Asymptotics of Asynchronicity

Markus Bibinger[1]

Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

Abstract

In this article we focus on estimating the quadratic covariation of continuous semimartingales from discrete observations that take place at asynchronous observation times. The Hayashi-Yoshida estimator serves as synchronized realized covolatility for that we give our own distinct illustration based on an iterative synchronization algorithm. We consider high-frequency asymptotics and prove a feasible stable central limit theorem. The characteristics of non-synchronous observation schemes affecting the asymptotic variance are captured by a notion of asymptotic covariations of times. These are precisely illuminated and explicitly deduced for the important case of independent time-homogeneous Poisson sampling.

Keywords: non-synchronous observations, quadratic covariation, Hayashi-Yoshida estimator, stable limit theorem, asymptotic distribution

MSC Classification: 62M10, 62G05, 62G20, 91B84

JEL Classification: C14, C32, C58, G10

1. Introduction

Nonparametric estimation methods for the quadratic variation of semimartingales have become an issue of great interest in recent years. One reason is the interpretation of the quadratic variation of the continuous part as integrated volatility in financial modeling.

If a semimartingale is observed discretely at times $t_i, 0 \leq i \leq n$ on a finite time horizon $[0, T]$, the sum of squared returns (increments of the semimartingale), called realized volatility, converges to the quadratic variation as $\sup (t_i - t_{i-1}) \to 0$ as $n \to \infty$. The same fact pertains to the multi-dimensional case where the realized covolatilities of two processes converge to the quadratic covariations. More usually multivariate data, in particular financial time series, are recorded at times following non-synchronous observation schemes. Therefore, realized covolatility estimates most commonly incorporate a previous-tick interpolation approach. Though, this machinery leads to the so-called Epps effect [8] that realized covolatilities tend to zero as the sampling frequency increases. Especially for the more and more available ultra high-frequency financial tick-data this issue poses problems.

A solution for the asynchronous estimation problem has been proposed in [13]. We call this estimator which arises as realized covolatility from all products of returns with overlapping observation time instants Hayashi-Yoshida estimator. Our investigation of that estimation approach leads to several useful rewritings and interpretations. The final representation is based on an iterative synchronization procedure which has been used first in [19]. This synchronized realized covolatility and the data aggregation technique for synchronization can serve as a basis for combined approaches in various generalizations of the underlying statistical model. A very important enhancement of the model in that we take market microstructure noise into account is covered in [4] and [5] by extending the synchronized realized covolatility to a generalized multiscale estimator.

The asymptotic theory developed in the article on hand is grounded on stable limit theorems for semimartingales from [16]. We obtain a stable limit theorem for the process associated with the estimation

error of a Hayashi-Yoshida estimator for the quadratic covariation at time $t \in [0, T]$ and for the overall estimator by the marginal distribution at $t = T$ the stable weak convergence to a centred mixed normal limiting distribution. The random asymptotic variance splits up in two terms induced by an idealized synchronous approximation and an additional error due to the lack of synchronicity.

The article is arranged in six sections. In the following Section 2 we give insight into the concept of stable weak convergence and a short review on the essential theory from the literature. In Section 3 the Hayashi-Yoshida estimator and our related synchronization algorithm that we have first presented in [4] is revisited and the asymptotic theory including the key result is provided in Section 4. The detailed proof of the central Theorem 2 is postponed to the Appendix A. Following some simple illustrative and motivating examples before, Section 5 comes up with the analysis for the important time-homogeneous independent Poisson sampling case for which we evaluate all ingredients of the asymptotic variance explicitly. To benefit from the stable central limit theorem and provide a basis for statistical inference, we give a consistent estimator for the asymptotic variance in Section 6.

2. Stable convergence and Jacod's stable limit theorem revisited

This section is devoted to the notion of stable weak convergence which will be an essential concept for the development of our limit theory throughout this article. The concept of stable convergence goes back to [21] and results about stable limit theorems were extended in [1] and [9]. The reason what makes stable weak convergence a key element of our asymptotic considerations, is that it allows to conclude joint weak convergence when we derive results about asymptotic mixed normality. In the case that a sequence of random variables (X_n) weakly converges to a mixed Gaussian limiting random variable VZ, with Z being standard normally distributed, $Z \sim \mathbf{N}(\mathbf{0}, \mathbf{1})$, and a strictly positive random variable V, independent of Z, we cannot derive confidence intervals if the distribution of V is unknown. However, if a consistent estimator V_n^2 for the asymptotic variance V^2 is available (in the sense that $V_n^2 \xrightarrow{p} V^2$), the stable weak convergence will assure that $(X_n, V_n^2) \rightsquigarrow (VZ, V^2)$ jointly and also that $X_n/V_n \rightsquigarrow Z$. The last implication also holds in a stable version: $X_n/V_n \xrightarrow{st} Z$. For that reason, if we are in situations as described above we gain from proving stable weak convergence which paves the way towards statistical inference.

Next, we present the formal definition and the main properties of stable weak convergence of sequences of random variables.

Let (X_n) be a sequence of random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in a Polish space $(E, \mathcal{E})$. We say that the sequence (X_n) converges weakly in $\mathbb{L}^1$ to X if for any bounded random variable Z

$$\lim_{n \to \infty} \mathbb{E}\left[ZX_n \right] = E\left[ZX \right]$$

holds.

Definition 1. For a sub-σ-field $\mathcal{G} \subseteq \mathcal{F}$ the sequence of random variables (X_n) is said to converge $\mathcal{G}$-stably, if there is a random probability measure μ on $(\Omega \times E, \mathcal{G} \otimes \mathcal{E})$ such that

$$\lim_{n \to \infty} \mathbb{E}\left[Zf(X_n) \right] = \int_{\Omega \times E} \mu(d\omega, dx) Z(\omega) f(x)$$

for all $f \in \mathcal{C}_b(E)$ (continuous and bounded) and $\mathcal{G}$-measurable bounded random variables Z.
If $\mathcal{G} = \mathcal{F}$, we say (X_n) converges stably in law to X ($X_n \xrightarrow{st} X$).

Remark 1. $\mathcal{G}$-stable convergence is the weak convergence in $\mathbb{L}^1$ of $\mathbb{E}\left[f(X_n)|\mathcal{G} \right]$ for all $f \in \mathcal{C}(E)$ to $\mu \circ f$. This implies convergence in distribution to the probability measure ν defined by

$$\nu(B) = \int \mu(d\omega, B) \mathbb{1}_{\{X(\omega) \in B\}} \mathbb{P}(d\omega) .$$

If (X_n) converges stably, the limiting law is $\mu(\Omega, \cdot)$.

The following proposition states some useful equivalent characterizations of $(\mathcal{G}-)$stable convergence.

Proposition 2.1. (X_n) *converges* $\mathcal{G}$-*stably is equivalent to:*

(i) *For every* $\mathcal{G}$-*measurable random variable* Z *on* Ω, (Z, X_n) *converges in law.*
(ii) *For every* $\mathcal{G}$-*measurable random variable* Z *on* Ω, (Z, X_n) *converges* $\mathcal{G}$-*stably.*
(iii) *The sequence* (X_n) *is tight, and for all* $G \in \mathcal{G}$ *and* $f \in \mathcal{C}(E)$, *the sequence* $\mathbb{E}\left[\mathbb{1}_G f(X_n)\right]$ *converges.*

This proposition is proved in [17] as part of Proposition IX.1.4. Stable (weak) convergence is a stronger mode of ordinary convergence in distribution. It is weaker than convergence in probability, but we emphasize that the limit depends on the limiting random variable X itself and not only on the distribution of X. If (X_n) converges stably to X, X is defined on an extension $(\Omega', \mathcal{F}', \mathbb{P}' = \mu)$ of the original probability space, so that

$$\forall f \in \mathcal{C}(E): \quad \lim_{n\to\infty} \mathbb{E}\left[f(X_n)Z\right] = \mathbb{E}'\left[f(X)Z\right] .$$

In the situation that we face in this article, a Gaussian random variable which is independent of $\mathcal{F}$ will appear as limiting law. In this case we call the extension of the original probability space orthogonal. In particular, if $X_n \overset{st}{\rightsquigarrow} X$ holds with an $\mathcal{F}$-measurable random variable X ($\mathbb{L}^1-$convergence), the foregoing proposition yields that $(X_n, X) \rightsquigarrow (X, X)$ and, hence $(X_n - X) \rightsquigarrow 0$ holds, which implies convergence in probability. Therefore, in all cases where stable weak convergence is a suitable adequate concept, limiting laws are defined on a genuine extension of the original probability space.

The following proposition gives the result that stable convergence is the suitable concept to derive feasible central limit theorems and confidence intervals if the asymptotic variances in limit theorems are unknown random, but can be estimated consistently.

Proposition 2.2. *Let* (X_n, V_n) *be real-valued random variables defined on* $(\Omega, \mathcal{F}, \mathbb{P})$. *If* $X_n \overset{st}{\rightsquigarrow} X$ *with a mixed normal limiting random variable* $X \sim \mathbf{N}(0, V^2)$ *and* $V_n \overset{p}{\longrightarrow} V$ *with* V *being* $\mathcal{F}$-*measurable. Then*

$$X_n/V_n \overset{st}{\rightsquigarrow} \mathbf{N}(0,1)$$

holds true.

Note that we use the same denotation expression for mixed normal laws and common normal laws and the difference becomes clear out of the context and by the specific variances. On the assumptions of the proposition $(X_n, V_n) \overset{st}{\rightsquigarrow} (X, V)$ is implied and the convergence of X_n/V_n follows by the continuous mapping theorem. This proposition is part of Proposition 2.5 in [20]. We restricted ourselves to real-valued random variables in the last proposition. A more general version can be found in [17].

The concept of stable convergence also carries over to stochastic processes. For this extension of stable convergence to stochastic processes, or more precisely to semimartingales, the Polish space E in Definition 1 is chosen to be the Skorohod space. The following limit theorem for stable convergence of continuous local martingales will be the foundation for our later deduced limit theorem in this article:

Theorem 1 (**Jacod's theorem: A martingale version**). *If* $(M_t, \mathcal{F}_t)$ *with* $0 \le t < \infty$ *is a continuous local martingale defined on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, *we denote by* $\mathcal{M}^\perp$ *the set of bounded* $(\mathcal{F}_t)$-*adapted martingales orthogonal to* $M = (M_t, \mathcal{F}_t)$ *what means that* $\left[M, M^\perp\right] \equiv 0$. *If* (X^n) *is a sequence of continuous* $(\mathcal{F}_t)$-*adapted local martingales for which*

$$\left[X^n\right]_t \overset{p}{\longrightarrow} V_t \ \forall t \in [0, \infty) \tag{1}$$

with a continuous process V *holds, the following two conditions*

$$\left[X^n, M\right]_t \overset{p}{\longrightarrow} 0 \ \forall t \in [0, \infty) \tag{2a}$$

$$\left[X^n, N\right]_t \overset{p}{\longrightarrow} 0 \ \forall t \in [0, \infty) \ and \ \forall N \in \mathcal{M}^\perp \tag{2b}$$

are sufficient that (X^n) *converges* $(\mathcal{F})$-*stably in law to* W_{V_t}, *where* W *is a standard Brownian motion independent of* $\mathcal{F}$.

This theorem is a simplified martingale version of the more general theorem 2–1 in [16]. A similar special version of the theorem is also used in [10]. A comprehensive illustrative overview on Jacod's stable limit theory and further motivation and applications of this result can be found in [20]. A discrete-time version of that theorem (cf. 3–1 in [16]) is the following:

Corollary 2.3. *Assume that $Z_t^n = \sum_{T_{n,i} \leq t} X_{n,i}$ is the endpoint of a discrete martingale and the $X_{n,i}$ are $\mathcal{F}_{T_{n,i}}$-measurable square integrable random variables and $(W_t, \mathcal{F}_t)$ a Brownian motion and $\Delta T_{n,i} = T_{n,i+1} - T_{n,i} \to 0$ as $n \to \infty$. If there exists a predictable process $(v_s)_{s \geq 0}$ such that*

$$\sum_{T_{n,i} \leq t} \mathbb{E}\left[X_{n,i}^2 | \mathcal{F}_{T_{n,i-1}}\right] \xrightarrow{p} \int_0^t v_s^2 \, ds , \tag{3a}$$

$$\forall \epsilon > 0 : \sum_{T_{n,i} \leq t} \mathbb{E}\left[X_{n,i}^2 \mathbb{1}_{\{X_{n,i} > \epsilon\}} | \mathcal{F}_{T_{n,i-1}}\right] \xrightarrow{p} 0 , \tag{3b}$$

$$\sum_{T_{n,i} \leq t} \mathbb{E}\left[X_{n,i}(W_{T_{n,i}} - W_{T_{n,i-1}}) | \mathcal{F}_{T_{n,i-1}}\right] \xrightarrow{p} 0 , \tag{3c}$$

$$\sum_{T_{n,i} \leq t} \mathbb{E}\left[X_{n,i}(M_{T_{n,i}} - M_{T_{n,i-1}}) | \mathcal{F}_{T_{n,i-1}}\right] \xrightarrow{p} 0 , \tag{3d}$$

for all bounded $\mathcal{F}_t$-martingales with $M_0 = 0$ and $[W, M] \equiv 0$. Then the following stable convergence of the process Z_t^n holds true:

$$Z_t^n \xrightarrow{st} Z_t = \int_0^t v_s \, dW_s^\perp \tag{4}$$

where $W^\perp$ is a Brownian motion defined on an orthogonal extension of the original probability space.

The limiting process in the foregoing Theorem 1 is a time-changed Brownian motion. The Brownian motion is of central importance in the theory of continuous local martingales, since every continuous local martingale M_t corresponds to a Dambis, Dubins-Schwarz time-changed Brownian motion $B_{[M]_t}$. For each (X_t^n) we have a representation as Dambis, Dubins-Schwarz Brownian motion $W_{[X^n]_t}^n$ and the sequence converges weakly to a limiting Brownian motion W_V by the asymptotic Knight-theorem. We refer to Theorem 7.7 in [6] for a proof. The conditions (2a) and (2b) about the quadratic covariations converging to zero in probability ensure that the weak convergence to W_V is stable.
For one fixed $0 < T < \infty$ we have the result that X_T^n converges stably in law to a centred mixed normal distribution:

$$X_T^n \xrightarrow{st} \mathbf{N}(0, V_T) . \tag{5}$$

The independence of the limiting Brownian motion W and (V, Y) for any $\mathcal{F}$-measurable random variable Y assures that (W_{V_T}, Y) has the same law as $(\sqrt{V_T} Z, Y)$ with $Z \sim \mathbf{N}(0, 1)$ and independent of (V_T, Y).
Note, that in the original theorem 2–1 in [16] for semimartingales the same conditions as in our Theorem 1 are imposed for the predictable quadratic (co-)variation processes that coincide with the quadratic (co-)variations for continuous semimartingales. Additionally, a condition that the drift can be neglected asymptotically is imposed. Compared to Theorem 3–1 in [16], we allow for non-equidistant discrete partitions which does not harm the deduction of Theorem 3–1 from Theorem 2–1 in [16]. A conditional Lindeberg-condition (3b) and a convergence condition on the conditional variances (3a) are analogous as in central limit theorems for triangular martingale arrays. The main difference to the stable limit theorem Corollary 3.1 in [12] (page 58 ff.) is that a certain nesting condition on the filtrations is replaced by conditions (3c) and (3d). Usually the reference Brownian motion W is given and "fully generates" the $X_{n,i}$s in the sense that (3d) holds.

The theorem also extends to a multi-dimensional setting which is formulated separately in the next corollary. For this purpose let M^* denote the transpose of a vector M and the $(d \times r)$-dimensional quadratic covariation $[M, N^*]_t := \left(\left[M^i, N^j \right]_t \right)_{ij}$ with $1 \leq i \leq d$ and $1 \leq j \leq r$ for a d-dimensional M and r-dimensional N. Recall that convergence in probability of a vector is equivalent to convergence in probability for every component.

Corollary 2.4. *Let $(M_t, \mathcal{F}_t)$ be a d-dimensional continuous local martingale and $\mathcal{M}^\perp$ again the set of $(\mathcal{F}_t)$-adapted bounded martingales orthogonal to M (to all components). A sequence of r-dimensional continuous $(\mathcal{F}_t)$-adapted local martingales (X^n) with*

$$[X^n, X^{n*}]_t \xrightarrow{p} V_t = \int_0^t w_s w_s^* \, ds \,, \tag{6}$$

where w_s is a predictable $\mathbb{R}^r \otimes \mathbb{R}^r$ process, and

$$[X^n, M^*]_t \xrightarrow{p} 0 \ \forall t \in [0, \infty) \,, \tag{7a}$$

$$[X^n, N]_t \xrightarrow{p} 0 \ \forall t \in [0, \infty) \text{ and } \forall N \in \mathcal{M}^\perp \,, \tag{7b}$$

converges stably in law to the process $\int_0^t w_s dW_s$, where W is a r-dimensional standard Brownian motion independent of $\mathcal{F}$.

Jacod's theorem provides a convenient stable central limit theorem for our purpose. Nesting conditions on the sequence of filtrations that are required for other stable limit theorems as in [12] and [22] are not satisfied here.

Furthermore, the concept of stable convergence enables us to prove the stable weak convergence to mixed Gaussian limiting random variables under an equivalent martingale measure $\tilde{\mathbb{P}}$ after a Girsanov transformation, where the drift processes are zero. Stable convergence guarantees that the asymptotic law carries over to the case with drift under the original measure $\mathbb{P}$. It is in this sense commutative with measure change (cf. [18]). If we have the result that $Z_n \xrightarrow{st} m + \mathbf{AVAR} \cdot \mathbf{N}(0, 1)$ under $\tilde{\mathbb{P}}$ with a standard Gaussian distribution independent of $\mathcal{F}$, defined on an orthogonal extension of the original probability space and $\mathcal{F}$-measurable bounded random variables m and $\mathbf{AVAR}$, the same convergence holds true under $\mathbb{P}$. Since stable convergence $Z_n \xrightarrow{st} Z$ implies for all $f \in \mathcal{C}(\mathcal{E})$ and $\mathcal{F}$-measurable bounded random variables X

$$\mathbb{E}\left[Xf(Z_n) \right] = \tilde{\mathbb{E}}\left[(d\mathbb{P}/d\tilde{\mathbb{P}}) Xf(Z_n) \right] \to \tilde{\mathbb{E}}'\left[(d\mathbb{P}/d\tilde{\mathbb{P}}) Xf(Z) \right] = \mathbb{E}'\left[Xf(Z) \right] \,,$$

by uniform integrability of $Xf(Z_n)(d\mathbb{P}/d\tilde{\mathbb{P}})$ with

$$d\mathbb{P}/d\tilde{\mathbb{P}} = \exp\left(-\int_0^t \gamma_s dB_s + \frac{1}{2} \int_0^t \gamma_s^2 \, ds \right) \,,$$

where $\sigma_s \gamma_s + \mu_s = 0$.

3. A synchronized realized covolatility estimator

Assumption 1. *On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, $X = (X_t)_{t \in \mathbb{R}^+}$ and $Y = (Y_t)_{t \in \mathbb{R}^+}$ are continuous semimartingales defined by the following stochastic differential equations:*

$$dX_t = \mu_t^X \, dt + \sigma_t^X \, dB_t^X \,,$$
$$dY_t = \mu_t^Y \, dt + \sigma_t^Y \, dB_t^Y \,,$$

with two $(\mathcal{F}_t)$–adapted standard Brownian motions B^X and B^Y and $\rho_t \, dt = d\left[B^X, B^Y \right]_t$. The drift processes μ_t^X and μ_t^Y are $(\mathcal{F}_t)$–adapted locally bounded stochastic processes and the spot volatilities σ_t^X and σ_t^Y and ρ_t are assumed to be $(\mathcal{F}_t)$–adapted with continuous paths. We assume strictly positive volatilities and the Novikov condition $\mathbb{E}\left[\exp\left((1/2) \int_0^T (\mu^\cdot / \sigma^\cdot)_t^2 \, dt \right) \right] < \infty$ for X and Y.

We consider the estimation of the quadratic covariation $[X, Y]_T$ of two continuous semimartingales, also called Itô processes, X and Y as defined in Assumption 1 from discrete observations following non-synchronous sampling schemes.

