\sqrt{T\int_0^T \sigma_u^4 du}$ corresponds to the noise-to-signal ratio.

Sampling frequency			Tick by tick		15 s			30 s		
	a_0^2	ξ^2	S_1	S_2	S_3	S_4	S_5	S_3	S_4	S_5
Roll model										
Constant volatility										
$\mathcal{H}_0$	0	0	0.05	0.07	0.05	0.05	0.05	0.04	0.04	0.04
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.05	0.05	0.05	0.03	0.03	0.03
$\mathcal{H}_1$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.33	0.33	0.34	0.08	0.08	0.08
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	1.00	1.00	1.00	0.98	0.99	0.99
Time-varying volatility and no price jump										
$\mathcal{H}_0$	0	0	0.05	0.07	0.18	0.05	0.09	0.16	0.06	0.09
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.18	0.07	0.11	0.17	0.05	0.10
$\mathcal{H}_1$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.63	0.52	0.55	0.25	0.15	0.19
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	1.00	1.00	1.00	0.98	0.97	0.97
Time-varying volatility and price jump										
$\mathcal{H}_0$	0	0	0.03	0.05	0.09	0.01	0.05	0.08	0.03	0.06
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.09	0.03	0.06	0.08	0.03	0.07
$\mathcal{H}_1$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.30	0.22	0.29	0.10	0.05	0.09
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	0.99	0.41	0.88	0.75	0.35	0.57
Signed spread model										
Constant volatility										
$\mathcal{H}_0$	0	0	0.05	0.08	0.05	0.05	0.05	0.04	0.04	0.04
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.05	0.06	0.06	0.03	0.04	0.04
$\mathcal{H}_1$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.30	0.30	0.30	0.07	0.07	0.07
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	1.00	1.00	1.00	0.98	0.98	0.98
Time-varying volatility and no price jump										
$\mathcal{H}_0$	0	0	0.05	0.08	0.14	0.04	0.07	0.16	0.05	0.08
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.17	0.06	0.09	0.16	0.06	0.09
$\mathcal{H}_1$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.62	0.52	0.53	0.26	0.15	0.18
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	1.00	1.00	1.00	0.98	0.97	0.97
Time-varying volatility and price jump										
$\mathcal{H}_0$	0	0	0.04	0.05	0.06	0.01	0.04	0.08	0.02	0.06
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.07	0.03	0.05	0.08	0.02	0.05
$\mathcal{H}_1$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.28	0.21	0.26	0.13	0.07	0.11
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	1.00	0.38	0.87	0.77	0.37	0.59

We first implement the tests of Aït-Sahalia and Xiu (2016) on the observed price on tick-by-tick data. The results can be found on Table 8. The six tests consistently indicate that we reject those tests almost all the time, indicating that there seems to be MMN at the highest frequency for the stocks and days considered. Accordingly, we report in Table 6 the measure of goodness of fit of several leading models: Roll, Glosten–Harris, signed timestamp, signed spread model, signed quoted depth, the order flow imbalance, a linear combination of all the aforementioned models and a non-linear signed spread. The signed spread model incontestably dominates with an astonishing proportion of variance explained estimated around 99%. This dominance is in fact consistent across sampling frequencies, stocks and over time, although not fully reported. Finally, this measure stagnates when sampling at sparser frequencies, and we argue that this is because there is (almost) no remaining noise.

Among the concurrent models, the goodness of fit of Roll model is very decent with a four fifth proportion of variance explained at the highest frequency and increasing when diminishing the sampling frequency. Yet this feature hints that the model cannot be considered as reasonably free of residual noise when using tick by tick data. The measure related to Glosten–Harris model is slightly bigger, suggesting that the information on the volume helps to improve the fit to a certain extent. Those estimated values are in line with the results discussed in the empirical study of Li et al. (2016). The fit is not as good on other models. We have tried many other alternative models (such as linear combinations of the aforementioned models) but have not found any significant improvement in the fit of the signed spread model, as reported in Table 6. In particular adding a Roll component to it was not found to improve the fit much.

We further investigate if the signed spread model can fairly be considered as free from residual noise by implementing the two tick-by-tick-robust Hausman tests. For each individual stock and test, the fraction of rejection at the 0.05 level is reported on Table 7. Although not reported, the results are very similar when using the other three statistics. Over the thirty one stocks, the averaged-across-tests fraction lies within 0.00 and 0.11 for twenty eight constituents (hereafter denoted as the main

Table 3
Simulation study results: fraction of rejections of the null hypothesis when the true scenario is the alternative with serially dependent, endogenous and heteroskedastic residual noise $\mathcal{H}_2$ at the 0.05 level. Note that $\xi^2 = a_0^2/\sqrt{T\int_0^T \sigma_u^4 du}$ corresponds to the noise-to-signal ratio.

Sampling frequency			Tick by tick		15 s			30 s		
	a_0^2	ξ^2	S_1	S_2	S_3	S_4	S_5	S_3	S_4	S_5
Roll model										
Constant volatility										
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.07	0.07	0.07	0.06	0.07	0.07
$\mathcal{H}_2$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.43	0.43	0.43	0.14	0.14	0.14
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	1.00	1.00	1.00	0.99	0.99	0.99
Time-varying volatility and no price jump										
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.22	0.09	0.13	0.21	0.08	0.14
$\mathcal{H}_2$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.75	0.63	0.56	0.30	0.19	0.22
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	1.00	1.00	1.00	0.99	0.99	0.99
Time-varying volatility and price jump										
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.12	0.05	0.08	0.10	0.05	0.08
$\mathcal{H}_2$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.34	0.25	0.33	0.12	0.07	0.10
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	1.00	0.50	0.99	0.83	0.41	0.63
Signed spread model										
Constant volatility										
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.07	0.08	0.08	0.07	0.07	0.07
$\mathcal{H}_2$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.40	0.40	0.40	0.14	0.15	0.15
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	1.00	1.00	1.00	0.99	0.99	0.99
Time-varying volatility and no price jump										
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.23	0.11	0.15	0.21	0.10	0.14
$\mathcal{H}_2$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.73	0.62	0.63	0.32	0.20	0.24
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	1.00	1.00	1.00	0.99	0.99	0.99
Time-varying volatility and price jump										
	10^{-9}	$\approx 10^{-6}$	1.00	1.00	0.10	0.05	0.07	0.11	0.04	0.07
$\mathcal{H}_2$	10^{-8}	$\approx 10^{-5}$	1.00	1.00	0.34	0.26	0.31	0.16	0.10	0.15
	10^{-7}	$\approx 10^{-4}$	1.00	1.00	1.00	0.46	0.99	0.95	0.50	0.72

Table 4
Simulation study results: comparison of eight volatility estimators. Note that $\xi^2 = a_0^2/\sqrt{T\int_0^T \sigma_u^4 du}$ corresponds to the noise-to-signal ratio, and that e^x stands for 10^x.

a_0^2	0			e^{-9}			mix		
ξ^2	0			$\approx e^{-6}$			mix		
Estimator	Bias	Stdv.	RMSE	Bias	Stdv.	RMSE	Bias	Stdv.	RMSE
Time-varying volatility and no price jump, *tick by tick*									
Roll model									
S	$1.71e^{-7}$	$3.89e^{-6}$	$3.89e^{-6}$	$1.33e^{-7}$	$6.30e^{-6}$	$6.30e^{-6}$	$1.52e^{-7}$	$5.24e^{-6}$	$5.24e^{-6}$
QMLEexp	$-5.65e^{-8}$	$3.10e^{-6}$	$3.10e^{-6}$	$4.67e^{-5}$	$3.48e^{-6}$	$4.68e^{-5}$	$2.38e^{-5}$	$2.31e^{-5}$	$3.32e^{-5}$
QMLEerr	$7.60e^{-8}$	$5.93e^{-6}$	$5.93e^{-6}$	$1.02e^{-7}$	$6.12e^{-6}$	$6.12e^{-6}$	$8.90e^{-8}$	$6.03e^{-6}$	$6.03e^{-6}$
EQMLE	$7.61e^{-8}$	$5.93e^{-6}$	$5.93e^{-6}$	$1.04e^{-7}$	$6.12e^{-6}$	$6.12e^{-6}$	$8.91e^{-8}$	$6.03e^{-6}$	$6.03e^{-6}$
QMLE	$6.13e^{-5}$	$1.50e^{-5}$	$6.31e^{-5}$	$5.76e^{-5}$	$1.53e^{-5}$	$5.96e^{-5}$	$5.94e^{-5}$	$1.55e^{-5}$	$6.14e^{-5}$
PAE	$-2.88e^{-6}$	$2.99e^{-5}$	$3.00e^{-5}$	$-2.88e^{-6}$	$2.99e^{-5}$	$3.00e^{-5}$	$-2.88e^{-6}$	$2.99e^{-5}$	$3.00e^{-5}$
RK	$6.10e^{-5}$	$1.40e^{-5}$	$6.25e^{-5}$	$5.86e^{-5}$	$1.46e^{-5}$	$6.04e^{-5}$	$5.98e^{-5}$	$1.39e^{-5}$	$6.14e^{-5}$
RV	$3.27e^{-4}$	$5.87e^{-6}$	$3.27e^{-4}$	$3.74e^{-4}$	$6.48e^{-6}$	$3.74e^{-4}$	$3.51e^{-4}$	$6.33e^{-6}$	$3.51e^{-4}$
Spread model									
S	$1.71e^{-7}$	$3.89e^{-6}$	$3.89e^{-6}$	$1.33e^{-7}$	$6.30e^{-6}$	$6.30e^{-6}$	$1.52e^{-7}$	$5.24e^{-6}$	$5.24e^{-6}$
QMLEexp	$-5.66e^{-8}$	$3.11e^{-6}$	$3.11e^{-6}$	$4.67e^{-5}$	$3.48e^{-6}$	$4.68e^{-5}$	$2.38e^{-5}$	$2.31e^{-5}$	$3.32e^{-5}$
QMLEerr	$7.62e^{-8}$	$5.93e^{-6}$	$5.93e^{-6}$	$1.02e^{-7}$	$6.12e^{-6}$	$6.12e^{-6}$	$8.91e^{-8}$	$6.03e^{-6}$	$6.03e^{-6}$
EQMLE	$7.62e^{-8}$	$5.93e^{-6}$	$5.93e^{-6}$	$1.04e^{-7}$	$6.11e^{-6}$	$6.11e^{-6}$	$8.92e^{-8}$	$6.03e^{-6}$	$6.03e^{-6}$
QMLE	$6.15e^{-5}$	$1.51e^{-5}$	$6.33e^{-5}$	$5.78e^{-5}$	$1.53e^{-5}$	$5.98e^{-5}$	$5.96e^{-5}$	$1.56e^{-5}$	$6.16e^{-5}$
PAE	$-2.87e^{-6}$	$2.99e^{-5}$	$3.00e^{-5}$	$-2.87e^{-6}$	$2.99e^{-5}$	$3.00e^{-5}$	$-2.87e^{-6}$	$2.99e^{-5}$	$3.00e^{-5}$
RK	$6.11e^{-5}$	$1.41e^{-5}$	$6.27e^{-5}$	$5.88e^{-5}$	$1.47e^{-5}$	$6.06e^{-5}$	$6.00e^{-5}$	$1.44e^{-5}$	$6.17e^{-5}$
RV	$3.31e^{-4}$	$5.95e^{-6}$	$3.31e^{-4}$	$3.78e^{-4}$	$6.51e^{-6}$	$3.78e^{-4}$	$3.55e^{-4}$	$6.38e^{-6}$	$3.55e^{-4}$

Table 5
Descriptive statistics.

| Name	Ticker	Nb. Trades (daily)	Price	Tick	Price/Tick ($\times 10^{-3}$)	Spread (in ticks)
Accor	ACCP	1900	31.87	0.005	6.37	3.04
Air Liquide	AIRP	4116	94.64	0.01	9.46	2.53
Alstom	ALSO	2818	42.55	0.005	8.51	2.75
Alcatel	ALUA	3667	4.13	0.001	4.13	1.52
AXA	AXAF	3156	15.03	0.005	3.01	1.25
Bouygues	BOUY	2275	34.13	0.005	6.83	1.80
Credit Agricole	CAGR	3014	11.74	0.005	2.35	1.17
Cap Gemini	CAPP	2002	40.83	0.005	8.17	2.04
Danone	DANO	3525	46.30	0.005	9.26	2.07
Airbus	EAD	1589	20.95	0.005	4.19	2.01
Essilor International	ESSI	1090	52.80	0.01	5.28	2.43
France Telecom	FTE	2327	15.76	0.005	3.16	1.16
Engie	GSZ	3513	28.26	0.005	5.65	1.68
Lafarge	LAFP	3344	44.36	0.005	8.87	3.58
Louis Vuitton	LVMH	2073	112.51	0.05	2.25	1.15
L'Oreal	OREP	2582	82.99	0.01	8.30	2.55
Peugeot	PEUP	2129	28.17	0.005	5.63	2.77
Kering	PRTP	961	109.15	0.05	2.18	1.22
Publicis Groupe	PUBP	2412	40.14	0.005	8.03	1.98
Renault	RENA	3728	39.43	0.005	7.89	2.78
Sanofi	SASY	3622	50.14	0.01	5.01	1.35
Schneider Electric	SCHN	2638	121.66	0.05	2.43	1.15
Suez	SEVI	1115	14.52	0.005	2.90	1.57
Vinci	SGEF	3133	44.44	0.005	8.89	2.76
Societe Generale	SOGN	6710	46.40	0.005	9.28	3.14
STMicroelectronics	STM	2921	8.72	0.001	8.72	3.15
Technip	TECF	3192	75.78	0.01	7.58	2.13
Total	TOTF	5909	43.09	0.005	8.62	1.80
Unibail Rodamco	UNBP	1440	153.95	0.05	3.08	1.28
Veolia Environnement	VIE	1724	21.97	0.005	4.39	1.51
Vivendi	VIV	2710	20.35	0.005	4.07	1.26
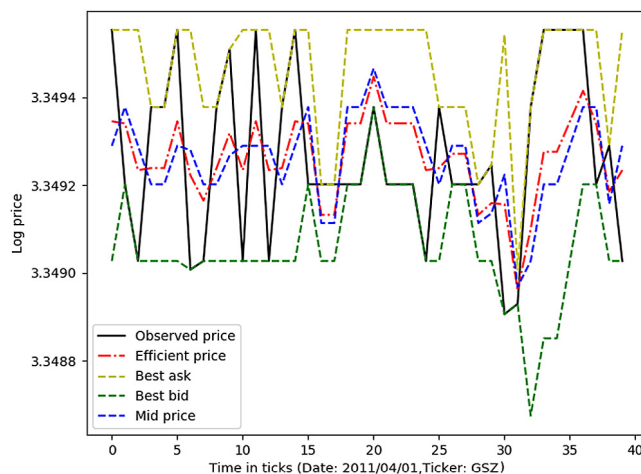
Fig. 1. An example of estimated efficient price related to the signed spread model.

group, and which features a proportion of variance explained bigger than 99%), whereas picking at 0.11, 0.16 and 0.42 for the three remaining components (henceforth referred as the minor group, which features a proportion of variance explained slightly below 95%). The observed fraction of rejections of the main group constituents can be considered as reasonably close to the theoretical threshold 0.05 hence we can not reject the null hypothesis for them and this indicates that stocks from the main group can be considered as fairly free from residual noise. On the contrary the rejection is clear for the three stocks from the minor group.

An example of estimated efficient price can be seen on Fig. 1. To explore the properties of the estimated price, we recognize it as the given observed price to be tested in Aït-Sahalia and Xiu (2016). Although their tests are by nature related to our

Table 6

Measure of goodness of fit of several leading models.[a]

Model for the explicative part	MMN variance	residual noise variance	Proportion of variance explained
Sampling frequency: tick by tick			
null	1.63×10^{-8}	1.63×10^{-8}	0.00%
Roll	2.27×10^{-8}	3.89×10^{-9}	83.84%
Glosten–Harris	2.28×10^{-8}	3.40×10^{-9}	85.97%
signed timestamp	1.82×10^{-8}	1.45×10^{-8}	20.98%
signed spread	2.78×10^{-8}	2.45×10^{-10}	99.19%
signed quote depth	2.02×10^{-8}	7.95×10^{-9}	59.71%
order flow imbalance	1.61×10^{-8}	1.55×10^{-8}	3.86%
NL signed spread	2.77×10^{-8}	2.44×10^{-10}	99.17%
general	2.79×10^{-8}	1.88×10^{-10}	99.36%
Sampling frequency $\approx$ 15 s			
null	1.51×10^{-8}	1.51×10^{-8}	0.00%
Roll	2.08×10^{-8}	2.67×10^{-9}	88.67%
Glosten–Harris	2.11×10^{-8}	2.35×10^{-9}	90.09%
signed timestamp	1.58×10^{-8}	1.39×10^{-8}	15.27%
signed spread	2.61×10^{-8}	2.84×10^{-10}	99.04%
signed quote depth	1.87×10^{-8}	7.02×10^{-9}	64.68%
order flow imbalance	1.50×10^{-8}	1.44×10^{-8}	4.11%
NL signed spread	2.61×10^{-8}	2.84×10^{-10}	99.04%
general	2.63×10^{-8}	2.13×10^{-10}	99.22%
Sampling frequency $\approx$ 30 s			
null	1.24×10^{-8}	1.24×10^{-8}	0.00%
Roll	1.98×10^{-8}	3.87×10^{-9}	92.17%
Glosten–Harris	2.03×10^{-8}	1.63×10^{-9}	93.34%
signed timestamp	1.33×10^{-8}	1.20×10^{-8}	21.80%
signed spread	2.46×10^{-8}	3.99×10^{-10}	98.01%
signed quote depth	1.71×10^{-8}	5.66×10^{-9}	71.77%
order flow imbalance	1.19×10^{-8}	1.12×10^{-8}	5.63%
NL signed spread	2.45×10^{-8}	3.98×10^{-10}	98.00%
general	2.47×10^{-8}	3.75×10^{-10}	98.41%

[a]The MMN variance, the residual noise variance and the proportion of variance explained are estimated using $\widehat{\pi}_V$ and averaged across the thirty one constituents of the CAC 40 and the 19 days in April 2011.

tests, the estimators they are using differ from the ones considered in our work, and thus their tests can be regarded as a sensibly independent check on the efficiency of the estimated price. The fraction of rejections of their null hypothesis at the 0.05 level can be consulted on Table 8. When restricting to the main group constituents, the six tests range from 0.05 to 0.09 and are equal to 0.06 on average. When considering the minor group stocks, the same tests range from 0.33 to 0.42. This largely corroborates the fact that the main group stocks are most likely free from residual noise whereas the minor group constituents cannot be considered as such. One feature common to the minor group constituents, namely France Telecom, Louis Vuitton and Schneider Electric, is that they are stocks with large ticks and small spread, which is almost always equal to 1 tick. This feature can be seen on Table 5, as the three stocks share the top 3 in terms of smallest spread, and part of the top 5 when looking at the smallest ratio of price over tick size. Even for those stocks, our findings strongly indicate that the estimated price is much closer from efficiency than the observed price on which the tests are rejected 100% of the time.

Another way to question the efficiency of the estimated price consists in inspecting the first lag of the autocorrelation function and the visual "signature plot" procedure of Andersen et al. (2000) (see also Patton (2011)). This can be seen on Figs. 2–3. A satisfactory amelioration of the first lag of the autocorrelation function is noticeable, as it averages 0.02 when looking at the estimated price whereas -0.28 when taking the observed price. The signature plot is also acceptable as it is relatively flat.

Finally, the maximum likelihood estimation in both settings delivers very similar (the difference is at most equal to 10^{-3} even when considering the three stocks from the minor group) and stable estimates of the parameter which lies systematically between 0.60 and 0.90 with an average around 0.79 and a standard deviation slightly above 0.03 when sampling at the highest frequency. The estimation is also stable across sampling frequencies as the average values are 0.77 and 0.76 when considering the 15-second and 30-second frequency, respectively. Moreover, Fig. 4 documents that the daily estimates averaged across stocks are also relatively stable over time.

Note that we can also consider as in Chaker (2017, p. 15) the finite sample correction by instrumental variables. This slightly shifts estimation of the parameter towards the origin, with a decrease of magnitude one quarter of its value in

Table 7
Fraction of rejections of the null hypothesis at the 0.05 level for the signed spread model.[a]

Ticker	Est. Par.	Var. Ex.	Aver.	S_1	S_2
Sampling frequency: tick by tick					
ACCP	0.84	99.78%	0.00	0.00	0.00
AIRP	0.79	100.00%	0.00	0.00	0.00
ALSO	0.84	99.69%	0.00	0.00	0.00
ALUA	0.82	100.00%	0.05	0.05	0.05
AXAF	0.73	99.63%	0.00	0.00	0.00
BOUY	0.79	100.00%	0.05	0.05	0.05
CAGR	0.74	99.70%	0.05	0.05	0.05
CAPP	0.85	100.00%	0.11	0.11	0.11
DANO	0.82	100.00%	0.11	0.11	0.11
EAD	0.81	99.96%	0.03	0.00	0.05
ESSI	0.82	99.89%	0.00	0.00	0.00
FTE	0.74	93.92%	0.16	0.16	0.16
GSZ	0.80	99.56%	0.00	0.00	0.00
LAFP	0.82	100.00%	0.00	0.00	0.00
LVMH	0.72	94.45%	0.11	0.11	0.11
OREP	0.80	99.90%	0.03	0.00	0.05
PEUP	0.84	99.48%	0.00	0.00	0.00
PRTP	0.72	98.61%	0.00	0.00	0.00
PUBP	0.82	100.00%	0.00	0.00	0.00
RENA	0.85	99.89%	0.00	0.00	0.00
SASY	0.75	99.72%	0.00	0.00	0.00
SCHN	0.68	92.47%	0.42	0.42	0.42
SEVI	0.75	99.97%	0.00	0.00	0.00
SGEF	0.85	99.75%	0.00	0.00	0.00
SOGN	0.87	99.87%	0.00	0.00	0.00
STM	0.87	100.00%	0.00	0.00	0.00
TECF	0.75	99.99%	0.05	0.05	0.05
TOTF	0.78	100.00%	0.00	0.00	0.00
UNBP	0.69	99.23%	0.00	0.00	0.00
VIE	0.77	99.95%	0.00	0.00	0.00
VIV	0.72	99.63%	0.00	0.00	0.00
Average	0.79	99.19%	0.06	0.06	0.06

[a]The values are averaged across the 19 days in April 2011. The column "Est. Par." stands for the estimated parameter and "Var. Ex." for "Proportion of variance explained". The column "Aver." corresponds to the fraction of rejections averaged across the two tests computed using tick by tick data.

Table 8
Fraction of rejections for the estimated efficient price related to the signed spread model at the 0.05 level.[a]

Price	Aver.	H_1	H_2	H_3	T	AC	H_4
Main group							
Est. price	0.06	0.08	0.05	0.05	0.05	0.09	0.04
Minor group							
Est. price	0.38	0.42	0.35	0.39	0.39	0.39	0.33
Overall							
Est. price	0.09	0.11	0.07	0.08	0.08	0.12	0.07
Mid price	0.25	0.31	0.21	0.25	0.25	0.29	0.19
Obs. price	1.00	1.00	1.00	1.00	1.00	1.00	0.99

[a]The values are averaged across the 19 days in April 2011. The column "Aver." corresponds to the fraction of rejections averaged across the six tests. The six tests from (Aït-Sahalia and Xiu, 2016) are implemented at the tick by tick frequency considering the estimated efficient price as the given observed price to be tested. We report the proportion of rejections for each constituent when considering the estimated price. We also report in "Obs. price" (respectively "Mid price") the proportion of rejections when considering the observed price (resp. the mid price). Finally, note that the main group corresponds to stocks a priori free of residual noise given our implemented tests, whereas the minor group stands for stocks with residual noise.

the estimates. As it can improve finite sample properties, we implemented and chose to work with such finite sample correction, including in our implemented tests. Finally, we have considered variance estimators based on raw returns instead of estimated price returns for stability reason.
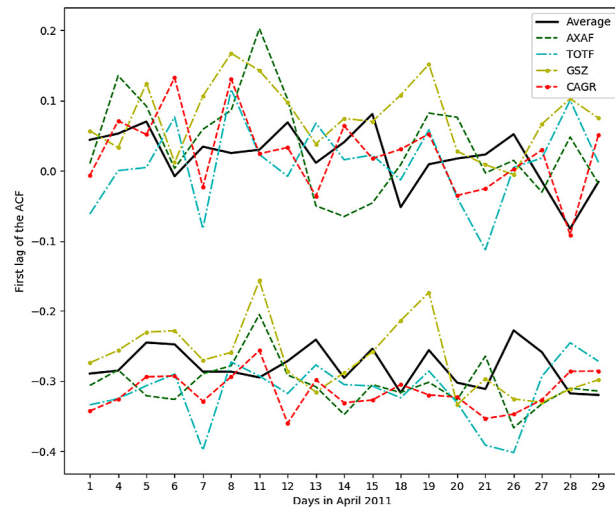
Fig. 2. First lag of the autocorrelation function over time related to the estimated efficient price with the signed spread model (above) and the observed price (below) for the across-the-assets average and 4 random assets.
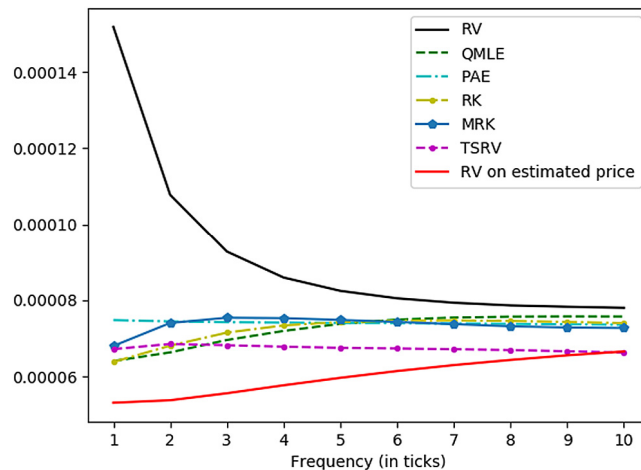

Fig. 3. Average daily signature plot for the CAC 40 constituents. Seven volatility estimators are implemented: RV, QMLE, PAE, RK, MRK (i.e. a modified realized kernel-based estimator robust to autocorrelated-noise considered in Varneskov (2017)), TSRV and RV on the estimated efficient price related to the signed spread model.

7. Conclusion

The paper introduces tests to assess if the market microstructure noise can be fully explained by the informational content of some variables from the limit order book. Two novel quasi-maximum likelihood estimators are extensively studied in the development. Subsequently, based on a common procedure the paper proposes an efficient price estimator.

We emphasize that the method can be easily implemented to assist anyone who is working with high frequency data. The empirical study should be taken as a reference for repeating the exercise, i.e. first testing among a class of candidates and then choosing one specific model based on the measure of goodness of fit. We hope this provides an alternative and reliable solution to the common dilemma between subsampling and using sophisticated noise-robust estimators.

We also call attention to the fact that when the market microstructure noise is fully explained by the limit order book, other quantities beyond quadratic variation, such as pure integrated volatility (by truncation), integrated powers of volatility, high-frequency covariance or even volatility of volatility can be estimated following the same procedure as investigated in our recent work (Clinet and Potiron, 2017).

