Estimation of integrated quadratic covariation with endogenous sampling times

Yoann Potiron a,*, Per A. Mykland b,1

a Faculty of Business and Commerce, Keio University, 2-15-45 Mita, Minato-ku, Tokyo, 108-8345, Japan
b Department of Statistics, The University of Chicago, 5734 S. University Avenue Chicago, IL 60637, United States

ARTICLE INFO

Article history:
Received 30 January 2016
Received in revised form 21 June 2016
Accepted 27 October 2016
Available online 10 November 2016

JEL classification:
C01
C02
C13
C14
C22
C32
C58

Keywords:
Asymptotic bias
Asynchronous times
Endogenous model
Hayashi–Yoshida estimator
High-frequency data
Quadratic covariation
Time endogeneity

ABSTRACT

When estimating high-frequency covariance (quadratic covariation) of two arbitrary assets observed asynchronously, simple assumptions, such as independence, are usually imposed on the relationship between the prices process and the observation times. In this paper, we introduce a general endogenous two-dimensional nonparametric model. Because an observation is generated whenever an auxiliary process called the hitting boundary process with time process (HBT) model, we establish a central limit theorem for the Hayashi–Yoshida (HY) estimator under HBT in the case where the price process and the observation price process follow a continuous Itô process. We obtain an asymptotic bias. We provide an estimator of the latter as well as a bias-corrected HY estimator of the high-frequency covariance. In addition, we give a consistent estimator of the associated standard error.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Covariation between two assets is a crucial quantity in finance. Fundamental examples include optimal asset allocation and risk management. In the past few years, using the increasing amount of high-frequency data available, many papers have been published about estimating this covariance. Suppose that the latent log-price of two arbitrary assets $X_t = (X_t^{(1)}, X_t^{(2)})$ follows a continuous Itô process

\[ \begin{align*}
    dX_t^{(1)} &= \mu_t^{(1)} dt + \sigma_t^{(1)} dW_t^{(1)}, \\
    dX_t^{(2)} &= \mu_t^{(2)} dt + \sigma_t^{(2)} dW_t^{(2)},
\end{align*} \]

where $\mu_t^{(1)}, \mu_t^{(2)}, \sigma_t^{(1)}, \sigma_t^{(2)}$ are random processes, and $W_t^{(1)}$ and $W_t^{(2)}$ are standard Brownian motions, with (random) high-frequency correlation $d\langle W^{(1)}, W^{(2)} \rangle_t = \rho_t dt$. Econometrics usually seeks to infer the integrated covariation

\[ \langle X_t^{(1)}, X_t^{(2)} \rangle_t = \int_0^t \rho_u \sigma_u^{(1)} \sigma_u^{(2)} du. \]

Earlier results were focused on estimating the integrated variance of a single asset, starting from the probabilistic point of view (Genon-Catalot and Jacod, 1993; Jacod, 1994). Barndorff-Nielsen and Shephard (2001, 2002) introduced the problem in econometrics. Adapted to two dimensions, if each process is observed simultaneously at (possibly random) times $\tau_{0,n} := 0, \tau_{1,n}, \ldots, \tau_{n,n}$ the
realized covariance \( \{X^{(1)}, X^{(2)}\}_t \) is defined as the sum of cross log returns
\[
\{X^{(1)}, X^{(2)}\}_t = \sum_{i, j, k} \Delta X_{t_i}^{(1)} \Delta X_{t_j}^{(2)},
\]
where for any positive integer \( i \), \( \Delta X_{t_i}^{(k)} = X_{t_i}^{(k)} - X_{t_{i-1}}^{(k)} \) corresponds to the increment of the \( k \)th process between the last two sampling times. As the observation intervals \( \Delta t_{i,n} \) get closer (and the number of observations \( N \) goes to infinity), \( \{X^{(1)}, X^{(2)}\}_t \) converges in distribution (see e.g. Theorem 1.4.47 in Jacod and Shiryaev (2003)). Furthermore, when the observation times \( t_{i,n} \) are independent of the prices processes \( X_{i} \), its estimation error follows a mixed normal distribution (Jacod and Trotter, 1958; Zhang, 2001; Mykland and Zhang, 2006). This gives us insight on how to estimate the integrated covariation. However, in practice, these two assumptions are usually not satisfied. The observation times of the two assets are rarely synchronous and there is endogeneity in the price sampling times.

The first issue has been studied for a long time. The lack of synchronicity often creates undesirable effects in inference. If we sample at very high frequencies, we observe the Epps effect (Epps, 1979), i.e. the correlation estimates are drastically decreased compared to an estimate with sparse observations. Hayashi and Yoshida (2005) introduced the so-called Hayashi–Yoshida estimator (HY)
\[
\{X^{(1)}, X^{(2)}\}_{HY} = \sum_{i, j, k} \Delta X_{t_i}^{(1)} \Delta X_{t_j}^{(2)} 1_{|t_i - t_j| < \epsilon},
\]
where \( t_{i,k} \) are the observation times of the \( k \)th asset. Note that if the observations of both processes occur simultaneously, \((3)\) and \((4)\) are equal. The consistency of this estimator was achieved in Hayashi and Yoshida (2005) and Hayashi and Kusuoka (2008). The corresponding central limit theorems were investigated in Hayashi and Yoshida (2008, 2011) under strong predictability of observation times, which is a more restrictive assumption than only assuming they are stopping times but still allows some dependence between prices and observation times. Recently, Koike (2014, 2015) extended the pre-averaged Hayashi–Yoshida estimator first under predictability of observation times, and then under a more general endogenous setting of stopping times. Other examples of high-frequency covariance estimators can be found in Zhang (2011), Barndorff-Nielsen et al. (2011), Aït-Sahalia et al. (2010), Christensen et al. (2010, 2013).

In a general one-dimensional endogenous model, the asymptotic behavior of the realized volatility \((3)\) has been investigated in the case of sampling times given by hitting times on a grid ( Fukasawa, 2010a; Robert and Rosenbaum, 2011, 2012; Fukasawa and Rosenbaum, 2012). Due to the regularity of those three models (see the discussion in the latter paper), they do not obtain any bias in the limit distribution of the normalized error. Also, the case of strongly predictable stopping times is treated in Hayashi and Yoshida (2011). Finally, two general results (Fukasawa, 2010b; Li et al., 2014) showed that we can identify and estimate the asymptotic bias.

The primary goal of this paper is to bias-correct the HY. Note that estimating the bias is more challenging than in the volatility case because observations are asynchronous. In particular, the estimator will involve a quantity that can be considered as the tricity of Li et al. (2014), but with a more intricate definition because of the asynchronicity in sampling times. This new definition can be seen as an analogy with the generalization of the RV estimator \((3)\) by the HY estimator \((4)\).

Another very important issue to address is the estimation of the asymptotic standard deviation. First, because the model is more general than in the no-endogeneity work, the theoretical asymptotic variance will be different. Consequently, a new variance estimator, which takes into proper account the endogeneity, will be given.

The authors want to take no position on the joint distribution of the log-return and the next observation time that corresponds to an asset price change because they know that their unknown relationship is most likely contributing to the bias and the variance of the high-frequency covariance’s estimate when we (wrongly) assume full independence between the price process and observation times. For this purpose, they introduce the hitting boundary process with time process (HBT) model.

Finally, techniques developed in the proofs are innovative in the sense that they reduce the normalized error of the Hayashi–Yoshida estimator to a discrete process, which is locally a uniformly ergodic homogeneous Markov chain. Thus, the problem can be solved locally, and because we assume that the volatility of assets is continuous, the error of approximation between the local Markov structure and the real structure of the normalized error vanishes asymptotically. This technique is not problem-specific, and it can very much be applied to other estimators dealing with temporal data.

The paper is organized as follows. We introduce the HBT model in Section 2. Examples covered by this model are given in Section 3. The main theorem of this work, the limit distribution of the normalized error is given in Section 4. Estimators of the asymptotic bias and variance are provided in Section 5. We carry out numerical simulations in Section 6 to corroborate the theory. Proofs are developed in the Appendix.

2. Definition of the HBT model

We first introduce the model in 1-dimension. We assume that for any positive integer \( i \), \( \tau_{i+1} \) is the next arrival time (after \( \tau_i \)) that corresponds to an actual change of price. In particular, several trades can occur at the same price \( Z_t \) between \( \tau_i \) and \( \tau_{i+1} \), but no trade can occur with a price different than \( Z_t \) before \( \tau_{i+1} \). We also assume that \( X_t \) is the efficient (log) price of the security of interest. In addition, we assume that the observations are noisy and that we observe \( Z_t := X_t + \epsilon_t \), where the microstructure noise \( \epsilon_t \) can be expressed as a known function of the observed prices \( Z_0, \ldots, Z_t \).

As an example, Robert and Rosenbaum (2012) showed in (2.3) in p. 5 that the model with uncertainty zones can be written with that noise structure if we assume that we know the friction parameter \( \eta \). Finally, we define \( \alpha > 0 \) as the tick size, and we assume that the observed price \( Z_t \) lies on the tick grid, i.e. there exist positive integers \( m \) such that \( Z_t := m \alpha t \).

Empirically, no economical model based on rational behaviors of agents on the stock markets, that shed light on the relationship between the efficient return \( \Delta X_t \) and time before the next price change \( \Delta t_i = \tau_i - \tau_{i-1} \), has won unanimous support. When arrival times are independent of the asset price, it follows directly from the continuous Itô-assumption that the dependence structure is such that the return \( \Delta X_t \) is a function of \( \Delta t \). The longer we wait, the bigger the variance of the return is expected to be. In this paper, we take the opposite point of view by building a model in which \( \tau_i \) is defined as a function of the efficient price path. For that purpose, we define the observation time process \( \tau_i^{(1)} \) that will drive the observation times. We also define the down process \( d_i(s) \) and the up process \( u_i(s) \). Note that for any \( t \geq 0 \), we assume that \( d_i \) and \( u_i \) are functions on \( \mathbb{R}^+ \). We also assume that the down process takes only negative values and that the up process takes only positive values. A new observation time will be generated whenever one of those two processes is hit by the increment of the observation time process. Then, the increment of the observation time process will start again from \( 0 \), and the next observation time
will be generated whenever it hits the up or the down process. Fig. 1 illustrates the HBT model. Formally, we define \( r_\alpha := 0 \) and for any positive integer \( i \) as

\[
\tau_i := \inf\left\{ t > \tau_{i-1} : \Delta X^{(i)}_t \not\in [d_i (t - \tau_{i-1}), u_i (t - \tau_{i-1})] \right\},
\]

where \( \Delta Y_{[a,b]} := Y_b - Y_a \). Note that if the observation time process \( X^{(i)}_t \) is equal to the price process \( X_t \), then the time process will hit the up process (down process). Note also that if the time process, the up process and the down process are independent of the efficient price process, then the arrival times are independent of the efficient price process. We assume that the two-dimensional process \( (X_t, X^{(i)}_t) \) is an Itô-process. Section 3.1 provides examples of the literature identifying the observation time process, the down process and the up process.

Generalizing to two dimensions is straightforward. We define \( X^{(k,i)}_t \) for \( k = 1, 2 \) to be the observation process associated with the \( k \)th price process, \( u^{(k)}_i \) the up process, \( d^{(k)}_i \) the down process, and the arrival times \( \tau^{(k,i)}_i \) generated by (5). We also define the four dimensional process \( Y_t := (X^{(1)}_t, X^{(2)}_t, X^{(1,1)}_t, X^{(1,2)}_t) \), and assume \( Y_t \) follows an Itô-process with volatility

\[
\sigma_t := \begin{bmatrix}
\sigma^{1,1}_t & \sigma^{1,2}_t & \sigma^{1,3}_t & \sigma^{1,4}_t \\
\sigma^{2,1}_t & \sigma^{2,2}_t & \sigma^{2,3}_t & \sigma^{2,4}_t \\
\sigma^{3,1}_t & \sigma^{3,2}_t & \sigma^{3,3}_t & \sigma^{3,4}_t \\
\sigma^{4,1}_t & \sigma^{4,2}_t & \sigma^{4,3}_t & \sigma^{4,4}_t
\end{bmatrix}.
\]

In particular, we have \( dY_t = \mu_t dt + \sigma_t dW_t \), where \( W_t \) is a four dimensional standard Brownian motion (for \( i = 1, \ldots, 4 \) and \( j = 1, \ldots, 4 \) such that \( i \neq j \), \( W_t^{(ij)} \) is independent of \( W_t^{(ij)} \). If we set \( \zeta_i := \sigma_i \sigma_i^t \), then the integrated covariance (or quadratic covariation) process is given by \( \langle Y, Y \rangle_t = \int_0^t \zeta_s ds \). Let \( \rho_t \) be the associated correlation process of \( Y_t \), i.e. for \( i = 1, \ldots, 4 \) and \( j = 1, \ldots, 4 \) we set \( \rho^{(ij)}_t = \zeta^{(i)}_t \zeta^{(j)}_t^{-1} \). Finally, it is useful sometimes to see \( Y_t \) as a four dimensional vector expressed as in Eqs. (1) and (2). For \( k = 1, \ldots, 4 \) we define the volatility of the \( k \)th process as \( \sigma^{(k)}_t := (\zeta^{(k)}_t)^{1/2} \), we can thus express \( Y_t \) as

\[
dY^{(k)}_t = \mu^{(k)}_t dt + \sigma^{(k)}_t dB^{(k)}_t
\]

where \( B^{(k)}_t \) is a standard Brownian motion, which typically depends on \( B^{(i)}_t \) for \( i = 1, \ldots, 4 \).

3. Examples

We insist on the fact that estimators of covariance and associated asymptotic variance given in this paper do not require any knowledge of the structure of the observation time process, the up process and the down process. Nonetheless, for financial and economic interpretation purposes, the reader might be interested in getting an idea on how those processes behave in practice. We provide in this section several examples from the literature as well as possible extensions of the model with uncertainty zones of Robert and Rosenbaum (2011) that can be expressed as HBT models.

3.1. Endogenous models contained in the HBT class

Example 1 (Hitting Constant Boundaries). The simplest endogenous semi-parametric model we can think of is a model where the time process \( X^{(i)}_t \) is equal to the price process \( X_t \), times are generated by hitting a constant barrier. Formally, it means that there exists a two-dimensional parameter \( (\theta_0, \theta_2) \) such that the up process is equal to \( \theta_0 \) and the down process is equal to \( \theta_2 \). We do not assume noise in that model.

Example 2 (Hitting Constant Boundaries of the Tick Size). One issue with Example 1 is that the efficient price \( X_t \), which is observed because no microstructure noise is assumed in the model, is not necessarily a modulo of the tick size \( \alpha \) if \( \theta_0 \) and \( \theta_2 \) are not multiples of \( \alpha \). To make Example 1 feasible in practice, we assume here that the constant barriers \( \theta_0 \) and \( \theta_2 \) are respectively equal to the tick size \( \alpha \) and its additive inverse \(-\alpha\). We also assume that \( Z_{\tau_i} := X_{\tau_i} \).

Example 3 (Hitting Constant Boundaries of the Jump Size). The issue with Example 2 is that the absolute jump size of the observed price \( Z_t \) is \( \alpha \). On the contrary, in practice the absolute jump size can actually be bigger than the tick size \( \alpha \). In the notation of Robert and Rosenbaum (2011), for any positive integer \( i \), we introduce discrete variables \( L_i \) which correspond to the observed price jump’s tick number between \( \tau_i \) and \( \tau_{i+1} \), with \( L_i \geq 1 \). We assume that \( L_i \) is bounded. The arrival times are defined recursively as \( \tau_0 = 0 \) and for any positive integer \( i \) as

\[
\tau_i := \inf\left\{ t > \tau_{i-1} : X_{\tau_{i-1}} - L_{i-1} \alpha \leq X_t < X_{\tau_{i-1}} + L_{i-1} \alpha \right\}.
\]

We assume that \( L_i \) are IID and independent of the other quantities. We finally assume that \( Z_{\tau_i} := X_{\tau_i} \). The up and down processes are piecewise constant in \( r \) and constant in \( s \), defined for any \( s \geq 0 \) as

\[
d_i(s) = -L_{i-1} \alpha \quad \text{for } t \in (\tau_{i-1}, \tau_i]
\]

\[
u_i(s) = L_{i-1} \alpha \quad \text{for } t \in (\tau_{i-1}, \tau_i].
\]

Example 4 (Model with Uncertainty Zones). We go one step further than Example 3 and introduce now the model with uncertainty zones of Robert and Rosenbaum (2011). In a frictionless market, we can assume that a trade with change of price \( Z_t \) will occur whenever the efficient price process crosses one of the mid-tick values \( Z_{\tau_i} + \frac{\alpha}{2} \) or \( Z_{\tau_i} - \frac{\alpha}{2} \). In that case, if the efficient price process hits the former value, we would observe an increment of the observed price \( Z_{\tau_i} = Z_{\tau_i} + \alpha \) and if it hits the former value, we would observe a decrement \( Z_{\tau_i} = Z_{\tau_i} - \alpha \). There are two reasons why in practice such a frictionless model is too simplistic. The first reason is that the absolute value of the increment (or the decrement) of the observed price can be bigger than the tick size \( \alpha \) and was already pointed out in Example 3. We will thus keep the notation \( L_i \) in this example. The second reason is that the frictions induce that the transaction will not exactly occur.
when the efficient process is equal to the mid-tick values. For this purpose in the notation of Robert and Rosenbaum (2012), let $0 < \eta < 1$ be a parameter that quantifies the aversion to price changes of the market participants. If we let $X_t^{(\omega)}$ be the value of $X_t$ rounded to the nearest multiple of $\omega$, the sampling times are defined recursively as $\tau_0 := 0$ and for any positive integer $i$ as

$$\tau_i := \inf\{ t > \tau_{i-1} : X_t = X_t^{(\omega)} - \alpha \left( I_{t-1} - \frac{1}{2} + \eta \right)$$

or $X_t = X_t^{(\omega)} + \alpha \left( I_{t-1} - \frac{1}{2} + \eta \right) \}$

The observed price is equal to the rounded efficient price $Z_t := X_t^{(\omega)}$. The time process $X_t^{(\omega)}$ is again equal to the price process $X_t$ itself in this model. The up and down processes are piecewise constant in $t$ and constant in $\omega$, defined for any $s \geq 0$ as

$$d_i(s) = -L_{i-1} \alpha 1_{[X_{t,i-1} < X_{t,i}]} - (2\eta + L_{i-1} - 1) \alpha 1_{[X_{t,i} < X_{t,i-1}]}$$

for $t \in (\tau_{i-1}, \tau_i)$

$$u_i(s) = L_{i-1} \alpha 1_{[X_{t,i-1} < X_{t,i}]} + (2\eta + L_{i-1} - 1) \alpha 1_{[X_{t,i} < X_{t,i-1}]}$$

for $t \in (\tau_{i-1}, \tau_i)]$

where $1_k$ is the indicator function of $A$. Note that in the case where $\eta = \frac{1}{2}$, we are back to Example 3.

