SFB 649 Discussion Paper 2011-034

An estimator for the quadratic covariation of asynchronously observed Itô processes with noise: Asymptotic distribution theory

Markus Bibinger*

* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

An estimator for the quadratic covariation of asynchronously observed Itô processes with noise: Asymptotic distribution theory

Markus Bibinger[1]

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

Abstract

The article is devoted to the nonparametric estimation of the quadratic covariation of non-synchronously observed Itô processes in an additive microstructure noise model. In a high-frequency setting, we aim at establishing an asymptotic distribution theory for a generalized multiscale estimator including a feasible central limit theorem with optimal convergence rate on convenient regularity assumptions. The inevitably remaining impact of asynchronous deterministic sampling schemes and noise corruption on the asymptotic distribution is precisely elucidated. A case study for various important examples, several generalizations of the model and an algorithm for the implementation warrant the utility of the estimation method in applications.

Keywords: non-synchronous observations, microstructure noise, integrated covolatility, multiscale estimator, stable limit theorem

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

JEL Classification: C14, C32, C58, G10

1. Introduction

The nonparametric estimation of the univariate quadratic variation of a latent semimartingale from n observations in a high-frequency setting with additive observation noise has been studied intensively in recent years. It is known from [13] that $n^{1/4}$ constitutes a lower bound for the rate of convergence. An important motivation which has stimulated an alliance of economists and statisticians to establish estimation techniques for this kind of latent semimartingale models is their utility in estimating daily integrated (co-)volatilities from high-frequency intraday returns that serve as a basis for risk management as well as portfolio optimization and hedging strategies. The last years have seen an enormous increase of the amount of trading activities for many liquid securities. Paradoxically, the availability of high-frequency data necessitated a new angle on financial modeling. In fact, for every semimartingale the discrete realized (co-)volatilities converge in probability to the integrated measures. However, realized volatilities of typical high-frequency financial time series data explode for very high frequencies. This effect is ascribed to market microstructure frictions. Sources of the market microstructure noise are manifold. One important role plays the occurrence of bid-ask spreads. Aside from that transaction costs, strategic trading, limited market depths and discreteness of prices spread out the structure of the long-run dynamics that can be characterized by semimartingales.

This strand of literature followed [31] that has attracted a lot of attention to this estimation problem. The so-called two-scales realized volatility by [31] is based on subsampling and a bias-correction and a stable central limit theorem with $n^{1/6}$-rate has been proved. A refinement of the subsample approach using multiple scales in [30] and related alternative techniques in [4], [23] and [29] have led to rate-optimal estimators and feasible stable central limit theorems. For the more specific nonparametric model with

[1]Financial support from the Deutsche Forschungsgemeinschaft via SFB 649 'Ökonomisches Risiko', Humboldt-Universität zu Berlin, is gratefully acknowledged.

Gaussian noise, asymptotic equivalence in the Le Cam sense to a Gaussian shift experiment is shown in [26] and an asymptotically efficient estimator whose asymptotic variance equals the parametric efficiency bound is constructed.
In the article on hand we are concerned with a multivariate stetting and apart from taking additive microstructure noise into account, we focus on a way to deal with non-synchronous observation schemes. This is also a central theme in financial applications. When realized covolatilities are calculated for fixed time distances and a previous-tick interpolation is applied, the phenomenon of the so-called Epps effect described in [10] appears that the realized covolatility tends to zero at the highest frequencies.
A methodology to deal with non-synchronous observations in a bivariate Itô processes model has been proposed by [15]. The so-called Hayashi-Yoshida estimator has superseded simpler previous-tick interpolation methods setting the standard for the estimation of the quadratic covariation from asynchronous observations in the absence of microstructure noise effects.
Our estimation approach, first proposed in [8], for the most general case in the presence of noise and non-synchronicity arises as a combination of the multiscale estimator to handle noise contamination on the one hand and a synchronization algorithm in accordance with the Hayashi-Yoshida estimator to cope with non-synchronicity on the other hand. A first attempt in the same direction, combining one-scale subsampling and the Hayashi-Yoshida estimator, has been given in [22].
In [8] it has been shown in the spirit of [13] that the optimal convergence rate $n^{1/4}$ carries over to the general multidimensional setup. The mathematical analysis of our generalized multiscale estimator in [8] shows that it is rate-optimal.
Alternative approaches to similar statistical models has been suggested by [5], [9] and [1]. In [5] a kernel-based method with a previous-tick interpolation to so-called refresh times is proposed and a stable central limit theorem with sub-optimal $n^{1/5}$-rate is established for a multivariate non-synchronous design. This estimator, furthermore, ensures that the estimated covariance matrix is positive semi-definite. [9] and [1] come up with combinations of pre-averaging ([23],[19]) and the Hayashi-Yoshida estimator and of the univariate quasi-maximum-likelihood method by [29], the polarization identity and a generalized synchronization scheme which is different from the Hayashi-Yoshida ansatz that we use, respectively, both also attaining the optimal rate.
In this article we aim at providing an asymptotic distribution theory for the generalized multiscale estimator. In distinction from alternative methods, the influence of non-synchronicity effects on the expectation is null and on the variance limited up to an interaction of interpolation steps and microstructure noise. The main result is a feasible stable central limit theorem for its estimation error with optimal rate and a closed-form asymptotic variance that does not hinge on interpolation errors in the signal term. The stable weak convergence of the estimation error to a centred mixed Gaussian limit and the consistent estimation of the random unknown asymptotic variance are the essential steps towards statistical inference and confidence sets. The theory is grounded on stable limit theorems for semimartingales from [18].
In Section 2 we present the model and our main findings. Section 3 comes up with a concise overview on the construction of the estimator and in Section 4 we develop the asymptotic theory. In Section 5 we propose a consistent estimator for the asymptotic variance and Section 6 comprises various extensions and and a concluding discussion. The proofs are postponed to the Appendix.

2. Model and key result

The considered statistical model of noisy latently observed Itô processes at deterministic observation times is precisely described by Assumptions 1-3 in this section.

Assumption 1 (efficient processes). *On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, the efficient processes $X = (X_t)_{t\in\mathbb{R}^+}$ and $Y = (Y_t)_{t\in\mathbb{R}^+}$ are Itô processes 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.

Assumption 2 (**observations**). *The deterministic observation schemes $\mathcal{T}^{X,n} = \{0 \leq t_0^{(n)} < t_1^{(n)} < \ldots < t_n^{(n)} \leq T\}$ of X and $\mathcal{T}^{Y,m} = \{0 \leq \tau_0^{(m)} < \tau_1^{(m)} < \ldots < \tau_m^{(m)} \leq T\}$ of Y are assumed to be regular in the following sense: There exists a constant $0 < \alpha \leq 1/9$ such that*

$$\delta_n^X = \sup_{i\in\{1,\ldots,n\}} \left(\left(t_i^{(n)} - t_{i-1}^{(n)}\right), t_0^{(n)}, T - t_n^{(n)}\right) \quad = \mathcal{O}\left(n^{-8/9-\alpha}\right) , \tag{1a}$$

$$\delta_m^Y = \sup_{j\in\{1,\ldots,m\}} \left(\left(\tau_j^{(m)} - \tau_{j-1}^{(m)}\right), \tau_0^{(m)}, T - \tau_m^{(m)}\right) = \mathcal{O}\left(m^{-8/9-\alpha}\right) . \tag{1b}$$

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 $n \sim m$. The efficient processes X and Y which satisfy Assumption 1 are discretely observed at the times $\mathcal{T}^{X,n}$ and $\mathcal{T}^{Y,m}$ with additive observation noise:

$$\tilde{X}_{t_i^{(n)}} = \int_0^{t_i^{(n)}} \mu_t^X\,dt + \int_0^{t_i^{(n)}} \sigma_t^X\,dB_t^X + \epsilon^X_{t_i^{(n)}} , 0 \leq i \leq n , \tag{2a}$$

$$\tilde{Y}_{\tau_j^{(m)}} = \int_0^{\tau_j^{(m)}} \mu_t^Y\,dt + \int_0^{\tau_j^{(m)}} \sigma_t^Y\,dB_t^Y + \epsilon^Y_{\tau_j^{(m)}} , 0 \leq j \leq m . \tag{2b}$$

Although we consider sequences of deterministic observation times, the case of random sampling that is independent of the observed processes is included when regarding the conditional law.
It turns out that it is accurate to prove the key result of the article on the following i. i. d. assumption on the microstructure noise since a closed-form expression for the asymptotic variance is not available for a combination of general asynchronous observation schemes and serially dependent observation errors. Since an extension to non-i. i. d. noise is crucial for the utility in financial applications, we comment on the robustness of our estimator to that case in Section 6.

Assumption 3 (**microstructure noise**). *The discrete microstructure noise processes*

$$\epsilon^X_{t_i^{(n)}}, \epsilon^Y_{\tau_j^{(m)}}, 0 \leq i \leq n, 0 \leq j \leq m .$$

are centred i. i. d., independent of each other and independent of the efficient processes X and Y. We assume that the observation errors have finite fourth moments and denote the variances

$$\eta_X^2 = \mathbb{V}\mathrm{ar}\left(\epsilon^X_{t_1^{(n)}}\right) , \eta_Y^2 = \mathbb{V}\mathrm{ar}\left(\epsilon^Y_{\tau_1^{(m)}}\right) .$$

The number of synchronized observations $N \sim n \sim m$ which appears in the rate of our feasible stable central limit theorem is introduced in Section 3.

Theorem 1 (**feasible stable central limit theorem**). *The generalized multiscale estimator* (12) *specified by the later given weights* (A.1)*, with $M_N = c_{multi} \cdot \sqrt{N}$ converges on the Assumptions 1, 2, 3 and further mild regularity conditions on the asymptotics of the sampling schemes, stated below in Assumptions 4 and 5, $\mathcal{F}-$stably in law with optimal rate $N^{1/4} \sim n^{1/4} \sim m^{1/4}$ to a mixed Gaussian limiting distribution:*

$$N^{1/4}\left(\widehat{[X,Y]}_T^{multi} - [X,Y]_T\right) \overset{st}{\rightsquigarrow} \mathbf{N}\left(0, \mathbf{AVAR}_{multi}\right)$$

with an almost surely finite random asymptotic variance given in (18) *in Theorem 3. With the consistent estimator for the asymptotic variance $\widehat{\mathbf{AVAR}}_{multi}$ in Proposition 5.1 , the feasible central limit theorem*

$$N^{1/4}\frac{\left(\widehat{[X,Y]}_T^{multi} - [X,Y]_T\right)}{\widehat{\mathbf{AVAR}}_{multi}} \overset{st}{\rightsquigarrow} \mathbf{N}(0,1) , \tag{3}$$

holds true.

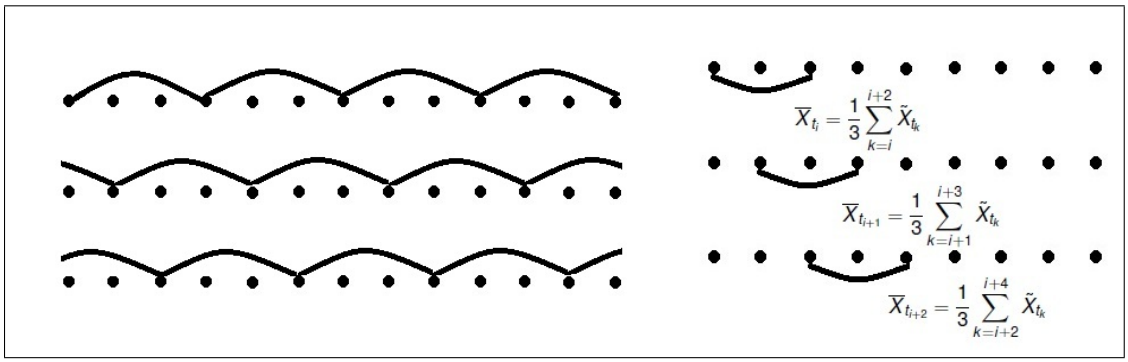


Figure 1: Sketch of the subsampling approach.

The notion of stable weak convergence going back to [27] is essential for our asymptotic theory. Stable weak convergence $X_n \overset{st}{\rightsquigarrow} X$ is the joint weak convergence of (X_n, Z) to (X, Z) for every measurable bounded random variable Z. The limiting random variables in stable limit theorems are defined on extensions of the original underlying probability spaces. The reason for us to involve this concept of a stronger mode of weak convergence is that mixed normal limiting distributions are derived where asymptotic variances are themselves strictly positive random variables. Provided we have a consistent estimator V_n^2 for such a random asymptotic variance V^2 on hand, the stable central limit theorem $X_n \overset{st}{\rightsquigarrow} VZ$ with Z distributed according to a standard Gaussian law, yields the joint weak convergence $(X_n, V_n^2) \rightsquigarrow (VZ, V^2)$ and also $X_n/V_n \rightsquigarrow Z$ and hence allows to perform statistical inference providing tests or confidence intervals.
In the proofs of our limit theorems we will 'remove' the drifts in the sense that after a transformation to an equivalent martingale measure stable central limit theorems for Itô processes without drift are proved and, as illustrated in [21], stability of the weak convergence ensures that the asymptotic law holds true under the original measure. In this sense stable convergence is commutative with measure change.
From now on, we often omit the superscripts of observation times for a shorter notation.