Finally, although we have checked that there is no major distortion of our tests in finite sample and that they can be useful to improve the precision of volatility estimation, challenging and interesting avenues for future research include a possible improvement of the method checking whether, or not, the testing procedure suffers from the classical post-model selection issue as presented in Leeb and Pötscher (2005).
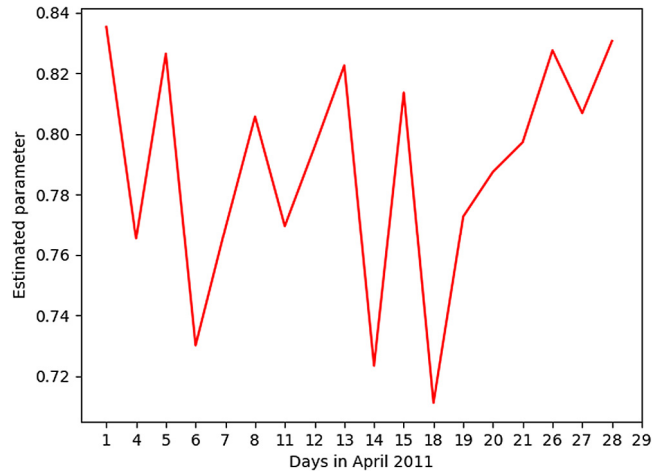
Fig. 4. Estimated parameter related to the signed spread model over time average across the CAC 40 constituents.

Appendix A. Definition of supplementary variance estimators when observations are regular

In this section, we provide supplementary variance estimators in the case of regular observations. We consider the three aforementioned scenarios, i.e.

  (i) constant volatility
  (ii) time-varying volatility and no price jump
  (iii) time-varying volatility and price jump

In the case (i), we have from Theorem 3.1 that $AVAR\big(\widehat{\sigma}^2_{exp} - \widehat{\sigma}^2_{err}\big) = 4\sigma_0^4$. This can be simply estimated by

$$\widehat{V}_3 = 4(\widehat{\sigma}^2_{exp})^2. \tag{A.1}$$

Under (ii), we have $AVAR\big(\widehat{\sigma}^2_{exp} - \widehat{\sigma}^2_{err}\big) = 4T^{-1}\int_0^T \sigma_s^4 ds$ which can be estimated by:

$$\widehat{V}_4 = \frac{4n}{3T^2} \sum_{i=1}^{n} \Delta\widehat{X}_i^4, \text{ with } \Delta\widehat{X}_i = \widehat{X}_{t_i} - \widehat{X}_{t_{i-1}}, \tag{A.2}$$

where $\widehat{X}_{t_i}$ was previously defined in (1.4). When assuming (iii), we have $AVAR\big(\widehat{\sigma}^2_{exp} - \widehat{\sigma}^2_{err}\big) = 4T^{-1}\big\{\int_0^T \sigma_s^4 ds + \sum_{0<s\leq T} \Delta J_s^2(\sigma_s^2 + \sigma_{s-}^2)\big\}$. If we introduce $k \to \infty$ such that $k\Delta \to 0$ and $u = \widetilde{\alpha}\Delta^\omega$ with $0 < \omega < 1/2$, and $\widetilde{\alpha} > 0$, the asymptotic variance can be estimated via

$$\widehat{V}_5 = \frac{4}{T}\left\{\frac{1}{3\Delta} \sum_{i=1}^{n} \Delta\widehat{X}_i^4 \mathbf{1}_{\{|\Delta\widehat{X}_i|\leq u\}} + \sum_{i=k+1}^{n-k} \Delta\widehat{X}_i^2 \mathbf{1}_{\{|\Delta\widehat{X}_i|>u\}}\big(\breve{\sigma}^2_{t_i} + \breve{\sigma}^2_{t_i-}\big)\right\}, \text{ where} \tag{A.3}$$

$$\breve{\sigma}^2_{t_i} = \frac{1}{k\Delta} \sum_{j=i+1}^{i+k} \Delta\widehat{X}_j^2 \mathbf{1}_{\{|\Delta\widehat{X}_j|\leq u\}}, \breve{\sigma}^2_{t_i-} = \breve{\sigma}^2_{t_{i-k-1}}.$$

The estimator $\widehat{V}_3$ is based on the truncation method considered in Mancini (2009). The three variances estimators $\widehat{V}_i$ for $i = 3, 4, 5$ are identical to the ones introduced in Aït-Sahalia and Xiu (2016), up to a scaling factor T. It is due to the fact that the authors scale their Hausman test statistics by Δ_n^{-1} whereas we used N instead. Those three estimators satisfy the conditions of Proposition 3.3 and Corollary 3.4. In the corresponding proofs, we also show for the case $i = 3, 4, 5$.

Appendix B. Expression of the asymptotic variance terms defined in Theorem 3.1

We recall that

$$\overline{\sigma}_0^2 = T^{-1}\left(\int_0^T \sigma_s^2 ds + \sum_{0<s\leq T} \Delta J_s^2\right), \tilde{\eta}_0 = \left(T^{-1}\int_0^T \alpha_s^{-1} ds\right)\eta_0,$$

$$\widetilde{\mathcal{K}} = \left(T^{-1}\int_0^T \alpha_s^{-1} ds\right)^2 \mathcal{K},$$

$$\mathcal{Q} = T^{-1} \int_0^T \alpha_s^{-1} ds \left\{ \int_0^T \sigma_s^4 \alpha_s ds + \sum_{0 < s \le T} \Delta J_s^2 (\sigma_s^2 \alpha_s + \sigma_{s-}^2 \alpha_{s-}) \right\}.$$

Moreover, we also have

$$\phi_0 = 1 - \frac{1}{2\tilde{\eta}_0} \left\{ \sqrt{\overline{\sigma}_0^2 (4\tilde{\eta}_0 + \overline{\sigma}_0^2)} - \overline{\sigma}_0^2 \right\}, \tag{B.1}$$

$$\gamma_0^2 = \frac{1}{2} \left\{ 2\tilde{\eta}_0 + \overline{\sigma}_0^2 + \sqrt{\overline{\sigma}_0^2 (4\tilde{\eta}_0 + \overline{\sigma}_0^2)} \right\}, \tag{B.2}$$

so that $\overline{\sigma}_0^2 = \gamma_0^2 (1 - \phi_0)^2$ and $\tilde{\eta}_0 = \gamma_0^2 \phi_0$. The components of **V** are expressed as

$$\begin{aligned}
\mathbf{V}_{11}(\overline{\sigma}_0^2, \tilde{\eta}_0, \widetilde{\mathcal{K}}) &= 4\widetilde{\mathcal{K}} + 12\tilde{\eta}_0^2 + 8\tilde{\eta}_0 \overline{\sigma}_0^2, \\
\mathbf{V}_{13}(\gamma_0^2, \phi_0) &= 4\gamma_0^4 \phi_0^2 (1 - \phi_0^2), \\
\mathbf{V}_{15}(\gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) &= 2\widetilde{\mathcal{K}}T + 4\gamma_0^4 T \phi_0 \\
\mathbf{V}_{33}(\mathcal{Q}, \gamma_0^2, \phi_0) &= \frac{-2\phi_0 (3\phi_0^5 - 10\phi_0^3 - 2\phi_0^2 + \phi_0 - 6)\mathcal{Q}}{(1 - \phi_0^2)^3 T} + \frac{8\gamma_0^4 \phi_0 (1 - \phi_0)^3 (\phi_0^2 + \phi_0 + 1)}{(1 + \phi_0)^3}, \\
\mathbf{V}_{35}(\mathcal{Q}, \gamma_0^2, \phi_0) &= \frac{2\phi_0^2 (\phi_0^3 + \phi_0^2 - 2\phi_0 - 4)\mathcal{Q}}{(1 - \phi_0)^2 (1 + \phi_0)^3} - \frac{2\gamma_0^4 T \phi_0 (1 - \phi_0)^2 (\phi_0^2 + \phi_0 + 2)}{(1 + \phi_0)^3} \\
\mathbf{V}_{55}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) &= \frac{\phi_0^2 (2 - \phi_0^2)\mathcal{Q}T}{(1 - \phi_0^2)^2} + \widetilde{\mathcal{K}}T^2 + \frac{\gamma_0^4 T^2 \phi_0 (1 + \phi_0^2)(\phi_0^2 + 3\phi_0 + 4)}{(1 + \phi_0)^3}.
\end{aligned}$$

Remark 3. When the volatility is constant, the observation times are regular and there are no jumps in the price process, we have $\overline{\sigma}_0^2 = \sigma^2$, $\mathcal{Q} = \sigma^4 T$, and thus replacing ϕ_0 and γ_0^2 using (B.1) and (B.2), we have that the variance matrix for $(\widehat{\sigma}_{err}^2, \widehat{a}_{err}^2)$ is of the form

$$\begin{pmatrix} 6\sigma^4 + \mathbf{V}_{33} & -2\sigma^4 + \mathbf{V}_{35} \\ -2\sigma^4 + \mathbf{V}_{35} & \sigma^4 + \mathbf{V}_{55} \end{pmatrix},$$

equal to

$$\begin{pmatrix} 2\sigma^4 + 4\sqrt{\sigma^6 (4\eta_0 + \sigma^2)} & -(\sigma^4 + 2\sigma^2 \eta_0 + \sqrt{\sigma^2 (4\eta_0 + \sigma^2)})T \\ -(\sigma^4 + 2\sigma^2 \eta_0 + \sqrt{\sigma^2 (4\eta_0 + \sigma^2)})T & (2\eta_0 + \sigma^2)\left(2\eta_0 + \sigma^2 + \sqrt{\sigma^2 (4\eta_0 + \sigma^2)}\right)T^2 + \mathcal{K}T^2 \end{pmatrix},$$

which corresponds to the limit variance of Theorem 2 p.371 of Aït-Sahalia et al. (2005) in the abovementioned framework.

For the information part, we define for any $k \in \mathbb{N}$ the quantity

$$\tilde{\rho}_k = \frac{1}{2} \mathbb{E} \left[\frac{\partial W_0(\theta_0)}{\partial \theta} \frac{\partial W_k(\theta_0)^T}{\partial \theta} + \frac{\partial W_k(\theta_0)}{\partial \theta} \frac{\partial W_0(\theta_0)^T}{\partial \theta} \right],$$

along with the matrices

$$\begin{aligned}
U_{\theta_0} &= 2(\tilde{\rho}_0 - \tilde{\rho}_1), \\
P_{\theta_0} &= 2(1 + \phi_0)^{-1} \left\{ \tilde{\rho}_0 - (1 - \phi_0) \sum_{k=1}^{+\infty} \phi_0^{k-1} \tilde{\rho}_k \right\}.
\end{aligned}$$

Then, the asymptotic variance and covariance terms can be expressed as

$$\begin{aligned}
\mathbf{V}_{22}(\theta_0, \tilde{\eta}_0) &= 3\tilde{\eta}_0 T U_{\theta_0}^{-1}, \\
\mathbf{V}_{24}\Big(\theta_0, \gamma_0^2, \phi_0, \sum_{0 < s \le T} \Delta J_s^2\Big) &= 2\gamma_0^2 T (1 - \phi_0^2)^{-1} U_{\theta_0}^{-1} ((1 - \phi_0)\{(\phi_0^4 - 4\phi_0^3 + 5\phi_0^2 - \phi_0 + 1)\tilde{\rho}_0 \\
&\quad + (\phi_0^3 - \phi_0^2 + 3\phi_0 - 1)\tilde{\rho}_1\} \\
&\quad + 2\phi_0 (1 - \phi_0)^2 \tilde{\rho}_2 + (2 - \phi_0)(1 - \phi_0)^4 \sum_{k=2}^{+\infty} \phi_0^k \tilde{\rho}_k) P_{\theta_0}^{-1} \\
&\quad - \Bigg[\frac{2 U_{\theta_0}^{-1}}{1 - \phi_0^2} \left\{ (1 - \phi_0)^3 \tilde{\rho}_0 - (1 - \phi_0)^2 \tilde{\rho}_1 + (1 - \phi_0)^3 \sum_{k=2}^{+\infty} \phi_0^k \tilde{\rho}_k \right\} P_{\theta_0}^{-1} - U_{\theta_0}^{-1} \Bigg] \\
&\quad \times \sum_{0 < s \le T} \Delta J_s^2,
\end{aligned}$$

$$\mathbf{V}_{44}\left(\theta_0, \gamma_0^2, \phi_0, \sum_{0<s\leq T} \Delta J_s^2\right) = \gamma_0^2 T \left(P_{\theta_0}^{-1} - (1-\phi_0)^2 U_{\theta_0}^{-1}\right)$$

$$- \left[\frac{2(1-\phi_0)^2}{(1-\phi_0^2)^2} P_{\theta_0}^{-1} \left(\frac{\tilde{\rho}_0}{1-\phi_0^2} + \sum_{k=1}^{+\infty} \left\{\frac{2\phi_0^k}{1-\phi_0^2} - k\phi_0^{k-1}\right\} \tilde{\rho}_k\right) P_{\theta_0}^{-1} - U_{\theta_0}^{-1}\right]$$

$$\times \sum_{0<s\leq T} \Delta J_s^2.$$

Finally, introducing $B_{\theta_0} = \sum_{0<s\leq T} \sum_{k=1}^{N_n} \phi_0^{|k-i_n(s)|} \frac{\partial \mu_k(\theta_0)}{\partial \theta} \Delta J_s$, and $i_n(s)$ is the only index such that $t_{i-1}^n < t \leq t_i^n$ the bias terms are expressed as

$$B_{\theta_0,exp} = U_{\theta_0}^{-1} B_{\theta_0} \text{ and } B_{\theta_0,err} = P_{\theta_0}^{-1} B_{\theta_0}.$$

In particular, under $\mathcal{H}_0$, $\phi_0 = 0$ and $P_{\theta_0} = U_{\theta_0}$ so that

$$B_{\theta_0,exp} = B_{\theta_0,err} = \sum_{0<s\leq T} \frac{\partial \mu_{i_n(s)}(\theta_0)}{\partial \theta} \Delta J_s U_{\theta_0}^{-1}.$$

Appendix C. Proofs

C.1. Simplification of the problem

We give an additional assumption which is harmless (see e.g. the discussion in Clinet and Potiron (2018a, Section A.1). Note that we can apply Girsanov theorem as all the assumptions on the information also hold on the risk-neutral probability.

(H) We have $b = \tilde{b} = 0$. Moreover $\sigma, \sigma^{-1}, \tilde{\sigma}^{(1)}, (\tilde{\sigma}^{(1)})^{-1}, \tilde{\sigma}^{(2)}, (\tilde{\sigma}^{(2)})^{-1}, \alpha, \alpha^{-1}$ are bounded. Given an *a priori* number $\gamma > 0$, we also have $\sup_{0\leq i\leq N_n} U_i^n \leq n^\gamma$.

From now on, to avoid confusion as much as possible, we explicitly write the exponent n in the expressions Q_i^n, U_i^n, etc. Note also that, by virtue of Lemma 14.1.5 in Jacod and Protter (2011), recalling the definition $\pi_t^n := \sup_{i\geq 1} t_i^n - t_{i-1}^n$, and $N_n(t) = \sup\{i \in \mathbb{N} - \{0\}|t_i^n \leq t\}$ we have

$$c > 0 \implies n^{1-c}\pi_t^n \to^{\mathbb{P}} 0, \tag{C.1}$$

We sometimes refer to the continuous part of X_t defined as

$$\tilde{X}_t := X_t - J_t. \tag{C.2}$$

We define $\mathcal{U} := \sigma\left\{U_i^n|i, n \in \mathbb{N}\right\} \vee \sigma\left\{\alpha_s|0 \leq s \leq T\right\}$ the σ-field that generates the observation times and which is independent of X and Q. We will often have to use the conditional expectation $\mathbb{E}[.|\mathcal{U}]$, that we hereafter denote for convenience by $\mathbb{E}_{\mathcal{U}}$. We also define the discrete filtration $\mathcal{G}_i^n := \mathcal{F}_{t_i^n}^X \vee \mathcal{U}$, and the continuous version $\mathcal{G}_t := \mathcal{F}_t^X \vee \mathcal{U}$. Note that by independence from α, X admits the same Itô semi-martingale dynamics in the extension $\mathbf{G} = (\mathcal{G}_t)_{0\leq t\leq T}$.

Finally, all along the proofs, we recall that we write N_n in place of $N_n(T)$, we define $K_n = N_n^{1/2+\delta}$, for some $\delta > 0$ to be adjusted, and we let K be a positive constant that may vary from one line to the next.

C.2. Estimates for Ω^{-1}

We start this appendix by giving some useful estimates for the matrix $\Omega^{-1} := [\omega^{i,j}]_{i,j}$ which was defined in (3.5). Let us define $u_0 = \sqrt{\frac{\sigma^2 T}{a^2}}$. Note that $\frac{\partial u_0}{\partial \sigma^2} = \frac{u_0}{2\sigma^2}$, and $\frac{\partial u_0}{\partial a^2} = -\frac{u_0}{2a^2}$. In all this section, the expression $O(1)$ means a (possibly random) function $f : (i, j, n, \xi) \to f(i, j, n, \xi)$, where $\xi = (\sigma^2, a^2, \theta) \in \Xi$, which is bounded uniformly in all its arguments and in $\omega \in \Omega$ under the constraint $i, j \leq N_n$, and which is C^∞ on the compact Ξ, such that its partial derivatives $\frac{\partial^\alpha O(1)}{\partial^\alpha \xi}$ are also bounded. In particular, for any multi-index α we have the useful property $\frac{\partial^\alpha O(1)}{\partial^\alpha \xi} = O(1)$. Finally, we define for $n \in \mathbb{N}$ the function $g_n : k \in \{1, \ldots, 2N_n\} \to k \wedge (2N_n - k)$. Note that $g_n(k)$ is always dominated by N_n.

Lemma C.1 (*Expansions for Ω^{-1}*). *There exists $s > 0$, such that uniformly in $(i, j) \in \{1, \ldots, N_n\}^2$,*

$$\omega^{i,j} = \frac{\sqrt{N_n}}{2u_0 a^2} \left\{\left(1 - \frac{u_0^3}{24N_n^{3/2}}|i-j| + O\left(N_n^{-1}\right)\right) e^{-u_0 \frac{|i-j|}{\sqrt{N_n}}} - \left(1 - \frac{u_0^3}{24N_n^{3/2}} g_n(i+j) + O\left(N_n^{-1}\right)\right) e^{-u_0 \frac{g_n(i+j)}{\sqrt{N_n}}}\right\}$$

$$+ O\left(e^{-s\sqrt{N_n}}\right).$$

Proof. From the relation $\omega^{i,j} = \omega^{N_n-i,N_n-j}$, it is sufficient to show the result with $2 \leq i+j \leq N_n$. Under the change of variable

$$\phi = 1 - \frac{1}{2a^2}\left\{\sqrt{\sigma^2 \Delta_{N_n}(4a^2 + \sigma^2 \Delta_{N_n})} - \sigma^2 \Delta_{N_n}\right\} \text{ and } \gamma^2 = \frac{1}{2}\left\{2a^2 + \sigma^2 \Delta_{N_n} + \sqrt{\sigma^2 \Delta_{N_n}(4a^2 + \sigma^2 \Delta_{N_n})}\right\},$$

we recall the expression of $\omega^{i,j}$

$$\omega^{i,j} = \frac{\phi^{|i-j|} - \phi^{i+j} - \phi^{2N_n-i-j+2} + \phi^{2N_n-|i-j|+2}}{\gamma^2(1-\phi^2)(1-\phi^{2N_n+2})}, \tag{C.3}$$

taken from Xiu (2010), eq. (28) p 245. By a short calculation we also have the expansions

$$\phi = 1 - \frac{u_0}{\sqrt{N_n}} + \frac{u_0^2}{2N_n} - \frac{u_0^3}{8N_n^{3/2}} + O\left(N_n^{-5/2}\right), \tag{C.4}$$

and

$$\gamma^2 = a^2 + \frac{a^2 u_0}{\sqrt{N_n}} + O\left(N_n^{-1}\right). \tag{C.5}$$

Now, for the first term in the numerator, we can write

$$\begin{aligned}
\phi^{|i-j|} &= \exp\left\{|i-j|\log\left(1 - \frac{u_0}{\sqrt{N_n}} + \frac{u_0^2}{2N_n} - \frac{u_0^3}{8N_n^{3/2}} + O\left(N_n^{-5/2}\right)\right)\right\} \\
&= \exp\left\{|i-j|\left(-\frac{u_0}{\sqrt{N_n}} + \frac{u_0^2}{2N_n} - \frac{u_0^3}{8N_n^{3/2}} - \frac{1}{2}\left(-\frac{u_0}{\sqrt{N_n}} + \frac{u_0^2}{2N_n} - \frac{u_0^3}{8N_n^{3/2}}\right)^2\right.\right. \\
&\qquad\left.\left. + \frac{1}{3}\left(-\frac{u_0}{\sqrt{N_n}} + \frac{u_0^2}{2N_n} - \frac{u_0^3}{8N_n^{3/2}}\right)^3 + O\left(N_n^{-2}\right)\right)\right\} \\
&= \exp\left\{-\frac{u_0}{\sqrt{N_n}}|i-j| - \frac{u_0^3}{24N_n^{3/2}}|i-j| + \underbrace{O\left(|i-j|N_n^{-2}\right)}_{O(N_n^{-1})}\right\} \\
&= \exp\left\{-\frac{u_0}{\sqrt{N_n}}|i-j|\right\}\left(1 - \frac{u_0^3}{24N_n^{3/2}}|i-j| + O\left(N_n^{-1}\right)\right).
\end{aligned}$$

Moreover, as $i+j \leq N_n$, similar calculations lead to the estimates

$$\phi^{i+j} = \exp\left\{-\frac{u_0}{\sqrt{N_n}}(i+j)\right\}\left(1 - \frac{u_0^3}{24N_n^{3/2}}(i+j) + O\left(N_n^{-1}\right)\right),$$

$$\phi^{2N_n-i-j} = O\left(\exp\left\{-u_0\sqrt{N_n}\right\}\right),$$

and finally

$$\phi^{2N_n-|i-j|+2} = O\left(\exp\left\{-\frac{3u_0}{2}\sqrt{N_n}\right\}\right).$$

We also have the expansion

$$\gamma^2(1-\phi^2)(1-\phi^{2N_n+2}) = 2a^2\frac{u_0}{\sqrt{N_n}} + O\left(N_n^{-3/2}\right)$$

by direct calculation. Overall we thus get

$$\omega^{i,j} = \frac{\sqrt{N_n}}{2u_0 a^2}\left\{\left(1 - \frac{u_0^3}{24N_n^{3/2}}|i-j| + O\left(N_n^{-1}\right)\right)e^{-u_0\frac{|i-j|}{\sqrt{N_n}}} - \left(1 - \frac{u_0^3}{24N_n^{3/2}}(i+j) + O\left(N_n^{-1}\right)\right)e^{-u_0\frac{i+j}{\sqrt{N_n}}}\right\},$$

up to the terms related to ϕ^{2N_n-i-j} and $\phi^{2N_n-|i-j|+2}$, which we gather in $O\left(e^{-s\sqrt{N_n}}\right)$. $\quad\square$

For a matrix $A = \left[a_{i,j}\right]_{1\leq i\leq j\leq n} \in \mathbb{R}^{N_n \times N_n}$, we associate the matrix $\dot{A} = [\dot{a}_{i,j}]_{0\leq i\leq N_n, 1\leq j\leq N_n} \in \mathbb{R}^{(N_n+1)\times N_n}$ and $\ddot{A} = [\ddot{a}_{i,j}]_{0\leq i\leq N_n, 0\leq j\leq N_n} \in \mathbb{R}^{(N_n+1)\times(N_n+1)}$ whose components respectively satisfy

$$\dot{a}_{i,j} = a_{i+1,j} - a_{i,j}, \tag{C.6}$$

and

$$\ddot{a}_{i,j} = \dot{a}_{i,j+1} - \dot{a}_{i,j} = a_{i+1,j+1} - a_{i,j+1} + a_{i,j} - a_{i+1,j}, \tag{C.7}$$

with the convention $a_{i,j} = 0$ when $i = 0$ or $j = 0$. We recall the following lemma taken from Clinet and Potiron (2018c).

Lemma C.2. *Let $y, z \in \mathbb{R}^{N_n+1}$, with $y = (y_0, \ldots, y_{N_n})^T$, $z = (z_0, \ldots, z_{N_n})^T$, and define*

$$\Delta y = (\Delta y_1, \ldots, \Delta y_{N_n}) := (y_1 - y_0, \ldots, y_{N_n} - y_{N_n-1})^T \in \mathbb{R}^{N_n},$$

and Δz the same way. Then we have the by-part summation identities

$$\Delta y^T A \Delta z = -y^T \dot{A} \Delta z = y^T \ddot{A} z.$$

We define accordingly $\dot{\Omega}^{-1}$, and $\ddot{\Omega}^{-1}$. In the next lemmas we derive some estimates for such matrices.

Lemma C.3 (*Expansions for $\dot{\Omega}^{-1}$*). *We have the approximation, uniform in $i \in \{0, \ldots, N_n\}, j \in \{1, \ldots, N_n\}$*

$$
\dot{\omega}^{i,j} = -\frac{1}{2a^2} \left\{ \left(\text{sgn}(i-j) - \frac{u_0}{2\sqrt{N_n}} - \frac{u_0^3}{24N_n^{3/2}} |i-j| + O\left(N_n^{-1}\right) \right) e^{-u_0 \frac{|i-j|}{\sqrt{N_n}}} \right.
$$
$$
\left. - \left(1 - \frac{u_0}{2\sqrt{N_n}} - \frac{u_0^3}{24N_n^{3/2}} g_n(i+j) + O\left(N_n^{-1}\right) \right) e^{-u_0 \frac{g_n(i+j)}{\sqrt{N_n}}} \right\} + O\left(e^{-s\sqrt{N_n}}\right),
$$

where $\text{sgn}(x) = \mathbb{1}_{\{x \geq 0\}} - \mathbb{1}_{\{x < 0\}}$.

