Local Parametric Estimation in High Frequency Data

Yoann Potiron & Per Mykland


To link to this article: https://doi.org/10.1080/07350015.2019.1566731

View supplementary material

Accepted author version posted online: 13 Feb 2019.
Published online: 06 May 2019.

Submit your article to this journal

Article views: 196

View related articles

View Crossmark data

Citing articles: 2 View citing articles
Local Parametric Estimation in High Frequency Data

Yoann POTIRON
Faculty of Business and Commerce, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
(potiron@fbc.keio.ac.jp)

Per MYKLAND
Department of Statistics, The University of Chicago, 5734 S. University Avenue, Chicago, IL
(mykland@pascal.uchicago.edu)

We give a general time-varying parameter model, where the multidimensional parameter possibly includes jumps. The quantity of interest is defined as the integrated value over time of the parameter process

\[ \Theta_1 = T^{-1} \int_0^T \theta^* dt. \]

We provide a local parametric estimator (LPE) of \( \Theta_1 \) and conditions under which we can show the central limit theorem. Roughly speaking those conditions correspond to some uniform limit theory in the parametric version of the problem. The framework is restricted to the specific convergence rate \( n^{1/2} \). Several examples of LPE are studied: estimation of volatility, powers of volatility, volatility when incorporating trading information and time-varying MA(1).

KEY WORDS: Integrated volatility; Market microstructure noise; Powers of volatility; Quasi-maximum likelihood estimator

1. INTRODUCTION

Modeling dynamics is essential in various fields, including finance, economics, physics, environmental engineering, geology, and sociology. Time-varying parametric models can deal with a specific problem in dynamics, namely, the temporal evolution of systems. The extensive literature on time-varying parameter models and local parametric methods include and are not limited to Fan and Gijbels (1996), Hastie and Tibshirani (1993), or Fan and Zhang (1999) when regression and generalized regression models are involved, locally stationary processes following the work of Dahlhaus (1997, 2000), Dahlhaus and Rao (2006), or any other time-varying parameter models, for example, Stock and Watson (1998) and Kim and Nelson (2006).

In this paper, we propose to specify local parametric methods in the particular context of high-frequency statistics for a broad class of problems. Local methods have been used extensively in the high-frequency data literature, see, for example, Mykland and Zhang (2009, 2011), Kristensen (2010), Reiß (2011), or Jacod and Rosenbaum (2013), among many others. If we define \( T \) as the horizon time, the (random) target quantity in this monograph is defined as the integrated parameter

\[ \Theta := \frac{1}{T} \int_0^T \theta^* ds, \]

which can be equal to the volatility, the covariation between several assets, the variance of the microstructure noise, the friction parameter of the model with uncertainty zones (see Example 4.4 for more details), the time-varying parameters of the MA(1) model, etc. To estimate the integrated parameter, we estimate the local parameter on each block by using the parametric estimator on the observations within the block and take a weighted sum of the local parameter estimates, where each weight is equal to the corresponding block length. We call the obtained estimator the local parametric estimator (LPE).

In Section 3, we investigate conditions under which we can establish the related central limit theorem with convergence rate \( n^{1/2} \), where \( n \) is the (possibly expected) number of observations. The framework is such that the local block length vanishes asymptotically. Basically, we aim to provide the statistician with a transparent and as simple as possible device to tackle the time-varying parameter problem based on central limit theory in the parametric version of the problem. The original key probabilistic step of the proof, which formally allows for switching from random to deterministic parameter, is the use of regular conditional distribution theory (see, e.g., Breiman 1992). The price to pay is some kind of uniformity in the parametric limit theory results, and to show that some deviation between the parametric and the time-varying parameter model vanishes asymptotically.

In Section 4, the technology is used on five distinct examples to derive the related central limit theorems. As far as the authors know, all those results are new. Depending on the considered example, the LPE is useful for one or several of the following reasons:

- **Robustness**: The LPE is robust to time-varying parameters (such as the noise variance, \( \eta \) from the model with uncertainty zones, the parameters of the MA(1) process) which...
are usually assumed constant. This is the case of all our examples, except for Example 4.3.

- **Efficiency:** The LPE turns out to be more efficient than the global estimator or existing concurrent approaches. This is the case of Example 4.3. In addition, the LPE is conjectured to be efficient in all our examples except for Example 4.4.

- **Definition of new estimators:** It can be the case that the estimator does not work globally but that the LPE provides a good candidate as in Examples 4.2 and 4.3.

We describe the five examples in what follows. To estimate integrated volatility under noisy observations, Xiu (2010) studied the quasi-maximum likelihood-estimator (QMLE) originally examined in Aït-Sahalia, Mykland, and Zhang (2005), showed the corresponding asymptotic theory when the variance of the noise is fixed and obtained a convergence rate \( n^{1/4} \), which is optimal (see Gloter and Jacob 2001). More recently, Aït-Sahalia and Xiu (2019) establish that it is robust to shrinking noise satisfying \( O_p(1/n^{3/4}) \) and Da and Xiu (2017) obtain central limit theorem with rate ranging from \( n^{1/4} \) to \( n^{1/2} \) depending on the magnitude of the noise. When assuming that it is \( O_p(1/\sqrt{n}) \), we show that the LPE of the QMLE is optimal (with rate \( n^{1/2} \)) and furthermore robust to time-varying noise variance.

Another important problem, which goes back to Barndorff-Nielsen and Shephard (2002), is the estimation of higher powers of volatility. To do that, we define a LPE where the latent parameter \( \theta \) is the case of Example 4.3. In addition, a modification of the local estimator as in Examples 4.2 and 4.3.

As an example, the observations can satisfy \( Z_{t_n} = X_{t_n} + \epsilon_{t_n} \), where \( X_t = \sigma_d W_t \) stands for the efficient price, \( W_t \) is a standard Brownian motion, \( \epsilon_{t_n} \) corresponds to the market microstructure noise (which will be restricted to be of order \( \epsilon_{t_n} = O_p(1/\sqrt{n}) \) due to the limitation of the technology developed in Section 3), is iid and independent from \( X_t \), and the latent parameter is equal to the volatility, that is, \( \theta^* = \sigma^2 \).

We assume that the parameter process \( \theta^* \) takes values in \( K \), a (not necessarily compact) subset of \( \mathbb{R}^p \). We do not assume any independence between \( \theta^* \) and the other quantities driving the observations, such as the Brownian motion of the efficient price process. In particular, there can be leverage effect (see, e.g., Wang and Mykland 2014; Aït-Sahalia et al. 2017). Also, the arrival times \( \tau_{i,n} \) and the parameter \( \theta^* \) can be correlated, that is, there is (some kind of) endogeneity in sampling times.

2.2. Asymptotics

There are commonly two choices of asymptotics in the literature: the high-frequency asymptotics, which makes the number of observations explode on \([0,T]\), and the low-frequency asymptotics, which takes \( T \) to infinity. We choose the former one. Investigating the low-frequency implementation case is beyond the scope of this article.\(^1\)

2.3. Estimation

The approach taken here is frequent in high-frequency data. We define the block size (i.e., the number of observations in a block) as \( h_n \), and the number of blocks as \( B_n := \lceil N_n h_n^{-1} \rceil \). For

\(^1\)If we set down the asymptotic theory in the same way as in Dahlhaus (1997, p. 3), we conjecture that the results of this article would stay true.
We consider first the simple case when observations are regular, that is, \( \tau_{i,n} = iT/n \) and \( N_n = n \). We assume that \( \mathcal{J}_T \) is a (continuous-time) filtration on \((\Omega', \mathcal{F}, P)\) such that \( \theta^*_n \) is adapted to it. In the following of this paper, when using the conditional expectation \( \mathbb{E}_r[Z] \), we will refer to the conditional expectation of \( Z \) knowing \( \mathcal{J}_T \). We define the discrete-time version of the filtration as \( \mathcal{I}_{i,n} = \mathcal{J}_{\tau_{i,n}} \). Finally, if we denote the returns of the observations as

\[
R_{i,n} = Z_{\tau_{i,n}} - Z_{\tau_{i-1,n}},
\]

we assume that the returns can be expressed as

\[
R_{i,n} = F_n[(P_{s,n})_{0 \leq s \leq \tau_{i,n}}, U_{i,n}, \{\theta^*_n\}_{s \leq \tau_{i,n}}],
\]

where \( F_n(x, y, z) \) is a \( \mathbb{R}^d \)-dimensional nonrandom function, the random innovation \( U_{i,n} \) are iid (although with distribution which can depend on \( n \)) adapted to \( \mathcal{I}_{i,n} \) and independent of the past information \( \mathcal{I}_{i-1,n} \). \( P_{s,n} \) is a (possibly multidimensional) process adapted to \( \mathcal{J}_T \) which stands for the past that matters in the model. We further assume that \( P_{s,n} \) is independent from \( \theta^*_n \).