3. Brief review on the foundation

3.1. Subsampling and the multiscale estimator

In the model imposed by Assumption 1, Assumption 2 with synchronous observations, $n = m$ and $t_j^{(n)} = \tau_j^{(n)}, 1 \leq j \leq n$, and Assumption 3, the realized (co-)volatilities do not provide consistent estimators for the quadratic (co-)variations any more. The variance due to noise conditional on the paths of the efficient processes

$$\mathbb{V}\mathrm{ar}_{X,Y}\left(\sum_{j=1}^{n}\left(\tilde{X}_{t_j} - \tilde{X}_{t_{j-1}}\right)\left(\tilde{Y}_{t_j} - \tilde{Y}_{t_{j-1}}\right)\right) = 4n\,\eta_X^2\eta_Y^2\ ,$$

increases linearly with n. The error due to noise perturbation can be reduced by the following estimator, which has been proposed for the univariate estimation of integrated volatility as the "second best approach" in [31] and which is called one-scale subsampling estimator in [8] and throughout this article. It can be motivated from two perspectives that are both sketched in Figure 1. On the left-hand side we have visualized that one can calculate simultaneously lower frequent realized covolatilities using subsamples, e. g. to the lag three in Figure 1, and (post-)average them.

$$\widehat{[X,Y]}_T^{sub} = \frac{1}{i}\sum_{j=i}^{n}\left(\tilde{X}_{t_j} - \tilde{X}_{t_{j-i}}\right)\left(\tilde{Y}_{t_j} - \tilde{Y}_{t_{j-i}}\right)\ . \tag{4a}$$

This motivation given in [31] is in line with the former common practice of a sparse-sampled low-frequency realized (co-)volatility estimator and proposes to use an average instead of one single lower frequent realized measure.
The same estimator arises as the usual realized covolatility calculated from the time series on that a linear filter is run before, what means that non-noisy observations at a time t_j are estimated with a (pre-)average of noisy observations at times $t_j, \ldots, t_{j+i}$ for some i. This is sketched on the right-hand side of Figure 1 for $i = 3$. Passing over to increments leads to telescoping sums and we end up finally with the one-scale subsampling estimator.
Since on the Assumption 3 there is no bias due to noise for the bivariate estimator, it already corresponds to the "first best approach" from [31] whereas in the univariate case a bias-correction completes the two scales realized volatility (TSRV):

$$\widehat{[X]}_T^{TSRV} = \frac{1}{i}\sum_{j=i}^{n}\left(\tilde{X}_{t_j} - \tilde{X}_{t_{j-i}}\right)^2 - \frac{1}{2n}\sum_{i=1}^{n}\left(\tilde{X}_{t_j} - \tilde{X}_{t_{j-1}}\right)^2 . \tag{4b}$$

There is a trade-off between the signal term and the error due to noise. Choosing $i = c_{sub}n^{2/3}$ dependent on n with a constant c_{sub}, the overall mean square error is minimized and of order $n^{-1/3}$. The one-scale subsampling estimator (4a) is hence a consistent and asymptotically unbiased estimator. The rate of convergence $n^{1/6}$, however, is slow and does not attain the optimal rate $n^{1/4}$ determined in [8]. For this reason, we focus on a multiscale extension of the subsampling approach on which the methods developed in [8] are based on. The multiscale realized covolatility (MSRC), and the univariate multiscale realized volatility (MSRV) introduced in [30], are linear combinations of one-scale subsampling estimators with M_n different subsampling frequencies $i = 1, \ldots, M_n$:

$$\widehat{[X,Y]}_T^{multi} = \sum_{i=1}^{M_n}\frac{\alpha_{i,M_n}^{opt}}{i}\sum_{j=i}^{n}\left(\tilde{X}_{t_j} - \tilde{X}_{t_{j-i}}\right)\left(\tilde{Y}_{t_j} - \tilde{Y}_{t_{j-i}}\right) , \tag{5a}$$

$$\widehat{[X]}_T^{multi} = \sum_{i=1}^{M_n}\frac{\alpha_{i,M_n}^{opt}}{i}\sum_{j=i}^{n}\left(\tilde{X}_{t_j} - \tilde{X}_{t_{j-i}}\right)^2 . \tag{5b}$$

The weights are chosen such that the estimator is asymptotically unbiased and the error due to noise minimized. They are given later in (A.1) and can be chosen equally for the bivariate and the univariate case. Those are the standard discrete weights of [30] and we abstain from giving a more general class of possible weight functions.
The mean square error of the multiscale realized covolatility (5a) can be split in uncorrelated addends that stem from discretization, microstructure noise and cross terms and end-effects. They are of orders M_n/n, n/M_n^3, and M_n^{-1}, respectively. Hence, a choice $M_n = c_{multi}\sqrt{n}$ leads to a rate-optimal $n^{1/4}$-consistent estimator.
The following stable central limit theorems for the multiscale realized covolatility (5a) and the one-scale estimator (4a) are implied by Theorem 3 and Corollary 4.1:

Proposition 3.1. *On Assumptions 1, 2 and 3 in the synchronous setup and if* $(n/T)\sum_i (t_i^{(n)} - t_{i-1}^{(n)})^2$ *converges to a continuously differentiable limiting function* G *and the difference quotients converge uniformly to* G' *on* $[0,T]$*, the multiscale realized covolatility* (5a) *and the subsampling estimator* (4a) *converge stably in law to mixed normal limiting random variables:*

$$n^{1/4}\left(\widehat{[X,Y]}_T^{multi} - [X,Y]_T\right) \overset{st}{\rightsquigarrow} \mathbf{N}\left(0\,,\,\mathbf{AVAR}_{multi,syn}\right) , \tag{6a}$$

$$n^{1/6}\left(\widehat{[X,Y]}_T^{sub} - [X,Y]_T\right) \overset{st}{\rightsquigarrow} \mathbf{N}\left(0\,,\,\mathbf{AVAR}_{sub,syn}\right) , \tag{6b}$$

with

$$\mathbf{AVAR}_{multi,syn} = c_{multi}^{-3} 24\eta_X^2\eta_Y^2 + c_{multi}\frac{26}{35}T\int_0^T G'(t)(\rho_t^2+1)(\sigma_t^X\sigma_t^Y)^2\,dt + c_{multi}^{-1}\frac{12}{5}\left(\eta_X^2\eta_Y^2 + \eta_X^2\int_0^T(\sigma_t^Y)^2\,dt + \eta_Y^2\int_0^T(\sigma_t^X)^2\,dt\right), \tag{6c}$$

$$\mathbf{AVAR}_{sub,syn} = c_{sub}^{-2} 4\eta_X^2\eta_Y^2 + c_{sub}\frac{2}{3}T\int_0^T G'(t)(\rho_t^2+1)(\sigma_t^X\sigma_t^Y)^2\,dt\,. \tag{6d}$$

3.2. Synchronization and the Hayashi–Yoshida estimator

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. The Hayashi-Yoshida estimator

$$\widehat{[X,Y]}_T^{(HY)} = \sum_{i=1}^n\sum_{j=1}^m \Delta X_{t_i}\Delta Y_{\tau_j}\mathbb{1}_{[\min(t_i,\tau_j)>\max(t_{i-1},\tau_{j-1})]}\,, \tag{7}$$

where the product terms include all increments of the processes with overlapping observation time instants, has been proved in [15] to be consistent in a model of asynchronously observed Itô processes with deterministic correlation, drift and volatility functions in the absence of observation noise and on further regularity conditions to be asymptotically normally distributed in [16].

For a combination of the strategy of the Hayashi-Yoshida estimator with techniques to handle noise contamination, we use an iterative algorithm introduced in [22] as 'pseudo-aggregation'. Incorporating telescoping sums there are the following rewritings of the estimator (7):

$$\begin{aligned}\widehat{[X,Y]}_T^{(HY)} &= \sum_{i=1}^n \Delta X_{t_i}\left(Y_{t_{i,+}} - Y_{t_{i-1,-}}\right)\\ &= \sum_{i=1}^N \left(X_{g_i} - X_{l_i}\right)\left(Y_{\gamma_i} - Y_{\lambda_i}\right)\\ &= \sum_{i=1}^N \left(X_{T_{i,+}^X} - X_{T_{i-1,-}^X}\right)\left(Y_{T_{i,+}^Y} - Y_{T_{i-1,-}^Y}\right),\end{aligned} \tag{8}$$

with the notion of next-tick interpolated times $t_{i,+} := \min_{0\le j\le m}(\tau_j|\tau_j \ge t_i)$ and previous-tick interpolated ones $t_{i,-} := \max_{0\le j\le m}(\tau_j|\tau_j \le t_i)$ in the first equality. This rewriting can be as well done in the symmetric way.

The illustration of (8) that serves as a basis for the construction of the generalized multiscale estimator relies on an aggregation of the observations according to Algorithm 1. This algorithm, which is a concise version of the construction in [8], stops after $(N+1) \le \min(n,m)+1$ steps when all observation times are grouped. Summation in (8) can start with $i = 0$ or $i = 1$. In the last equality only the denotation expressions $g_i, \gamma_i, l_i, \lambda_i$ are substituted emphasizing that those sampling times obtained by Algorithm 1 can be interpreted as previous- and next-tick interpolations again with respect to a synchronous sampling scheme $T_k := \min(g_k, \gamma_k)$, $1 \le k \le N$, which we call the closest synchronous approximation. Increments in (8) are taken from previous-tick interpolations at left-end points of instants $[T_{k-1}, T_k]$, $2 \le k \le N$ to next-tick interpolated sampling times at right-end points. Since $T_k = \max(l_{k+1}, \lambda_{k+1})$, $1 \le k \le (N-1)$ holds true, we split the estimation error of (8) in two uncorrelated parts $D_T^N + A_T^N$ with

$$D_T^N := \sum_{i=1}^N\left(\left(X_{T_i} - X_{T_{i-1}}\right)\left(Y_{T_i} - Y_{T_{i-1}}\right) - \int_{T_{i-1}}^{T_i}\rho_t\sigma_t^X\sigma_t^Y\,dt\right) - \int_0^{t_0\wedge\tau_0}\rho_t\sigma_t^X\sigma_t^Y\,dt - \int_{t_n\wedge\tau_m}^T\rho_t\sigma_t^X\sigma_t^Y\,dt \tag{9}$$

Define $t^+(s) := \min\{t_i \in \mathcal{T}^{X,n} | t_i \geq s\}$, $\tau^+(s) := \min\{\tau_j \in \mathcal{T}^{Y,n} | \tau_j \geq s\}$;
$t^-(s) := \max\{t_i \in \mathcal{T}^{X,n} | t_i < s\}$, $\tau^-(s) := \max\{\tau_j \in \mathcal{T}^{Y,n} | \tau_j < s\}$.

first step:

- For $t_0 \leq \tau_0$ $\dashrightarrow$ $g_0 = t^+(\tau_0), l_0 = t_0, \gamma_0 = \tau_0, \lambda_0 = \tau_0$.
- For $t_0 > \tau_0$ $\dashrightarrow$ $g_0 = t_0, l_0 = t_0, \gamma_0 = \tau^+(t_0), \lambda_0 = \tau_0$.

ith step (given g_{i-1} and γ_{i-1}):

- If $g_{i-1} = \gamma_{i-1}$
  - and $t^+(g_{i-1}) \leq \tau^+(\gamma_{i-1})$ $\dashrightarrow$ $g_i = t^+(\tau^+(\gamma_{i-1})), l_i = g_{i-1}, \gamma_i = \tau^+(\gamma_{i-1}), \lambda_i = \gamma_{i-1}$.
  - and $t^+(g_{i-1}) > \tau^+(\gamma_{i-1})$ $\dashrightarrow$ $g_i = t^+(g_{i-1}), l_i = g_{i-1}, \gamma_i = \tau^+(t^+(g_{i-1})), \lambda_i = \gamma_{i-1}$.
- If $g_{i-1} < \gamma_{i-1}$
  - and $\gamma_{i-1} \leq t^+(g_{i-1})$ $\dashrightarrow$ $g_i = t^+(g_{i-1}), l_i = g_{i-1}, \gamma_i = \tau^+(t^+(g_{i-1})), \lambda_i = \tau^-(\gamma_{i-1})$.
  - and $\gamma_{i-1} > t^+(g_{i-1})$ $\dashrightarrow$ $g_i = t^+(\gamma_{i-1}), l_i = g_{i-1}, \gamma_i = \gamma_{i-1}, \lambda_i = \tau^-(\gamma_{i-1})$.
- If $g_{i-1} > \gamma_{i-1}$: symmetrically as for $g_{i-1} < \gamma_{i-1}$.