Proof. Once again we assume without loss of generality that $i + j \leq N_n$. From Lemma C.1 and the definition of $\omega^{i,j}$, some calculation gives, up to the term $O\left(e^{-s\sqrt{N_n}}\right)$,

$$
\dot{\omega}^{i,j} = \frac{\sqrt{N_n}}{2u_0 a^2} \left\{ \left(1 - \frac{u_0^3}{24N_n^{3/2}} |i+1-j| + O\left(N_n^{-1}\right) \right) \left(e^{-u_0 \frac{|i+1-j|}{\sqrt{N_n}}} - e^{-u_0 \frac{|i-j|}{\sqrt{N_n}}} \right) \right.
$$
$$
\left. - \left(1 - \frac{u_0^3}{24N_n^{3/2}} (i+j+1) + O\left(N_n^{-1}\right) \right) \left(e^{-u_0 \frac{i+j+1}{\sqrt{N_n}}} - e^{-u_0 \frac{i+j}{\sqrt{N_n}}} \right) \right\}
$$
$$
= \frac{\sqrt{N_n}}{2u_0 a^2} \left\{ \left(1 - \frac{u_0^3}{24N_n^{3/2}} |i-j| + O\left(N_n^{-1}\right) \right) e^{-u_0 \frac{|i-j|}{\sqrt{N_n}}} \left(e^{-u_0 \frac{\text{sgn}(i-j)}{\sqrt{N_n}}} - 1 \right) \right.
$$
$$
\left. - \left(1 - \frac{u_0^3}{24N_n^{3/2}} (i+j) + O\left(N_n^{-1}\right) \right) e^{-u_0 \frac{i+j}{\sqrt{N_n}}} \left(e^{-\frac{u_0}{\sqrt{n}}} - 1 \right) \right\}
$$
$$
= \frac{\sqrt{N_n}}{2u_0 a^2} \left\{ \left(1 - \frac{u_0^3}{24N_n^{3/2}} |i-j| + O\left(N_n^{-1}\right) \right) e^{-u_0 \frac{|i-j|}{\sqrt{N_n}}} \left(e^{-u_0 \frac{\text{sgn}(i-j)}{\sqrt{N_n}}} - 1 \right) \right.
$$
$$
\left. - \left(1 - \frac{u_0^3}{24N_n^{3/2}} (i+j) + O\left(N_n^{-1}\right) \right) e^{-u_0 \frac{i+j}{\sqrt{N_n}}} \left(e^{-\frac{u_0}{\sqrt{n}}} - 1 \right) \right\}
$$
$$
= \frac{\sqrt{N_n}}{2u_0 a^2} \left\{ \left(1 - \frac{u_0^3}{24N_n^{3/2}} |i-j| + O\left(N_n^{-1}\right) \right) e^{-u_0 \frac{|i-j|}{\sqrt{N_n}}} \left(-u_0 \frac{\text{sgn}(i-j)}{\sqrt{N_n}} + \frac{u_0^2}{2N_n} + O\left(N_n^{-3/2}\right) \right) \right.
$$
$$
\left. - \left(1 - \frac{u_0^3}{24N_n^{3/2}} (i+j) + O\left(N_n^{-1}\right) \right) e^{-u_0 \frac{i+j}{\sqrt{N_n}}} \left(-\frac{u_0}{\sqrt{N_n}} + \frac{u_0^2}{2N_n} + O\left(N_n^{-3/2}\right) \right) \right\},
$$

and expanding the terms in parenthesis we get the result. $\quad\square$

From the previous lemma we deduce by similar computations an expansion for $\ddot{\Omega}^{-1}$.

Lemma C.4 (*Expansions for $\ddot{\Omega}^{-1}$*). *If $i \neq j$, we have the approximation, uniform in $(i, j) \in \{0, \ldots, N_n\}^2$*

$$
\ddot{\omega}^{i,j} = -\frac{u_0}{2a^2 \sqrt{N_n}} \left(1 + O\left(N_n^{-1/2}\right) \right) \left\{ e^{-u_0 \frac{|i-j|}{\sqrt{N_n}}} + e^{-u_0 \frac{g_n(i+j)}{\sqrt{N_n}}} \right\} + O\left(e^{-s\sqrt{N_n}}\right). \tag{C.8}
$$

Moreover, uniformly in $i \in \{0, \ldots, N_n\}$,

$$
\ddot{\omega}^{i,i} = \frac{1}{a^2} \left(1 - \frac{u_0}{2\sqrt{N_n}} + O\left(N_n^{-1}\right) \right) + \frac{u_0}{2a^2 \sqrt{N_n}} \left(1 + O\left(N_n^{-1/2}\right) \right) e^{-u_0 \frac{g_n(2i)}{\sqrt{N_n}}} + O\left(e^{-s\sqrt{N_n}}\right). \tag{C.9}
$$

C.3. Estimates for the efficient price X

Hereafter, we adopt the same notation convention as in Clinet and Potiron (2018a), Section A.3. For a process V, and $t \in [0, T]$ we write $\Delta V_t = V_t - V_{t-}$. We also write $\Delta V_i^n := V_{t_i^n} - V_{t_{i-1}^n}$. Finally, for interpolation purpose we sometimes

write the continuous version $\Delta V_{i,t}^n := V_{t_i^n \wedge t} - V_{t_{i-1}^n \wedge t}$. Let us define

$$\zeta_{i,t}^n := (\Delta \tilde{X}_{i,t}^n)^2 - \sigma_{t_{i-1}^n}^2 (t_i^n \wedge t - t_{i-1}^n \wedge t), \text{ and } \overline{\zeta}_{i,t}^n := \mathbb{E}\left[\zeta_{i,t}^n | \mathcal{G}_{i-1}^n\right]. \tag{C.10}$$

We recall the following standard estimates.

Lemma C.5. *We have, for some constant $K > 0$ independent of i,*

$$\mathbb{E}\left[\sup_{t \in]t_{i-1}^n, t_i^n]} |\Delta \tilde{X}_{i,t}^n|^p \,\middle|\, \mathcal{G}_{i-1}^n\right] \leq Kn^{-p/2}(U_i^n)^{p/2}, \tag{C.11}$$

$$\left|\tilde{\zeta}_{i,t}^n\right| \leq Kn^{-3/2}(U_i^n)^{3/2}, \tag{C.12}$$

$$\mathbb{E}\left[\left.(\zeta_{t,i}^n)^p\,\right|\,\mathcal{G}_{i-1}^n\right] \leq Kn^{-p}(U_i^n)^p, \tag{C.13}$$

$$\mathbb{E}\left[\left.\left|\int_{t_{i-1}^n \wedge t}^{t_i^n \wedge t} \sigma_s^2 ds - \sigma_{t_{i-1}^n}^2 (t_i^n \wedge t - t_{i-1}^n \wedge t)\right|^p\,\right|\,\mathcal{G}_{i-1}^n\right] \leq Kn^{-3p/2}(U_i^n)^{3p/2}. \tag{C.14}$$

C.4. Estimates for the information part

In this section we derive some asymptotic results for the information part. We define for $\xi = (\sigma^2, a^2, \theta) \in \Xi$

$$\mathbf{G}_n(\xi) = (\mu(\theta_0) - \mu(\theta))^T \Omega^{-1}(\mu(\theta_0) - \mu(\theta)), \tag{C.15}$$

along with the asymptotic fields

$$\mathbf{G}_{\infty,1}(\xi) = -\frac{u_0}{2a}\left\{\rho_0(\theta) + 2\sum_{k=1}^{+\infty}\rho_k(\theta)\right\}, \tag{C.16}$$

and

$$\mathbf{G}_{\infty,2}(\xi) = \frac{\rho_0(\theta)}{a^2}. \tag{C.17}$$

By Lemma C.4, we have the following matrix decomposition for $\ddot{\Omega}^{-1}$. Let

$$\mathcal{E}^- = \left[e^{-u_0 \frac{|i-j|}{\sqrt{N_n}}}\right]_{0 \leq i,j \leq N_n}, \text{ and } \mathcal{E}^+ = \left[e^{-u_0 \frac{g_n(i+j)}{\sqrt{N_n}}}\right]_{0 \leq i,j \leq N_n}. \tag{C.18}$$

Then we have

$$\ddot{\Omega}^{-1} = \frac{1}{a^2}\mathbf{I}_{N_n} - \frac{u_0}{2a^2\sqrt{N_n}}\left(1 + O\left(N_n^{-1/2}\right)\right)\left\{\mathcal{E}^- - \mathcal{E}^+\right\} + O\left(e^{-s\sqrt{N_n}}\right)\mathbf{J}_{N_n}, \tag{C.19}$$

where $\mathbf{I}_{N_n}, \mathbf{J}_{N_n} \in \mathbb{R}^{N_n \times N_n}$ are respectively the identity matrix and the matrix whose components are all equal to 1.

Lemma C.6. *Let $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2)$ be a multi-index such that $|\boldsymbol{\alpha}| \leq m$. if $\alpha_0 > 0$, then we have*

$$\sup_{\xi \in \Xi} \mathbb{E}_{\mathcal{U}}\left[\left(\frac{1}{\sqrt{N_n}}\frac{\partial^{\boldsymbol{\alpha}}\mathbf{G}_n(\xi)}{\partial \xi^{\boldsymbol{\alpha}}} - \frac{\partial^{\boldsymbol{\alpha}}\mathbf{G}_{\infty,1}(\xi)}{\partial \xi^{\boldsymbol{\alpha}}}\right)^2\right] \to^{\mathbb{P}} 0. \tag{C.20}$$

Moreover, if $\alpha_0 = 0$, then we have

$$\sup_{\xi \in \Xi} \mathbb{E}_{\mathcal{U}}\left[\left(\frac{1}{N_n}\frac{\partial^{\boldsymbol{\alpha}}\mathbf{G}_n(\xi)}{\partial \xi^{\boldsymbol{\alpha}}} - \frac{\partial^{\boldsymbol{\alpha}}\mathbf{G}_{\infty,2}(\xi)}{\partial \xi^{\boldsymbol{\alpha}}}\right)^2\right] \to^{\mathbb{P}} 0. \tag{C.21}$$

Proof. First note that by Lemma C.2, $\mathbf{G}_n(\xi)$ has the representation

$$\mathbf{G}_n(\xi) = W(\theta)^T \ddot{\Omega}^{-1} W(\theta), \tag{C.22}$$

and thus by (C.19) $\mathbf{G}_n(\xi)$ admits the decomposition

$$\mathbf{G}_n(\xi) = \frac{1}{a^2}\text{Tr}\left(W(\theta)W(\theta)^T\right) - \frac{u_0}{2a^2\sqrt{N_n}}\left(1 + O\left(\frac{1}{\sqrt{N_n}}\right)\right)W(\theta)^T\left\{\mathcal{E}^- - \mathcal{E}^+\right\}W(\theta)$$
$$+ O\left(e^{-s\sqrt{N_n}}\right)W(\theta)^T\mathbf{J}_{N_n}W(\theta).$$

Consider now some multi-index $\boldsymbol{\alpha}$ such that $|\boldsymbol{\alpha}| \leq m$, and first assume that $\alpha_0 > 0$. Let us denote

$$J_{\boldsymbol{\alpha}}^+ = \frac{\partial^{\boldsymbol{\alpha}}}{\partial\xi^{\boldsymbol{\alpha}}}\left\{\frac{u_0}{2a^2}\mathcal{E}^+\right\}, \tag{C.23}$$

and a similar definition for $J_{\boldsymbol{\alpha}}^-$. We show (C.20). First note that in that case $\frac{\partial^{\boldsymbol{\alpha}}}{\partial\xi^{\boldsymbol{\alpha}}}\left(\frac{1}{a^2}\text{Tr}\left(W(\theta)W(\theta)^T\right)\right) = 0$ since $\alpha_0 > 0$ and $\frac{1}{a^2}\text{Tr}\left(W(\theta)W(\theta)^T\right)$ does not depend on σ^2. Moreover, it is immediate to see that the term in $O\left(e^{-s\sqrt{N_n}}\right)\frac{\partial^{\boldsymbol{\alpha}}}{\partial\xi^{\boldsymbol{\alpha}}}\left\{W(\theta)^T\mathbf{J}_{N_n}W(\theta)\right\}$ is negligible given the factor $e^{-s\sqrt{N_n}}$. Let us show now that we have

$$\sup_{\xi\in\Xi}\mathbb{E}_{\mathcal{U}}\left[\left(\frac{\partial^{\boldsymbol{\alpha}}}{\partial\xi^{\boldsymbol{\alpha}}}\frac{u_0}{2a^2N_n}\left(1 + O\left(\frac{1}{\sqrt{N_n}}\right)\right)W(\theta)^T\mathcal{E}^+W(\theta)\right)^2\right] \to^{\mathbb{P}} 0, \tag{C.24}$$

which after some straightforward calculation is equivalent to showing that

$$\sup_{\xi\in\Xi}\frac{1}{N_n^2}\mathbb{E}_{\mathcal{U}}\left[\left(\frac{\partial^{\alpha_2}}{\partial\theta^{\alpha_2}}\left\{W(\theta)^TJ_{\boldsymbol{\alpha}}^+W(\theta)\right\}\right)^2\right] \to^{\mathbb{P}} 0. \tag{C.25}$$

By the classical variance–bias decomposition, (C.25) will be proved if we can show uniformly in $\xi \in \Xi$ that we have

$$\frac{1}{N_n}\frac{\partial^{\alpha_2}}{\partial\theta^{\alpha_2}}\mathbb{E}_{\mathcal{U}}\left[W(\theta)^TJ_{\boldsymbol{\alpha}}^+W(\theta)\right] \to^{\mathbb{P}} 0 \tag{C.26}$$

on the one hand, and

$$\frac{1}{N_n^2}\frac{\partial^{\alpha_2}}{\partial\theta^{\alpha_2}}\text{Var}_{\mathcal{U}}\left[W(\theta)^TJ_{\boldsymbol{\alpha}}^+W(\theta)\right] \to^{\mathbb{P}} 0 \tag{C.27}$$

on the other hand. We start by (C.26). Recall that $K_n = N_n^{1/2+\delta}$ for some $0 < \delta < 1/2$. From the definition of $\mathcal{E}^+$ in (C.18), it is straightforward to see that $J_{\boldsymbol{\alpha},i,j}^+ = O\left(e^{-\frac{u_0}{2}N_n^\delta}\right)$ as soon as $i+j \geq K_n$, and $J_{\boldsymbol{\alpha},i,j}^+ = O(1)$ otherwise. Therefore, we have

$$\frac{1}{N_n}\frac{\partial^{\alpha_2}}{\partial\theta^{\alpha_2}}\mathbb{E}_{\mathcal{U}}\left[W(\theta)^TJ_{\boldsymbol{\alpha}}^+W(\theta)\right] = \frac{1}{N_n}\sum_{i,j=0}^{N_n}J_{\boldsymbol{\alpha},i,j}^+\frac{\partial^{\alpha_2}\rho_{|i-j|}(\theta)}{\partial\theta^{\alpha_2}}. \tag{C.28}$$

From the symmetry $J_{\boldsymbol{\alpha},N_n-i,N_n-j}^+ = J_{\boldsymbol{\alpha},i,j}^+$ we split the sum as

$$\frac{1}{N_n}\sum_{i,j=0}^{N_n}J_{\boldsymbol{\alpha},i,j}^+\frac{\partial^{\alpha_2}\rho_{|i-j|}(\theta)}{\partial\theta^{\alpha_2}} = \frac{2}{N_n}\sum_{0\leq i+j<N_n}J_{\boldsymbol{\alpha},i,j}^+\frac{\partial^{\alpha_2}\rho_{|i-j|}(\theta)}{\partial\theta^{\alpha_2}} + \underbrace{\frac{1}{N_n}\sum_{i+j=N_n}J_{\boldsymbol{\alpha},i,j}^+\frac{\partial^{\alpha_2}\rho_{|i-j|}(\theta)}{\partial\theta^{\alpha_2}}}_{O\left(e^{-\frac{u_0}{2}N_n^\delta}\right)},$$

$$= \frac{2}{N_n}\sum_{0\leq i+j<K_n}J_{\boldsymbol{\alpha},i,j}^+\frac{\partial^{\alpha_2}\rho_{|i-j|}(\theta)}{\partial\theta^{\alpha_2}} + \frac{2}{N_n}\sum_{K_n\leq i+j<N_n}J_{\boldsymbol{\alpha},i,j}^+\frac{\partial^{\alpha_2}\rho_{|i-j|}(\theta)}{\partial\theta^{\alpha_2}}$$
$$+ O\left(e^{-\frac{u_0}{2}N_n^\delta}\right),$$

and first we have

$$\frac{2}{N_n}\left|\sum_{K_n\leq i+j<N_n}J_{\boldsymbol{\alpha},i,j}^+\frac{\partial^{\alpha_2}\rho_{|i-j|}(\theta)}{\partial\theta^{\alpha_2}}\right| = O\left(e^{-\frac{u_0}{2}N_n^\delta}\right) \times \frac{1}{N_n}\left|\sum_{K_n\leq i+j<N_n}\frac{\partial^{\alpha_2}\rho_{|i-j|}(\theta)}{\partial\theta^{\alpha_2}}\right|$$

$$\leq O\left(e^{-\frac{u_0}{2}N_n^\delta}\right) \times \underbrace{\sum_{0\leq k<N_n}\left(1 - \frac{k}{N_n}\right)\left|\frac{\partial^{\alpha_2}\rho_k(\theta)}{\partial\theta^{\alpha_2}}\right|}_{O(1)},$$

$$= O\left(e^{-\frac{u_0}{2}N_n^\delta}\right),$$

where the estimates $\sum_{0 \leq k < N_n} \left(1 - \frac{k}{N_n}\right) \left| \frac{\partial^{\alpha_2} \rho_k(\theta)}{\partial \theta^{\alpha_2}} \right| = O(1)$ uniformly in $\theta \in \Theta$ is a consequence of Assumption (2.11). Now we also have

$$\frac{2}{N_n} \left| \sum_{0 \leq i+j < K_n} J_{\boldsymbol{\alpha},i,j}^+ \frac{\partial^{\alpha_2} \rho_{|i-j|}(\theta)}{\partial \theta^{\alpha_2}} \right| \leq O(1) \times \frac{K_n}{N_n} \sum_{0 \leq k < K_n} \left(1 - \frac{k}{K_n}\right) \left| \frac{\partial^{\alpha_2} \rho_k(\theta)}{\partial \theta^{\alpha_2}} \right|$$

$$= O\left(N_n^{\delta - 1/2}\right),$$

by the same argument. Thus we have proved (C.26). Now, using a similar formula as for McCullagh (1987), Section 3.3 p. 61, (C.27) can be expressed as the sum

$$\frac{1}{N_n^2} \mathrm{Var}_{\mathcal{U}} \left[\frac{\partial^{\alpha_2}}{\partial \theta^{\alpha_2}} \left\{ W(\theta)^T J_{\boldsymbol{\alpha}}^+ W(\theta) \right\} \right] = V_1^+ + V_2^+, \tag{C.29}$$

where by the Leibniz rule,

$$V_1^+ = \frac{1}{N_n^2} \sum_{r_1, r_2 = 0}^{\alpha_2} \binom{\alpha_2}{r_1} \binom{\alpha_2}{r_2} \sum_{i,j,k,l=0}^{N_n} J_{\boldsymbol{\alpha},i,j}^+ J_{\boldsymbol{\alpha},k,l}^+ \kappa_{j-i,k-i,l-i}^{\boldsymbol{\beta(r)}}(\theta), \tag{C.30}$$

with $\boldsymbol{\beta(r)} = (r_1, \alpha_2 - r_1, r_2, \alpha_2 - r_2), \boldsymbol{r} = (r_1, r_2)$, where r_1 and r_2 are d dimensional multi-indices such that $r_1, r_2 \leq \alpha_2$, and

$$V_2^+ = \frac{1}{N_n^2} \sum_{i,j,k,l=0}^{N_n} J_{\boldsymbol{\alpha},i,j}^+ J_{\boldsymbol{\alpha},k,l}^+ \left\{ \rho_{|k-i|}^{(r_1,r_2)}(\theta) \rho_{|l-i|}^{(\alpha_2-r_1,\alpha_2-r_2)}(\theta) + \rho_{|l-i|}^{(r_1,\alpha_2-r_2)}(\theta) \rho_{|k-j|}^{(\alpha_1-r_1,r_2)}(\theta) \right\}. \tag{C.31}$$

First, we have

$$\frac{4}{N_n^2} \left| \sum_{0 \leq i,j,k,l \leq N_n} J_{\boldsymbol{\alpha},i,j}^+ J_{\boldsymbol{\alpha},k,l}^+ \kappa_{j-i,k-i,l-i}^{\boldsymbol{\beta(r)}}(\theta) \right| \leq O(1) \times \frac{1}{N_n^2} \sum_{0 \leq i,j,k,l \leq N_n} \left| \kappa_{j-i,k-i,l-i}^{\boldsymbol{\beta(r)}}(\theta) \right|.$$

Now, since we can swap the elements of $\boldsymbol{\beta(r)}$ without loss of generality we can assume that $\kappa_{j-i,k-i,l-i}^{\boldsymbol{\beta(r)}}(\theta)$ is symmetric in i, j, k, l so that we have up to a multiplicative constant

$$\left| V_1^+ \right| \leq O(1) \times \frac{1}{N_n^2} \sum_{r_1, r_2 = 0}^{\alpha_2} \sum_{0 \leq i \leq j,k,l < K_n} \left| \kappa_{j-i,k-i,l-i}^{\boldsymbol{\beta(r)}}(\theta) \right|$$

$$\leq O(1) \times \frac{1}{N_n} \sum_{0 \leq p,q,r < N_n} \left(1 - \frac{p \wedge q \wedge r}{N_n}\right) \left| \kappa_{p,q,r}^{\boldsymbol{\beta(r)}}(\theta) \right|$$

$$= O\left(N_n^{-1}\right),$$

by Assumption (2.12). On the other hand, following a similar path as for the bias case, we can reduce up to negligible terms the elementary terms of $\left| V_2^+ \right|$ to

$$\frac{4}{N_n^2} \left| \sum_{0 \leq i+j < K_n, 0 \leq k+l < K_n} J_{\boldsymbol{\alpha},i,j}^+ J_{\boldsymbol{\alpha},k,l}^+ \left\{ \rho_{|k-i|}^{(r_1,r_2)}(\theta) \rho_{|l-j|}^{(\alpha_2-r_1,\alpha_2-r_2)}(\theta) \right\} \right|$$

$$\leq O(1) \times \sum_{0 \leq i+j < K_n, 0 \leq k+l < K_n} \left| \left\{ \rho_{|k-i|}^{(r_1,r_2)}(\theta) \rho_{|l-j|}^{(\alpha_2-r_1,\alpha_2-r_2)}(\theta) \right\} \right|$$

$$\leq O(1) \times \frac{K_n^2}{n^2} \sum_{p,q=0}^{K_n-1} \left(1 - \frac{p}{K_n}\right) \left(1 - \frac{q}{K_n}\right) \left| \left\{ \rho_p^{(r_1,r_2)}(\theta) \rho_q^{(\alpha_2-r_1,\alpha_2-r_2)}(\theta) \right\} \right|$$

$$= O\left(N_n^{2\delta-1}\right),$$

where the last estimate is obtained by application of Assumption (2.11). Similar reasoning also shows the negligibility of the terms in $\rho_{|l-i|}^{(r_1,\alpha_2-r_2)}(\theta) \rho_{|k-j|}^{(\alpha_1-r_1,r_2)}(\theta)$. Thus we have proved (C.27), so that (C.24) holds true. To complete the proof of (C.20), it remains to show

$$\sup_{\xi \in \Xi} \mathbb{E}_{\mathcal{U}} \left[\left(\frac{\partial^{\boldsymbol{\alpha}}}{\partial \xi^{\boldsymbol{\alpha}}} \left\{ \frac{u_0}{2a^2 N_n} \left(1 + O\left(\frac{1}{\sqrt{N_n}}\right)\right) W(\theta)^T \mathcal{E}^- W(\theta) - \mathbf{G}_{\infty,1}(\xi) \right\} \right)^2 \right] \to^{\mathbb{P}} 0, \tag{C.32}$$

which can be expressed as

$$\sup_{\xi \in \Xi} \frac{1}{N_n^2} \frac{\partial^{\alpha_2}}{\partial \theta^{\alpha_2}} \mathbb{E}_{\mathcal{U}} \left[\left(W(\theta)^T J_{\boldsymbol{\alpha}}^- W(\theta) - \frac{\partial^{\boldsymbol{\alpha}} \mathbf{G}_{\infty,1}(\xi)}{\partial \xi^{\boldsymbol{\alpha}}} \right)^2 \right] \to^{\mathbb{P}} 0. \tag{C.33}$$

We adopt the same bias–variance approach as before, and first compute using the symmetry $J^-_{\boldsymbol{\alpha},i,j} = J^-_{\boldsymbol{\alpha},j,i}$,

$$\frac{1}{N_n}\frac{\partial^{\alpha_2}}{\partial\theta^{\alpha_2}}\mathbb{E}_{\mathcal{U}}\left[W(\theta)^T J^-_{\boldsymbol{\alpha}} W(\theta)\right] = \frac{2}{N_n}\sum_{0\le i<j\le N_n} J^-_{\boldsymbol{\alpha},i,j}\frac{\partial^{\alpha_2}\rho_{|i-j|}(\theta)}{\partial\theta^{\alpha_2}} + \frac{1}{N_n}\sum_{i=0}^{N_n} J^-_{\boldsymbol{\alpha},i,i}\frac{\partial^{\alpha_2}\rho_0(\theta)}{\partial\theta^{\alpha_2}}. \tag{C.34}$$

Now, we have immediately $J^-_{\boldsymbol{\alpha},i,i} = \frac{\partial^{\boldsymbol{\alpha}}}{\partial\xi^{\boldsymbol{\alpha}}}\left(\frac{u_0}{2a^2}\right)$, so that

$$\frac{1}{N_n}\sum_{i=0}^{N_n} J^-_{\boldsymbol{\alpha},i,i}\frac{\partial^{\alpha_2}\rho_0(\theta)}{\partial\theta^{\alpha_2}} - \frac{\partial^{\boldsymbol{\alpha}}}{\partial\xi^{\boldsymbol{\alpha}}}\left(\frac{u_0}{2a^2}\rho_0(\theta)\right) \to^{\mathbb{P}} 0 \tag{C.35}$$

uniformly in $\xi\in\Xi$. Thus, by (C.34) and (C.35) we have

$$\mathbb{E}_{\mathcal{U}}\left[\frac{1}{N_n}\frac{\partial^{\alpha_2}}{\partial\theta^{\alpha_2}}\left\{W(\theta)^T J^-_{\boldsymbol{\alpha}} W(\theta)\right\} - \frac{\partial^{\boldsymbol{\alpha}}\mathbf{G}_{\infty,1}(\xi)}{\partial\xi^{\boldsymbol{\alpha}}}\right] = \frac{2}{N_n}\sum_{0\le i<j\le N_n}\left(J^-_{\boldsymbol{\alpha},i,j} - \frac{\partial^{\boldsymbol{\alpha}}}{\partial\xi^{\boldsymbol{\alpha}}}\left(\frac{u_0}{2a^2}\right)\right)\frac{\partial^{\alpha_2}\rho_{|i-j|}(\theta)}{\partial\theta^{\alpha_2}}$$
$$+ o_{\mathbb{P}}(1),$$

and noticing that $\sup_{\xi\in\Xi}\left\{J^-_{\boldsymbol{\alpha},i,j} - \frac{\partial^{\boldsymbol{\alpha}}}{\partial\xi^{\boldsymbol{\alpha}}}\left(\frac{u_0}{2a^2}\right)\right\} = o_{\mathbb{P}}(1)$, we easily conclude as before by the dominated convergence theorem along with Assumption (2.11) that

$$\sup_{\xi\in\Xi}\frac{1}{N_n}\frac{\partial^{\alpha_2}}{\partial\theta^{\alpha_2}}\mathbb{E}_{\mathcal{U}}\left[W(\theta)^T J^-_{\boldsymbol{\alpha}} W(\theta) - \frac{\partial^{\boldsymbol{\alpha}}\mathbf{G}_{\infty,1}(\xi)}{\partial\xi^{\boldsymbol{\alpha}}}\right] \to^{\mathbb{P}} 0. \tag{C.36}$$

Moreover, the variance term is treated as previously.