**Example 5** (Times Generated by Hitting an Irregular Grid Model). The fourth model we are looking at is called times generated by hitting an irregular grid model. We follow the notation of Fukasawa and Rosenbaum (2012) and consider the irregular grid $g_i = \lceil pk \rceil_{k \in \mathbb{Z}}$, with $p_k = p_{k+1}$. We set $\tau_0 = 0$ and for $i \geq 1$

$$\tau_i := \inf\{ t > \tau_{i-1} : X_t \in g - \{ X_{t,i-1} \} \},$$

where $g - \{ X_{t,i-1} \}$ is the set obtained by removing $\{ X_{t,i-1} \}$ from $g$. We can rewrite it as an element of the HBT model where the time process is equal to the price process, and for all $s \geq 0$ the up and down processes are defined as

$$d_i(s) = p_{k-1} - p_k \text{ for } t \in (\tau_{i-1}, \tau_i)$$

$$u_i(s) = p_{k+1} - p_k \text{ for } t \in (\tau_{i-1}, \tau_i),$$

where $k$ is the (random) index such that $p_k = X_{t,i-1}$.

**Example 6** (Structural Autoregressive Conditional Duration Model). There have been several drafts for this model. We follow here a former version (Renault et al., 2009), because we can directly express it as an element of the HBT model. In the structural autoregressive conditional duration model, the time $\tau_i$ when the next event occurs is given by $\tau_0 = 0$ and for $i > 0$

$$\tau_i := \inf\{ t > \tau_{i-1} : A_t - A_{t,i-1} = \tilde{d}_{t,i-1} \text{ or } A_t - A_{t,i-1} = \tilde{c}_{t,i-1} \}$$

(6)

where $A_t$ is a standard Brownian motion (not necessarily independent of $X_t$). Expressed as an element of the HBT model, we have that the time process $X_t^{(\omega)}$ is equal to the Brownian motion $A_t$ for all $s \geq 0$

$$d_i(s) = \tilde{d}_{t,i-1} \text{ for } t \in (\tau_{i-1}, \tau_i)$$

$$u_i(s) = \tilde{c}_{t,i-1} \text{ for } t \in (\tau_{i-1}, \tau_i).$$

---

3.2. Possible extensions of the model with uncertainty zones

The model with uncertainty zones of Robert and Rosenbaum (2011) introduced in Example 4, which is semi-parametric, assumes that the observed price is the efficient price rounded to the nearest tick value $Z_t := X_t^{(\omega)}$ and thus the noise is equal to $\epsilon_i := \alpha(\frac{1}{2} - \eta)$ if the last trade increased the price and $\epsilon_i := -\alpha(\frac{1}{2} - \eta)$ if the last trade decreased the price. In particular, the noise is auto-correlated and correlated to the efficient price. Because of this specific noise distribution, it is directly possible to estimate the underlying friction parameter $\eta$ without any data pre-processing such as preaveraging (see Robert and Rosenbaum (2012)).

We believe the model with uncertainty zones is a very interesting starting point, because all the endogenous and noise structure of the model is reduced to the estimation of the 1-dimensional friction parameter $\eta$. Nevertheless, as this semi-parametric model wants to be the simplest, it suffers from several issues. We will investigate two of them in the following.

First, the model does not allow for asymmetric information between the buyers and the sellers. Define $\eta^+$ and $\eta^-$, which are respectively the aversion to a positive price change and a negative price change. As a positive price change means that a buyer decided to put an order at the best ask price and a negative price change corresponds to a seller that puts an order at the best bid price (if we assume that cancel and repost orders are not the reason why the price changed), the difference $\eta^+ - \eta^-$ can be seen as a measure of information asymmetry. We define $\tau_0 := 0$ and recursively for $i$ any positive integer

$$\tau_i := \inf\{ t > \tau_{i-1} : X_t = X_t^{(\omega)} - \alpha \left( I_t - \frac{1}{2} + \eta^+ \right)$$

or $X_t = X_t^{(\omega)} + \alpha \left( I_t - \frac{1}{2} + \eta^- \right) \}$

Note that the HBT class contains this model and that it can be directly fitted if we slightly modify $\eta$ in Robert and Rosenbaum (2012) to estimate $\eta^+$ and $\eta^-$. One possible application would be to build a test of asymmetric information $\eta^+ := \eta^-$. This is beyond the scope of this paper.

One other issue is that the authors do not do any model checking in their work. According to their empirical work (see pp. 359–361 of Robert and Rosenbaum (2011)), the estimated values for $\eta$ are stable across days for the ten French assets tested. Stability of $\eta$ favors their model but by doing so, the model does not allow any other structure than the full-endogeneity for the sampling times. Even if the true structure of sampling times is (mostly) independent of the asset price, we will still estimate an $\eta$ that will be stable across days. If we allow the time process to be different from the price process itself, we can estimate the correlation $\rho^{1,1}$ between them and see how endogenous the sampling times are (the bigger $|\rho^{1,1}|$, the more endogenous the sampling times are). We would need to add more general microstructure noise in the model, and thus this is left for further work.

4. Main result

4.1. Assumptions and theorem

Without loss of generality, we fix the horizon time $T := 1$, and we consider $[0, 1]$ to represent the course of an economic event, such as a trading day. We first introduce the definition of stable convergence, which is a little bit stronger than usual convergence in distribution and needed for statistical purposes of inference, such as the prediction value of the high-frequency covariance and the construction of a confidence interval at a given confidence level.
Definition 1. We suppose that the random processes $Y_t$, $\mu_t$, and $\sigma_t$ are adapted to a filtration $(\mathcal{F}_t)$. Let $Z_t$ be a sequence of $\mathcal{F}_t$-measurable random variables. We say that $Z_t$ converges stably in distribution to $Z$ as $n \to \infty$ if $Z$ is measurable with respect to an extension of $\mathcal{F}_t$ so that for all $A \in \mathcal{F}_1$ and for all bounded continuous$^3$ functions $f$, $\mathbb{E}[f(Z_n)] \to \mathbb{E}[f(Z)]$ as $n \to \infty$.

In the setting of Section 2, the target of inference, the integrated covariation, can be written for all $t \in [0, 1]$ as

$$\langle X^1(t), X^2(t) \rangle_{\mathcal{F}_t} := \int_0^t \sigma_s^{(1)} \sigma_s^{(2)} \rho_s^{1,2} \, ds.$$  

We are providing now the asymptotics. We want to make the number of observations go to infinity asymptotically. The idea is to scale and thus keep the structure that drives the next return and the next time observation, while making the tick size vanish (and thus the number of observations explode on $[0, 1]$). Formally, we let the tick size $\alpha > 0$ and we define the observation times

$$t_{i,j} := \inf\{t > t_{i-1,j} : \Delta X(t) \neq 0\}$$

and define the HY estimator when the tick size is equal to $\alpha$ as

$$\langle X^1(t), X^2(t) \rangle_{\mathcal{F}_t}^{HY} := \sum_{0 < t_{i,j} < t_{i,j+1}} \Delta X^1(t_{i,j}) \Delta X^2(t_{i,j}) \mathbf{1}_{\{|t_{i,j+1} - t_{i,j}| < \alpha\}}.$$  

We now give the assumptions needed to prove the central limit theorem of (7). We need to introduce some definitions for this purpose. In view of the different models introduced in Section 3, there are three different possible assumptions regarding the correlation between the time processes $X^i(t)$ and the price processes $X_t$. The first possibility is that they can be equal for all $0 \leq t \leq T$. In this case we define $\lambda_{min}$ as the smallest eigen-value of $\sigma_{i,j}^{(1)}, \sigma_{i,j}^{(2)}$. The second scenario is that for one $k \in [1, 2]$ we have $X_{k}(t) := X_{t,k}$, but the other time process is different from its associated price process. In that case, we define $\lambda_{min}$ as the smallest eigen-value of $\sigma_{i,j}^{(1)}, \sigma_{i,j}^{(2)}$. The third possible setting is that the time process is different from its associated asset price for both assets, and we let $\lambda_{min}$ the smallest eigen-value of $\sigma_{i,j}^{(1)}, \sigma_{i,j}^{(2)}$. The idea is that the time process is independent of its associated asset price process. The covariance matrix $\sigma_t$ is continuous. We assume that the volatility is continuous.

Assumption (A1). The drift $\mu_t$, the volatility matrix $\sigma_t$ and the (four dimensional) Brownian motion $W_t$ are adapted to a filtration $(\mathcal{F}_t)$. Also, $\mu_t$ is integrable and locally bounded. Furthermore, $\sigma_t$ is continuous. Finally, we assume that $\inf_{t \in [0,1]} \lambda_{min} > 0$ a.s.

Remark 1 (Robustness to Jumps in Volatility). The proof techniques, holding the volatility constant on small blocks, require the “continuity of volatility”. This is the same strategy as in Mykland and Zhang (2009) and Mykland (2012) where the volatility process follows a continuous Itô process. Nonetheless, following the same line of reasoning as for the proof of Remark 6, we can add a finite number of jumps in the volatility matrix. The proof of Theorem 1 will break in the case of infinite number of jumps in $\sigma_t$.

The following condition roughly assumes that both time processes cannot be equal to each other, even on a very small time interval. Specifically, we will assume that there is a constant strictly smaller than 1 such that the module of the instantaneous high-frequency correlation $\rho^{3,4}$ cannot be bigger than this constant. In practice, Assumption (A2) is harmless.

Assumption (A2). For all $t \in [0, 1]$ we have

$$\rho^{3,4} \in [\rho^{3,4} - \delta, \rho^{3,4}]$$

where $\max(|\rho^{3,4} - \delta|, |\rho^{3,4}|) < 1$.

The next assumption deals with the down process $d_t$ and the up process $u_t$. It is clear that $d_t$ and $u_t$ have to be known with information at time $t$, which is why we assume that they are adapted to $(\mathcal{F}_t)$. The rest of Assumption (A3) is very technical and we only try to be as general as we can with respect to the proof techniques we will use. The reader should understand Assumption (A3) as “assume the worst dependence structure possible between the return $\Delta X_t$ and the time increment $\Delta t$, knowing that they follow the HBT model”. We insist once again on the fact that we only make the dependence structure as bad as we can in our model so that we can investigate how biased the HY estimator can be in practice, and how much the estimates of the variance assuming no endogeneity are wrong.

Assumption (A3). For both assets $k = 1, 2$, define the couple of the down process and the up process $g_t^{(1,2)} := (d_t^{(1,2)}, u_t^{(1,2)})$ and let $g_t := (g_t^{(1)}, g_t^{(2)})$. We assume that

$$g^{(k)} : \mathbb{R}^+ \to \mathbb{R}^+ \times \mathbb{R}^- \times \mathbb{R}^+$$

to be integrable and locally bounded. Furthermore, there exist two non-random constants $0 < g^- < g^+$ such that a.s. for any $t \in [0, 1]$ and for any $s \geq 0$

$$g^- \leq \min(-d_t^{(k)}(s), u_t^{(k)}(s)) \leq \max(-d_t^{(k)}(s), u_t^{(k)}(s)) \leq g^+.$$  

Further, there exist two non-random constants $K > 0$ and $d > 0$ such that a.s.

$$\forall s \geq 0, g_t^{(k)}(s) = g_t^{(k)}(K).$$

$$\forall t \geq 0, g_t^{(k)}(s) \leq \min(g_t^{(1)}(s), |g_t^{(2)}(s)|) \leq K,$$

$$\forall (u,v) \in [0,1]^2 \text{ s.t. } 0 < u < v,$$

$$\|g_t^{(1)} - g_t^{(2)}\| \leq K|u - v|^d,$$

where $\|f_1, f_2\| := \sup_{u \in [0,1]} \max(|f_1(u)|, |f_2(u)|)$.  

Remark 2. Consider the space $C$ of constants defined in Assumption (A3)

$$C := \left\{ (g^-, g^+, K, d) \text{ s.t. } 0 < g^- < g^+, K > 0, d > \frac{1}{2} \right\}.$$  

For any $c \in C$, we define $g(c)$ to be the functional subspace of $\mathbb{R}^+ \to \left(\mathbb{R}^+ \to \mathbb{R}^\times \mathbb{R}^+\right)^2$ such that $\forall g \in g$, $g$ satisfies (9),

\[\text{Note that the continuity of } f \text{ refers to continuity with respect to the Skorokhod topology of } D[0,1]. \text{ Nevertheless, we can also use continuity given by the sup-norm, because all our limits are in } C[0,1]. \text{ One can look at Chapter VI of Jacod and Shiryayev (2003) as a reference. For further definition of stable convergence, one can look at Rényi (1963), Aldous and Eagleston (1978), Chapter 3 p.56 of Hall and Heyde (1980), Rootzén (1980), and Section 2 (pp. 169–170) of Jacod and Protter (1998).} \]
Remark 6. The advised reader will have noticed that Examples 3–5 and Example 6, where time processes are piecewise-constant and may depend on \( n \), do not follow Assumption (A3). The adaptation of Theorem 1 proofs in those examples is discussed in Appendix A.5. We have made the choice not to state more general conditions to keep tractability of Assumption (A3).

The last assumption is only technical, and also appears in the literature (Mykland and Zhang, 2012; Li et al., 2014).

Assumption (A4). The filtration \((\mathcal{F}_t)\) is generated by finitely many Brownian motions.

We can now state the main theorem.

**Theorem 1.** Assume (A1)–(A4). Then, there exist processes \( AB_t \) and \( AV_t \) adapted to \((\mathcal{F}_t)\) such that stably in law as the tick size \( \alpha \to 0 \),

\[
\alpha^{-1} \left( (X^{(1)}, X^{(2)})_{1,t}^{HY} - (X^{(1)}, X^{(2)})_t \right) \to AB_t + \int_0^t (AV_s)^{1/2} dZ_s,
\]

where \( Z_t \) is a Brownian motion independent of the underlying \( \sigma \)-field. The asymptotic bias \( AB_t \) and the asymptotic variance \( AV_t \) are defined in Section 4.3 and estimated in Section 5.

**Remark 4 (Path-Bias).** Note that the asymptotic bias term \( AB_t \) on the right-hand side of (13) does not mean that the Hayashi–Yoshida estimator is biased, but rather path-biased. The latter is a weaker statement which means that once we have seen a path, there is a bias for the HY estimator on this specific path of value \( AB_t \). In practice, we only get to see one path and thus bias and path-bias can be confused easily. When doing simulations, we can observe many paths and the reader should keep in mind that the path-bias will be different for each path. In addition, note that if we assume that \( \sigma_1 \) is bounded and bounded away from \( 0 \) on \([0,T]\), there is no bias in Theorem 1 because \( E[AB_t] = 0 \).

**Remark 5 (Convergence Rate).** At first glance, the convergence rate \( \alpha^{-1} \) looks different from the optimal rate of convergence \( n^{1/2} \) we obtain in the no-endogeneity case. This is merely a change of perspective because we are looking from the tick size point-of-view. Actually, if for \( k = 1, 2 \) we define \( N_{t,a}^{(k)} \) as the number of observations before \( t \) of the \( k \)th asset and the sum of observations of both processes \( N_{t,a}^{(S)} := N_{t,a}^{(1)} + N_{t,a}^{(2)} \), we have that \( N_{t,a}^{(S)} \) is exactly of order \( O_p(\alpha^{-2}) \). Thus, if we define the expected number of observations \( n := E[N_{t,a}^{(S)}] \), we obtain the optimal rate of convergence \( n^{1/2} \) in (13).

**Remark 6 (Robustness to Jumps in Price Processes).** We assume that we add a jump component to the price process

\[
dX^{(k)}_t = \mu^{(k)}_t dt + \sigma^{(k)}_t dB^{(k)}_t + dJ^{(k)}_t
\]

for \( k = 1, 2 \), where \( J_t \) denotes a 2-dimensional finite activity jump process and \( dB^{(k)}_t \) is either zero (no jump) or a real number indicating the size of the jump at time \( t \). We follow exactly the setting of p. 2 in Andersen et al. (2012). We assume that \( J_t \) is a general Poisson process independent of the other quantities. Under the same assumptions the conclusion of Theorem 1 remains valid. The proof can be found in Appendix A.6. The infinitely many jumps case is complex and beyond the scope of this paper. This was already the case in the 1-dimensional case (see Remark 4 in p. 586 of Li et al. (2014)).

**Remark 7 (Grid on the Original Non-Log Scale).** Theorem 1 covers the particular case where \( X_t \) corresponds to the log-price and observations are obtained when the price on the original scale hits a boundary. This can be done by a parametrization of \( g_t^{(k)} \) by \( \tilde{g}_t^{(k)}(s) := (-\exp(-d_t^{(k)}), \exp(u_t^{(k)}) \).

**Remark 8 (Arbitrary Number of Assets).** The authors chose for simplicity to work only with two assets, but they conjecture that this result would stay true for an arbitrary number of assets, and that our proofs would adapt to show it, at the cost of more involved notations and definitions.

### 4.2. Definition of the Bias-corrected HY estimator

Assume that we have a consistent estimator \( \tilde{AB}_{t,a} \) of the bias \( AB_{t,a} := \alpha AB_t \). Such estimator will be provided in Section 5.

We define the new estimator \( (X^{(1)}, X^{(2)})_{t,a}^{BC} \) of high-frequency covariance as the estimate obtained when removing the bias estimate \( AB_{t,a} \) from the Hayashi–Yoshida estimator

\[
(X^{(1)}, X^{(2)})_{t,a}^{BC} := (X^{(1)}, X^{(2)})_{t,a}^{HY} - \tilde{AB}_{t,a}. \tag{15}
\]

With the bias-corrected estimator \( (X^{(1)}, X^{(2)})_{t,a}^{BC} \), we get rid of the asymptotic bias and keep the same asymptotic variance as we can see in the following corollary.

**Corollary 2.** Assume (A1)–(A4). Then, stably in law as \( \alpha \to 0 \),

\[
\alpha^{-1} \left( (X^{(1)}, X^{(2)})_{t,a}^{BC} - (X^{(1)}, X^{(2)})_t \right) \to \int_0^t (AV_s)^{1/2} dZ_s. \tag{16}
\]

### 4.3. Computation of the theoretical asymptotic bias and asymptotic variance

We warn the reader interested in implementing the bias-corrected estimator that this section is highly technical and we advise her to go directly to Section 5 and refer to this section only for the definitions. On the contrary, if the reader wants to understand the main ideas of the proofs, she should take this section as a reference. We also want to emphasize on the fact that the theoretical values of asymptotic bias and asymptotic variance found at the end of this section are rather abstract and do not shed easily light on how the change of parameters \( \sigma_1 \) and \( g \) in the model would influence the asymptotic bias and asymptotic variance. The main purpose of this paper is that we do not need to know the theoretical values in order to compute the estimators in Section 5.