We impose the following regularity assumptions on the underlying asynchronous sampling schemes:

Assumption 2. *The deterministic observation times $\mathcal{T}^{X,n} = \{0 \le t_0^{(n)} < t_1^{(n)} < \ldots < t_n^{(n)} \le T\}$ of X and $\mathcal{T}^{Y,m} = \{0 \le \tau_0^{(m)} < \tau_1^{(m)} < \ldots < \tau_m^{(m)} \le T\}$ of Y are assumed to be regular in the following sense: There exists a constant $0 < \alpha \le 1/3$ such that*

$$\delta_n^X = \sup_{i \in \{1,\ldots,n\}} \left(\left(t_i^{(n)} - t_{i-1}^{(n)} \right), t_0^{(n)}, T - t_n^{(n)} \right) = \mathcal{O}\left(n^{-2/3-\alpha} \right), \tag{8a}$$

$$\delta_m^Y = \sup_{j \in \{1,\ldots,m\}} \left(\left(\tau_j^{(m)} - \tau_{j-1}^{(m)} \right), \tau_0^{(m)}, T - \tau_m^{(m)} \right) = \mathcal{O}\left(m^{-2/3-\alpha} \right). \tag{8b}$$

We consider asymptotics where the number of observations of X and Y are assumed to be of the same asymptotic order $n = \mathcal{O}(m)$ and $m = \mathcal{O}(n)$ and express that shortly by $n \sim m$.

For synchronous data $n = m$ and $t_i^{(n)} = \tau_i^{(n)}$ for all $i \in \{0, \ldots, n\}$ holds. In the non-synchronous case the number of observations $(n+1)$ of X and $(m+1)$ of Y may differ and the sets of observation times $\mathcal{T}^{X,n}$ also contain times $t_i^{(n)} \notin \mathcal{T}^{Y,m}$ and $\tau_j^{(n)} \notin \mathcal{T}^{X,n}$. We work within the general model where also synchronous observation times can take place and hence $\mathcal{T}^{Y,m}$ and $\mathcal{T}^{X,n}$ are not assumed to be disjoint. In the following, we omit the superscripts (n) and (m) for observation times to increase the readability.

Although the sequences of observation times are modeled deterministically, we remark that the case of random sampling times that are independent of the observed processes is included in that analysis regarding the conditional law given the observation times.

We use the short notation $\Delta X_{t_i}, i = 1, \ldots, n$ from now on for increments $X_{t_i} - X_{t_{i-1}}$ and analogously for Y. In [13] the consistency of the estimator

$$\widehat{[X,Y]}_T^{(HY)} = \sum_{i=1}^n \sum_{j=1}^m \Delta X_{t_i} \Delta Y_{\tau_j} \mathbb{1}_{[\min(t_i,\tau_j) > \max(t_{i-1},\tau_{j-1})]},$$

is proved, where the product terms include all increments of the processes with overlapping observation time intervals, for a similar model of discretely observed Itô diffusions with deterministic correlation, drift and volatility functions. Consistency directly carries over to our setting including random correlation, drift and volatility processes. The estimator is also in our setting, furthermore, unbiased if drift terms are zero and else asymptotically unbiased. In [14] it has further been shown that on stronger regularity assumptions on the observation schemes this Hayashi-Yoshida estimator is asymptotically distributed according to a Gaussian law.

For a general strategy leading to a synchronization mechanism that keeps to the Hayashi-Yoshida approach and its valuable properties, we focus on an alternative useful method to handle the asynchronicity of the data. It has been introduced in [19], where it was called pseudo-aggregation. The method translates the Hayashi-Yoshida estimator into an iterative algorithm that allows to rewrite the estimator without indicator functions. This can be done by aggregation of addends for which partial sums are telescoping. A first simple rewriting of the Hayashi-Yoshida estimator is obtained by taking the sum of the products of all increments of X with the telescoping sums of aggregated observed increments of Y for that observation time instants overlap with the according observation time instant of X (or in the symmetric way):

$$\widehat{[X,Y]}_T^{(HY)} = \sum_{i=1}^n \Delta X_{t_i} \left(\sum_{j \in \{1,\ldots,m\}} \Delta Y_{\tau_j} \mathbb{1}_{[\min(t_i,\tau_j) > \max(t_{i-1},\tau_{j-1})]} \right)$$

$$= \sum_{j=1}^m \Delta Y_{\tau_j} \left(\sum_{i \in \{1,\ldots,n\}} \Delta X_{t_i} \mathbb{1}_{[\min(t_i,\tau_j) > \max(t_{i-1},\tau_{j-1})]} \right).$$

first step:

- for $t_0 < \tau_0$ and $\mu_0 = \min\left(w \in \{1, \ldots, n\} | \tau_0 \leq t_w\right)$:

$$\mathcal{H}^0 = \{t_0, \ldots, t_{\mu_0}\} \text{ and } \mathcal{G}^0 = \{\tau_0\}$$

$$q_1 = \begin{cases} \mu_0 + 1 & \text{if } \tau_0 = t_{\mu_0} \\ \mu_0 & \text{if } \tau_0 < t_{\mu_0} \end{cases} \quad \text{and } r_1 = 1$$

- for $t_0 = \tau_0$:

$$\mathcal{H}^0 = \{t_0\} \text{ and } \mathcal{G}^0 = \{\tau_0\}$$

$$q_1 = 1 \text{ and } r_1 = 1$$

- for $t_0 > \tau_0$ and $w_0 = \min\left(l \in \{1, \ldots, m\} | t_0 \leq \tau_l\right)$:

$$\mathcal{H}^0 = \{t_0\} \text{ and } \mathcal{G}^0 = \{\tau_0, \ldots, \tau_{w_0}\}$$

$$q_1 = 1 \text{ and } r_1 = \begin{cases} w_0 + 1 & \text{if } t_0 = \tau_{w_0} \\ w_0 & \text{if } t_0 < \tau_{w_0} \end{cases}$$

ith step (given $\mathcal{H}^{i-1}$ and $\mathcal{G}^{i-1}$):

- for $t_{q_i} < \tau_{r_i}$ and $\mu_i = \min\left(w \in \{q_i + 1, \ldots, n\} | \tau_{r_i} \leq t_w\right)$:

$$\mathcal{H}^i = \{t_{q_i}, \ldots, t_{\mu_i}\} \text{ and } \mathcal{G}^i = \{\tau_{r_i}\}$$

$$q_i \dashrightarrow \begin{cases} q_{i+1} = \mu_i + 1 & \text{if } \tau_{r_i} = t_{\mu_i} \\ q_{i+1} = \mu_i & \text{if } \tau_{r_i} < t_{\mu_i} \end{cases} \quad \text{and } r_i \dashrightarrow r_{i+1} = r_i + 1$$

- for $t_{q_i} = \tau_{r_i}$:

$$\mathcal{H}^i = \{t_{q_i}\} \text{ and } \mathcal{G}^i = \{\tau_{r_i}\}$$

$$q_i \dashrightarrow q_{i+1} = q_i + 1 \text{ and } r_i \dashrightarrow r_{i+1} = r_i + 1$$

- for $t_{q_i} > \tau_{r_i}$ and $w_i = \min\left(l \in \{r_i + 1, \ldots, m\} | t_{q_i} \leq \tau_l\right)$:

$$\mathcal{H}^i = \{t_{q_i}\} \text{ and } \mathcal{G}^i = \{\tau_{r_i}, \ldots, \tau_{w_i}\}$$

$$q_i \dashrightarrow q_{i+1} = q_i + 1 \text{ and } r_i \dashrightarrow \begin{cases} r_{i+1} = w_i + 1 & \text{if } t_{q_i} = \tau_{w_i} \\ r_{i+1} = w_i & \text{if } t_{q_i} < \tau_{w_i} \end{cases}$$

Algorithm 1: Iterative algorithm for construction of the joint grid from asynchronous data.

Defining the next-tick interpolation $t_{i,+} := \min_{0 \leq j \leq m}\left(\tau_j | \tau_j \geq t_i\right)$ and the previous-tick interpolation $t_{i,-} := \max_{0 \leq j \leq m}\left(\tau_j | \tau_j \leq t_i\right)$, the last expression can be illustrated

$$\widehat{[X, Y]}_T^{(HY)} = \sum_{i=1}^n \Delta X_{t_i}\left(Y_{t_{i,+}} - Y_{t_{i-1,-}}\right) .$$

The algorithm which we will use is a more enhanced method to aggregate the data in an adequate way. For this purpose $(N + 1)$ sets $\mathcal{H}^i$ and $\mathcal{G}^i$ are constructed, where $N < \min(n, m)$, each set including one or more than one observation time of X and Y, respectively. This method to construct a joint grid for the observations of the two processes is described by Algorithm 1.

The Algorithm 1 that we have first presented in [4] stops after $(N + 1)$ steps when the last observation time is reached. We pass over from the original observations to the sums of observed increments $X^{\mathcal{H}^i}$ over sets $\mathcal{H}^i$ and $Y^{\mathcal{G}^i}$ over sets $\mathcal{G}^i$, respectively. The observations are grouped together so that the resulting realized

covolatility estimator

$$\sum_{i=0}^{N} X^{\mathcal{H}^i} Y^{\mathcal{G}^i} = \sum_{i=1}^{n} \sum_{j=1}^{m} \Delta X_{t_i} \Delta Y_{\tau_j} \mathbb{1}_{[\min{(t_i,\tau_j)} > \max{(t_{i-1},\tau_{j-1})}]}$$

calculated from the 'synchronized' observations

$$X^{\mathcal{H}^i} = \sum_{t_j \in \mathcal{H}^i} \Delta X_{t_j} , \ \ Y^{\mathcal{G}^i} = \sum_{\tau_j \in \mathcal{G}^i} \Delta Y_{\tau_j} , \ \ i \in \{0, \ldots, N\} .$$

for the integrated covolatility will coincide with the one by [13] stated above. We use a different illustration of this estimator compared to [19] making use of telescoping sums.

With the denotation expressions from Algorithm 1

$$\mu_i = \max{(k|t_k \in \mathcal{H}^i)}, \qquad\qquad w_i = \max{(k|\tau_k \in \mathcal{G}^i)} \ \text{ and}$$
$$q_i = \min{(k|t_k \in \mathcal{H}^i)}, \qquad\qquad r_i = \min{(k|\tau_k \in \mathcal{G}^i)} \ \ , i \in \{0, \ldots, N\}$$

and for the purpose of a simpler notation

$$X_{g_i} = X_{t_{\mu_i}}, \qquad\qquad Y_{\gamma_i} = Y_{\tau_{w_i}} \ \ , i \in \{0, \ldots, N\} \ \text{and}$$
$$X_{l_i} = X_{t_{q_i-1}}, \qquad\qquad Y_{\lambda_i} = Y_{\tau_{r_i-1}} \ \ , i \in \{1, \ldots, N\}$$

with $l_0 := t_0$, $\lambda_0 := \tau_0$, $X^{\mathcal{H}^i}$ and $Y^{\mathcal{G}^i}$ can be written as telescoping sums $X^{\mathcal{H}^i} = (X_{g_i} - X_{l_i})$, $Y^{\mathcal{G}^i} = (Y_{\gamma_i} - Y_{\lambda_i})$. This leads to

$$\widehat{[X,Y]}_T^{(HY)} = \sum_{i=1}^{N} (X_{g_i} - X_{l_i})(Y_{\gamma_i} - Y_{\lambda_i}) , \tag{9}$$

where summation starts with $i = 0$ or $i = 1$ since the addend for $i = 0$ is always zero. Although we use this specific new illustration throughout this article, we will call this realized covolatility of our synchronized observations also Hayashi-Yoshida estimator in the following. In this notation g_i denotes the greatest and l_i the last observation time before the least element of the set $\mathcal{H}^i$ and analogously γ_i and λ_i of $\mathcal{G}^i$.

Example
An illustration of the application of Algorithm 1 to observations is given in Figure 1. In this example, we have $\mathcal{H}^0 = \{t_0\}, \mathcal{G}^0 = \{\tau_0\}, \mathcal{H}^1 = \{t_1, t_2, t_3\}, \mathcal{G}^1 = \{\tau_1\}, \mathcal{H}^2 = \{t_3\}, \mathcal{G}^2 = \{\tau_2, \tau_3\}, \mathcal{H}^3 = \{t_4, t_5, t_6\}, \mathcal{G}^3 = \{\tau_4\}, \mathcal{H}^4 = \{t_6, t_7\}, \mathcal{G}^4 = \{\tau_5\}, \mathcal{H}^5 = \{t_7, t_8\}, \mathcal{G}^5 = \{\tau_6\}, \mathcal{H}^6 = \{t_8\}, \mathcal{G}^6 = \{\tau_7, \tau_8\}, \mathcal{H}^7 = \{t_9\}, \mathcal{G}^7 = \{\tau_8, \tau_9\}, \mathcal{H}^8 = \{t_{10}\}, \mathcal{G}^8 = \{\tau_9, \tau_{10}\}$.

The example highlights the important features of the synchronization procedure. The sets $\mathcal{H}^i$ and $\mathcal{G}^i$ are in general not disjoint and the maxima of consecutive sets can be the same time points. The minimum of a successive set can as well equal the maximum of the prevenient. Contrarily, consecutive minima are not equal. For further examples we refer to [19]. Of course the example is just for illustration and the number of observations is much smaller than in practice. The synchronization of $n + 1 = 11$ and $m + 1 = 11$ observations leads to $N + 1 = 9$ synchronized observations in this example.

The fact that we obtain $(N + 1) < \min{(n, m)} + 1$ synchronized observations indicates heuristically that the efficiency of such techniques of covariance estimation mainly depends on the number of observations available for the less liquid process which is observed at a lower frequency. By Assumption 2 we restrict us to the case that n and m are of the same order. Thus for the suprema of times between two observations

$$\delta_n^X = \mathcal{O}\left(N^{-2/3-\alpha}\right) \ \text{ and } \ \delta_n^Y = \mathcal{O}\left(N^{-2/3-\alpha}\right)$$

holds with a constant $0 < \alpha \le 1/3$.
In the next section, we show that on Assumption 1 and 2 the estimator (9) is $\sqrt{N}$-consistent and, on further
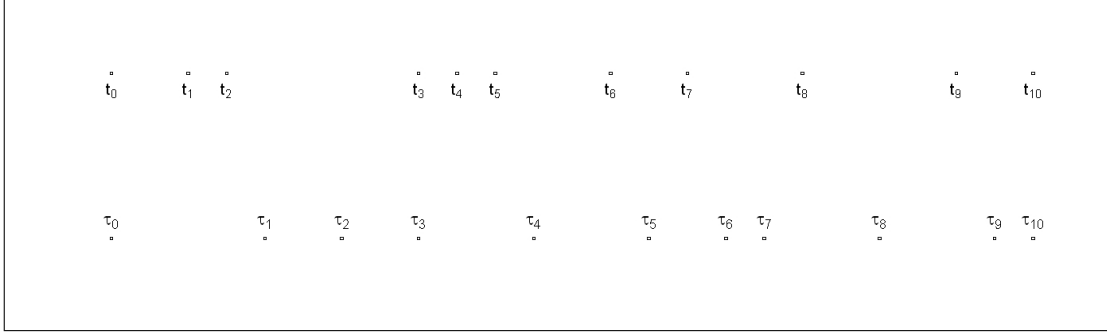
Figure 1: Example for synchronization using Algorithm 1.

assumptions on the asymptotic behavior of the asynchronous sampling schemes, asymptotically normally distributed. Using standard interpolation methods such an estimator cannot be obtained.

Another recent approach to deal with non-synchronous discrete observations in a general setting including market microstructure noise has been proposed by [2]. This method is also related to our approach. The so-called refresh times are the cumulative sums of waiting times until both processes are observed. Assume that in the ith step of Algorithm 1 $t_{q_i} < \tau_{r_i}$ holds. Then the next observation times of X are grouped together ending with the first observation time $t_{\mu_i-1} < \tau_{r_i} \le t_{\mu_i}$ greater or equal than τ_{r_i}. Then we start the next comparison step and compare this last observation time grouped to the set $\mathcal{H}^i$ to τ_{r_i+1}, except for the case where two synchronous observations appeared, where we compare the two following times. Since in the completely asynchronous case at the refresh times only one of the two processes is observed, the refresh time method used in [2] includes a previous-tick interpolation for the unobserved process at the refresh times. Refresh times provide the 'closest synchronous approximation' to the asynchronous sampling schemes that we define in Proposition 11 below. The number of refresh times which are denoted in this work by $T_i, i = 0, \ldots, N$, equals the number of sets constructed by pseudo-aggregation. In a setting that also takes microstructure noise into account, a consistent estimator requires smoothing techniques to reduce the noise perturbation and the optimal convergence rate is slower (cf. [4]). The previous-tick interpolation, however, causes a negative bias due to asynchronicity when calculating the simple realized covolatility estimator based on the refresh time and previous-tick approach and it does not equal the estimator of Hayashi-Yoshida. The reason for this bias is that, due to the previous-tick interpolation, products of increments with overlapping observation time instants fall out of the realized covolatility. The pseudo-aggregation Algorithm 1 used in this work corresponds to the refresh time method when replacing the previous-tick interpolation by a next-tick interpolation for the right end points of refresh time instants. Then, the resulting realized covolatility of 'synchronized observations'

$$\widehat{[X,Y]}_T^{(HY)} = \sum_{i=1}^{N} \left(X_{g_i} - X_{l_i} \right) \left(Y_{\gamma_i} - Y_{\lambda_i} \right)$$
$$= \sum_{i=1}^{N} \left(X_{T_{i,+}^X} - X_{T_{i-1,-}^X} \right) \left(Y_{T_{i,+}^Y} - Y_{T_{i-1,-}^Y} \right) \tag{10}$$

coincides with the Hayashi-Yoshida estimator and has no bias due to asynchronicity. As figured out in the simulation study of [4] the asymptotically vanishing influence of the bias due to pure previous-tick interpolation also shows up in the setting with noise for finite sample sizes and mild noise variances for that combined estimators are constructed in [2] and [4], among others.

Figure 2 visualizes refresh times T_i, $i = 0, \ldots, 8$ for our above given example. For this example the
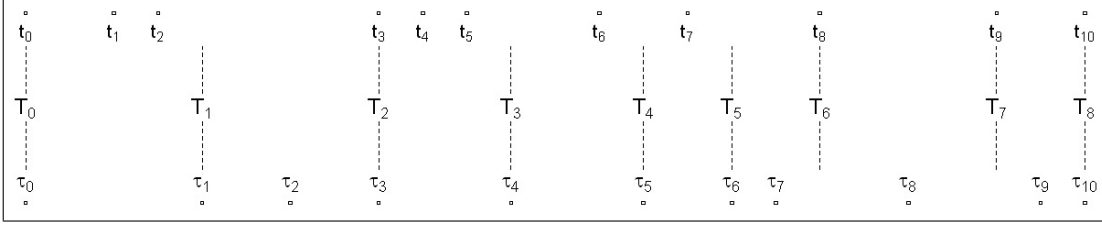
Figure 2: Example for synchronization using Algorithm 1 including refresh times.

realized covolatility calculated with refresh time previous-tick interpolated values equals

$$
(X_{t_2} - X_{t_0})(Y_{\tau_1} - Y_{\tau_0}) + (X_{t_3} - X_{t_2})(Y_{\tau_3} - Y_{\tau_1}) + (X_{t_5} - X_{t_3})(Y_{\tau_4} - Y_{\tau_3}) +
$$
$$
(X_{t_6} - X_{t_5})(Y_{\tau_5} - Y_{\tau_4}) + (X_{t_7} - X_{t_6})(Y_{\tau_6} - Y_{\tau_5}) + (X_{t_8} - X_{t_7})(Y_{\tau_7} - Y_{\tau_6}) +
$$
$$
(X_{t_9} - X_{t_8})(Y_{\tau_8} - Y_{\tau_7}) + (X_{t_{10}} - X_{t_9})(Y_{\tau_{10}} - Y_{\tau_8})
$$

and is biased downwards due to non-synchronicity, whereas (9) yields

$$
(X_{t_3} - X_{t_0})(Y_{\tau_1} - Y_{\tau_0}) + (X_{t_3} - X_{t_2})(Y_{\tau_3} - Y_{\tau_1}) + (X_{t_6} - X_{t_3})(Y_{\tau_4} - Y_{\tau_3}) +
$$
$$
(X_{t_7} - X_{t_5})(Y_{\tau_5} - Y_{\tau_4}) + (X_{t_8} - X_{t_6})(Y_{\tau_6} - Y_{\tau_5}) + (X_{t_8} - X_{t_7})(Y_{\tau_8} - Y_{\tau_6}) +
$$
$$
(X_{t_9} - X_{t_8})(Y_{\tau_9} - Y_{\tau_7}) + (X_{t_{10}} - X_{t_9})(Y_{\tau_{10}} - Y_{\tau_8}) \, ,
$$

which is an unbiased estimator for observations of processes according to Assumption 1, when drift terms are assumed to be zero.