Finally, for (C.21), similar computations yield the result. Note that when $\alpha_0 = 0$, the diagonal terms of $\frac{\partial^{\boldsymbol{\alpha}}\ddot{\Omega}^{-1}}{\partial\xi^{\boldsymbol{\alpha}}}$ predominate, hence there is absence of higher order correlations $\rho_k(\theta)$, $k\ge 1$, in the limit. $\quad\square$

To deal with the cross terms, we define in the same fashion as before

$$\mathbf{K}_n(\xi) = (\mu(\theta_0) - \mu(\theta))^T \Omega^{-1}\widetilde{Y}(\theta_0), \tag{C.37}$$

Lemma C.7. *Let $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2)$ be a multi-index such that $|\boldsymbol{\alpha}| \le m$. Then, if $\alpha_0 = 0$, we have*

$$\sup_{\xi\in\Xi}\mathbb{E}_{\mathcal{U}}\left[\left(\frac{1}{\sqrt{N_n}}\frac{\partial^{\boldsymbol{\alpha}}\mathbf{K}_n(\xi)}{\partial\xi^{\boldsymbol{\alpha}}}\right)^2\right] \to^{\mathbb{P}} 0. \tag{C.38}$$

If $\alpha_0 > 0$, we have

$$\sup_{\xi\in\Xi}\mathbb{E}_{\mathcal{U}}\left[\left(\frac{1}{N_n}\frac{\partial^{\boldsymbol{\alpha}}\mathbf{K}_n(\xi)}{\partial\xi^{\boldsymbol{\alpha}}}\right)^2\right] \to^{\mathbb{P}} 0. \tag{C.39}$$

Proof. Let us first show (C.38). By Lemma C.2, we have the representation

$$\mathbf{K}_n(\xi) = W(\theta)^T \dot{\Omega}^{-1}\Delta X + W(\theta)^T \ddot{\Omega}^{-1}\epsilon. \tag{C.40}$$

Accordingly, we start by showing that

$$\sup_{\xi\in\Xi}\mathbb{E}_{\mathcal{U}}\left[\left(\frac{1}{\sqrt{N_n}}\frac{\partial^{\boldsymbol{\alpha}}}{\partial\xi^{\boldsymbol{\alpha}}}\left\{W(\theta)^T\dot{\Omega}^{-1}\Delta X\right\}\right)^2\right] \to^{\mathbb{P}} 0. \tag{C.41}$$

In the case where X is continuous, that is $J = 0$, we have

$$\mathbb{E}_{\mathcal{U}}\left[\left(\frac{1}{\sqrt{N_n}}\frac{\partial^{\boldsymbol{\alpha}}}{\partial\xi^{\boldsymbol{\alpha}}}\left\{W(\theta)^T\dot{\Omega}^{-1}\Delta X\right\}\right)^2\right] = \frac{1}{N_n}\mathbb{E}_{\mathcal{U}}\left[\mathbb{E}\left[\left(\frac{\partial^{\boldsymbol{\alpha}}}{\partial\xi^{\boldsymbol{\alpha}}}\left\{W(\theta)^T\dot{\Omega}^{-1}\Delta X\right\}\right)^2\middle|\mathcal{U}\vee\sigma(X)\right]\right]$$

$$= \frac{1}{N_n}\mathbb{E}_{\mathcal{U}}\left[\sum_{0\le i,k\le N_n, 1\le j,l\le N_n}\frac{\partial^{\alpha_2}\rho_{|i-k|}(\theta)}{\partial\theta^{\alpha_2}}\frac{\partial^{\boldsymbol{\alpha}}\dot{\omega}^{i,j}}{\partial\xi^{\boldsymbol{\alpha}}}\frac{\partial^{\boldsymbol{\alpha}}\dot{\omega}^{k,l}}{\partial\xi^{\boldsymbol{\alpha}}}\Delta X^n_j\Delta X^n_l\right]$$

$$= \frac{1}{N_n}\sum_{0\le i,k\le N_n, 1\le j\le N_n}\frac{\partial^{\alpha_2}\rho_{|i-k|}(\theta)}{\partial\theta^{\alpha_2}}\frac{\partial^{\boldsymbol{\alpha}}\dot{\omega}^{i,j}}{\partial\xi^{\boldsymbol{\alpha}}}\frac{\partial^{\boldsymbol{\alpha}}\dot{\omega}^{k,j}}{\partial\xi^{\boldsymbol{\alpha}}}\mathbb{E}_{\mathcal{U}}\int_{t^n_{j-1}}^{t^n_j}\sigma^2_s ds.$$

Now, using $\frac{\partial^{\alpha}\dot{\omega}^{k,j}}{\partial\xi^{\alpha}} = O(1)$, and that $\mathbb{E}_{\mathcal{U}}\int_{t_{j-1}^{n}}^{t_{j}^{n}}\sigma_{s}^{2}ds = O_{\mathbb{P}}\left(n^{-1+\gamma}\right)$ by Assumption (**H**) where $\gamma > 0$ can be taken arbitrary small, we get

$$\mathbb{E}_{\mathcal{U}}\left[\left(\frac{1}{\sqrt{N_n}}\frac{\partial^{\alpha}}{\partial\xi^{\alpha}}\left\{W(\theta)^{T}\dot{\Omega}^{-1}\Delta X\right\}\right)^{2}\right] = O_{\mathbb{P}}\left(N_n^{-1}n^{-1+\gamma}\right)\times\sum_{0\leq i,k\leq N_n,1\leq j\leq N_n}\left|\frac{\partial^{\alpha}\dot{\omega}^{i,j}}{\partial\xi^{\alpha}}\right|\left|\frac{\partial^{\alpha_2}\rho_{|i-k|}(\theta)}{\partial\theta^{\alpha_2}}\right|$$

and summing first over k and using Assumption (2.11), we get

$$\mathbb{E}_{\mathcal{U}}\left[\left(\frac{1}{\sqrt{N_n}}\frac{\partial^{\alpha}}{\partial\xi^{\alpha}}\left\{W(\theta)^{T}\dot{\Omega}^{-1}\Delta X\right\}\right)^{2}\right] = O_{\mathbb{P}}\left(N_n^{-1}n^{-1+\gamma}\right)\times\sum_{0\leq i\leq N_n,1\leq j\leq N_n}\left\{e^{-u_0\frac{|i-j|}{\sqrt{N_n}}}+e^{-u_0\frac{i+j}{\sqrt{N_n}}}\right\}$$
$$= O_{\mathbb{P}}\left(N_n^{1/2}n^{-1+\gamma}\right)\to^{\mathbb{P}}0,$$

uniformly in $\xi\in\Xi$ by direct calculation for the last estimate. Now, when $J\neq 0$ for n sufficiently large, using the finite activity property of the jump process it is easy to see that an additional term appears in the quadratic expression,

$$\frac{1}{N_n}\sum_{0\leq i,k\leq N_n}\frac{\partial^{\alpha_2}\rho_{|i-k|}(\theta)}{\partial\theta^{\alpha_2}}\sum_{j=1}^{N_J}\frac{\partial^{\alpha}\dot{\omega}^{i,i_j}}{\partial\xi^{\alpha}}\frac{\partial^{\alpha}\dot{\omega}^{k,i_j}}{\partial\xi^{\alpha}}\mathbb{E}_{\mathcal{U}}\Delta J_{\tau_j}^{2}, \tag{C.42}$$

where N_J is the finite number of jumps of J on $[0, T]$, $(\tau_j)_{1\leq j\leq N_J}$ the related jump times, and i_j is the only index such that $t_{i_j-1}^{n} < \tau_j \leq t_{i_j}^{n}$. Given the estimate of $\dot{\omega}^{i,j}$ of the previous section and the definition of i_j, we immediately see that the coefficients $\frac{\partial^{\alpha}\dot{\omega}^{i,i_j}}{\partial\xi^{\alpha}} = O\left(e^{-v\sqrt{N_n}}\right)$ for some $v > 0$, and thus (C.42) is negligible so that we have (C.41). Now we show that we have

$$\sup_{\xi\in\Xi}\mathbb{E}_{\mathcal{U}}\left[\left(\frac{1}{\sqrt{N_n}}\frac{\partial^{\alpha}}{\partial\xi^{\alpha}}\left\{W(\theta)^{T}\ddot{\Omega}^{-1}\epsilon\right\}\right)^{2}\right]\to^{\mathbb{P}}0. \tag{C.43}$$

By independence, we immediately have

$$\mathbb{E}_{\mathcal{U}}\left[\left(\frac{1}{\sqrt{N_n}}\frac{\partial^{\alpha}}{\partial\xi^{\alpha}}\left\{W(\theta)^{T}\ddot{\Omega}^{-1}\epsilon\right\}\right)^{2}\right] = \frac{a_0^2}{N_n}\sum_{0\leq i,j,k\leq N_n}\frac{\partial^{\alpha_2}\rho_{|i-k|}(\theta)}{\partial\theta^{\alpha_2}}\frac{\partial^{\alpha}\ddot{\omega}^{i,j}}{\partial\xi^{\alpha}}\frac{\partial^{\alpha}\ddot{\omega}^{k,j}}{\partial\xi^{\alpha}}, \tag{C.44}$$

and from here using that $\frac{\partial^{\alpha}\ddot{\omega}^{k,j}}{\partial\xi^{\alpha}} = O\left(\frac{1}{\sqrt{N_n}}\right)$ when $\alpha_0 > 0$ on the one hand, then summing over k and applying Assumption (2.11) and finally computing the explicit sum of exponential terms leads to

$$\mathbb{E}_{\mathcal{U}}\left[\left(\frac{1}{\sqrt{N_n}}\frac{\partial^{\alpha}}{\partial\xi^{\alpha}}\left\{W(\theta)^{T}\ddot{\Omega}^{-1}\epsilon\right\}\right)^{2}\right] = O_{\mathbb{P}}\left(N_n^{-1/2}\right) \tag{C.45}$$

uniformly in $\xi\in\Xi$. Thus we have proved (C.38). Finally, (C.39) is proved similarly. □

C.5. Proof of Theorem 4.1

We derive the limit theory for $\widehat{\xi}_{n,err}$ (hereafter denoted by $\widehat{\xi}_n$) when $a_0^2 > 0$. In our terminology, we recall that we have for any $\xi\in\Xi$ and up to an additive constant term

$$l_n(\xi) = \underbrace{-\frac{1}{2}\log\det(\Omega) - \frac{1}{2}\widetilde{Y}(\theta_0)^{T}\Omega^{-1}\widetilde{Y}(\theta_0)}_{l_n^{(\sigma^2,a^2)}(\xi)}\underbrace{-\frac{1}{2}\mathbf{G}_n(\xi) - \mathbf{K}_n(\xi)}_{l_n^{(\theta)}(\xi)}, \tag{C.46}$$

with $\Omega^{-1} = [\omega^{i,j}]_{1\leq i\leq N_n,1\leq j\leq N_n}$, where we also recall that the definition of $\omega^{i,j}$ can be found in (C.3). Moreover, note that $l_n^{(\sigma^2,a^2)}$ does not depend on θ and corresponds precisely to the quasi log-likelihood with no information as studied in Xiu (2010) and extended to a more general setting in Clinet and Potiron (2018a). On the other hand, $l_n^{(\theta)}$ is the additional part incorporating θ and depends on the whole vector $\xi = (\sigma^2, a^2, \theta)$. Following a similar procedure as in Xiu (2010) and Clinet and Potiron (2018a), we introduce the approximate log-likelihood random field as

$$\bar{l}_n(\xi) = \underbrace{-\frac{1}{2}\log\det(\Omega) - \frac{1}{2}\mathrm{Tr}\left(\Omega^{-1}\left\{\Sigma_0^c + \Sigma_0^d\right\}\right)}_{\bar{l}_n^{(\sigma^2,a^2)}(\xi)}\underbrace{-\frac{1}{2}\mathbf{G}_n(\xi)}_{\bar{l}_n^{(\theta)}(\xi)}, \tag{C.47}$$

with

$$
\Sigma_0^c =
\begin{pmatrix}
\int_0^{t_1^n} \sigma_s^2 ds + 2a^2 & -a^2 & 0 & \cdots & 0 \\
-a^2 & \int_{t_1^n}^{t_2^n} \sigma_s^2 ds + 2a^2 & -a^2 & \ddots & \vdots \\
0 & -a^2 & \int_{t_2^n}^{t_3^n} \sigma_s^2 ds + 2a^2 & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & -a^2 \\
0 & \cdots & 0 & -a^2 & \int_{t_{N_n-1}^n}^{t_{N_n}^n} \sigma_s^2 ds + 2a^2
\end{pmatrix},
$$

and

$$
\Sigma_0^d = \operatorname{diag} \left(\sum_{0 < s \le t_1^n} \Delta J_s^2, \sum_{t_1^n < s \le t_2^n} \Delta J_s^2, \ldots, \sum_{t_{N_n-1}^n < s \le t_{N_n}^n} \Delta J_s^2 \right).
$$

Consider the diagonal scaling matrix

$$
\Phi_n = \operatorname{diag}(N_n^{1/2}, N_n, N_n). \tag{C.48}
$$

Define also for $\xi \in \Xi$ the scaled scores $\Psi_n^{(\sigma^2,a^2)}(\xi) = -\Phi_n^{-1} \frac{\partial l_n^{(\sigma^2,a^2)}(\xi)}{\partial \xi}$, $\Psi_n^{(\theta)}(\xi) = -\Phi_n^{-1} \frac{\partial l_n^{(\theta)}(\xi)}{\partial \xi}$, and $\Psi_n = \Psi_n^{(\sigma^2,a^2)} + \Psi_n^{(\theta)}$. Accordingly, the approximate scores $\overline{\Psi}_n$, $\overline{\Psi}_n^{(\sigma^2,a^2)}$, and $\overline{\Psi}_n^{(\theta)}$ admit the same definition replacing l_n by $\bar{l}_n$. We start by a technical lemma to ensure the uniform convergence of some random fields.

Lemma C.8. *Let $X_n(\xi)$ be a sequence of random variables of class C^m in $\xi \in \Xi_n \subset \mathbb{R}^d$, each Ξ_n convex compact, such that $2m > d$, and $\mathcal{U}$ a sub σ-field of the general σ-field. For any multi-index $\boldsymbol{\alpha}$ such that $0 \le |\boldsymbol{\alpha}| \le m$, we assume that*

$$
\sup_{\xi \in \Xi_n} \mathbb{E}_{\mathcal{U}} \left[\partial_\xi^{\boldsymbol{\alpha}} X_n(\xi)^2 \right] = o_\mathbb{P}(\operatorname{Leb}(\Xi_n)).
$$

Then we have the uniform convergence

$$
\mathbb{E}_{\mathcal{U}} \left[\sup_{\xi \in \Xi_n} |X_n(\xi)| \right] \to^\mathbb{P} 0. \tag{C.49}
$$

Proof. By Theorem 4.12 Part I case A (taking $j = 0$, $p = 2$) from Adams and Fournier (2003), we apply Sobolev's inequality and define some constant M such that we have

$$
\mathbb{E}_{\mathcal{U}} \left[\sup_{\xi \in \Xi_n} |X_n(\xi)| \right] \le M \sum_{\boldsymbol{\alpha} || \boldsymbol{\alpha}| \le m} \left(\int_{\Xi_n} \mathbb{E} \left[\partial_\xi^{\boldsymbol{\alpha}} X_n(\xi)^2 \right] d\xi \right)^{1/2}
$$

$$
\le M \operatorname{Leb}(\Xi_n)^{1/2} \sum_{\boldsymbol{\alpha} || \boldsymbol{\alpha}| \le m} \left(\sup_{\xi \in \Xi_n} \mathbb{E} \left[\partial_\xi^{\boldsymbol{\alpha}} X_n(\xi)^2 \right] \right)^{1/2} \to^\mathbb{P} 0. \quad \square
$$

Lemma C.9 (*Asymptotic Score*)**.** *For any $\xi \in \Xi$, let*

$$
\Psi_\infty(\xi) =
\begin{pmatrix}
-\dfrac{\sqrt{T}}{8a\sigma^3} \left(\overline{\sigma}_0^2 - \sigma^2 \right) - \dfrac{\sqrt{T}}{8a^3\sigma} \left(a^2 - a_0^2 - \rho_0(\theta) - 2 \sum_{k=1}^{+\infty} \rho_k(\theta) \right) \\
\dfrac{1}{2a^4} \left(a^2 - a_0^2 - \rho_0(\theta) \right) \\
\dfrac{1}{2a^2} \dfrac{\partial \rho_0(\theta)}{\partial \theta}
\end{pmatrix}.
$$

We have

$$
\sup_{\xi \in \Xi} |\Psi_n(\xi) - \Psi_\infty(\xi)| \to^\mathbb{P} 0. \tag{C.50}
$$

Proof. Since $\Psi_n = \Psi_n^{(\sigma^2, a^2)} + \Psi_n^{(\theta)}$, the lemma will be proved if we can show that

$$\Psi_n^{(\sigma^2, a^2)}(\xi) \to^{\mathbb{P}} \begin{pmatrix} -\dfrac{\sqrt{T}}{8a\sigma^3}\left(\overline{\sigma}_0^2 - \sigma^2\right) - \dfrac{\sqrt{T}}{8a^3\sigma}\left(a^2 - a_0^2\right) \\[2mm] \dfrac{1}{2a^4}\left(a^2 - a_0^2\right) \\[2mm] 0 \end{pmatrix} \tag{C.51}$$

and

$$\Psi_n^{(\theta)}(\xi) \to^{\mathbb{P}} \begin{pmatrix} \dfrac{\sqrt{T}}{8a^3\sigma}\left(\rho_0(\theta) + 2\sum_{k=1}^{+\infty}\rho_k(\theta)\right) \\[2mm] -\dfrac{\rho_0(\theta)}{2a^4} \\[2mm] \dfrac{1}{2a^2}\dfrac{\partial\rho_0(\theta)}{\partial\theta} \end{pmatrix} \tag{C.52}$$

uniformly in $\xi \in \Xi$. Note that (C.51) is a consequence of Lemma A.3 in Clinet and Potiron (2018a). For (C.52), by Lemma C.7 combined with Lemma C.8, we have uniformly in $\xi \in \Xi$ that $\Psi_n^{(\theta)}(\xi) - \overline{\Psi}_n^{(\theta)}(\xi) = o_{\mathbb{P}}(1)$. Thus it is sufficient to show that we have the convergence (C.52) for $\overline{\Psi}_n^{(\theta)}$. Combining Lemmas C.6 and C.8, we obtain

$$\sup_{\xi\in\Xi}\left\{\frac{1}{2\sqrt{N_n}}\frac{\partial \mathbf{G}_n(\xi)}{\partial\sigma^2} + \frac{u_0}{8a^2\sigma^2}\left\{\rho_0(\theta) + 2\sum_{k=1}^{+\infty}\rho_k(\theta)\right\}\right\} \to^{\mathbb{P}} 0, \tag{C.53}$$

$$\sup_{\xi\in\Xi}\left\{\frac{1}{2N_n}\frac{\partial \mathbf{G}_n(\xi)}{\partial a^2} + \frac{\rho_0(\theta)}{2a^4}\right\} \to^{\mathbb{P}} 0, \tag{C.54}$$

and

$$\sup_{\xi\in\Xi}\left\{\frac{1}{2N_n}\frac{\partial \mathbf{G}_n(\xi)}{\partial\theta} - \frac{1}{2a^2}\frac{\partial\rho_0(\theta)}{\partial\theta}\right\} \to^{\mathbb{P}} 0, \tag{C.55}$$

and recalling that $\overline{\Psi}_n^{(\theta)} = \frac{1}{2}\Phi_n^{-1}\frac{\partial \mathbf{G}_n(\xi)}{\partial\xi}$, we get (C.52). □

Theorem C.10 (*Consistency*). *If $\widehat{\xi}_n = (\widehat{\sigma}_n^2, \widehat{a}_n^2, \widehat{\theta}_n)$ is the QMLE, we have*

$$\widehat{\xi}_n \to^{\mathbb{P}} \xi_0 := \left(\overline{\sigma}_0^2, a_0^2, \theta_0\right). \tag{C.56}$$

Proof. We extend the proof of Lemma A.4 in Clinet and Potiron (2018a), that is, since we already have

$$\sup_{\xi\in\Xi}|\Psi_n(\xi) - \Psi_\infty(\xi)| \to^{\mathbb{P}} 0 \tag{C.57}$$

by Lemma C.9, we show that for any $\epsilon > 0$

$$\inf_{\xi\in\Xi:\|\xi-\xi_0\|\geq\epsilon}\|\Psi_\infty(\xi)\|^2 > 0 = \|\Psi_\infty(\xi_0)\|^2 \; \mathbb{P}\text{-a.s.} \tag{C.58}$$

Given the form of Ψ_∞, the equality $\Psi_\infty(\xi_0) = 0$ is immediate. Note also that the left hand side inequality of (C.58) will be automatically satisfied if we show that $\|\Psi_\infty(\xi)\|^2 > 0$ as soon as $\xi \neq \xi_0$ by a continuity argument since Ξ is compact. Let us then take $\xi \in \Xi - \{\xi_0\}$ such that $\Psi_\infty(\xi) = 0$, and assume first that $\theta \neq \theta_0$. In that case, we have

$$\|\Psi_\infty(\xi)\|^2 \geq \frac{1}{4a^4}\left\|\frac{\partial\rho_0(\theta)}{\partial\theta}\right\|^2 > 0,$$

by (2.14), which leads to a contradiction. We thus get $\theta = \theta_0$ and in a similar way, we also have

$$0 = \|\Psi_\infty(\sigma^2, a^2, \theta_0)\|^2 \geq \frac{1}{4a^8}\left(a^2 - a_0^2\right)^2,$$

that implies $a^2 = a_0^2$. Finally, the first component of Ψ_∞ leads to the domination

$$0 = \|\Psi_\infty(\sigma^2, a_0^2, \theta_0)\|^2 \geq \frac{T}{64a_0^2\sigma^6}\left(\overline{\sigma}_0^2 - \sigma^2\right)^2,$$

so that we can conclude $\sigma^2 = \overline{\sigma}_0^2$. □

Let H_n be the scaled Hessian matrix of the likelihood field, defined as

$$H_n(\xi) = -\Phi_n^{-1/2} \frac{\partial^2 l_n(\xi)}{\partial \xi^2} \Phi_n^{-1/2},$$

(C.59)

and similarly $H_n^{(\sigma^2,a^2)}, H_n^{(\theta)}, \overline{H}_n, \overline{H}_n^{(\sigma^2,a^2)}, \overline{H}_n^{(\theta)}$.

Lemma C.11 (*Fisher Information*). *For $\xi_0 = (\sigma_0^2, a_0^2, \theta_0)$, let $\Gamma(\xi_0)$ be the matrix*

$$\Gamma(\xi_0) = \begin{pmatrix} \dfrac{\sqrt{T}}{8a_0\sigma_0^3} & 0 & 0 \\ 0 & \dfrac{1}{2a_0^4} & 0 \\ 0 & 0 & a_0^2 V_{\theta_0} \end{pmatrix}.$$

(C.60)

We have, for any ball V_n, centered on ξ_0 and shrinking to $\{\xi_0\}$,

$$\sup_{\xi_n \in V_n} \|H_n(\xi_n) - \Gamma(\xi_0)\| \to^{\mathbb{P}} 0.$$

(C.61)

Proof. As for the proof of Lemma C.9, we have

$$H_n^{(\sigma^2,a^2)}(\xi) \to^{\mathbb{P}} \begin{pmatrix} \dfrac{\sqrt{T}}{8a_0\sigma_0^3} & 0 & 0 \\ 0 & \dfrac{1}{2a_0^4} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(C.62)

uniformly in $\xi \in \Xi$ by Lemma A.5 in Clinet and Potiron (2018a), so that by the identity $H_n = H_n^{(\sigma^2,a^2)} + H_n^{(\theta)}$ the lemma will be proved if we can show

$$H_n^{(\theta)}(\xi) \to^{\mathbb{P}} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a_0^2 V_{\theta_0} \end{pmatrix},$$

(C.63)

uniformly in $\xi \in \Xi$. Since $H_n^{(\theta)} - \overline{H}_n^{(\theta)} = -\Phi_n^{-1/2} \frac{\partial^2 \mathbf{K}_n}{\partial \xi^2} \Phi_n^{-1/2}$, an immediate application of Lemmas C.7 and C.8 yields $\sup_{\xi \in \Xi} \left\{ H_n^{(\theta)}(\xi) - \overline{H}_n^{(\theta)}(\xi) \right\} \to^{\mathbb{P}} 0$. Moreover, by Lemmas C.6 and C.8, we have

$$\sup_{\xi \in \Xi} \left\{ \overline{H}_n^{(\theta)}(\xi) - \overline{H}_\infty^{(\theta)}(\xi) \right\} \to^{\mathbb{P}} 0,$$

(C.64)

with

$$\overline{H}_\infty^{(\theta)}(\xi) = \frac{1}{2} \begin{pmatrix} \dfrac{\partial^2 \mathbf{G}_{\infty,1}(\xi)}{\partial (\sigma^2)^2} & 0 & 0 \\ 0 & \dfrac{\partial^2 \mathbf{G}_{\infty,2}(\xi)}{\partial (a^2)^2} & \dfrac{\partial^2 \mathbf{G}_{\infty,2}(\xi)}{\partial a^2 \partial \theta} \\ 0 & \dfrac{\partial^2 \mathbf{G}_{\infty,2}(\xi)}{\partial \theta \partial a^2} & \dfrac{\partial^2 \mathbf{G}_{\infty,2}(\xi)}{\partial \theta^2} \end{pmatrix}.$$

(C.65)

Thus, by continuity of $\overline{H}_\infty^{(\theta)}$, we deduce

$$\sup_{\xi_n \in V_n} \left\{ \overline{H}_n^{(\theta)}(\xi_n) - \overline{H}_\infty^{(\theta)}(\xi_0) \right\} \to^{\mathbb{P}} 0,$$

(C.66)

and moreover it is immediate to check that

$$\overline{H}_\infty^{(\theta)}(\xi_0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a_0^2 V_{\theta_0} \end{pmatrix},$$