We need to introduce some definitions in order to compute the theoretical asymptotic bias \( AB_t \) and the asymptotic variance term \( AV_t \). We first need to rewrite the HY estimator (7) in a different way. For any positive integer \( i \), consider the \( i \)th sampling time of the first asset \( t_{1,a}^{(1)} \). We define two random times, \( \tau_{i-1,a} \) and \( \tau_i^{(1,a)} \), which are functions of \( t_{1,a}^{(1)} \) and all the observation times of the second asset \( \{s_{N_{t,a}^{(2)}}\}_{i=0} \), and which correspond respectively to the closest sampling time of the second asset that is strictly smaller

\[ AB_{t,a} \text{ is consistent means that } \alpha^{-1} AB_{t,a} = \alpha^{-1} AB_{t,a} + O_p(1). \]
than \( r_{\tau-1,\alpha}^{(1)} \) and the closest sampling time of the second asset that is (not necessarily strictly) bigger than \( r_{\tau-1,\alpha}^{(1)} \) as
\[
\tau_{\alpha,\tau}^{-} = 0, \tag{17}
\]
\[
\tau_{\alpha,\tau}^{-} = \max\{\tau_{\alpha,\tau}^{(2)} : \tau_{\alpha,\tau}^{(2)} < \tau_{\alpha,\tau}^{-}\} \quad \text{for } i \geq 2, \tag{18}
\]
\[
\tau_{\alpha,\tau}^{+} = \min\{\tau_{\alpha,\tau}^{(2)} : \tau_{\alpha,\tau}^{(2)} \geq \tau_{\alpha,\tau}^{-}\} \quad \text{for } i \geq 1. \tag{19}
\]
We consider \( \Delta X_{\tau-1,\alpha}^{(2)} \), the increment of the second asset between \( \tau_{\alpha,\tau}^{-} \) and \( \tau_{\alpha,\tau}^{+} \)
\[
\Delta X_{\tau-1,\alpha}^{(2)} := \Delta X_{\tau-1,\alpha}^{+(2)} - \Delta X_{\tau-1,\alpha}^{-}\tag{20}
\]
Rearranging the terms in (7) gives us (except for a few terms at the edge)
\[
(X_{\tau-1,\alpha}^{(1)}, X_{\tau-1,\alpha}^{(2)})_{\tau,\alpha} = \sum_{\tau_{\alpha,\tau}^{-} < t} \Delta X_{\tau,\alpha}^{(1)} X_{\tau,\alpha}^{(2)}, \tag{21}
\]

The representation in (21) is very useful in the sense that it gives a natural order between the terms in the sum. Nevertheless, any term of this sum is a priori correlated with the other terms. We will rearrange once again the terms in (21), so that each term is only correlated with the previous and the next term of the sum. In this case, we say that they are 1-correlated. For this purpose, we need to introduce some notation. We remind the reader that \( T_{\alpha} \) is the two-dimensional vector of sampling times, where for each \( k = 1, 2 \) the \( k \)th component \( T_{\alpha}^{k} \) is equal to the sequence of sampling times associated with the \( k \)th asset. We will construct a subsequence \( T_{\alpha}^{IC} \) of \( T_{\alpha}^{IC} \) that also depends on the observation times of the second asset \( T_{\alpha}^{IC} \), and will be such that we can write the Hayashi–Yoshida estimator as a 1-correlated sum similar to (21), except the new sampling times \( T_{\alpha}^{IC} \) will replace the original observation times \( T_{\alpha}^{IC} \). The new sampling times \( T_{\alpha}^{IC} \) are obtained using the following algorithm. We define \( T_{\alpha}^{IC} := T_{\alpha}^{(1)} \), and recursively for \( i \) any nonnegative integer
\[
T_{\alpha}^{IC+i} := \min\{T_{\alpha}^{IC}: \text{there exists } j \in \mathbb{N} \}
\]
\[
\text{such that } T_{\alpha}^{IC+i} \leq T_{\alpha}^{(2)} < T_{\alpha}^{IC+i-1}. \tag{22}
\]
In words, if we sit at the observation time \( T_{\alpha}^{IC} \) of the first asset, we wait first to hit an observation time of the second asset, and we then choose the next strictly bigger observation time of the first asset. In analogy with (17)–(19) and (20), we define the following times
\[
T_{\alpha}^{IC-}, \quad \text{for } i \geq 2, \tag{23}
\]
\[
T_{\alpha}^{IC-} := \max\{T_{\alpha}^{IC+i} : T_{\alpha}^{IC+i} < T_{\alpha}^{IC}\} \quad \text{for } i \geq 2, \tag{24}
\]
\[
\Delta X_{\tau-1,\alpha}^{(2)} := \Delta X_{\tau-1,\alpha}^{+(2)} - \Delta X_{\tau-1,\alpha}^{-(2)} \quad \text{for } i \geq 1. \tag{26}
\]
First, observe that, except for maybe a few terms at the edge, we can rewrite (21) as
\[
(X_{\tau-1,\alpha}^{(1)}, X_{\tau-1,\alpha}^{(2)})_{\tau,\alpha} = \sum_{\tau_{\alpha,\tau}^{-} < t} \Delta X_{\tau,\alpha}^{(1)} X_{\tau,\alpha}^{(2)\alpha+\alpha}, \tag{27}
\]
Also, we define the following compensated increments of the HY estimator
\[
N_{\tau,\alpha} = \Delta X_{\tau,\alpha}^{IC} \Delta X_{\tau,\alpha}^{IC} - \int_{\tau_{\alpha,\alpha}^{IC} - \tau_{\alpha,\alpha}^{IC}}^{\tau_{\alpha,\alpha}^{IC}} \Delta X_{\tau,\alpha}^{IC} - ds. \tag{28}
\]
Note that they are compensated in the sense that they are centered (if we decompose \( \Delta X_{\tau,\alpha}^{IC} \) into a left (−), a central and a right (+) part and condition the expectation, this is straightforward to show). Similarly, we can show that they are 1-correlated.

The idea of the proof is the following. If we consider the volatility matrix \( \sigma_{t} \) and the grid function \( g_{i} \) to be constant over time, we can express the conditional returns of the normalized error of HY as a homogeneous Markov chain (of order 1), show that the Markov chain is uniformly ergodic and thus use results in the limit theory of Markov chains (see, e.g., Meyn and Tweedie (2009)) to show that it has a stationary distribution. Then, we prove that we can approximate locally the returns of the normalized error when the volatility matrix and grid function are not constant by the returns when holding them constant on a small block. Finally, using limit theory techniques developed in Mykland and Zhang (2012) together with standard probability results of conditional distribution (see, e.g., Breiman (1992)), we can bound uniformly in time the error of the returns when holding the volatility matrix and grid function constant.

Based on the definitions introduced in Appendix A.1, we can define the instantaneous variance of the normalized HY estimate’s error (29), which depends on the volatility matrix \( \sigma \) and the grid \( g \). Similarly, we also define the instantaneous covariance between the normalized HY’s error and the first asset price (30), and the instantaneous covariance between the error and the second asset price (31). Finally, we define the instantaneous 1-correlated time, which is the approximation of \( \frac{\mathbb{E}_{\tau_{\alpha}^{IC}}[\Delta X_{\tau_{\alpha}^{IC}}^{(2)\alpha+\alpha]}}{\Delta X_{\tau_{\alpha}^{IC}}^{(2)\alpha+\alpha]}} \), where if \( \tau \) is a \( (\mathcal{F}_{\tau}) \)-stopping time, \( \mathbb{E}_{\tau} \{Y\} \) is defined as the conditional distribution of \( Y \) given \( \mathcal{F}_{\tau} \).

\[
\psi^{AV}(\sigma, g, x, u) := \mathbb{E}[N_{\tau}^{2} + 2N_{\tau}N_{\tau}], \tag{29}
\]
\[
\psi^{AC1}(\sigma, g, x, u) := \mathbb{E}[\hat{N}_{\tau} \Delta X_{\tau}^{(1)}], \tag{30}
\]
\[
\psi^{AC2}(\sigma, g, x, u) := \mathbb{E}[\hat{N}_{\tau} \Delta X_{\tau}^{(2)}], \tag{31}
\]
\[
\psi^{x}(\sigma, g, x, u) := \mathbb{E}[\Delta X_{\tau}^{(2)}]. \tag{32}
\]

Remark 9. The reader might expect \( \bar{N}_{\tau} \) in lieu of \( \tilde{N}_{\tau} \) in (29)–(31) and (32). Actually, we cannot use \( \bar{N}_{\tau} \) directly from the definition because the corresponding time \( \tau_{\alpha}^{IC} < 0 \). We would need to set it to \( -u \) to alter the definition of (29)–(31) and (32), which we have chosen not to do for the sake of clarity.

Set \( \tilde{Z}_{0} := (x, u) \) and for any positive integer \( i \)
\[
\tilde{Z}_{i} := \left( \Delta X_{\tau_{\alpha}^{IC}}^{(4)\alpha+\alpha}, \tilde{\tau}_{\alpha}^{IC} - \tilde{\tau}_{\alpha}^{IC}\right), \tag{33}
\]
For any nonnegative integer \( i \), we consider \( \tilde{\pi}_{i}(\tilde{\sigma}, \tilde{g}, x, u) \) the distribution of \( \tilde{Z}_{i} \). We also introduce the notation \( \Pi(\tilde{\sigma}, \tilde{g}, x, u) := \{\tilde{\pi}_{i}(\tilde{\sigma}, \tilde{g}, x, u)\}_{i \geq 0} \). By the strong Markov property of Brownian motion, we can show that \( \tilde{Z}_{i} \) is a homogeneous Markov chain (of order 1) on the state space \( \mathbb{S}_{\tilde{g}} \). In the following lemma, we show that there exists a stationary distribution of \( \tilde{\pi}_{i}(\tilde{\sigma}, \tilde{g}, x, u) \).

Lemma 3. Let \( c := (g_{-}, g_{+}, K, d) \) be a four-dimensional vector such that \( c \in \mathbb{C} \) and consider \( \tilde{\sigma} \) a constant volatility matrix such that \( \tilde{\sigma}_{\min} > 0 \) and \( \tilde{g} \in \mathbb{C}^{c} \). Then, there exists a stationary distribution \( \tilde{\pi}_{i}(\tilde{\sigma}, \tilde{g}, x, u) \).
The proof of Lemma 3 can be found in the Appendix (proof of Lemma 14). The next definition is the average (regarding the stationary distributions) of the instantaneous variance, covariances and 1-correlated time. For any $\theta \in \{AV, AC1, AC2, \tau\}$, \[
\phi^\theta(\hat{\sigma}, \hat{g}) := \int_{x_2} \psi^\theta(\hat{\sigma}, \hat{g}, y, v) \, d\hat{\pi}(\hat{\sigma}, \hat{g})(y, v).
\]

We introduce the notation $\phi^\theta_k(\sigma, g)$ and consider the following quantities needed to compute the asymptotic bias and variance.
\[
k^1 := (\sigma^2_k)^{-2} \phi^AC1_k(\phi^r)^{-1}, \quad (34)
\]
\[
k_k := (1 - \rho_k^{-1})^{-1} (\sigma^2_k)^{-2} \phi^AC2_k
\]
\[
- (\sigma^2_k \rho_k^{-2})^{-1} \rho_k^{-1} \phi^AC1_k(\phi^r)^{-1}. \quad (35)
\]
We express now $AV_s$ the quantity integrated to obtain the asymptotic variance.
\[
AV_k := \phi^AW_k + 2(k^1_k \sigma^2_k) \rho^2_k \phi^AC1_k
\]
\[
- (k^1_k + k^1_k \phi^AC2_k) \phi^r_k
\]
\[
+ (\phi^r_k)^1 \rho^2_k (\phi^r_k)^2 (1 - \rho_k^{-1} \phi^AC2_k)^2. \quad (36)
\]

The asymptotic bias is defined as $AB_k := \int_0^1 AB^1_k \, dx_k^1 + \int_0^1 AB^2_k \, dx_k^2$ where
\[
AB^1_k := (\frac{1}{k^1_k} - \rho^2_k \sigma^2_k)(\sigma^2_k)^{-1}, \quad (37)
\]
\[
AB^2_k := \frac{1}{k^1_k}. \quad (38)
\]

**Remark 10 (Asymptotic Bias).** Looking at the expressions for $AB^1_k$ and $AB^2_k$, one can be tempted to think that because of the $(1 - \rho^{-2}_k \phi^AC2_k^2)$ term in $k^{1,1}$, the bias will increase drastically when both assets are highly correlated. In this case, the reader should keep in mind that the second term of $AB^1_k$, when integrated with respect to $X^1_k$, and $AB^2_k$, when integrated with respect to $X^2_k$, will be roughly of the same magnitude, with opposite signs, and thus there is no explosion of asymptotic bias. We chose the above asymptotic bias’ representation because it is straightforward to build estimators from it. We can also express the asymptotic bias differently. For this purpose, we can rewrite the log-price process as
\[
dx^1_k = \sigma^1_k \db_k^1,
\]
\[
dx^2_k = \rho^2_k \sigma^2_k \db_k^2 + (1 - \rho^2_k \phi^AC2_k^2)^{1/2} \sigma^2_k \db_k^1, \]
where $\db^1_k$ and $\db^1_k$ are independent Brownian motions. Let \[
\dx_k^1 := (1 - \rho^2_k \phi^AC2_k^2)^{1/2} \sigma^2_k \db_k^1, \quad (39)
\]
be the part of $X_k^2$ that is not correlated with $X_k^1$. We can express the asymptotic bias as $AB_k := \int_0^1 AB^1_k \, dx_k^1 + \int_0^1 AB^2_k \, dx_k^1, \quad (1)$.

In this case, $AB_k = k^{1,1}$ and \[
AB_k := \lim_{n \to \infty} (M^n, B^{1,1}),
\]
where $M^n$ is defined in the proofs. We can show that this limit exists, and does not explode when both assets are highly correlated.

## 5. Estimation of the bias and variance

We need to introduce some new notations. We recall that $N^{1,1}$ the number of observations corresponding to the first asset before 1 and we also define $N_{IC}$ the number of 1-correlated observations before 1, i.e., $N_{IC} := \max\{i \in \mathbb{N} \text{ s.t. } \tau_{IC} < 1\}$. In practice, the first step is to transform the returns of the first asset
\[
\left\{(\Delta x_k^{1,1}, \Delta t_{IC})\right\}_{i=1}^{N_{IC}}
\]
into 1-correlated returns
\[
\left\{(\Delta x_k^{IC}, \Delta t_{IC})\right\}_{i=1}^{N_{IC}}
\]
using algorithm (22). Then, for each asset, we will chop the data into $B_k$ blocks and on each block $i = 1, \ldots, B_k$ we will estimate $\deltaV_{t, \alpha}^{AB_{IC}}$ and $\deltaV_{t, \alpha}^{AB}$, pretending that the volatility matrix $\sigma_t$ and grid $g_t$ are block-constant.

Because there is asynchronicity in the observation times, the blocks of each asset are not exactly equal. Let $h_n$ be the block size. For the first asset, we consider block $1^{1,1} := [0, \tau_{h_1}^{1,1}]$, block $2^{1,1} := [\tau_{h_1}^{1,1}, \tau_{h_2}^{1,1}]$, etc. For the second asset, we let block $1^{2,1} := [\tau_{h_1}^{2,1}, \tau_{h_2}^{2,1}]$, block $2^{2,1} := [\tau_{h_2}^{2,1}, \tau_{h_3}^{2,1}]$, etc. In the following, we will say $j \in \text{block } i$ when $\tau_j^{1,1} \in \text{block } i$. Similarly, we say $j \in \text{block } i$ when $\tau_j^{2,1} \in \text{block } i$. Finally, we define $j \in \text{block } i$ if $j \in [(1 - h_n) + 1, \ldots, h_n]$. First, we estimate the volatility of both assets using the corrected estimator in Li et al. (2014). To do this, we need to define an estimate of the spot volatility on each block for each asset $k = 1, 2$ by
\[
\hat{a}_{k,a}^{(k)} := \left( \sum_{j \in \text{block } k} (\Delta x_j^{(k)})^2 \right)^{1/2}.
\]

Then, we estimate the asymptotic bias of the volatility via
\[
\hat{\deltaV}_{t, \alpha}^{(k)} = \frac{2}{3(\hat{a}_{k,a}^{(k)})^2} \sum_{j \in \text{block } k} (\Delta x_j^{(k)})^3.
\]

We obtain the bias-corrected estimators of volatility on each block:
\[
\hat{a}_{k,a}^{(k)} = \hat{a}_{k,a}^{(k)} - \hat{\deltaV}_{t, \alpha}^{(k)}.
\]

Then, we estimate the correlation between both assets using the naive HV estimator
\[
\hat{\rho}_{t, \alpha}^{1,2} = \frac{1}{\hat{a}_{1,a}^{(1)} \hat{a}_{2,a}^{(2)}} \sum_{j \in \text{block } i} \Delta x_j^{1,1} \Delta x_j^{2,1}.
\]

We then build an estimator of the compensated increments of the HV estimator, following the definition in (28).
\[
\hat{\deltaV}_{t, \alpha}^{1,2} = \frac{2}{3(\hat{a}_{1,a}^{(1)} \hat{a}_{2,a}^{(2)})^2} \sum_{j \in \text{block } i} (\Delta x_j^{1,1} \Delta x_j^{2,1})^3.
\]

The next step is to estimate the instantaneous variance (29), both instantaneous covariances (30) and (31) and the instantaneous 1-correlated time (32) on each block. This is done by taking the sample average of the corresponding estimated quantities. Note that we do not directly estimate $\psi^{AW}_{\psi AC1}$, $\psi^{AC2}$, and $\psi^{AC}$, but rather a scaling version of them, i.e., $a^2_{\psi^{AW}}$, $a_{\psi^{AC1}}$, $a_{\psi^{AC2}}$ and $a_{\psi^{AC}}$. In practice, we can always assume $a_t := 1$ by scaling $g_t$ by the tick size, and thus we match the definitions of the following estimators with (29)–(32). For the sake of simplicity, we assume that the number of 1-correlated observations of the last block $b_n$ is also $h_n$. In practice, this will be most likely different from $h_n$, and thus the denominator of (40)–(43) will have to be changed so that it is equal to the number of 1-correlated observations in this last
block. The estimates are given by
\[
\hat{\phi}_{i,a}^{\text{AV}} = h_n^{-1} \sum_{j \in \text{block } i} \hat{N}_{j,a}^2 + 2 \hat{N}_{j,a} \hat{N}_{j+1,a},
\]
(40)
\[
\hat{\phi}_{i,a}^{\text{AC}1} = h_n^{-1} \sum_{j \in \text{block } i} \hat{N}_{j,a} \Delta X_{i,a}^{(1)},
\]
(41)
\[
\hat{\phi}_{i,a}^{\text{AC}2} = h_n^{-1} \sum_{j \in \text{block } i} \hat{N}_{j,a} \Delta X_{i,a}^{(2)},
\]
(42)
\[
\hat{\phi}_{i,a}^{\text{AR}} = h_n^{-1} \sum_{j \in \text{block } i} \Delta \tau_{i,a}^{(1)},
\]
(43)
We estimate now the quantities (34) and (35) as
\[
\hat{k}_{i,a}^{(1)} := \left( \sigma_{i,a}^{(1)} \right)^2 \hat{\phi}_{i,a}^{\text{AC}1} (\hat{\phi}_{i,a}^{\text{AR}})^{-1},
\]
(44)
\[
\hat{k}_{i,a}^{(1,2)} := \left( 1 - \hat{\rho}_{i,a}^{(1,2)} \hat{\rho}_{i,a}^{(1)} \hat{\rho}_{i,a}^{(2)} \hat{\rho}_{i,a}^{(2)} \right)^{-1} \left( \hat{\rho}_{i,a}^{(1,2)} \left( \hat{\phi}_{i,a}^{\text{AC}1} \left( \hat{\phi}_{i,a}^{\text{AR}} \right)^{-1} \right) - \hat{\rho}_{i,a}^{(1,2)} \hat{\rho}_{i,a}^{(1)} \hat{\rho}_{i,a}^{(2)} \hat{\rho}_{i,a}^{(2)} \hat{\phi}_{i,a}^{\text{AC}2} \left( \hat{\phi}_{i,a}^{\text{AR}} \right)^{-1} \right).
\]
(45)
We follow (37) and (38) to estimate the bias integrated terms \(AB_t^{(1)}\) and \(AB_t^{(2)}\) on each block
\[
\hat{AB}_{i,a}^{(1)} := \left( \hat{k}_{i,a}^{(1)} \hat{X}_{i,a}^{(1)} - \hat{X}_{i,a}^{(1)} \right) \hat{\rho}_{i,a}^{(1,2)} \hat{\rho}_{i,a}^{(1)} \hat{\rho}_{i,a}^{(2)} \hat{\rho}_{i,a}^{(2)} \left( \hat{\phi}_{i,a}^{\text{AR}} \right)^{-1},
\]
\[
\hat{AB}_{i,a}^{(2)} := \hat{k}_{i,a}^{(1,2)}.\]
For the variance term \(AV_t\), we decide not to use the direct definition in (36) because it can provide negative estimates. Instead, we will be using the following estimator
\[
\hat{AV}_{i,a} := \left( \sum_{j \in \text{block } i} \hat{N}_{j,a} \right) \Delta \tau_{i,a}^{(1)} \left( \hat{X}_{i,a}^{(1)} - \hat{X}_{i,a}^{(1)} \right).
\]
(46)
We define the final estimators of asymptotic bias and asymptotic variance as
\[
\hat{AB}_a := \sum_{i=1}^{B_a} \hat{AB}_{i,a} \left( \hat{X}_{i,a}^{(1)} - \hat{X}_{i,a}^{(1)} \right),
\]
\[
\hat{AV}_a := \sum_{i=1}^{B_a} \hat{AV}_{i,a} \left( \hat{X}_{i,a}^{(1)} - \hat{X}_{i,a}^{(1)} \right).
\]
(47)
As a corollary of Theorem 1, we obtain the following result, which states the consistency of (46) and (47).