4. Asymptotic distribution theory

In this section the elements for an analysis of the asymptotic properties of the estimator (9) are developed where the emphasis is on the asymptotic distribution of the estimator.

The following technical Proposition constitutes the theoretical justification that the refresh times $T_i^{(n)}$, $1 \leq i \leq N$ introduced in the foregoing section can serve as a convenient basis to decompose the overall estimation error of the synchronized realized covolatility (9). For every N these times induce a partition of the time horizon $[0, T]$ that we call the closest synchronous approximation.

Proposition 4.1. *If we define $T_i^{(N)} := \min(g_i, \gamma_i)$, $i = 0, \ldots, N$, the set $\mathcal{T}^{syn,N} = \{T_0^{(N)}, \ldots, T_N^{(N)}\}$ induces a partition of the time span $[0, T]$ in the sense that $\dot{\bigcup}_i [T_i^{(N)}, T_{i+1}^{(N)}) = [T_0^{(N)}, T - T_N^{(N)})$. The following equality holds true:*

$$
T_i^{(N)} = \min\left(g_i^{(N)}, \gamma_i^{(N)}\right) = \max\left(l_{i+1}^{(N)}, \lambda_{i+1}^{(N)}\right), \; i = 1, \ldots, N-1 \tag{11}
$$

and on Assumption 2 $\delta_N := \sup_{i \in \{1,\ldots,N\}} \left(T_i^{(N)} - T_{i-1}^{(N)}\right) = \mathcal{O}\left(N^{-2/3-\alpha}\right)$ holds.

In the following we frequently leave out superscripts indicating dependence on N to guarantee clarity and increase the readability.

Proof. Assume without loss of generality $g_i \leq \gamma_i$ for an arbitrarily fixed $i \in \{1, \ldots, N-1\}$. Taking Algorithm 1 into account, we proof that (11) holds true.

If $g_i < \gamma_i$, then the observation times γ_i and $g_{i,+} := \min\left(t_k \in \mathcal{T}^X | t_k > g_i\right)$ are compared in the $(i+1)$th step of the synchronization Algorithm 1 and $g_{i,+} = \min\left(t_k \in \mathcal{T}^X | t_k \in \mathcal{H}^{i+1}\right)$ holds true. Thus, $g_i = l_{i+1}$ and (11) holds true. We remark that in this case $\gamma_i \in \mathcal{G}^{i+1}$ and thus $\gamma_i > \lambda_{i+1} = \gamma_{i,-} := \max\left(\tau_k \in \mathcal{T}^Y | \tau_k < \gamma_i\right) \geq \gamma_{i-1}$.

If $g_i = \gamma_i$, then the observation times $g_{i,+}$ and $\gamma_{i,+}$ are compared in the $(i+1)$th step of Algorithm 1 and

10

$l_{i+1} = \lambda_{i+1} = g_i = \gamma_i$ what implies (11).

Equation (11) does not hold true for $i = 0, N$ and $T_0 = t_0 \wedge \tau_0$ because we have set $l_0 = t_0$ and $\lambda_0 = \tau_0$. Although consecutive maxima g_i of the sets $\mathcal{H}^i$ and γ_i of the sets $\mathcal{G}^i$, respectively, can be equal, $T_i > T_{i-1}$ holds for all $i \in \{1, \ldots, N\}$ because $g_{i+1} = g_i$ implies that $\gamma_{i+1} > \gamma_i$ and $\gamma_{i+1} = \gamma_i$ implies that $g_{i+1} > g_i$. Hence, the set $\mathcal{T}^{syn}$ induces a partition of the time span $[0, T]$. $\square$

The times $T_i, i = 0, \ldots, N$ defined through (11) equal the refresh times from [2] as has been mentioned in the last section. We use Proposition 4.1 to split the error of the estimator (9) for the integrated covolatility $[X, Y]_T$ in two asymptotically uncorrelated parts. The error of the estimator (9) can be written

$$\sum_{i=1}^{N} \left(X_{g_i} - X_{l_i} \right) \left(Y_{\gamma_i} - Y_{\lambda_i} \right) - \int_0^T \rho_t \sigma_t^X \sigma_t^Y \, dt = D_T^N + A_T^N$$

where

$$D_T^N := \sum_{i=1}^{N} \left(\left(X_{T_i} - X_{T_{i-1}} \right) \left(Y_{T_i} - Y_{T_{i-1}} \right) - \int_{T_{i-1}}^{T_i} \rho_t \sigma_t^X \sigma_t^Y \, dt \right) \qquad (12)$$
$$- \int_0^{t_0 \wedge \tau_0} \rho_t \sigma_t^X \sigma_t^Y \, dt - \int_{t_n \wedge \tau_m}^{T} \rho_t \sigma_t^X \sigma_t^Y \, dt$$

is a synchronous-type discretization error of the realized covolatility estimator evaluated with synchronous observations at the times $T_i, i = 0, \ldots, N$, which is the closest synchronous approximation to the asynchronous sampling scheme, and

$$A_T^N = \sum_{i=1}^{N} \left(Y_{\gamma_i} - Y_{\lambda_i} \right) \left(X_{g_i} - X_{T_i} \right) \mathbb{1}_{\{T_i = \gamma_i\}} + \left(Y_{T_i} - Y_{T_{i-1}} \right) \left(X_{T_{i-1}} - X_{l_i} \right) \mathbb{1}_{\{T_{i-1} = \lambda_i\}}$$
$$+ \sum_{i=1}^{N} \left(X_{T_i} - X_{l_i} \right) \left(Y_{\gamma_i} - Y_{T_i} \right) \mathbb{1}_{\{T_i = g_i\}} + \left(X_{T_i} - X_{T_{i-1}} \right) \left(Y_{T_{i-1}} - Y_{\lambda_i} \right) \mathbb{1}_{\{T_{i-1} = l_i\}} \qquad (13)$$

is the remaining additional error due to the lack of synchronicity. When we write the increments involved in the estimator (9) in the way

$$\left(X_{g_j} - X_{l_j} \right) = X_j^+ + X_j^S + X_j^- , \quad \left(Y_{\gamma_j} - Y_{\lambda_j} \right) = Y_j^+ + Y_j^S + Y_j^- ,$$

where $X_j^+ = X_{g_j} - X_{T_j}$ denotes the next-tick interpolation error at right-end points, $X_j^- = X_{T_{j-1}} - X_{l_j}$ the previous-tick interpolation error at left-end points, $X_j^S = X_{T_j} - X_{T_{j-1}}$, $j = 1, \ldots, N$ the increment over the time instant of the closest synchronous approximation and analogously for Y, D_T^N and A_T^N can be expressed:

$$D_T^N = \sum_{i=1}^{N} X_i^S Y_i^S ,$$

$$A_T^N = \sum_{i=1}^{N} \left(X_i^+ (Y_i^S + Y_i^-) + Y_i^+ (X_i^S + X_i^-) + X_i^- Y_i^S + Y_i^- X_i^S \right) .$$

D_T^N is an usual synchronous-type realized covolatility but incorporates an idealized sampling design at the times of the closest synchronous approximation for which we do not have observations in an asynchronous setting. Nevertheless, this idealized approximation turns out to be helpful for our further analysis. The error due to non-synchronicity A_T^N hinges on the interpolations that have to be carried out since we do not observe X and Y at the times $T_i^N, 1 \leq i \leq N$. The term is asymptotically centred since only products of increments over disjoint time instants remain whereas D_T^N is an unbiased estimator for $[X, Y]_T$. Since either X or Y is observed at a certain $T_i, 1 \leq i \leq N$, one of each interpolation errors in the illustration above equals zero.
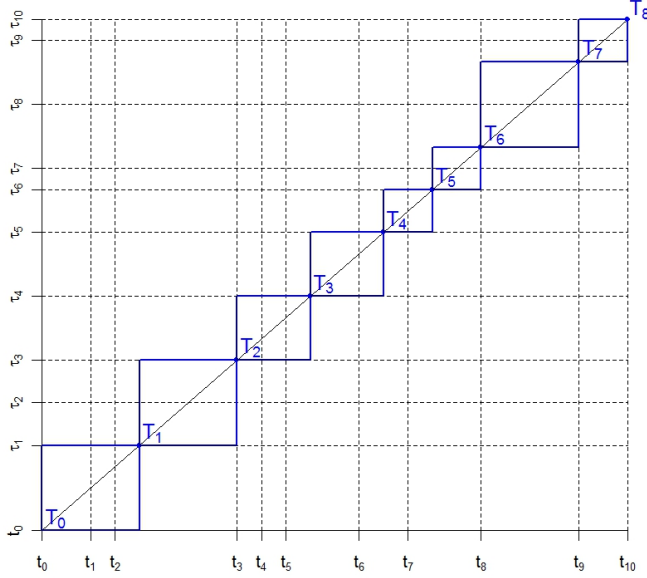
Figure 3: Illustration of the synchronous approximation for our example.

Proposition 4.2. *The Brownian parts of A_T^N and D_T^N are uncorrelated. This means, that if we assume the drift terms to be identically zero in Assumption 1, A_T^N and D_T^N are uncorrelated. If the drift terms are non-zero, A_T^N and D_T^N are asymptotically uncorrelated.*

Proof. A_T^N and D_T^N are both centred. If Assumption 1 holds with $\mu_t^X \equiv \mu_t^Y \equiv 0$, the expectation of the product of A_T^N and D_T^N is zero, since the previous- and next-tick interpolated increments in (13) are centred and uncorrelated to the other three factors in each addend of the inner sums.

If we allow for non-zero drift terms, Assumption 1 and Assumption 2 ensure that the increments over time intervals due to the drift induce terms at most of order δ_N in probability by products of drift terms and at most of order $\delta_N^{1/2}$ in probability by products of drift and Brownian increments in the overall correlation. □

In Figure 3 the observation times $\tau_j, j = 0, \ldots, 11$ of Y for our Example 1 from the last section are plotted against the observation times $t_i, i = 0, \ldots, 11$ of X. The dashed lines intersect for synchronous observation times $t_0 = \tau_0, t_3 = \tau_3$ and $t_{10} = \tau_{10}$ on the diagonal of the square in Figure 3. A similar visualization of the realized covolatility estimator for synchronous and equidistant data would yield coextensive squares around the diagonal, over which multiplied increments are summed up. Refresh times are (in general) not equidistant but provide a synchronous realized covolatility estimator as an approximation. The Hayashi-Yoshida estimator (9) is the sum of products of increments with overlapping observation time instants. The relation to the synchronous approximation D_T^N is that we have next-tick interpolations and previous-tick interpolations to the times $T_i, i = 0, \ldots, 8$ and take increments from previous-tick interpolated values to next-tick interpolated values. The time instants of D_T^N are visualized for our example in Figure 3. The previous- and next-tick interpolations are illustrated in Figure 4. The products of time instants leading to the error A_T^N are illustrated in the same picture by the grey rectangles. As can be seen for the example in Figure 4, A_T^N is the sum of the errors by the ith next-tick interpolation multiplied with the increments of the other process over $[\min(l_i, \lambda_i), T_i]$ and the sum of the errors of the ith previous-tick interpolation multiplied with the increments of the other process over $[T_{i-1}, T_i]$. The sum of the increments over the squares in Figure 3, D_T^8 for our example, and the grey rectangles in Figure 4, A_T^8 for our example, is the Hayashi-Yoshida estimator evaluated at the end of the last section.
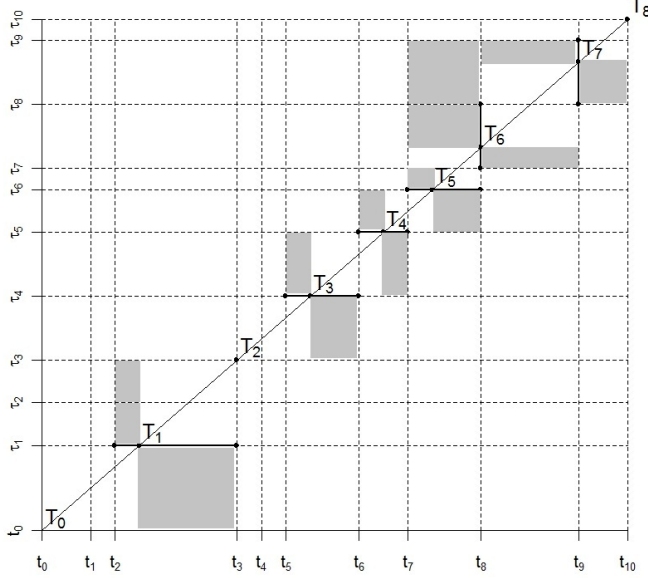
Figure 4: Illustration of the next- and previous-tick interpolated values and the error due to non-synchronicity for our example.

Definition 2 (quadratic (co-)variations of time). For any $N \in \mathbb{N}$ let $T_i^{(N)}$, $i = 0, \ldots, N$ be the times from the partition of $[0, T]$ defined in (11) above and $g_i^{(N)}, \gamma_i^{(N)}, l_i^{(N)}, \lambda_i^{(N)}$ the corresponding observation times designated by Algorithm 1 from the estimator (9). T/N is the mean of the time instants $\Delta T_i^{(N)} = T_i^{(N)} - T_{i-1}^{(N)}$, $i = 1, \ldots, N$. Define the following sequences of functions

$$G^N(t) = \frac{N}{T} \sum_{T_i^{(N)} \leq t} \left(\Delta T_i^{(N)} \right)^2 , \tag{14a}$$

$$F^N(t) = \frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} (T_i^{(N)} - \lambda_i^{(N)})(g_i^{(N)} - T_i^{(N)}) + \left(T_i^{(N)} - l_i^{(N)} \right) \left(\gamma_i^{(N)} - T_i^{(N)} \right)$$
$$+ \Delta T_{i+1}^{(N)} \left(T_i^{(N)} - l_{i+1}^{(N)} \right) + \Delta T_{i+1}^{(N)} \left(T_i^{(N)} - \lambda_{i+1}^{(N)} \right) , \tag{14b}$$

$$H^N(t) = \frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left(T_i^{(N)} - l_{i+1}^{(N)} \right) \left(g_i^{(N)} - T_i^{(N)} \right) + \left(T_i^{(N)} - \lambda_{i+1}^{(N)} \right) \left(\gamma_i^{(N)} - T_i^{(N)} \right) , \tag{14c}$$

for $t \in [0, T]$ that we call sequences of quadratic (co-)variations of times.

A stable central limit theorem for the estimation error is deduced on the assumption that the sequences defined by (14a), (14b) and (14c) converge pointwise and the sequences of difference quotients uniformly:

Assumption 3 (asymptotic quadratic (co-)variation of times). *Assume that for the sequences of sampling schemes and the times $T_i^{(N)}, g_i^{(N)}, \gamma_i^{(N)}, l_i^{(N)}, \lambda_i^{(N)}$ and the sequences of quadratic (co-) variations of times $G^N(t), F^N(t), H^N(t)$ defined in Definition 2 the following holds true:*

(i) $G^N(t) \to G(t)$, $F^N(t) \to F(t)$, $H^N(t) \to H(t)$ as $N \to \infty$, where $G(t), F(t), H(t)$ are continuously differentiable functions on $[0, T]$.

(ii) *For any null sequence* (h_N), $h_N = \mathcal{O}\left(N^{-1}\right)$

$$\frac{G^N(t+h_N) - G^N(t)}{h_N} \to G'(t) \tag{15a}$$

$$\frac{F^N(t+h_N) - F^N(t)}{h_N} \to F'(t) \tag{15b}$$

$$\frac{H^N(t+h_N) - H^N(t)}{h_N} \to H'(t) \tag{15c}$$

uniformly on [0,T] as $N \to \infty$.

Assumption 3 is necessary to ensure that the sequence of variances of the estimator (9) converges as $n, m \to \infty$. The derivative of the asymptotic quadratic variation of refresh times (15a) will appear in the asymptotic variance of the discretization error D_T^N, since refresh times are (in general) not equidistant. For $\Delta T_i^{(N)} = T/N$ for all $i \in \{0, \ldots, N\}$, $G'(t) = \mathbb{1}_{[0,T]}$ holds true.
The uniform convergence of the difference quotients defined by (15b) and (15c) are necessary to ensure that the sequence of variances of A_T^N converges as $N \to \infty$. The assumptions imposed by (15a)-(15c) are weaker than assuming convergence of the joint sampling design of $\left(\mathcal{T}^{X,n}, \mathcal{T}^{Y,m}\right)$ and are not very restrictive. They hold true whenever the sequences of sampling schemes tend to a certain state of asynchronicity or have a uniform behaviour of non-synchronicity in the limit as $n, m \to \infty$. For homogeneous sampling schemes these (co-)variations of time converge to linear limiting functions.
The sequence of functions F^N describe an interaction of interpolation steps between the two processes. In contrast, H^N is defined to measure an impact of the in general non-zero correlations of next-tick and previous-tick interpolations to the same refresh time T_i, for each process separately.

Example:
Consider the synchronous equidistant sampling schemes with $N = n = m$ and $t_i^{(n)} = \tau_j^{(n)} = i/n, i = 0, \ldots, n$. The left-hand side of Figure 5 shows the quadratic (co)variations of time G^N, F^N and H^N for $N = 30000$. F^N and H^N are identically zero since there are no asynchronous observations and because $T_i^{(N)} = i/N, t_i^{(n)} = \tau_i^{(n)}, 0 \le i \le n$, interpolation steps are redundant and A_T^N equals zero. The function G^N is a step function that will tend to the identity on $[0, T]$ as $N \to \infty$.
Next, we consider a situation which originates from the complete synchronous equidistant one by shifting one time-scale half a time instant $1/2N$. Then we have completely non-synchronous sampling schemes and we will call this situation intermeshed sampling. In this case the synchronous approximation is still equidistant with instants $1/N$ and, hence, G is the identity function. F and H are linear limiting functions with slope 1 and 1/4, respectively. Interpolations are carried out for all $1 \le i \le N$ for the same process for which its first observation takes place after the first observation of the other process. All interpolation steps equal $1/2N$ and thus $H' = 1/4$ follows. Since for H interpolated time instants $1/2N$ are multiplied with refresh time instants $1/N$ in **both** addends due to the specific structure, F equals the identity on $[0, T]$. The functions G^N, F^N, H^N for intermeshed sampling are illustrated in Figure 5 on the right-hand side.