Algorithm 1: Iterative synchronization algorithm.

being the discretization error of a synchronous-type realized covolatility including in general non-observable idealized values at the times of the closest synchronous approximation and A_T^N an additional error due to the lack of synchronicity, in particular next- and previous-tick interpolations. The times T_k equal the so-called refresh times of [5] and thus our synchronization differs from the one in [5] by replacing pure previous-tick interpolation by the above given machinery of previous- and next-tick interpolations.
The asymptotic theory for the estimator (8) as $N \to \infty$, concisely repeated here, is separately proved and presented in a more elaborate way in [7].
First, we take up the illustrative example from [8] to motivate the synchronization procedure. For further details and examples we refer to [8] and [22]. Figure 2 visualizes the aggregation carried out by Algorithm 1 and the times T_i, $i = 0, \ldots, 8$ for a toy example. The example emphasizes the important fact that consecutive right-end points of increments can be the same time points. The realized covolatility calculated from previous-tick interpolated values to refresh times equals

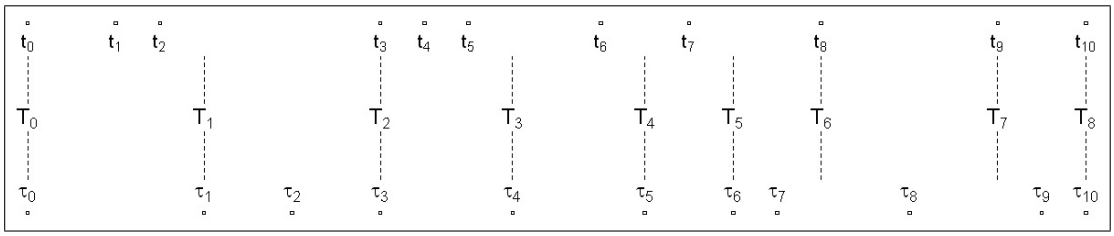
$$\begin{aligned}(X_{t_2} - X_{t_0})(Y_{\tau_1} - Y_{\tau_0}) + (X_{t_3} - X_{t_2})(Y_{\tau_3} - Y_{\tau_1}) + (X_{t_5} - X_{t_3})(Y_{\tau_4} - Y_{\tau_3}) + \\ (X_{t_6} - X_{t_5})(Y_{\tau_5} - Y_{\tau_4}) + (X_{t_7} - X_{t_6})(Y_{\tau_6} - Y_{\tau_5}) + (X_{t_8} - X_{t_7})(Y_{\tau_7} - Y_{\tau_6}) +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

and is systematically biased downwards by interpolations, whereas (8) yields

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

which is not biased due to interpolations.

Definition 1 (quadratic (co-)variations of time). For any $N \in \mathbb{N}$ let $T_i^{(N)}$, $i = 0, \ldots, N$ be the times of the closest synchronous approximation and $g_i^{(N)}, \gamma_i^{(N)}, l_i^{(N)}, \lambda_i^{(N)}$ the corresponding observation times that appear in the estimator (8) defined above by Algorithm 1 . T/N is the mean of the time instants



i	0	1	2	3	4	5	6	7	8
$\mathcal{H}^i$	$\{t_0\}$	$\{t_1,t_2,t_3\}$	$\{t_3\}$	$\{t_4,t_5,t_6\}$	$\{t_6,t_7\}$	$\{t_7,t_8\}$	$\{t_8\}$	$\{t_9\}$	$\{t_{10}\}$
$\mathcal{G}^i$	$\{\tau_0\}$	$\{\tau_1\}$	$\{\tau_2,\tau_3\}$	$\{\tau_4\}$	$\{\tau_5\}$	$\{\tau_6\}$	$\{\tau_7,\tau_8\}$	$\{\tau_8,\tau_9\}$	$\{\tau_9,\tau_{10}\}$
l_i	t_0	t_0	t_2	t_3	t_5	t_6	t_7	t_8	t_9
g_i	t_0	t_3	t_3	t_6	t_7	t_8	t_8	t_9	t_{10}
λ_i	τ_0	τ_0	τ_1	τ_3	τ_4	τ_5	τ_6	τ_7	τ_8
γ_i	τ_0	τ_1	τ_3	τ_4	τ_5	τ_6	τ_8	τ_9	τ_{10}
T_i	t_0=τ_0	τ_1	t_3=τ_3	τ_4	τ_5	τ_6	t_8	t_9	t_{10}=τ_{10}

Figure 2: Example for non-synchronous sampling design with the sets constructed by the synchronization algorithm, interpolated observation times occurring in the estimators and the synchronous approximation.

$\Delta T_i^{(N)} = T_i^{(N)} - T_{i-1}^{(N)},\ i = 1,\dots,N$. Define the following sequences of functions

$$G^N(t) = \frac{N}{T} \sum_{T_i^{(N)} \le t} \left(\Delta T_i^{(N)}\right)^2 , \tag{10a}$$

$$F^N(t) = \frac{N}{T} \sum_{T_{i+1}^{(N)} \le t} (T_i^{(N)} - \lambda_i^{(N)})(g_i^{(N)} - T_i^{(N)}) + \left(T_i^{(N)} - l_i^{(N)}\right)\left(\gamma_i^{(N)} - T_i^{(N)}\right) + \Delta T_{i+1}^{(N)}\left(T_i^{(N)} - l_{i+1}^{(N)}\right) + \Delta T_{i+1}^{(N)}\left(T_i^{(N)} - \lambda_{i+1}^{(N)}\right) , \tag{10b}$$

$$H^N(t) = \frac{N}{T} \sum_{T_{i+1}^{(N)} \le t} \left(T_i^{(N)} - l_{i+1}^{(N)}\right)\left(g_i^{(N)} - T_i^{(N)}\right) + \left(T_i^{(N)} - \lambda_{i+1}^{(N)}\right)\left(\gamma_i^{(N)} - T_i^{(N)}\right) , \tag{10c}$$

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 in [7] on the assumption that the sequences defined by (10a), (10b) and (10c) converge pointwise to continuous differentiable limiting functions G, F, H and the sequences of difference quotients uniformly. The asymptotic quadratic variation of time G of the $T_i^{(N)}$s influences the asymptotics of D_T^N. The covariation of times F^N measures an interaction of interpolation errors between the two processes and H^N the impact of the in general non-zero correlations of the products involving previous- and next-tick interpolations at the same $T_i^{(N)}$s for each process separately.

Consider as easiest example the synchronous equidistant sampling schemes with $N = n = m$ and $t_i^{(n)} = \tau_j^{(n)} = i/n, i = 0,\dots,n$. In this case F^N and H^N are identically zero since interpolations are redundant. The function G^N is a step function that will tend to the identity on $[0,T]$ as $N \to \infty$.

Then, consider a situation of completely non-synchronous sampling schemes that originates from the complete synchronous equidistant one by shifting one time-scale half a time instant $1/2N$. We will call this situation intermeshed sampling. For this example, 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.

In [7] we show for an important special case, independent homogeneous Poisson sampling, that the convergence assumptions on (10a)-(10c) are fulfilled when replacing deterministic convergence by convergence

in probability. Furthermore, the stochastic limits $G'(t), F'(t), H'(t)$ are calculated explicitly.
The main result for the estimator (8) is Theorem 2. It serves as preparation to prove the stable limit theorem for the generalized multiscale estimator in Theorem 3 and gives insight into the asymptotic distribution of (8). For the proof we refer to [7]. A similar stable limit theorem for the original Hayashi-Yoshida estimator is provided in [17].

Theorem 2. *The estimation error of* (8) *converges on the Assumptions 1, 2 and convergence assumptions on* (10a)-(10c) *and the difference quotients stably in law to a centred, mixed Gaussian distribution:*

$$\sqrt{N}\left(\sum_{i=1}^{N}\left(X_{g_i}-X_{l_i}\right)\left(Y_{\gamma_i}-Y_{\lambda_i}\right)-[X\,,\,Y]_T\right)\overset{st}{\rightsquigarrow}\mathbf{N}\left(0\,,\,v_{D_T}+v_{A_T}\right)\,, \tag{11}$$

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.

3.3. Hybrid approach to non-synchronous and noisy observations

In [8] we have proposed the following combined estimation method for the quadratic covariation or integrated covolatility from noisy asynchronous observations. After applying Algorithm 1 to the observation times, the generalized multiscale estimator is defined by

$$\widehat{[X,Y]}_T^{multi}=\sum_{i=1}^{M_N}\frac{\alpha_{i,M_N}^{opt}}{i}\sum_{j=i}^{N}\left(\tilde{X}_{g_j^{(N)}}-\tilde{X}_{l_{j-i+1}^{(N)}}\right)\left(\tilde{Y}_{\gamma_j^{(N)}}-\tilde{Y}_{\lambda_{j-i+1}^{(N)}}\right)\,. \tag{12}$$

It is a weighted sum of M_N one-scale subsampling estimators of the type

$$\widehat{[X,Y]}_T^{sub}=\frac{1}{i_N}\sum_{j=i_N}^{N}\left(\tilde{X}_{g_j^{(N)}}-\tilde{X}_{l_{j-i_N+1}^{(N)}}\right)\left(\tilde{Y}_{\gamma_j^{(N)}}-\tilde{Y}_{\lambda_{j-i_N+1}^{(N)}}\right) \tag{13}$$

with subsampling frequencies $i=1,\ldots,M_N$ and optimal weights given later in (A.1). Owing to the aggregation of non-synchronous observation times before applying subsampling and the multiscale approach, the resulting estimator has a conformable appearance as in the synchronous case (5a). Recall that in the synchronous setting $g_j=\gamma_j=T_j$ and $l_{j-i+1}=\lambda_{j-i+1}=T_{j-i}$ holds.
Choosing $M_N=c_{multi}\cdot\sqrt{N}$ and $i_N=c_{sub}\cdot N^{2/3}$, both estimators above provide consistent and asymptotically unbiased estimators with convergence rate $N^{1/4}$ and $N^{1/6}$, respectively.

4. Asymptotics and a stable central limit theorem for the generalized multiscale estimator

A comprehensive analysis of the asymptotic distribution of the estimation error necessitates an elaborate screening of the conjunction of Algorithm 1 and the joint sampling design $\left(\mathcal{T}^{X,n},\mathcal{T}^{Y,m}\right)$.
Note that the generalized multiscale estimator (12) differs from the other plausible Hayashi-Yoshida version of a multiscale estimator

$$\sum_{i=1}^{M_N}\frac{\beta_{i,M_N}^{opt}}{i}\sum_{j=i}^{n}\sum_{k\in\mathbb{Z}}\left(\tilde{X}_{t_j^{(n)}}-\tilde{X}_{t_{j-i}^{(n)}}\right)\left(\tilde{Y}_{\tau_{j+k\cdot i}^{(m)}}-\tilde{Y}_{\tau_{j+(k-1)\cdot i}^{(m)}}\right)\mathbb{1}_{\{\max\left(t_{j-i}^{(n)},\tau_{j+(k-1)\cdot i}^{(m)}\right)<\min\left(t_j^{(n)},\tau_{j+k\cdot i}^{(m)}\right)\}}\,, \tag{14}$$

which arises as natural Hayashi-Yoshida multiscale estimator when, on the basis of (non-synchronized) observations of $\tilde{X}$ and $\tilde{Y}$, sparse-sample Hayashi-Yoshida estimators are averaged to one-scale subsample estimators and those extended to a linear combination using different time lags. We state without proof that this estimator is consistent, asymptotically unbiased and will attain the optimal rate of convergence. Nevertheless, we benefit from the data aggregation method and applying subsampling to the synchronized

i	1	2	3	4	5	6	7	8
case X	②	①	③	③	②	①	①	①
case Y	①	①	①	①	①	③	④	①
relations X	$g_1 = g_2 = l_3$	$g_2 = l_3$	$g_3 = l_5$	$g_4 = l_6$	$g_5 = l_7$	$g_6 = l_7$	$g_7 = l_8$	–
relations Y	$\gamma_1 = \lambda_2$	$\gamma_2 = \lambda_3$	$\gamma_3 = \lambda_4$	$\gamma_4 = \lambda_5$	$\gamma_5 = \lambda_6$	$\gamma_6 = \lambda_8$	–	–