The key example stands as follows. We assume that the observations are following the additive model \( Z_{\tau_{i,n}} = X_{\tau_{i,n}} + \epsilon_{i,n} \), where \( X_t = \sigma dW_t \) is the efficient price and \( \epsilon_{i,n} \) (the (shrinking) iid noise independent from \( X_t \)), and that the parameter is \( \theta^*_n = \sigma^2 \). In that case \( U_{i,n} = \{(W_s)_{s \leq \tau_{i,n}} - W_{\tau_{i-1,n}}, \epsilon_{i,n}\} \), and \( P_{s,n} = \epsilon_{i,n} \) if \( \tau_{i,n} \leq s < \tau_{i+1,n} \). The function \( F_n \) takes on the form

\[
F_n = \int_{\tau_{i-1,n}}^{\tau_{i,n}} \sigma dW_s + \epsilon_{i,n} - \epsilon_{i-1,n}.
\]

Crucial to the expression (8) is that the dependence in the past is only through the past noise \( \epsilon_{i-1,n} \), that is, we do not need to know the whole past of \( P_{s,n} \), but rather only the current value. This will be very useful in what follows.

We provide now the outline of the method. Our goal is to investigate the limit distribution of (5) using prior limit result on the parametric version of the problem. A common approach in high frequency statistics proofs consists in decomposing \( (\hat{\Theta}_{i,n} - \Theta_{i,n}) \) into

\[
(\hat{\Theta}_{i,n} - \Theta_{i,n}) + (\hat{\Theta}_{i,n} - \hat{\Theta}_{i-1,n}) + (\hat{\Theta}_{i-1,n} - \Theta_{i,n}),
\]

where \( \hat{\Theta}_{i,n} \) stands for the estimator when we hold the parameter constant on each block. Then, one can usually deal with the first term and the third term (most likely using Burkholder–Davis–Gundy and Markov type of inequalities) and eventually show that they vanish asymptotically. The main work lies in establishing the central limit theory of the second term in (9). A typical proof consists in using locally parametric results along with some Riemann sum argument. But this can be cumbersome as the parameter on each block, although

\[
\frac{d}{ds} \int_0^s f(X_u) \sigma dW_u = f(X_s) \sigma dW_s + \sigma^2 \int_0^s f'(X_u) \sigma dW_u,
\]

where \( f \) is a \( \mathbb{R}^d \)-valued function. The related assumption is that \( \tau \) is a \( \mathcal{J}_T \)-stopping time.
constant, is random. Instead, we propose to look at the further decomposition of \((\hat{\Theta}_{i,n} - \theta_{T_{i-1,n}}^*)\) into
\[
(\hat{\Theta}_{i,n} - \hat{\Theta}_{P,i,n}^\theta) + \hat{\Theta}_{P,i,n}^\theta - \theta_{T_{i-1,n}}^*
\]
where \(\hat{\Theta}_{P,i,n}^\theta := \hat{\Theta}_{i,n} \mid \{P_{s,n}\}_{0 \leq s \leq T_{i-1,n}} = P\). (10)

and \(P\) is a fixed nonrandom past. In the case of (8), we can choose \(P = 0\). From this new decomposition, it is expected as relatively accessible to show that the first term goes to 0, so that the central limit theory will be investigated on the second term of the decomposition. By conditioning on one particular past in (11), we got rid of some randomness, although the parameter is still random. Using conditional regular distribution results in our proofs, we actually show that we can also take the parameter nonrandom. The price to pay for such method is to show some kind of uniformity in the parameter value when showing the limit results, and that the first term in (10) vanishes asymptotically.

We introduce some definition. For \(i = 1, \ldots, B_n\) we define the returns on the \(i\)th block \(R_{i,n}^j := R_{i-1,n}h_{i-1,n}, j, n\) for \(j = 1, \ldots, h_n\), and similarly \(U_{i,n}^j, W_{i,n}^j, \tilde{W}_{i,n}^j, e_{i,n}^j\). We assume that
\[
R_{i,n}^j := \tilde{\Theta}_{h_{i,n}}(R_{i-1,n}^j, \ldots, R_{i,n}^j),
\]
where \(\hat{\Theta}_{h_{i,n}}\) is a function on \(\mathbb{R}^{dh_n}\). The approximated returns and the approximated estimates are defined as
\[
\tilde{R}_{i,n}^j := F_n\{\{P_{s,n}\}_{0 \leq s \leq T_{i-1,n}}, U_{i,n}^j, \theta_{T_{i-1,n}}^*\}.
\]

We also introduce the conditional parametric version as
\[
\tilde{R}_{i,n}^{1,P} := [\tilde{R}_{i,n}^j \mid \{U_{i,n}^j\}_{0 \leq j \leq b_i}, \{P_{s,n}\}_{0 \leq s \leq T_{i-1,n}} = P],
\]
where \(\tilde{R}_{i,n}^{1,P} := \tilde{\Theta}_{h_{i,n}}(\tilde{R}_{i,n}^{1,P}, \ldots, \tilde{R}_{i,n}^{P})\). (17)

Here, we fix the past equal to \(P\) in (16), which removes some randomness compared with (13). In the key example, we can (arbitrarily) choose \(P = 0\), and this past will only “affect” the first conditional parametric version of the return on the block equal to
\[
\tilde{R}_{i,n}^{1,P} = \sigma_{T_{i-1,n}}(W_{i,n}^1 - W_{i,n}^0) + \epsilon_{i,n},
\]

whereas for \(j = 2, \ldots, h_n\), we have \(\tilde{R}_{i,n}^{j,P} = \tilde{R}_{i,n}^j\). This key example is an instance where the model is 1-Markovian in the sense that the past only affects the value of the first return on the block. This is quite mild assumption, and we will see that more sophisticated models, such as the model with uncertainty zones, naturally exhibit longer past time-dependence. Moreover, we introduce a parametric version of the returns and the estimators when the parameter is equal to \(\theta\) and the past fixed to \(P\). Accordingly, the randomness is further reduced in the following expressions. This will be useful in Condition (E).

\[
\tilde{R}_{i,n}^{1,P} := \mathbb{E}[\tilde{R}_{i,n}^j \mid \{U_{i,n}^j\}_{0 \leq j \leq b_i}, \theta_{T_{i-1,n}} = \theta, \{P_{s,n}\}_{0 \leq s \leq T_{i-1,n}} = P].
\]

\[
\tilde{\Theta}_{i,n}^{P,\theta} := \tilde{\Theta}_{h_{i,n}}(\tilde{R}_{i,n}^{1,P}; \ldots; \tilde{R}_{i,n}^{b_i,P}).
\]

We provide now the assumptions on \(\theta_*\). The first assumption considers the continuous Itô-semimartingale case.

**Condition (P1).** The parameter \(\theta_*\) is of the form
\[
\theta_* = a_i^\theta dt + \sigma_i^\theta dW^\theta_i,
\]
where \(a_i^\theta\) is adapted locally bounded (of dimension \(p\)) and \(\sigma_i^\theta\) is a nonnegative continuous Itô-process adapted locally bounded (of dimension \(p \times p\)), and \(W^\theta_i\) is a standard \(p\)-dimensional Brownian motion.

We introduce a norm for
\[
u \in \mathbb{R}^p \text{ as } |\nu| := \sqrt{(u(1))^2 + \cdots + (u(p))^2}.
\]

The following assumption allows for a more general process than semi-martingales. Nonetheless, this assumption is quite restrictive, in particular since \(h_n\) does not show up on the right hand-side of (22). In practice this is useful when considering a smooth parameter which cannot be expressed as a “pure drift.”

**Condition (P2).** \(\theta_*\) satisfies uniformly in \(i = 1, \ldots, B_n\) that
\[
\mathbb{E}_{T_{i-1,n}}\left[\sup_{T_{i-1,n} \leq s \leq T_{i,n}} |\theta_* - \theta_*^T_{T_{i-1,n}}|^2\right] = o_p(n^{-1})
\]

As the uniformity of limit results on the whole space \(K\) might be impossible to obtain, we allow to work on the compact subspace \(K_M\), which grows to \(K\) as \(M\) increases. Accordingly, we assume that \(\theta_*\) is locally bounded on a compact set \(K_M\) in the sense that there exists \(\tau_m \xrightarrow{p} T\) such that for any \(m\), there exists \(M_m > 0\) which satisfies \(\theta_*^T \in K_{M_m}\) for any \(t \in [0, \tau_m]\).