In the next section we will show that for an important special case, independent homogeneous Poisson sampling, (15a)-(15c) are fulfilled when replacing deterministic convergence by convergence in probability. Furthermore, the stochastic limits $G'(t), F'(t), H'(t)$ are calculated explicitly and are again constant on $[0, T]$. For data applications one can calculate easily empirical versions $\tilde{G}'_{n,m}(t), \tilde{F}'_{n,m}(t), \tilde{H}'_{n,m}(t)$ of $G'(t), F'(t), H'(t)$ and use those as estimators for (15a)-(15c).
The key result of this section is the following Theorem 2. The detailed proof is postponed to the Appendix A. This result gives insight into the asymptotic distribution of the Hayashi-Yoshida estimator. It improves on the asymptotic normality result in [14], since the weak convergence is stable in the setting where we allow for random correlation, drift and volatility processes. The representation of the asymptotic variance using (15a)-(15c) differs from that in [15], where a similar stable convergence result is established, by
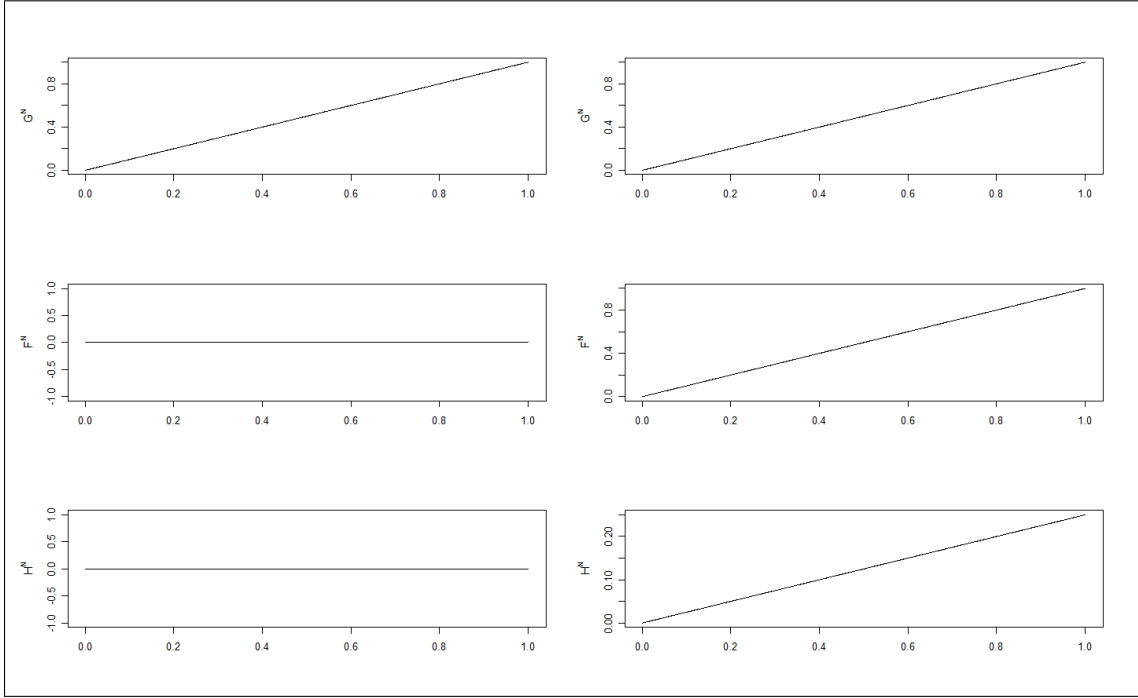
Figure 5: Quadratic (co-)variations of time for synchronous equidistant (left) and intermeshed (right) sampling.

the decomposition of the estimation error in (12) and (13) and the notion of (co-)variations of times. The latter provide helpful tools to describe the stylized facts and features of non-synchronous data and build the ground work for combined approaches for widespread generalizations and extensions of the underlying model. A very important one is the generalized multiscale estimator in [5] when market microstructure noise effects are taken into account.

Theorem 2. *The estimation error of* (9) *converges on the Assumptions 1, 2 and 3 stably in law to a centred, mixed Gaussian distribution:*

$$\sqrt{N}\left(\sum_{i=1}^{N}\left(X_{g_i}-X_{l_i}\right)\left(Y_{\gamma_i}-Y_{\lambda_i}\right)-[X\,,\,Y]_T\right)\overset{st}{\rightsquigarrow}\mathbf{N}\left(0\,,\,v_{D_T}+v_{A_T}\right)\,,\tag{16}$$

with the asymptotic variance

$$v_{D_T}+v_{A_T}=T\int_0^T G'(t)\left(\sigma_t^X\sigma_t^Y\right)^2\left(\rho_t^2+1\right)dt+T\int_0^T\left(F'(t)\left(\sigma_t^X\sigma_t^Y\right)^2+2H'(t)\left(\rho_t\sigma_t^X\sigma_t^Y\right)^2\right)dt$$

where the two addends come from the asymptotic variances of D_T^N *and* A_T^N*, respectively.*

5. Independent Poisson sampling

In this section, we consider the model in which the sequences of observation times are supposed to be realizations of two homogeneous Poisson processes that are mutually independent and independent of the processes X and Y.

Thereto, let $\tilde{n}^{(n)}(t)$ and $\tilde{m}^{(n)}(t)$ be sequences of two independent homogeneous Poisson processes with parameters Tn/θ_1 and Tn/θ_2 ($n\in\mathbb{N}$), such that the waiting times between jumps of $\tilde{n}^{(n)}$ and $\tilde{m}^{(n)}$ are exponentially distributed with expectations $\mathbb{E}\left[\Delta t_i^{(n)}\right]=\theta_1/n$ and $\mathbb{E}\left[\Delta\tau_j^{(n)}\right]=\theta_2/n$, $i\in\mathbb{N}, j\in\mathbb{N}$. Thus, $\tilde{n}^{(n)}(T)$ and $\tilde{m}^{(n)}(T)$ correspond to the sequences giving the numbers of observation times of X and

15

Y in the time span $[0, T]$. The increments of the sampling times of the closest synchronous approximation (11) are maxima of the exponentially distributed waiting times and we obtain:

$$\Delta T_k^{(n)} \sim F(t) = 1 - \exp\left(-\frac{tn}{\theta_1}\right) - \exp\left(-\frac{tn}{\theta_2}\right) + \exp\left(-tn\left(\frac{1}{\theta_1} + \frac{1}{\theta_2}\right)\right), k \in \mathbb{N}.$$

Denote $\tilde{N}(T)^{(n)} = \max_{N \in \mathbb{N}} \{\sum_{k=0}^{N} \Delta T_k^{(n)} \leq T\}$. We focus on the characteristics of the sampling schemes affecting the asymptotics of the synchronized realized covolatility estimator (9). In particular our interest is in the quadratic (co-)variations of times defined in Definition 2.

Proposition 5.1. *In the independent homogeneous Poisson model for sampling schemes, it holds true that*

$$G^N(t) \xrightarrow{p} 2\left(1 - \frac{2\theta_1^2\theta_2^2}{\theta_1^2\theta_2^2 + (\theta_1^2 + \theta_2^2)(\theta_1 + \theta_2)^2}\right)\frac{t}{T} \quad \left(= \frac{14}{9}\frac{t}{T} \text{ if } \theta_1 = \theta_2 = \theta\right), \tag{17a}$$

$$F^N(t) \xrightarrow{p} \left(\frac{2\theta_1\theta_2}{(\theta_1^2 + \theta_1\theta_2 + \theta_2^2)} + \frac{4\theta_1^2\theta_2^2}{\left(\theta_1 + \theta_2 - \frac{\theta_1\theta_2}{\theta_1+\theta_2}\right)^2(\theta_1 + \theta_2)^2}\right)\frac{t}{T} \tag{17b}$$
$$\left(= \frac{10}{9}\frac{t}{T} \text{ if } \theta_1 = \theta_2 = \theta\right),$$

$$H^N(t) \xrightarrow{p} 2\left(\frac{1}{\left(\theta_1 + \theta_2 - \frac{\theta_1\theta_2}{\theta_1+\theta_2}\right)^2}\frac{\theta_1^2\theta_2^2}{(\theta_1 + \theta_2)^2}\right)\frac{t}{T} \quad \left(= \frac{2}{9}\frac{t}{T} \text{ if } \theta_1 = \theta_2 = \theta\right). \tag{17c}$$

Proof. Poisson processes are Markovian and the exponential distribution of the increments between arrival times is memoryless. Wald's identity ensures that $\mathbb{E}\left[\sum_{k=0}^{\tilde{N}(T)^{(n)}} \Delta T_k^{(n)}\right] = \mathbb{E}\left[\tilde{N}(T)^{(n)}\right]\mathbb{E}\left[\Delta T_1^{(n)}\right]$. For the proofs of these attributes and further information on properties of mutually independent homogeneous Poisson processes we refer interested readers to [7].

First of all we ascertain that $t_i^{(n)} \neq \tau_j^{(n)} \forall (i, j) \in \{1, \ldots, \tilde{n}^{(n)}(T)\} \times \{1, \ldots, \tilde{m}^{(n)}(T)\}$ almost surely. For an arbitrarily fixed i, the expected values of next-tick, previous-tick and refresh time instants yield

$$\mathbb{E}\left[g_i^{(n)} - T_i^{(n)}\right] = \mathbb{E}\left[\left(g_i^{(n)} - T_i^{(n)}\right) \mid T_i^{(n)} = \gamma_i^{(n)}\right]\mathbb{P}\left(T_i^{(n)} = \gamma_i^{(n)}\right) = \frac{\theta_1}{n}\frac{\theta_2}{\theta_1 + \theta_2},$$

$$\mathbb{E}\left[\gamma_i^{(n)} - T_i^{(n)}\right] = \frac{\theta_2}{n}\frac{\theta_1}{\theta_1 + \theta_2},$$

$$\mathbb{E}\left[T_i^{(n)} - l_{i+1}^{(n)}\right] = \int_0^\infty y\frac{n}{\theta_2}e^{-\frac{yn}{\theta_2}}e^{-\frac{yn}{\theta_1}}\,dy = \frac{1}{n}\frac{\theta_1^2\theta_2}{(\theta_1 + \theta_2)^2},$$

$$\mathbb{E}\left[T_i^{(n)} - \lambda_{i+1}^{(n)}\right] = \frac{1}{n}\frac{\theta_1\theta_2^2}{(\theta_1 + \theta_2)^2},$$

$$\mathbb{E}\left[T_{i+1}^{(n)} - T_i^{(n)}\right] = \frac{\theta_1}{n} + \frac{\theta_2}{n} - \frac{1}{n}\frac{\theta_1\theta_2}{\theta_1 + \theta_2}.$$

The conditional expectations given that the ith refresh time $T_i^{(n)} = \gamma_i^{(n)}$ is an arrival time of $\tilde{m}^{(n)}$ yield
$$\mathbb{E}\left[T_{i+1}^{(n)} - T_i^{(n)}|T_i^{(n)} = \gamma_i^{(n)}\right] = \mathbb{E}\left[T_{i+1}^{(n)} - T_i^{(n)}\right] \text{ and } \mathbb{E}\left[T_i^{(n)} - l_{i+1}^{(n)}|T_i^{(n)} = \gamma_i^{(n)}\right] = \mathbb{E}\left[T_i^{(n)} - l_{i+1}^{(n)}\right],$$

since the latter previous-tick interpolation is zero with probability 1 if $T_i^{(n)} \neq \gamma_i^{(n)}$. Only for $(T_i^{(n)} - \lambda_i^{(n)})$ the conditional expectation differs from the unconditional and can be calculated by further conditioning

$$\mathbb{E}\left[T_i^{(n)} - \lambda_i^{(n)} | T_i^{(n)} = \gamma_i^{(n)}\right] =$$

$$\mathbb{E}\left[T_i^{(n)} - \lambda_i^{(n)} | T_i^{(n)} = \gamma_i^{(n)}, T_{i-1}^{(n)} = \lambda_i^{(n)}\right] \mathbb{P}\left(T_{i-1}^{(n)} = \lambda_i^{(n)} | T_i^{(n)} = \gamma_i^{(n)}\right)$$

$$+ \mathbb{E}\left[T_i^{(n)} - \lambda_i^{(n)} | T_i^{(n)} = \gamma_i^{(n)}, T_{i-1}^{(n)} = l_i^{(n)}\right] \mathbb{P}\left(T_{i-1}^{(n)} = l_i^{(n)} | T_i^{(n)} = \gamma_i^{(n)}\right)$$

$$= \left(\theta_1 + \theta_2 - \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}\right) \frac{\theta_1}{\theta_1 + \theta_2} + 2\theta_1 \frac{\theta_2}{\theta_1 + \theta_2} \,,$$

where the factor $2\theta_1$ in the second addend is simply the expectation of the waiting time for two jumps of $\tilde{n}$. Here, we have used some simplifying symmetry aspects, a rigorous proof using the density functions is obtained by calculation of

$$\mathbb{E}\left[T_i^{(n)} - \lambda_i^{(n)} \mathbb{1}_{\{T_i^{(n)} = \gamma_i^{(n)}, T_{i-1}^{(n)} = \lambda_i^{(n)}\}}\right] = \int_0^\infty \int_x^\infty x \frac{n}{\theta_1} e^{-x\frac{n}{\theta_1}} e^{-y\frac{n}{\theta_2}} y \frac{n}{\theta_2} e^{-x\frac{n}{\theta_1}} e^{-y\frac{n}{\theta_2}} \, dx \, dy = \frac{2\theta_1 \theta_2}{\theta_1 + \theta_2} \,.$$

The conditional expectations on $T_i^{(n)} = g_i^{(n)}$ are deduced analogously. Since $\mathbb{E}\left[T_i^{(n)} - l_i^{(n)}\right] =$ $\mathbb{E}\left[T_i^{(n)} - T_{i-1}^{(n)}\right] + \mathbb{E}\left[T_{i-1}^{(n)} - l_i^{(n)}\right]$ and the (conditional) expectations of the products occurring in G^N, F^N, H^N equal the products of (conditional) expectations thanks to the memorylessness of exponential distributions, the latter results suffice to apply the law of large numbers to the empirical (co-)variations of times. For the asymptotics of $G^N(T), F^N(T)$ and $H^N(T)$, we conclude for the number of addends $\tilde{N}(T)^{(n)}$, that $\mathbb{E}\tilde{N}(T)^{(n)} = (T/\theta)n + o(n)$ with $\theta = \theta_1 + \theta_2 - (\theta_1 \theta_2)/(\theta_1 + \theta_2)$ what follows from $\mathbb{E}\tilde{N}(T)^{(n)}\mathbb{E}\left[\Delta T_1^{(N)}\right] = T + \mathcal{O}_p(n^{-1})$ and $\mathbb{V}\mathrm{ar}\left(\tilde{N}(T)^{(n)}\right) = \mathcal{O}(n^{-1})$ since

$$\mathbb{V}\mathrm{ar}\left(\sum_{k=0}^{\tilde{N}(T)^{(n)}} \Delta T_k^{(n)}\right) = \mathbb{V}\mathrm{ar}\left(\tilde{N}(T)^{(n)}\right) \mathbb{E}\left[\left(\Delta T_1^{(n)}\right)^2\right] + \mathbb{E}\left[\tilde{N}(T)^{(n)}\right] \mathbb{V}\mathrm{ar}\left(\Delta T_1^{(n)}\right) \,.$$

The exact probability mass functions of the counting processes $\tilde{N}(t)^{(n)}$ associated with the maxima of the waiting times $\Delta t_i^{(n)}, \Delta \tau_j^{(n)}$ have a quite complicated form, so that we only give the last two results on the expectation and the variance that are necessary for the proof of the proposition.
From the preceding conclusions, it follows that

$$G^N(t) = \frac{\tilde{N}(T)^{(n)}}{T} \sum_{T_i^{(n)} \leq t} \left(\Delta T_i^{(n)}\right)^2 \xrightarrow{p} \frac{n^2}{\theta^2} \left(\frac{2\theta_1^2}{n^2} + \frac{2\theta_2^2}{n^2} - 2\left(\frac{\theta_1 \theta_2}{(\theta_1 + \theta_2)}\right)^2 \frac{1}{n^2}\right) \frac{t}{T} \,,$$

$$F^N(t) = \frac{\tilde{N}(T)^{(n)}}{T} \sum_{T_{i+1}^{(n)} \leq t} (T_i^{(n)} - \lambda_i^{(n)})(g_i^{(n)} - T_i^{(n)}) + \left(T_i^{(n)} - l_i^{(n)}\right)\left(\gamma_i^{(n)} - T_i^{(n)}\right)$$

$$+ \Delta T_{i+1}^{(n)} \left(T_i^{(n)} - l_{i+1}^{(n)}\right) + \Delta T_{i+1}^{(n)} \left(T_i^{(n)} - \lambda_{i+1}^{(n)}\right)$$

$$\xrightarrow{p} \frac{t}{T\theta^2} \left(\frac{\theta_1 \theta_2}{(\theta_1 + \theta_2)} \left(2\theta_1 + 2\theta_2 - 2\frac{\theta_1 \theta_2}{(\theta_1 + \theta_2)} + \frac{2\theta_1 \theta_2}{(\theta_1 + \theta_2)}\right)\right.$$

$$\left. + \left(\theta_1 + \theta_2 - \frac{\theta_1 \theta_2}{(\theta_1 + \theta_2)}\right) \frac{\theta_1^2 \theta_2 + \theta_1 \theta_2^2}{(\theta_1 + \theta_2)^2}\right) \,,$$

$$H^N(t) = \frac{\tilde{N}(T)^{(n)}}{T} \sum_{T_{i+1}^{(n)} \leq t} \left(T_i^{(n)} - l_{i+1}^{(n)}\right)\left(g_i^{(n)} - T_i^{(n)}\right) + \left(T_i^{(n)} - \lambda_{i+1}^{(n)}\right)\left(\gamma_i^{(n)} - T_i^{(n)}\right)$$

$$\xrightarrow{p} \frac{t}{T\theta^2} \frac{\theta_1^2 \theta_2^2 (\theta_1 + \theta_2)}{(\theta_1 + \theta_2)^3} \,.$$
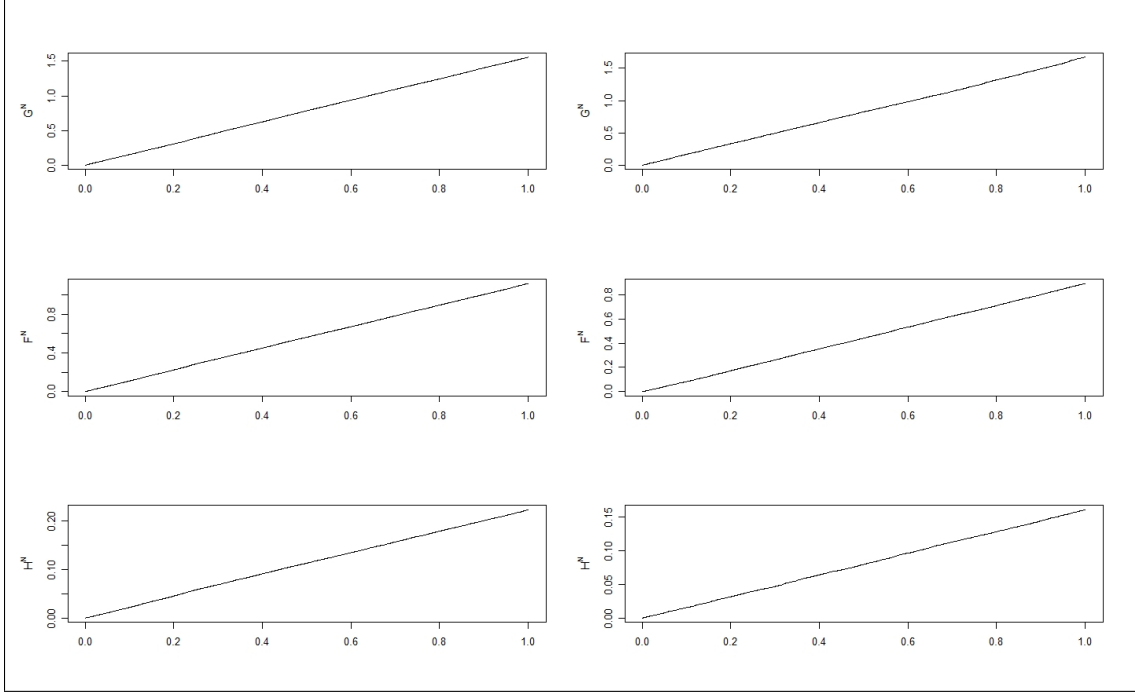
Figure 6: Quadratic (Co-)variations of times for homogeneous Poisson sampling.

Inserting θ we obtain formulae (17a)-(17c). In the evaluation of G^N we have also used the second moment of $\Delta T_1^{(n)}$ which can be calculated using the above given distribution function. $\qquad\square$

Figure 6 depitcs the quadratic (co-)variations of times for simulated mutually independent homogeneous Poisson processes. On the left-hand side both parameters have been set $\theta = 1$ for $T = 1$ and $n = 30000$. The stochastic limits are linear increasing functions on $[0, 1]$ with slope $14/9, 10/9, 2/9$ and $1/4$, respectively. On the right-hand side we see the (co-)variations of times for $T = 1, n = 30000, \theta_1 = 1, \theta_2 = 0.5$. Those tend in probability to linear limiting functions with slope $82/49, 44/49, 8/49$ and $2/9$, respectively.