(C.67)

in view of definitions (2.7), (2.9), (C.16) and (C.17). □

We now extend the notations of Clinet and Potiron (2018a) (A.27)-(A.30) p. 128, and we define a few processes involved in the derivation of the central limit theorem. For $(\beta) \in \{(\sigma^2), (a^2), (\theta)\}$, and $t \in [0, T]$,

$$M_1^{(\beta)}(t) := \sum_{i=1}^{N_n(t)} \frac{\partial \omega^{i,i}}{\partial \beta} \left\{ \left(\Delta X_{i,t}^n \right)^2 - \int_{t_{i-1}^n \wedge t}^{t_i^n \wedge t} \sigma_s^2 ds - \sum_{t_{i-1}^n \wedge t < s \le t_i^n \wedge t} \Delta J_s^2 \right\}, \tag{C.68}$$

$$M_2^{(\beta)}(t) := \sum_{i=1}^{N_n(t)} \left\{ \sum_{1 \le j < i} \frac{\partial \omega^{i,j}}{\partial \beta} \Delta X_{j,t}^n \right\} \Delta X_{i,t}^n, \tag{C.69}$$

$$M_3^{(\beta)}(t) := -2 \sum_{i=0}^{N_n(t)} \left\{ \sum_{j=1}^{N_n(t)} \frac{\partial \dot\omega^{i,j}}{\partial \beta} \Delta X_{j,t}^n \right\} \epsilon_{t_i^n}, \tag{C.70}$$

$$M_4^{(\beta)}(t) := \sum_{i=0}^{N_n(t)} \frac{\partial \ddot\omega^{i,i}}{\partial \beta} \left\{ \epsilon_{t_i^n}^2 - a_0^2 \right\} + 2 \sum_{i=0}^{N_n(t)} \left\{ \sum_{0 \le j < i} \frac{\partial \ddot\omega^{i,j}}{\partial \beta} \epsilon_{t_j^n} \right\} \epsilon_{t_i^n}, \tag{C.71}$$

$$M_5^{(\beta)}(t) := \sum_{i=0}^{N_n(t)} \sum_{j=0}^{N_n(t)} \ddot\omega^{i,j} \frac{\partial W_j(\theta_0)}{\partial \beta} \epsilon_{t_i^n} + \sum_{i=1}^{N_n(t)} \sum_{j=0}^{N_n(t)} \omega^{i,j} \frac{\partial W_j(\theta_0)}{\partial \beta} \Delta X_{i,t}^n. \tag{C.72}$$

We also define the three-dimensional vectors $M_i(t) := \left(M_i^{(\sigma^2)}(t), M_i^{(a^2)}(t), M_i^{(\theta)}(t) \right)^T$ for $i \in \{1, \dots, 5\}$. In all the definitions (C.68)–(C.72), the terms involving the parameters such as $\Omega^{-1}, \dot\Omega^{-1}, \ddot\Omega^{-1}...$ are evaluated at point $\xi := (\sigma^2, a_0^2, \theta_0)$, for some $\sigma^2 \in [\underline{\sigma}^2, \overline{\sigma}^2]$. For $i \in \{1, \dots, 4\}$, when properly scaled, the processes $M_i(T)$ admit limit distributions whose expressions can be found in Lemma A.6, A.7, A.9 and A.10 of Clinet and Potiron (2018a). We complete these results and show that $M_5(T)$ tends in distribution conditioned on $\mathcal{G}_T$ to a normal distribution. Before stating the results, we recall that for a σ-field $\mathcal{H}$, a random vector Z and a sequence of random vectors Z_n in $\mathbb{R}^b$, Z_n is said to converge in law towards Z conditioned on $\mathcal{H}$ if we have for any $u \in \mathbb{R}^b$

$$\mathbb{E}\left[e^{iu^T Z_n} \,\middle|\, \mathcal{H} \right] \stackrel{\mathbb{P}}{\to} \mathbb{E}\left[e^{iu^T Z} \,\middle|\, \mathcal{H} \right]. \tag{C.73}$$

Lemma C.12.
  We have, conditionally on $\mathcal{G}_T$, the convergence in distribution

$$N_n^{-1/2} M_5(T) \to \mathcal{N}\left(0, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a_0^{-2} V_{\theta_0} \end{pmatrix} \right). \tag{C.74}$$

Proof. Since $M_5^{(\sigma^2)}(T) = M_5^{(a^2)}(T) = 0$, it is sufficient to prove the marginal CLT for $M_5^{(\theta)}(T)$. The proof is conducted in two steps.
  Step 1. We introduce

$$\tilde{M}_5^{(\theta)}(T) := a_0^{-2} \sum_{i=0}^{N_n} \frac{\partial W_i(\theta_0)}{\partial \theta} \epsilon_{t_i^n}. \tag{C.75}$$

We show that $N_n^{-1/2} \left\{ M_5^{(\theta)}(T) - \tilde{M}_5^{(\theta)}(T) \right\} \to^{\mathbb{P}} 0$. We decompose

$$N_n^{-1/2} \left\{ M_5^{(\theta)}(T) - \tilde{M}_5^{(\theta)}(T) \right\} = R_n^{(1)} + R_n^{(2)} + R_n^{(3)}, \tag{C.76}$$

where

$$R_n^{(1)} = N_n^{-1/2} \sum_{i=1}^{N_n} \sum_{j=0}^{N_n} \dot\omega^{i,j} \frac{\partial W_j(\theta_0)}{\partial \theta} \Delta X_{i,t}^n, \tag{C.77}$$

$$R_n^{(2)} = N_n^{-1/2} \sum_{i=0}^{N_n} \sum_{j \ne i} \ddot\omega^{i,j} \frac{\partial W_j(\theta_0)}{\partial \theta} \epsilon_{t_i^n}, \tag{C.78}$$

and

$$R_n^{(3)} = N_n^{-1/2} \sum_{i=0}^{N_n} \left\{ \ddot\omega^{i,i} - a_0^{-2} \right\} \frac{\partial W_i(\theta_0)}{\partial \theta} \epsilon_{t_i^n}. \tag{C.79}$$

Now, by a similar proof as in Lemma C.7, we have easily $R_n^{(1)} \to^{\mathbb{P}} 0$. Moreover, by application of Lemma C.4, (2.11) and the independence of ϵ and $W(\theta_0)$, we easily get that $R_n^{(2)} \to^{\mathbb{P}} 0$. Finally, by Lemma C.4 we have $\ddot{\omega}^{i,i} - a_0^{-2} = O\left(N_n^{-1/2}\right)$ so that we obtain directly $\mathbb{E}_\mathcal{U} R_n^{(3)} = 0$ by independence of $\frac{\partial W(\theta_0)}{\partial \theta}$ and ϵ and for each component $R_{n,k}^{(3)}$ of $R_n^{(3)}$, $1 \le k \le d$,

$$
\begin{aligned}
\mathrm{Var}_\mathcal{U} R_{n,k}^{(3)} &\le O\left(N_n^{-2}\right) \times \sum_{i,j=0}^{N_n} \mathbb{E}_\mathcal{U}\left[\frac{\partial W_i(\theta_0)}{\partial \theta_k} \epsilon_{t_i^n} \frac{\partial W_j(\theta_0)}{\partial \theta_k} \epsilon_{t_j^n}\right] \\
&= O\left(N_n^{-1}\right),
\end{aligned}
$$

and thus $R_n^{(3)} \to^{\mathbb{P}} 0$.

Step 2. Now we show that $N_n^{-1/2}\tilde{M}_5^{(\theta)}(T) \to \mathcal{N}\left(0, a_0^{-2}V_{\theta_0}\right)$ conditionally on $\tilde{\mathcal{G}}_T = \mathcal{G}_T \vee \{Q_i^n | i, n, \in \mathbb{N}\}$ (and so conditionally on $\mathcal{G}_T$). To do so, we will apply a conditional version of Theorem 5.12 from Kallenberg (2006, p. 92) with respect to $\tilde{\mathcal{G}}_T$. Accordingly, we first remark that we have the representation $N_n^{-1/2}\tilde{M}_5^{(\theta)}(T) = \sum_{i=0}^{N_n} \chi_i^n$ where $\chi_i^n = N_n^{-1/2} a_0^{-2} \frac{\partial W_i(\theta_0)}{\partial \theta} \epsilon_{t_i^n}$ are rowwise conditionally independent and centered given $\tilde{\mathcal{G}}_T$. To get the desired convergence in distribution, it is thus sufficient to show that

$$
N_n^{-1} \sum_{i=0}^{N_n} \mathbb{E}\left[\chi_i^n (\chi_i^n)^T \,\middle|\, \tilde{\mathcal{G}}_T\right] \to^{\mathbb{P}} a_0^{-2} V_{\theta_0}, \tag{C.80}
$$

and Lindeberg's condition, for some $p > 0$,

$$
\sum_{i=0}^{N_n} \mathbb{E}\left[\|\chi_i^n\|^p \,\middle|\, \tilde{\mathcal{G}}_T\right] \to^{\mathbb{P}} 0.
$$

For (C.80), we immediately note that

$$
N_n^{-1} \sum_{i=0}^{N_n} \mathbb{E}\left[(\chi_i^n)^2 \,\middle|\, \tilde{\mathcal{G}}_T\right] = a_0^{-4} N_n^{-1} \sum_{i=0}^{N_n} \frac{\partial W_i(\theta_0)}{\partial \theta} \frac{\partial W_i(\theta_0)^T}{\partial \theta} \to^{\mathbb{P}} a_0^{-4} V_{\theta_0} \tag{C.81}
$$

where we have used the independence of ϵ and Q, and where the final step is a straightforward consequence of conditions (2.11) and (2.12). Moreover Lindeberg's condition (C.81) for $p = 4$ is also easily verified using moment conditions and stationarity of the information process. □

Lemma C.13. We have for any $\sigma^2 \in [\underline{\sigma}^2, \overline{\sigma}^2]$, taking $\xi = (\sigma^2, a_0^2, \theta_0)$, stably in $\mathcal{G}_T$, the convergence in distribution

$$
\Phi_n^{1/2}\left\{\Psi_n(\xi) - \overline{\Psi}_n(\xi)\right\} \to \mathcal{MN}\left(0, \begin{pmatrix} \frac{1}{4a_0}\left(\frac{5\mathcal{Q}}{16\sigma^7 T^{1/2}} + \frac{\overline{\sigma}_0^2 \sqrt{T}}{8\sigma^5} + \frac{\sqrt{T}}{16\sigma^3}\right) & 0 & 0 \\ 0 & \frac{1}{2a_0^4} + \frac{\mathrm{cum}_4[\epsilon]}{4a_0^8} & 0 \\ 0 & 0 & a_0^{-2}V_{\theta_0} \end{pmatrix}\right),
$$

where $\mathcal{Q} = T^{-1}\int_0^T \alpha_s^{-1} ds \left\{\int_0^T \sigma_s^4 \alpha_s ds + \sum_{0 < s \le T} \Delta J_s^2(\sigma_s^2 \alpha_s + \sigma_{s-}^2 \alpha_{s-})\right\}$.

Proof. Note that we have the decomposition

$$
2\Phi_n^{1/2}\left\{\Psi_n(\xi) - \overline{\Psi}_n(\xi)\right\} = \Phi_n^{-1/2}\left\{M_1(T) + 2M_2(T) + M_3(T) + M_4(T) + 2M_5(T)\right\}.
$$

First we show that $\Phi_n^{-1/2}\left\{M_3(T) + M_4(T) + 2M_5(T)\right\}$ converges in distribution conditioned on $\mathcal{G}_T$. By Lemma A.9 and A.10 in Clinet and Potiron (2018a), $\Phi_n^{-1/2}M_3(T)$ and $\Phi_n^{-1/2}M_4(T)$ both tend in distribution conditioned on $\mathcal{G}_T$ to mixed normal distributions of respective asymptotic variances

$$
V_3 = \begin{pmatrix} \frac{\overline{\sigma}_0^2 \sqrt{T}}{8\sigma^5 a_0} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{C.82}
$$

and

$$
V_4 = \begin{pmatrix} \frac{\sqrt{T}}{16a_0\sigma^2} & 0 & 0 \\ 0 & \frac{2}{a_0^4} + \frac{\mathrm{cum}_4[\epsilon]}{a_0^8} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{C.83}
$$

Moreover, by independence of ϵ with the other processes and the fact that $\frac{\partial W_j(\theta_0)}{\partial \theta}$ and ΔX are uncorrelated, we deduce that the conditional covariance terms between any pair $M_i(T), M_j(T)$ $i \neq j, i, j \in \{3, 4, 5\}$ are null, so that along with the marginal convergence obtained in Lemma C.12 this automatically yields the convergence in law conditioned on $\mathcal{G}_T$

$$\Phi_n^{1/2}\{M_3(T) + M_4(T) + 2M_5(T)\} \rightarrow \mathcal{MN}\left(0, \begin{pmatrix} \frac{1}{a_0}\left(\frac{\overline{\sigma}_0^2\sqrt{T}}{8\sigma^5} + \frac{\sqrt{T}}{16\sigma^3}\right) & 0 & 0 \\ 0 & \frac{2}{a_0^4} + \frac{\mathrm{cum}_4[\epsilon]}{a_0^8} & 0 \\ 0 & 0 & 4a_0^{-2}V_{\theta_0} \end{pmatrix}\right).$$

By Slutsky's lemma and Lemma A.6 and A.7 from Clinet and Potiron (2018a), we also have the $\mathcal{G}_T$-stable convergence in distribution of $M_1(T) + 2M_2(T)$ towards a mixed normal distribution of random variance

$$V_2 = \begin{pmatrix} \frac{5\mathcal{Q}}{16T^{1/2}\sigma^7 a_0} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{C.84}$$

Finally, by application of Proposition A.8 from Clinet and Potiron (2018a), we deduce the joint $\mathcal{G}_T$-stable convergence of $(M_1(T) + 2M_2(T), M_3(T) + M_4(T) + 2M_5(T))$, hence the convergence of the global term $\Phi_n^{1/2}\{\Psi_n(\xi) - \overline{\Psi}_n(\xi)\}$, and we are done. $\square$

We are now ready to prove the central limit theorem.

Proof of Theorem 4.1. The proof follows exactly the proof of Theorem A.12 in Clinet and Potiron (2018a) in the case $B = 1$, and replacing $\xi = (\sigma^2, a^2)$ by $\xi = (\sigma^2, a^2, \theta)$, and the score function along with the Fisher information by their three dimensional counterparts. $\square$

C.6. Proof of Theorem 3.1

In the case $a_0^2 = \eta_0 T/n$, that is in the small residual noise framework, we are interested in deriving the joint law of two estimators, which are $\widehat{\xi}_{n,err} = (\widehat{\sigma}_{n,err}^2, \widehat{\theta}_{n,err}, \widehat{a}_{n,err}^2)$ and $\widehat{\upsilon}_{n,exp} = (\widehat{\sigma}_{n,exp}^2, \widehat{\theta}_{n,exp})$, which is obtained under the no-residual noise constraint $a = 0$ in the quasi-likelihood function. For the former estimator, the limit theory is fairly different from the case of a fixed non-zero noise. Indeed, we are going to show that the rate of convergence $(N_n^{1/4}, N_n^{1/2}, N_n^{1/2})$ is changed to $(N_n^{1/2}, N_n^{3/2}, N_n)$. Moreover, the noise and price increments being of the same order, complex interaction terms now appear in the limit variance of the estimators. To derive a CLT for $(\widehat{\upsilon}_{n,exp}, \widehat{\xi}_{n,err})$, let us reformulate a bit the problem and, introducing $\widehat{\eta}_{n,err} := \Delta_{N_n}^{-1}\widehat{a}_{n,err}^2, \widehat{u}_{n,err} := N_n^{1/2}(\widehat{\theta}_{n,err} - \theta_0)$, and $\widehat{u}_{n,exp} := N_n^{1/2}(\widehat{\theta}_{n,exp} - \theta_0)$, we are now interested in showing that

$$\widehat{\upsilon}_n = \left(\widehat{w}_{n,exp}, \widehat{\zeta}_{n,err}\right) := (\widehat{\sigma}_{n,exp}^2, \widehat{u}_{n,exp}, \widehat{\sigma}_{n,err}^2, \widehat{u}_{n,err}, \widehat{\eta}_{n,err}) \tag{C.85}$$

admits a CLT with rate $N_n^{1/2}$. We note that $\widehat{\zeta}_{n,err}$ is the QMLE related to the new log-likelihood

$$\mathcal{L}_{n,err}(\sigma^2, u, \eta) := -\frac{1}{2}\log\det(\Lambda) - \frac{1}{2}(Y - \mu(\theta_0 + N_n^{-1/2}u))^T \Lambda^{-1}(Y - \mu(\theta_0 + N_n^{-1/2}u)), \tag{C.86}$$

where now Ω is replaced by Λ with

$$\Lambda = \Delta_{N_n}\begin{pmatrix} \sigma^2 + 2\eta & -\eta & 0 & \cdots & 0 \\ -\eta & \sigma^2 + 2\eta & -\eta & \ddots & \vdots \\ 0 & -\eta & \sigma^2 + 2\eta & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\eta \\ 0 & \cdots & 0 & -\eta & \sigma^2 + 2\eta \end{pmatrix}, \tag{C.87}$$

and $\widehat{w}_{n,exp}$ is one maximizer in the variables (σ^2, u) of

$$\mathcal{L}_{n,exp}(\sigma^2, u) := \mathcal{L}_{n,err}(\sigma^2, u, 0), \tag{C.88}$$

that is (C.86) with $\Lambda = \Delta_{N_n}\sigma^2\mathbf{I}_{N_n}$, where $\mathbf{I}_{N_n} \in \mathbb{R}^{N_n \times N_n}$ is the identity matrix. We now adopt somewhat similar notations to Aït-Sahalia and Xiu (2016), proof of Theorem 1, and we simplify the problem introducing the change of variables

$$\phi = 1 - \frac{1}{2\eta}\left\{\sqrt{\sigma^2(4\eta + \sigma^2)} - \sigma^2\right\} \text{ and } \gamma^2 = \frac{1}{2}\left\{2\eta + \sigma^2 + \sqrt{\sigma^2(4\eta + \sigma^2)}\right\},$$

and $\phi = 0$ when $\eta = 0$. We have

$$\lambda^{i,j} = \Delta_{N_n}^{-1} \gamma^{-2} \frac{\phi^{|i-j|} - \phi^{i+j} - \phi^{2N_n+2-(i+j)} - \phi^{2N_n+2-|i-j|}}{(1-\phi^2)(1-\phi^{2n+2})}, \tag{C.89}$$

with the convention $0^0 = 1$. We are thus going to derive the asymptotic properties of

$$\tilde{\nu}_n = (\tilde{w}_{n,exp}, \tilde{\zeta}_{n,err}) = (\widehat{\sigma}^2_{n,exp}, \widehat{u}_{n,exp}, \widehat{\gamma}^2_{n,err}, \widehat{u}_{n,err}, \widehat{\phi}_{n,err}),$$

obtained by maximization of the log-likelihood functions seen respectively as functions of $\zeta := (\gamma^2, u, \phi)$ and $w := (\sigma^2, u)$. Given the form of ϕ and η, we see that there exists $\overline{\phi} \in (0, 1)$ and $\overline{\gamma}^2 > \underline{\gamma}^2 > 0$ such that the optimizations are respectively conducted on the sets $\Xi_{n,err} := \left[\underline{\gamma}^2, \overline{\gamma}^2\right] \times \{u \in \mathbb{R}^d_+ | \theta_0 + N_n^{-1/2}u \in \Theta\} \times \left[-\overline{\phi}, \overline{\phi}\right]$, and $\Xi_{n,exp} := \left[\underline{\gamma}^2, \overline{\gamma}^2\right] \times \{u \in \mathbb{R}^d_+ | \theta_0 + N_n^{-1/2}u \in \Theta\}$. Then, we will get back to $\widehat{\zeta}_{n,err}$ by the delta method.

We keep similar notations as in the previous part and we write $\Psi_{n,err}(\zeta) = -N_n^{-1} \frac{\partial \mathcal{L}_{n,err}}{\partial \zeta}$ the score function for the first experiment, and similarly $\Psi_{n,exp}(w) = -N_n^{-1} \frac{\partial \mathcal{L}_{n,exp}}{\partial w}$. Sometimes we will consider the joint process $\Psi_n = (\Psi_{n,exp}, \Psi_{n,err})$. We also naturally adapt the notations (C.15) and (C.37) to the small noise context

$$\mathbf{G}_n(\zeta) = (\mu(\theta_0) - \mu(\theta_0 + N_n^{-1/2}u))^T \Lambda^{-1}(\mu(\theta_0) - \mu(\theta_0 + N_n^{-1/2}u)), \tag{C.90}$$

and

$$\mathbf{K}_n(\zeta) = (\mu(\theta_0) - \mu(\theta_0 + N_n^{-1/2}u))^T \Lambda^{-1} \{\Delta X + \Delta \epsilon\}. \tag{C.91}$$

We first give the limit of both estimators when $a_0^2 = \eta_0 T/n$.

Theorem C.14 (*Consistency*). *Assume that $a_0^2 = \eta_0 T/n$. Define $\tilde{\eta}_0 := \eta_0 T^{-1} \int_0^T \alpha_s^{-1} ds$. Let $\phi_0 := 1 - \frac{1}{2\tilde{\eta}_0} \left\{\sqrt{\overline{\sigma}_0^2(4\tilde{\eta}_0 + \overline{\sigma}_0^2)} - \overline{\sigma}_0^2\right\}$, and $\gamma_0^2 := \frac{1}{2} \left\{2\tilde{\eta}_0 + \overline{\sigma}_0^2 + \sqrt{\overline{\sigma}_0^2(4\tilde{\eta}_0 + \overline{\sigma}_0^2)}\right\}$. Let*

$$\nu_0 := (w_{0,exp}, \zeta_{0,err}) = (\overline{\sigma}_0^2 + 2\tilde{\eta}_0, 0, \gamma_0^2, 0, \phi_0).$$

We have

$$\tilde{\nu}_n \to^{\mathbb{P}} \nu_0.$$

In particular, under the null hypothesis $\eta_0 = 0$, both estimators are consistent as we have

$$\nu_0 = (\overline{\sigma}_0^2, 0, \overline{\sigma}_0^2, 0, 0).$$

Proof. We show separately the consistency of $\tilde{\zeta}_{n,err}$ and $\tilde{w}_{n,exp}$. Let us start with $\tilde{\zeta}_{n,err}$. The methodology for the proof is the same as those of Lemma C.9 and Theorem C.10. First, let us define

$$\tilde{\rho}_k := \frac{1}{2} \mathbb{E} \left[\frac{\partial W_0(\theta_0)}{\partial \theta} \frac{\partial W_k(\theta_0)^T}{\partial \theta} + \frac{\partial W_k(\theta_0)}{\partial \theta} \frac{\partial W_0(\theta_0)^T}{\partial \theta} \right]. \tag{C.92}$$

Lemmas C.6 and C.7 are now easily adapted to

$$\sup_{\zeta \in \Xi_{n,err}} \mathbb{E}_{\mathcal{U}} \left[\left(\frac{1}{N_n} \frac{\partial^{\boldsymbol{\alpha}} \mathbf{G}_n(\zeta)}{\partial \zeta^{\boldsymbol{\alpha}}} - \frac{\partial^{\boldsymbol{\alpha}} \mathbf{G}_{\infty,err}(\zeta)}{\partial \zeta^{\boldsymbol{\alpha}}} \right)^2 \right] = o_{\mathbb{P}}(N_n^{-1/2}), \tag{C.93}$$

for any multi-index $\boldsymbol{\alpha}$ such that $|\boldsymbol{\alpha}| \leq m$, and with

$$\mathbf{G}_{\infty,err}(\zeta) := \frac{2}{T\gamma^2(1+\phi)} u^T \left\{ \tilde{\rho}_0 - (1-\phi) \sum_{k=1}^{+\infty} \phi^{k-1} \tilde{\rho}_k \right\} u$$

$$= \frac{1}{\gamma^2 T} u^T P_{\theta_0,\phi} u$$

with $P_{\theta_0,\phi} := 2(1+\phi)^{-1} \{\tilde{\rho}_0 - (1-\phi) \sum_{k=1}^{+\infty} \phi^{k-1} \tilde{\rho}_k\}$, and

$$\sup_{\zeta \in \Xi_{n,err}} \mathbb{E}_{\mathcal{U}} \left[\left(\frac{1}{N_n} \frac{\partial^{\boldsymbol{\alpha}} \mathbf{K}_n(\zeta)}{\partial \zeta^{\boldsymbol{\alpha}}} \right)^2 \right] = o_{\mathbb{P}}(N_n^{-1/2}). \tag{C.94}$$

Therefore, adapting the reasoning from Aït-Sahalia and Xiu (2016, p. 45) to our setting (by (C.1) we have that the step size of the observation grid $\pi_T^n \to^{\mathbb{P}} 0$) and combining them with (C.93), (C.94) and Lemma C.8, we get the convergence

$$\sup_{\zeta \in \Xi_{n,err}} \left| \Psi_{n,err}(\zeta) - \Psi_{\infty,err}(\zeta) \right| \to^{\mathbb{P}} 0, \tag{C.95}$$

where

$$\Psi_{\infty,err}(\zeta) = \begin{pmatrix} \dfrac{1}{2\gamma^2} - \dfrac{\overline{\sigma}_0^2 + 2(1-\phi)\tilde{\eta}_0}{2\gamma^4(1-\phi^2)} + \dfrac{1}{2}\dfrac{\partial \mathbf{G}_{\infty,err}(\zeta)}{\partial\gamma^2} \\ \dfrac{1}{2}\dfrac{\partial \mathbf{G}_{\infty,err}(\zeta)}{\partial u} \\ \phi\dfrac{\overline{\sigma}_0^2 + 2(1-\phi)\tilde{\eta}_0}{\gamma^2(1-\phi^2)^2} - \dfrac{\tilde{\eta}_0}{\gamma^2(1-\phi^2)} + \dfrac{1}{2}\dfrac{\partial \mathbf{G}_{\infty,err}(\zeta)}{\partial\phi} \end{pmatrix}.$$

Now, by a classical statistical argument (see e.g Van der Vaart (2000, Theorem 5.9) along with (C.95), the consistency of $\widehat{\zeta}_{n,err}$ will be proved if for any $\epsilon > 0$, $\inf_{\zeta \in \Xi_{\infty,err}:|\zeta-\zeta_0|>\epsilon} \|\Psi_{\infty,err}(\zeta)\| > 0$ where $\|x\| = \sqrt{\sum_i x_i^2}$, and where $\Xi_{\infty,err} = \left[\underline{\gamma}^2, \overline{\gamma}^2\right] \times \mathbb{R}^d \times \left[-\overline{\phi}, \overline{\phi}\right]$, and if $\Psi_{\infty,err}(\zeta_0) = 0$. The second assertion is immediate. To prove the former, let us take $b > 0$ an arbitrary number and consider $\Xi_{b,err} := \left[\underline{\gamma}^2, \overline{\gamma}^2\right] \times [-b, b]^d \times \left[-\overline{\phi}, \overline{\phi}\right]$. We are going to show that $\inf_{\zeta \in \Xi_{\infty,err}-\Xi_{b,err}:|\zeta-\zeta_0|>\epsilon} \|\Psi_{\infty,err}(\zeta)\| > 0$ on the one hand, and $\inf_{\zeta \in \Xi_{b,err}:|\zeta-\zeta_0|>\epsilon} \|\Psi_{\infty,err}(\zeta)\| > 0$ on the other hand. In the first case, by hypothesis $u^T u \geq b^2$, and writing $\tilde{M}(\phi) = \tilde{\rho}_0 - (1-\phi)\sum_{k=1}^{+\infty}\phi^{k-1}\tilde{\rho}_k$, we automatically have that $\tilde{M}(\phi)$ is a symmetric positive matrix for any $\phi \in [-\overline{\phi}, \overline{\phi}]$ as a simple consequence of the Cauchy–Schwarz inequality and the stationarity of the information process. Thus, writing $\tilde{c}(\phi) > 0$ the minimal eigenvalue of $M(\phi)^2$ and $\tilde{c} = \min_{\phi \in [-\overline{\phi},\overline{\phi}]} \tilde{c}(\phi)$, we get that in the first case

$$\|\Psi_{\infty,err}(\zeta)\| > \frac{\tilde{c}}{T\overline{\gamma}^2(1+\overline{\phi})}b^2 > 0.$$

In the second case, since $\Xi_{b,err}$ is a compact space a continuity argument shows that it is sufficient to prove that $\Psi_{\infty,err}(\zeta) = 0$ if and only if $\zeta = \zeta_{0,err}$. Let thus $\zeta \in \Xi_{b,err}$ such that the score at point ζ is null. Given the shape of $\mathbf{G}_{\infty,err}$ and the positivity of $M(\phi)^2$ for any ϕ, $\Psi^2_{\infty,err}(\zeta) = 0$ yields $u = 0$. Then

$$0 = \Psi^3_{\infty,err}(\zeta) = \phi\frac{\overline{\sigma}_0^2 + 2(1-\phi)\tilde{\eta}_0}{\gamma^2(1-\phi^2)^2} - \frac{\tilde{\eta}_0}{\gamma^2(1-\phi^2)}$$

yields the second order equation $\tilde{\eta}_0\phi^2 - (\overline{\sigma}_0^2 + 2\tilde{\eta}_0)\phi + \tilde{\eta}_0 = 0$, which in turn implies $\phi = \phi_0$ (the other root being non-admissible). Finally

$$0 = \Psi^1_{\infty,err}(\zeta) = \frac{1}{2\gamma^2} - \frac{\overline{\sigma}_0^2 + 2(1-\phi)\tilde{\eta}_0}{2\gamma^4(1-\phi^2)}$$

gives $\gamma^2 = \gamma_0^2$ replacing ϕ_0 by its expression, and thus $\zeta = \zeta_{0,err}$. In particular, when $\eta_0 = \tilde{\eta}_0 = 0$, this gives $\phi_0 = 0$ and $\gamma_0^2 = \overline{\sigma}_0^2$.