Table 1: Allocation of sampling times to cases ① − ④ for the example.

scheme, since the variance of our estimator (12) is smaller than the one of this alternative estimator and we are able to find a feasible closed-form expression of the asymptotic variance.
The crucial difference between both approaches is that for the alternative method next- and previous-tick interpolation errors take place on sparse-sampling time intervals in average of order i/N whereas the interpolation errors of the generalized multiscale estimator (12) take place on the highest-frequency-scale and hence on intervals in average of order $1/N$. In particular the decomposition

$$\left(X_{g_j^{(N)}} - X_{l_{j-i_N+1}^{(N)}}\right) = \big(\underbrace{X_{g_j^{(N)}} - X_{T_j^{(N)}}}_{=\mathcal{O}_p(N^{-1/2})} + \underbrace{X_{T_j^{(N)}} - X_{T_{j-i}^{(N)}}}_{=\mathcal{O}_p((i/N)^{(1/2)})} + \underbrace{X_{T_{j-i}^{(N)}} - X_{l_{j-i_N+1}^{(N)}}}_{=\mathcal{O}_p(N^{-1/2})} \big)$$

of the increments of X and analogously for Y, give an heuristic that the interpolation errors driving the error due to non-synchronicity asymptotically not affect the variance of the signal term. The stochastic orders are given for time instants of average order N^{-1}.
For a rigorous clarification of the asymptotic error due to noise and the cross term, both influenced by the i. i. d. observation errors at times $g_i, l_i, \gamma_i, \lambda_i$, we figure out the times $g_i = g_{i+1}$ and the right-end points $g_i^{(N)} = l_{i+1}^{(N)}, g_i^{(N)} = l_{i+2}^{(N)}$ that are as well preceding left-end points and analogously for the sampling times of $\tilde{Y}$.
All observation times γ_i, λ_i are characterized through one of the following four mutually exclusive cases. Denote $\gamma_{j,-}$ the last observation time of $\tilde{Y}$ before γ_j and $\gamma_{j,+}$ the first one after γ_j. We illustrate the allocation of the observation times for $\mathcal{T}^{Y,m}$ and γ_j , $j = 1, \ldots, N-2$:

$$\begin{aligned}
&① \ \gamma_j \le g_j && \Rightarrow \gamma_j \ne \gamma_{j+1} \ , \ \gamma_j = \lambda_{j+1} \ , \ \gamma_j \ne \lambda_{j+2} \ , \\
&② \ \gamma_j > g_j \ , \ \gamma_j \ge g_{j,+} && \Rightarrow \gamma_j = \gamma_{j+1} \ , \ \gamma_j \ne \lambda_{j+1} \ , \ \gamma_j = \lambda_{j+2} \ , \\
&③ \ \gamma_j > g_j \ , \ \gamma_j < g_{j,+} \ , \ \gamma_{j,+} > g_{j,+} && \Rightarrow \gamma_j \ne \gamma_{j+1} \ , \ \gamma_j \ne \lambda_{j+1} \ , \ \gamma_j = \lambda_{j+2} \ , \\
&④ \ \gamma_j > g_j \ , \ \gamma_j < g_{j,+} \ , \ \gamma_{j,+} \le g_{j,+} && \Rightarrow \gamma_j \ne \gamma_{j+1} \ , \ \gamma_j \ne \lambda_{j+1} \ , \ \gamma_j \ne \lambda_{j+2} \ , \ \gamma_{j,+} = \lambda_{j+2} \ .
\end{aligned}$$

Only sampling times distributed to case ② lead to repeated $\gamma_i = \gamma_{i+1}$. In cases ①, ② and ③ a subsequent left-end point $\lambda_k, k = i+1$ or $k = i+2$ of observation time instants incorporated in the subsampling estimators is designated by γ_i. All other λ_k, $k = 2, \ldots, N$ appear in an allocation of sampling times of the type ④, where $\lambda_{j+2} = \gamma_{j,+} \ne \gamma_l \, \forall l$. Recall that $\lambda_i \ne \lambda_k$ for all $i \ne k$ holds true.
If ② holds for γ_j with fixed $j \in \{1, \ldots, N-2\}$ and if $k := \arg\min_{k \in \{j, \ldots, N-1\}} (\gamma_k > g_k \ , \ \gamma_k \ge g_{k,+})$ exists, then ② holds necessarily for one $g_l, l \in \{j+1, \ldots, k-1\}$ or $g_l = \gamma_l$.
In Table 1 we list the relations for the sampling design of our previous example.

Assumption 4 (asymptotic quadratic variation of time). *Assume that for the sequences of sampling schemes and the times $T_i^{(N)}$ of the closest synchronous approximations and for the sequence of quadratic variations of time $G^N(t)$ defined in Definition 1, the following holds true:*

(i) $G^N(t) \to G(t)$ as $N \to \infty$, where $G(t)$ is a continuously differentiable function on $[0, T]$.
(ii) For any null sequence (h_N), $h_N = \mathcal{O}\left(N^{-1}\right)$

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

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

(iii) The derivative $G'(t)$ is bounded away from zero.

Definition 2 (degree of regularity of asynchronicity). For $N \in \mathbb{N}$ and times $g_i^{(N)}, \gamma_i^{(N)}, i = 0, \dots, N$ constructed from aggregated sampling schemes $\mathcal{T}^{X,n}, \mathcal{T}^{Y,m}$ that fulfill Assumption 2, define the following sequences of functions:

$$I_X^N(t) = \frac{1}{N} \sum_{g_j^{(N)} \leq t} \mathbb{1}_{\{g_j^{(N)} = g_{j-1}^{(N)}\}} , \tag{16a}$$

$$I_Y^N(t) = \frac{1}{N} \sum_{\gamma_j^{(N)} \leq t} \mathbb{1}_{\{\gamma_j^{(N)} = \gamma_{j-1}^{(N)}\}} , \tag{16b}$$

which describe the degree of regularity of asynchronicity between observation times $\mathcal{T}^{X,n}$ and $\mathcal{T}^{Y,m}$.

In the completely asynchronous case, we can directly conclude that $|I_X^N(t) - I_Y^N(t)| \leq T/N$ for all $t \in [0, T]$ and one sequence suffices to reflect the regularity of the non-synchronous sampling schemes.

Assumption 5 (asymptotic degree of regularity of asynchronicity). *Assume that for the sequences of sampling schemes and for the sequences of functions I_X^N, I_Y^N defined in Definition 2, the following holds true:*

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

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

$$\frac{I_X^N(t + h_N) - I_X^N(t)}{h_N} \to I_X'(t) , \tag{17a}$$

$$\frac{I_Y^N(t + h_N) - I_Y^N(t)}{h_N} \to I_Y'(t) \tag{17b}$$

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

For both, synchronous and intermeshed sampling which have been introduced in the last section, the sequences of functions I_X^N, I_Y^N are identically zero. The functions defined in Definition 2 are non-negative and bounded above by 1. In Section 6 we explicitly deduce the asymptotic degree of regularity of asynchronicity for mutually independent homogeneous Poisson sampling schemes. The term (asymptotic) degree of regularity of asynchronicity has been chosen since Assumption 5 holds for all non-degenerate sequences where observation times conforming to one of the cases ① − ④ from above tend to be distributed according to some regular pattern and it gives information on the interaction of allocations of observation times.

It is interesting and might seem surprising at first glance that the asymptotics of the estimator (12) hinges on this asymptotic feature whereas, as indicated before, the asymptotic interpolations to the closest synchronous approximation are asymptotically immaterial. This circumstance is caused by the fact that for the construction of an estimator with Algorithm 1, as for the original Hayashi-Yoshida estimator (8), observed values of the processes at next-tick interpolated observation times can appear twice. If there is observation noise, the number of observations allocated conforming to case ② has an impact on the asymptotics. The influence of interpolations is asymptotically vanishing for the combined method in contrast to the estimator (8) with faster convergence rate $\sqrt{N}$ since interpolation steps take place on the time-scale of high-frequency observations, but lower-frequency sparse-sampled increments of the synchronous approximation are involved to reduce the error due to noise. We continue with the central result of this article:

Theorem 3 (Central limit theorem for the generalized multiscale estimator). *On the Assumptions 1, 2, 3, 4 and 5, the generalized multiscale estimator* (12) *with noise-optimal weights* $\alpha_{i,M_N}^{opt} = (12i^2/M_N^3) - (6i/M_N^2)\,(1+\mathcal{o}(1))$*, that are explicitly given in* (A.1)*, and* $M_N = c_{multi} \cdot \sqrt{N}$ *converges* $\mathcal{F}-$*stably in law with optimal rate* $N^{1/4}$ *to a mixed Gaussian limiting distribution:*

$$N^{1/4}\left(\widehat{[X,Y]}_T^{multi} - [X,Y]_T\right) \overset{st}{\rightsquigarrow} \mathbf{N}\left(0, \mathbf{AVAR}_{multi}\right)$$

with the asymptotic variance

$$\begin{aligned}\mathbf{AVAR}_{multi} = c_{multi}^{-3} \underbrace{\left(24 + 12\left(I_X(T) + I_Y(T)\right)\right)\eta_X^2\eta_Y^2}_{=\mathbf{AVAR}_{noise}} + c_{multi}^{-1}\,\frac{12\eta_X^2\eta_Y^2}{5} \\ + c_{multi}\,\underbrace{\frac{26}{35}T\int_0^T G'(t)(\sigma_t^X\sigma_t^Y)^2(1+\rho_t^2)\,dt}_{=\mathbf{AVAR}_{dis}} \\ + c_{multi}^{-1}\,\underbrace{\frac{12}{5}\left(\eta_Y^2\int_0^T (1+T\,I_Y'(t))(\sigma_t^X)^2\,dt + \eta_X^2\int_0^T (1+T\,I_X'(t))(\sigma_t^Y)^2\,dt\right)}_{=\mathbf{AVAR}_{cross}}.\end{aligned} \tag{18}$$

The weak convergence is proved to be stable with respect to the σ-algebra $\mathcal{F}$ associated with the efficient processes. As a side result, we also obtain a stable central limit theorem for a simpler one-scale subsampling estimator:

Corollary 4.1 (Central limit theorem for the one-scale subsampling estimator). *On the Assumptions 1, 2, 3 and 4, the one-scale subsampling estimator with subsampling frequency* $i_N = c_{sub} \cdot N^{2/3}$ *converges* $\mathcal{F}$*-stably in law with rate* $N^{1/6}$ *to a mixed Gaussian limiting distribution:*

$$N^{1/6}\left(\widehat{[X,Y]}_T^{sub} - [X,Y]_T\right) \overset{st}{\rightsquigarrow} \mathbf{N}\left(0, \mathbf{AVAR}_{sub}\right)\ , \tag{19}$$

with the asymptotic variance

$$\mathbf{AVAR}_{sub} = c_{sub}^{-2} \underbrace{4\eta_X^2\eta_Y^2}_{=\mathbf{AVAR}_{noise,sub}} + c_{sub}\,\underbrace{\frac{2}{3}T\int_0^T G'(t)(\sigma_t^X\sigma_t^Y)^2(1+\rho_t^2)\,dt}_{=\mathbf{AVAR}_{dis,sub}}\ . \tag{20}$$

For the proof of Theorem 3, we split the total estimation error of the generalized multiscale estimator in three asymptotically uncorrelated addends due to noise, cross terms and the signal term. For the one-scale subsampling estimator we follow the same ansatz. The orders of the errors have been derived in [8] and we focus on the asymptotic distribution here.
The error due to microstructure noise of the one-scale subsampling estimator has expectation zero and the variance yields

$$i_N^{-2}\sum_{j=i_N}^{N}\mathbb{E}\left[\left(\epsilon_{g_j}^X - \epsilon_{l_{j-i_N+1}}^X\right)^2\left(\epsilon_{\gamma_j}^Y - \epsilon_{\lambda_{j-i_N+1}}^Y\right)^2\right] = 4Ni_N^{-2}\eta_X^2\eta_Y^2 + \mathcal{o}\left(Ni_N^{-2}\right)\ ,$$

since observation noises of $\tilde{X}$ and $\tilde{Y}$ are independent of each other by Assumption 3 and $l_k \neq l_r$ for $k \neq r$, $\lambda_k \neq \lambda_r$ for $k \neq r$ and if $g_k = g_{k+1} \Rightarrow \gamma_k < \gamma_{k+1}$, $0 \leq k \leq (N_1)$. Hence, the error due to noise is a sum of uncorrelated centred random variables with equal variances and the standard central limit theorem applies.
For the generalized multiscale estimator, we further decompose the error due to noise in a main part of order $N^{1/2}M_N^{-3/2}$ and two terms due to end-effects of orders $M_N^{-1/2}$, where all three terms are asymptotically

uncorrelated. In Propositions A.1 and A.2 we prove central limit theorems for these terms. Asymptotic normality holds both, conditionally and unconditionally on the paths of the efficient processes.
The error due to noise of the one-scale estimator does not depend on any further influence of the sampling schemes except the number of constructed sets N and G'. Cross terms are, in contrast to the multiscale case, asymptotically negligible since

$$\begin{aligned}
&\mathbb{E}\left[\left(i_N^{-1}\sum_{j=i_N}^{N}\left(\left(X_{g_j}-X_{l_{j-i_N+1}}\right)\left(\epsilon^Y_{\gamma_j}-\epsilon^Y_{\lambda_{j-i_N+1}}\right)+\left(Y_{\gamma_j}-Y_{\lambda_{j-i_N+1}}\right)\left(\epsilon^X_{g_j}-\epsilon^X_{l_{j-i_N+1}}\right)\right)\right)^2\right]\\
&=i_N^{-2}\sum_{j=i_N}^{N}\left(\int_{l_{j-i_N+1}}^{g_j}(\sigma_t^X)^2\,dt\;2\eta_Y^2+\int_{\lambda_{j-i_N+1}}^{\gamma_j}(\sigma_t^Y)^2\,dt\;2\eta_X^2\right)+\mathcal{o}\left(i_N^{-2}\right)\\
&=\mathcal{O}\left(i_N^{-2}\right)=o(1)\,.
\end{aligned}$$