In the model of non-synchronously observed Itô processes X and Y which fulfill Assumption 1 and observation times following an independent Poisson sampling scheme of the above given form, we derive the following stable central limit theorem as special case of Theorem 2:

Corollary 5.2. *The estimation error of the synchronized realized covolatility estimator* (9) *converges on the Assumption 1 conditionally on the independent Poisson sampling scheme with* $0 < \theta_1 < \infty$ *and* $0 < \theta_2 < \infty$ *stably in law to a centred mixed Gaussian distribution:*

$$\sqrt{\tilde{N}(T)^{(n)}} \left(\sum_{i=0}^{\tilde{N}(T)^{(n)}} \left(X_{g_i^{(n)}} - X_{l_i^{(n)}} \right) \left(Y_{\gamma_i^{(n)}} - Y_{\lambda_i^{(n)}} \right) - [X, Y]_T \right) \overset{st}{\rightsquigarrow} \mathbf{N}\left(0, v_T\right) , \qquad (18)$$

with the asymptotic variance

$$v_T = 2 \int_0^T \left(\rho_t \sigma_t^X \sigma_t^Y\right)^2 \, dt + \left(2\frac{\theta_1 \theta_2}{\theta(\theta_1 + \theta_2)} + 1 \right) \int_0^T \left(\sigma_t^X \sigma_t^Y\right)^2$$

where the two addends come from the asymptotic variances of the discretization error D_T^N of the closest synchronous approximation (12) *and the additional error A_T^N due to interpolations* (13)*, respectively, and* $\theta = \theta_1 + \theta_2 - \frac{\theta_1 \theta_2}{\theta_1 + \theta_2}$.

18

Proof. It is a basic result in the theory of extreme values that for the supremum of n i. i. d. exponentially distributed waiting times ΔT_i with $\mathbb{E}\Delta T_i = n^{-1}$, it holds true that $\sup_i (\Delta T_i) = \mathcal{O}_p(\log(n)/n)$. We refer to [11] for a proof. In the setting of mutually independent homogeneous Poisson processes with parameters Tn/θ_1 and Tn/θ_2, we conclude that $\sup_{i \in \{1,\ldots,\tilde{N}(T)^{(n)}\}} = \mathcal{O}_p\left(\log \tilde{N}(T)^{(n)}/\tilde{N}(T)^{(n)}\right)$. Hence, Assumption 2 holds for the sampling design where the orders of δ_n^X, δ_m^Y hold in probability. Then all findings in the proofs of Propositions A.2 and A.5 stay valid when we insert the (co-)variations of time deduced above in the limits of the variances. $\qquad\square$

The stable convergence holds conditionally given the observation times, what means that endogenous observation times are not covered but Poisson sampling independent of the processes X and Y.
The asymptotic variance of the mixed Gaussian limit is in line with the results by [14] and [15]. We remark that one has to pay attention to the proportionality to θ in the rate $\tilde{N}(T)^{(n)}$ when comparing the asymptotic variance to the one in [15].
From an applied point of view, the model considered in this section could be criticized for its flaw that sampling schemes of two correlated processes are modeled to follow two independent processes and for time homogeneity. Both seems to be rather unrealistic in financial time series. However, independent and homogeneous Poisson sampling times designs constitute the most commonly used model in this research area (cf. [23], [13] among others) because they are handy and allow for explicit calculations while the model is not too far away from the real world.

6. Asymptotic variance estimation

Finally, we state a consistent estimator for the asymptotic variance of the Hayashi-Yoshida estimator (9) from Theorem 2. Since in [14] a central limit theorem for the case of deterministic correlation and volatility functions has been proved, the asymptotic variance is non-random in that setting. In a recent publication [15], in that the authors also generalize the asymptotic distribution result to a stable central limit theorem in the setting of random volatility and correlation functions, a consistent estimation method for the asymptotic variance is provided using kernel estimates. Our estimator differs from this method and we incorporate only one time transformed histogram-type estimator.

Proposition 6.1. *Define the estimator*

$$\widehat{\mathbf{AVAR}}_{HY} := N \sum_{j=1}^{N-1} (X_{g_j} - X_{l_j})(Y_{\gamma_j} - Y_{\lambda_j}) \left[(X_{g_j} - X_{l_j})(Y_{\gamma_j} - Y_{\lambda_j}) \right.$$
$$\left. + 2(X_{g_{j+1}} - X_{l_{j+1}})(Y_{\gamma_{j+1}} - Y_{\lambda_{j+1}}) \right] - 3T\tilde{I}_1$$

with

$$\tilde{I}_1 := \sum_{j=1}^{K_N} \left(\frac{\widehat{\Delta[X,Y]}_{G_j^N}^{HY}}{\Delta G_j^N} \right)^2 \frac{G^N(T)}{K_N}$$

being a histogram-based estimator for $\int_0^T (\rho_t \sigma_t^X \sigma_t^Y)^2 G'(t)dt$. The estimators for the increase of the quadratic covariation on bins are Hayashi-Yoshida estimators of the type

$$\widehat{\Delta[X,Y]}_{G_j^N}^{HY} := \sum_{r \in [G_j^N, G_{j+1}^N)} (X_{g_r} - X_{l_r})(Y_{\gamma_r} - Y_{\lambda_r}).$$

It holds true that

$$\widehat{\mathbf{AVAR}}_{HY} \xrightarrow{p} T \int_0^T G'(t) (\sigma_t^X \sigma_t^Y)^2 (\rho_t^2 + 1)\, dt + T \int_0^T \left(F'(t)(\sigma_t^X \sigma_t^Y)^2 dt + 2H'(t)(\rho_t \sigma_t^X \sigma_t^Y)^2 dt \right)$$

19

on the Assumptions 1, 2 and 3. Thus, we have on hand a consistent estimator for the asymptotic variance of the Hayashi-Yoshida estimator and the feasible stable central limit theorem

$$\frac{\widehat{[X,Y]}_T^{(HY)}}{\sqrt{\widehat{\mathbf{AVAR}}_{HY}}} \overset{st}{\leadsto} \mathbf{N}(0,1) \ . \tag{19}$$

For more motivation and details on the construction of histogram estimators for which bins are chosen equispaced according to a transformed timescale associated with a certain monotonic function, as the asymptotic quadratic variation of refresh times here, we refer to [5]. Proposition 6.1 is proved in Appendix B.

A. Proof of Theorem 2

A.1. Discretization error of the synchronous approximation

Proposition A.1. *On the Assumptions 1, 2 and* (15a) *the discretization error of the closest synchronous approximation converges stably in law to a centred mixed Gaussian distribution:*

$$\sqrt{\frac{N}{T}} D_T^N \overset{st}{\leadsto} \mathbf{N} \left(0, \int_0^T G'(t) \left(\sigma_t^X \sigma_t^Y \right)^2 \left(\rho_t^2 + 1 \right) dt \right) \ . \tag{A.1}$$

Proof. In the proofs superscripts of the sampling times are frequently omitted to increase the readability. First note that on Assumption 1, by Girsanov's theorem we may without loss of generality further suppose that $\mu_t^X = \mu_t^Y = 0$ identically since we have learned in Section 2 that stable convergence is commutative with measure change. Let M_t and L_t be the continuous martingales $L_t = \int_0^t \sigma_s^X \, dW_s^X$, $M_t = \int_0^t \sigma_s^Y \, dW_s^Y$ where W^X, W^Y are two standard Brownian motions with quadratic covariation $\left[W^X, W^Y \right]_t = \int_0^t \rho_s \sigma_s^X \sigma_s^Y \, ds$ and denote $L_i = \int_0^{T_i} \sigma_s^X \, dW_s^X$, $M_i = \int_0^{T_i} \sigma_s^Y \, dW_s^Y$.

Proposition A.2. *On the same Assumptions as in Proposition A.1, the process $\mathcal{D}_t^N$ defined by*

$$\mathcal{D}_t^N := \sqrt{\frac{N}{T}} \sum_{T_i^{(N)} \leq t} (L_i - L_{i-1})(M_i - M_{i-1}) - \int_0^t \rho_s \sigma_s^X \sigma_s^Y \, ds$$

for $0 \leq t \leq T$ converges as $N \to \infty$ stably in law:

$$\mathcal{D}_t^N \overset{st}{\leadsto} \int_0^t \sqrt{v_{\mathcal{D}_s}} dW_s^\perp \tag{A.2}$$

where $W^\perp$ is a Brownian motion independent of $\mathcal{F}$ and

$$v_{\mathcal{D}_s} = G'(s)(\sigma_s^X \sigma_s^Y)^2 (\rho_s^2 + 1) \ . \tag{A.3}$$

Proof. We will prove this stable convergence of the process associated with the transformed discretization error by application of Jacod's stable limit Theorem 1. It is also possible to use the discrete-time version of this Theorem from Corollary 2.3 which we apply in the next subsection.
Using the definition of the quadratic covariation process of martingales or integration by parts formula, we find an illustration of the discretization error by a sum of stochastic integrals and an asymptotically

negligible term:

$$\sum_{T_i^{(N)} \leq t} \left(L_{T_i} - L_{T_{i-1}}\right) \left(M_{T_i} - M_{T_{i-1}}\right) = \sum_{T_i^{(N)} \leq t} \left(L_i - L_{i-1}\right) \left(M_i - M_{i-1}\right)$$

$$= \sum_{T_i^{(N)} \leq t} \left(L_i M_i - L_i M_{i-1} - M_i L_{i-1} + L_{i-1} M_{i-1}\right)$$

$$= \sum_{T_i^{(N)} \leq t} \Bigg(\int_{T_{i-1}}^{T_i} L_s dM_s + \int_{T_{i-1}}^{T_i} M_s dL_s + \Delta\left[L, M\right]_{T_i}$$

$$- M_{i-1}(L_i - L_{i-1}) - L_{i-1}(M_i - M_{i-1}) \Bigg)$$

$$= \left[L, M\right]_{\mathfrak{T}(t)} - \left[L, M\right]_{T_0} + \sum_{T_i^{(N)} \leq t} \left(\int_{T_{i-1}}^{T_i} (L_s - L_{i-1}) dM_s + \int_{T_{i-1}}^{T_i} (M_s - M_{i-1}) dL_s \right)$$

where we denote $\mathfrak{T}(t) := \max_i (T_i^{(N)} \leq t)$.
Thus, we obtain

$$\sqrt{\frac{N}{T}} \mathcal{D}_t^N = \sqrt{\frac{N}{T}} \sum_{T_i^{(N)} \leq t} \left(\int_{T_{i-1}}^{T_i} (L_s - L_{i-1}) dM_s + \int_{T_{i-1}}^{T_i} (M_s - M_{i-1}) dL_s \right) + \mathcal{O}_p \left(\sqrt{\frac{N}{T}} \right),$$

since $\left[L, M\right]_{\mathfrak{T}(t)} - \left[L, M\right]_{T_0} = \left[L, M\right]_t + \mathcal{O}_p(1)$. Consider the centred continuous martingale

$$\phi_\tau^{(N)} := \sqrt{\frac{N}{T}} \Bigg(\sum_{T_i^{(N)} \leq t} \left(\int_{T_{i-1}}^{T_i} (L_s - L_{i-1}) dM_s + \int_{T_{i-1}}^{T_i} (M_s - M_{i-1}) dL_s \right)$$

$$+ \int_{\mathfrak{T}(t)}^{\tau} (L_s - L_{\mathfrak{T}(t)}) dM_s + \int_{\mathfrak{T}(t)}^{\tau} (M_s - M_{\mathfrak{T}(t)}) dL_s \Bigg), \quad \tau \in [\mathfrak{T}(t), t].$$

We calculate the corresponding quadratic variation process at time t:

$$\left[\phi^{(N)}\right]_t = \frac{N}{T} \left[\sum_{T_i^{(N)} \leq t} \left(\int_{T_{i-1}}^{T_i} (L_s - L_{i-1}) dM_s + \int_{T_{i-1}}^{T_i} (M_s - M_{i-1}) dL_s \right) \right]_t + \left[\phi^{(N)}\right]_t - \left[\phi^{(N)}\right]_{\mathfrak{T}(t)}$$

$$= \frac{N}{T} \sum_{T_i^{(N)} \leq t} \Bigg(\int_{T_{i-1}}^{T_i} (L_s - L_{i-1})^2 d\left[M\right]_s + \int_{T_{i-1}}^{T_i} (M_s - M_{i-1})^2 d\left[L\right]_s$$

$$+ 2 \int_{T_{i-1}}^{T_i} (L_s - L_{i-1})(M_s - M_{i-1}) d\left[M, L\right]_s \Bigg) + \left[\phi^{(N)}\right]_t - \left[\phi^{(N)}\right]_{\mathfrak{T}(t)}$$

$$\underset{\text{(Lemma A.3)}}{=} \frac{N}{T} \sum_{T_i^{(N)} \leq t} \Bigg(\int_{T_{i-1}}^{T_i} \left[L - L_{i-1}\right]_s d\left[M\right]_s + \int_{T_{i-1}}^{T_i} \left[M - M_{i-1}\right]_s d\left[L\right]_s$$

$$+ 2 \int_{T_{i-1}}^{T_i} \left[L - L_{i-1}\right]_s \left[M - M_{i-1}\right]_s d\left[M, L\right]_s \Bigg) + \mathcal{O}_p(1)$$

$$= \frac{N}{T} \sum_{i=1}^{N} \Bigg(\int_{T_{i-1}}^{T_i} d\left(\left[L - L_{i-1}\right]_s \left[M - M_{i-1}\right]_s\right)$$

$$+ 2 \int_{T_{i-1}}^{T_i} \left[L - L_{i-1}\right]_s \left[M - M_{i-1}\right]_s d\left[M, L\right]_s \Bigg) + \mathcal{O}_p(1)$$

$$= \frac{N}{T} \sum_{T_i^{(N)} \leq t} \left(\left(\int_{T_{i-1}}^{T_i} (\sigma_s^X)^2 \, ds \int_{T_{i-1}}^{T_i} (\sigma_s^Y)^2 \, ds \right) + \left(\int_{T_{i-1}}^{T_i} \rho_s \sigma_s^X \sigma_s^Y \, ds \right)^2 + \mathcal{O}_p(1) \right.$$

$$= \frac{N}{T} \sum_{T_i^{(N)} \leq t} \left(\left((\overline{\rho \sigma^X \sigma^Y})_i \right)^2 (\Delta T_i)^2 + \left((\overline{\sigma^X})_i (\overline{\sigma^Y})_i \right)^2 (\Delta T_i)^2 \right) + \mathcal{O}_p(1)$$

$$= \sum_{T_i^{(N)} \leq t} \left(\frac{(G^{(N)}(T_i) - G^{(N)}(T_{i-1}))}{\Delta T_i} \left(\left(\sigma_{T_{i-1}}^X \sigma_{T_{i-1}}^Y \right)^2 \left(1 + \rho_{T_{i-1}}^2 \right) \right) \Delta T_i \right) + \mathcal{O}_p(1)$$

$$\xrightarrow{p} \int_0^t G'(s)(\rho_s^2 + 1)(\sigma_s^X \sigma_s^Y)^2 \, ds \ .$$

In this calculation we have used integration by parts and the change of variables Theorem for the integrals with quadratic covariation integrators that are of finite variation. The second last equality is an application of the mean value theorem (the volatility and the correlation processes are continuous and thus also bounded on compact sets) where the constants $(\overline{\sigma^X})_i$, $(\overline{\sigma^Y})_i$ and $(\overline{\rho \sigma^X \sigma^Y})_i$ come from. The Riemann sum converges and with Definition 2 and Assumption 3 this yields the convergence in probability of the quadratic variation to $\int_0^t G'(s)(\rho_s^2 + 1)(\sigma_s^X \sigma_s^Y)^2 \, ds = \int_0^t v_{\mathcal{D}_s}$. The third equality above is proved in:

Lemma A.3. *It holds true that the approximation error terms*

$$\sum_{T_i^{(N)} \leq t} \left(\int_{T_{i-1}}^{T_i} (L_s - L_{i-1})^2 d \left[M - M_{i-1} \right]_s - \int_{T_{i-1}}^{T_i} [L - L_{i-1}]_s \, d \left[M - M_{i-1} \right]_s \right) \tag{A.4a}$$

$$\sum_{T_i^{(N)} \leq t} \left(\int_{T_{i-1}}^{T_i} (M_s - M_{i-1})^2 d \left[L - L_{i-1} \right]_s - \int_{T_{i-1}}^{T_i} [M - M_{i-1}]_s \, d \left[L - L_{i-1} \right]_s \right) \tag{A.4b}$$

$$\sum_{T_i^{(N)} \leq t} \left(\int_{T_{i-1}}^{T_i} (M_s - M_{i-1})(L_s - L_{i-1}) d \left[M, L \right]_s - \int_{T_{i-1}}^{T_i} [M - M_{i-1}, L - L_{i-1}]_s \, d \left[M, L \right]_s \right) \tag{A.4c}$$

$$\frac{N}{T} \sum_{T_i^{(n)} \leq t} (\Delta T_i)^2 \left((\overline{\rho \sigma^X \sigma^Y})_i^2 + (\overline{\sigma^X})_i^2 (\overline{\sigma^Y})_i^2 - \left(\left(\rho_{T_{i-1}} \sigma_{T_{i-1}}^X \sigma_{T_{i-1}}^Y \right)^2 + \left(\sigma_{T_{i-1}}^X \sigma_{T_{i-1}}^Y \right)^2 \right) \right) \tag{A.4d}$$

converge to zero in probability.

Proof. The proofs for (A.4a) and (A.4b) are completely analogous and we restrict ourselves to prove it for (A.4a). By Itô's formula

$$(L_s - L_{i-1})^2 = 2 \int_{T_{i-1}}^s (L_r - L_{i-1}) dL_r + [L - L_{i-1}]_s$$

holds. The left-hand side of (A.4a) equals

$$\sum_{T_i^{(N)} \leq t} \left(\int_{T_{i-1}}^{T_i} \left(2 \int_{T_{i-1}}^s (L_r - L_{i-1}) dL_r \right) d \left[M - M_{i-1} \right]_r \right)$$

$$= \sum_{T_i^{(N)} \leq t} \left(2 \int_{T_{i-1}}^{T_i} (L_s - L_{i-1})([M - M_{i-1}]_{T_i}) dL_s - 2 \int_{T_{i-1}}^{T_i} (L_s - L_{i-1})([M - M_{i-1}]_s) dL_s \right)$$

by application of the integration by parts formula in the way

$$Z_{T_i} [M - M_{i-1}]_{T_i} = \int_0^{T_i} Z_t d [M - M_{i-1}]_t + \int_0^{T_i} [M - M_{i-1}]_t dZ_t$$

with $Z_t := \int_{T_{i-1}}^t 2(L_s - L_{i-1}) dL_s$ for $T_{i-1} \le t \le T$ to the addends. Therefore, we can write the left-hand side of (A.4a) in the way $\mathcal{M}_1^{(N)} + \mathcal{M}_2^{(N)}$ with two centred continuous martingales $\mathcal{M}_1^{(N)}, \mathcal{M}_2^{(N)}$ defined in the fashion of $\phi^{(N)}$ above and calculate the quadratic covariation processes at time t:

$$\left[\mathcal{M}_2^{(N)} \right]_t = 4 \sum_{T_i^{(N)} \le t} \left(\int_{T_{i-1}}^{T_i} (L_s - L_{i-1})^2 ([M - M_{i-1}]_s)^2 d [L]_s \right) + \mathcal{o}_p(1)$$

$$\le 4 \max_i \sup_{s \in (T_{i-1}, T_i]} (L_s - L_{i-1})^2 \max_i \sup_{s \in (T_{i-1}, T_i]} [M - M_{i-1}]_s^2 \sum_{T_i^{(N)} \le t} \int_{T_{i-1}}^{T_i} d [L]_s + \mathcal{o}_p(1) \,.$$