We now derive the limit of $\widehat{w}_{n,exp}$. By the relation (C.88) we also immediately deduce that

$$\sup_{\zeta \in \Xi_{n,exp}} \left|\Psi_{n,exp}(w) - \Psi_{\infty,exp}(w)\right| \to^{\mathbb{P}} 0,$$

where

$$\Psi_{\infty,exp}(w) = \begin{pmatrix} \dfrac{1}{2\sigma^2} - \dfrac{\overline{\sigma}_0^2 + 2\tilde{\eta}_0}{2\sigma^4} + \dfrac{1}{2}\dfrac{\partial \mathbf{G}_{\infty,exp}(w)}{\partial\sigma^2} \\ \dfrac{1}{2}\dfrac{\partial \mathbf{G}_{\infty,exp}(w)}{\partial u} \end{pmatrix},$$

and

$$\mathbf{G}_{\infty,exp}(w) = \frac{2}{T\sigma^2}\left\{u^T\tilde{\rho}_0 u - u^T\tilde{\rho}_1 u\right\} = \frac{1}{T\sigma^2}u^T U_{\theta_0}u.$$

From there, a similar reasoning to the case applied to $\widehat{\zeta}_{n,err}$ yields the convergence in probability of $\widehat{w}_{n,exp}$ towards $(\overline{\sigma}_0^2 + 2\tilde{\eta}_0, 0)$. In particular, the volatility component is consistent under the null hypothesis $\eta_0 = 0$ and inconsistent otherwise whereas the information estimator is consistent in both situations. $\square$

We now introduce, as in the previous section,

$$H_{n,err}(\zeta) = -N_n^{-1}\frac{\partial^2 \mathcal{L}_{n,err}(\zeta)}{\partial\zeta^2}, \tag{C.96}$$

and

$$H_{n,exp}(w) = -N_n^{-1}\frac{\partial^2 \mathcal{L}_{n,exp}(w)}{\partial w^2}. \tag{C.97}$$

Lemma C.15 (*Fisher Information*). *Let $\Gamma_{err}(\zeta_{0,err})$ and $\Gamma_{exp}(w_{0,exp})$ be the matrices*

$$\Gamma_{err}(\zeta_{0,err}) = \begin{pmatrix} \frac{1}{2}\gamma_0^{-4} & 0 & 0 \\ 0 & \gamma_0^{-2}T^{-1}P_{\theta_0} & 0 \\ 0 & 0 & (1-\phi_0^2)^{-1} \end{pmatrix} \tag{C.98}$$

and

$$\Gamma_{exp}(w_{0,exp}) = \begin{pmatrix} \frac{1}{2}(\overline{\sigma}_0^2 + 2\tilde{\eta}_0)^{-2} & 0 \\ 0 & (\overline{\sigma}_0^2 + 2\tilde{\eta}_0)^{-1}T^{-1}U_{\theta_0} \end{pmatrix}, \tag{C.99}$$

where $P_{\theta_0} := P_{\theta_0,\phi_0}$. We have, for any balls $V_{n,err}$ and $V_{n,exp}$ respectively centered on $\zeta_{0,err}$ and $w_{0,exp}$ and shrinking to $\{\zeta_{0,err}\}$ and $\{w_{0,err}\}$,

$$\sup_{\zeta_n \in V_{n,err}} \left\| H_{n,err}(\zeta_n) - \Gamma_{err}(\zeta_{0,err}) \right\| \to^{\mathbb{P}} 0 \tag{C.100}$$

and

$$\sup_{w_n \in V_{n,exp}} \left\| H_{n,exp}(w_n) - \Gamma_{exp}(w_{0,exp}) \right\| \to^{\mathbb{P}} 0. \tag{C.101}$$

Proof. Adapting the reasoning to get the third and fourth equations (p. 45) in Aït-Sahalia and Xiu (2016) (replacing $a_0^2 \Delta_n^{\gamma_0-1}$ by $\tilde{\eta}_0$), taking the first derivatives of the score functions, and using (C.93) and (C.94) for the information part, we directly obtain the convergence for any $b > 0$

$$\sup_{\zeta \in \Xi_{b,err}} \left\| H_{n,err}(\zeta) - H_{\infty,err}(\zeta) \right\| \to^{\mathbb{P}} 0, \tag{C.102}$$

where

$$H_{\infty,err}(\zeta) := \begin{pmatrix} -\frac{1}{2\gamma^4} + \frac{\overline{\sigma}_0^2 + 2(1-\phi)\tilde{\eta}_0}{\gamma^6(1-\phi^2)} & 0 & -\frac{\overline{\sigma}_0^2\phi - \tilde{\eta}_0(\phi-1)^2}{\gamma^4(1-\phi^2)^2} \\ 0 & 0 & 0 \\ -\frac{\overline{\sigma}_0^2\phi - \tilde{\eta}_0(\phi-1)^2}{\gamma^4(1-\phi^2)^2} & 0 & \frac{1}{\gamma^2}\left\{ \frac{\overline{\sigma}_0^2 + 2(1-3\phi)\tilde{\eta}_0}{(1-\phi^2)^2} + \frac{4\phi^2\left(\overline{\sigma}_0^2 + 2(1-\phi)\tilde{\eta}_0\right)}{(1-\phi^2)^3} \right\} \end{pmatrix}$$
$$+ \frac{1}{2}\frac{\partial^2 \mathbf{G}_{\infty,err}(\zeta)}{\partial\zeta^2}.$$

By continuity of $H_{\infty,err}$, we immediately deduce that

$$\sup_{\zeta_n \in V_{n,err}} \left\| H_{n,err}(\zeta_n) - H_{\infty,err}(\zeta_{0,err}) \right\| \to^{\mathbb{P}} 0, \tag{C.103}$$

and moreover from the definition of $H_{\infty,err}(\zeta)$ we have

$$H_{\infty,err}(\zeta_{0,err}) = \Gamma_{err}(\zeta_{0,err}), \tag{C.104}$$

since

$$\frac{1}{2}\frac{\partial^2 \mathbf{G}_{\infty,err}(\zeta_{0,err})}{\partial\zeta^2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \gamma_0^{-2}T^{-1}P_{\theta_0} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and using the relations $\overline{\sigma}_0^2 = \gamma_0^2(1-\phi_0)^2$ and $\tilde{\eta}_0 = \gamma_0^2\phi_0$. Convergence (C.101) is proved in the same way. $\square$

Let $\alpha^{i,j} = \phi^{|i-j|} - \phi^{i+j} - \phi^{2N_n+2-(i+j)} - \phi^{2N_n+2-|i-j|}$ and $\beta^{i,j} = \partial\alpha^{i,j}/\partial\phi$. We define, similarly to (C.68)–(C.72) for $t \in [0,T]$, the martingales

$$S_1(t) := \sum_{i=1}^{N_n(t)} \alpha^{i,i} \left\{ \left(\Delta X_{i,t}^n\right)^2 - \int_{t_{i-1}^n \wedge t}^{t_i^n \wedge t} \sigma_s^2 ds - \sum_{t_{i-1}^n \wedge t < s \le t_i^n \wedge t} \Delta J_s^2 \right\},$$

$$S_2^{(\mathbf{a})}(t) := \sum_{i=1}^{N_n(t)} \left\{ \sum_{1 \le j < i} \mathbf{a}^{i,j} \Delta X_{j,t}^n \right\} \Delta X_{i,t}^n, \quad \mathbf{a} \in \{\alpha, \beta\},$$

$$S_3^{(\mathbf{a})}(t) := -2 \sum_{i=0}^{N_n(t)} \left\{ \sum_{j=1}^{N_n(t)} \dot{\mathbf{a}}^{i,j} \Delta X_{j,t}^n \right\} \epsilon_{t_i^n}, \, \mathbf{a} \in \{\alpha, \beta\},$$

$$S_4^{(\mathbf{a})}(t) := \sum_{i=0}^{N_n(t)} \ddot{\mathbf{a}}^{i,i} \left\{ \epsilon_{t_i^n}^2 - n^{-1}\eta_0 \right\} + 2 \sum_{i=0}^{N_n(t)} \left\{ \sum_{0 \le j < i} \ddot{\mathbf{a}}^{i,j} \epsilon_{t_j^n} \right\} \epsilon_{t_i^n}, \, \mathbf{a} \in \{\alpha, \beta\},$$

$$S_5(t) := -N_n^{-1/2} \sum_{j=1}^{N_n(t)} \left\{ \sum_{i=0}^{N_n(t)} \dot{\alpha}^{i,j} \frac{\partial W_i(\theta_0)}{\partial \theta} \right\} \Delta X_{j,t}^n + N_n^{-1/2} \sum_{i=0}^{N_n(t)} \sum_{j=0}^{N_n(t)} \ddot{\alpha}^{i,j} \frac{\partial W_j(\theta_0)}{\partial \theta} \epsilon_{t_i^n},$$

and note that, up to exponentially negligible terms, we have the representation

$$\Psi_n(\nu_0) = \begin{pmatrix} \frac{1}{2\gamma_0^4(1+\phi_0^2)^2 T}(S_1(T) + 2S_2^{(\alpha)}(T) + S_3^{(\alpha)}(T) + S_4^{(\alpha)}(T))_{|\phi=0} \\ \frac{1}{\gamma_0^2(1+\phi_0^2)T} S_5(T)_{|\phi=0} \\ \frac{1}{2\gamma_0^4(1-\phi_0^2)T}(S_1(T) + 2S_2^{(\alpha)}(T) + S_3^{(\alpha)}(T) + S_4^{(\alpha)}(T))_{|\phi=\phi_0} \\ \frac{1}{\gamma_0^2(1-\phi_0^2)T} S_5(T)_{|\phi=\phi_0} \\ -\frac{1}{2\gamma_0^2(1-\phi_0^2)T} \left(\frac{2\phi_0}{1-\phi_0^2} \left\{ S_1(T) + 2S_2^{(\alpha)}(T) + S_3^{(\alpha)}(T) + S_4^{(\alpha)}(T) \right\} + 2S_2^{(\beta)}(T) + S_3^{(\beta)}(T) + S_4^{(\beta)}(T) \right)_{|\phi=\phi_0} \end{pmatrix}.$$

Lemma C.16. *Let $\phi \in]-1, 1[$. We have, $\mathcal{G}_T$-stably in law, that*

$$N_n^{1/2} S_1(T) \to \mathcal{MN}\left(0, 2T\mathcal{Q}\right).$$

Proof. Note that $\alpha^{i,i} = 1 - \phi^{2i} - \phi^{2N_n+2-2i} - \phi^{2N_n+2}$. Since $|\phi| < 1$, standard calculations involving Burkholder–Davis–Gundy inequalities, assumption **(H)** and the finite activity property of the jumps easily yield

$$N_n^{1/2} S_1(T) = N_n^{1/2} \widetilde{S}_1(T) + o_{\mathbb{P}}(1),$$

where

$$\widetilde{S}_1(T) = \sum_{i=1}^{N_n(t)} \left\{ \left(\Delta X_{i,t}^n\right)^2 - \int_{t_{i-1}^n \wedge t}^{t_i^n \wedge t} \sigma_s^2 ds - \sum_{t_{i-1}^n \wedge t < s \le t_i^n \wedge t} \Delta J_s^2 \right\}.$$

Now, by a straightforward adaptation of Lemma 1 (in the case $q = 0$) in the proof of Theorem 2 Appendix A.3 p.40 in Aït-Sahalia and Xiu (2016) in the case of an irregular grid of the form (2.4), and following the same line of reasoning as the proof of Lemma A.7 in Clinet and Potiron (2018a), we conclude that $\mathcal{G}_T$-stably in law,

$$N_n^{1/2} \widetilde{S}_1(T) \to \mathcal{MN}\left(0, 2T\mathcal{Q}\right),$$

and we are done. $\square$

Lemma C.17. *Let $\phi \in]-1, 1[$ and define $S_2(T) = (S_2^{(\alpha)}(T), S_2^{(\beta)}(T))$. We have, $\mathcal{G}_T$-stably in law, that*

$$N_n^{1/2} S_2(T) \to \mathcal{MN}\left(0, T\mathcal{Q} \begin{pmatrix} \frac{\phi^2}{1-\phi^2} & \frac{\phi}{(1-\phi^2)^2} \\ \frac{\phi}{(1-\phi^2)^2} & \frac{1+\phi^2}{(1-\phi^2)^3} \end{pmatrix}\right).$$

Proof. Again, given the shapes of α and β, by standard calculations on martingale increments, introducing $\widetilde{\alpha}^{i,j} = \phi^{|i-j|}$ and $\widetilde{\beta}^{i,j} = |i - j|\phi^{|i-j|-1}$, we easily have

$$N_n^{1/2} S_2^{(\alpha)}(T) = N_n^{1/2} \widetilde{S}_2^{(\alpha)}(T) + o_{\mathbb{P}}(1)$$

where

$$\widetilde{S}_2^{(\alpha)}(T) = \sum_{i=1}^{N_n(t)} \left\{ \sum_{1 \le j < i} \widetilde{\alpha}^{i,j} \Delta X_{j,t}^n \right\} \Delta X_{i,t}^n$$

and a similar statement for $S_2^{(\beta)}(T)$. Proving a central limit theorem for $\widetilde{S}_2(T)$ now boils down to following exactly the same calculations as for $M_2^{(\sigma^2)}(T)$ in the large noise case (see the proof of Lemma A.7 in Clinet and Potiron (2018a)) but replacing $\frac{\partial \omega^{i,j}}{\partial \sigma^2}$ by $\widetilde{\alpha}^{i,j} = \phi^{|i-j|}$ and $\widetilde{\beta}^{i,j} = |i-j|\phi^{|i-j|-1}$. In particular, a careful inspection of the proof shows that all the calculations remain valid replacing the scalar $\frac{5}{64T^{3/2}\sigma^7 a_0} = \lim_n N_n^{-3/2} \Delta_n \sum_{j=1}^{N_n} \left(\frac{\partial \omega^{N_n,j}}{\partial \sigma^2} \right)^2$ in the expression of the asymptotic variance, by the 2×2 matrix

$$
\lim_n T \begin{pmatrix} \sum_{j=1}^{N_n-1} \left(\widetilde{\alpha}^{N_n,j} \right)^2 & \sum_{j=1}^{N_n-1} \widetilde{\alpha}^{N_n,j}\widetilde{\beta}^{N_n,j} \\ \sum_{j=1}^{N_n-1} \widetilde{\alpha}^{N_n,j}\widetilde{\beta}^{N_n,j} & \sum_{j=1}^{N_n-1} \left(\widetilde{\beta}^{N_n,j} \right)^2 \end{pmatrix} = T \begin{pmatrix} \sum_{k=1}^{+\infty} \phi^{2k} & \sum_{k=1}^{+\infty} k\phi^{2k-1} \\ \sum_{k=1}^{+\infty} k\phi^{2k-1} & \sum_{k=1}^{+\infty} k^2 \phi^{2k-2} \end{pmatrix}
$$

$$
= T \begin{pmatrix} \dfrac{\phi^2}{1-\phi^2} & \dfrac{\phi}{(1-\phi^2)^2} \\ \dfrac{\phi}{(1-\phi^2)^2} & \dfrac{1+\phi^2}{(1-\phi^2)^3} \end{pmatrix},
$$

which yields the $\mathcal{G}_T$-stable convergence in distribution

$$
N_n^{1/2}\widetilde{S}_2(T) \to \mathcal{MN}\left(0, T\mathcal{Q} \begin{pmatrix} \dfrac{\phi^2}{1-\phi^2} & \dfrac{\phi}{(1-\phi^2)^2} \\ \dfrac{\phi}{(1-\phi^2)^2} & \dfrac{1+\phi^2}{(1-\phi^2)^3} \end{pmatrix} \right). \quad \square
$$

Lemma C.18. *Let $\phi \in]-1, 1[$ and define $S_3(T) = (S_3^{(\alpha)}(T), S_3^{(\beta)}(T))$. We have, conditioned on $\mathcal{G}_T$ the convergence in distribution*

$$
N_n^{1/2}S_3(T) \to \mathcal{MN}\left(0, 4\overline{\sigma}_0^2\tilde{\eta}_0 T^2 \begin{pmatrix} \dfrac{2(1-\phi)^2}{1-\phi^2} & \dfrac{-2(1-\phi)^2}{(1-\phi^2)^2} \\ \dfrac{-2(1-\phi)^2}{(1-\phi^2)^2} & \dfrac{4(1-\phi)^2}{(1-\phi^2)^3} \end{pmatrix} \right).
$$

Proof. As in the previous lemma, we introduce the coefficients $\widetilde{\alpha}^{i,j} = \phi^{|i-j|}$ and $\widetilde{\beta}^{i,j} = |i-j|\phi^{|i-j|-1}$. Note that $\overset{\approx}{\alpha}{}^{i,j} = \phi^{|i-j|}(\phi-1)$ for $i \geq j$ and $\overset{\approx}{\alpha}{}^{i,j} = \phi^{|i-j|}(\phi^{-1}-1)$ for $j \geq i+1$. Moreover, if $i \geq j$, $\overset{\approx}{\beta}{}^{i,j} = \phi^{|i-j|} + |i-j|\phi^{|i-j|-1}(\phi-1)$ and if $j \geq i+1$, $\overset{\approx}{\beta}{}^{i,j} = -\phi^{|j-i|-2} + |j-i|\phi^{|j-i|-2}(1-\phi)$. Given the exponential shape of the coefficients, we easily show as for the previous lemma that

$$
N_n^{1/2}S_3^{(\alpha)}(T) = N_n^{1/2}\widetilde{S}_3^{(\alpha)}(T) + o_\mathbb{P}(1)
$$

where

$$
\widetilde{S}_3^{(\alpha)}(T) = -2 \sum_{i=0}^{N_n(t)} \left\{ \sum_{j=1}^{N_n(t)} \overset{\approx}{\alpha}{}^{i,j} \Delta X_{j,t}^n \right\} \epsilon_{t_i^n}
$$

and a similar definition for β. Now, as for $S_2(T)$, we adapt the proof of $M_3^{(\sigma^2)}(T)$ from the large noise case (proof of Lemma A.9 in Clinet and Potiron (2018a)). Again, all the calculations remain valid except that now $\frac{\partial \dot{\omega}^{i,j}}{\partial \sigma^2}$ should be replaced by $\overset{\approx}{\alpha}{}^{i,j}$ and $\overset{\approx}{\beta}{}^{i,j}$, and accordingly, in the limiting variance the scalar $\frac{1}{8\sqrt{T}\sigma^5 a_0} = \lim_n a_0^2 N_n^{-1/2} \sum_{j=1}^{N_n} \left(\frac{\partial \dot{\omega}^{N_n/2,j}}{\partial \sigma^2} \right)^2$ is replaced by the 2×2 matrix

$$
\lim_n \frac{4\eta_0 T N_n}{n} \begin{pmatrix} \sum_{j=1}^{N_n-1} \left(\overset{\approx}{\alpha}{}^{N_n/2,j} \right)^2 & \sum_{j=1}^{N_n-1} \overset{\approx}{\alpha}{}^{N_n/2,j}\overset{\approx}{\beta}{}^{N_n/2,j} \\ \sum_{j=1}^{N_n-1} \overset{\approx}{\alpha}{}^{N_n/2,j}\overset{\approx}{\beta}{}^{N_n/2,j} & \sum_{j=1}^{N_n-1} \left(\overset{\approx}{\beta}{}^{N_n/2,j} \right)^2 \end{pmatrix} = 4\tilde{\eta}_0 T \begin{pmatrix} \dfrac{2(1-\phi)^2}{1-\phi^2} & \dfrac{-2(1-\phi)^2}{(1-\phi^2)^2} \\ \dfrac{-2(1-\phi)^2}{(1-\phi^2)^2} & \dfrac{4(1-\phi)^2}{(1-\phi^2)^3} \end{pmatrix},
$$

where the last step is obtained by direct calculation on the coefficients, and because $\eta_0 T N_n/n \to^\mathbb{P} \tilde{\eta}_0 T$ by definition of $\tilde{\eta}_0$. $\square$

Lemma C.19. *Let $\phi \in \,]-1, 1[$, and define $S_4(T) = (S_4^{(\alpha)}(T), S_4^{(\beta)}(T))$. We have, conditioned on $\mathcal{G}_T$ the convergence in distribution*

$$N_n^{1/2} S_4(T) \to \mathcal{N}\left(0, 4\left(\widetilde{\mathcal{K}} + 2\tilde{\eta}_0^2\right) T^2 \begin{pmatrix} (1-\phi)^2 & -(1-\phi) \\ -(1-\phi) & 1 \end{pmatrix} + 4\tilde{\eta}_0^2 T^2 (1-\phi)^4 \begin{pmatrix} \dfrac{1}{1-\phi^2} & -\dfrac{(\phi+2)}{(1-\phi^2)^2} \\ -\dfrac{(\phi+2)}{(1-\phi^2)^2} & \dfrac{\phi^2+4\phi+5}{(1-\phi^2)^3} \end{pmatrix}\right).$$

Proof. As for the previous lemmas, introducing $\tilde{\alpha}^{i,j} = \phi^{|i-j|}$ and $\tilde{\beta}^{i,j} = |i-j|\phi^{|i-j|-1}$, we have by standard calculation the approximation

$$N_n^{1/2} S_4^{(\alpha)}(T) = N_n^{1/2} \widetilde{S}_4^{(\alpha)}(T) + o_{\mathbb{P}}(1)$$

where

$$\widetilde{S}_4^{(\alpha)}(t) = \underbrace{\sum_{i=0}^{N_n(t)} \overset{\approx}{\alpha}{}^{i,i} \left\{ \epsilon_{t_i^n}^2 - n^{-1}\eta_0 T \right\}}_{U^{(\alpha)}(t)} + \underbrace{2\sum_{i=0}^{N_n(t)} \left\{ \sum_{0 \le j < i} \overset{\approx}{\alpha}{}^{i,j} \epsilon_{t_j^n} \right\} \epsilon_{t_i^n}}_{V^{(\alpha)}(t)},$$

and a similar statement for $S_4^{(\beta)}(T)$. Moreover, we have $\overset{\approx}{\alpha}{}^{i,j} = -(1-\phi)^2 \phi^{|i-j|-1}$ for $i \ne j$, and $\overset{\approx}{\alpha}{}^{i,j} = 2(1-\phi)$ for $i = j$. Similarly, we have $\overset{\approx}{\beta}{}^{i,j} = (1-\phi^2)\phi^{|i-j|-2} - |i-j|(1-\phi)^2\phi^{|i-j|-2}$ for $i \ne j$ and $\overset{\approx}{\beta}{}^{i,j} = -2$ for $i = j$. Now, defining $U = (U^{(\alpha)}, U^{(\beta)})$ and $V = (V^{(\alpha)}, V^{(\beta)})$, we have that U and V are uncorrelated sums of martingale increments so that it is sufficient to prove that conditionally on $\mathcal{G}_T$

$$N_n^{1/2} U(T) \to \mathcal{MN}\left(0, 4\left(\widetilde{\mathcal{K}} + 2\tilde{\eta}_0^2\right) T^2 \begin{pmatrix} (1-\phi)^2 & -(1-\phi) \\ -(1-\phi) & 1 \end{pmatrix}\right)$$

and

$$N_n^{1/2} V(T) \to \mathcal{MN}\left(0, 4\tilde{\eta}_0^2 T^2 (1-\phi)^4 \begin{pmatrix} \dfrac{1}{1-\phi^2} & -\dfrac{(\phi+2)}{(1-\phi^2)^2} \\ -\dfrac{(\phi+2)}{(1-\phi^2)^2} & \dfrac{\phi^2+4\phi+5}{(1-\phi^2)^3} \end{pmatrix}\right).$$

The first limit is an immediate consequence of the fact that $U(T)$ is a sum of centered independent and identically distributed variables and the fact that $\overset{\approx}{\alpha}{}^{i,i} = 2(1-\phi)$ and $\overset{\approx}{\beta}{}^{i,i} = -2$, and that ϵ admits a finite fourth order moment. As for $V(T)$ a similar argument to that of $\widetilde{S}_2^{(\alpha)}$ in the proof of Lemma C.17 yields the convergence in distribution of $N_n^{1/2} V(T)$ to a normal limit with variance matrix

$$\lim_n 4\tilde{\eta}_0^2 T^2 \begin{pmatrix} \sum_{j=0}^{N_n-1} \left(\overset{\approx}{\alpha}{}^{N_n,j}\right)^2 & \sum_{j=0}^{N_n-1} \overset{\approx}{\alpha}{}^{N_n,j} \overset{\approx}{\beta}{}^{N_n,j} \\ \sum_{j=0}^{N_n-1} \overset{\approx}{\alpha}{}^{N_n,j} \overset{\approx}{\beta}{}^{N_n,j} & \sum_{j=0}^{N_n-1} \left(\overset{\approx}{\beta}{}^{N_n,j}\right)^2 \end{pmatrix} = 4\tilde{\eta}_0^2 T^2 \begin{pmatrix} \dfrac{(1-\phi)^4}{1-\phi^2} & -\dfrac{(1-\phi)^4(\phi+2)}{(1-\phi^2)^2} \\ -\dfrac{(1-\phi)^4(\phi+2)}{(1-\phi^2)^2} & \dfrac{(1-\phi)^4(\phi^2+4\phi+5)}{(1-\phi^2)^3} \end{pmatrix}$$

by direct calculation on the coefficients. Finally, the convergences are both conditional on $\mathcal{G}_T$ because the process ϵ is independent of $\mathcal{G}_T$. $\quad\square$