We provide in what follows sufficient conditions to the bias condition (3.10), the increment condition (3.11) and the Lindeberg condition (3.13) in Theorem 3.2 from Jacod (1997). (Almost) equivalently, Theorem IX.7.3 and Theorem IX.7.28 in Jacod and Shiryaev (2003) or Theorem 2.2.15 in Jacod and Protter (2011) could have been used. Those conditions are based on the parametric version of the problem.

**Condition (E).** For any (nonrandom) parameter \(\theta \in K\), we assume that there exists a (nonrandom) covariance matrix \(V_0\) positive definite such that for any \(M > 0\), we have \(V_0\) is bounded for any \(\theta \in K_M\) and uniformly in \(\theta \in K_M\) and in \(i = 1, \ldots, B_n\) we have

\[
\mathbb{E}[\hat{\Theta}_{h_{i,n}}^{P,\theta} - \theta]\bigg| = o(n^{-\frac{1}{2}})
\]

\[
\text{var}\left[h_n^{-\frac{1}{2}}(\hat{\Theta}_{h_{i,n}}^{P,\theta} - \theta)\right] = V_0T + o(1)
\]

\[
\mathbb{E}[\left|h_n^{-\frac{1}{2}}(\hat{\Theta}_{h_{i,n}}^{P,\theta} - \theta)\right|^2] = o(1), \forall \epsilon > 0.
\]
We let $B_n(t)$ be the number of blocks before $t$, and $M_b$ the set of all bounded martingales. We now provide the central limit theorem.

**Theorem 1 (Central limit theorem with regular observation times).** We assume Condition (E). Moreover, we assume Condition (P1) and that the block size $h_n$ is such that

$$n^{-\frac{1}{2}}h_n = o(1),$$

(26)

or Condition (P2). Let $M_t$ be a $p$-dimensional square-integrable continuous martingale. Furthermore, we assume that for all $t \in [0, T]$ we have

$$n^{-\frac{1}{2}}h_n \sum_{i=1}^{B_n(t)} \left[ (\hat{\Theta}_{t,n} - \Theta_{t_0,n}) (M_{T_{t,n}} - M_{T_{t_1,n}}) \right]^T \xrightarrow{p} 0,$$

(27)

$$n^{-\frac{1}{2}}h_n \sum_{i=1}^{B_n(t)} \left[ (\hat{\Theta}_{t,n} - \Theta_{t_0,n}) (N_{T_{t,n}} - N_{T_{t_1,n}}) \right]^T \xrightarrow{p} 0,$$

(28)

for all $N \in M_b(M^\perp)$, where $M_b(M^\perp)$ is the class of all elements of $M_b$ which are orthogonal to $M$ (i.e., to all components of $M$). Finally, we assume that

$$n^{-\frac{1}{2}}h_n \sum_{i=1}^{B_n(t)} (\hat{\Theta}_{i,n} - \hat{\Theta}_{i,n}) \xrightarrow{p} 0.$$  

(29)

Then, stably in law as $n \to \infty$, we have

$$n^\frac{1}{2} (\hat{\Theta}_n - \Theta) \to \tilde{Z},$$

(30)

where $(\tilde{Z}, \tilde{Z}_t) = T^{-1} \int_0^T V_{0t}^2 ds$, and $(\tilde{Z}, \tilde{M}_t) = 0$. In particular, we have

$$n^\frac{1}{2} (\hat{\Theta}_n - \Theta) \to \left( T^{-1} \int_0^T V_{0t}^2 ds \right)^{\frac{1}{2}} N(0, 1).$$

(31)

**Remark 1 (Parametric model).** Note that in the case where the time-varying parameter model is equal to the parametric model with parameter equal to $\theta^*$, the VAR of $\Theta_n$ is equal to the variance of the parametric model, that is,

$$n^\frac{1}{2} (\hat{\Theta}_n - \Theta) \to V_{\theta^*}^{\frac{1}{2}} N(0, 1).$$

**Remark 2 (Estimating the VAR).** If the statistician does not have a (parametric) variance estimator at hand and that her parametric estimator can be written as in Mykland and Zhang (2017), one can use the techniques of the cited paper to obtain a variance estimate. Investigating if such techniques would work in our setting is beyond the scope of this paper. If she has a variance estimator $\hat{\eta}_{h_n,n}$, then for any $i = 1, \ldots, B_n$ she can estimate the $i$th block variance $\hat{V}_{i,n}$ as $\hat{V}_{i,n} := \hat{\eta}_{h_n,n}(R^1_{i,n}, \ldots, R^B_{i,n})$, and the VAR as the weighted sum

$$\hat{V}_n = T^{-1} \sum_{i=1}^{B_n} \hat{V}_{i,n} \Delta T_{i,n}.$$  

(32)

This estimator will be consistent under mild uniformity assumptions.

**Remark 3 (Nonzero asymptotic bias).** If we further assume that in place of condition (27) there is a nonzero continuous process $G_t$ such that

$$n^{-\frac{1}{2}}h_n \sum_{i=1}^{B_n(t)} \mathbb{E}_{T_{i-1,n}} \left[ (\hat{\Theta}_{i,n} - \Theta_{i_{-1},n}) (M_{T_{i,n}} - M_{T_{i-1,n}}) \right]^T \xrightarrow{p} G_t,$$

(33)

then (30) still holds, where $(\tilde{Z}, \tilde{Z}_t) = T^{-1} \int_0^T V_{0t}^2 ds$ and $(\tilde{Z}, \tilde{M}_t) = G_t$, but (31) no longer holds.

### 3.2. Nonregular Observation Case

We consider now the case when observations can be random (even endogenous). We define the increment of time as $\Delta \tau_{i,n} := \tau_{i,n} - \tau_{i-1,n}$ and make the first natural assumption.

**Condition (T).** The observation times are such that

$$\mathbb{E} \left[ N_n \right] = O(n),$$

(34)

$$\sup_{1 \leq i \leq N_n} \mathbb{E} \left[ (\Delta \tau_{i,n})^3 \right] = O_p(n^{-3}).$$

(35)

**Remark 4 (Block length).** As an obvious consequence of (35), we have that the block length satisfies $\mathbb{E} \left[ \Delta T_{i,n} \right] = O(h_n n^{-1})$.

The observation times are related to $\theta^*$, as are the returns. We assume that $(R_{i,n}, \Delta \tau_{i,n})$ satisfies (7), and that all the definitions (12)–(20) follow. Finally, we define $\Delta \tilde{T}_{i,n}^\theta = \Delta \tilde{T}_{i,n}^\theta - \tilde{T}_{(h_n-1)i,n}$ and $\Delta \tilde{T}_{i,n}^\theta = \tilde{T}_{h_n i,n} - \tilde{T}_{(h_n-1)i,n}$. We adapt Condition (E) in this case.

**Condition (E*).** For any (nonrandom) parameter $\theta \in K$, we assume that there exists a (nonrandom) covariance matrix $V_{\theta} > 0$ such that for any $M > 0$, we have $V_{\theta}$ is bounded for any $\theta \in K_M$ and uniformly in $\theta \in K_M$ and in $i = 1, \ldots, B_n$ we have

$$\mathbb{E} \left[ (\hat{\Theta}_{i,n} - \Theta_{i,n}) \Delta \tilde{T}_{i,n}^\theta \right] = o(h_n n^{-\frac{1}{2}}),$$

(36)

$$\text{var} \left[ \hat{h}_n^\frac{1}{2} (\hat{\Theta}_{i,n} - \Theta_{i,n}) \Delta \tilde{T}_{i,n}^\theta \right] = V_{\theta} \mathbb{E} \left[ (\Delta \tilde{T}_{i,n}^\theta)^2 \right] h_n n^{-1} + o(h_n^2 n^{-2}).$$

(37)

$$\mathbb{E} \left[ n^2 h_n^{-1} A_{i,n}^\theta \right] = o(1), \quad \forall \epsilon > 0,$$

(38)

where $A_{i,n}^\theta = | \hat{\Theta}_{i,n} - \Theta_{i,n} | \Delta \tilde{T}_{i,n}^\theta$.

We also adopt the central limit theorem.