For the generalized multiscale estimator instead the cross terms are of order $M_N^{-1/2}$ and will have effect upon the asymptotic distribution. In Proposition A.11 a limit theorem is stated where the weak convergence also holds conditionally and unconditionally on the paths of the efficient processes. The asymptotic variance $\mathbf{AVAR}_{cross}$ includes the influence of the asymptotic degree of regularity of asynchronicity.
The error due to discretization of the one-scale subsampling estimator yields:

$$\begin{aligned}
&\frac{1}{i}\sum_{j=i}^{N}\left(X_{g_j}-X_{l_{j-i+1}}\right)\left(Y_{\gamma_j}-Y_{\lambda_{j-i+1}}\right)-[X,Y]_T\\
&=\frac{1}{i}\sum_{j=i}^{N}\left(X_{T_j}-X_{T_{j-i}}\right)\left(Y_{T_j}-Y_{T_{j-i}}\right)-[X,Y]_T\\
&+\frac{1}{i}\sum_{j=i}^{N}\left[\left(X_{g_j}-X_{T_j}\right)\left(Y_{T_j}-Y_{T_{j-i}}\right)+\left(X_{g_j}-X_{T_j}\right)\left(Y_{T_{j-i}}-Y_{\lambda_{j-i+1}}\right)\right.\\
&\quad+\left(X_{T_{j-i}}-X_{l_{j-i+1}}\right)\left(Y_{\gamma_j}-Y_{T_j}\right)+\left(X_{T_{j-i}}-X_{l_{j-i+1}}\right)\left(Y_{T_j}-Y_{T_{j-i}}\right)\\
&\quad\left.+\left(Y_{\gamma_j}-Y_{T_j}\right)\left(X_{T_j}-X_{T_{j-i}}\right)+\left(Y_{T_{j-i}}-Y_{\lambda_{j-i+1}}\right)\left(X_{T_j}-X_{T_{j-i}}\right)\right]\\
&=\frac{1}{i}\sum_{j=i}^{N}\Big(\sum_{k=j-i+1}^{j}\Delta X_{T_k}\Big)\Big(\sum_{k=j-i+1}^{j}\Delta Y_{T_k}\Big)-[X,Y]_T\\
&+\frac{1}{i}\sum_{j=i}^{N-1}\left[\left(X_{g_j}-X_{T_j}\right)\Big(\sum_{k=j-i+1}^{j}\Delta Y_{T_k}\Big)+\left(X_{g_j}-X_{T_j}\right)\left(Y_{T_{j-i}}-Y_{\lambda_{j-i+1}}\right)\right.\\
&\quad+\left(X_{T_{j-i}}-X_{l_{j-i+1}}\right)\left(Y_{\gamma_j}-Y_{T_j}\right)+\left(X_{T_j}-X_{l_{j+1}}\right)\Big(\sum_{k=j-i+1}^{j}\Delta Y_{T_k}\Big)\\
&\quad\left.+\left(Y_{\gamma_j}-Y_{T_j}\right)\Big(\sum_{k=j-i+1}^{j}\Delta X_{T_k}\Big)+\left(Y_{T_j}-Y_{\lambda_{j+1}}\right)\Big(\sum_{k=j-i+1}^{j}\Delta X_{T_k}\Big)\right]+\mathcal{O}_p\left(N^{-1}\right)\\
&=\frac{1}{i}\sum_{j=i}^{N}\Big(\sum_{k=j-i+1}^{j}\Delta X_{T_k}\Delta Y_{T_k}+\sum_{\substack{l\neq r\\ l,r\in\{j-i+1,\ldots,j\}}}\Delta X_{T_l}\Delta Y_{T_r}\Big)-[X,Y]_T\\
&+\frac{1}{i}\sum_{j=i}^{N-1}\left[\left(X_{g_j}-X_{T_j}\right)\Big(\sum_{k=j-i+1}^{j}\Delta Y_{T_k}\Big)+\left(Y_{\gamma_j}-Y_{T_j}\right)\Big(\sum_{k=j-i+1}^{j}\Delta X_{T_k}\Big)\right.\\
&\quad\left.+\Delta X_{T_{j+1}}\Big(\sum_{k=j-i+1}^{j}\left(Y_{T_k}-Y_{\lambda_{k+1}}\right)\Big)+\Delta Y_{T_{j+1}}\Big(\sum_{k=j-i+1}^{j}\left(X_{T_k}-X_{l_{k+1}}\right)\Big)\right]
\end{aligned}$$

$$+ \mathcal{O}_p\left(i^{-1}N^{-\frac{1}{2}}\right) + \mathcal{O}_p\left(N^{-1}\right) .$$

We have written the overall discretization error of the one-scale estimator as the sum of a discretization error of the closest synchronous approximation

$$\sum_{j=1}^{N}\left(\Delta X_{T_j}\sum_{l=1}^{i\wedge j}\left(1-\frac{l}{i}\right)\Delta Y_{T_{j-l}} + \Delta Y_{T_j}\sum_{l=1}^{i\wedge j}\left(1-\frac{l}{i}\right)\Delta X_{T_{j-l}}\right) + \mathcal{O}_p\left(iN^{-1}\right) + \mathcal{O}_p\left(N^{-1/2}\right) \quad (21)$$

and the asymptotically negligible error due to the lack of synchronicity. A stable central limit theorem using the theory of [18] for the leading term of order $i^{1/2}N^{-1/2}$ that will drive the asymptotic distribution is postponed to Proposition A.4. The error due to asynchronicity is treated in Proposition A.10.
The discretization error of the generalized multiscale estimator is of order $M_N^{1/2}N^{-1/2}$ and that of the one-scale estimator of order $i_N^{1/2}N^{-1/2}$. There is a trade-off between the error due to noise and the discretization error for both estimators. For the generalized multiscale estimator these are of orders $N^{1/2}M_N^{-3/2}$ and $M_N^{1/2}N^{-1/2}$, respectively. Remaining other terms are of orders $M_N^{-1/2}$. Thus, choosing $M_N = c_{multi}\cdot N^{1/2}$, the total estimation error is minimized and of order $M_N^{-1/2} = N^{-1/4}$ which constitutes the optimal rate of convergence in Theorem 3.
The weak convergence of the discretization error is proved to be stable, so it converges jointly in law with every bounded $\mathcal{F}$-measurable random variable defined on the same probability space. Since the asymptotic normality of the cross term and the error due to noise holds both, conditionally and unconditionally given the efficient processes, and the discretization error is independent of ϵ^X and ϵ^Y we can apply a central limit theorem for mixing triangular arrays as in [28] to the sum that is adapted with respect to $\mathcal{A}_j = \sigma\left(\epsilon^X_{t_k}, t_k < T_{j+1}, \epsilon^Y_{\tau_k}, \tau_k < T_{j+1}, \mathcal{F}_{T_j}\right)$ where $\mathcal{F}$ is the σ-algebra associated with the efficient processes. The asymptotic variance is the sum of those of the uncorrelated addends. With the Cramér-Wold device joint normality and asymptotic independence of the different errors can be concluded.
This is likewise for the one-scale estimator and Corollary 4.1. Choosing the subsampling frequency $i_N = c_{sub}\cdot N^{2/3}$ balances the variance of the error due to noise which is of order Ni^{-2} and the discretization variance of order iN^{-1}.

5. Asymptotic variance estimation

The asymptotic variances (18) and (20) of the generalized multiscale estimator (12) and the one-scale subsampling estimator (13), appearing in the stable central limit theorems in Theorem 3 and Corollary 4.1, are random and depend on unknown quantities. In this section, we aim at estimating these asymptotic variances consistently to make our limit theorems feasible.
It is a known result that a consistent estimator of the noise variance is given by (24) (cf. [31]). Furthermore, the estimators for η_X^2 and η_Y^2 are asymptotically uncorrelated on Assumption 3, since the uncorrelated noise terms dominate the correlated Brownian parts. The constant $I_X(T) + I_Y(T)$ in the noise part of (18) can be estimated with the empirical version $I_X^N(T) + I_Y^N(T)$ that converges as $N \to \infty$ on Assumption 5. Eventually, consistent estimators for the discretization variances and the variance due to cross terms for the multiscale estimator are required.
We propose histogram-type estimators using bins according to timescales associated with the quadratic variation of synchronized sampling times and associated with the degree of regularity of asynchronicity, respectively. For this purpose, given a chosen number of bins K_N, with $K_N \to \infty$ and $K_N^{-1}N \to \infty$ as $N \to \infty$, we determine the assigned non-equispaced bin-widths $\Delta G_j^N = G_j^N - G_{j-1}^N$, $\Delta(I_X)_j^N = (I_X)_j^N - (I_X)_{j-1}^N$ and $\Delta(I_Y)_j^N = (I_Y)_j^N - (I_Y)_{j-1}^N$, $j \in \{1,\dots,K_N\}$, where

$$G_j^N := \inf\left\{t \in [0,T] \,\middle|\, G^N(t) = (N/T)\sum_{T_k^{(N)}\le t}\left(\Delta T_k^{(N)}\right)^2 \ge (j/K_N)\cdot G^N(T)\right\},$$

$j \in \{1,\dots,K_N\}$, and analogously for the functions I_X^N and I_Y^N if $I_X^N(T) > 0$ and $I_Y^N(T) > 0$. Set $G_0^N = (I_X)_0^N = (I_Y)_0^N := 0$ and recall that those functions are monotone increasing on $[0,T]$. On each

bin we calculate multiscale estimators in the same spirit as (12) and its univariate version from [30] for the increase of the quadratic (co-) variations that are denoted $\widehat{\Delta\,[X]}_{G_j^N}$, $\widehat{\Delta\,[Y]}_{G_j^N}$, $\widehat{\Delta\,[X,Y]}_{G_j^N}$, $\widehat{\Delta\,[X]}_{(I_Y)_j^N}$ and $\widehat{\Delta\,[Y]}_{(I_X)_j^N}$ in the following. The underlain idea is to approximate the continuous random processes $(\sigma_t^X\sigma_t^Y\rho_t)^2$, $(\sigma_t^X)^2$ and $(\sigma_t^Y)^2$, or rather their time-transformed versions, by locally constant functions. This construction leads to time-adjusted histogram estimators

$$\hat{I}_1 = \sum_{j=1}^{K_N}\left(\frac{\widehat{\Delta\,[X,Y]}_{G_j^N}}{\Delta G_j^N}\right)^2\frac{G^N(T)}{K_N} \qquad \text{for} \quad \int_0^T G'(t)(\sigma_t^X\sigma_t^Y\rho_t)^2\,dt, \tag{22a}$$

$$\hat{I}_2 = \sum_{j=1}^{K_N}\left(\frac{\widehat{\Delta\,[X]}_{G_j^N}\widehat{\Delta\,[Y]}_{G_j^N}}{\left(\Delta G_j^N\right)^2}\right)\frac{G^N(T)}{K_N} \qquad \text{for} \quad \int_0^T G'(t)(\sigma_t^X\sigma_t^Y)^2\,dt, \tag{22b}$$

$$\hat{I}_3 = \sum_{j=1}^{K_N}\left(\frac{\widehat{\Delta\,[X]}_{(I_Y)_j^N}}{\Delta(I_Y)_j^N}\right)\frac{I_Y^N(T)}{K_N} \qquad \text{for} \quad \int_0^T I_Y'(t)(\sigma_t^X)^2\,dt, \tag{22c}$$

$$\hat{I}_4 = \sum_{j=1}^{K_N}\left(\frac{\widehat{\Delta\,[Y]}_{(I_X)_j^N}}{\Delta(I_X)_j^N}\right)\frac{I_X^N(T)}{K_N} \qquad \text{for} \quad \int_0^T I_X'(t)(\sigma_t^Y)^2\,dt. \tag{22d}$$