Lemma C.20. *Let $\phi \in \,]-1, 1[$. We have the stable convergence in distribution*

$$N_n^{1/2}\left(S_5(T) - N_n^{-1/2} B_{\theta_0,\phi}\right) \to \mathcal{MN}\left(0, A_{\theta_0,\phi}\right)$$

where $B_{\theta_0,\phi} = \sum_{0 < s \le T} \sum_{k=1}^{N_n} \phi^{|k-i_n(s)|} \dfrac{\partial \mu_k(\theta_0)}{\partial \theta} \Delta J_s$, and $i_n(s)$ is the only index such that $t_{i-1}^n < t \le t_i^n$, and

$$\begin{aligned} A_{\theta_0,\phi} &= 2\int_0^T \sigma_s^2 ds (1-\phi)^2 \left(\dfrac{\tilde{\rho}_0}{1-\phi^2} + \sum_{k=1}^{+\infty}\left\{\dfrac{2\phi^k}{1-\phi^2} - k\phi^{k-1}\right\}\tilde{\rho}_k\right) \\ &\quad + 2\tilde{\eta}_0 T(1-\phi)^3 \left(\dfrac{\phi+3}{1-\phi^2}\tilde{\rho}_0 + \sum_{k=1}^{+\infty}\left\{\dfrac{2(1-\phi)\phi^k}{1-\phi^2} - 4\phi^{k-1} + (k-1)(1-\phi)\phi^{k-2}\right\}\tilde{\rho}_k\right). \end{aligned}$$

Proof. We apply the same line of reasoning as for $S_1, \ldots, S_4$. Let $\tilde{\alpha}^{i,j} = \phi^{|i-j|}$ and $\tilde{\beta}^{i,j} = |i-j|\phi^{|i-j|-1}$. By standard moment calculation, we have

$$N_n^{1/2} S_5(T) = N_n^{1/2} \widetilde{S}_5(T) + o_{\mathbb{P}}(1)$$

where

$$
\widetilde{S}_5(t) = \underbrace{-N_n^{-1/2} \sum_{j=1}^{N_n(t)} \left\{ \sum_{i=0}^{N_n(t)} \dot{\tilde{\alpha}}^{i,j} \frac{\partial W_i(\theta_0)}{\partial \theta} \right\} \Delta X_{j,t}^n}_{U(t)} + \underbrace{N_n^{-1/2} \sum_{i=0}^{N_n(t)} \sum_{j=0}^{N_n(t)} \dot{\tilde{\alpha}}^{i,j} \frac{\partial W_j(\theta_0)}{\partial \theta} \epsilon_{t_i^n}}_{V(t)} .
$$

Let us assume for now that there are no jumps in the price, i.e. $J = 0$. U and V being uncorrelated sums of martingale increments, all we have to prove is that we have the $\mathcal{G}_T$-stable marginal convergences in distribution

$$
N_n^{1/2} U(T) \to \mathcal{MN} \left(0, 2 \int_0^T \sigma_s^2 ds (1 - \phi)^2 \left(\frac{\tilde{\rho}_0}{1 - \phi^2} + \sum_{k=1}^{+\infty} \left\{ \frac{2\phi^k}{1 - \phi^2} - k\phi^{k-1} \right\} \tilde{\rho}_k \right) \right)
$$

and

$$
N_n^{1/2} V(T) \to \mathcal{N} \left(0, 2\tilde{\eta}_0 T (1 - \phi)^3 \left(\frac{\phi + 3}{1 - \phi^2} \tilde{\rho}_0 + \sum_{k=1}^{+\infty} \left\{ \frac{2(1-\phi)\phi^k}{1 - \phi^2} - 4\phi^{k-1} + (k - 1)(1 - \phi)\phi^{k-2} \right\} \tilde{\rho}_k \right) \right).
$$

We start with U. In that case, we are going to apply Theorem 2–1 from Jacod (1997) to the continuous $\tilde{\mathcal{G}}_t$-martingale $N_n^{1/2} U$, where $\tilde{\mathcal{G}}_t := \mathcal{G}_t \vee \{Q_i^n, i, n \in \mathbb{N}\}$ (Note that X and W are still respectively an Itô process and a Brownian motion under $\tilde{\mathcal{G}}$ in view of the assumptions). Condition (2.8) is satisfied with $B = 0$. For condition (2.9), note that we have for any $t \in [0, T]$

$$
N_n \langle U, U \rangle_t = \sum_{j=1}^{N_n} \left(\sum_{i,i'=0}^{N_n} \dot{\tilde{\alpha}}^{i,j} \dot{\tilde{\alpha}}^{i',j} \frac{\partial W_i(\theta_0)}{\partial \theta} \frac{\partial W_{i'}(\theta_0)^T}{\partial \theta} \right) \int_{t_{j-1}^n \wedge t}^{t_j^n \wedge t} \sigma_s^2 ds,
$$

which, by similar calculations as for the proof of Lemma C.6 converges in probability to the limit

$$
C_t = \lim_n \left(\sum_{i,i'=0}^{N_n} \dot{\tilde{\alpha}}^{i,N_n/2} \dot{\tilde{\alpha}}^{i',N_n/2} \mathbb{E} \left[\frac{\partial W_i(\theta_0)}{\partial \theta} \frac{\partial W_{i'}(\theta_0)^T}{\partial \theta} \right] \right) \int_0^t \sigma_s^2 ds,
$$

$$
= \left(\frac{2(1 - \phi^2)}{1 - \phi^2} \tilde{\rho}_0 + 2(1 - \phi)^2 \sum_{k=1}^{+\infty} \left\{ \frac{2\phi^k}{1 - \phi^2} - k\phi^{k-1} \right\} \tilde{\rho}_k \right) \int_0^t \sigma_s^2 ds
$$

by direct calculation on the coefficients $\dot{\tilde{\alpha}}^{i,j}$, assumptions **(H)**, (2.11) and (2.12), and recalling that $\tilde{\rho}_k = \frac{1}{2}\mathbb{E}\left[\frac{\partial W_0(\theta_0)}{\partial \theta} \frac{\partial W_k(\theta_0)^T}{\partial \theta} + \frac{\partial W_k(\theta_0)}{\partial \theta} \frac{\partial W_0(\theta_0)^T}{\partial \theta} \right]$ for any $k \in \mathbb{N}$. Similarly, we also have Condition (2.10), i.e that $N_n^{1/2} \langle U, W \rangle_t \to^{\mathbb{P}} 0$. Finally Condition (2.11) comes from the continuity of U and Condition (2.12) is automatically satisfied since for any bounded martingale $\overline{N}$ orthogonal to W we have $\langle U, \overline{N} \rangle_t = 0$, which yields the $\tilde{\mathcal{G}}_T$ (and so $\mathcal{G}_T$) stable convergence. Now we turn to V. Note that proving the central limit theorem for $V(T)$ boils down to adapting the reasoning of the proof of $M_5(T)$ in Lemma C.12 replacing $\ddot{\omega}^{i,j}$ by $\dot{\tilde{\alpha}}^{i,j}$. A careful inspection of the proof shows that the calculation remains valid but the limiting variance is now expressed as

$$
\lim_n \frac{\eta_0 T}{n} N_n \sum_{i,i'=0}^{N_n} \dot{\tilde{\alpha}}^{i,N_n/2} \dot{\tilde{\alpha}}^{i',N_n/2} \mathbb{E} \left[\frac{\partial W_i(\theta_0)}{\partial \theta} \frac{\partial W_{i'}(\theta_0)^T}{\partial \theta} \right]
$$

$$
= 2\tilde{\eta}_0 T (1 - \phi)^3 \left(\frac{\phi + 3}{1 - \phi^2} \tilde{\rho}_0 + \sum_{k=1}^{+\infty} \left\{ \frac{2(1-\phi)\phi^k}{1 - \phi^2} - 4\phi^{k-1} + (k - 1)(1 - \phi)\phi^{k-2} \right\} \tilde{\rho}_k \right)
$$

by direct calculation on the coefficients $\dot{\tilde{\alpha}}^{i,N_n/2}$.

When $J \neq 0$, we immediately see that there is an additional term in $U(T)$ of the form

$$
- \sum_{j=1}^{N_n(t)} \left\{ \sum_{i=0}^{N_n(t)} \dot{\tilde{\alpha}}^{i,j} \frac{\partial W_i(\theta_0)}{\partial \theta} \right\} \Delta J_{j,t}^n
$$

which, by the finite activity property of J is equal to $- \sum_{j=1}^{N_J} \left\{ \sum_{i=0}^{N_n} \dot{\tilde{\alpha}}^{i,i(\tau_j)} \frac{\partial W_i(\theta_0)}{\partial \theta} \right\} \Delta J_{\tau_j}$ for n sufficiently large and where N_J is the random number of times J jumps on $[0, T]$, and $\tau_1, \ldots \tau_J$ are the associated jumping times. Now, note that by the by-part summation formula of Lemma C.2 and the definition of $\dot{\tilde{\alpha}}^{i,j}$, this term is equal to $B_{\theta_0,\phi}$ up to exponentially negligible terms and we are done. □

Proof of Theorem 3.1. Let us first derive the limit of $N_n^{1/2}(\tilde{\nu}_n - \nu_0)$. To do so, note that we have the martingale representation, up to exponentially negligible terms,

$$
\Psi_n(\nu_0) = \begin{pmatrix}
\dfrac{1}{2\gamma_0^4(1+\phi_0^2)^2 T}(S_1(T) + 2S_2^{(\alpha)}(T) + S_3^{(\alpha)}(T) + S_4^{(\alpha)}(T))_{|\phi=0} \\[2mm]
\dfrac{1}{\gamma_0^2(1+\phi_0^2)T}S_5(T)_{|\phi=0} \\[2mm]
\dfrac{1}{2\gamma_0^4(1-\phi_0^2)T}(S_1(T) + 2S_2^{(\alpha)}(T) + S_3^{(\alpha)}(T) + S_4^{(\alpha)}(T))_{|\phi=\phi_0} \\[2mm]
\dfrac{1}{\gamma_0^2(1-\phi_0^2)T}S_5(T)_{|\phi=\phi_0} \\[2mm]
-\dfrac{1}{2\gamma_0^2(1-\phi_0^2)T}\left(\dfrac{2\phi_0}{1-\phi_0^2}\left\{S_1(T) + 2S_2^{(\alpha)}(T) + S_3^{(\alpha)}(T) + S_4^{(\alpha)}(T)\right\} + 2S_2^{(\beta)}(T) + S_3^{(\beta)}(T) + S_4^{(\beta)}(T)\right)_{|\phi=\phi_0}
\end{pmatrix}.
$$

Moreover, the first order condition on Ψ_n yields

$$0 = \Psi_n(\tilde{\nu}_n) = \Psi_n(\nu_0) + H_n(\bar{\nu}_n)(\tilde{\nu}_n - \nu_0),\tag{C.105}$$

where $\bar{\nu}_n \in [\tilde{\nu}_n, \nu_0]$. We reformulate (C.105) as

$$\Gamma(\nu_0)^{-1}H_n(\bar{\nu}_n)N_n^{1/2}(\tilde{\nu}_n - \nu_0) = -\Gamma(\nu_0)^{-1}N_n^{1/2}\Psi_n(\nu_0).\tag{C.106}$$

Thus, by Lemma C.15, Slutsky's Lemma for stable convergence, the above martingale representation for $\Psi_n(\nu_0)$ along with lemmas C.16–C.20 (the brackets $\langle S_i, S_j \rangle_T$ for $i \neq j$ are all negligible so that we have the joint $\mathcal{G}_T$-stable convergence of the family $(S_i)_{i=1\ldots5}$), we deduce that $\mathcal{G}_T$-stably in law

$$N_n^{1/2}\left(\tilde{\nu}_n - \nu_0 - N_n^{-1/2}\widetilde{B}_{\theta_0}\right) \to \mathcal{MN}(0, \mathbf{U}),$$

where $\widetilde{B}_{\theta_0} = (0, U_{\theta_0}^{-1}B_{\theta_0}, 0, (1-\phi_0^2)^{-1}P_{\theta_0}^{-1}B_{\theta_0}, 0, 0)^T$, $\mathbf{U}$ is the matrix

$$
\begin{pmatrix}
\mathbf{U}_{11}(\mathcal{Q}, \overline{\sigma}_0^2, \tilde{\eta}_0, \widetilde{\mathcal{K}}) & 0 & \mathbf{U}_{13}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) & 0 & \mathbf{U}_{15}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) \\
0 & \left(\int_0^T \sigma_s^2 ds + 3\tilde{\eta}_0 T\right)U_{\theta_0}^{-1} & 0 & \mathbf{U}_{24}(\theta_0, \gamma_0^2, \phi_0, \sum_{0<s\leq T}\Delta J_s^2) & 0 \\
\mathbf{U}_{13}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) & 0 & \mathbf{U}_{33}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) & 0 & \mathbf{U}_{35}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) \\
0 & \mathbf{U}_{24}(\theta_0, \gamma_0^2, \phi_0, \sum_{0<s\leq T}\Delta J_s^2) & 0 & \mathbf{U}_{24}(\theta_0, \gamma_0^2, \phi_0, \sum_{0<s\leq T}\Delta J_s^2) & 0 \\
\mathbf{U}_{15}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) & 0 & \mathbf{U}_{35}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) & 0 & \mathbf{U}_{55}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}})
\end{pmatrix},
$$

with

$$\mathbf{U}_{11}(\mathcal{Q}, \overline{\sigma}_0^2, \tilde{\eta}_0, \widetilde{\mathcal{K}}) = \frac{2\mathcal{Q}}{T} + 4\widetilde{\mathcal{K}} + 12\tilde{\eta}_0^2 + 8\tilde{\eta}_0\overline{\sigma}_0^2,$$

$$\mathbf{U}_{13}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) = \frac{2\mathcal{Q}}{(1-\phi_0^2)T} + \frac{4\widetilde{\mathcal{K}}}{1+\phi_0} + \frac{4\gamma_0^4\phi_0(\phi_0^2 - \phi_0 + 2)}{1+\phi_0},$$

$$\mathbf{U}_{15}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) = -\frac{2\phi_0\mathcal{Q}}{\gamma_0^2(1-\phi_0^2)T} + \frac{2(1-\phi_0)^2\widetilde{\mathcal{K}}}{1-\phi_0^2} + \frac{4\gamma_0^2\phi_0(1-\phi_0)^2(\phi_0^2+1)}{1-\phi_0^2}$$

$$\mathbf{U}_{33}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) = \frac{2(1+\phi_0^2)\mathcal{Q}}{(1-\phi_0^2)^3 T} + \frac{4\widetilde{\mathcal{K}}}{(1+\phi_0)^2} + \frac{4\gamma_0^4\phi_0(\phi_0^2 + \phi_0 + 2)}{(1+\phi_0)^3},$$

$$\mathbf{U}_{35}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) = -\frac{2(2+\phi_0^2)\mathcal{Q}}{(1-\phi_0^2)^3\gamma_0^2 T} + \frac{2(1-\phi_0)\widetilde{\mathcal{K}}}{(1+\phi_0)^2\gamma_0^2} + \frac{2\gamma_0^2\phi_0(1-\phi_0)(\phi_0^2+2)}{(1+\phi_0)^3}$$

$$\mathbf{U}_{55}(\mathcal{Q}, \gamma_0^2, \phi_0, \widetilde{\mathcal{K}}) = \frac{(2\phi_0^4 + 7\phi_0^2 + 1)\mathcal{Q}}{(1-\phi_0^2)^3\gamma_0^4 T} + \frac{(1-\phi_0)^2\widetilde{\mathcal{K}}}{(1+\phi_0)^2\gamma_0^4} + \frac{\phi_0(4-\phi_0)(1-\phi_0)(\phi_0^2+1)}{(1+\phi_0)^3}$$

and for the terms involving θ_0,

$$
\begin{aligned}
\mathbf{U}_{24}(\theta_0, \gamma_0^2, \phi_0, \sum_{0<s\leq T}\Delta J_s^2) = {}& 2\gamma_0^2 T(1-\phi_0^2)^{-1}U_{\theta_0}^{-1}((1-\phi_0)\{(\phi_0^4 - 4\phi_0^3 + 5\phi_0^2 - \phi_0 + 1)\tilde{\rho}_0 \\
& + (\phi_0^3 - \phi_0^2 + 3\phi_0 - 1)\tilde{\rho}_1\} \\
& + 2\phi_0(1-\phi_0)^2\tilde{\rho}_2 + (2-\phi_0)(1-\phi_0)^4\sum_{k=2}^{+\infty}\phi_0^k\tilde{\rho}_k)P_{\theta_0}^{-1} \\
& - 2(1-\phi_0^2)^{-1}U_{\theta_0}^{-1}\left\{(1-\phi_0)^3\tilde{\rho}_0 - (1-\phi_0)^2\tilde{\rho}_1 + (1-\phi_0)^3\sum_{k=2}^{+\infty}\phi_0^k\tilde{\rho}_k\right\} \\
& \times P_{\theta_0}^{-1}\sum_{0<s\leq T}\Delta J_s^2,
\end{aligned}
$$

$$\mathbf{U}_{44}(\theta_0, \gamma_0^2, \phi_0, \sum_{0<s\leq T} \Delta J_s^2) = \gamma_0^2 T P_{\theta_0}^{-1} - \frac{2(1-\phi_0)^2}{(1-\phi_0^2)^2} P_{\theta_0}^{-1} \left(\frac{\tilde{\rho}_0}{1-\phi_0^2} + \sum_{k=1}^{+\infty} \left\{ \frac{2\phi_0^k}{1-\phi_0^2} - k\phi_0^{k-1} \right\} \tilde{\rho}_k \right)$$

$$\times P_{\theta_0}^{-1} \sum_{0<s\leq T} \Delta J_s^2 .$$

Finally, recalling that $(\widehat{\sigma}_{n,exp}^2, \widehat{\theta}_{n,exp}, \widehat{\sigma}_{n,err}^2, \widehat{\theta}_{n,err}, \widehat{a}_{n,err}^2)$ is equal to

$$(\widehat{\sigma}_{n,exp}^2, \theta_0 + N_n^{-1/2}\widehat{u}_{n,exp}, \widehat{\gamma}_{n,err}^2(1-\widehat{\phi}_{n,err})^2, \theta_0 + N_n^{-1/2}\widehat{u}_{n,err}, n^{-1}T\widehat{\gamma}_{n,err}^2\widehat{\phi}_{n,err}),$$

a straightforward application of the delta method yields Theorem 3.1. □

C.7. Proofs related to the test

Proof of Proposition 3.3. We start by showing our claim under $\mathcal{H}_0$. Note that the case $\widehat{V}_3$ is a consequence of Theorem 3.1. For $k \neq 3$, we conduct the proof for $\widehat{V}_k$ in several steps.

Step 1. We show that in the expressions of the variance estimators, we can replace the estimate returns $\Delta\widehat{X}_i^n$ by the efficient ones ΔX_i^n. For the sake of brevity we prove it in the case $k = 1$. Note that the cases $k = 2, 4, 5$ can be proved following the same line of reasoning. Introducing $\overline{V}_1 := \frac{4N_n}{T^2} \sum_{i=2}^{N_n-1} (\Delta X_i^n)^2 (\Delta X_{i-1}^n)^2$, we have to show that $\widehat{V}_1 - \overline{V}_1 \to^{\mathbb{P}} 0$. Since by definition $\widehat{X}_{t_i^n} = X_{t_i^n} + \phi(Q_i^n, \theta_0) - \phi(Q_i^n, \widehat{\theta}_{n,exp})$, if we introduce $b_i(\theta) = \mu_i(\theta_0) - \mu_i(\theta)$, we have the representation

$$\Delta\widehat{X}_i^n = \Delta X_i^n + b_i(\widehat{\theta}_{n,exp}). \tag{C.107}$$

Developing $\widehat{V}_1$ and using (C.107), we get

$$\widehat{V}_1 - \overline{V}_1 = A_{i,i-1}^n + A_{i-1,i}^n,$$

where

$$A_{i,i-1}^n = \frac{4N_n}{T^2} \sum_{i=2}^{N_n-1} \left\{ 2b_{i-1}(\widehat{\theta}_{n,exp})b_i(\widehat{\theta}_{n,exp})\Delta X_{i-1}^n \Delta X_i^n + 2b_i(\widehat{\theta}_{n,exp})^2 b_{i-1}(\widehat{\theta}_{n,exp})\Delta X_{i-1}^n \right.$$

$$\left. + 2b_i(\widehat{\theta}_{n,exp})\left(\Delta X_{i-1}^n\right)^2 \Delta X_i^n + b_i(\widehat{\theta}_{n,exp})^2 \left(\Delta X_{i-1}^n\right)^2 + \frac{1}{2}b_{i-1}(\widehat{\theta}_{n,exp})^2 b_i(\widehat{\theta}_{n,exp})^2 \right\},$$

$$= \sum_{j=1}^5 A_{i,i-1}^n[j],$$

and $A_{i-1,i}^n$ has the same expression as above inverting the role of i and $i-1$. Now, using the expansion

$$b_i(\widehat{\theta}_{n,exp}) = \frac{\partial \mu_i(\theta_0)}{\partial \theta}\left(\widehat{\theta}_{n,exp} - \theta_0\right) + \frac{1}{2}\frac{\partial^2 \mu_i(\tilde{\theta}_n)}{\partial \theta^2}\left(\widehat{\theta}_{n,exp} - \theta_0\right)^{\otimes 2}, \tag{C.108}$$

for some $\tilde{\theta}_n \in [\theta_0, \widehat{\theta}_{n,exp}]$, along with the fact that $\widehat{\theta}_{n,exp} - \theta_0 = O_{\mathbb{P}}(N_n^{-1})$ by Theorem 3.1 and that $\mathbb{E}\left[\sup_{\theta\in\Theta}\left|\left|\frac{\partial^j\mu_i(\theta)}{\partial\theta^j}\right|\right|^p X\right] < \infty$ independent of n and for any $p \geq 1, j \leq 2$, we easily deduce by direct calculation that each $A_{i,i-1}^n[j] = o_{\mathbb{P}}(1)$.

Step 2. Now we have to show that for any $k \in \{1, 2, 4, 5\}$, we have the convergence $\overline{V}_k \to^{\mathbb{P}} AVAR(\widehat{\sigma}_{exp}^2 - \widehat{\sigma}_{err}^2)$, where $\overline{V}_k$ has the same expression as $\widehat{V}_k$ except that each $\Delta\widehat{X}_i^n$ is replaced by the efficient return ΔX_i^n. The cases $k = 4, 5$ have already been tackled in Aït-Sahalia and Xiu (2016), so that it remains to show it for $k = 1, 2$. In this step we show the case $k = 1$, that is when there is no jump in the price process ($J = 0$), and when $AVAR(\widehat{\sigma}_{exp}^2 - \widehat{\sigma}_{err}^2) = 4T^{-2}\int_0^T \alpha_s^{-1}ds\int_0^T \sigma_s^4\alpha_s ds$. Let us introduce $\tilde{V}_1 = \frac{4N_n}{T^2}\sum_{i=2}^{N_n}\sigma_{t_{i-2}^n}^4 \Delta t_{i-1}^n \Delta t_i^n$. We first show that $\overline{V}_1 - \tilde{V}_1 = o_{\mathbb{P}}(1)$. Note that

$$\overline{V}_1 - \tilde{V}_1 = P_n^{(1)} + P_n^{(2)},$$

with

$$P_n^{(1)} = \frac{4N_n}{T^2}\sum_{i=2}^{N_n}\left(\Delta X_{i-1}^n\right)^2 \left\{\left(\Delta X_i^n\right)^2 - \sigma_{t_{i-2}^n}^2 \Delta t_i^n\right\},$$

and

$$P_n^{(2)} = \frac{4N_n}{T^2}\sum_{i=2}^{N_n}\left\{\left(\Delta X_{i-1}^n\right)^2 - \sigma_{t_{i-2}^n}^2 \Delta t_{i-1}^n\right\}\sigma_{t_{i-2}^n}^2 \Delta t_i^n.$$

We show that $P_n^{(1)} \to^{\mathbb{P}} 0$. Assume first that the volatility process has no jumps ($\tilde{J} = 0$). Note that $P_n^{(1)} = \sum_{i=2}^{N_n-1}\chi_i^n$ with $\chi_i^n = \frac{4N_n}{T^2}\left(\Delta X_{i-1}^n\right)^2\left\{\left(\Delta X_i^n\right)^2 - \sigma_{t_{i-2}^n}^2 \Delta t_i^n\right\}$, so that by Lemma 2.2.11 in Jacod and Shiryaev (2003), we only need to show

$\sum_{i=2}^{N_n-1} \mathbb{E}[\chi_i^n | \mathcal{G}_{i-1}^n] \to^{\mathbb{P}} 0$ on the one hand, and $\sum_{i=2}^{N_n-1} \mathbb{E}[(\chi_i^n)^2 | \mathcal{G}_{i-1}^n] \to^{\mathbb{P}} 0$ on the other hand. We have

$$\left| \sum_{i=2}^{N_n-1} \mathbb{E}[\chi_i^n | \mathcal{G}_{i-1}^n] \right| \leq \frac{4N_n}{T^2} \sum_{i=2}^{N_n-1} \left(\Delta X_{i-1}^n \right)^2 \mathbb{E}\left[\int_{t_{i-1}^n}^{t_i^n} \left| \sigma_s^2 - \sigma_{t_{i-2}}^2 \right| ds \, \bigg| \, \mathcal{G}_{i-1}^n \right]$$

$$\leq KN_n n^{-3/2+3/2\gamma} \underbrace{\sum_{i=2}^{N_n-1} \left(\Delta X_{i-1}^n \right)^2}_{O_{\mathbb{P}}(1)}$$

$$\to^{\mathbb{P}} 0,$$

since by Assumption **(H)** and (C.11) for the continuous Itô semi-martingale σ^2 we have

$$\mathbb{E}\left[\int_{t_{i-1}^n}^{t_i^n} \left| \sigma_s^2 - \sigma_{t_{i-2}}^2 \right| ds \, \bigg| \, \mathcal{G}_{i-1}^n \right] \leq \Delta t_i^n \mathbb{E}\left[\sup_{s \in [t_{i-1}^n, t_i^n]} \left| \sigma_s^2 - \sigma_{t_{i-2}}^2 \right| \, \bigg| \, \mathcal{G}_{i-1}^n \right] \leq Kn^{-3/2+3/2\gamma}.$$

Moreover,

$$\mathbb{E}_{\mathcal{U}}\left[\sum_{i=2}^{N_n-1} \mathbb{E}[(\chi_i^n)^2 | \mathcal{G}_{i-1}^n] \right] = \frac{16N_n^2}{T^4} \mathbb{E}_{\mathcal{U}}\left[\sum_{i=2}^{N_n-1} \left(\Delta X_{i-1}^n \right)^4 \mathbb{E}\left[\left\{ \left(\Delta X_i^n \right)^2 - \sigma_{t_{i-2}}^2 \Delta t_i^n \right\}^2 \, \bigg| \, \mathcal{G}_{i-1}^n \right] \right]$$

$$\leq KN_n^2 n^{-2+2\gamma} \sum_{i=2}^{N_n-1} \mathbb{E}_{\mathcal{U}} \left(\Delta X_{i-1}^n \right)^4$$

$$\leq KN_n^3 n^{-4+4\gamma} \to^{\mathbb{P}} 0,$$

where again we have used **(H)** and (C.13). Finally, when $\tilde{J} \neq 0$, by the finite activity property, only a finite number of terms are affected in the above sums, and it is easy to see that the convergence still holds in that case. Thus we have proved $P_n^{(1)} \to^{\mathbb{P}} 0$, and $P_n^{(2)} \to^{\mathbb{P}} 0$ is proved similarly. Now, recalling (2.4), we decompose

$$\tilde{V}_1 - 4T^{-2} \int_0^T \alpha_s^{-1} ds \int_0^T \sigma_s^4 \alpha_s ds = Q_n^{(1)} + Q_n^{(2)},$$

with

$$Q_n^{(1)} = \frac{4N_n}{T^2} \Delta_n \sum_{i=2}^{N_n-1} \sigma_{t_{i-2}}^4 \alpha_{t_{i-2}}^n \left\{ U_{i-1}^n - 1 \right\} \Delta t_i^n,$$

and

$$Q_n^{(2)} = \frac{4N_n}{T^2} \Delta_n \sum_{i=2}^{N_n-1} \sigma_{t_{i-2}}^4 \alpha_{t_{i-2}}^n \Delta t_i^n - 4T^{-2} \int_0^T \alpha_s^{-1} ds \int_0^T \sigma_s^4 \alpha_s ds.$$

Using assumption **(H)** and the fact that the U_i^n's are i.i.d, independent of the other quantities such that $\mathbb{E}[U_i^n] = 1$, we easily deduce that $\mathbb{E}_{\mathcal{U}} \left(Q_n^{(1)} \right)^2 \leq KN_n n^{-2+2\gamma} \to^{\mathbb{P}} 0$. Moreover, $Q_n^{(2)} \to^{\mathbb{P}} 0$ is a direct consequence of $N_n \Delta_n \to^{\mathbb{P}} \int_0^T \alpha_s^{-1} ds$, and the convergence of the Riemann sum $\sum_{i=2}^{N_n-1} \sigma_{t_{i-2}}^4 \alpha_{t_{i-2}}^n \Delta t_i^n \to^{\mathbb{P}} \int_0^T \sigma_s^4 \alpha_s ds$.