**Corollary 4.** There exists a choice of the block size \(h_n\) such that when \(\alpha \to 0\), we have
\[
\alpha^{-1} \hat{AB}_a \Rightarrow AB_1,
\]
(48)
\[
\alpha^{-1} \hat{AV}_a \Rightarrow \int_0^1 AV_t ds.
\]
(49)

In particular, in view of Corollary 2, the bias-corrected estimator \((X_1^{(1)}, X_2^{(2)})^{\text{RC}}_{1,a} := (X_1^{(1)}, X_2^{(2)})^{\text{ASC}}_{1,a} - AB_a\) is such that
\[
\frac{(X_1^{(1)}, X_2^{(2)})^{\text{RC}}_{1,a}}{\sqrt{\hat{\alpha}}^{1/2}} \to N(0, 1).
\]
(50)

\[^6\text{The exact assumptions on } h_n \text{ can be found in the proofs of Theorem 1.}\]
In the fourth setting, we consider another model of arrival times, namely Example 4. We set the tick size $\alpha = 0.0001$ and the friction parameter $\eta = 0.15$. Price and volatility processes are assumed to follow the same model as in the second setting. We simulate price processes and observation times for 10 years of 252 business days. We choose $h_n = n^{-1}$ for Settings 2–4. We provide in Table 1 a summary of the comparison results between HY and the bias-corrected HY. As expected from the theory, the RMSE is improved when using the bias-corrected estimator in Example 1. In Example 4, the bias-corrected HY does not seem to perform better than HY. We conjecture that there is no asymptotic bias in Example 4, and that this is the reason why we do not observe any difference between the two estimators in that simple model. In addition, the sample bias is almost the same when using HY and the bias-corrected estimator for the four different settings, which is also expected from Remark 4. Furthermore, this sample bias tends to 0, which comes from the fact that both estimators are consistent. Finally, the standardized feasible statistic (50) in the first setting is reported in Table 2 and plotted in Fig. 2.

Table 1
Summary statistics based on simulated endogenous data of 1, 5 and 10 years. The RMSE in the table corresponds to the square root of the squared distance between the estimated value and the true value $6.4e-05$. HY stands for the usual Hayashi–Yoshida estimator (4), and BCHY represents the bias-corrected estimator (15).

| No. years | Estim | Setting | Sample Bias | RMSE  | % Reduced RMSE |
|-----------|-------|---------|-------------|-------|----------------|
| 1         | HY    | 1       | 5.41e-07    | 1.36e-05 | -              |
| 1         | BCHY  | 1       | 5.43e-07    | 1.19e-05 | 13%            |
| 5         | HY    | 1       | 1.10e-07    | 1.42e-05 | -              |
| 5         | BCHY  | 1       | 1.07e-07    | 1.26e-05 | 11%            |
| 10        | HY    | 1       | 5.54e-08    | 1.39e-05 | -              |
| 10        | BCHY  | 1       | 5.53e-08    | 1.20e-05 | 14%            |
| 1         | HY    | 2       | 5.47e-07    | 1.66e-05 | -              |
| 1         | BCHY  | 2       | 5.44e-07    | 1.50e-05 | 9%             |
| 5         | HY    | 2       | 1.13e-07    | 1.71e-05 | -              |
| 5         | BCHY  | 2       | 1.15e-07    | 1.58e-05 | 8%             |
| 10        | HY    | 2       | 5.56e-08    | 1.70e-05 | -              |
| 10        | BCHY  | 2       | 5.60e-08    | 1.57e-05 | 8%             |
| 1         | HY    | 3       | 5.61e-07    | 1.80e-05 | -              |
| 1         | BCHY  | 3       | 5.62e-07    | 1.67e-05 | 7%             |
| 5         | HY    | 3       | 1.14e-07    | 1.81e-05 | -              |
| 5         | BCHY  | 3       | 1.12e-07    | 1.68e-05 | 7%             |
| 10        | HY    | 3       | 5.56e-08    | 1.80e-05 | -              |
| 10        | BCHY  | 3       | 5.55e-08    | 1.68e-05 | 7%             |
| 1         | HY    | 4       | 4.41e-07    | 1.10e-05 | -              |
| 1         | BCHY  | 4       | 4.44e-07    | 1.11e-05 | - 1%           |
| 5         | HY    | 4       | 8.81e-08    | 1.10e-05 | -              |
| 5         | BCHY  | 4       | 8.80e-08    | 1.09e-05 | 1%             |
| 10        | HY    | 4       | 4.39e-08    | 1.08e-05 | -              |
| 10        | BCHY  | 4       | 4.43e-08    | 1.08e-05 | 0%             |

7. Conclusion

We have introduced in this paper the HBT model, and we have shown that it is more general than some of the endogenous models of the literature. This model can be extended to a model including more general noise structure in observations, and even noise in sampling times. This is investigated in Potiron (2016).

Under this model, we have proved the central limit theorem of the Hayashi–Yoshida estimator. Our main theorem states that there is an asymptotic bias. Accordingly, we built a bias-corrected HY estimator. We also computed the theoretical
standard deviation, and we provided consistent estimates of it. Numerical simulations corroborate the theory.

The techniques used for the proof of the main theorem could be applied to more general models and to other problems such as the estimation of the integrated variance of noise, integrated betas, etc. In particular, independence between the efficient price process and the noise is not needed in the model. As long as we can approximate the joint distribution of the noise and the returns by a Markov chain, ideas of our proof can be used.

Appendix

A.1. Definition of some quantities of approximation

We define in this section some quantities assuming the volatility matrix $\sigma_t$ and the grid function $g_t$ are constant. For that purpose, let $\tilde{W}$ be a four dimensional Wiener process, $c := (g^-, g^+, K, d)$ a four-dimensional vector such that $c \in C$ and $\tilde{\sigma}$ a constant volatility matrix such that the associated $\tilde{\lambda}_{min}$, which is the analog of $\lambda_{min}$ defined in Section 4.1 when we replace $\sigma_t$ by $\tilde{\sigma}$, is strictly bigger than 0 and $g \in \mathcal{G}(c)$ a constant grid function. In analogy with the definition of the grid function $g_t$ in (A3), we assume that $\tilde{g}$ can be written in terms of the down and up functions of both assets, i.e. $\tilde{g} := (\tilde{g}^1, \tilde{g}^2)$ where for each $k = 1, 2$ we have $\tilde{g}^k := (\tilde{u}^k, \tilde{d}^k)$. Also, we introduce $\tilde{d}_{k} := (y, v) \in \mathbb{R} \times \mathbb{R}^5$ s.t. $\tilde{d}^2(u) \leq y \leq \tilde{u}^2(v)$.

If we set $\tilde{X} = \tilde{\sigma} \tilde{W}$ and the corresponding sampling times of both assets $\tilde{T} := (\tilde{T}^{(1)}, \tilde{T}^{(2)})$, where for $k = 1, 2$ we have $\tilde{T}^{(k)} := \{\tilde{t}_i\}_{i \geq 0}$, we define the observation times of the first asset as $\tilde{t}_{0}^{(1)} := 0$ and recursively for any positive integer $i$ we have

$$\tilde{t}_{i}^{(1)} := \inf\{t > \tilde{t}_{i-1}^{(1)} : \Delta \tilde{X}_t^{(1)} \not\in [\tilde{d}^1(t - \tilde{t}_{i-1}^{(1)}), \tilde{u}^1(t - \tilde{t}_{i-1}^{(1)})]\}.$$

These stopping times will be seen as approximations of the observation times of the first asset when we hold the volatility matrix $\sigma_t$ and the grid $g_t$ constant. We will always start our approximation at a 1-correlated observation time $\tilde{t}_{i,n}$, which corresponds to an observation time of the first asset. As the sampling times of the second asset are not synchronized with the ones from the first asset, we need two more quantities $x(u, v) \in \tilde{d}_{k}$ to approximate the observation times of the second asset. They correspond respectively to the increment of the second asset's time process $X_t^{(2)}$ since the last observation of the second asset occurred and the time elapsed since the last observation time of the second asset. We define $\tilde{t}_{i}^{(2)} := 0$

$$\tilde{t}_{i}^{(2)} := \inf\{t > 0 : x + \Delta \tilde{X}_t^{(2)} \not\in [\tilde{d}_2(t + u), \tilde{u}_2(t + u)]\},$$

and for any integer $i \geq 2$

$$\tilde{t}_{i}^{(2)} := \inf\{t > \tilde{t}_{i-1}^{(2)} : \Delta \tilde{X}_t^{(2)} \not\in [\tilde{d}_2(t - \tilde{t}_{i-1}^{(2)}), \tilde{u}_2(t - \tilde{t}_{i-1}^{(2)})]\}.$$

Similarly, we define the analogs of (17)–(18), (19), (20), (23)–(24), (25), (26) and (28) respectively as $\tilde{t}_{i-1}^{+}, \tilde{t}_{i}^{+}, \Delta \tilde{X}_{\tilde{t}_{i-1}^{+}}^{(1)}, \Delta \tilde{X}_{\tilde{t}_{i}^{+}}^{(1)}, \tilde{t}_{i-1}^{+}, \tilde{t}_{i}^{+}, \tilde{t}_{i-1}^{+}, \tilde{t}_{i}^{+}, \Delta \tilde{X}_{\tilde{t}_{i-1}^{+}}^{(1)}, \Delta \tilde{X}_{\tilde{t}_{i}^{+}}^{(1)}$, and $N_i$ by putting tildes on the quantities in the definitions.

A.2. Preliminary lemmas

Without loss of generality, we choose to work under the third scenario defined in Section 4.1, i.e. the asset price is different from the time process for both assets. Because we shall prove stable convergence, and because of the local boundedness of $\sigma$ (because by (A1) $\sigma$ is continuous), and that $\inf_{\epsilon \in (0, 1)} \lambda_{min}^{\epsilon} > 0$ we can without loss of generality assume that for all $t \in [0, 1]$ there exist some nonrandom constants $\sigma^-$ and $\sigma^+$ such that for any eigenvalue $\lambda_t$ of $\sigma_t$ we have

$$0 < \sigma^- \leftarrow \lambda_t \leftarrow \sigma^+,$$

by using a standard localization argument such that the one used in Section 2.4.5 of Mykland and Zhang (2012). One can further suppress $\mu$ as in Section 2.2 (pp. 1407–1409) of Mykland and Zhang (2009), and act as if $X$ is a martingale.

We define the subspace $\mathcal{M}$ of matrices of dimension $4 \times 4$ such that $V \in \mathcal{M}$, for any eigenvalue $\lambda_M$ of $V$, we have

$$\sigma^- \leftarrow \lambda_M \leftarrow \sigma^+$$

and $\frac{(MM^*)^{-1/4}}{(MM^*)^{1/4}} \in [\rho^{-3/4}, \rho^{1/4}]$. By (8) of (A2) and (51), we will assume in the following that $V \in [0, 1], \sigma_t \in \mathcal{M}$.

We define $\sigma^p$ the process of dimension $4 \times 4$ on $\mathbb{R}^+$ such that

$$\sigma^p := \begin{cases} \sigma & \forall t \in [0, 1], \\ 0 & \forall t \in [1, \infty). \end{cases}$$

Define now $X^p$ the process such that for all $t \geq 0$

$$dX^p_t = \sigma^p_t dW_t, \quad X^p_0 = x_0.$$

Because $X^p$ and $X$ have the same initial value and follow the same stochastic differential equation on $[0, 1]$, they are equal for all $t \in [0, 1]$. For simplicity, we keep from now on the notation $X$ for $X^p$.

In the following, $C$ will be a constant which does not depend on $\epsilon$ or $n$, but that can vary from a line to another. Also, we are going to use the notation $\tau^i_n$ as a substitute of $\tau^i_{h_n}$, where $\theta$ can take various names, such that (1), (2) and so on. Let $h : \mathbb{N} \to \mathbb{N}$ (not strictly) increasing non-random sequence such that

$$h_n \to +\infty, \quad h_n \alpha_n \to 0.$$

To keep notation as simple as possible, we define $\tilde{t}^i_{h_n} := \tau^i_{h_n}$, $\tilde{t}^i_{h_n} := \tau^i_{h_n} - \tilde{t}^i_{h_n} := \tilde{t}^i_{h_n}$. We also let $A_n := (i \geq 1 s.t. \tilde{t}^i_{h_n} \leq t)$, where $t \in [0, 1]$. Also, we recall the notation $(X^{(e)}_t, X^{(s)}_t) := (X^{(e)}_{\tilde{t}^i_{h_n}}, X^{(s)}_{\tilde{t}^i_{h_n}})$, and $\Delta \tilde{X}_{\tilde{t}^i_{h_n}}$. Finally, for $\theta \in \{1, (2), 1C, h\}$, we define $\tilde{C}_n := \sup_{\theta \in \Delta \tilde{X}_{\tilde{t}^i_{h_n}}}$. We show that these quantities tend to 0 almost surely in the following lemma.

Lemma 5. We have $\tilde{C}_n \alpha_n \to 0$.

Proof. We can follow the proof of Lemma 4.5 in Robert and Rosenbaum (2012) to prove that for $k \in \{1; 2\}, s^{(k)} \alpha_n \to 0$. Then, we can notice that a.s. $s^{(1)} \alpha_n < \tilde{a}_n + \tilde{s}^{(1)} \alpha_n$, to deduce that $s^{(1)} \alpha_n \to 0$. To show that $\tilde{C}_n \to 0$, define the process $Z$ such that $Z_0 = 0$ and $\forall t > 0$ we have

$$Z_t := \begin{cases} \Delta X^{(2)}_{\tilde{t}^i_{h_n}} + Z^{(e)}_{\tilde{t}^i_{h_n}} & \forall t \in [\tilde{t}^i_{h_n}, \tilde{t}^i_{h_n}], \\ \Delta X^{(1C)}_{\tilde{t}^i_{h_n}} + Z^{(s)}_{\tilde{t}^i_{h_n}} & \forall t \in [\tilde{t}^i_{h_n}, \tilde{t}^i_{h_n}], \end{cases}$$

Substituting $X$ in Lemma 4.5 of Robert and Rosenbaum’s proof by our $Z$, we can follow the same reasoning. The only main change will be that in their notation $M_n \leq C_n h_n \alpha_n$, but this tends to 0 by (54).

Let $f$ be a random process, $s$ a random number, we define

$$S(f, s) := \sup_{0 \leq u, v \leq 1, \epsilon = \epsilon} |f_u - f_v|.$$
Lemma 6. Let $f$ be a bounded random process such that for all non-random sequences $(q_n)_{n \geq 0}$ if $q_n \rightarrow 0$, then $S(f, q_n) \rightarrow 0$. Let also a random sequence $(s_n)_{n \geq 0}$ such that $s_n \rightarrow 0$. Then we have $\forall i \geq 1$ that $S(f, s_n) \rightarrow 0$.

Proof. As $f$ is bounded, convergence in $\mathbb{P}$ implies convergence in $\mathbb{L}^1$ for any $i \geq 1$. Hence it is sufficient to show that $S(f, s_n) \rightarrow 0$. Let $\eta > 0$ and $\epsilon > 0$, we want to show that $\exists N > 0$ such that $\forall n \geq N$, we have

$$\mathbb{P}(S(f, s_n) > \eta) < \epsilon.$$

We need some definitions first. Let $\mathcal{F}_i$ be a sequence of independent 4-dimensional Brownian motions (i.e. for each $i$, $\mathcal{F}_i$ is a 4-dimensional Brownian motion), independent of everything we have defined so far. We define $\forall i, n \geq 0$,

$$S_n^i := \left\{ g_n^i \left( \Delta W_{I_n^i, I_{n+1}^i} \right) \right\} \forall t \in [0, \Delta t_{I_{n+1}^i}],$$

and

$$(\tilde{t}_{i,j,n}^k)_{i,j,n,k \geq 0} = \tilde{t}_{i,j,n} = t_i^{(2)} - t_j^{(1)} + t_{i,n}^{(1)} - t_{j,n}^{(2)}.$$
Proof. For any Brownian motion \((W_t)_{t \geq 0}\), by the scale property we have that \((W_t)_{t \geq 0} \overset{d}{=} (\alpha^{-1} W_{\alpha t})_{t \geq 0}\). Thus, if we define \(\tau = \inf\{t > 0 \text{ s.t. } W_t \not\in [d(t), u(t)]\}\) and \(\tau_\alpha = \inf\{t > 0 \text{ s.t. } W_t \not\in [\alpha d(t), \alpha u(t)]\}\), we have that
\[
\tau \overset{d}{=} \inf\{t > 0 \text{ s.t. } W_{\alpha^2 t} \not\in [\alpha d(t), \alpha u(t)]\} \overset{d}{=} \alpha^2 \tau_\alpha.
\]
We deduce that
\[
(\tau, W_\tau) \overset{d}{=} (\alpha^2 \tau_\alpha, W_{\alpha^2 \tau_\alpha}) \overset{d}{=} (\alpha^2 \tau_\alpha, \alpha W_\tau).
\]
We can prove the lemma based on the way we proved \((58)\), at the cost of 2-dimension definitions that would be more involved and straightforward applications of Strong Markov property of Brownian motions that we will not write, so that we do not lose ourselves in the technicality of this proof. \(\square\)

We introduce the number of points in the \(i\)th block in the \(k\)th process as the following:
\[
N_{1,n}^{(k)} = \max\{j \geq 0 \text{ s.t. } \tau_{1,n}^j + \tau_{1,n}^{(k)} \leq \tau_{1,n}^{(k)}\}.
\]
We also introduce the total number of points in the block \(N_{1,n} = N_{1,n}^{(1)} + N_{1,n}^{(2)}\). We show now that we can control uniformly the error of the approximations of the observation times.