The first addend is up to a logarithmic factor $\mathcal{O}_p(\delta_N^3)$ and hence $\mathcal{M}_2^{(N)} = \mathcal{o}_p(1)$ on Assumption 2. That $\mathcal{M}_1^{(N)} = \mathcal{o}_p(1)$ is proved analogously. This implies that (A.4a) is $\mathcal{o}_p(1)$.
The strategy of the proof for (A.4c) follows the same approach, starting with the equation

$$(L_s - L_{i-1})(M_s - M_{i-1}) = \int_{T_{i-1}}^s (L_r - L_{i-1}) d(M - M_{i-1})_r + \int_{T_{i-1}}^s (M_r - M_{i-1}) d(L - L_{i-1})_r$$

$$+ [L - L_{i-1}, M - M_{i-1}]_s$$

and applying integration by parts as above with $Z_t = \int_{T_{i-1}}^t (L_s - L_{i-1}) d(M - M_{i-1})_s + \int_{T_{i-1}}^t (M_s - M_{i-1}) d(L - L_{i-1})_s$ for $T_{i-1} \le t \le T_i$.
We complete the proof of the convergence of the quadratic variation with the proof for (A.4d). Denote $\overline{(\rho \sigma^X \sigma^Y)}_i^2 = \widetilde{(\sigma^X)^2}_i \cdot \widetilde{(\sigma^Y)^2}_i \cdot \widetilde{(\rho)^2}_i$ to distinguish between the values from the application of the mean value theorems to the two different addends. An upper bound of the left-hand side of (A.4d) can be found by elementary algebra and the triangle inequality for the absolute value:

$$\frac{N}{T} \sum_{T_i^{(N)} \le t} (\Delta T_i)^2 \left((\widetilde{\rho} \widetilde{\sigma^X} \widetilde{\sigma^Y})_i^2 + (\overline{\sigma^X})_i^2 (\overline{\sigma^Y})_i^2 - \left(\left(\rho_{T_{i-1}} \sigma_{T_{i-1}}^X \sigma_{T_{i-1}}^Y \right)^2 + \left(\sigma_{T_{i-1}}^X \sigma_{T_{i-1}}^Y \right)^2 \right) \right)$$

$$\le \frac{N}{T} \sum_{T_i^{(N)} \le t} (\Delta T_i)^2 \left(\widetilde{(\sigma^X)^2}_i \widetilde{(\sigma^Y)^2}_i \left| \widetilde{(\rho)^2}_i - \rho_{T_{i-1}}^2 \right| + \widetilde{(\sigma^Y)^2}_i \rho_{T_{i-1}}^2 \left| \widetilde{(\sigma^X)^2}_i - (\sigma_{T_{i-1}}^X)^2 \right| \right.$$

$$\left. + \rho_{T_{i-1}}^2 (\sigma_{T_{i-1}}^X)^2 \left| \widetilde{(\sigma^Y)^2}_i - (\sigma_{T_{i-1}}^Y)^2 \right| + (\overline{\sigma^Y})_i^2 \left| (\overline{\sigma^X})_i^2 - (\sigma_{T_{i-1}}^X)^2 \right| + (\sigma_{T_{i-1}}^X)^2 \left| (\overline{\sigma^Y})_i^2 - (\sigma_{T_{i-1}}^Y)^2 \right| \right)$$

$$= \mathcal{o}_p(1) \,.$$

$\square$

The martingales $\phi^{(N)}$ can be written for every N as time-changed Brownian motions $B_{[\phi^{(N)}]_t}^{(DDS,N)} = \phi_t^{(N)}$ by the Dambis-Dubins-Schwarz theorem. The sequence of martingales $\phi^{(N)}$ or associated time-changed Dambis-Dubins-Schwarz Brownian motions converges weakly to a limiting Brownian motion by the asymptotic Knight-theorem. The limiting Brownian motion will be defined on an orthogonal extension of the original probability space. To obtain the stable convergence result, we apply Jacod's Theorem 1 and thus, we are left to verify conditions (2a) and (2b).
Consider the quadratic covariation process of $\phi^{(N)}$ and the reference martingale L

$$\left[L, \phi^{(N)} \right]_t = \sqrt{\frac{N}{T}} \sum_{T_i^{(N)} \le t} \left(\int_{T_{i-1}}^{T_i} (L_s - L_{i-1}) d [M, L]_s + \int_{T_{i-1}}^{T_i} (M_s - M_{i-1}) d [L]_s \right) + \mathcal{o}_p(1) \,.$$

The term of smaller order than 1 in probability comes from the increment of the covariation process on $[\mathfrak{T}(t), t]$. As before, this equality holds true for all t, since for $t < T_1$ the covariation is $o_p(1)$. Integration by parts yields:

$$\left[L, \phi^{(N)}\right]_t = \sqrt{\frac{N}{T}} \sum_{T_i^{(N)} \leq t} \left[([M,L]_{T_i} - [M,L]_{T_{i-1}})(L_i - L_{i-1}) - \int_{T_{i-1}}^{T_i} [M,L]_s \, d(L_s - L_{i-1}) \right.$$
$$\left. + ([L]_{T_i} - [L]_{T_{i-1}})(M_i - M_{i-1}) - \int_{T_{i-1}}^{T_i} [L]_s \, d(M_s - M_{i-1}) \right] .$$

It remains to show that this term converges to zero in probability. The term is centred and using Itô isometry we find the following upper bound for the second moment:

$$\mathbb{E}\left[\left(\left[L, \phi^{(N)}\right]_t \right)^2 \right] \leq 2 \frac{N}{T} \mathbb{E}\left[\sum_{T_i^{(N)} \leq t} \left(([M,L]_{T_i} - [M,L]_{T_{i-1}})^2 (L_i - L_{i-1})^2 \right. \right.$$
$$+ ([L]_{T_i} - [L]_{T_{i-1}})^2 (M_i - M_{i-1})^2 \Big)$$
$$+ \max_{i \in \{1,\dots,N\}} \sup_{s \in (T_{i-1}, T_i]} ([M,L]_s - [M,L]_{T_{i-1}})^2 \sum_i \int_{T_{i-1}}^{T_i} d\left[L - L_{i-1}\right]_t$$
$$\left. + \max_{i \in \{1,\dots,N\}} \sup_{s \in (T_{i-1}, T_i]} ([L]_s - [L]_{T_{i-1}})^2 \sum_i \int_{T_{i-1}}^{T_i} d\left[M - M_{i-1}\right]_t \right]$$
$$= \mathcal{O}\left(N \delta_N^2\right) .$$

The term is bounded by a constant times $N \delta_N^2$ since squared increments, cross products of increments and increments of the quadratic (co-)variations of L and M over time instants $\Delta T_i^{(N)}$ are bounded by $\Delta T_i^{(N)}$ times a constant. To sums with products of time instants we can apply Hölder's inequality with the supremum norm to obtain upper bounds. There are at most order δ_N^{-1} time instants $\Delta T_i^{(N)}$ of order $\sup_i \Delta T_i^{(N)} = \delta_N$ since $\sum_i \Delta T_i^{(N)} \leq T$ and the time span T is fixed.
Hence, $\left[L, \phi^{(N)}\right]_t = o_p(1) \quad \forall t \in [0,T]$. With the same strategy $\left[M, \phi^{(N)}\right]_t = o_p(1) \quad \forall t \in [0,T]$ can be shown.
For every bounded $\mathcal{F}_t$-martingale $L^\perp$ satisfying $\left[L, L^\perp\right] \equiv 0$ the covariation

$$\left[L^\perp, \phi^{(N)}\right]_t = \frac{N}{T} \sum_{T_i^{(N)} \leq t} \left(\int_{T_{i-1}}^{T_i} (L_s^\perp - L_{i-1}^\perp) d\left[M, L^\perp\right]_s \right) + o_p(1) = o_p(1)$$

converges to zero. The same holds true for every bounded $\mathcal{F}_t$-martingale orthogonal to M. Applying Theorem 1, we deduce that Proposition A.2 holds true. □

Proposition A.1 is a direct consequence of the stronger result in Proposition A.2 since for $t = T$ the marginal distribution is a mixed normal distribution which is independent of $\mathcal{F}$. The stable convergence assures that the convergence also holds under the original probability measure and non-zero drift terms with the same asymptotic law. □

A.2. Error due to non-synchronicity

Proposition A.4. *Let Assumptions 1, 2 and* (15b)-(15c) *from Assumption 3 be satisfied. The error A_T^N due to the lack of synchronicity converges stably in law to a centred mixed Gaussian distribution:*

$$\sqrt{\frac{N}{T}} A_T^N \overset{st}{\rightsquigarrow} \mathbf{N}\left(0, v_{A_T}\right) , \tag{A.5}$$

with asymptotic variance

$$v_{A_T} = \int_0^T F'(t) \left(\sigma_t^X \sigma_t^Y \right)^2 dt + \int_0^T 2H'(t) \left(\rho_t \sigma_t^X \sigma_t^Y \right)^2 dt \ . \tag{A.6}$$

Proof. First, we write the ith increments occurring as factors in the addends of the estimator (9) as the sum of the next-tick interpolation at T_i, the increments $\Delta X_{T_i} = X_{T_i} - X_{T_{i-1}}$ and $\Delta Y_{T_i} = Y_{T_i} - Y_{T_{i-1}}$, respectively, and the previous-tick interpolation at T_{i-1} and multiply out the addends.

$$\widehat{[X,Y]}_T = \sum_{i=1}^N \left(X_{g_i} - X_{T_i} + X_{T_i} - X_{T_{i-1}} + X_{T_{i-1}} - X_{l_i} \right) \left(Y_{\gamma_i} - Y_{T_i} + Y_{T_i} - Y_{T_{i-1}} + Y_{T_{i-1}} - Y_{\lambda_i} \right)$$

$$= \sum_{i=1}^N \left((X_{g_i} - X_{T_i}) \Delta Y_{T_i} + (Y_{\gamma_i} - Y_{T_i}) \Delta X_{T_i} + (X_{T_{i-1}} - X_{l_i}) \Delta Y_{T_i} + (Y_{T_{i-1}} - Y_{\lambda_i}) \Delta X_{T_i} \right.$$

$$\left. + (X_{g_i} - X_{T_i})(Y_{T_{i-1}} - Y_{\lambda_i}) + (Y_{\gamma_i} - Y_{T_i})(X_{T_{i-1}} - X_{l_i}) \right) + D_T^N$$

The indicator functions in (13) have been dropped since the corresponding addends are zero if the indicator functions were zero. Since at least one of the next-tick interpolation errors is zero and as well one of the previous-tick interpolation errors, too, two addends, namely the products of next-tick interpolation errors and the product of previous-tick interpolation errors, equal zero. Thus, the error due to asynchronicity can be written as the sum of the remaining six terms (where at least another three equal zero in each addend). We conclude, that the error A_T^N can be expressed in the following way:

$$A_T^N = \sum_{i=1}^{N-1} \left((X_{g_i} - X_{T_i})(Y_{T_i} - Y_{\lambda_i}) + (Y_{\gamma_i} - Y_{T_i})(X_{T_i} - X_{l_i}) \right.$$

$$\left. (X_{T_{i+1}} - X_{T_i})(Y_{T_i} - Y_{\lambda_{i+1}}) + (X_{T_i} - X_{l_{i+1}})(Y_{T_{i+1}} - Y_{T_i}) \right) + o_p(1) \ .$$

In this equality an index shift has been applied to the partial sum of previous-tick interpolated errors multiplied with ΔX_{T_i} and ΔY_{T_i}, respectively, leading to the structure that in the ith addend the factors contain next- and previous-tick interpolated errors to the same T_i. The $o_p(1)$-term emerges from end-effects when shifting the original sum.

In the last illustration of A_T^N consecutive addends of the sum are uncorrelated in contrast to the non-shifted illustration. The reason is that, if without loss of generality $\gamma_i = T_i$ holds, $(X_{g_i} - X_{T_i}) \Delta Y_{T_i}$ and $(X_{T_i} - X_{l_{i+1}}) \Delta Y_{T_{i+1}}$ have in general a non-zero correlation whereas $(X_{g_i} - X_{T_i}) \Delta Y_{T_i}$ and $(X_{T_{i-1}} - X_{l_i}) \Delta Y_{T_i}$ are uncorrelated. Furthermore, the fact that $\gamma_i = T_i \Rightarrow \lambda_{i+1} = T_i$ assures that the addends in the last illustration of A_T^N are uncorrelated. Roughly speaking we capture correlation between subsequent addends of the outer sum and transfer it into additional correlation in the inner sum.

As in the foregoing proof of Proposition A.1, it is sufficient to prove the stable convergence result for the zero-drift case. We denote, as before, the corresponding transformed processes $L_t = \int_0^t \sigma_s^X dW_s^X$ and $M_t = \int_0^t \sigma_s^Y dW_s^Y$.

Consider the sum

$$\mathcal{A}_t^N = \sum_{T_{i+1}^{(N)} \leq t} \Delta A_i^N := \sqrt{\frac{N}{T}} \sum_{T_{i+1}^{(N)} \leq t} \left((L_{g_i} - L_{T_i})(M_{T_i} - M_{\lambda_i}) + (M_{\gamma_i} - M_{T_i})(L_{T_i} - L_{l_i}) \right.$$

$$\left. + (L_{T_i} - L_{l_{i+1}})(M_{T_{i+1}} - M_{T_i}) + (M_{T_i} - M_{\lambda_{i+1}})(L_{T_{i+1}} - L_{T_i}) \right) \tag{A.7}$$

for fixed $0 \leq t \leq T$.

Proposition A.5. *Assume the same conditions as in Proposition A.4. For fixed $0 \leq t \leq T$ the transformed error due to non-synchronicity $\mathcal{A}_t^N$ is the endpoint of a discrete, centred, square-integrable martingale with respect to the filtration $\mathcal{F}_{i,N} := \mathcal{F}_{T_{i+1}^{(N)}}$. The process $\mathcal{A}_t^N$ converges as $N \to \infty$ stably in law:*

$$\mathcal{A}_t^N \overset{st}{\rightsquigarrow} \mathcal{A}_t = \int_0^t \sqrt{v_{\mathcal{A}_s}} dW_s^\perp \tag{A.8}$$

25

where $W^\perp$ is a Brownian motion independent of $\mathcal{F}$ and

$$v_{\mathcal{A}_s} = F'(s)\left(\sigma_s^X \sigma_s^Y\right)^2 + 2H'(s)\left(\rho_s \sigma_s^X \sigma_s^Y\right)^2 . \tag{A.9}$$

Proof. The expectation of the absolute value of the sum is bounded for all $t \in [0,T]$ and ΔA_i^N, $i = 0, \ldots, N$ are $\mathcal{F}_{i,N} = \mathcal{F}_{T_{i+1}^{(N)}}$-measurable. Since

$$
\begin{aligned}
\mathbb{E}\left[\Delta A_i^N | \mathcal{F}_{i-1,N}\right] &= \mathbb{E}\left[\Delta A_i^N | \mathcal{F}_{T_i^{(N)}}\right] \\
&= \mathbb{E}\left[(L_{g_i} - L_{T_i})(M_{T_i} - M_{\lambda_i}) + (M_{\gamma_i} - M_{T_i})(L_{T_i} - L_{l_i}) \right. \\
&\qquad \left. + (L_{T_i} - L_{l_{i+1}})\Delta M_{T_{i+1}} + (M_{T_i} - M_{\lambda_{i+1}})\Delta L_{T_{i+1}} | \mathcal{F}_{T_i^{(N)}}\right] \\
&= \mathbb{E}\left[L_{g_i} - L_{T_i}\right](M_{T_i} - M_{\lambda_i}) + \mathbb{E}\left[M_{\gamma_i} - M_{T_i}\right](L_{T_i} - L_{l_i}) \\
&\qquad + (L_{T_i} - L_{l_{i+1}})\mathbb{E}\left[\Delta M_{T_{i+1}}\right] + (M_{T_i} - M_{\lambda_{i+1}})\mathbb{E}\left[\Delta L_{T_{i+1}}\right] = 0
\end{aligned}
$$

for the conditional expectation of the increments holds, $\mathcal{A}_t^N$ is the endpoint of a $\mathcal{F}_{i,N}$-martingale.
The stable weak convergence to a limiting Brownian motion is proven with Corollary 2.3 to Jacod's Theorem 1.
First, we verify the conditional Lindeberg condition that is implied by the stronger conditional Lyapunov condition. It is sufficient to proof the following:

Lemma A.6. *The sum of the conditional fourth moments of the martingale increments A_i^N converges to zero in probability:*

$$\mathbb{E}\left[\sum_{T_{i+1}^{(N)} \leq t}\left(\Delta A_i^N\right)^4 \Big| \mathcal{F}_{i-1,N}\right] = o_p(1) .$$

Proof. Throughout the proof C denotes a generic constant that does not depend on N. We consider different addends of the fourth conditional moments consecutively. The sum of conditional fourth moments incorporates addends of the following types:

- fourth-order moments:

$$\frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[(L_{g_i} - L_{T_i})^4\right](M_{T_i} - M_{\lambda_i})^4 ,$$

- second-order moments:

$$\frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[(L_{g_i} - L_{T_i})^2 (\Delta M_{T_{i+1}})^2\right](L_{T_i} - L_{l_{i+1}})^2 (M_{T_i} - M_{\lambda_i})^2 ,$$

- third- and first-order moments:

$$\frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} 4(M_{T_i} - M_{\lambda_i})^3 (L_{T_i} - L_{l_{i+1}})^3 \mathbb{E}\left[\Delta M_{T_{i+1}}(L_{g_i} - L_{T_i})\right] .$$

For the partial sum with addends of the first type an application of the Burkholder-Davis-Gundy (BDG)

inequalities yields

$$\frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[(L_{g_i} - L_{T_i})^4\right](M_{T_i} - M_{\lambda_i})^4$$

$$\leq C \frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[\left(\int_{T_i}^{g_i} (\sigma_s^X)^2 ds\right)^2\right](M_{T_i} - M_{\lambda_i})^4$$

$$\leq C \frac{N^2}{T^2} \sup_{s \in [0,T]} (\sigma_s^X)^2 \sum_{T_{i+1}^{(N)} \leq t} (M_{T_i} - M_{\lambda_i})^4 (g_i - T_i)^2 \leq \mathcal{O}_p\left(N\delta_N^2\right) = o_p(1) \,.$$

The last inequality can be deduced by the result that the convergence $(N/(3T)) \sum_i (\Delta M_{T_i})^4 \to \int_0^t (\sigma_s^Y)^4 ds$ holds almost surely as $N \to \infty$ for the so-called realized quarticity ([3]) and that $(g_i - T_i) \leq \delta_N$. Without the result about the convergence of the realized quarticity, the asymptotic order in probability can be derived by the convergence to zero of the expectation of the above sum and calculating the second moment that is bounded from above by a constant times $N^4 \delta_N^7$.

For the partial sum including addends that incorporate second-order moments we obtain an upper bound by application of the Cauchy-Schwarz inequality and the BDG inequalities:

$$\frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} 6 \,\mathbb{E}\left[(L_{g_i} - L_{T_i})^2 (\Delta M_{T_{i+1}})^2\right](L_{T_i} - L_{l_{i+1}})^2 (M_{T_i} - M_{\lambda_i})^2$$

$$\leq \frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} 6 \sqrt{\mathbb{E}\left[(L_{g_i} - L_{T_i})^4\right]} \sqrt{\mathbb{E}\left[(\Delta M_{T_{i+1}})^4\right]}(L_{T_i} - L_{l_{i+1}})^2 (M_{T_i} - M_{\lambda_i})^2$$

$$\leq C \frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} 6 \left(\mathbb{E}\left[\left(\int_{T_i}^{g_i} (\sigma_s^X)^2 ds\right)^2\right] \mathbb{E}\left[\left(\int_{T_i}^{T_{i+1}} (\sigma_s^Y)^2 ds\right)^2\right]\right)^{\frac{1}{2}} (L_{T_i} - L_{l_{i+1}})^2 (M_{T_i} - M_{\lambda_i})^2$$

$$= o_p(1) \,.$$

The stochastic order follows, since the term has the expectation

$$C \frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} 6 \left(\mathbb{E}\left[\left(\int_{T_i}^{g_i} (\sigma_s^X)^2 ds\right)^2\right] \mathbb{E}\left[\left(\int_{T_i}^{T_{i+1}} (\sigma_s^Y)^2 ds\right)^2\right]\right)^{\frac{1}{2}} \mathbb{E}\left[(L_{T_i} - L_{l_{i+1}})^2 (M_{T_i} - M_{\lambda_i})^2\right]$$

$$\leq C \frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} 6 \left(\mathbb{E}\left(\int_{T_i}^{g_i} (\sigma_s^X)^2 ds\right)^2 \mathbb{E}\left(\int_{T_i}^{T_{i+1}} (\sigma_s^Y)^2 ds\right)^2 \mathbb{E}\left(\int_{l_{i+1}}^{T_i} (\sigma_s^X)^2 ds\right)^2 \mathbb{E}\left(\int_{\lambda_i}^{T_i} (\sigma_s^Y)^2 ds\right)^2\right)^{\frac{1}{2}}$$

$$\leq C N^2 \delta_N^3 = o(1) \,,$$

where again the Cauchy-Schwarz and BDG inequalities have been applied. The variance is bounded from above by a constant times $N^4 \delta_N^7$, what can be shown by a similar calculation where thanks to the fact that $T_i = \gamma_i \Rightarrow \lambda_{i+1} = T_i$ the addends are uncorrelated and the variance of the sum equals the sum of variances.