**Theorem 2 (Central limit theorem with nonregular observation times).** We assume Condition (T) and Condition (E*). Moreover, we assume Condition (P1) and (26), or Condition (P2). Let $M_t$ be a $p$-dimensional square-integrable continuous
martingale. Furthermore, we assume that for all $t \in [0, T]$ we have
\[
\frac{n^2}{T} \sum_{i=1}^{B_n(t)} \mathbb{E}_{T_{i-1}, n} \left[ \left( \hat{\Theta}_{i,n}^\text{P} - \Theta_{i,n}^*) T_{i,n} \right) \hat{M}_{T_{i,n}} - M_{T_{i-1}, n} \right]^T \to 0,
\]  
(39)

and for all $N \in \mathcal{M}_p(M^\perp)$. Finally, we assume that
\[
\frac{n^2}{T} \sum_{i=1}^{B_n(t)} \left[ \left( \hat{\Theta}_{i,n}^\text{P} - \Theta_{i,n}^*) T_{i,n} \right) \hat{M}_{T_{i,n}} - M_{T_{i-1}, n} \right]^T \to 0,
\]  
(40)

uniformly in $i = 1, \ldots, B_n$. Then, stably in law as $n \to \infty$, we have
\[
\frac{n^2}{2} (\hat{\Theta}_{n} - \Theta) \to \mathcal{Z},
\]  
(43)

where $(\mathcal{Z}, \mathcal{Z})_t = T^{-1} \int_0^T V_{0,t}^T ds$, and $(\mathcal{Z}, M)_t = 0$. In particular, we have
\[
\frac{n^2}{2} (\hat{\Theta}_{n} - \Theta) \to \left( T^{-1} \int_0^T V_{0,t}^T ds \right)^\frac{1}{2} \mathcal{N}(0, 1).
\]  
(44)

Remark 5 (Nonzero asymptotic bias). More generally, if there is a nonzero continuous process $G_t$ such that for all $t \in [0, T]$ we have
\[
\frac{n^2}{T} \sum_{i=1}^{B_n(t)} \left[ \left( \hat{\Theta}_{i,n}^\text{P} - \Theta_{i,n}^*) T_{i,n} \right) \hat{M}_{T_{i,n}} - M_{T_{i-1}, n} \right]^T \to G_t,
\]  
(45)

instead of (39), then (43) still holds, where $(\mathcal{Z}, \mathcal{Z})_t = T^{-1} \int_0^T V_{0,t}^T ds$, and $(\mathcal{Z}, M)_t = G_t$, but (44) no longer holds.

3.3. Bias Correction

As the parametric estimator must satisfy the bias condition (36), it is useful to consider in some instances a bias-corrected (BC) version of it which provides the estimate on the ith block $\hat{\Theta}_{i,n}^\text{(BC)}$. The BC LPE is then constructed as
\[
\hat{\Theta}_{n}^\text{(BC)} = \frac{1}{T} \sum_{i=1}^{B_n} \hat{\Theta}_{i,n}^\text{(BC)} T_{i,n}.
\]

4. EXAMPLES

This section provides some applications of the theory introduced in Section 3. The central limit theorems provided in this section are all new. We choose four examples with regular observations in which it is sufficient to show the conditions of Theorem 1. We further consider the model with uncertainty zones where there is endogeneity in observation times implying that we have to verify the more general conditions of Theorem 2.

4.1. Estimation of Volatility With the QMLE

4.1.1. Central Limit Theorem. We assume that the noise has the form
\[
\epsilon_{i,t} := n^{-\alpha} v_{i,t}^\gamma \gamma_{i,t},
\]
where $\alpha \geq 1/2$, the noise variance $v_{i,t}$ is time-varying, and $\gamma_i$ are iid with null-mean and unity variance. In other words we have $\epsilon_{i,t} = O_p(1/\sqrt{n})$. The parameter process is defined as the two-dimensional volatility and noise variance process $\Theta_{i} = (\sigma_i^2, v_i)$ and thus $\Theta = (T^{-1} \int_0^T \sigma_i^2 dt, T^{-1} \int_0^T v_i dt)$. Correspondingly we work locally with the QMLE considered in Xiu (2010, p. 236) and we introduce the notation for the corresponding LPE $\hat{\Theta}_n = (\sigma_n^2, \bar{v}_n)$.

We also consider the bias-corrected version of the QMLE $\hat{\Theta}_n^\text{(BC)}$, where the procedure to construct the unbiased estimator is given in Section 4.1.2. In numerical simulations under a realistic framework, this bias is not observed even with small values of $n$ (see Section 6 in Xiu (2010) and Section 5 in Clinet and Potiron (2018b)), and thus it is safe to use $\hat{\Theta}_n = (\sigma_n^2, \bar{v}_n)$ in practice.

The assumption of $\alpha \geq 1/2$ is quite restrictive in view of the related literature on the QMLE. Unfortunately in the case $\alpha < 1/2$, the techniques of this article do not apply. Xiu (2010) showed the CLT of the QMLE when $v_{i,t}$ is non time-varying and $\alpha = 0$. In the same setting, Clinet and Potiron (2018b) showed that the AVAR can be smaller when using the LPE with $B_n = 2$ fixed and documented that in finite sample the LPE was advantageous over the global QMLE. Ait-Sahalia and Xiu (2019) actually establish that the MLE is robust to noise of the form $O_p(1/n^{\gamma/2})$. Da and Xiu (2017) show the central limit theory with rate of convergence ranging from $n^{1/2}$ to $n^{1/4}$ depending on the magnitude of the noise.

However, the techniques allow us to investigate how the LPE behaves in a different asymptotics, that is, when the noise variance is $O_p(1/\sqrt{n})$ and $B_n$ tends to $+\infty$. Moreover, we allow for heteroscedasticity in noise variance. Finally, in the case where the noise variance goes to 0 at the same speed as the variance of the returns, that is, $\alpha = 1/2$, we can also retrieve the integrated variance noise. In accordance with the setting of this paper, the convergence rate of both the volatility and the noise is $n^{1/2}$.

To verify the conditions for the CLT, we use heavily the asymptotic results of the QMLE (see Theorem 6 in Xiu (2010)) and the MLE in the low-frequency asymptotics (see Proposition 1 in p. 369 of Ait-Sahalia, Mykland, and Zhang (2005)). The result is formally embedded in the following theorem.

Theorem 3 (QMLE). We define $\mathcal{F}^X_t$ the filtration generated by $X_t$.

(i) We assume that $\alpha > \frac{1}{2}$. Then, $\mathcal{F}^X_t$-stably in law as $n \to \infty$,
\[
\frac{n^{\alpha}}{T} \left( \hat{\sigma}_n^2 - T^{-1} \int_0^T \sigma_t^2 ds \right) \rightarrow \left( 6T^{-1} \int_0^T \sigma_t^4 ds \right)^\frac{1}{2} \mathcal{N}(0, 1).
\]  
(46)

(ii) When $\alpha = \frac{1}{2}$, we have $\mathcal{F}^X_t$-stable convergence in law of
\[
\frac{n^{1/2}}{T} \left( \hat{\Theta}_n^\text{(BC)} - \Theta \right) \rightarrow \text{mixed normal random variable with}
\]

zero mean and AVAR given by
\[
T^{-1} \left( \begin{array}{c}
\frac{1}{2} \int_0^T (2 v_i + \sigma_i^2) \left( \sigma_i^2 + 2 v_i + \sigma_x \sqrt{4 v_i + \sigma_i^2} \right) ds \\
\frac{1}{2} \int_0^T (\sigma_i^4 + 4 \sigma_x^3 \sqrt{4 v_i + \sigma_i^2}) ds
\end{array} \right),
\]
where \( A = \int_0^T (2 \sigma_i^4 + 4 \sigma_x^3 \sqrt{4 v_i + \sigma_i^2}) ds. \)

Remark 6 (Estimation of high-frequency covariance with the QMLE). To estimate integrated covariance under noisy observations and asynchronous observations, Aït-Sahalia, Fan, and Xiu (2010) introduced a QMLE based on a synchronization adjustment.

From Theorem 6 (p. 241) in Xiu (2010), we compute the bias-corrected estimates when using formulas (21) and (22) in the aforementioned theorem.

Theorem 4 (Powers of volatility). Let \( g \) be a nonnegative function such that
\[
| g^{(j)}(x) | \leq K (1 + | x |^{p-j}), \quad j = 0, 1, 2, 3,
\]
for some constants \( K > 0, p \geq 3. \)

(i) We assume that \( \alpha > \frac{1}{2}. \) Then, \( F_T^X \)-stably in law as \( n \to \infty,
\[
n^{1/2} \left( \hat{\Theta}_n^{(BC,1)} - \Theta \right) \to \left( 6 T^{-1} \int_0^T (g'(\sigma_i^2))^2 \sigma_i^4 ds \right)^{1/2} \times N(0, 1).
\]

(ii) When \( \alpha = \frac{1}{2}, \) we have \( F_T^X \)-stably in law that
\[
n^{1/2} \left( \hat{\Theta}_n^{(BC,2)} - \Theta \right) \to \left( T^{-1} \int_0^T (g'(\sigma_i^2))^2 \right)^{1/2} \times \left( 2 \sigma_i^4 + 4 \sigma_x^3 \sqrt{4 v_i + \sigma_i^2} \right) ds \times N(0, 1).
\]
consider the estimation of quarticity (i.e., with $g(v) = v^2$) and we note that a global QMLE would estimate $g(\int_0^T \sigma_t^2 dt)$, which is except when volatility is constant different from $\int_0^T \sigma_t^4 dt$. The extensive empirical work in Andersen, Dobrislav, and Schaumburg (2014) also indicates that the two quantities are very different in practice.