**Lemma 11.** Let \(i \geq 1\), we have that
\[
\sup_{i \geq 0} \mathbb{E}\left[\left|\Delta \hat{t}_{1,n}^{IC} - \Delta \hat{t}_{1,n}^{IC}\right|^p\right] = o_p\left(\alpha_n^2\right)
\]
and
\[
\sup_{i \geq 0} \mathbb{E}\left[\left|\Delta \hat{t}_{1,n}^{IC} - \Delta \hat{t}_{1,n}^{IC}\right|^p\right] = o_p\left(\alpha_n^2\right).
\]

**Proof.** We introduce the notation \(o_p^{(U)}\) where \(U\) stands for "uniformly in \(i \geq 0\)\), meaning that the sup of the rests is of the given order.

**First step:** We define \(\hat{t}_{1,n}^{IC} = \sup_{i \in \mathbb{N}_0} \hat{t}_{1,n}^{IC}\). We show in this step that
\[
\hat{t}_{1,n}^{IC} \overset{p}{\rightarrow} 0.
\]
We define the accumulated time of approximated durations, i.e.
\[
\hat{t}_{1,n}^{IC} = \sum_{l=0}^{\infty} \hat{t}_{1,n}^{IC}.
\]
Using **Lemma 7** together with **Lemma 8**, \(\exists \theta > 0\) such that
\[
P\left(\hat{t}_{1,n}^{IC} \leq \theta\right) \rightarrow 1.
\]
We define \(Z_{1,n}^{(0)} = 0 \text{ and } \forall t \in \left[\hat{t}_{1,n}^{IC}, \hat{t}_{1,n}^{IC}\right]\)
\[
Z_{1,n}^{(0)} = Z_{1,n}^{(0)} + c_{1-n,1}^{(0)} - c_{1-n,1}^{(0)}.
\]
A slight modification of the proof of **Lemma 5** will conclude.

**Second step:** We show that we can do a localization in the number of observations in the \(i\)th block, i.e. there exists a non-random \(M_0\) such that
\[
P\left(\max\{N_{1,n}^{(1)}, N_{1,n}^{(2)}\} > M_0\right) \rightarrow 1.
\]
Converges uniformly (in \(i\)) towards 0 and \(M_0\) increasing at most linearly with \(h_n\), i.e. we have \(M_0 \leq \beta h_n\) where \(\beta > 0\).

To prove (62), we need some definitions. Define for \(i \geq 0\) the order of observation times \(O_{i,k,n}\) and the order of the approximated observation times \(\hat{O}_{i,k,n}\) in the following way. Let \(\hat{T}_{i,n}^{O} := \{\hat{t}_{i,n}^{(j)}\}_{j \geq 0}\) be the sorted set of all observation times (corresponding to processes 1 and 2) strictly greater than \(\hat{t}_{i,n}^{(j)}\). Then for \(j \geq 1\), we will set \(O_{i,k,n} = 1\) if the \(j\)th observation time in \(\hat{T}_{i,n}^{O}\) corresponds to an observation of the first process and \(O_{i,k,n} = 2\) if it corresponds to an observation of the second process. Similarly, we set \(\hat{T}_{i,n}^{k}\) the sorted set of all approximated times \(\{\hat{t}_{i,n}^{(j)}\}_{j \geq 0}\). \(\hat{O}_{i,k,n}\) are defined in the same way. There exists a \(p > 0\) such that for all integers \(i, j, n:\)
\[
P\left(O_{i,j+1,n} = 1\right) O_{i,j,n} = 2 \geq p \text{ and } P\left(O_{i,j+1,n} = 2\right) O_{i,j,n} = 1 > p.
\]
Indeed, let \(l\) be the (random) index such that \(\hat{t}_{i,n}^{(1)} = \hat{t}_{i,n}^{(2)}\).

Conditionally on \(O_{i,j,n} = 1\), we know that \(O_{i,j+1,n} = 2\) if \(\Delta X^{(1)}_{\hat{t}_{i,n}^{(1)} + \hat{t}_{i,n}^{(2)}} > g^+\) or \(-g^-\) before \(\Delta X^{(2)}_{\hat{t}_{i,n}^{(2)} + \hat{t}_{i,n}^{(1)}} > g^+\) or \(-g^-\). Using (8) of (A2) and (51), we can easily bound away from 0 this probability, thus we deduce (63). Now, using (22) together with (63) and strong Markov property of Brownian motions, we deduce (62).

**Third step:** let \(g = (d, u)\) such that \((g, g) \in \mathbb{G}, \sigma \in [\sigma^-, \sigma^+]\) and \(\epsilon \leq \frac{\epsilon}{2}\). We define \(\tau(g, \sigma, \epsilon) = \inf\{t > 0 : \sigma W_t = u(t) + \epsilon \text{ or } \sigma W_t = d(t) - \epsilon\}\), where \(W_t\) is a standard Brownian motion. We show that
\[
E\left[\tau(g, \sigma, \epsilon) - \tau(g, \sigma, 0)\right] = o_p\left(\alpha_n^2\right)
\]
where \(\gamma^{(i)}(\epsilon) \stackrel{\epsilon \rightarrow 0}{\longrightarrow} 0\).

In order to show (64), let
\[
\gamma^{(i)}(\epsilon) \stackrel{\epsilon \rightarrow 0}{\longrightarrow} 0.
\]

By (9) and (11) of (A3), we have \(\tau(g, \sigma, \epsilon) \leq \tau(g, \sigma, 0)\). Conditionally on \(\tau(g, \sigma, \epsilon)\), using strong Markov property of Brownian motions, we can show that \(E_{(g, \sigma, \epsilon)}\)
\[
\left[\tau(g, \sigma, \epsilon)\right] \stackrel{\epsilon \rightarrow 0}{\longrightarrow} 0 \text{ using Theorem 2 in Pötzlberger and Wang (2001) for instance.}
\]

**Fourth step:** Let \(k \in \{1, 2\}\). We show here that
\[
\sum_{j \leq M_0} E\left[\left|\hat{t}_{1,n}^{(k)} - \hat{t}_{1,n}^{(k)}\right|^p\right] = o_p\left(\alpha_n^2\right).
\]

The idea is to show that by recurrence in \(j\), \(E\left[\left|\hat{t}_{1,n}^{(k)} - \hat{t}_{1,n}^{(k)}\right|^p\right]\) can be arbitrarily small when \(n\) grows. It is then a straightforward analysis exercise to use the localization in second step and choose a different sequence \(h\) if necessary, that will still be non-random increasing and following (53) and (54), so that the sum in (65) will be also arbitrarily small. Let us start with \(j = 1\) and \(k = 1\).
where $E_{i,n} = E_{i,n}^{(1)} \cap E_{i,n}^{(2)}$ with

$$E_{i,n}^{(1)} = \left\{ \sup_{s \geq [\tau_{i,n}^{(b)} + T_{i,n}^{(b)}], t_i(n,1a)} |\Delta X_{i,n}^{(1)}(s) - \Delta \tilde{X}_{i,n}^{(1)}(s)| < \eta_{i,n,1} \right\},$$

$$E_{i,n}^{(2)} = \left\{ \sup_{s \geq [\tau_{i,n}^{(b)} + T_{i,n}^{(b)}], t_i(n,1a)} \| g_s^{(1)} - g_{\tilde{s}}^{(1)} \|_{L_1} < \eta_{i,n,1} \right\},$$

$\eta_{i,n} = \eta_0 \alpha_n, \quad \eta_n = \max \left\{ \eta_0, \frac{d - 1/2}{2}, \frac{z_n}{2} \right\}$ and $z_n = \sup_{1 \leq u, v \leq 4} \left( E \left( \left( G \sigma_{n,\eta,0}^{(u,v)} \right)^2 \right) \right)^{1/2}$. By (58) and (64),

$$E \left[ \left| t_{i,n}^{(k)} - \tilde{t}_{i,n}^{(k)} \right| 1_{E_{i,n}} \right] \leq C \alpha_n^2 \left( \gamma(2q_n) + \gamma(-2q_n) \right).$$

Using Cauchy–Schwarz inequality and Lemma 7,

$$E \left[ \left| t_{i,n}^{(k)} - \tilde{t}_{i,n}^{(k)} \right| 1_{E_{i,n}} \right] \leq C \alpha_n^2 \left( \left( E \left( \left| t_{i,n}^{(k)} - \tilde{t}_{i,n}^{(k)} \right| ^2 \right) \right)^{1/2} \right).$$

On the other hand,

$$E \left( \left( E_{i,n}^{(1)} \right)^C \right) \leq (\eta_{i,n,1}^{-1}) \left[ \sup_{s \geq [\tau_{i,n}^{(b)} + T_{i,n}^{(b)}], t_i(n,1a)} |\Delta X_{i,n}^{(1)}(s) - \Delta \tilde{X}_{i,n}^{(1)}(s)| \right]$$

$$\leq C (\eta_{i,n,1}^{-1}) \left( \left( \int_{\tau_{i,n}^{(b)} + T_{i,n}^{(b)}}^{t_i(n,1a)} (\sigma_{n,\eta,0}^{(u,v)} - \sigma_{n,\eta,0}^{(u,v)})^2 ds \right)^{1/2} \right)$$

$$\leq C (\eta_{i,n,1}^{-1}) \left( \left( \int_{\tau_{i,n}^{(b)} + T_{i,n}^{(b)}}^{t_i(n,1a)} \sigma_{n,\eta,0}^{(u,v)} \sigma_{n,\eta,0}^{(u,v)} ds \right)^{1/2} \right)$$

$$\leq C (\eta_{i,n,1}^{-1}) \left( \left( \int_{\tau_{i,n}^{(b)} + T_{i,n}^{(b)}}^{t_i(n,1a)} \sigma_{n,\eta,0}^{(u,v)} \sigma_{n,\eta,0}^{(u,v)} ds \right)^{1/2} \right)$$

$$\leq C \alpha_n^{-d/2},$$

where we used Markov inequality in the first inequality, conditional Burkholder–Davis–Gundy inequality in the second inequality, Cauchy–Schwarz inequality in the fourth inequality, and Lemma 7 in the last inequality. On the other hand,

$$E \left( \left( E_{i,n}^{(2)} \right)^C \right) \leq (\eta_{i,n,1}^{-1}) \left[ \sup_{s \geq [\tau_{i,n}^{(b)} + T_{i,n}^{(b)}], t_i(n,1a)} \| g_s^{(1)} - g_{\tilde{s}}^{(1)} \|_{L_1} \right]$$

$$\leq C (\eta_{i,n,1}^{-1}) \left( \left( \int_{\tau_{i,n}^{(b)} + T_{i,n}^{(b)}}^{t_i(n,1a)} \sigma_{n,\eta,0}^{(u,v)} \sigma_{n,\eta,0}^{(u,v)} ds \right)^{1/2} \right)$$

$$\leq C \alpha_n^{-d/2},$$

where we used Markov inequality in the first inequality, (12) of (A3) in the second inequality, and Lemma 7 in the last inequality. In summary, we have

$$E \left[ \left| t_{i,n}^{(k)} - \tilde{t}_{i,n}^{(k)} \right| \right] \leq C \alpha_n^2 \left( \gamma(2q_n) + \gamma(-2q_n) + z_n + d \right)$$

which we can make arbitrarily small, because $z_n \to 0$ by first step together with Lemma 6 and the continuity of $\sigma$ (A1). The case with $k = 2$ is very similar. Finally, for $j > 1$, the same kind of computation techniques, using in addition (11) of (A3), will work.

Fifth step: Prove that uniformly (in $i$)

$$P \left( \forall j \leq M, 0_{i,j,n} = 0_{i,j,n} \right) \to 1.$$

To show (66), let $i \leq M$. We define the (random) index $v$ such that $t_{i,j,n}^{(k)} = t_{i,j,n}^{(k,v)}$. Modifying suitably $h$ if needed, there exists (using fourth step) a sequence $(\epsilon_n)$ such that

$$P \left( \left| t_{i,j,n}^{(k,v)} - t_{i,j,n}^{(k,v)} \right| \leq \epsilon_n \right) \to 1,$$

$$P \left( \left| t_{i,j,n}^{(k,v)} \right| \leq \epsilon_n \right) \to 0.$$

Using (67) and (68), we can verify (66) by recurrence.

Sixth step: We prove here (59) and (60). Using Lemma 7 and (66) we obtain

$$E \left[ \Delta t_{i,j,n}^{(k)} - \Delta t_{i,j,n}^{(k)} \right] = E \left[ \Delta t_{i,j,n}^{(k)} - \Delta t_{i,j,n}^{(k)} \right] + \alpha_n^2 \left( \alpha_n^2 \right).$$

The first term on the right part of the inequality can be bounded by

$$E \left[ \Delta t_{i,j,n}^{(k)} - \Delta t_{i,j,n}^{(k)} \right] = E \left[ \Delta t_{i,j,n}^{(k)} - \Delta t_{i,j,n}^{(k)} \right] + \alpha_n^2 \left( \alpha_n^2 \right).$$

Both terms can be treated with the same trick. Using the second step and Lemma 7, the first term is equal to

$$\sum_{1 \leq M} E \left[ \left| t_{i,j,n}^{(k)} - \tilde{t}_{i,j,n}^{(k)} \right| \right]$$

and using (65), we prove (59). We can deduce (60) with the same kind of computations. \qed

Let $M^n$ be the interpolated normalized error, i.e.

$$M^n = \alpha_n^{-1} \left( \sum_{1 \leq i} \Delta X_{i,n}^{(1)} - t_{i,n}^{(1)} \right)^2 \Delta X_{i,n}^{(2)} - t_{i,n}^{(2)} \right)^2 \right)^{1/2} \right)$$

$$M^n$$ corresponds exactly to the normalized error of the Hayashi–Yoshida estimator if we observe the price of both assets at time $t$. We recall the definition of

$$N_{i,n} = \Delta X_{i,n}^{(1)} - \Delta X_{i,n}^{(1)} - \sum_{1 \leq i} \sigma_{n,\eta,0}^{(1)} \sigma_{n,\eta,0}^{(2)} \rho_{i,1} \left( \Delta X_{i,n}^{(2)} - t_{i,n}^{(2)} \right)^2 ds.$$

\textbf{Lemma 12.} We have

$$\sum_{1 \leq i} E_{i,n}^{(1)} \left[ \left( \Delta M_{i,n}^{(k)} \right)^2 \right]$$

$$= n^{-2} \sum_{1 \leq i} E_{i,n}^{(1)} \left[ \sum_{u=2}^n \left( (\tilde{N}_{i(n-1)u} + u) \right)^2 + 2N_{(i(n-1)u) + u} + N_{(i(n-1)u) + u} \right]$$

$$+ o_i(1).$$

\textbf{Proof.} We obtain this equality noting that $(N_{i,n})$ are centered and 1-correlated, and that the terms left converge to 0 in probability. \qed
We introduce the observation time at the start of a block, where “s” stands for “start”

\[ \tau_{s,n}^l = \sup \{ \tau_{r,n}^h \text{ s.t. } \tau_{r,n}^h < \tau_{s,n}^C \} . \]

**Lemma 13.** We have
\[
\alpha_n^{-2} \sum_{i=0}^{h_n} \mathbb{E}_{\tau_{1-n}^i} \left[ \sum_{u=2}^{h_n} \left( N_{(i-1)h_n+u}^2 + 2N_{(i-1)h_n+u}N_{(i-1)h_n+u+1} \right) \right] = \alpha_n^{-2} \sum_{i=0}^{h_n} \sum_{u=2}^{h_n} \int_{0}^{\tau_{1-n}^i} \psi^{(i)}(\sigma_{\tau_{1-n}^i}^i, g_{\tau_{1-n}^i}^i, \alpha_n^{-2} v, \alpha_n^{-2} v) \times d\xi_{\tau_{1-n}^i}(x, v) + o_p(1) .
\]