Step 3. Finally we show the case $k = 2$. We are going to show both convergences

$$A_n^{(1)} := \frac{4N_n}{T^2} \sum_{i=2}^{N_n-1} \left(\Delta X_i^n \right)^2 \left(\Delta X_{i-1}^n \right)^2 \mathbf{1}_{\{|\Delta X_i^n| \leq \tilde{u}_i\}} \mathbf{1}_{\{|\Delta X_{i-1}^n| \leq \tilde{u}_{i-1}\}} \to^{\mathbb{P}} 4T^{-2} \int_0^T \alpha_s^{-1} ds \int_0^T \sigma_s^4 \alpha_s ds$$

and

$$A_n^{(2)} := \frac{4}{T} \sum_{i=\tilde{k}_n}^{N_n-1-\tilde{k}_n} \left(\Delta X_i^n \right)^2 \mathbf{1}_{\{|\Delta X_i^n| > \tilde{u}_i\}} \left(\widehat{\sigma_{t_i}^2 \alpha_{t_i}} + \widehat{\sigma_{t_i-}^2 \alpha_{t_i-}} \right)$$

$$\to^{\mathbb{P}} 4T^{-2} \int_0^T \alpha_s^{-1} ds \left\{ \sum_{0 < s \leq T} \Delta J_s^2 (\sigma_s^2 \alpha_s + \sigma_{s-}^2 \alpha_{s-}) \right\}.$$

For $A_n^{(1)}$, we first show that we can replace ΔX_i^n by its continuous part $\Delta \tilde{X}_i^n$ in the square increments of the formula. To do so, define

$$B_n^{(1)} = \frac{4N_n}{T^2} \sum_{i=2}^{N_n-1} \left(\Delta \tilde{X}_i^n\right)^2 \left(\Delta X_{i-1}^n\right)^2 \mathbf{1}_{\{|\Delta X_i^n| \leq \widetilde{u}_i\}} \mathbf{1}_{\{|\Delta X_{i-1}^n| \leq \widetilde{u}_{i-1}\}},$$

$$B_n^{(2)} = \frac{4N_n}{T^2} \sum_{i=2}^{N_n-1} \left(\Delta \tilde{X}_i^n\right)^2 \left(\Delta \tilde{X}_{i-1}^n\right)^2 \mathbf{1}_{\{|\Delta X_i^n| \leq \widetilde{u}_i\}} \mathbf{1}_{\{|\Delta X_{i-1}^n| \leq \widetilde{u}_{i-1}\}}.$$

Let us define $(\tau_q)_{1 \leq q \leq N_J < +\infty}$ a.s be the successive jump times of J (i.e $\Delta J_{\tau_q} \neq 0$ a.s), and for $1 \leq q \leq N_J$, i_q is such that $t_{i_q-1}^n < \tau_q \leq t_{i_q}^n$. By the finite activity property of J, for n sufficiently large we have

$$|A_n^{(1)} - B_n^{(1)}| = \frac{4N_n}{T^2} \sum_{q=1}^{N_J} \left(\left(\Delta J_{\tau_q}\right)^2 + 2\left|\Delta J_{\tau_q} \Delta \tilde{X}_{i_q}^n\right|\right) \left(\Delta X_{i-1}^n\right)^2 \mathbf{1}_{\{|\Delta X_i^n| \leq \widetilde{u}_i\}} \mathbf{1}_{\{|\Delta X_{i-1}^n| \leq \widetilde{u}_{i-1}\}},$$

$$\leq \frac{4N_n}{T^2} \sum_{q=1}^{N_J} \underbrace{\left(\left(\Delta J_{\tau_q}\right)^2 + 2\left|\Delta J_{\tau_q} \Delta \tilde{X}_{i_q}^n\right|\right)}_{O_{\mathbb{P}}(1)} \underbrace{\left(\Delta X_{i-1}^n\right)^2}_{O_{\mathbb{P}}(n^{-1+\gamma})} \underbrace{\tilde{\alpha} \left(\Delta t_{i_q}^n\right)^\omega}_{O_{\mathbb{P}}(n^{-\omega+\gamma\omega})} \underbrace{\left|\Delta \tilde{X}_{i_q}^n + \Delta J_{\tau_q}\right|^{-1}}_{O_{\mathbb{P}}(1)},$$

where we have used assumption **(H)**, and the fact that $\Delta J_{\tau_q} \neq 0$ whereas $\Delta \tilde{X}_{i_q}^n = o_{\mathbb{P}}(1)$ for the estimate $\left|\Delta \tilde{X}_{i_q}^n + \Delta J_{\tau_q}\right|^{-1} = O_{\mathbb{P}}(1)$. Since the sum is finite almost surely, overall we deduce $|A_n^{(1)} - B_n^{(1)}| = O_{\mathbb{P}}(n^{\gamma-\omega+\gamma\omega})$ and since γ can be considered arbitrary close to 0 we deduce that $A_n^{(1)} - B_n^{(1)} \to^{\mathbb{P}} 0$. Similarly we show that $B_n^{(1)} - B_n^{(2)} \to^{\mathbb{P}} 0$. Now we get rid of the indicator functions in $B_n^{(2)}$. Define

$$B_n^{(3)} = \frac{4N_n}{T^2} \sum_{i=2}^{N_n-1} \left(\Delta \tilde{X}_i^n\right)^2 \left(\Delta \tilde{X}_{i-1}^n\right)^2.$$

Then easy calculation gives

$$\mathbb{E}_{\mathcal{U}} |B_n^{(2)} - B_n^{(3)}| = \frac{4N_n}{T^2} \mathbb{E}_{\mathcal{U}} \sum_{i=2}^{N_n-1} \left(\Delta \tilde{X}_i^n\right)^2 \left(\Delta \tilde{X}_{i-1}^n\right)^2 \mathbf{1}_{\{|\Delta X_i^n| > \widetilde{u}_i\} \cup \{|\Delta X_{i-1}^n| > \widetilde{u}_{i-1}\}}$$

$$\leq \frac{4N_n}{T^2} \mathbb{E}_{\mathcal{U}} \sum_{i=2}^{N_n-1} \left(\Delta \tilde{X}_i^n\right)^2 \left(\Delta \tilde{X}_{i-1}^n\right)^2 \left(\frac{|\Delta X_i^n|}{\tilde{\alpha} \left(\Delta t_i^n\right)^\omega} + \frac{|\Delta X_{i-1}^n|}{\tilde{\alpha} \left(\Delta t_i^n\right)^\omega}\right)$$

$$\leq K N_n^2 \times n^{-5/2+5/2\gamma+\omega-\omega\gamma} \to^{\mathbb{P}} 0$$

since $\omega < 1/2$. To conclude, note that by Step 2 of this proof $B_n^{(3)} \to^{\mathbb{P}} 4T^{-2} \int_0^T \alpha_s^{-1} ds \int_0^T \sigma_s^4 \alpha_s ds$ so that combined with $A_n^{(1)} - B_n^{(3)} \to^{\mathbb{P}} 0$, this yields the desired convergence for $A_n^{(1)}$ as well. For $A_n^{(2)}$, by similar techniques as above, defining

$$C_n^{(1)} = \frac{4}{T} \sum_{q=1}^{N_J} \Delta J_{\tau_q}^2 \left(\widehat{\sigma_{t_{i_q}}^2 \alpha_{t_{i_q}}} + \widehat{\sigma_{t_{i_q}-}^2 \alpha_{t_{i_q}-}}\right), \tag{C.109}$$

we easily deduce $A_n^{(2)} = C_n^{(1)} + o_{\mathbb{P}}(1)$. Moreover, we also easily deduce by assumption **(H)** along with the fact that $\tilde{k}_n \Delta_n \to 0$ and $\tilde{k}_n \to \infty$,

$$\widehat{\sigma_{t_{i_q}}^2 \alpha_{t_{i_q}}} = \frac{N_n}{\tilde{k}_n T} \sum_{j=i_q}^{i_q+\tilde{k}_n-1} \left(\Delta X_j^n\right)^2 \mathbf{1}_{\{|\Delta X_j^n| \leq \widetilde{u}_j\}}$$

$$= \frac{N_n}{\tilde{k}_n T} \sum_{j=i_q}^{i_q+\tilde{k}_n-1} \left(\Delta \tilde{X}_j^n\right)^2 + o_{\mathbb{P}}(1)$$

$$= \frac{N_n}{\tilde{k}_n T} \sum_{j=i_q}^{i_q+\tilde{k}_n-1} \sigma_{\tau_q}^2 \Delta t_j^n + o_{\mathbb{P}}(1)$$

$$= \frac{N_n \Delta_n}{T} \sigma_{\tau_q}^2 \alpha_{\tau_q} \underbrace{\frac{1}{\tilde{k}_n} \sum_{j=i_q}^{i_q + \tilde{k}_n - 1} U_j^n}_{\to^{\mathbb{P}} \mathbb{E}[U_1^1] = 1} + o_{\mathbb{P}}(1)$$

$$\to^{\mathbb{P}} T^{-1} \int_0^T \alpha_s^{-1} ds \sigma_{\tau_q}^2 \alpha_{\tau_q},$$

where we have used that $N_n \Delta_n \to^{\mathbb{P}} \int_0^T \alpha_s^{-1} ds$, and the law of large numbers for the i.i.d sequence $(U_i^n)_{i,n}$. Similarly, we also have $\widehat{\sigma_{t_{i_q}-}^2 \alpha_{t_{i_q}-}} \to^{\mathbb{P}} T^{-1} \int_0^T \alpha_s^{-1} ds \sigma_{\tau_q-}^2 \alpha_{\tau_q-}$. Finally, combined with (C.109), we deduce the desired convergence for $A_n^{(2)}$.

Under the alternative, similar techniques yield that when $\eta_0 > 0$ the variance estimators $\widehat{V}_k, k \in \{1, \dots, 5\}$ remain of order $O_{\mathbb{P}}(1)$ (although they become inconsistent).

We show in what follows the consistency of the test.

Proof of Corollary 3.4. Under the null hypothesis, the corollary is a direct consequence of Theorem 3.1 (with $\eta_0 = \mathcal{K} = 0$) and Proposition 3.3. Under the alternative $\eta_0 > 0$, by Theorem 3.1, we have

$$\widehat{\sigma}_{n,err}^2 - \widehat{\sigma}_{n,exp}^2 \to^{\mathbb{P}} 2\tilde{\eta}_0 > 0 \mathbb{P} - a.s.$$

so that this yields $\frac{N_n (\widehat{\sigma}_{n,err}^2 - \widehat{\sigma}_{n,exp}^2)^2}{\widehat{V}_k} \to^{\mathbb{P}} +\infty$ (by Proposition 3.3 we have $\widehat{V}_k = O_{\mathbb{P}}(1)$), which completes the proof. □

Finally, we show Corollary 3.5.

Proof of Corollary 3.5. We start by showing Formula (3.16) along with the fact that $\widehat{\sigma}_{n,exp}^2$ is equal to the least square estimator of Li et al. (2016), introduced in (9). Indeed, (3.16) is obtained solving directly the first order condition $\frac{\partial l_{n,exp}}{\partial \sigma^2}(\widehat{\sigma}_{n,exp}^2, \widehat{\theta}_{n,exp}) = 0$, using Definition (3.2) for $l_{n,exp}$. Moreover, the first order condition for $\widehat{\theta}_{n,exp}$ reads $\partial_\theta \mu(\widehat{\theta}_{n,exp})^T (Y - \mu(\widehat{\theta}_{n,exp})) = 0$, which is also the first order condition related to the quadratic loss introduced in equation (9) in Li et al. (2016). This proves that $\widehat{\theta}_{n,exp}$ and the estimator (9) from Li et al. (2016) coincide. Now, the convergence (3.15) is a straightforward consequence of the consistency of $\widehat{\theta}_{n,exp}$ under $\mathcal{H}_0$ along with the right continuity of the efficient price X. Therefore, the convergences stated in (3.18) and (3.19) are particular cases of Theorem 3.1. □

Proof of Proposition 3.6. When the noise is large (i.e. $a_0^2 > 0$ is fixed), similar calculation to that of the proof of Proposition 3.3 yields

$$\widehat{V}_1 = \frac{4N_n}{T^2} \sum_{i=2}^{N_n} \underbrace{\Delta \epsilon_{t_i^n}^2 \Delta \epsilon_{t_{i-1}^n}^2}_{O_{\mathbb{P}}(1)} + o_{\mathbb{P}}(N_n^2)$$

$$= O_{\mathbb{P}}(N_n^2).$$

Similarly, we have $\widehat{V}_i = O_{\mathbb{P}}(N_n^2)$ for $i = 2, \dots, 5$. □

Proof of Corollary 3.7. Under the large noise alternative, adapting the proof of Theorem C.14 for a fixed $a_0^2 > 0$ easily yields $\widehat{\sigma}_{n,err}^2 - \widehat{\sigma}_{n,exp}^2 = 2a_0^2 T^{-1} N_n + o_{\mathbb{P}}(N_n)$, that is $N_n (\widehat{\sigma}_{n,err}^2 - \widehat{\sigma}_{n,exp}^2)^2 = 4a_0^4 T^{-2} N_n^3 + o_{\mathbb{P}}(N_n^3)$ so that, by Proposition 3.6, for any $i = 1, \dots, 5$ we have $S_i = N_n (\widehat{\sigma}_{n,err}^2 - \widehat{\sigma}_{n,exp}^2)^2 / \widehat{V}_i \to^{\mathbb{P}} +\infty$ since $\widehat{V}_i = O_{\mathbb{P}}(N_n^2)$. □

Proof of Proposition 3.8. When $\phi = 0$, note that we are in the situation where the model (1.3) remains true with $\theta_0 = \tilde{\theta}$. In particular, under the small noise assumption, we still have $\widehat{\theta}_{exp} - \tilde{\theta} = O_{\mathbb{P}}(1/N_n)$ and thus all the calculations in the proof of Proposition 3.3 remain true. In particular for $i = 1, 2$, $V_i = O_{\mathbb{P}}(1)$. Under the large noise alternative, a similar argument yields $V_i = O_{\mathbb{P}}(N_n^2)$ as in the proof of Proposition 3.6. □

Proof of Corollary 3.9. Again, when $\phi = 0$, this is equivalent to assume that the model of log-returns (1.3) remains true with $\theta_0 = \tilde{\theta}$. In particular, under the small noise assumption, Theorem C.14 remains valid replacing θ_0 by $\tilde{\theta}$, and in particular we still have $\widehat{\sigma}_{n,err}^2 - \widehat{\sigma}_{n,exp}^2 \to^{\mathbb{P}} 2\tilde{\eta}_0$. Combined with Proposition 3.8, this yields for $i = 1, 2$, $S_i \to^{\mathbb{P}} +\infty$. For the large noise case, a similar argument yields $\widehat{\sigma}_{n,err}^2 - \widehat{\sigma}_{n,exp}^2 = 2T^{-1} a_0^2 N_n + o_{\mathbb{P}}(N_n)$ as in the proof of Proposition 3.6 and thus, by Proposition 3.8 we can conclude for $i = 1, 2$ that $S_i \to^{\mathbb{P}} +\infty$. □

Proof of Lemma 4.6. The fact that $\widehat{\pi}_V = \pi_V + o_{\mathbb{P}}(1)$ in the fixed noise $a_0^2 > 0$ is a direct consequence of the consistency of $\widehat{\theta}_{n,err}$ and $\widehat{a}_{n,err}^2$ by Theorem 4.1 along with the mixing condition (2.11). Now we prove the second claim.

For $i \in \{1, \ldots, N_n\}$, we use the notation $\phi_i(\theta) = \phi(Q_i, \theta)$. By the mean value theorem, there exist $A_n \in [\widehat{a}_{n,err}^2, \eta_0 T/n]$ and $B_n \in [(N_n + 1)^{-1} \sum_{i=0}^{N_n} \phi_i(\widehat{\theta}_{n,err})^2, \mathbb{E}[\phi_0(\theta_0)^2]]$ such that

$$\widehat{\pi}_V - \pi_V = (A_n + B_n)^{-2} \left\{ -B_n \left(\widehat{a}_{n,err}^2 - \frac{\eta_0 T}{n} \right) + A_n \left(\frac{\sum_{i=0}^{N_n} \phi_i(\widehat{\theta}_{n,err})^2}{N_n + 1} - \mathbb{E}[\phi_0(\theta_0)^2] \right) \right\},$$

and since by Theorem 3.1 we have the estimate $\widehat{a}_{err}^2 = \eta_0 T/n + O_{\mathbb{P}}(N^{-3/2})$ and $\widehat{\theta}_{n,err} = \theta_0 + O_{\mathbb{P}}(N^{-1})$, we easily deduce, again using the mixing condition (2.11), $(N_n + 1)^{-1} \sum_{i=0}^{N_n} \phi_i(\widehat{\theta}_{n,err})^2 = \mathbb{E}[\phi(\theta_0)^2] + O_{\mathbb{P}}(N^{-1/2})$, and thus $(A_n + B_n)^{-2} = \mathbb{E}[\phi_0(\theta_0)^2]^{-2} + o_{\mathbb{P}}(1)$, $B_n \left(\widehat{a}_{n,err}^2 - \eta_0 T/n \right) = O_{\mathbb{P}}(N_n^{-3/2})$, and finally $A_n \left((N + 1)^{-1} \sum_{i=0}^{N_n} \phi_i(\widehat{\theta}_{n,err})^2 - \mathbb{E}[\phi_0(\theta_0)^2] \right) = O_{\mathbb{P}}(N_n^{-3/2})$ so that

$$\widehat{\pi}_V - \pi_V = O_{\mathbb{P}}(N_n^{-3/2})$$

and we are done. □

References

Adams, R.A., Fournier, J.J., 2003. Sobolev Spaces, vol. 140. Academic press.

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

Aït-Sahalia, Y., Mykland, P.A., Zhang, L., 2011. Ultra high frequency volatility estimation with dependent microstructure noise. J. Econometrics 160 (1), 160–175.

Aït-Sahalia, Y., Xiu, D., 2016. A Hausman test for the presence of market microstructure noise in high frequency data. J. Econometrics http://dx.doi.org/10.1016/j.jeconom.2018.12.013, (in press).

Almgren, R., Chriss, N., 2001. Optimal execution of portfolio transactions. J. Risk 3, 5–40.

Altmeyer, R., Bibinger, M., 2015. Functional stable limit theorems for quasi-efficient spectral covolatility estimators. Stochastic Process. Appl. 125 (12), 4556–4600.

Andersen, T., Bollerslev, T., Diebold, F., Labys, P., 2000. Great realizations. Risk 105–108.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001. The distribution of realized exchange rate volatility. J. Amer. Statist. Assoc. 96 (453), 42–55.

Andersen, T.G., Cebiroglu, G., Hautsch, N., 2017. Volatility, Information Feedback and Market Microstructure Noise: a Tale of Two Regimes. Working paper available at SSRN 2921097.

Andersen, T.G., Dobrev, D., Schaumburg, E., 2012. Jump-robust volatility estimation using nearest neighbor truncation. J. Econometrics 169 (1), 75–93.

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.

Barndorff-Nielsen, O.E., Shephard, N., 2002. Estimating quadratic variation using realized variance. J. Appl. Econometrics 17 (5), 457–477.

Barndorff-Nielsen, O.E., Shephard, N., 2004a. Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. Econometrica 72 (3), 885–925.

Barndorff-Nielsen, O.E., Shephard, N., 2004b. Power and bipower variation with stochastic volatility and jumps. J. Financ. Econ. 2 (1), 1–37.

Black, F., 1986. Noise. J. Finance 41 (3), 528–543.

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

Chen, F., Hall, P., 2013. Inference for a nonstationary self-exciting point process with an application in ultra-high frequency financial data modeling. J. Appl. Probab. 50 (4), 1006–1024.

Christensen, K., Hounyo, U., Podolskij, M., 2018. Is the diurnal pattern sufficient to explain intraday variation in volatility? A nonparametric assessment. J. Econometrics 205 (2), 336–362.

Clinet, S., Potiron, Y., 2017. Estimation for High-Frequency Data Under Parametric Market Microstructure Noise. Working paper available at arXiv: 1712.01479.

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

Clinet, S., Potiron, Y., 2018b. A Relation Between the Efficient, Transaction and Mid Prices: Disentangling Sources of High Frequency Market Microstructure Noise. Working paper available at SSRN:3167014.

Clinet, S., Potiron, Y., 2018c. Statistical inference for the doubly stochastic self-exciting process. Bernoulli 24 (4B), 3469–3493.

Cont, R., Kukanov, A., Stoikov, S., 2014. The price impact of order book events. J. Financ. Econ. 12 (1), 47–88.

Da, R., Xiu, D., 2017. When Moving-Average Models Meet High-Frequency Data: Uniform Inference on Volatility. Working paper available on Dacheng Xiu's website.

Diebold, F.X., Strasser, G., 2013. On the correlation structure of microstructure noise: A financial economic approach. Rev. Econom. Stud. 80 (4), 1304–1337.

Engle, R.F., Russell, J.R., 1998. Autoregressive conditional duration: a new model for irregularly spaced transaction data. Econometrica 1127–1162.

Ghysels, E., Harvey, A.C., Renault, E., 1996. 5 Stochastic volatility. Handbook of Statist. 14, 119–191.

Glosten, L.R., Harris, L.E., 1988. Estimating the components of the bid/ask spread. J. Financ. Econ. 21 (1), 123–142.

Hansen, P.R., Lunde, A., 2006. Realized variance and market microstructure noise. J. Bus. Econom. Statist. 24 (2), 127–161.

Hasbrouck, J., 1993. Assessing the quality of a security market: A new approach to transaction-cost measurement. Rev. Financ. Stud. 6 (1), 191–212.

Hasbrouck, J., 2007. Empirical Market Microstructure: the Institutions, Economics, and Econometrics of Securities Trading. Oxford University Press.

Hausman, J.A., 1978. Specification tests in econometrics. Econometrica 1251–1271.

Huang, X., Tauchen, G., 2005. The relative contribution of jumps to total price variance. J. Financ. Econ. 3 (4), 456–499.

Jacod, J., 1997. On continuous conditional gaussian martingales and stable convergence in law. Sém. Probab. Strasbourg 31, 232–246, http://eudml.org/doc/113958.

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

Jacod, J., Protter, P.E., 2011. Discretization of Processes. Springer Science & Business Media.

Jacod, J., Shiryaev, A., 2003. Limit Theorems for Stochastic Processes, second ed. Springer-Verlag, Berlin.

Kallenberg, O., 2006. Foundations of Modern Probability. Springer Science & Business Media.

Kalnina, I., Linton, O., 2008. Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error. J. Econometrics 147 (1), 47–59.

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

Lee, C., Ready, M.J., 1991. Inferring trade direction from intraday data. J. Finance 46 (2), 733–746.

Leeb, H., Pötscher, B.M., 2005. Model selection and inference: Facts and fiction. Econometric Theory 21 (1), 21–59.

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

Madhavan, A., 2000. Market microstructure: A survey. J. Financial Mark. 3 (3), 205–258.

Madhavan, A., Richardson, M., Roomans, M., 1997. Why do security prices change? A transaction-level analysis of NYSE stocks. Rev. Financ. Stud. 10 (4), 1035–1064.

Mancini, C., 2009. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. Scand. J. Stat. 36 (2), 270–296.

McCullagh, P., 1987. Tensor Methods in Statistics, vol. 161. Chapman and Hall London.

Muni Toke, I., 2016. Reconstruction of order flows using aggregated data. Mark. Microstruct. Liquid. 2 (02), 1650007.

O'hara, M., 1995. Market Microstructure Theory, vol. 108. Blackwell Cambridge, MA.

Patton, A.J., 2011. Data-based ranking of realised volatility estimators. J. Econometrics 161 (2), 284–303.

Potiron, Y., Mykland, P.A., 2016. Local parametric estimation in high frequency data. J. Bus. Econom Statist. (forthcoming).

Reiss, M., 2011. Asymptotic equivalence for inference on the volatility from noisy observations. Ann. Statist. 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. J. Financ. Econ. 9 (2), 344–366.

Robert, C.Y., Rosenbaum, M., 2012. Volatility and covariation estimation when microstructure noise and trading times are endogenous. Math. Finance 22 (1), 133–164.

Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. J. Finance 39 (4), 1127–1139.

Stoll, H.R., 2000. Presidential address: friction. J. Finance 55 (4), 1479–1514.

Todorov, V., Tauchen, G., 2011. Volatility jumps. J. Bus. Econom. Statist. 29 (3), 356–371.

Van der Vaart, A.W., 2000. Asymptotic Statistics. Cambridge university press.

Varneskov, R.T., 2017. Estimating the quadratic variation spectrum of noisy asset prices using generalized flat-top realized kernels. Econometric Theory 33 (6), 1457–1501.

Xiu, D., 2010. Quasi-maximum likelihood estimation of volatility with high frequency data. J. Econometrics 159 (1), 235–250.

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. J. Amer. Statist. Assoc. 100 (472), 1394–1411.

Further Reading

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.