### 4.3. Estimation of Volatility and Higher Powers of Volatility Incorporating Trading Information

To incorporate all the information available in high frequency data (e.g., in addition to transaction prices, we also observe the trading volume, the type of trade, that is, buyer or seller initiated, more generally bid/ask information from the limit order book), Li, Xie, and Zheng (2016) considered the model where the noise is partially observed through a parametric function

$$Z_{t_{i,n}} = X_{t_{i,n}} + \epsilon_{i,n} = X_{t_{i,n}} + h(l_{i,n}, v) + \tilde{\epsilon}_{i,n},$$

where $l_{i,n}$ is the vector of information at time $t_{i,n}$ and $\tilde{\epsilon}_{i,n}$ is the noisy part of the original noise $\epsilon_{i,n}$. See also the related papers Chaker (2017) and Clinet and Potiron (2017, 2018c, 2018d). Here again the observation times are assumed to be regular, that is, $t_{i,n} = iT/n$.

The authors assumed that $\tilde{\epsilon}_{i,n}$ is with mean 0, finite SD and that $n \text{var}(\tilde{\epsilon}_{i,n}) \to v$, which in turn implies that $\tilde{\epsilon}_{i,n} = O_p(1/\sqrt{n})$. To embed this assumption in our LPM framework, there is no harm assuming that

$$\tilde{\epsilon}_{i,n} = n^{-\alpha} v^{1/2} \gamma_{t_{i,n}},$$

where $\alpha \geq 1/2$ and $\gamma$ are iid with null-mean and unity variance. They estimated $v$ and the underlying price as

$$\hat{v} = \arg \min_v \frac{1}{2} \sum_{i=1}^{N_n} ((Z_{t_{i,n}} - Z_{t_{i-1,n},n}) - (h(l_{i,n}, v) - h(l_{i-1,n}, v)))^2,$$

$$\hat{X}_{t_{i,n}} = Z_{t_{i,n}} - h(l_{i,n}, \hat{v}).$$

The authors then estimated the integrated volatility with

$$\text{ERV}_{\text{ext}} = \sum_{i=1}^{N_n} (\Delta \hat{X}_{t_{i,n}})^2 + 2 \sum_{i=0}^{N_n} \Delta \hat{X}_{t_{i,n}} \Delta \hat{X}_{t_{i-1,n}},$$

where $\Delta \hat{X}_{t_{i,n}} = \hat{X}_{t_{i,n}} - \hat{X}_{t_{i-1,n}}$, and show the according central limit theory. Under suitable assumptions, they obtain the optimal convergence rate $n^{1/2}$ and the AVAR when $T = 1$

$$\text{AVAR}(\text{ERV}) = 6 \int_0^1 \sigma_t^4 dt + 8v \int_0^1 \sigma_t^2 dt + 8v^2.$$

They also considered another estimator (which they call E-QMLE) which consists in using the QMLE from Xiu (2010), which we considered as a local estimator in Example 4.1, on the estimated observations $\hat{X}_{t_{i,n}}$. They indicated that the E-QMLE might yield a smaller AVAR (see their discussion on p. 38), and they report in their numerical study that its finite sample performance is comparable to ERV_{ext} (see Table 2 in p. 41). They did not investigate the corresponding central limit theory.

With the theory provided in our article, we cannot investigate the E-QMLE, but rather the E-(LPE of QMLE), that is, we apply Example 4.1 on $\hat{X}_{t_{i,n}}$. To keep notation of our paper, we denote $\tilde{\Theta}_n$ the E-(LPE of QMLE) estimator of volatility and $\tilde{\Theta}_n^{(\text{BC})}$ its bias-corrected version (i.e., E-(BC LPE of QMLE)). The AVARs obtained in Theorem 5 are the same as in Theorem 3. This is due to the fact that the estimation of $v$ is very accurate featuring $n$ as a rate of convergence and thus the pre-estimation does not impact the AVAR. This was already the case for the ERV_{ext} (see the proof of Theorem 3 in pp. 46–47 of Li, Xie, and Zheng (2016)).

Recalling that the LPE of QMLE is conjectured to be more efficient than the QMLE, in particular this implies that E-(LPE of QMLE) is also conjectured to be more efficient than E-QMLE. In Figure 1, we can see that E-(LPE of QMLE) highly improves the AVAR compared to the ERV_{ext}. The improvement gets bigger as the noise of $\tilde{\epsilon}_{i,n}$ increases. When setting the volatility and the noise variance as in the setting of the numerical study in Li, Xie, and Zheng (2016), the ratio of AVARs is equal to 0.7. When we further assume no jumps in volatility, this ratio goes to 0.2. When choosing a bigger noise variance 1.44e–07 which remains reasonable, this ratio is lower than 0.01. The overall picture is clearly in favor of the E-(LPE of QMLE). We provide the theorem of this estimator in what follows.

**Theorem 5 (E-(LPE of QMLE)).** Under Assumption A in Li, Xie, and Zheng (2016, p. 7):

(i) We assume that $\alpha > 1/2$. Then, stably in law\(^8\) as $n \to \infty$,

$$n^{1/2} \left( \tilde{\Theta}_n - T^{-1} \int_0^T \sigma_s^2 ds \right) \to \left( 6T^{-1} \int_0^T \sigma_s^4 ds \right)^{1/2} N(0, 1).$$

(ii) When $\alpha = 1/2$, we have stable convergence in law of

$$n^{1/2} \left( \tilde{\Theta}_n^{(\text{BC})} - \Theta \right) \to \left( T^{-1} \int_0^T (2\sigma_s^4 + 4\sigma_s^3 \sqrt{4v + \sigma_s^2}) ds \right)^{1/2} N(0, 1).$$

We discuss now briefly how to estimate the higher powers of volatility, that is, when $\theta_n^\ast = g(\sigma_n^2)$ with $g$ not being the identity function. We consider the estimators from Example 4.2. The difference with Example 4.2 is that this estimator is used on the estimated price $\hat{X}_{t_{i,n}}$ based on the information rather than on the raw price. The related theorem is given in what follows.

**Theorem 6 (Powers of volatility).** Under Assumption A in Li, Xie, and Zheng (2016, p.7):

(i) We assume that $\alpha > 1/2$. Then, stably in law as $n \to \infty$,

$$n^{1/2} \left( \tilde{\Theta}_n^{(\text{BC},1)} - \Theta \right) \to \left( 6T^{-1} \int_0^T (g'(\sigma_s^2))^2 \sigma_s^4 ds \right)^{1/2} N(0, 1).$$

\(^8\)Here and in the following statements, the stable convergence in law is with respect to the filtration considered in Li, Xie, and Zheng (2016).
Figure 1. A VAR of ERVext and E-(LPE of QMLE) as a function of the noise variance, that is, the variance of $\hat{\epsilon}_{t,n}$. The horizon time is set to $T = 1$ (which corresponds to 6.5 hr of intraday trading). On the left hand-side, we follow exactly the setting of the numerical study in Li, Xie, and Zheng (2016), where $\sigma_t^2 = 0.000125$ if $0.05 \leq t < 0.95$ and $\sigma_t^2 = 15 \times 0.000125$ otherwise. There is on average one observation a second, which corresponds to $n = 23,400$. On the right-hand side, the setting is the same except that we remove the jumps in volatility and consider $\sigma_t^2 = 0.000125$ for $0 \leq t \leq 1$.

(ii) When $\alpha = 1/2$, we have

$$n^{\frac{1}{2}}(\hat{\Theta}_n^{BC,2} - \Theta) \rightarrow \left(T^{-1} \int_0^T (g'(\sigma_s^2))^2 \right)^{\frac{1}{2}} \cdot \mathcal{N}(0,1).$$

4.4. Estimation of Volatility Using the Model With Uncertainty Zones

We introduce a time-varying friction parameter extension to the model with uncertainty zones introduced in Robert and Rosenbaum (2011). To incorporate microstructure noise in the model, we define $\alpha_n$ as the tick size, and the related asymptotics is such that $\alpha_n \rightarrow 0$. Correspondingly we assume that the observed price $Z_{\alpha,n}$ takes values on the tick grid (i.e., modulo of size $\alpha_n$).