Proposition 5.1. *The asymptotic variances* (18) *and* (20) *of the generalized multiscale estimator* (12) *and the one-scale subsampling estimator* (13) *with* $M_N = c_{multi}N^{1/2}$ *and* $i_N = c_{sub}N^{2/3}$*, can be estimated consistently by*

$$\begin{aligned}\widehat{\mathbf{AVAR}}_{multi} = &\left(c_{multi}^{-3}\left(24 + 12\,\left(I_X^N(T) + I_Y^N(T)\right)\right) + \frac{12}{5}c_{multi}^{-1}\right)\widehat{\eta_X^2}\widehat{\eta_Y^2}\\ &+ c_{multi}\frac{26}{35}T\left(\hat{I}_1 + \hat{I}_2\right) + c_{multi}^{-1}\frac{12}{5}\left(\widehat{\eta_Y^2}(1 + T\hat{I}_3) + \widehat{\eta_X^2}(1 + T\hat{I}_4)\right)\,,\end{aligned} \tag{23a}$$

$$\widehat{\mathbf{AVAR}}_{sub} = c_{sub}^{-2}4\widehat{\eta_X^2}\widehat{\eta_Y^2} + c_{sub}\frac{2}{3}\left(\hat{I}_1 + \hat{I}_2\right)\,, \tag{23b}$$

where $\hat{I}_1$-$\hat{I}_4$ *are the estimators* (22a)-(22d) *and*

$$\widehat{\eta_X^2} = (2n)^{-1}\sum_{i=1}^{n}(\Delta X_{t_i})^2\,,\ \widehat{\eta_Y^2} = (2m)^{-1}\sum_{j=1}^{m}(\Delta Y_{\tau_j})^2\,. \tag{24}$$

Remark 1. Convergence rates of the estimators (23a) and (23b) for the asymptotic variances depend on the smoothness of σ^X, σ^Y and ρ. For current stochastic volatility models as the Heston model, they are $N^{1/5}$-consistent when choosing $K_N = c_K N^{1/5}$ for a constant c_K and $M_N \sim N^{3/5}$ for the binwise multiscale estimators.

In the absence of noise, a consistent estimator for the asymptotic variance $2T\int_0^T G'_X(t)(\sigma_t^X)^4dt$ of the realized volatility has been proposed in [6] as $(2n/3)\sum_{i=1}^{n}(\Delta X_{t_i})^4$. In the bivariate synchronous setting $(n/2)\sum_{i=1}^{n-1}(\Delta X_{t_i})^2\left((\Delta Y_{t_i})^2 + (\Delta Y_{t_{i+1}})^2\right)$ is a convenient estimator. Consistency can be proved with Itô's formula and partial integration and comprehended by the analogy to a bivariate Gaussian distribution $(X,Y) \sim \mathbf{N}(0,\Sigma)$ with a covariance matrix Σ with entries $\sigma_X^2, \sigma_Y^2, \rho\sigma_X\sigma_Y$. Then, $\mathbb{E}X^4 = 3\sigma_X^4$ and $\mathbb{E}\left[X^2Y^2\right] = 2\rho^2\sigma_X^2\sigma_Y^2 + \sigma_X^2\sigma_Y^2$ hold true.

In the noisy case smoothed versions of the estimators (using multiscale or alternative methods) are adequate (cf. [9]). However, in the non-synchronous non-noisy setting, there is no direct extension available and for that reason we have made the effort to construct the consistent histogram-based estimators (22a)-(22d) above.

6. Discussion and application

6.1. A case study

We have learned by now that in a synchronous setting the special version of the central limit Theorem 3 from Proposition 3.1 holds true. Since asymptotics of the estimator (12) not hinge on interpolations in the signal term the same central limit theorem applies in the case of intermeshed sampling introduced in Section 3. Now we focus on observation schemes that arise as realizations of two homogeneous Poisson processes that are mutually independent and independent of the processes $\tilde{X}$ and $\tilde{Y}$. Although this model can be criticized for its flaw that sampling schemes of two correlated processes follow two independent processes and time homogeneity, what might seem to be unrealistic in financial applications, independent and homogeneous Poisson sampling constitutes the most commonly used model in this research area (cf. [30], [15] among others) and appertains to show that the general form of (18) is tractable.
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 $\mathbb{E}\left[\Delta \tau_j^{(n)}\right] = \theta_2/n$, $i \in \mathbb{N}, j \in \mathbb{N}$. In this case

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

$$I_X^N(t) \xrightarrow{p} \frac{\theta_1\theta_2 t}{(\theta_1+\theta_2)^2}\ ,\ I_Y^N(t) \xrightarrow{p} \frac{\theta_1\theta_2 t}{(\theta_1+\theta_2)^2} \quad \left(= \frac{z\,t}{(z+1)^2} \ \text{ if } \ \theta_1 = z\theta_2\right)\ , \tag{25}$$

hold true and we derive the following Poisson sampling version of Theorem 3:

Corollary 6.1. *On the Assumptions 1 and 3, the generalized multiscale estimator* (12) *with noise-optimal weights* (A.1)*, and $M_N = c_{multi} \cdot \sqrt{N}$, converges conditionally on the independent Poisson sampling scheme with $0 < \theta_1 < \infty$ and $0 < \theta_2 < \infty$ stably in law with rate $N^{1/4}$ to a mixed normal limit:*

$$N^{1/4}\left(\widehat{[X,Y]}_T^{multi} - [X,Y]_T\right) \overset{st}{\rightsquigarrow} \mathbf{N}\left(0, \mathbf{AVAR}_{multi}^{poiss}\right)$$

with the asymptotic variance

$$\begin{aligned}\mathbf{AVAR}_{multi}^{poiss} = {} & c_{multi}^{-3}\left(24 + 12\,\frac{2\theta_1\theta_2}{(\theta_1+\theta_2)^2}\right)\eta_X^2\eta_Y^2 + c_{multi}^{-1}\,\frac{12\eta_X^2\eta_Y^2}{5} \\ & + c_{multi}\,\frac{26}{35}\int_0^T 2\left(1 - \frac{2\theta_1^2\theta_2^2}{\theta_1^2\theta_2^2 + (\theta_1^2+\theta_2^2)(\theta_1+\theta_2)^2}\right)(\sigma_t^X\sigma_t^Y)^2(1+\rho_t^2)\,dt \\ & + c_{multi}^{-1}\,\frac{12}{5}\left(\eta_Y^2\int_0^T (1 + \frac{\theta_1\theta_2}{\theta_1+\theta_2})(\sigma_t^X)^2\,dt + \eta_X^2\int_0^T (1 + \frac{\theta_1\theta_2}{\theta_1+\theta_2})(\sigma_t^Y)^2\,dt\right).\end{aligned} \tag{26}$$

The asymptotic variance of the $N^{1/6}$-consistent one-scale estimator becomes

$$\begin{aligned}\mathbf{AVAR}_{sub}^{poiss} = {} & c_{sub}^{-2}\,4\eta_X^2\eta_Y^2 \\ & + c_{sub}\,\frac{2}{3}\int_0^T 2\left(1 - \frac{2\theta_1^2\theta_2^2}{\theta_1^2\theta_2^2 + (\theta_1^2+\theta_2^2)(\theta_1+\theta_2)^2}\right)(\sigma_t^X\sigma_t^Y)^2(1+\rho_t^2)\,dt\,.\end{aligned} \tag{27}$$

The order for the supremum of time instants in Assumption 2 holds in probability and the proof in Appendix A stays valid. A Poisson sampling version of Theorem 2 is given in [7], where the stochastic limit of G^N is deduced using the distribution of a maximum of two exponentials stated above.

6.2. A bridge between the noisy and the non-noisy setup

So far we considered noise variances not dependent on N, but from an applied point of view there is interest in the case where the noise level may vary with $N \sim n \sim m$. The primary motivation to accommodate dependence of the noise on the sample size in the model originates from the economic background. Empirical studies of (ultra) high-frequency financial data suggest to rather model the observed log-prices as sum of a latent semimartingale and noise for that the variance decreases in N as reported in [20] and [3], among others.

If an estimation approach uses previous-tick interpolations, as the one proposed in [5], these methods are not accurate for that setting any more. This also becomes apparent in the simulation study in [8] when the performance of the estimators is compared for varying noise levels. The generalized multiscale estimator is not biased due to asynchronicity and passes over to the Hayashi-Yoshida estimator (8) for $M_N = 1$, a $\sqrt{N}$-consistent estimator in the complete absence of noise. For that reason, our estimation method achieves an improved convergence rate in the model with decreasing noise variances. The next Corollary is obtained by a direct extension of the proof of Theorem 3 in Appendix A when replacing the moments of the noise processes. A similar extension for the one-scale estimator where we obtain the rate $N^{\frac{1}{6}+\frac{\alpha}{3}}$ for a subsample frequency $i_N = c_{sub} N^{\frac{2}{3}(1-\alpha)}$ holds analogously.

Corollary 6.2. *Consider the model imposed by Assumption 1, 2 and 3, but with noise variances* $\eta_X^2(N) = \zeta_X N^{-\alpha}$, $\eta_Y^2(N) = \zeta_Y N^{-\alpha}$, $0 < \alpha < 1$ *and constants* $0 < \zeta_X < \infty$, $0 < \zeta_Y < \infty$. *The generalized multiscale estimator* (12) *with* $M_N = c_{multi} N^{\frac{1}{2}-\frac{\alpha}{2}}$ *and optimal weights* (A.1) *converges stably in law to a mixed Gaussian limit:*

$$N^{\frac{1}{4}+\frac{\alpha}{4}} \left(\widehat{[X,Y]}_T^{multi} - [X,Y]_T \right) \overset{st}{\rightsquigarrow} \mathbf{N}\left(0, \mathbf{AVAR}^*_{multi}\right) \tag{28}$$

with the asymptotic variance

$$\begin{aligned}
\mathbf{AVAR}^*_{multi} = {} & c_{multi}^{-3} \left(24 + 12 \, \frac{I_X(T) + I_Y(T)}{T} \right) \zeta_X \zeta_Y + c_{multi}^{-1} \, \frac{12 \zeta_X \zeta_Y}{5} \\
& + c_{multi} \frac{26}{35} T \int_0^T G'(t) (\sigma_t^X \sigma_t^Y)^2 (1 + \rho_t^2) \, dt \\
& + c_{multi}^{-1} \frac{12}{5} \left(\zeta_Y \int_0^T (1 + I'_Y(t)) (\sigma_t^X)^2 \, dt \, + \zeta_X \int_0^T (1 + I'_X(t)) (\sigma_t^Y)^2 \, dt \right) .
\end{aligned}$$

Incorporating a pure previous-tick interpolation strategy as in [5], one cannot gain an improved convergence rate in that setting due to the bias by non-synchronicity effects.

6.3. Application

6.3.1. Note on modeling assumptions

For an application to financial time series data the conditions imposed by Assumptions 1-3 seem to be restrictive and the model will not describe stylized facts of the data in an adequate way. In particular relaxing the i. i. d. assumption on the noise is important. On the other hand, the underlying model in the sections before is convenient to establish an asymptotic distribution theory and ascertains a closed-form expression for the asymptotic variance. Yet, the generalization for serially dependent observation errors as carried out in [2] for the one-dimensional case is possible for the generalized multiscale estimator (12) as well. On the assumption of stationary strong mixing noise processes the multiscale estimator remains consistent and rate-optimal without any further adjustment. The analysis for the synchronous case can be adopted from [2], but for the general non-synchronous noisy setting a closed-form expression of the asymptotic variance and a corresponding limit theorem is not available. Furthermore, the condition that noise processes are mutually independent can be relaxed if one wants to allow some correlation $\mathbb{E}\left[\epsilon_{t_i}^X \epsilon_{\tau_j}^Y\right] = \eta_{X,Y}^{i,j}$ for t_i and τ_j located near each other. In any case the generalized multiscale estimator remains asymptotically unbiased and $N^{1/4}$-consistent. This does not necessarily hold true for the one-scale estimator that would have to be bias-corrected as the TSRV estimator by [31] in the one-dimensional case.