We treat the third type of addends occurring in the sum of conditional fourth moments in the same way. Itô isometry yields

$$\frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} 4(M_{T_i} - M_{\lambda_i})^3 (L_{T_i} - L_{l_{i+1}})^3 \mathbb{E}\left[\Delta M_{T_{i+1}}(L_{g_i} - L_{T_i})\right]$$

$$= \frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} 4(M_{T_i} - M_{\lambda_i})^3 (L_{T_i} - L_{l_{i+1}})^3 \mathbb{E}\left[\int_{T_i}^{g_i} \rho_s \sigma_s^X \sigma_s^Y ds\right] \,.$$

This term has expectation

$$\frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} 4\mathbb{E}\left[(M_{T_i} - M_{\lambda_i})^3 (L_{T_i} - L_{l_{i+1}})^3\right] \mathbb{E}\left[\int_{T_i}^{g_i} \rho_s \sigma_s^X \sigma_s^Y ds\right]$$

$$\leq \frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} 4\sqrt{\mathbb{E}\left[(M_{T_i} - M_{\lambda_i})^6\right] \mathbb{E}\left[(L_{T_i} - L_{l_{i+1}})^6\right]} \mathbb{E}\left[\int_{T_i}^{g_i} \rho_s \sigma_s^X \sigma_s^Y ds\right]$$

$$\leq C\frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} 4\left(\mathbb{E}\left[\left(\int_{\lambda_i}^{T_i} (\sigma_s^Y)^2 ds\right)^3\right] \mathbb{E}\left[\left(\int_{l_{i+1}}^{T_i} (\sigma_s^X)^2 ds\right)^3\right]\right)^{1/2} \mathbb{E}\left[\int_{T_i}^{g_i} \rho_s \sigma_s^X \sigma_s^Y ds\right]$$

$$\leq CN^2 \delta_N^3 = o(1),$$

and an analogous calculation as before yields that the variance is of order $N^4 \delta_N^7$.
Thereby, the sum converges to zero in probability. $\qquad\square$

Next, we consider the sum of conditional variances of the increments of the discrete martingale.

Lemma A.7.

$$\mathbb{E}\left[\sum_{T_{i+1}^{(N)} \leq t} \left(\Delta A_i^N\right)^2 \Big| \mathcal{F}_{T_i^{(N)}}\right] \xrightarrow{p} \int_0^t F'(s) \left(\sigma_s^X \sigma_s^Y\right)^2 ds + \int_0^t 2H'(s) \left(\rho_s \sigma_s^X \sigma_s^Y\right)^2 ds. \qquad (A.10)$$

It holds true that

Proof.

$$\mathbb{E}\left[\sum_{T_{i+1}^{(N)} \leq t} \left(\Delta A_i^N\right)^2 \Big| \mathcal{F}_{T_i^{(N)}}\right]$$

$$= \frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[(L_{g_i} - L_{T_i})^2 (M_{T_i} - M_{\lambda_i})^2 + (M_{\gamma_i} - M_{T_i})^2 (L_{T_i} - L_{l_i})^2 + (L_{T_i} - L_{l_{i+1}})^2 (\Delta M_{T_{i+1}})^2\right.$$

$$+ (M_{T_i} - M_{\lambda_{i+1}})^2 (\Delta L_{T_{i+1}})^2 + 2(L_{g_i} - L_{T_i})(M_{T_i} - M_{\lambda_i})(L_{T_i} - L_{l_{i+1}})\Delta M_{T_{i+1}}$$

$$\left. + 2(M_{\gamma_i} - M_{T_i})(L_{T_i} - L_{l_i})(M_{T_i} - M_{l_{i+1}})\Delta L_{T_{i+1}} \Big| \mathcal{F}_{T_i^{(N)}}\right]$$

$$\underset{\text{(Itô isometry)}}{=} \frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left(\mathbb{E}\left[\int_{T_i}^{g_i} (\sigma_s^X)^2 ds\right](M_{T_i} - M_{\lambda_i})^2 + \mathbb{E}\left[\int_{T_i}^{\gamma_i} (\sigma_s^Y)^2 ds\right](L_{T_i} - L_{l_i})^2\right.$$

$$+ (L_{T_i} - L_{l_{i+1}})^2 \mathbb{E}\left[\int_{T_i}^{T_{i+1}} (\sigma_s^Y)^2 ds\right] + (M_{T_i} - M_{\lambda_{i+1}})^2 \mathbb{E}\left[\int_{T_i}^{T_{i+1}} (\sigma_s^X)^2 ds\right]$$

$$+ 2(M_{T_i} - M_{\lambda_i})(L_{T_i} - L_{l_{i+1}})\mathbb{E}\left[\int_{T_i}^{g_i} \rho_s \sigma_s^X \sigma_s^Y ds\right]$$

$$\left. + 2(L_{T_i} - L_{l_i})(M_{T_i} - M_{l_{i+1}})\mathbb{E}\left[\int_{T_i}^{\gamma_i} \rho_s \sigma_s^X \sigma_s^Y ds\right]\right)$$

$$
\underset{\text{(Lemma A.8)}}{=} \frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left(\mathbb{E}\left[\int_{T_i}^{g_i} (\sigma_s^X)^2 ds \right] \int_{\lambda_i}^{T_i} (\sigma_s^Y)^2 ds + \mathbb{E}\left[\int_{T_i}^{\gamma_i} (\sigma_s^Y)^2 ds \right] \int_{l_i}^{T_i} (\sigma_s^X)^2 ds \right.
$$

$$
+ \int_{l_{i+1}}^{T_i} (\sigma_s^X)^2 ds \, \mathbb{E}\left[\int_{T_i}^{T_{i+1}} (\sigma_s^Y)^2 ds \right] + \int_{\lambda_{i+1}}^{T_i} (\sigma_s^Y)^2 ds \, \mathbb{E}\left[\int_{T_i}^{T_{i+1}} (\sigma_s^X)^2 ds \right]
$$

$$
\left. + \int_{l_{i+1}}^{T_i} \rho_s \sigma_s^X \sigma_s^Y ds \, \mathbb{E}\left[\int_{T_i}^{g_i} \rho_s \sigma_s^X \sigma_s^Y ds \right] + \int_{l_{i+1}}^{T_i} \rho_s \sigma_s^X \sigma_s^Y ds \, \mathbb{E}\left[\int_{T_i}^{\gamma_i} \rho_s \sigma_s^X \sigma_s^Y ds \right] \right) + \mathcal{O}_p(1)
$$

$$
\underset{\text{(Lemma A.8)}}{=} \frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left(\int_{T_i}^{g_i} (\sigma_s^X)^2 ds \int_{\lambda_i}^{T_i} (\sigma_s^Y)^2 ds + \int_{T_i}^{\gamma_i} (\sigma_s^Y)^2 ds \int_{l_i}^{T_i} (\sigma_s^X)^2 ds \right.
$$

$$
+ \int_{l_{i+1}}^{T_i} (\sigma_s^X)^2 ds \int_{T_i}^{T_{i+1}} (\sigma_s^Y)^2 ds + \int_{\lambda_{i+1}}^{T_i} (\sigma_s^Y)^2 ds \int_{T_i}^{T_{i+1}} (\sigma_s^X)^2 ds
$$

$$
\left. + 2 \int_{l_{i+1}}^{T_i} \rho_s \sigma_s^X \sigma_s^Y ds \int_{T_i}^{g_i} \rho_s \sigma_s^X \sigma_s^Y ds + 2 \int_{l_{i+1}}^{T_i} \rho_s \sigma_s^X \sigma_s^Y ds \int_{T_i}^{\gamma_i} \rho_s \sigma_s^X \sigma_s^Y ds \right) + \mathcal{O}_p(1)
$$

$$
\underset{\text{(Lemma A.9)}}{=} \frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left((\sigma_{T_i}^X \sigma_{T_i}^Y)^2 \left((T_i - \lambda_i)(g_i - T_i) + (\gamma_i - T_i)(T_i - l_i) + (T_i - l_{i+1})\Delta T_{i+1} \right. \right.
$$

$$
+ (T_i - \lambda_{i+1})\Delta T_{i+1}) + (\rho_{T_i} \sigma_{T_i}^X \sigma_{T_i}^Y)^2 (2(T_i - l_{i+1})(g_i - T_i) + 2(T_i - \lambda_{i+1})(\gamma_i - T_i))) + \mathcal{O}_p(1)
$$

$$
= \sum_{T_{i+1}^{(N)} \leq t} \frac{F(T_{i+1}) - F(T_i)}{T_{i+1} - T_i} (\sigma_{T_i}^X \sigma_{T_i}^Y)^2 \Delta T_{i+1} + 2 \frac{H(T_{i+1}) - H(T_i)}{T_{i+1} - T_i} (\rho_{T_i} \sigma_{T_i}^X \sigma_{T_i}^Y)^2 \Delta T_{i+1} + \mathcal{O}_p(1)
$$

$$
\xrightarrow{p} \int_0^t F'(s) \left(\sigma_s^X \sigma_s^Y \right)^2 ds + \int_0^t 2 H'(s) \left(\rho_s \sigma_s^X \sigma_s^Y \right)^2 ds .
$$

In the last step we have involved Definition 2. The Riemann sum converges on the Assumption 3 (in particular (15b) and (15c)) in probability as $N \to \infty$ to the expression $\int_0^t v_{\mathcal{A}_s} ds$ with $v_{\mathcal{A}_s}$ given in Proposition A.5.

The detailed proofs of the approximations are postponed in the following two lemmas.

Lemma A.8. *On the assumptions as before, the following equations hold true:*

$$
\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left((M_{T_i} - M_{\lambda_i})^2 - \int_{\lambda_i}^{T_i} (\sigma_s^Y)^2 ds \right) \mathbb{E}\left[\int_{T_i}^{g_i} (\sigma_s^X)^2 ds \right] = \mathcal{O}_p(1) ,
$$

$$
\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \int_{\lambda_i}^{T_i} (\sigma_s^Y)^2 ds \left(\mathbb{E}\left[\int_{T_i}^{g_i} (\sigma_s^X)^2 ds \right] - \int_{T_i}^{g_i} (\sigma_s^X)^2 ds \right) = \mathcal{O}_p(1) ,
$$

$$
\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left((L_{T_i} - L_{l_i})^2 - \int_{l_i}^{T_i} (\sigma_s^X)^2 ds \right) \mathbb{E}\left[\int_{T_i}^{\gamma_i} (\sigma_s^Y)^2 ds \right] = \mathcal{O}_p(1) ,
$$

$$
\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \int_{l_i}^{T_i} (\sigma_s^X)^2 ds \left(\mathbb{E}\left[\int_{T_i}^{\gamma_i} (\sigma_s^Y)^2 ds \right] - \int_{T_i}^{\gamma_i} (\sigma_s^Y)^2 ds \right) = \mathcal{O}_p(1) ,
$$

$$
\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left((M_{T_i} - M_{\lambda_{i+1}})^2 - \int_{\lambda_{i+1}}^{T_i} (\sigma_s^Y)^2 ds \right) \mathbb{E}\left[\int_{T_i}^{T_{i+1}} (\sigma_s^X)^2 ds \right] = \mathcal{O}_p(1) ,
$$

$$\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \int_{\lambda_{i+1}}^{T_i} (\sigma_s^Y)^2 ds \left(\mathbb{E}\left[\int_{T_i}^{T_{i+1}} (\sigma_s^X)^2 ds \right] - \int_{T_i}^{T_{i+1}} (\sigma_s^X)^2 ds \right) = \mathcal{o}_p(1) \,,$$

$$\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left((L_{T_i} - L_{l_{i+1}})^2 - \int_{l_{i+1}}^{T_i} (\sigma_s^X)^2 ds \right) \mathbb{E}\left[\int_{T_i}^{T_{i+1}} (\sigma_s^Y)^2 ds \right] = \mathcal{o}_p(1) \,,$$

$$\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \int_{l_{i+1}}^{T_i} (\sigma_s^X)^2 ds \left(\mathbb{E}\left[\int_{T_i}^{T_{i+1}} (\sigma_s^Y)^2 ds \right] - \int_{T_i}^{T_{i+1}} (\sigma_s^Y)^2 ds \right) = \mathcal{o}_p(1) \,,$$

$$\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left((M_{T_i} - M_{\lambda_i})(L_{T_i} - L_{l_{i+1}}) - \int_{l_{i+1}}^{T_i} \rho_s \sigma_s^X \sigma_s^Y ds \right) \mathbb{E}\left[\int_{T_i}^{g_i} \rho_s \sigma_s^X \sigma_s^Y ds \right] = \mathcal{o}_p(1) \,,$$

$$\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \int_{l_{i+1}}^{T_i} \rho_s \sigma_s^X \sigma_s^Y ds \left(\mathbb{E}\left[\int_{T_i}^{g_i} \rho_s \sigma_s^X \sigma_s^Y ds \right] - \int_{T_i}^{g_i} \rho_s \sigma_s^X \sigma_s^Y ds \right) = \mathcal{o}_p(1) \,,$$

$$\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left((L_{T_i} - L_{l_i})(M_{T_i} - L_{\lambda_{i+1}}) - \int_{\lambda_{i+1}}^{T_i} \rho_s \sigma_s^X \sigma_s^Y ds \right) \mathbb{E}\left[\int_{T_i}^{\gamma_i} \rho_s \sigma_s^X \sigma_s^Y ds \right] = \mathcal{o}_p(1) \,,$$

$$\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \int_{\lambda_{i+1}}^{T_i} \rho_s \sigma_s^X \sigma_s^Y ds \left(\mathbb{E}\left[\int_{T_i}^{\gamma_i} \rho_s \sigma_s^X \sigma_s^Y ds \right] - \int_{T_i}^{\gamma_i} \rho_s \sigma_s^X \sigma_s^Y ds \right) = \mathcal{o}_p(1) \,.$$

Proof. We restrict ourselves to the proof of the first two equalities, since all other terms can shown to converge to zero in probability in an analogous way. The left-hand side of the first equality has an expectation equal to zero which can be concluded directly by Itô isometry:

$$\mathbb{E}\left[\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left((M_{T_i} - M_{\lambda_i})^2 - \int_{\lambda_i}^{T_i} (\sigma_s^Y)^2 ds \right) \mathbb{E}\left[\int_{T_i}^{g_i} (\sigma_s^X)^2 ds \right] \right] = 0 \,.$$

In order to derive the stochastic order of the term, consider the second moment:

$$\mathbb{E}\left[\left(\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left((M_{T_i} - M_{\lambda_i})^2 - \int_{\lambda_i}^{T_i} (\sigma_s^Y)^2 ds \right) \mathbb{E}\left[\int_{T_i}^{g_i} (\sigma_s^X)^2 ds \right] \right)^2 \right]$$

$$= \frac{N^2}{T^2} \sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[(M_{T_i} - M_{\lambda_i})^4 - 2(M_{T_i} - M_{\lambda_i})^2 \int_{\lambda_i}^{T_i} (\sigma_s^Y)^2 ds + \left(\int_{\lambda_i}^{T_i} (\sigma_s^Y)^2 ds \right)^2 \right]$$

$$\times \left(\mathbb{E}\left[\int_{T_i}^{g_i} (\sigma_s^X)^2 ds \right] \right)^2 = \mathcal{o}(1) \,,$$

where the asymptotic order is deduced by Itô isometry and the BDG inequalities. Since the error induced by this term in the approximation of the conditional variance before is centred and has a variance converging

to zero as $N \to \infty$, the error is asymptotically negligible.

In the second equality we consider the error when the expected increment of the quadratic variation of X over the next-tick interpolated time interval is substituted by the integral itself. We proceed as before for the first approximation. Since

$$\mathbb{E}\left[\frac{N}{T}\sum_{T_{i+1}^{(N)}\leq t}\int_{\lambda_i}^{T_i}(\sigma_s^Y)^2 ds\left(\mathbb{E}\left[\int_{T_i}^{g_i}(\sigma_s^X)^2 ds\right]-\int_{T_i}^{g_i}(\sigma_s^X)^2 ds\right)\right]=0$$

and

$$\mathbb{E}\left[\left(\frac{N}{T}\sum_{T_{i+1}^{(N)}\leq t}\int_{\lambda_i}^{T_i}(\sigma_s^Y)^2 ds\left(\mathbb{E}\left[\int_{T_i}^{g_i}(\sigma_s^X)^2 ds\right]-\int_{T_i}^{g_i}(\sigma_s^X)^2 ds\right)\right)^2\right]$$

$$=\frac{N^2}{T^2}\sum_{T_{i+1}^{(N)}\leq t}\mathbb{V}\mathrm{ar}\left(\int_{\lambda_i}^{T_i}(\sigma_s^Y)^2 ds\right)\mathbb{E}\left[\left(\int_{\lambda_i}^{T_i}(\sigma_s^Y)^2 ds\right)^2\right]=o(1)\,,$$

the approximation error is asymptotically negligible. The fact that $\gamma_i = T_i \Rightarrow \lambda_{i+1} = T_i$ has been used that guarantees that the addends of the sum are uncorrelated. $\square$

Lemma A.8 has been applied in the second and third equality in the evaluation of the sum of conditional variances and the proof of Lemma A.7 is completed by the following

Lemma A.9. *On the same assumptions as before, the following equation holds true*

$$\frac{N}{T}\sum_{T_{i+1}^{(N)}\leq t}\left(\int_{T_i}^{g_i}(\sigma_s^X)^2 ds\int_{\lambda_i}^{T_i}(\sigma_s^Y)^2 ds-(\sigma_{T_i}^X\sigma_{T_i}^Y)^2(T_i-\lambda_i)(g_i-T_i)\right)=o_p(1)$$

and analogously the errors in the five other addends converge to zero in probability when replacing the product of increments of quadratic (co-)variations by the values of $\rho_{T_i}, \sigma_{T_i}^X, \sigma_{T_i}^Y$ multiplied with the corresponding times increments.

Proof. We prove the equality explicitly given in the lemma. The five remaining terms can be handled by the same strategy. By an application of the mean value theorem, elementary algebra and the triangle inequality for the absolute values, we deduce that

$$\frac{N}{T}\sum_{T_{i+1}^{(N)}\leq t}\left(\int_{T_i}^{g_i}(\sigma_s^X)^2 ds\int_{\lambda_i}^{T_i}(\sigma_s^Y)^2 ds-(\sigma_{T_i}^X\sigma_{T_i}^Y)^2(T_i-\lambda_i)(g_i-T_i)\right)$$

$$=\frac{N}{T}\sum_{T_{i+1}^{(N)}\leq t}\left((\sigma_{\varsigma_i}^X\sigma_{\xi_i}^Y)^2-(\sigma_{T_i}^X\sigma_{T_i}^Y)^2\right)(T_i-\lambda_i)(g_i-T_i)$$

$$\leq\frac{N}{T}\sum_{T_{i+1}^{(N)}\leq t}\left|(\sigma_{\varsigma_i}^X\sigma_{\xi_i}^Y)^2-(\sigma_{T_i}^X\sigma_{T_i}^Y)^2\right|(T_i-\lambda_i)(g_i-T_i)$$

$$\leq\frac{N}{T}\sum_{T_{i+1}^{(N)}\leq t}\left(\left((\sigma_{\xi_i}^Y)^2\left|(\sigma_{\varsigma_i}^X)^2-(\sigma_{T_i}^X)^2\right|+(\sigma_{T_i}^X)^2\left|(\sigma_{\xi_i}^Y)^2-(\sigma_{T_i}^Y)^2\right|\right)(T_i-\lambda_i)(g_i-T_i)\right)$$

$$\leq C\frac{N}{T}\sum_{T_{i+1}^{(N)}\leq t}\left(\left(\sup_{s\in[\lambda_i,T_i]}\left|(\sigma_s^X)^2-(\sigma_{T_i}^X)^2\right|+\sup_{s\in[T_i,g_i]}\left|(\sigma_s^Y)^2-(\sigma_{T_i}^Y)^2\right|\right)(T_i-\lambda_i)(g_i-T_i)\right)$$

$$=o_p(1)$$

holds on Assumption 1. $\square$

$\square$

To prove the stability of the convergence in Proposition A.5, we show in the following that the discrete covariations of $\mathcal{A}_t^N$ with the $\mathcal{F}$-generating underlying martingales L and M converge to zero in probability.