**Proof.** First step: approximating with holding volatility constant.

Set
\[ \tilde{N}_{n,i} = (\sigma_{\tau_{1-n}^i}^i \Delta W_{i,1}^{(1)})(\sigma_{\tau_{1-n}^i}^i \Delta W_{i,1}^{(2)} - \int_{0}^{\tau_{1-n}^i} \zeta_{i,1-n}^{1,2} ds) \]
where \( A^{(i)} \) is the \( i \)th component of the vector \( A \). We want to show that
\[
\alpha_n^{-2} \sum_{i=0}^{h_n} \mathbb{E}_{\tau_{1-n}^i} \left[ \sum_{u=2}^{h_n} \left( N_{(i-1)h_n+u}^2 + 2N_{(i-1)h_n+u}N_{(i-1)h_n+u+1} \right) \right] = \alpha_n^{-2} \sum_{i=0}^{h_n} \sum_{u=2}^{h_n} \left( \tilde{N}_{n,i}^2 \right) + o_p(1) .
\]
Noting \( F_{i,n} = (\tilde{N}_{n,i})^2 + 2N_{i,n}N_{i+1,n} \) and \( \tilde{F}_{i,n} = (\tilde{N}_{n,i})^2 + 2N_{i,n}N_{i+1,n} \), it is sufficient to show that
\[
\alpha_n^{-2} \sum_{i=1}^{\infty} \mathbb{E}_{\tau_{1-n}^i} \left[ \left| F_{i,n} - \tilde{F}_{i,n} \right| \mathbb{1}_{[\tau_{1-n}^i < t]} \right] \xrightarrow{p} 0.
\]
that we can rewrite as \( \alpha_n^{-2} \sum_{i=1}^{\infty} \mathbb{E}_{\tau_{1-n}^i} \left[ \left| F_{i,n} - \tilde{F}_{i,n} \right| \mathbb{1}_{[\tau_{1-n}^i < t]} \right] \xrightarrow{p} 0 \). Using **Lemma 9**, it is sufficient to show that \( \forall u > 0 \):
\[
\alpha_n^{-2} \sum_{i=1}^{\infty} \mathbb{E}_{\tau_{1-n}^i} \left[ \left| F_{i,n} - \tilde{F}_{i,n} \right| \mathbb{1}_{[\tau_{1-n}^i < t]} \right] \xrightarrow{p} 0.
\]
Thus, it is sufficient to show the convergence \( L^1 \) of this quantity, i.e. that
\[
\alpha_n^{-2} \sum_{i=1}^{\infty} \mathbb{E}_{\tau_{1-n}^i} \left[ \left| F_{i,n} - \tilde{F}_{i,n} \right| \mathbb{1}_{[\tau_{1-n}^i < t]} \right] \xrightarrow{L^1} 0.
\]
We have that
\[
\left| F_{i,n} - \tilde{F}_{i,n} \right| \leq B_{(1)}^{(i)} + B_{(2)}^{(i)},
\]
where \( B_{(1)}^{(i)} = |N_{n,i}^2 - \tilde{N}_{n,i}^2| \) and \( B_{(2)}^{(i)} = |N_{i-1,n}N_{i,n} - \tilde{N}_{i-1,n}\tilde{N}_{i,n}| \). We have that
\[
B_{(1)}^{(i)} \leq C_{(1)}^{(i)} + C_{(2)}^{(i)} + C_{(3)}^{(i)} ,
\]
where
\[
C_{(1)}^{(i)} = \left( \frac{\Delta X_{1-n}^{(1)} \Delta X_{1-n}^{(2)}}{\tau_{1-n}^{(1)} \tau_{1-n}^{(2)}} \right)^2 - \left( \sigma_{\tau_{1-n}^{(1)}}^i \Delta W_{1-n}^{(1)} \right)^2 \left( \sigma_{\tau_{1-n}^{(2)}}^i \Delta W_{1-n}^{(2)} + \int_{0}^{\tau_{1-n}^{(2)}} \zeta_{i,1-n}^{1,2} ds \right)^2 ,
\]
\[
C_{(2)}^{(i)} = \left( \sigma_{\tau_{1-n}^{(1)}}^i \Delta W_{1-n}^{(1)} \right)^2 \left( \sigma_{\tau_{1-n}^{(2)}}^i \Delta W_{1-n}^{(2)} + \int_{0}^{\tau_{1-n}^{(2)}} \zeta_{i,1-n}^{1,2} ds \right)^2 ,
\]
\[
C_{(3)}^{(i)} = \left( \sigma_{\tau_{1-n}^{(1)}}^i \Delta W_{1-n}^{(1)} \right)^2 \left( \sigma_{\tau_{1-n}^{(2)}}^i \Delta W_{1-n}^{(2)} + \int_{0}^{\tau_{1-n}^{(2)}} \zeta_{i,1-n}^{1,2} ds \right)^2 .
\]
Let us show that \( \alpha_n^{-2} \sum_{i=1}^{\infty} \mathbb{E}_{[\tau_{1-n}^i < t]} \left| C_{(1)}^{(i)} \right| \xrightarrow{p} 0 \). We can write it as \( C_{(1)}^{(i)} \leq D_{(1)}^{(i)} + D_{(2)}^{(i)} \), where
\[
D_{(1)}^{(i)} = \left( \sigma_{\tau_{1-n}^{(1)}}^i \Delta W_{1-n}^{(1)} \right)^2 \left( \sigma_{\tau_{1-n}^{(2)}}^i \Delta W_{1-n}^{(2)} + \int_{0}^{\tau_{1-n}^{(2)}} \zeta_{i,1-n}^{1,2} ds \right)^2 ,
\]
\[
D_{(2)}^{(i)} = \left( \sigma_{\tau_{1-n}^{(1)}}^i \Delta W_{1-n}^{(1)} \right)^2 \left( \sigma_{\tau_{1-n}^{(2)}}^i \Delta W_{1-n}^{(2)} + \int_{0}^{\tau_{1-n}^{(2)}} \zeta_{i,1-n}^{1,2} ds \right)^2 .
\]
\[
\alpha_n^{-2} \sum_{i=1}^{\infty} \mathbb{E}_{[\tau_{1-n}^i < t]} \left| D_{(1)}^{(i)} \right| \xrightarrow{p} 0 \]. We define:
\[
E_{(1)}^{(i)} = \Delta X_{1-n}^{(1)} \Delta X_{1-n}^{(2)} ,
\]
\[
E_{(2)}^{(i)} = \left( \sigma_{\tau_{1-n}^{(1)}}^i \Delta W_{1-n}^{(1)} \right)^2 \left( \sigma_{\tau_{1-n}^{(2)}}^i \Delta W_{1-n}^{(2)} + \int_{0}^{\tau_{1-n}^{(2)}} \zeta_{i,1-n}^{1,2} ds \right)^2 .
\]
Using Cauchy–Schwarz inequality, we deduce:
\[
\mathbb{E} \left[ \left( E_{(1)}^{(i)} + E_{(2)}^{(i)} \right)^2 \right] \leq \mathbb{E} \left[ \left( E_{(1)}^{(i)} + E_{(2)}^{(i)} \right)^2 \right] \mathbb{E} \left[ \left( E_{(1)}^{(i)} + E_{(2)}^{(i)} \right)^2 \right] .
\]
Using Cauchy–Schwarz inequality together with Burkholder–Davis–Gundy inequality and **Lemma 7**, we obtain that:
\[
\mathbb{E} \left[ \left( E_{(1)}^{(i)} + E_{(2)}^{(i)} \right)^2 \right] = o(U) .
\]
where \( U \) stands for “uniformly in \( 1 \leq i \leq \alpha_n^2 \). Another application of Cauchy–Schwarz inequality gives us
\[
\mathbb{E} \left[ \left( E_{(1)}^{(i)} + E_{(2)}^{(i)} \right)^2 \right] \leq \mathbb{E} \left[ \left( \Delta X_{1-n}^{(1)} \Delta W_{1-n}^{(1)} \right)^4 \right] .
\]
Using once again Cauchy–Schwarz inequality together with Burkholder–Davis–Gundy inequality and **Lemma 7**, we obtain that:
\[
\mathbb{E} \left[ \left( \Delta X_{1-n}^{(2)} \right)^4 \right] = o(U) .
\]
Similarly, we compute using conditional Burkholder–Davis–Gundy in first inequality, Cauchy–Schwarz in third inequality, Lemmata 5–7 together with the continuity of $\sigma$ (A1) in last equality.

\[
E \left[ \left( \Delta X_{t,n}^{(1)} - (\sigma_{t-n} \Delta W_{t,n}^{(i)})^{(1)} \right)^2 \right] = E \left[ \left( \Delta X_{t,n}^{(1)} - (\sigma_{t-n} \Delta W_{t,n}^{(i)})^{(1)} \right)^4 \right] \\
= E \left[ \left( \Delta X_{t,n}^{(1)} - (\sigma_{t-n} \Delta W_{t,n}^{(i)})^{(1)} \right)^2 \right] \\
\leq C \sup_{1 \leq j \leq 4} E \left[ \left( \Delta X_{t,n}^{(1)} - (\sigma_{t-n} \Delta W_{t,n}^{(i)})^{(1)} \right)^2 \right] \\
= C \sup_{1 \leq j \leq 4} E \left[ \left( \Delta X_{t,n}^{(1)} - (\sigma_{t-n} \Delta W_{t,n}^{(i)})^{(1)} \right)^2 \right] \\
\leq C \left( E \left[ (\Delta X_{t,n}^{(1)})^2 \right] \right) \sup_{1 \leq j \leq 4} E \left[ \left( (\sigma_{t-n} \Delta W_{t,n}^{(i)})^{(1)} \right)^2 \right]^{1/2} + o^3(\alpha_n^4) \\
= o^3(\alpha_n^4).
\]

With the same kind of computations, we show that $\alpha_n^{-2} \sum_{i=1}^{n-2} E \left[ C_{t,n}^{(2)} \right] \to 0$, and we also show $\alpha_n^{-2} \sum_{i=1}^{n-2} E \left[ C_{t,n}^{(2)} \right] \to 0$ (thus we have also that $\alpha_n^{-2} \sum_{i=1}^{n-2} E \left[ B_{t,n}^{(2)} \right] \to 0$ and $\alpha_n^{-2} \sum_{i=1}^{n-2} E \left[ B_{t,n}^{(2)} \right] \to 0$).

Second step: approximating using $(\tilde{\tau}_{i,n})_{i,n \geq 0}$. We set

\[
\tilde{N}_{i,n} = (\sigma_{t-n} \Delta W_{t,n}^{(i)})^{(1)} - \int_{t-n}^{t} (\sigma_{u-n} \Delta W_{u,n}^{(i)})^{(1)} \, du.
\]

We want to show that

\[
\alpha_n^{-2} \sum_{i=1}^{n-2} E \left[ B_{t,n}^{(2)} \right] \left( \tilde{N}_{i,n}^{(1)} \right)^2 + 2 \tilde{N}_{i,n-2} \tilde{N}_{i,n} + \tilde{N}_{i,n-1} \tilde{N}_{i,n+1} + 1 \to 0.
\]

Using the same kind of computations as in the first step together with Lemma 11, we conclude.

Third step: express the result as a function of $\psi^{AV}$. Using Lemma 10 in last equality, we deduce for any integer $u$ such that $2 \leq u \leq \tilde{h}_n$ that

\[
E \left[ \left( \tilde{N}_{i,n-1,n}^{(1)} \right)^2 \right] = \int_{t-1}^{t} \psi^{AV} \left( \sigma_{t-n} \Delta W_{t,n}^{(i)} \right) \tilde{N}_{i,n-1} \tilde{N}_{i,n} + \tilde{N}_{i,n-2} \tilde{N}_{i,n} + \tilde{N}_{i,n-1} \tilde{N}_{i,n+1} + 1 \to 0.
\]

Lemma 14. $\forall \sigma \in \mathcal{M}, \forall g \in \mathcal{G}, \exists \pi(\sigma, g)$ distribution such that:

\[
\alpha_n^2 \sum_{i=1}^{n-2} \sum_{u=0}^{n-3} \int \psi^{AV} \left( \sigma_{t-n} \Delta W_{t,n}^{(i)} \right) \tilde{N}_{i,n-1} \tilde{N}_{i,n} + \tilde{N}_{i,n-2} \tilde{N}_{i,n} + \tilde{N}_{i,n-1} \tilde{N}_{i,n+1} + 1 \to 0.
\]

Proof. We define the transition functions of the Markov chains $(\tilde{Z}_i(\sigma, g))_{i \geq 0}$ defined in (33). For $(x, u) \in \delta \mu$, $B \in \mathcal{B}_\mu$ (borelans of $\delta \mu$)

\[
P(\sigma, g) (x, u), B = \mathbb{E} \left[ \tilde{Z}_i(\sigma, g) \right] \right|_{i \geq 0} (x, u) \cdot B(\sigma, g).
\]

First step: We prove that $\forall \sigma \in \mathcal{M}, \forall g \in \mathcal{G}$, the state space $\delta \mu$ is $\nu$-small, i.e. there exists a non-trivial measure $\nu$ on $\mathcal{B}(\mathbb{R}^2)$ such that $\forall (\sigma, g) \in \mathcal{M}$, $\forall B \in \mathcal{B}_\mu$, $\mathbb{E} \left[ \tilde{Z}_i(\sigma, g) \right] (x, u) \geq \nu (B)$. Let $B = [x_0, x_1] \times [u_0, u_1]$. We are choosing $\nu$ such that $\nu (B) = 0$ outside $[\varepsilon, \varepsilon]^2$.

Thus, without loss of generality, we have that $[x_0, x_1] \times [u_0, u_1] \subseteq [\varepsilon, \varepsilon] \times [\varepsilon, \varepsilon]^2$. We want to show that $3 \varepsilon > 0$ such that uniformly

\[
P(\sigma, g) (x, u), B \geq c (x_0 - x_1) (u_0 - u_1).
\]

There are two ways to rewrite $(\tilde{X}^{(i)}, \tilde{X}^{(ii)})$. The first one is:

\[
\tilde{X}^{(i)} := \sigma^{(i)} \tilde{B}^{(i)},
\]

\[
\tilde{X}^{(ii)} := \rho^{(i)} \sigma^{(i)} \tilde{B}^{(ii)} + \left( 1 - \left( \rho^{(i)} \sigma^{(i)} \right)^2 \right)^{1/2} \sigma^{(i)} \tilde{B}^{(ii)} \cdot
\]

where $\tilde{B}^{(i)}$ and $\tilde{B}^{(ii)}$ are independent, $\rho^{(i)} \sigma^{(i)} \in [\rho^{(i)} \sigma^{(i)}, \rho^{(i)} \sigma^{(i)}]$, and $\delta = \left( 1 - \left( \rho^{(i)} \sigma^{(i)} \right)^2 \right)^{1/2}$.

The other way to rewrite it is:

\[
\tilde{X}^{(i)} := \sigma^{(i)} \tilde{B}^{(i)},
\]

\[
\tilde{X}^{(ii)} := \rho^{(i)} \sigma^{(i)} \tilde{B}^{(ii)} + \left( 1 - \left( \rho^{(i)} \sigma^{(i)} \right)^2 \right)^{1/2} \sigma^{(i)} \tilde{B}^{(ii)}.
\]

where $\tilde{B}^{(i)}$ and $\tilde{B}^{(ii)}$ are independent. For $(B_t)_{t \geq 0}$ a standard Brownian motion, $a < x < b$, we denote the exiting-zone time of the Brownian motion

\[
\inf \{ t > 0 : x + B_t = a \text{ or } x + B_t = b \}
\]

and $p(x, a, b, t)$ the density of $r_x^{(a,b)}$. We also define $p_2(x, a, b, s, y)$ the distribution of $B_t + x$ conditioned on $t^{(a,b)} = s$. Finally, let $p_3(x, a, b, t)$ be the distribution of $r_x^{(a,b)}$ conditioned on $B_t = b$. All the formulas can be found in Borodin and Salminen (2002). Consider the spaces $C_1 = C_2 = \{(x, a, b, t) \in \mathbb{R}^4 \mid a \leq x < b, t > 0\}$, $C_3 = \{(x, a, b, t, y) \in \mathbb{R}^4 \mid a \leq x < b, a < y < b, t > 0\}$. The functions $p_i$ are continuous on $C_i$ and positive. Thus, for all compact set $K \subseteq C_i$, we have

\[
\inf_{k \in K} p_i(k) > 0.
\]

We can bound below

\[
P(\sigma, g) (x, u), B \\
\geq \frac{C}{E_{0} \bigcap E_{1} \bigcap E_{2} \bigcap E_{3} \bigcap E_{4} \bigcap Z_0 = (x, u)}.
\]
where
\[ E_0 = \left\{ \sup_{0 \leq t \leq T_1} \| \dot{X}^{(3)}_t \| < \frac{\epsilon \sigma^- \min(\sigma^-, 1)}{15 \sigma^+}, \quad t_1 \leq T \right\}, \]
\[ E_1 = \left\{ \sup_{t_1 \leq t \leq T_1} \left| \dot{X}^{(3)}_t \right| < \frac{\epsilon \sigma^-}{10 \sigma^+}, \quad \sup_{t_1 \leq t \leq T_1} \| \Delta \hat{B}^{1, \perp}_{[t_1, t]} \| < \frac{g^*}{2 \sigma^+} \right\}, \]
\[ E_2 = \left\{ \sup_{K+1 \leq t \leq T_1} \left| \dot{X}^{(3)}_t \right| \leq \frac{\epsilon}{5} \right\}, \]
\[ E_3 = \left\{ \exists s \in [T_2, K+4], \dot{X}^{(3)}_s \in [d_1(K), u_1(K), \dot{X}^{(3)}_{K+4} \in [u_1(K) - 2\epsilon, u_1(K) + \epsilon] \right\}, \]
\[ \bigcup \left\{ \sup_{t_1 \leq t \leq T_1} \| \Delta \dot{X}^{(4)}_{[t_1, t]} \| \leq \frac{g^-}{12} \right\}, \]
\[ E_4 = \left\{ \frac{1}{K+1 \leq t \leq K+4}, \inf_{K+1 \leq t \leq T_1} \| \dot{X}^{(3)}_{[t_1, t]} \| > -2\epsilon \right\}, \]
\[ \bigcup \left\{ \sup_{K+1 \leq t \leq T_1} \| \Delta \dot{X}^{(4)}_{[t_1, t]} \| < \frac{g^-}{2 \sigma^+} \right\} \]
where \(\epsilon = \frac{\epsilon^*}{60\sigma^-}\). Using extensively Bayes formula, we can rewrite
\[ \mathbb{P} \left( E_0 \cap E_1 \cap E_2 \cap E_3 \cap E_4 \cap [\tilde{Z}_1 \in B] \right| \tilde{Z}_0 = (x, u) \right) = I \times II \times III \times IV \times V, \]
where I \(\equiv \mathbb{P} \left( E_0 \right| \tilde{Z}_0 = (x, u) \right), \quad II = \mathbb{P} \left( E_1 \cap E_0 \right| \tilde{Z}_0 = (x, u) \right), \quad III = \mathbb{P} \left( E_2 \cap E_1 \cap E_0 \right| \tilde{Z}_0 = (x, u) \right), \quad IV = \mathbb{P} \left( E_3 \cap E_2 \cap E_1 \cap E_0 \right| \tilde{Z}_0 = (x, u) \right), \quad V = \mathbb{P} \left( E_4 \cap E_3 \cap E_2 \cap E_1 \cap E_0 \right| \tilde{Z}_0 = (x, u) \right).

We prove that I is uniformly bounded away from 0. Using (52), (69), (70) and (71), we deduce that \(E_0^{(1)} \cap E_0^{(2)} \subset E_0\) where
\[ E_0^{(1)} = \left\{ \sup_{0 \leq t \leq K} \| \hat{B}^{(3)}_t \| < \frac{\epsilon \sigma^- \min(\sigma^-, 1)}{15 \sigma^+} \right\}, \]
\[ E_0^{(2)} = \left\{ \sup_{0 \leq t \leq K} \frac{x}{\sigma^+(1 - (3\rho^2)^2) + \hat{B}^{1, \perp}_t} \right\}, \]
\[ \geq \frac{g^*}{\delta \sigma^-} \left\| \frac{\epsilon \sigma^- \min(\sigma^-, 1)}{15 \sigma^+} \right\| \]
Conditionally on \([\tilde{Z}_0 = (x , u)], E_0^{(1)}\) and \(E_0^{(2)}\) are independent. Thus, we deduce
\[ I \geq \mathbb{P} \left( E_0^{(1)} \right) \mathbb{P} \left( E_0^{(2)} \right) \mathbb{P} \left( \tilde{Z}_0 = (x, u) \right). \]

Using Markov property of Brownian motions, we obtain that the right part of the inequality is equal to
\[ \left( 1 - \int_0^K p_1 \left( 0, \frac{\epsilon \sigma^- \min(\sigma^-, 1)}{15 \sigma^+}, \frac{\epsilon \sigma^- \min(\sigma^-, 1)}{15 \sigma^+}, t \right) dt \right) \times \int_0^K p_1 \left( y^{(1)}, -y^{(2)}, y^{(2)}, t \right) dt, \]
where \(y^{(1)} = \frac{x}{\sigma^+(1 - (3\rho^2)^2) + \hat{B}^{1, \perp}_t} + \frac{g^*}{\delta \sigma^-} \frac{\epsilon \sigma^- \min(\sigma^-, 1)}{15 \sigma^+}, y^{(2)} = \frac{g^*}{\delta \sigma^-} + \frac{\epsilon \sigma^- \min(\sigma^-, 1)}{15 \sigma^+}\), which is uniformly (in \(x, \sigma\) and \(g\) bounded away from 0 using (52) and (74).