More general efficient processes including jumps can be covered in the model when we combine the method to a two-stage approach as presented by [11] for the one-dimensional estimation approach. Considering semimartingales

$$X_t = \int_0^t \mu_s^X \, ds + \int_0^t \sigma_s^X \, dB_s^X + \sum_{l=1}^{J_t^X} L_l^X \ , \ Y_t = \int_0^t \mu_s^Y \, ds + \int_0^t \sigma_s^Y \, dB_s^Y + \sum_{l=1}^{J_t^Y} L_l^Y \ ,$$

with locally bounded drifts, continuous volatilities and counting processes J_t^X, J_t^Y counting the jumps of X and Y with jump sizes $L_l^X, l = 1, \ldots, J_t^X$ and $L_l^Y, l = 1, \ldots, J_t^Y$, respectively, the generalized multiscale estimator converges in probability to the total quadratic covariation

$$\int_0^T \rho_t \sigma_t^X \sigma_t^Y + \sum_{0 \le s \le t} \Delta X_s \Delta Y_s$$

where $\Delta X_s = X_s - X_{s,-}, \Delta Y_s = Y_s - Y_{s,-}$, and the second addend is the sum of the simultaneous co-jumps. When one is interested in disentangling the continuous part from the jumps, one convincing possibility following [11] is to use wavelet methods that locate the jumps in the sample paths, estimate jump sizes and afterwards use our estimation approach for the validated observations.
In conclusion the generalized multiscale estimator (12) is capable for usage in various applications. In [8], we have approved that the estimator performs well and has satisfying finite sample size features in simulations including serially dependent noise and typical stochastic volatility models.

6.3.2. Choice of tuning parameters

An implementation of the estimation approach requires first a rule to choose tuning parameters. We provide an accurate algorithm to implement the estimators and to obtain estimates for their asymptotic variances.
One plausible selection of the constants $c_{multi} = \sqrt{N}/M_N$ and $c_{sub} = N^{2/3}/i_N$ can be derived as solutions of the minimization problems of the asymptotic variances. This leads to formula (29) in Algorithm 2. The tactic of Algorithm 2 is the following: Evaluate a pilot estimate $\hat{c}_{multi}^{(p)}$ for c_{multi} as solution of formula (29) inserting a sparse-sampled estimator for the signal term. Then set up the estimation of the asymptotic variances involving the estimators (22a)-(22d). Take $M_N^b = c_{multi}^b \sqrt{NK_N}$ fixed for the multiscale estimators on all bins and set $M_N^b = c_{multi}^b \sqrt{N c_K N^{1/5}}$ where $c_K = K_N^{-1} N^{1/5}$. This selection is optimal for common volatility models. We obtain $c_{multi}^b c_K^{-1/2} = \hat{c}_{multi}^{(p)}$ and from the orders of the different errors of the histogram estimators $c_{multi}^b = c_K^{5/2}$. Hence, $c_{multi}^b = \left(c_{multi}^{(p)}\right)^{5/4}$ and $c_K = \sqrt{c_{multi}^{(p)}}$ is derived. Using estimators (22a)-(22d), we calculate estimates for the addends of the asymptotic variance and $\hat{c}_{multi}$ according to formula (29) again and $M_N = \lceil \hat{c}_{multi} \sqrt{N} \rceil$ is used for the final estimator. It turns out that this strategy it quite robust to the a priori chosen sparse-sample frequency that can be chosen under the impression of usual diagnostic tools as signature plots and acfs.

6.3.3. Simulation and data analysis check

As completion to the detailed simulation study in the supplementary material to [8] we investigate the performance of Algorithm 2 and the histogram-estimators (22a)-(22d) here. For this purpose we simulate from an simple Brownian motion model with zero drifts and constant volatilities $\sigma^X = \sigma^Y = 1$ and $\rho = 1/2$ with equal noise variances η^2. Sampling schemes are generated by independent time-homogeneous Poisson sampling with 30.000 expected observation for both processes on $[0, 1]$. Results of the estimates are listed in Table 2. In addition, our approach is tested in an application to EUREX future tick-data taken from a database provided by the Research Data Center (RDC) of the CRC 649 'Economic Risk' in Berlin. We aim at estimating integrated covolatilities between the four financial securities with the highest tick-frequencies in the database. These are the Euro-Bund Future (FGBL), that is based on a notional long-duration debt instrument issued by the Federal Republic of Germany, the Euro-Bobl Future (FGBM), a likewise medium-duration contract, and futures on the EURO STOXX 50 (FESX) and the German DAX (FDAX).

- Choose a priori L and calculate pilot estimator $\widehat{\mathbf{AVAR}}^{p}_{multi}$ with

$$\widehat{\mathbf{AVAR}}^{p}_{dis}=\frac{N}{2}\sum_{k=1}^{\lfloor N/L\rfloor}\left(\left(\tilde{X}_{g_{kL}}-\tilde{X}_{l_{(k-1)L+1}}\right)^2+\left(\tilde{X}_{g_{kL+2}}-\tilde{X}_{l_{(k-1)L+3}}\right)^2\right)\left(\tilde{Y}_{\gamma_{kL}}-\tilde{Y}_{\lambda_{(k-1)L+1}}\right)^2$$

and $\widehat{[X]}^{p}_{T}=\sum_{j=kL,k\geq 1}\left(\tilde{X}_{t_j}-\tilde{X}_{t_{j-L}}\right)^2$ and $\widehat{[Y]}^{p}_{T}$ analogously and $I_X(t)\equiv I^N_X(T)$ and $I_Y(t)\equiv I^N_Y(T)$. Calculate $\widehat{\eta^2_X}$ and $\widehat{\eta^2_Y}$ according to (24).

- Use pilot estimates to estimate optimal constant(s)

$$\hat{c}^{(p)}_{multi}=\left(\frac{-\widehat{\mathbf{AVAR}}^{p}_{cross,\mathrm{n}}+\sqrt{\left(\widehat{\mathbf{AVAR}}^{p}_{cross,\mathrm{n}}\right)^2+12\widehat{\mathbf{AVAR}}^{p}_{dis}\widehat{\mathbf{AVAR}}^{p}_{\mathrm{n}}}}{6\widehat{\mathbf{AVAR}}^{p}_{\mathrm{n}}}\right)^{-1/2} \tag{29}$$

and $\hat{c}^{(p)}_{sub}=\sqrt[3]{2\widehat{\mathbf{AVAR}}^{p}_{\mathrm{n},sub}/\widehat{\mathbf{AVAR}}^{p}_{dis,sub}}$.

- Calculate $\hat{I}_1-\hat{I}_4$, given in (22a)-(22d), with

$$K_N=\sqrt{\hat{c}^{(p)}_{multi}}N^{1/5}\ \text{ bins and }\ M_N(j)=\left(\hat{c}^{(p)}_{multi}\right)^{5/4}N^{3/5}\,\forall\, j\ .$$

- Estimate asymptotic variance with $\hat{I}_1-\hat{I}_4$ and $\widehat{\eta^2_X},\widehat{\eta^2_Y}$ and determine $\hat{c}_{multi}$ and $\hat{c}_{sub}$ with the above given formulae.
- Calculate the generalized multiscale estimator (12) with optimal weights (A.1) (and the one-scale subsample estimator) with $M_N=\hat{c}_{multi}\sqrt{N}$ (and $i_N=\hat{c}_{sub}N^{2/3}$)

Algorithm 2: Algorithm for the estimation procedure.

We apply the procedure with Algorithm 1 and Algorithm 2 to a suitable filtered dataset and give the results for two days in Table 3 with associated estimated optimal multiscale frequencies. The estimates for the quadratic covariations are given $\pm$ estimated standard deviation from the estimated asymptotic variances and the pertaining N for each pair. Although there are characteristics of the dataset not in accordance with the model assumptions, above all price discreteness and the fact that most returns are zero, the estimation approach passes this intuition check. Since the ESX and the DAX share 13 companies constituting c. 28.5% weighting in the ESX and c. 72.4% in the DAX there is a big systematic positive correlation between both and we presume that there is as well a high correlation between the two debt instruments which is both revealed by the estimates. On 09/11/2001, there has been a tremendous impact so that FGBL/FGBM have increased and the FESX/FDAX decreased. For that day we have significantly negative integrated covolatilities between debt instruments and stock indices which is not the case for the ordinary trading day in comparison. As answer to the great amount of zero returns the only adjustment of the method that we undertake is to estimate noise variances in (24) by dividing the realized volatility by twice the number of non-zero returns instead of all returns.
To sum up, the estimation approach based on Algorithm 1, a multiscale extension of subsampling and Algorithm 2 provides a convincing method to obtain integrated covolatility estimates for very general high-frequency data.

Table 2: Estimators (22a)-(22d), estimators for the asymptotic variances of the multiscale (18) and the one-scale estimator (20), calculated asymptotic variances and estimates for the quadratic covariation. The estimates are given $\pm$ empirical standard deviations.

noise var. η^2	0.0001	$(0.001/\sqrt{10})$	0.001	$(0.01/\sqrt{10})$	0.01
$\hat{I}_1$	0.392 ± 0.038	0.390 ± 0.047	0.394 ± 0.073	0.413 ± 0.144	0.423 ± 0.128
$\hat{I}_2$	1.557 ± 0.067	1.552 ± 0.085	1.538 ± 0.141	1.529 ± 0.276	1.462 ± 0.230
$\hat{I}_3$	0.250 ± 0.007	0.249 ± 0.010	0.249 ± 0.016	0.247 ± 0.031	0.234 ± 0.085
$\hat{I}_4$	0.250 ± 0.007	0.249 ± 0.009	0.249 ± 0.016	0.247 ± 0.030	0.233 ± 0.083
$\widehat{\mathbf{AVAR}}_{multi}$	0.090 ± 0.003	0.143 ± 0.005	0.246 ± 0.015	0.434 ± 0.050	0.778 ± 0.082
$\mathbf{AVAR}_{multi}$	0.0663	0.1185	0.2159	0.3774	0.6737
$\widehat{\mathbf{AVAR}}_{sub}$	0.0086 ± 0.0002	0.017 ± 0.001	0.037 ± 0.002	0.080 ± 0.009	0.157 ± 0.017
$\mathbf{AVAR}_{sub}$	0.0077	0.0166	0.0357	0.0768	0.1656
$\widehat{[X,Y]}_T^{multi}$	0.501 ± 0.024	0.499 ± 0.029	0.498 ± 0.038	0.499 ± 0.049	0.501 ± 0.065
$\widehat{[X,Y]}_T^{sub}$	0.500 ± 0.022	0.499 ± 0.028	0.499 ± 0.042	0.500 ± 0.058	0.503 ± 0.074

$\widehat{[X,Y]}_T^{multi}$ (MSFR)	FGBL	FGBM	FESX	FDAX
FGBL	$5.46 \pm .87$ (7)	$3.32 \pm .74$ (7)	$0.98 \pm .99$ (9)	$0.52 \pm .74$ (6)
FGBM		$2.42 \pm .47$ (6)	$0.78 \pm .69$ (7)	$0.64 \pm .51$ (4)
FESX			68.26 ± 7.1 (10)	29.39 ± 2.4 (7)
FDAX				61.43 ± 3.5 (5)

$\widehat{[X,Y]}_T^{multi}$ (MSFR)	FGBL	FGBM	FESX	FDAX
FGBL	27.89 ± 4.1 (15)	12.94 ± 1.7 (12)	-52.55 ± 18 (8)	-34.13 ± 11 (7)
FGBM		18.10 ± 2.5 (8)	-26.44 ± 14 (6)	-25.01 ± 28 (3)
FESX			3070 ± 172 (6)	757 ± 15 (4)
FDAX				1870 ± 94 (4)

Table 3: Estimates for integrated covolatilities, standard deviations ($\cdot\, 10^6$) and used multiscale frequencies for 10/01/2008 (top) and 09/11/2001 (bottom).

Appendix A. Proof of Theorem 3

Appendix A.1. Error due to noise and choosing the weights

The error due to microstructure noise of the generalized multiscale estimator is given by

$$\sum_{i=1}^{M_N} \frac{\alpha_{i,M_N}}{i} \sum_{j=i}^{N} \left(\epsilon^X_{g_j} - \epsilon^X_{l_{j-i+1}}\right)\left(\epsilon^Y_{\gamma_j} - \epsilon^Y_{\lambda_{j-i+1}}\right) = \sum_{i=1}^{M_N} \frac{\alpha_{i,M_N}}{i} \Big(\sum_{j=1}^{N} \left(\epsilon^X_{g_j}\epsilon^Y_{\gamma_j} + \epsilon^X_{l_j}\epsilon^Y_{\lambda_j}\right) - \sum_{j=i}^{N} \left(\epsilon^X_{g_j}\epsilon^Y_{\lambda_{j-i+1}} + \epsilon^Y_{\gamma_j}\epsilon^X_{l_{j-i+1}}\right) - \sum_{j=1}^{i-1} \epsilon^X_{g_j}\epsilon^Y_{\gamma_j} - \sum_{j=N-i+1}^{N} \epsilon^X_{l_j}\epsilon^Y_{\lambda_j} \Big).$$

Additionally to the standardization condition

$$\sum_{i=1}^{M_N} \alpha_{i,M_N} = 1\,, \tag{C1}$$

that is necessary for asymptotic unbiasedness and consistency, we now impose the auxiliary condition

$$\sum_{i=1}^{M_N} \frac{\alpha_{i,M_N}}{i} = 0\,, \tag{C2}$$

on the weights which assures that the leading term of the noise error equals zero. Hence, there remain three uncorrelated addends in the error induced by microstructure noise.