Lemma A.10.

$$\sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[\Delta A_i^N \Delta L_{T_{i+1}^{(N)}}\Big|\mathcal{F}_{T_i^{(N)}}\right] \xrightarrow{p} 0\,,$$

$$\sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[\Delta A_i^N \Delta M_{T_{i+1}^{(N)}}\Big|\mathcal{F}_{T_i^{(N)}}\right] \xrightarrow{p} 0\,.$$

Proof. Both relations are proven similarly and we restrict ourselves to the proof of the first one. The left-hand side equals

$$\sqrt{\frac{N}{T}} \sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\Big[\Delta L_{T_{i+1}}\left((L_{g_i} - L_{T_i})(M_{T_i} - M_{\lambda_i}) + (M_{\gamma_i} - M_{T_i})(L_{T_i} - L_{l_i})\right.$$

$$\left. + (L_{T_i} - L_{l_{i+1}})(M_{T_{i+1}} - M_{T_i}) + (M_{T_i} - M_{\lambda_{i+1}})(L_{T_{i+1}} - L_{T_i})\right)\Big|\mathcal{F}_{T_i^{(N)}}\Big]$$

$$= \sqrt{\frac{N}{T}} \sum_{T_{i+1}^{(N)} \leq t} \left(\mathbb{E}\left[\int_{T_i}^{g_i} (\sigma_s^X)^2 ds\right](M_{T_i} - M_{\lambda_i}) + \mathbb{E}\left[\int_{T_i}^{\gamma_i} \rho_s \sigma_s^X \sigma_s^Y ds\right](L_{T_i} - L_{l_i})\right.$$

$$\left. + (L_{T_i} - L_{l_{i+1}})\mathbb{E}\left[\int_{T_i}^{T_{i+1}} \rho_s \sigma_s^X \sigma_s^Y ds\right] + (M_{T_i} - M_{\lambda_{i+1}})\mathbb{E}\left[\int_{T_i}^{T_{i+1}} (\sigma_s^X)^2 ds\right]\right) =: \Gamma\,.$$

Γ is centred and using Itô isometry the variance is shown to converge to zero:

$$\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \left(\left(\mathbb{E}\left[\int_{T_i}^{g_i} (\sigma_s^X)^2 ds\right]\right)^2 \mathbb{E}\left[\int_{\lambda_i}^{T_i} (\sigma_s^Y)^2 ds\right] + \left(\mathbb{E}\left[\int_{T_i}^{\gamma_i} \rho_s \sigma_s^X \sigma_s^Y ds\right]\right)^2 \mathbb{E}\left[\int_{l_i}^{T_i} (\sigma_s^X)^2 ds\right]\right.$$

$$+ \mathbb{E}\left[\int_{l_{i+1}}^{T_i} (\sigma_s^X)^2 ds\right]\left(\mathbb{E}\left[\int_{T_i}^{T_{i+1}} \rho_s \sigma_s^X \sigma_s^Y ds\right]\right)^2 + \mathbb{E}\left[\int_{\lambda_{i+1}}^{T_i} (\sigma_s^Y)^2 ds\right]\left(\mathbb{E}\left[\int_{T_i}^{T_{i+1}} (\sigma_s^X)^2 ds\right]\right)^2$$

$$+ 2\mathbb{E}\left[\int_{l_{i+1}}^{T_i} \rho_s \sigma_s^X \sigma_s^Y ds\right] \mathbb{E}\left[\int_{T_i}^{g_i} (\sigma_s^X)^2 ds\right] \mathbb{E}\left[\int_{T_i}^{T_{i+1}} \rho_s \sigma_s^X \sigma_s^Y ds\right]$$

$$\left. + 2\mathbb{E}\left[\int_{\lambda_{i+1}}^{T_i} \rho_s \sigma_s^Y \sigma_s^Y ds\right] \mathbb{E}\left[\int_{T_i}^{T_{i+1}} (\sigma_s^X)^2 ds\right] \mathbb{E}\left[\int_{T_i}^{\gamma_i} \rho_s \sigma_s^X \sigma_s^Y ds\right]\right)$$

$$\leq CN\delta_N^2 = o(1)\,.$$

Once more we can conclude that the addends are uncorrelated since $\gamma_i = T_i \Rightarrow \lambda_{i+1} = T_i$ and $g_i = T_i \Rightarrow l_{i+1} = T_i$, respectively. $\square$

Finally, we prove that the discrete covariation of our considered martingale with every bounded $\mathcal{F}_t$-martingale that is orthogonal to L_t or M_t, converges to zero in probability. Hence, this lemma will complete the proof of Proposition A.5.

Lemma A.11. *Assume that $L_t^\perp$ and $M_t^\perp$ are bounded $\mathcal{F}_t$-martingales, with $[L, L^\perp] \equiv 0$ and $[M, M^\perp] \equiv 0$, respectively. It holds true that*

$$\sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[\Delta A_i^N \Delta L_{T_{i+1}^{(N)}}^\perp \Big| \mathcal{F}_{T_i^{(N)}}\right] \xrightarrow{p} 0\,,$$

$$\sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[\Delta A_i^N \Delta M_{T_{i+1}^{(N)}}^\perp \Big| \mathcal{F}_{T_i^{(N)}}\right] \xrightarrow{p} 0\,.$$

Proof. As in the preceding lemma, we only prove the first part of the result. The left-hand side of the first equation equals

$$\sqrt{\frac{N}{T}} \sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[\Delta L_{T_{i+1}}^\perp \left((L_{g_i} - L_{T_i})(M_{T_i} - M_{\lambda_i}) + (M_{\gamma_i} - M_{T_i})(L_{T_i} - L_{l_i})\right.\right.$$

$$\left.\left. + (L_{T_i} - L_{l_{i+1}})(M_{T_{i+1}} - M_{T_i}) + (M_{T_i} - M_{\lambda_{i+1}})(L_{T_{i+1}} - L_{T_i})\right) \Big| \mathcal{F}_{T_i^{(N)}}\right]$$

$$= \sqrt{\frac{N}{T}} \sum_{T_{i+1}^{(N)} \leq t} \left(\mathbb{E}\left[\int_{T_i}^{g_i} d\left[L, L^\perp\right]_s\right] (M_{T_i} - M_{\lambda_i}) + \mathbb{E}\left[\int_{T_i}^{\gamma_i} d\left[M, L^\perp\right]_s\right] (L_{T_i} - L_{l_i})\right.$$

$$\left. + (L_{T_i} - L_{l_{i+1}})\mathbb{E}\left[\int_{T_i}^{T_{i+1}} d\left[M, L^\perp\right]_s\right] + (M_{T_i} - M_{\lambda_{i+1}})\mathbb{E}\left[\int_{T_i}^{T_{i+1}} d\left[L, L^\perp\right]_s\right]\right)$$

$$= \sqrt{\frac{N}{T}} \sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[\int_{T_i}^{T_{i+1}} d\left[M, L^\perp\right]_s\right] (L_{T_i} - L_{l_{i+1}}) + \mathbb{E}\left[\int_{T_i}^{\gamma_i} d\left[M, L^\perp\right]_s\right] (L_{T_i} - L_{l_i})\,.$$

This term is centred and the has the variance

$$\frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[\left(\int_{T_i}^{T_{i+1}} d\left[M, L^\perp\right]_s\right)^2\right] \mathbb{E}\left[(L_{T_i} - L_{l_{i+1}})^2\right] + \mathbb{E}\left[\left(\int_{T_i}^{\gamma_i} d\left[M, L^\perp\right]_s\right)^2\right]$$

$$\times \mathbb{E}\left[(L_{T_i} - L_{l_i})^2\right] + \mathbb{E}\left[(L_{T_i} - L_{L_{i+1}})^2\right] \mathbb{E}\left[\int_{T_i}^{T_{i+1}} d\left[M, L^\perp\right]_s\right] \mathbb{E}\left[\int_{T_i}^{\gamma_i} d\left[M, L^\perp\right]_s\right]$$

$$\underset{\text{Itô isometry}}{=} \frac{N}{T} \sum_{T_{i+1}^{(N)} \leq t} \mathbb{E}\left[\left(\int_{T_i}^{T_{i+1}} d\left[M, L^\perp\right]_s\right)^2\right] \mathbb{E}\left[\int_{l_{i+1}}^{T_i} (\sigma_s^X)^2 ds\right]$$

$$+ \mathbb{E}\left[\left(\int_{T_i}^{\gamma_i} d\left[M, L^\perp\right]_s\right)^2\right] \mathbb{E}\left[\int_{l_i}^{T_i} (\sigma_s^X)^2 ds\right]$$

$$+ \mathbb{E}\left[\int_{l_{i+1}}^{T_i} (\sigma_s^X)^2 ds\right] \mathbb{E}\left[\int_{T_i}^{T_{i+1}} d\left[M, L^\perp\right]_s\right] \mathbb{E}\left[\int_{T_i}^{\gamma_i} d\left[M, L^\perp\right]_s\right] = \mathcal{O}(1)\,.$$

Thus, the covariations converge to zero in probability. $\qquad\square$

The Lemma completes the proof of Proposition A.5. $\qquad\square$

The mixed normal limit in Proposition A.4 can be obtained as the marginal distribution of $\mathcal{A}_T^N$ in $t = T$. $\qquad\square$

Proposition A.4 for the error of the approximation by the discretization error of the closest synchronous approximation (13) and the stable limit theorem for this synchronous discretization error (12) given in Proposition A.1 suffice to imply Theorem 2. That is because the multivariate stable convergence Theorem

2.4 applies to the vector of the two uncorrelated terms and since the covariations converge to zero , the stable convergence to the mixed Gaussian limit with the sum of the two asymptotic variances is concluded.$\square$

B. Proof of Proposition 6.1

The proof will be divided into three parts in that the sum of squared products, products of consecutive increments and the histogram estimator are considered, respectively. Denote $X_j^+ = X_{g_j} - X_{T_j}$, $X_j^- = X_{T_{j-1}} - X_{l_j}$ and $X_j^S = X_{T_j} - X_{T_{j-1}}$, $j = 1, \ldots, N$. In the first step it is proved that

$$N \sum_{j=1}^{N-1} \left(X_{g_j} - X_{l_j} \right)^2 \left(Y_{\gamma_j} - Y_{\lambda_j} \right)^2 \xrightarrow{p} T \int_0^T G'(t) \left(\sigma_t^X \sigma_t^Y \right)^2 \left(2\rho_t^2 + 1 \right) dt + T \int_0^T F'(t) \left(\sigma_t^X \sigma_t^Y \right)^2 dt \ .$$

$$N \sum_{j=1}^{N-1} \left(X_j^+ + X_j^S + X_j^- \right)^2 \left(Y_j^+ + Y_j^S + Y_j^- \right)^2 = N \sum_{j=1}^{N-1} \left((X_j^+)^2 (Y_j^S + Y_j^-)^2 + (Y_j^+)^2 (X_j^S + X_j^-)^2 \right.$$
$$\left. + (X_j^-)^2 (Y_j^S)^2 + (Y_j^-)^2 (X_j^S)^2 + (X_j^S Y_j^S)^2 \right) + o_p(1)$$

All centred addends have a variance tending to zero as $N \to \infty$ and converge to zero in probability. The sum of the first four addends times the factor N/T has been proved to converge in probability to $\int_0^T F'(t) (\sigma_t^X \sigma_t^Y)^2 dt$ in Lemma A.7 where this term has appeared in the sequence of conditional variances of the error due to non-synchronicity.

Hence, it remains to prove that $N \sum_{j=1}^{N-1} (X_j^S Y_j^S)^2 \xrightarrow{p} T \int_0^T (2\rho_t^2 + 1)(\sigma_t^X \sigma_t^Y)^2 G'(t) dt$. For this purpose recall the notation from the proof of Proposition A.1. With $L_t = \int_0^t \sigma_s^X dW_s^X$, $M_t = \int_0^t \sigma_s^Y dW_s^Y$, $L_i = L_{T_i}, M_i = M_{T_i}$, we can write the term

$$N \sum_{j=1}^{N-1} \left((L - L_{i-1})_{T_i} (M - M_{i-1})_{T_i} \right)^2 = N \sum_{i=1}^{N-1} \left(2 \int_0^{T_i} (L - L_{i-1})_t (M - M_{i-1})_t^2 d(L - L_{i-1})_t \right.$$
$$+ 2 \int_0^{T_i} (L - L_{i-1})_t^2 (M - M_{i-1})_t d(M - M_{i-1})_t$$
$$+ 4 \int_0^{T_i} (L - L_{i-1})_t (M - M_{i-1})_t d \left[M - M_{i-1}, L - L_{i-1} \right]_t$$
$$\left. + \int_0^{T_i} (M - M_{i-1})_t^2 d \left[M - M_{i-1} \right]_t + \int_0^{T_i} (L - L_{i-1})_t^2 d \left[L - L_{i-1} \right]_t \right) \ ,$$

where we have applied Itô's formula. The sum of the first two addends converges to zero in probability since it is centred and the variance converges to zero. Since

$$\int_0^{T_i} (L - L_{i-1})_t (M - M_{i-1})_t d \left[M - M_{i-1}, L - L_{i-1} \right]_t = \int_{T_{i-1}}^{T_i} (L - L_{i-1})_t (M - M_{i-1})_t d \left[M, L \right]_t \ ,$$

the sum of the third addends has been considered in the proof of Proposition A.2 as part of the quadratic variation of the discretization error of the closest synchronous approximation and converges in probability to $2 T \int_0^T G'(t) (\rho_t \sigma_t^X \sigma_t^Y)^2 dt$. The remaining sum of the fourth addends is also similar to the other part of the quadratic variation in the proof of Proposition A.2. An analogous approximation and integration by

parts yields

$$\int_{T_{i-1}}^{T_i} (M - M_{i-1})_t^2 d[M]_t + \int_{T_{i-1}}^{T_i} (L - L_{i-1})_t^2 d[L]_t$$

$$= \int_{T_{i-1}}^{T_i} [M - M_{i-1}]_t \, d[M]_t + \int_{T_{i-1}}^{T_i} [L - L_{i-1}]_t \, d[L]_t + \mathcal{O}_p(1)$$

$$= \int_{T_{i-1}}^{T_i} d([L - L_{i-1}]_t [M - M_{i-1}]_t) + \mathcal{O}_p(1)$$

and the convergence of the above given term to $T \int_0^T (\sigma_t^X \sigma_t^Y)^2 G'(t) dt$.
In the second part of the proof we are concerned with the term

$$2N \sum_{j=1}^{N-1} (X_j^+ + X_j^S + X_j^-)(X_{j+1}^+ + X_{j+1}^S + X_{j+1}^-)(Y_j^+ + Y_j^S + Y_j^-)(Y_{j+1}^+ + Y_{j+1}^S + Y_{j+1}^-)$$

$$= 2N \sum_{j=1}^{N-1} \left(X_j^S Y_j^S X_{j+1}^S Y_{j+1}^S + X_j^+ Y_j^S X_{j+1}^- Y_{j+1}^S + Y_j^+ X_j^S Y_{j+1}^- X_{j+1}^S \right) + \mathcal{O}_p(1).$$

The sum incorporating all centred addends converges to zero in probability. The last two addends capture the only dependence between consecutive addends in the error due to non-synchronicity (13), namely when next-tick interpolations and previous-tick interpolations at the same $T_i, i = 1, \ldots, N$ are included. Those have appeared in the proof of Lemma A.7 and have been proved to converge to $T \int_0^T 2H'(t)(\sigma_t^X \sigma_t^Y)^2 dt$ in probability. That $2N \sum(X_j^S Y_j^S X_{j+1}^S Y_{j+1}^S) \xrightarrow{p} 2 \int_0^T G'(t)(\rho_t \sigma_t^X \sigma_t^Y)^2 dt$ follows with the methodology from [5] and Lemma 1 from [23] using the concept of a time-change in the asymptotic quadratic variation of refresh times such that $\sum_i (\Delta T_i - T/N)^2 = \mathcal{O}(N^{-1})$ holds true. Using the mean value theorem and $(\Delta T_j)^2 - \Delta T_j \Delta T_{j+1} = \Delta T_j (\Delta T_j - T/N) + \Delta T_j (T/N - \Delta T_{j+1})$ together with the Cauchy-Schwarz inequality

$$N \left| \sum_{j=1}^{N-1} \left(\Delta T_j \left(\Delta T_j - \frac{T}{N} \right) \right) \right| \leq N \sqrt{\sum_{j=1}^{N-1} (\Delta T_j)^2} \sqrt{\sum_{j=1}^{N-1} \left(\Delta T_j - \frac{T}{N} \right)^2}$$

yields the result.
The Hayashi-Yoshida estimators on the bins in the histogram-based estimator (19) fulfill

$$\widehat{\Delta[X,Y]}_{G_j^N}^{(HY)} = \int_{G_{j-1}^N}^{G_j^N} \rho_t \sigma_t^X \sigma_t^Y \, dt + \mathcal{O}_p \left(K_N^{1/2} N^{-1/2} \right)$$

so that the estimation error of the sum is of order $K_N^{3/4} N^{-1/2}$ in probability and for $K_N \to \infty$, $N \to \infty$, $K_N N^{-2/3} \to 0$ consistency holds and we conclude consistency of the estimator of the asymptotic variance. $\square$

References

[1] D. Aldous, G. Eagleson, On mixing and stability of limit theorems, Annals of Probability 6 (1978) 325–331.

[2] O.E. Barndorff-Nielsen, P.R. Hansen, A. Lunde, N. Shephard, Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading, SSRN working paper 1154144, University of Aarhus (2008).

[3] O.E. Barndorff-Nielsen, N. Shephard, Econometric analysis of realized volatility and its use in estimating stochastic volatility models, Journal of the Royal Statistical Society 64 (2002) 253–280.

[4] M. Bibinger, Efficient covariance estimation for asynchronous noisy high-frequency data, Scandinavian Journal of Statistics 38 (2011) 23–45.

[5] M. Bibinger, An estimator for the quadratic covariation of asynchronously observed itô processes with noise: Asymptotic distribution theory, preprint version, Humboldt-Universität zu Berlin (2011) URL=http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2011-034.pdf.

[6] P. Billingsley, Probability and Measure, Springer, New York, 2 edition, 1991.

[7] D.R. Cox, V. Isham, Point processes, Monographs on applied probability and statistics series, Chapman and Hall, New York, 1980.

[8] T.W. Epps, Comovements in stock prices in the very short run, Journal of the American Statistical Association 74 (1979) 291–298.

[9] P.D. Feigin, Stable convergence of semimartingales, Stochastic Processes and their Applications 19 (1985) 125 – 134.

[10] M. Fukasawa, Realized volatility with stochastic sampling, Stochastic Processeses and their Applications 120 (2010) 209–233.

[11] L. de Haan, A. Ferreira, Extreme Value Theory: An Introduction, Springer, New York, 2006.

[12] P. Hall, C. Heyde, Martingale Limit Theory and its Application, Academic Press, Boston, 1980.

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

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

[15] T. Hayashi, N. Yoshida, Nonsynchronous covariation process and limit theorems, Stochastic Processes and their Applications In Press, Uncorrected Proof (2011).

[16] J. Jacod, On continuous conditional gaussian martingales and stable convergence in law, Séminaire de Probabilitiés (1997) 232–246.

[17] J. Jacod, A.N. Shiryaev, Limit Theorems for Stochastic Processes, Springer, New York, 2003.

[18] P. Mykland, L. Zhang, Inference for continuous semimartingales observed at high frequency, Econometrica 77 (2009) 1403–1445.

[19] A. Palandri, Consistent Realized Covariance for Asynchronous Observations Contaminated by Market Microstructure Noise, Technical Report, University of Copenhagen, 2006.

[20] M. Podolskij, M. Vetter, Understanding limit theorems for semimartingales: a short survey, Statistica Nederlandica 64 (2010) 329–351.

[21] A. Rényi, On stable sequences of events, Sankhya: The Indian Journal of Statistics, Series A 25 (1963) 293–302.

[22] H. van Zanten, A multivariate central limit theorem for continuous local martingales, Statistics and Probability Letters 50 (2000) 229 – 235.

[23] L. Zhang, Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach, Bernoulli 12 (2006) 1019–1043.