We discuss first a simple version of the model with uncertainty zones, which features endogeneity in arrival times. In a frictionless market, we can assume that all the returns (which we recall to be defined as $R_{t,n} = Z_{t,n,n} - Z_{t-1,n,n}$) have a magnitude of exactly one tick, and that the next transaction will occur when the latent price process crosses the mid-tick value $X_{\alpha,n} + \frac{\alpha_n}{2}$ in case of the price goes up (or $X_{\alpha,n} - \frac{\alpha_n}{2}$ when the price goes down). We extend this toy model in what follows.

The authors introduced the discrete variables $L_{t,n}$ that stands for the absolute size, in tick number, of the next return. In other words, the next observed price has the form $Z_{t+\alpha,n} = Z_{\alpha,n} + \alpha_n L_{t,n}$. They also introduced a continuous (possibly multidimensional) time-varying parameter $\chi_t$, and assume that conditional on the past, $L_{t,n}$ takes values on $\{1, \ldots, m\}$ with

$$\mathbb{P}_{\tau_{t,n}}(L_{t,n} = k) = p_k(\chi_{\tau_{t,n}})$$

for some unknown positive differentiable with bounded derivative functions $p_k$ such that $\sum_{k=1}^m p_k = 1$.

Also, the frictions induce that the transactions will not occur exactly when the efficient process crosses the mid-tick values. For this purpose, in the notation of Robert and Rosenbaum (2012), let $0 < \nu < 1$ be the parameter that quantifies the aversion to price change. The frictionless scenario corresponds to $\nu = 0$. Conversely, the agents are very averse to trade when $\nu$ is closer to 1. If we define $X_{t}^{(\alpha)}$ as the value of $X_t$ rounded to the nearest multiple of $\alpha$, the sampling times are defined recursively as $\tau_{t,n} := 0$ and for any positive integer $i$ as

$$\tau_{t,n} := \inf \left\{ t > \tau_{t-1,n} : X_t = X_{t-1,n}^{(\alpha)} - \alpha_n \left( L_{t-1,n} - \frac{1}{2} + \eta \right) \right\}$$

or

$$X_t = X_{t-1,n}^{(\alpha)} + \alpha_n \left( L_{t-1,n} - \frac{1}{2} + \eta \right).$$

(55)

Correspondingly, the observed price is assumed to be equal to the rounded efficient price $Z_{\alpha,n} := X_{\alpha,n}$. 
In the extension of (55) when \( \eta_t \) is time-varying, we assume that the sampling times are defined recursively as \( \tau_{t,n} := 0 \) and for any positive integer \( i \) as

\[
\tau_{i,t,n} := \inf \left\{ i > \tau_{i-1,n} : \chi_{\tau_{i-1,n}} = X_{\tau_{i-1,n}} - \alpha_n \left( L_{i,n} - \frac{1}{2} + \eta_{\tau_{i-1,n}} \right) \right\},
\]

or \( X_i = \chi_{\tau_{i-1,n}} + \alpha_n \left( L_{i,n} - \frac{1}{2} + \eta_{\tau_{i-1,n}} \right) \).

(56)

The idea behind the time-varying friction model with uncertainty zones is that we hold the parameter \( \eta_t \) constant between two observations.

To express the model with uncertainty zones as a LPM, we consider that \( \theta_t^\ast := (\sigma_t^2, \eta_t, \chi_t) \). Following the definition (p. 11) in Robert and Rosenbaum (2012), we further introduce a Brownian motion \( W_t^0 \) independent of all the other quantities, and let \( \Phi \) denote the cumulative distribution function of a standard Gaussian random variable. We specify the definition of \( L_{i,n} \) related to \( W_t^0 \)

\[
g_{i,n} = \sup \{ \tau_{j,n} : \tau_{j,n} < t \},
\]

\[
L_i = \sum_{k=1}^m k \left\{ \Phi \left( \sqrt{t - g_{i,n}} \right) - \sum_{j=1}^{k-1} \sum_{j=1}^k p_j(\chi_t) \right\},
\]

and \( L_{i,n} = L_{i,n}^0 \). If \( U_{i,n} := ((W_t - W_{t-1,n})_{t \geq \tau_{i,n}}) \), and the past as \( P_{t_i,n} := (L_{i,n}, \text{sign}(R_{i,n})) \), we can deduce the form of \( F_n \) in the model.9

We provide in what follows the definition of the estimators. We are not interested in estimating directly \( \chi_t \) and thus we consider the subparameter \( \Theta := (\int_0^T \sigma_t^2 dt, \int_0^T \eta_t dt) \) to be estimated. For \( k = 1, \ldots, m \), we define

\[
N_{t,k,n}^{(a)} = \sum_{i=1}^N \mathbb{1}_{\{R_{i,n} \leq \tau_{i,n} \leq 0, |R_{i,n}| = k \alpha_n \}},
\]

\[
N_{t,k,n}^{(c)} = \sum_{i=1}^N \mathbb{1}_{\{R_{i,n} \geq \tau_{i,n} > 0, |R_{i,n}| = k \alpha_n \}},
\]

as, respectively, the number of alternations and continuations of \( k \) ticks. By alternation (continuation) of \( k \) ticks, we mean that the return magnitude is \( k \) with a direction opposite to (with the same direction as) the previous return. We define the estimator of \( \eta_t \) as

\[
\hat{\eta}_{t,n} := \sum_{k=1}^m \lambda_{t,k,n} \eta_{t,k,n},
\]

(57)

\begin{align*}
\lambda_{t,k,n} := \frac{N_{t,k,n}^{(a)} + N_{t,k,n}^{(c)}}{\sum_{j=1}^m \left( N_{t,j,n}^{(a)} + N_{t,j,n}^{(c)} \right)}, \\
u_{t,k,n} := \max \left\{ 0, \min \left\{ 1, 2 \left( \frac{N_{t,k,n}^{(c)} - N_{t,k,n}^{(a)}}{N_{t,k,n}^{(a)}} - 1 \right) + 1 \right\} \right\}.
\end{align*}

where \( N_a(t) \) is defined as the integer satisfying \( Z_{T(t_a,t)} < t < Z_{T(t_a,t_k)} \), we assume that \( C/0 = \infty \), and in particular \( \mathbb{1}_{t_k,n} = 1 \) when \( N_a(t) = 0 \). The key idea is that \( \mathbb{1}_{t_k,n} \) is consistent estimators of \( \eta_t \). Based on the friction parameter estimate, we can construct a consistent latent price estimator as

\[
\hat{X}_{t,n} := Z_{t,n} - \alpha_n (1/2 - \hat{\eta}_{t,n}) \text{sign}(R_{t,n}).
\]

The estimator of integrated volatility is obtained using the usual realized volatility estimator on the estimated price defined as

\[
\hat{R}V_{t,n} = \sum_{i=1}^N \left( \hat{X}_{t,n} - \hat{X}_{t-1,n} \right)^2.
\]

The related local estimators \( \hat{\Theta}_{t,n} := (\hat{\sigma}_{t,n}, \hat{\eta}_{t,n}) \) are constructed from local versions of \((\hat{R}V_{t,n}, \hat{\eta}_{t,n})\).

Theorem 7 (Time-varying friction parameter model with uncertainty zones). Let \( \mathbb{G}_t \) be the filtration generated by \( X_t, \chi_t \), and \( \eta_t, \mathbb{G}_t \)-stably in law as \( n \to \infty \),

\[
\alpha_n^{-1}(\hat{\Theta}_{t,n} - \Theta) \to \left( T^{-1} \int_0^T V_{\mathbb{G}_t} ds \right)^{1/2} \times \mathcal{N}(0,1),
\]

(58)

where \( V_0 \) can be straightforwardly inferred from the definition of Lemma 4.19 in p. 26 of Robert and Rosenbaum (2012).

Remark 7 (Convergence rate). Note that, equivalently, the convergence rate in (58) is \( n^{1/2} \) when \( n \) corresponds to the expected number of observations. One can consult Remark 4 in Potiron and Mykland (2017) for more details about this.

4.5. Application in Time Series: The Time-Varying MA(1)

We first specify the LPM for a general one-dimensional time series. In that case, we assume that the observation times are regular. We further assume that the returns \( R_{t,n} \) stand for time series observations. Finally, we assume that the time-varying time series can be expressed as the interpolation of \( \theta_t^\ast \) via

\[
R_{t,n} = F_n \left( \left( H_{t,n} \right)_{0 \leq t \leq \tau_{t,n}}, U_{t,n}, \theta_t^\ast \chi_{\tau_{t,n}} \right),
\]

(59)

where \( \theta_t^\ast \) is assumed to be independent of all the innovations. When \( \theta_t^\ast \) is constant, numerous time series11 are of the form (59).