We prove that II is uniformly bounded away from 0. Conditionally on \(E_0 \cap [\tilde{Z}_0 = (x, u)]\), the two quantities of \(E_1\) are independent. Thus, we bound below II (the same way we did for I) by
\[ \left( 1 - \int_0^{K+1} p_1 \left( \hat{B}^{(3)}_{t_1}, \frac{\epsilon \sigma^-}{10 \sigma^+} + \frac{\epsilon \sigma^-}{10 \sigma^+} \right), t \right) dt \]
\[ \left( 1 - \int_0^{K+1} p_1 \left( 0, -\frac{g^*}{4 \sigma^+} - \frac{g^*}{4 \sigma^+} \right), t \right) dt, \]
which is uniformly bounded away from 0 using (52) together with (74).

We prove that III is uniformly bounded away from 0. Using (52), (69), (70) and (71), we deduce that \(E_0 \cap E_2^{(2)} \subset E_2\) where
\[ E_2^{(1)} = \left\{ \sup_{K+1 \leq t \leq K+4} \| \hat{B}^{(3)}_t \| < \frac{\epsilon}{5 \sigma^+} \right\}, \]
\[ E_2^{(2)} = \left\{ \sup_{K+1 \leq t \leq K+4} \| \Delta \hat{B}^{1, \perp}_{[t_1, t]} \| < \frac{g^*}{2 \sigma^+}, \sup_{K+2 \leq t \leq K+3} \| \Delta \hat{B}^{1, \perp}_{[t_1, t]} \| < \frac{g^*}{2 \sigma^+} \right\}, \]
\[ \geq \frac{g^*}{\delta \sigma^-} + \frac{\epsilon}{5 \sigma^+} \].

Conditionally on \([\tilde{Z}_0 = (x, u)]\), \(E_2^{(1)}\) and \(E_2^{(2)}\) are independent. Thus, we deduce
\[ III \geq \mathbb{P} \left( E_2^{(1)} \right) \mathbb{P} \left( E_2^{(2)} \right) \mathbb{P} \left( \tilde{Z}_0 = (x, u) \right). \]

Using Markov property of Brownian motions, we obtain that the right part of the inequality conditioned on \([\tilde{B}^{(3)}_{K+1}, \Delta \hat{B}^{1, \perp}_{[t_1, K+1]}\]
\[ E_0 \cap E_2^{(2)} \cap [\tilde{Z}_0 = (x,u)] \]
\[ E_0^{(1)} \cap E_0^{(2)} \subset E_2\] conditioned on
\[ \left\{ \frac{-E_{K+1}^{(0)}}{\Delta \hat{B}^{(3)}_{[t_1, K+1]}}, \min \right\} \geq 1 \].

Using the definition of \(E_1\) together with (52) and (74), we have III which is uniformly bounded away from 0.

We prove that IV is uniformly bounded away from 0. Using (72) and (73), we deduce that \(E_3^{(1)} \cap E_3^{(2)} \subset E_3\) where
\[ E_3^{(1)} = \left\{ \sup_{t_1 \leq t \leq T_1} \| \Delta \hat{B}^{(4)}_{[t_1, t]} \| < \frac{\epsilon \sigma^-}{5 \sigma^+} \right\}, \]
\[ E_3^{(2)} = \left\{ \exists s \in [T_2, K+4], \Delta \hat{B}^{1, \perp}_{[t_1, t]} \in [\overline{y}^{(1)}, \overline{y}^{(2)}], \Delta \hat{B}^{1, \perp}_{[t_1, K+4]} \in [\overline{y}^{(3)}, \overline{y}^{(4)}] \right\}. \]
with \( y_2^{(1)} = \frac{\sigma_4^{-1}(1-\rho^4)}{\sigma_4^{-1}(1-\rho^4)} \) as well as \( y_2^{(3)} = \frac{u_1(K)-2\epsilon/5}{\sigma_4^{-1}(1-\rho^4)} \). Conditionally on \( E_1 \cap E_2 \cap E_0 \cap \{ \tilde{Z}_0 = (x, u) \} \), \( E_3^{(1)} \) and \( E_3^{(2)} \) are independent. Thus, we deduce

\[
I_V \geq P\left( E_3^{(1)} \bigg| E_1 \cap E_2 \cap E_0 \cap \{ \tilde{Z}_0 = (x, u) \} \right)
\]

Using Markov property of Brownian motions, we obtain that the right part of the inequality conditioned on \( \{ F_2 \} \cap E_1 \cap E_2 \cap E_0 \cap \{ \tilde{Z}_0 = (x, u) \} \) is equal to

\[
\left(1 - \int_0^{K+\epsilon+1/2} p_2 \left( 0, y_2^{(3)} \right) dy \right)^\Delta \int_0^{y_2^{(1)}} p_1 \left( 0, y_2^{(1)} \right) dy + \int_{y_2^{(1)}}^{y_2^{(2)}} p_2 \left( 0, y_3^{(1)} \right) dy,
\]

which is uniformly bounded away from 0 using (52), (71) and (74).

We prove that \( V > c(x_0-x_2)u_2 u_1 \). Using (69) and (70), we deduce that \( E_4^{(1)} \cap E_4^{(2)} \subset E_4 \) where

\[
E_4^{(1)} = \left\{ \tilde{t} \in [u_0 + 1, 2]; \tilde{t} \neq 0, \tilde{X}_1^{(1)} = u_1(K) \right\},
\]

\[
E_4^{(2)} = \left\{ \tilde{t} \in [u_0 + 1, 2]; \tilde{X}_1^{(2)} = u_1(K) \right\},
\]

\( \tilde{X}_1^{(1)} = u_1(K) \) or \( \tilde{X}_1^{(2)} = u_1(K) \). We have

\[
y_4^{(1)} = \frac{x_0 - \tilde{X}_1^{(3)} \sigma_4^{-1}(1-\rho^4) u_1(K)-\tilde{X}_1^{(3)} \sigma_4^{-1}(1-\rho^4)}{\sigma_4^{-1}(1-\rho^4)^2}.
\]

\[
y_4^{(2)} = \frac{x_0 - \tilde{X}_1^{(3)} \sigma_4^{-1}(1-\rho^4) u_1(K)-\tilde{X}_1^{(3)} \sigma_4^{-1}(1-\rho^4)}{\sigma_4^{-1}(1-\rho^4)^2}
\]

\( \tilde{X}_1^{(3)} = u_1(K) \) and \( \tilde{X}_1^{(3)} = u_1(K) \). We have

\[
V = P\left( \tilde{X}_1^{(3)} = u_1(K) \right) \times P\left( E_4^{(1)} \cap E_4^{(2)} \bigg| E_1 \cap E_2 \cap E_0 \cap \{ \tilde{Z}_0 = (x, u) \} \right).
\]

The first term on the right part of the equation is uniformly bounded away from 0 (Borodin and Salminen, 2002). Because \( \tilde{t} \) is a function of \( \tilde{X}_1^{(3)} \) and \( B_0^{\tilde{t}} \) is independent with \( \tilde{X}_1^{(3)} \), \( \tilde{v} \) and \( B_0^{\tilde{t}} \) are independent. Thus the second term on the right part conditioned on

\[
\{ y_4^{(1)}, y_4^{(2)}, X_1^{(3)}, \tilde{Z}_0 = (x, u) \}
\]

can be expressed as:

\[
\sum_{x_0+\tilde{t}_0^{(2)} \geq K+\epsilon} \int_y^{y_4^{(2)}} p_4 \left( \frac{X_1^{(3)} - 2\epsilon}{\sigma_3}, t \right) u_1(K) dt dy,
\]

\[
\times \int_{y_4^{(1)}}^{y_3^{(2)}} p_2 \left( 0, y_3^{(1)} \right) dt dy,
\]

where \( y_4^{(3)} = 6 \sigma_4^{-1}(1-\rho^4)^{1/2} \). We have that \( y_4^{(1)} \) and \( y_4^{(2)} \) are dominated by \( 3^4 \left( 1-\rho^4 \right)^{1/2} \). Using this together with (52), (71) and (74), we deduce that \( V \geq c(x_0-x_2)u_2 u_1 \).

Second step: We prove that \( \| \psi^{AV} \|_\infty = \sup_{\sigma \in M, g \in \mathcal{G}, (x, u) \in D} | \psi^{AV} (\sigma, g, x, u) | < \infty \). To show this, we bound the term as

\[
\left[ \left( \Delta X_1^{(3)} \right)^2 + \left( \Delta X_1^{(2)} \right)^2 \right]^{1/2} \leq 2 \| \psi^{AV} \|_\infty \left[ \left( \Delta X_1^{(3)} \right)^2 + \left( \Delta X_1^{(2)} \right)^2 \right]^{1/2}.
\]

The second term in the right-hand side of the inequality is uniformly bounded using (52) and Lemma 7. Using successively Cauchy–Schwarz and Burkholder–Davis–Gundy inequality, (52) and Lemma 7, we can also bound uniformly the first term. The other term of (29) can be bounded in the same way.

Third step: Define \( \tilde{q} = (\sigma, g, x, u) \) and \( \mathcal{Q} = \{ (\sigma, g, x, u) \in \mathcal{F} | \sigma \in M, g \in \mathcal{G}, (x, u) \in D \} \).

Prove that \( \forall q \in \mathcal{Q} \), there exists a measure \( \tilde{\pi} (\sigma, g) \) such that

\[
\sum_{q \in \mathcal{Q}} \int_0^1 \int_{\mathbb{R}^2} \psi^{AV} (\sigma, g, x, u) d\tilde{\pi} (\sigma, g, x, u) (y, v) dy
\]

\[
\leq \int_0^1 \psi^{AV} (\sigma, g, x, u) d\tilde{\pi} (\sigma, g, x, u) dy
\]

\[
= no_{\tilde{q}} (1).
\]

To show this, we use first step together with Th. 16.0.2 (Meyn and Tweedie, 2009). We obtain that there exists \( \tilde{\pi} (\sigma, g) \) where

\[
\| \psi^{AV} (\sigma, g, x, u) \|_\infty \leq 2r^n
\]

and \( r = 1 - \nu (\mathbb{R}^2) \). Thus, we deduce:

\[
\int_{\mathbb{R}^2} \psi^{AV} (\sigma, g, x, u) d\tilde{\pi} (\sigma, g, x, u) (y, v) dy
\]

\[
\leq \int_{\mathbb{R}^2} \psi^{AV} (\sigma, g, x, u) d\tilde{\pi} (\sigma, g, x, u) (y, v) dy
\]

\[
\leq \int_{\mathbb{R}^2} \psi^{AV} (\sigma, g, x, u) d\tilde{\pi} (\sigma, g, x, u) (y, v) dy
\]

\[
= no_{\tilde{q}} (1).
\]

We want to show that \( \forall \epsilon > 0, \exists N > 0 \) such that \( \forall n \geq N : \)

\[
\sum_{n=0}^{N-1} \int_0^1 \psi^{AV} (\sigma, g, x, u) d\tilde{\pi} (\sigma, g, x, u) (y, v) dy
\]

\[
\leq \int_0^1 \psi^{AV} (\sigma, g, x, u) d\tilde{\pi} (\sigma, g, x, u) (y, v) dy
\]

\[
\leq \int_0^1 \psi^{AV} (\sigma, g, x, u) d\tilde{\pi} (\sigma, g, x, u) (y, v) dy
\]

\[
< \epsilon n.
\]

The rest is a straightforward analysis exercise. Let \( \epsilon > 0, \exists N_1 > 0 \) such that \( r^{N_1} < \frac{\epsilon}{2} \). Choosing \( N > 8N_1 \epsilon^{-1} \| \psi^{AV} \|_\infty^{-1} \), we first use
the triangular inequality, and then split the sum of the left part of (76) into two parts, one up to \( N_t \) and the other one up to \( N \). We use (75) in the second part to obtain (76).

Fourth step: Proving the Lemma. Let \( w > 0 \). From Lemma 9, we just have to show that
\[
\alpha_n \sum_{i=1}^{w n} \int_{\mathbb{R}^2} \psi \left( \sigma_{t_{i-1},t_i} \cdot \mathbf{g}_{t_{i-1},t_i} \cdot \alpha_n^{-1} x, \alpha_n^{-2} u \right) \times d\mathcal{N}_{t_{i-1},t_i}(x, u) + o_p(1),
\]
tends to 0 in probability. Using third step together with standard results on regular conditional distributions (see for instance Section 4.3 (pp. 77–80) in Breiman (1992)), we prove the lemma. □

**Lemma 15.** We have
\[
\alpha_n^2 \sum_{i=1}^{h_n} \mathbb{E}_{t_{i-1},t_i} \left[ \left( \psi \left( \sigma_{t_{i-1},t_i} \cdot \mathbf{g}_{t_{i-1},t_i} \right) \right)^2 \right] \times \left( \mathbb{E}_{t_{i-1},t_i} \left[ \Delta \mathbf{r}_{t_{i-1},t_i} \right] \right)^{-1} = \mathbb{E}_{t_{i-1},t_i} \left[ \phi \left( \sigma_{t_{i-1},t_i} \cdot \mathbf{g}_{t_{i-1},t_i} \right) \Delta \mathbf{r}_{t_{i-1},t_i} \left( \alpha_n^{-1} \right) \right] + o_p(1).
\]

**Proof.** First step: Defining
\[
\begin{align*}
\alpha_n^2 \sum_{i=1}^{h_n} \mathbb{E}_{t_{i-1},t_i} \left[ \left( \psi \left( \sigma_{t_{i-1},t_i} \cdot \mathbf{g}_{t_{i-1},t_i} \right) \right)^2 \right] & \times \left( \mathbb{E}_{t_{i-1},t_i} \left[ \Delta \mathbf{r}_{t_{i-1},t_i} \right] \right)^{-1} \\
\alpha_n & = \alpha_n^2 \sum_{i=1}^{h_n} \mathbb{E}_{t_{i-1},t_i} \left[ \psi \left( \sigma_{t_{i-1},t_i} \cdot \mathbf{g}_{t_{i-1},t_i} \right) \Delta \mathbf{r}_{t_{i-1},t_i} \left( \alpha_n^{-1} \right) \right], \\
\end{align*}
\]
we have that \( A_0 = A_1 + o_p(1) \). To show this, in light of Lemma 11, we have that
\[
\mathbb{E}_{t_{i-1},t_i} \left[ \Delta \mathbf{r}_{t_{i-1},t_i} \right] - u_i \leq h(n) c_t,
\]
where \( C_t \) tends to 0 in probability. From this, we can easily show that \( A_0 = A_1 + o_p(1) \).

Second step: We have that
\[
\begin{align*}
A_1 & = \alpha_n^2 \sum_{i=1}^{h_n} \mathbb{E}_{t_{i-1},t_i} \left[ \psi \left( \sigma_{t_{i-1},t_i} \cdot \mathbf{g}_{t_{i-1},t_i} \right) \Delta \mathbf{r}_{t_{i-1},t_i} \left( \alpha_n^{-1} \right) \right] + o_p(1).
\end{align*}
\]
To prove it, we can mimic the proof of Lemma 14, together with Lemma 11. □

**A.3. Computation of the limits of \( \langle M^\alpha \rangle_t \), \( \langle M^\alpha, X^{(1)} \rangle_t \), and \( \langle M^\alpha, X^{(2)} \rangle_t \)**

\[
\begin{align*}
\langle M^\alpha \rangle_t & = \alpha_n^2 \sum_{i=1}^{h_n} \mathbb{E}_{t_{i-1},t_i} \left[ \left( \mathcal{M}^\alpha_{t_{i-1},t_i} \right)^2 \right] + o_p(1) \\
& = \alpha_n^2 \sum_{i=1}^{h_n} \mathbb{E}_{t_{i-1},t_i} \left[ \left( N_{t_{i-1},t_i} + 2 N_{t_{i-1},t_i} \right)^2 \right] + o_p(1) \\
& = \alpha_n^2 \sum_{i=1}^{h_n} \mathbb{E}_{t_{i-1},t_i} \left[ \left( \mathbf{g}_{t_{i-1},t_i} \cdot \mathbf{g}_{t_{i-1},t_i} \right)^2 \right] + o_p(1) \\
& = \mathbb{E}_{t_{i-1},t_i} \left[ \phi \left( \mathbf{g}_{t_{i-1},t_i} \right) \Delta \mathbf{r}_{t_{i-1},t_i} \left( \alpha_n^{-1} \right) \right] + o_p(1),
\end{align*}
\]
where we used Lemma 2.2.11 of Jacod and Protter (2012) in first equality, Lemma 12 in second equality, and Lemma 13 in third equality.

We deduce (using Lemma 14 in first equality and Lemma 15 in third equality)
\[
\langle M^\alpha \rangle_t = \alpha_n^2 \sum_{i=1}^{h_n} \mathbb{E}_{t_{i-1},t_i} \left[ \mathbf{g}_{t_{i-1},t_i} \Delta \mathbf{r}_{t_{i-1},t_i} \right] + o_p(1) \\
= \alpha_n^2 \sum_{i=1}^{h_n} \mathbb{E}_{t_{i-1},t_i} \left[ \mathbf{g}_{t_{i-1},t_i} \Delta \mathbf{r}_{t_{i-1},t_i} \right] + o_p(1).
\]

Using Lemma 2.2.11 of Jacod and Protter (2012) again, we deduce
\[
\langle M^\alpha \rangle_t = \sum_{i=1}^{h_n} \phi_{t_{i-1},t_i} \Delta \mathbf{r}_{t_{i-1},t_i} \left( \alpha_n^{-1} \right) + o_p(1).
\]

Using Lemma 5 together with Prop. 1.4.44 (page 51) in Jacod and Shiryaev (2003), we obtain
\[
\langle M^\alpha \rangle_t \to \int_0^t \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right] ds.
\]

Using the same approximations and computations, we also compute
\[
\langle M^\alpha, X^{(1)} \rangle_t \to \int_0^t \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right] ds,
\]
\[
\langle M^\alpha, X^{(2)} \rangle_t \to \int_0^t \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right] ds.
\]

**A.4. Computation of the asymptotic bias and variance**

We follow the idea in 1-dimension in pp. 155–156 of Mykland and Zhang (2012), and define an auxiliary martingale
\[
\tilde{M}_t = M_t - \int_0^t \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right] ds - \int_0^t \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right] ds,
\]
where \( X^{(1)} \) is defined in (39). Using (78), we deduce
\[
\tilde{M}_t, X^{(1)} \to \int_0^t \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right] ds - \int_0^t \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right] ds.
\]
Hence, we choose
\[
k_{t}^{(1)} = (\sigma_{t}^{(1)})^{-2} \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right]^{-1}.
\]
By the same techniques that we used to compute (78), we have that
\[
\int_0^t \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right] ds - \int_0^t \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right] ds.
\]

Using (79) and (80) we compute
\[
\tilde{M}_t, X^{(1)} = M_t - \int_0^t \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right] ds - \int_0^t \phi_{s} \mathbb{E} \left[ \phi_{s}^{-1} \right] ds.
\]
\[
\begin{align*}
\frac{\delta}{t} & \int_0^t \left( \phi_s \mathcal{AC}_2 - (\sigma_s^{(1)})^{-1} \phi_s \mathcal{AC}_1 \right) (\phi_t) - ds \\
- \int_0^t k_s \left( 1 - (\rho_{s}^{(2)})^2 \right) (\sigma_s^{(2)})^2 ds.
\end{align*}
\]

Hence, we choose
\[
k_s = \left( 1 - (\rho_{s}^{(2)})^2 \right)^{-1} \times \left( (\sigma_s^{(2)})^2 \phi_s \mathcal{AC}_2 - (\sigma_s^{(1)})^{-1} \phi_s \mathcal{AC}_1 \right) (\phi_t) - ds.
\]

By (A4), there exists \( S > 0 \) such that the \( S \) Brownian motions \( \{D^{(1)}, \ldots, D^{(5)}\} \) generate the filtration \( (\mathcal{F}_t)_{t \geq 0} \). To show that \( \widehat{M}_n, D^{(i)} \) tends to 0 in probability, we decompose \( D^{(i)} = D^{(1)} + D^{(2)} \) where \( D^{(1)} \) belongs to the space spanned by \( \{X^{(1)}, X^{(2)}\} \) and \( D^{(2)} \) is orthogonal to this space. By what precedes, we have clearly \( \widehat{M}_n, D^{(1)} \) tends to 0 in probability. Also, \( D^{(2)} \) is a martingale that is, conditionally on the observations times of both processes, independent of \( \widehat{M}_n \). Thus we also deduce that \( \widehat{M}_n, D^{(2)} \) converges to 0 in probability.

We can now compute
\[
\widehat{M}_n = M_n - \int_0^t k_s dX^{(1)} - \int_0^t k_{s-} dX^{(1)-}.
\]

\[
\int_0^t \left( (\sigma_s^{(1)})^{2} \left( 1 - (\rho_{s}^{(2)})^2 \right) (k_s^{(1)})^2 \right) ds
\]

\[
- \int_0^t k_s^{(1)} dX^{(1)} - \int_0^t k_{s-} dX^{(1)-}.
\]

\[
\frac{\delta}{t} \inf_{t} \left( \phi_t^{(1)} + 2 \left( (\sigma_s^{(1)})^{-1} \phi_s \mathcal{AC}_1 \right) \right) (\sigma_t^{(1)})^{-1}
\]

\[
+ \left( (\sigma_s^{(1)})^{-1} \phi_s \mathcal{AC}_1 \right) (\sigma_t^{(1)})^{-1}
\]

\[
+ \left( (\sigma_s^{(1)})^2 \phi_s \mathcal{AC}_1 \right) (\phi_t) - ds.
\]

By letting
\[
AV_t = \left( \phi_t^{(1)} + 2 \left( (\sigma_s^{(1)})^{-1} \phi_s \mathcal{AC}_1 \right) \right) (\sigma_t^{(1)})^{-1}
\]

\[
+ \left( (\sigma_s^{(1)})^2 \phi_s \mathcal{AC}_1 \right) (\sigma_t^{(1)})^{-1}
\]

\[
+ \left( (\sigma_s^{(1)})^{-1} \phi_s \mathcal{AC}_1 \right) (\sigma_t^{(1)})^{-1}
\]

\[
+ \left( (\sigma_s^{(1)})^{2} \phi_s \mathcal{AC}_1 \right) (\phi_t) - ds.
\]

We deduce using Theorem 2.28 in Mykland and Zhang (2012) that stably in law as \( \rho \to 0 \),

\[
\sigma_t^{(1)} (\widehat{V}_{t,n} - \mathcal{V}_t) \to \int_0^t k_s^{(1)} dX^{(1)} + \int_0^t k_{s-} dX^{(1)-} + \int_0^t (AV_t)^{1/2} d\widehat{V}_t.
\]

We have just shown Theorem 1. Now, we express the asymptotic bias \( AB_t = \int_0^t k_s^{(1)} dX^{(1)} + \int_0^t k_{s-} dX^{(1)-} \) differently as

\[
AB_t = \int_0^t k_s^{(1)} dX^{(1)} + \int_0^t k_{s-} \left( 1 - (\rho_{s}^{(2)})^2 \right) (\sigma_s^{(2)})^2 db^{(1)}
\]

\[
- \int_0^t k_s^{(1)} \rho^{(2)} \sigma_s^{(2)} db_s^{(1)} + \int_0^t k_{s-} \rho^{(2)} \sigma_s^{(2)} db_s^{(1)}
\]

\[
\int_0^t k_{s-} \left( 1 - (\rho_{s}^{(2)})^2 \right) (\sigma_s^{(2)})^2 dw^{(1)} + \int_0^t \left( (\sigma_s^{(1)})^{-1} \phi_s \mathcal{AC}_1 \right) dX^{(1)} + \int_0^t k_{s-} dX^{(1)-}.
\]

We thus deduce the expression of \( AB_t^{(1)} \) and \( AB_t^{(2)} \).

The proof of Corollary 4 follows in the same way as the proof of Theorem 1. We hold constant the asymptotic variance and the asymptotic bias on blocks of size \( h_n \). Moreover, we can see that \( \mathcal{A}B_t^{(1)}, \mathcal{A}B_t^{(2)} \) and \( \mathcal{A}V_t \) are uniformly consistent estimators under the constant model.

A.5. Discussion on the adaptation of Theorem 1 proofs for more general models

We discuss in this section how to adapt the proofs of Theorem 1 when considering Example 3 up to Example 6. In that case, the HBT can be defined for each \( k = 1, 2 \) as \( \tau_{i,n}^{(k)} := 0 \) and recursively as

\[
\tau_{i,n}^{(k)} := \inf \left\{ t > \tau_{i-1,n}^{(k)} : \Delta X_{t}(k) \right\}
\]

\[

\geq \left[ \alpha_t d_{t,n}^{(k)}(t - \tau_{i-1,n}^{(k)}), \alpha_t u_{t,n}^{(k)}(t - \tau_{i-1,n}^{(k)}) \right]
\]

(81)

for any positive integer \( i \). In (81), the grid \( d_{t,n}^{(k)} \) and \( u_{t,n}^{(k)} \) depend on \( n \), thus the term \( g_t^{(k)} \) in the asymptotic variance obtained in Theorem 1 will have a different interpretation. Indeed, \( g_t^{(k)} \) will be seen as a (possibly multidimensional) continuous time-varying parameter which generates (81) instead of the scaled grid function itself. In particular, the approximations will not be carried with holding \( g_t^{(k)} \) constant on each block, but rather with holding \( g_t^{(k)} \) constant on each block. Also, for any fixed \( t \in [0, 1] \), \( g_t^{(k)} \) will not be a function on \( \mathbb{R}^+ \), but a simple vector. The reader can refer to Potiron (2016) for the notion of time-varying parameter. Note that Assumption (A3) is only used in Lemmas 11 and 14. Thus, Lemmas 11 and 14 are the only parts in the proof which need to be adapted.

A.5.1. Example 3 (hitting constant boundaries of the jump size)

For each asset \( k = 1, 2 \) we define the jump sizes as \( l^{(k)}_n \). We assume that \( l^{(1)}_n \) and \( l^{(2)}_n \) are independent of each other. We have that \( g_t^{(k)}(s) := (-l^{(1)}_{t-1,n}, l^{(1)}_{t-1,n}) \) for \( t \in (\tau_{i-1,n}, \tau_i^{(k)}) \). We also have a non-time-varying parameter \( \gamma_t := 1 \).

As the jump size \( l^{(k)}_n \) is IID and independent of the other quantities, we can consider the same \( l^{(k)}_n \) when making local approximations. Note that in Lemma 11, the proof is made recursively for each observation time of the block. Thus a “jump” of \( g_t^{(k)} \) is not a problem when it happens exactly at observation times, as long as the jump is also made in the approximation block. Since \( l^{(k)}_n \) is assumed to be bounded, it is straightforward to adapt the proof of Lemma 11.

We discuss now how to adapt the proof of Lemma 14. To do that, we consider the Markov chain \( \widehat{Z}_t := (\Delta X_t^{(1)}, \Delta X_t^{(2)}, \Delta X_t^{(1)} - \tau_t^{(1)}, \tau_t^{(1)}, l^{(1)}_t, l^{(2)}_t) \), where \( i \) is such that \( \tau_t^{(i)} = \tau_t^{(1)} \). \( \tau_t^{(1)} \) is not a problem when it happens exactly at observation times, as long as the jump is also made in the approximation block. Since \( l^{(k)}_n \) is assumed to be bounded, it is straightforward to adapt the proof of Lemma 11.

A.5.2. Example 4 (model with uncertainty zones)

This model is very similar to Example 3, except that the sequence \( \tau_{i,n}^{(k)} \) is obtained as a function of \( \chi^{(k)} \), where \( \chi^{(k)} \) corresponds to the continuous time-varying parameter \( \chi_t^{(k)} \) of the \( k \)-th asset introduced in p. 5 of Robert and Rosenbaum (2012). We thus consider \( \gamma_t^{(k)} := \chi^{(k)}_t \). The proof of Lemma 11 can be extended using the convenient construction of \( l^{(k)}_n \) provided in p. 11 of Robert and Rosenbaum (2012). We extend this construction in two-dimension assuming that \( (W_t^{(1)}) \) and \( (W_t^{(2)}) \) are independent. As
Example 4 is slightly more involved than Example 3, the Markov chain $Z_t$ needs to include also the type of previous price change (increment or decrement) for each asset. We thus consider $Z_t := \left( \Delta X_t^{(2)} \big|_{\tau_i^1 \leq \cdot \leq \tau_i^2} - \tau_i^1, \tau_i^2 \right) \cdot \text{sign} (\Delta X_t^{(1)}), \text{sign} (\Delta X_t^{(2)})$, and can follow the same line of reasoning as in Lemma 14.

A.5.3. Example 5 (times generated by hitting an irregular grid model)

In this case, the parameter $g_{(K)}^0 := 1$ is non time-varying. Lemma 11 can be adapted easily. To show Lemma 14, a further condition is needed on $q_{(K)}^0 := p_{(K)}^0 - p_{(K)}^{-1}$, we assume that there exists a positive number $Q_{(K)}$ such that for any non-negative number $j$ and any $l \in \{0, \ldots, Q_{(K)} - 1\}$ we have $q_{(K)}^0 (Q_{(K)} + 1 + l) = q_{(K)}^0$. We also define the Markov chain $Z_t := \left( \Delta X_t^{(2)} \big|_{\tau_i^1 \leq \cdot \leq \tau_i^2} - \tau_i^1, \tau_i^2 \right) \cdot \text{sign} (\Delta X_t^{(1)}), \text{sign} (\Delta X_t^{(2)})$, where $l_{(K)}$ is the index such that there exists a non-negative number $m$ with $p_{(K)}^0 (m + l_{(K)}) = \hat{X}_t^{(1)}$, and $l_{(K)}$ is the index such that there exists a non-negative number $m$ with $p_{(K)}^0 (m + l_{(K)} + 1) = \hat{X}_t^{(2)}$. Under this assumption, we can show Lemma 14.

A.5.4. Example 6 (structural autoregressive conditional duration model)

We assume that the mixing variables $d_{(K)}^0$ and $c_{(K)}^0$ are interpolated by time-varying continuous stochastic parameters $(d_{(K)}, c_{(K)}$). We have that $g_{(K)}^0 := (d_{(K)}, c_{(K)})$. The central limit theorem in Example 6 can be obtained as a straightforward corollary of Theorem 1. If we define for any $s \geq 0$ the grid functions $g_{(K)}^0 (s) := (d_{(K)}, c_{(K)})$, the only difference between the HBT model (5) and the structural ACD model (6) is that we hold the grid between two observations in the latter model. In view of this specific assumption which implicates that the quantities of approximation are closer to the approximated quantities than under the HBT model, the proof of Lemma 11 simplifies. The proof of Lemma 14 remains unchanged as it deals only with quantities of approximation.

A.6. Jump case: proof of Remark 6

We update in this section the proof in the jump case model (14). The idea is to exclude all the blocks where we observe a jump. Such blocks will be finitely counted, and we will have at most one jump (either for $Y_t$ or for $Y_{(K)}$ but not for both prices at the same time) in each block. This is the main difference with the one-dimensional case.

We introduce the notation $A_{(K)}^{(0)} := \left\{ l \geq 1 \text{ s.t. } t_{l-1,n}^h \leq t \text{ and there is no jumps on } [t_{l-1,n}, t_{l,n}] \right\}$. The proof of Lemma 5 can be adapted because of the finiteness of jumps. The proof of Lemma 6 remains unchanged. Lemma 7 and Lemma 8 remain true in view of the finiteness of jumps. Lemmas 9 and 10 do not need any change. We modify Lemma 11 as follows. Let $l \geq 1$, we have that

$$\sup_{i \in A_{(K)}^{(0)}, 2 \leq j \leq h} \mathbb{E} \left[ \left| \Delta X_t^{(1), j} - \Delta \hat{X}_t^{(1), j} \right|^m \right] = o_p (\sigma_{(K)}^2)$$

and

$$\sup_{i \in A_{(K)}^{(0)}, 2 \leq j \leq h} \mathbb{E} \left[ \left| \Delta X_t^{(2), j} - \Delta \hat{X}_t^{(2), j} \right|^m \right] = o_p (\sigma_{(K)}^2) .$$

The proof remains unchanged in view of the independence assumption between jumps and the other quantities. Lemma 12 stays true with no further change. We introduce the new following lemma to be inserted between Lemmas 12 and 13 in the proofs.

**Lemma 16.** We have

$$\alpha_{n^{-2}} \left( \sum_{i \in A_{(K)}^{(0)}} \sum_{u=2}^{h_n} (N_{(i-1)h_n+u})^2 + 2N_{(i-1)h_n+u}N_{(i-1)h_n+u+1} \right) = o_n (1).$$

**Proof.** This is a simple consequence to the fact that we have at most one jump in $\Delta X_t^{(1)}$ or $\Delta X_t^{(2)}$ asymptotically, together with the finiteness of jumps. \(\square\)

Starting from Lemma 13 up to the end of the proof of Theorem 1, in view of Lemma 16, we can use “$i \in A_{(K)}^{(0)}$” in lieu of “$i \in A_n$”. We have thus proved that Theorem 1 is robust to jumps.

**References**

Alt-Sahalia, Y., Fan, J., Xiu, D., 2010. High-frequency covariance estimates with noisy and asynchronous financial data. J. Amer. Statist. Assoc. 105, 1504–1517.

Alouise, D.J., Eagleson, G.K., 1978. On mixing and stability of limit theorems. Ann. Probab. 6, 325–331.

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., 2011. Multivariate realised kernels: consistent positive semi-definite estimators of the covariance of equity prices with noise and non-synchronous trading. J. Econometrics 162, 149–176.

Barndorff-Nielsen, O.E., Shephard, N., 2001. Non-gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. J. Roy. Statist. Soc. Ser. B 63, 167–241.

Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. Roy. Statist. Soc. Ser. B 64, 253–280.

Borodin, A.N., Salminen, P., 2002. Handbook of Brownian Motion—Facts and formulae. In: Probability and its applications. Birkhäuser, Basel.

Breiman, L., 1992. Probability. In: Classics in applied mathematics, vol. 7. Society for Industrial and Applied Mathematics, SIAM, Philadelphia, PA.

Christensen, K., Kinnebrock, S., Podolskij, M., 2010. Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. J. Econometrics 159, 116–133.

Christensen, K., Podolskij, M., Vetter, M., 2013. On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes. J. Multivariate Anal. 120, 59–84.

Epst, T.W., 1979. Conferences in stock prices in the very short run. J. Amer. Statist. Assoc. 74, 291–298.

Fukasawa, M., 2010a. Central limit theorem for the realized volatility based on tick time sampling. Finance Stoch. 14, 209–233.

Fukasawa, M., 2010b. Realized volatility with stochastic sampling. Stochastic Process. Appl. 120, 829–552.

Fukasawa, M., Rosenbaum, M., 2012. Central limit theorems for realized volatility under hitting times of an irregular grid. Stochastic Process. Appl. 122 (12), 3901–3920.

Genon-Catalot, V., Jacod, J., 1993. On the estimation of the diffusion coefficient for multidimensional diffusion. Ann. Inst. Henri Poincaré Probab. Stat. 29, 119–151.

Hall, P., Heyde, C.C., 1980. Martingale Limit Theory and its Application. Academic Press.

Hayashi, T., Yoshida, N., 2005. On covariation estimation of non-synchronously observed diffusion processes. Bernoulli 11, 359–379.

Hayashi, T., Yoshida, N., 2008. Asymptotic normality of a covariation estimator for non-synchronously observed diffusion processes. Ann. Inst. Statist. Math. 60, 357–396.

Hayashi, T., Yoshida, N., 2011. Non-synchronous covariation process and limit theorems. Stochastic Process. Appl. 121, 2416–2454.

Jacod, J., 1994. Limit of Random Measures Associated with the Increments of a Brownian Semi-martingale. Technical Report, Université de Paris VI.

Jacod, J., Protter, P., 1998. Asymptotic error distributions for the Euler method for stochastic differential equations. Ann. Probab. 26, 267–307.

Jacod, J., Protter, P., 2012. Discretization of Processes. Springer.

Jacod, J., Shiryaev, A., 2003. Limit Theorems for Stochastic Processes, second ed. Springer-Verlag, Berlin.
Koike, Y., 2014. Limit theorems for the pre-averaged Hayashi-Yoshida estimator with random sampling. Stochastic Process. Appl. 124 (8), 2699–2753.

Koike, Y., 2015. Time endogeneity and an optimal weight function in pre-averaging covariance estimation. Preprint, Available at arXiv: http://arxiv.org/abs/1403.7889v2.

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

Meyn, S.P., Tweedie, R.L., 2009. Markov Chains and Stochastic Stability. Cambridge University Press.

Mykland, P.A., 2012. A Gaussian calculus for inference from high frequency data. Ann. Finance 8 (2–3), 235–258.

Mykland, P.A., Zhang, L., 2006. ANOVA for diffusions and Itô processes. Ann. Statist. 34, 1931–1961.

Mykland, P.A., Zhang, L., 2009. Inference for continuous semimartingales observed at high frequency. Econometrica 77, 1403–1445.

Potiron, Y., 2016. Estimating the integrated parameter of the locally parametric model in high-frequency data. arXiv preprint arXiv:1603.05700.

Pötzelberger, K., Wang, L., 2001. Boundary crossing probability for Brownian motion. J. Appl. Probab. 38, 152–164.

Renault, E., van der Heijden, T., Werker, B.J., 2009. A structural autoregressive conditional duration model. Presented at the 2010 Winter Meetings of the Econometric Society in Atlanta.

Renault, E., van der Heijden, T., Werker, B.J., 2014. The dynamic mixed hitting-time model for multiple transaction prices and times. J. Econometrics 180, 233–250.

Rényi, A., 1963. On stable sequences of events. Sankhy Series A 25, 293–302.

Revuz, D., Yor, M., 1999. Continuous Martingales and Brownian Motion, third ed. Springer, Germany.

Rootzén, H., 1980. Limit distributions for the error in approximations of stochastic integrals. Ann. Probab. 8, 241–251.

Robert, C.Y., Rosenbaum, M., 2011. A new approach for the dynamics of ultra-high-frequency data: the model with uncertainty zones. J. Financ. Econ. 9, 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.

Zhang, L., 2001. From Martingales to ANOVA : Implied and Realized Volatility (Ph.D. thesis). The University of Chicago, Department of Statistics.

Zhang, L., 2011. Estimation of the integrated parameter of a diffusion process, with applications to high-frequency data. J. Econometrics 160, 33–47.