11We can actually show that any time series in state space form can be expressed with a corresponding \( F_n \) function.
We now discuss the specific MA(1) representation. Several time-varying extensions are possible and we choose to work with the time-varying parameter model

\[ R_{i,n} = \mu_{t_{i-1,n}} + \sqrt{k_{t_{i-1,n}}}\lambda_{i,n} + \beta_{t_{i-1,n}}\sqrt{k_{t_{i-1,n}}}\lambda_{i-1,n}, \]

where \( \lambda_{i,n} \) are standard normally-distributed white noise error terms, and \( k_i \) is the time-varying variance. The three-dimensional parameter is defined as \( \theta^*_n := (\mu_i, \beta_i, k_i) \in \mathbb{R}^2 \times \mathbb{R}_+^* \). We fix both the innovation and the past equal to the white noise \( U_{i,n} = \lambda_{i,n} \) and \( P_{t_{i,n}} = \lambda_{i,n} \). We have thus expressed the MA(1) as a LPM.

We discuss how to estimate the parameters in what follows. For simplicity, we assume that \( \mu_t = 0 \). The target quantity is thus equal to \( \Theta = (\int_0^T \beta_t dt, \int_0^T \kappa_t dt) \). The local estimator is the MLE (see Hamilton 1994, sec. 5.4). On each block (of size \( h_n \)), the MLE bias is of order \( h_n^{-1} \) (Tanaka 1984) and thus the bias condition (23) is not satisfied. Nonetheless, we can correct for the bias up to the order \( O(h_n^{-2}) \) as follows. We define the bias-corrected estimator as

\[ \hat{\Theta}_{i,n}^{(BC)} = \hat{\Theta}_{i,n} - \hat{b}(\hat{\Theta}_{i,n}, h_n), \]

where the bias function \( b(\hat{\theta}, h) \) can be derived following the techniques in Tanaka (1984). In particular this implies that the bias-corrected estimator satisfies the bias condition if \( h_n \) is chosen such that \( n^{1/4} = o(h_n) \). In practice this bias can be obtained by Monte Carlo simulations (see our simulation study).

In the parametric case and in a low frequency asymptotics where \( T \to \infty \) and observations times are \( 0, \Delta, \ldots, T = n\Delta \) with \( \Delta > 0 \), known results (see, e.g., the proof of Proposition I in pp. 391–393 of Ait-Sahalia, Mykland, and Zhang (2005)) show that the AVAR of the MA(1) is such that

\[ n^{1/2}(\tilde{\beta}, \tilde{\kappa}) - (\beta, \kappa) \to \begin{pmatrix} 1 - \beta^2 & 0 \\ 0 & 2\kappa^2 \end{pmatrix}^{1/2} N(0, 1). \]

The following theorem provides the time-varying version of the asymptotic theory when \( T \) is fixed.

**Theorem 8 (Time-varying MA(1)).** Let \( \mathcal{F}_{i}^T \) be the filtration generated by \( \theta^*_n \). We assume that \( n^{1/4} = o(h_n) \) and Condition (P2). Then, \( \mathcal{F}_{i}^T \)-stably in law as \( n \to \infty \),

\[ n^{1/2}(\hat{\Theta}_{i,n}^{(BC)} - \Theta) \to \left( T^{-1} \int_0^T (1 - \beta^2_s) ds, 0 \right) \left( T^{-1} \int_0^T 2\kappa^2_s ds \right)^{1/2} \times N(0, 1). \]

### 4.6. Further Examples

Two further examples include our own recent work. Potiron and Mykland (2017) introduced a bias-corrected Hayashi–Yoshida estimator (Hayashi and Yoshida 2005) of the high-frequency covariance and showed the corresponding CLT under endogenous and asynchronous observations. To model duration data, Clinet and Potiron (2018a) built a time-varying Hawkes self-exciting process, derived the bias-corrected MLE and showed the CLT of the corresponding LPE.

### 4.7. Discussion

We provide in what follows a discussion on the efficiency and robustness of the specific examples considered in this section. The subsequent techniques may also be useful to tackle other examples from the literature.

#### 4.7.1. Efficiency

There are many problems where \( n^{1/2} \) is rate-optimal from Gloeter and Jacod (2001), such as all the examples considered in this section. In addition, the local feature of the technology should yield efficiency in case the parametric estimator is efficient itself. This is the case of (47) in Example 4.1, Theorem 4(ii) in Example 4.2, Theorem 5(ii) and Theorem 6(ii) in Example 4.3, Theorem 8 in Example 4.5, where the parametric estimator achieves the Cramér–Rao bound of efficiency locally.

In the case of (46) in Example 4.1, that is, when estimating volatility assuming that the noise is very small \( \epsilon_{i,n} = o_p(1/\sqrt{n}) \), the AVAR is equal to \( 6T^{-1} \int_0^T \sigma^4_s ds \), whereas the efficient bound \( 2T^{-1} \int_0^T \sigma^4_s ds \) is attained by the RV . This increases the variance by a factor of 3, which is also observed on the MLE (when assuming the volatility is constant) when misspecified on a model which does not incorporate microstructure noise (see, e.g., Barndorff-Nielsen et al. 2008, sec. 2.4, pp. 1486–1487).

#### 4.7.2. Robustness to Drift and Jumps in the Latent Price Process

We focus on the specific case where the observations are related to a latent continuous-Itô price model \( dX_t = \int_0^t \sigma_u dW_u \), as in Examples 4.1–4.4 (Example 4.5 considers a time series without any underlying price process). We discuss how we can add a drift and jumps in \( X_t \) in those examples.

We first show how to add a drift component. By Girsanov theorem, in conjunction with local arguments (see, e.g., Mykland and Zhang 2012, pp. 158–161), we can weaken the price and volatility local-martingale assumption by allowing them to follow an Itô-process (of dimension 2 in case of volatility or powers of volatility estimation), with a volatility matrix locally bounded and locally bounded away from 0, and drift which is also locally bounded.

It is also easy to see that we can allow for finite activity jumps in \( X_t \). To do that, we assume that \( \hat{\Theta}_{i,n} \) is taking values on a compact set.\(^{12}\) Consider \( J_n \subset \{1, \ldots, B_n\} \) the set of blocks where there is at least one jump in \( X_t \). As the number of blocks \( B_n \to \infty \), the cardinality of \( J_n \) is at most finite, and thus we have that

\[ \frac{1}{T} \sum_{i=1}^N \hat{\Theta}_{i,n} \Delta T_{i,n} \approx \frac{1}{T} \sum_{i \in J_n} \hat{\Theta}_{i,n} \Delta T_{i,n}. \]

It is then immediate to adapt the proof of the CLT. On the other hand, if infinitely many jumps are possible in both the price process and the parameter, the theoretical development is beyond the scope of this paper.

#### 4.7.3. Robustness to Jumps in \( \theta_n^* \)

By a similar reasoning as for when adding jumps in \( X_t \), the techniques of this article are robust to jumps (of finite activity) in \( \theta_n^* \) in all our examples.

\(^{12}\)The MLE is always performed on a compact set, so the assumption is trivially satisfied in that case, which corresponds to Examples 4.1–4.3. Moreover, the estimator of \( \eta \) in Example 4.4 is bounded by definition, but one would need to bound the volatility estimator to apply the technique.
4.7.4. Nonregular Observation Times. We also assume here that there is a latent price process and reason about the type of observation times which falls into the LPM. We consider first the time deformation of Barndorff-Nielsen et al. (2008, sec. 5.3, pp. 1505–1507). To express their setting as a LPM, we assume that the observation times are of the form
\[ \tau_{i,n} = \Gamma_{i}/(nT), \]
where \( \Gamma_{i} \) is a stochastic process satisfying \( \Gamma_{i} = \int_{0}^{1} \Gamma_{2} du \), with \( \Gamma_{2} \) a strictly positive parameter of the LPM. We can then construct a (change of time) process \( \tilde{X}_{t} = X_{\Gamma_{t}} \) so that for \( \tilde{X}_{t} \) the observations are regular. In view of Dambis Dubins-Schwarz theorem (see, e.g., Revuz and Yor 1999, Theorem 1.6, p. 181) we have that \( [X]_{T} = [\tilde{X}]_{\Gamma_{T}} \). In addition, it is immediate to see that Condition (T) and (42) hold in that case.

Alternatively one can assume that the quadratic variation of time (see, e.g., Mykland and Zhang 2006, Assumption A, p. 1939) exists and that observation times are independent of the price process. Under suitable assumptions, we can also show that Condition (T) and (42) hold.

Our setting can actually allow for (some kind of) endogenous stopping times as in the case of the model with uncertainty zones detailed in Example 4.4. The type of endogeneity is such that there is no asymptotic bias in the related central limit theorem.

Finally, the model allows for endogenous observation times in the general multidimensional HBT model introduced in Potiron and Mykland (2017). In that case, the central limit theorem features an asymptotic bias.13

4.7.5. Estimating Time-Varying Functions of \( \theta_{t}^{*} \). Another nice corollary about the introduced theory is that we can obtain the central limit theorem of the powers of the integrated parameter \( g(t, \theta_{t}^{*}) \) for \( g \) smooth enough when using the local estimates \( g(\tilde{\Gamma}_{i−1:n}, \tilde{\Theta}_{i:n}) \). Essentially, the proof uses on each block a Taylor expansion as in the delta method. We apply the technique on the local QMLE in Example 4.2 and on an adapted estimator from Li, Xie, and Zheng (2016) in Example 4.3 to estimate the higher powers of volatility.

5. NUMERICAL STUDY: ESTIMATION OF VOLATILITY WITH THE QMLE

5.1. Goal of the Study

To investigate the finite sample performance of the LPE, we consider the local QMLE introduced in Section 4.1. The goal of the study is 2-fold. First, we want to investigate how the LPE performs compared to the global QMLE. Second, we want to discuss about the choice of the number of blocks \( B_{n} \) in practice. Complementary simulation results can be found in Clinet and Potiron (2018b).

5.2. Model Design

We perform Monte Carlo simulations of \( M = 1000 \) days of high-frequency observations where the related horizon time is set to \( T = 1/252 \) (i.e., annualized). One working day stands for 6.5 hr of trading activity, which can also be expressed as 23,400 sec. We consider three high-frequency sampling frequency scenarios: every second, every other second, and every 3 sec.

We perform local QMLE with number of blocks ranging from \( B_{n} = 1 \) (i.e., the global QMLE case) to \( B_{n} = 20 \). The corresponding number of observations per block ranges from \( h_{n} = 1170 \) to \( h_{n} = 23,400 \) in the case of 1-sec sampling frequency, from \( h_{n} = 585 \) to \( h_{n} = 11,700 \) if we sample ever other second, and from \( h_{n} = 390 \) to \( h_{n} = 7800 \) when subsampling every 3 sec. Note that the minimal number of observations per block remains reasonable in view of the finite sample performance of the global QMLE (see the numerical study in Xiu (2010)).

We bring forward the Heston model with U-shape intraday seasonality component and jumps in volatility as
\[ dX_{t} = bdt + \sigma_{d}dW_{t}, \]
\[ \sigma_{t} = \sigma_{t−U} \phi \sigma_{SV}, \]
where
\[ \sigma_{t,U} = C + A e^{-\alpha t/T} + D e^{-\beta (1−t/T)} - \beta \sigma_{t−U} 1_{\{t \geq \tau\}}, \]
\[ d\sigma_{SV} = \alpha (\bar{\sigma}^{2} - \sigma_{SV}^{2}) dt + \delta \sigma_{SV} dW_{t}, \]
where the parameters are set to \( b = 0.03, C = 0.75, A = 0.25, D = 0.89, a = 10, c = 10, \) the volatility jump size parameter \( \beta = 0.5, \) the volatility jump time \( \tau \) follows a uniform distribution on \([0, T]\), \( \alpha = 5, \bar{\sigma}^{2} = 0.1, \delta = 0.4, \) \( W_{t} \) is a standard Brownian motion such that \( d\langle W, W \rangle_{t} = \varphi dt \), \( \bar{\sigma} = -0.75, \sigma_{SV}^{2} \) is sampled from a Gamma distribution of parameters \((2a\bar{\sigma}^{2}/\delta^{2}, \delta^{2}/2\alpha)\), which corresponds to the stationary distribution of the CIR process. For further reference, see Clinet and Potiron (2018b). The model is almost the same as that of Andersen, Dobrev, and Schaumburg (2012). Finally, the noise is assumed normally distributed with zero-mean and constant variance \( \nu \) set so that the noise to signal ratio defined as
\[ \xi^{2} = \frac{\sigma_{0}^{2}}{\sqrt{T \int_{0}^{T} \sigma_{u}^{2} du}} \]
is equal to \( \xi^{2} = 0.0001 \).

5.3. Results

Table 1 reports the sample bias, SD, and the RMSE of the local quasi maximum likelihood volatility estimator. The number of blocks ranges from \( B_{n} = 1 \), which corresponds to the global QMLE, to \( B_{n} = 20 \). Regardless of the sampling frequency, the numerical experiment results are quite similar. There is a very small sample bias (the bias to SD ratio magnitude is around 0.03), which increases with the number of blocks while staying very small, all of which hinting that the it is not necessary to use a bias correction of the local QMLE in practice. The SD decreases and then stays (roughly)

---

13Details about the model can be found in a previous version of the manuscript circulated under the name “Estimating the Integrated Parameter of the Locally Parametric Model in High-Frequency Data.”
stable. The picture for the RMSE is the same, all of this very much in line with the fact that almost all the theoretical gain is already obtained in the case of $B_n = 1$ (i.e., the global QMLE case) to $B_n = 20$

<table>
<thead>
<tr>
<th>Samp. freq.</th>
<th>1 sec.</th>
<th>1 sec.</th>
<th>1 sec.</th>
<th>2 sec.</th>
<th>2 sec.</th>
<th>2 sec.</th>
<th>3 sec.</th>
<th>3 sec.</th>
<th>3 sec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. blocks</td>
<td>Bias</td>
<td>SD</td>
<td>RMSE</td>
<td>Bias</td>
<td>SD</td>
<td>RMSE</td>
<td>Bias</td>
<td>SD</td>
<td>RMSE</td>
</tr>
<tr>
<td>1</td>
<td>−2.398</td>
<td>8.158</td>
<td>8.162</td>
<td>−2.503</td>
<td>10.813</td>
<td>10.814</td>
<td>−0.492</td>
<td>11.798</td>
<td>11.798</td>
</tr>
<tr>
<td>2</td>
<td>−2.614</td>
<td>7.938</td>
<td>7.943</td>
<td>−3.604</td>
<td>10.634</td>
<td>10.640</td>
<td>−0.700</td>
<td>11.642</td>
<td>11.642</td>
</tr>
<tr>
<td>3</td>
<td>−2.882</td>
<td>7.820</td>
<td>7.825</td>
<td>−4.041</td>
<td>10.537</td>
<td>10.544</td>
<td>−0.600</td>
<td>11.615</td>
<td>11.615</td>
</tr>
</tbody>
</table>

NOTE: The number of seconds for one working day is 23,400. The number of Monte Carlo simulations is 1000. Three sampling frequencies are considered: every second, every 2 sec, and every 3 sec.

6. CONCLUSION

In this article, we have introduced a general framework to provide theoretical tools to build central limit theorems of convergence rate $n^{1/2}$ in a time-varying parameter model. We have applied successfully the method to investigate estimation of volatility (possibly under trading information), higher powers of volatility, the time-varying parameters of the model with uncertainty zones and the MA(1). This allowed us to obtain estimators robust to time-varying quantities, more efficient and/or new estimators of quantities (such as in the case of higher powers of volatility).

Subsequently, we believe that many other examples can be solved using the framework of our article, which is simple and natural. This was successfully done in our related papers Potiron and Mykland (2017) and Clinet and Potiron (2018a). In those instances, the regular conditional distribution trick significantly simplified the work of the proofs.

SUPPLEMENTARY MATERIALS

The supplementary materials consist of four distinct sections. First, we investigate consistency in a simple model. Second, the proofs are provided. Third, an additional numerical study, i.e. time-varying MA(1), is explored. Finally, one can find an empirical illustration in the model with uncertainty zones.

ACKNOWLEDGMENTS

We are indebted to the editor, Todd Clark, two anonymous referees, and an anonymous associate editor, Simon Clinet, Takaki Hayashi, Dacheng Xiu, participants of the seminars in Berlin and Tokyo and conferences in Osaka, Toyama, the SoFie annual meeting in Hong Kong, the PIMS meeting in Edmonton for valuable comments, which helped in improving the quality of the paper.

FUNDING

Financial support from the National Science Foundation under grant DMS 14-07812 and Japanese Society for the Promotion of Science Grant-in-Aid for Young Scientists (B) No. 60781119 are greatly acknowledged.

REFERENCES