Proposition A.1. *Let Assumptions 2 and 5 on the observation times and Assumption 3 on the observation errors hold true. The asymptotic variance of the term*

$$-\sum_{i=1}^{M_N}\frac{\alpha_{i,M_N}}{i}\sum_{j=i}^{N}\left(\epsilon^X_{g_j}\epsilon^Y_{\lambda_{j-i+1}}+\epsilon^Y_{\gamma_j}\epsilon^X_{l_{j-i+1}}\right)$$

is minimized by the weights

$$\alpha^{opt}_{i,M_N}=\left(\frac{12i^2}{(M_N^3-M_N)}-\frac{6i}{(M_N^2-1)}-\frac{6i}{(M_N^3-M_N)}\right)=\frac{12i^2}{M_N^3}-\frac{6i}{M_N^2}\left(1+\mathcal{O}(1)\right) \tag{A.1}$$

as $M_N, N\to\infty$ and $M_N/N\to 0$ with $N=\mathcal{O}\left(M_N^4\right)$. The following asymptotic normality result holds true:

$$\sqrt{\frac{M_N^3}{N}}\left(\sum_{i=1}^{M_N}\frac{\alpha^{opt}_{i,M_N}}{i}\sum_{j=i}^{N}\left(\epsilon^X_{g_j}\epsilon^Y_{\lambda_{j-i+1}}-\epsilon^Y_{\gamma_j}\epsilon^X_{l_{j-i+1}}\right)\right)\rightsquigarrow\mathbf{N}\left(0\,,\,\mathbf{AVAR}_{noise}\right)\,, \tag{A.2}$$

with the asymptotic variance

$$\mathbf{AVAR}_{noise}=(24+12(I_X(T)+I_Y(T)))\eta_X^2\eta_Y^2 \tag{A.3}$$

with the functions I_X and I_Y defined in Assumption 5. The weak convergence also holds true conditionally given the paths of the efficient processes.

Proof. The term is centred and we illustrate it in the way

$$\begin{aligned}-\sum_{i=1}^{M_N}\frac{\alpha_{i,M_N}}{i}\sum_{j=i}^{N}\left(\epsilon^X_{g_j}\epsilon^Y_{\lambda_{j-i+1}}+\epsilon^Y_{\gamma_j}\epsilon^X_{l_{j-i+1}}\right)=-\sum_{j=1}^{N}\sum_{i=1}^{M_N\wedge j}\frac{\alpha_{i,M_N}}{i}\Big(\epsilon^X_{g_j}\epsilon^Y_{\lambda_{j-i+1}}\\ (\mathbb{1}_{\{g_j=g_{j+1}\}}+\mathbb{1}_{\{g_j\neq g_{j+1}\}})+\epsilon^Y_{\gamma_j}\epsilon^X_{l_{j-i+1}}(\mathbb{1}_{\{\gamma_j=\gamma_{j+1}\}}+\mathbb{1}_{\{\gamma_j\neq\gamma_{j+1}\}})\Big)\,.\end{aligned}$$

For fixed i the addends of the inner sum are uncorrelated because $l_i\neq l_j$ and $\lambda_i\neq\lambda_j$ for all $i\neq j$. Consecutive right-end points g_i,γ_i can be the same observation times instead, so that the inner sums are 2-dependent random variables. Thus, the variance is given by

$$\begin{aligned}&\mathbb{V}\mathrm{ar}\left(\sum_{j=1}^{N}\frac{\alpha_{i,M_N}}{i}\sum_{i=1}^{M_N\wedge j}\left(\epsilon^X_{g_j}\epsilon^Y_{\lambda_{j-i+1}}-\epsilon^Y_{\gamma_j}\epsilon^X_{l_{j-i+1}}\right)\right)\\&=\sum_{j=1}^{N}\sum_{i=1}^{M_N\wedge j}\left(\frac{\alpha_{i,M_N}}{i}\right)^2 2\eta_X^2\eta_Y^2+\sum_{j=1}^{N}\sum_{i=1}^{(M_N-1)\wedge(j-1)}\frac{\alpha_{i,M_N}\alpha_{i+1,M_N}}{i(i+1)}\eta_X^2\eta_Y^2(1-\mathbb{1}_{\{g_j\neq g_{j+1}\,,\,\gamma_j\neq\gamma_{j+1}\}})\,.\end{aligned}$$

The weights that minimize the first addend of the above variance and also the total variance asymptotically have been determined in [8]. Those weights are in line with the standard weights from [30] in the univariate setting and correspond to a cubic kernel for the kernel estimator by [5].
Inserting the noise-optimal weights (A.1), we can apply a central limit theorem for strong mixing triangular arrays from [28] to

$$\sqrt{\frac{M_N^3}{N}}\sum_{j=1}^{N}\frac{\alpha^{opt}_{i,M_N}}{i}\sum_{i=1}^{M_N\wedge j}\left(\epsilon^X_{g_j}\epsilon^Y_{\lambda_{j-i+1}}+\epsilon^X_{l_{j-i+1}}\epsilon^Y_{\gamma_j}\right)\,.$$

The sequence of variances with the chosen weights according to (A.1)

$$\begin{aligned}
&\mathbb{V}\mathrm{ar}\left(\sqrt{\frac{M_N^3}{N}}\sum_{j=1}^{N}\frac{\alpha^{opt}_{i,M_N}}{i}\sum_{i=1}^{M_N\wedge j}\left(\epsilon^X_{g_j}\epsilon^Y_{\lambda_{j-i+1}}+\epsilon^X_{l_{j-i+1}}\epsilon^Y_{\gamma_j}\right)\right)\\
&=\frac{M_N^3}{N}\sum_{j=0}^{N}\sum_{i=1}^{M_N\wedge(j+1)}\left(\frac{\alpha^{opt}_{i,M_N}}{i}\right)^2\eta_X^2\eta_Y^2(3-\mathbb{1}_{\{g_j\neg_{j+1}\,,\,\gamma_j\neq\gamma_{j+1}\}})+\mathcal{O}(1)\\
&\longrightarrow 36\eta_X^2\eta_Y^2-12(1-(I_X(T)+I_Y(T)))\eta_X^2\eta_Y^2
\end{aligned}$$

converges to $\mathbf{AVAR}_{\text{noise}}$ on the Assumption 5.
Since the inner sums are 2-dependent and hence in particular ϕ-mixing, the Lyapunov condition that holds obviously suffices to apply the central limit theorem of [28]. This completes the proof of the proposition. □

Next, we consider the remaining addends of the error induced by microstructure frictions and insert the weights (A.1):

Proposition A.2. *On the Assumptions 2, 5 and 3, the following weak convergence to a centred normal distribution holds true:*

$$\sqrt{M_N}\left(\sum_{i=1}^{M_N}\frac{\alpha^{opt}_{i,M_N}}{i}\left(\sum_{j=1}^{i-1}\epsilon^X_{g_j}\epsilon^Y_{\gamma_j}+\sum_{j=N-i+1}^{N}\epsilon^X_{l_j}\epsilon^Y_{\lambda_j}\right)\right)\rightsquigarrow\mathbf{N}\left(0,\frac{12}{5}\eta_X^2\eta_Y^2\right)\,. \tag{A.4}$$

This convergence also holds conditionally on the paths of the efficient processes.

Proof.

$$\sum_{i=1}^{M_N}\frac{\alpha^{opt}_{i,M_N}}{i}\left(\sum_{j=1}^{i-1}\epsilon^X_{g_j}\epsilon^Y_{\gamma_j}+\sum_{j=N-i+1}^{N}\epsilon^X_{l_j}\epsilon^Y_{\lambda_j}\right)=\sum_{j=1}^{M_N-1}\left(\epsilon^X_{g_j}\epsilon^Y_{\gamma_j}+\epsilon^X_{l_{N-j}}\epsilon^Y_{\lambda_{N-j}}\right)\sum_{i=j+1}^{M_N}\frac{\alpha^{opt}_{i,M_N}}{i}$$

Both addends are uncorrelated and treated analogously. We restrict ourselves to the proof for the first term. $\sqrt{M_N}\sum_{j=1}^{M_N-1}\epsilon^X_{g_j}\epsilon^Y_{\gamma_j}\sum_{i=j+1}^{M_N}\alpha^{opt}_{i,M_N}/i$ is the endpoint of a discrete centred martingale with respect to the filtration $\overline{\mathcal{A}}^N_j:=\sigma\left(\epsilon^X_{t_k}|t_k\le g_j\,,\,X_{t_k}|0\le k\le n\right)\vee\sigma\left(\epsilon^Y_{\tau_k}|\tau_k\le\gamma_j\,,\,Y_{\tau_k}|0\le k\le m\right)$. Namely, since $g_j=g_{j-1}\Rightarrow\gamma_j>\gamma_{j-1}$ and $\gamma_j=\gamma_{j-1}\Rightarrow g_j>g_{j-1}$ analogously:

$$\begin{aligned}
&\mathbb{E}\left[\sqrt{M_N}\epsilon^X_{g_l}\epsilon^Y_{\gamma_l}\sum_{i=l+1}^{M_N}\frac{\alpha^{opt}_{i,M_N}}{i}\Big|\overline{\mathcal{A}}^N_{l-1}\right]\\
&=\sqrt{M_N}\Big(\mathbb{1}_{\{g_l=g_{l-1}\}}\mathbb{E}\left[\epsilon^Y_{\gamma_l}\right]\epsilon^X_{g_{l-1}}+\mathbb{1}_{\{\gamma_l=\gamma_{l-1}\}}\mathbb{E}\left[\epsilon^X_{g_l}\right]\epsilon^Y_{\gamma_{l-1}}+\mathbb{1}_{\{g_l\neq g_{l-1}\,,\,\gamma_l\neq\gamma_{l-1}\}}\mathbb{E}\left[\epsilon^X_{g_l}\right]\mathbb{E}\left[\epsilon^Y_{\gamma_l}\right]\Big)=0\,.
\end{aligned}$$

A central limit theorem for martingale triangular arrays from [14] is applied, in particular the non-stable version of Corollary 3. 1 (cf. the following remark in [14] and references cited therein). The conditional Lindeberg condition can be verified by the stronger conditional Lyapunov condition. The proof of it is obtained by a similar calculation as the following one and we omit it. The conditional variance equals

$$\begin{aligned}
M_N\sum_{j=1}^{M_N-1}\mathbb{V}\mathrm{ar}\left(\epsilon^X_{g_j}\epsilon^Y_{\gamma_j}\sum_{i=j+1}^{M_N}\frac{\alpha^{opt}_{i,M_N}}{i}\Big|\mathcal{A}_{N,j-1}\right)&=M_N\sum_{j=1}^{M_N-1}\left(\sum_{i=j+1}^{M_N}\frac{\alpha^{opt}_{i,M_N}}{i}\right)^2\\
&\quad\times\left(\eta_X^2\eta_Y^2\,\mathbb{1}_{\{g_j\neq g_{j-1}\,,\,\gamma_j\neq\gamma_{j-1}\}}+\left(\epsilon^X_{g_j}\right)^2\mathbb{1}_{\{g_j=g_{j-1}\}}\eta_Y^2+\left(\epsilon^Y_{\gamma_j}\right)^2\mathbb{1}_{\{\gamma_j=\gamma_{j-1}\}}\eta_X^2\right)\\
\xrightarrow{p}\frac{6}{5}\eta_X^2\eta_Y^2\,.&
\end{aligned}$$

We have used the formula $\sum_{j=1}^{M_N-1}\left(\sum_{i=j+1}M_N(\alpha^{opt}_{i,M_N}/i)\right)^2=(6/5)M_N^{-1}+\mathcal{O}(M_N^{-1})$.

□

A.2. Discretization errors of the estimators

Proposition A.3. *On the Assumptions 1, 2 and 4, the discretization error of the one-scale subsampling estimator with subsampling frequency i_N converges stably in law to a centred mixed normal limit as $i_N \to \infty, N \to \infty, i_N/N^\alpha \to 0$ for every $\alpha > 2/3$:*

$$\sqrt{\frac{N}{i_N}} \left(\sum_{j=i}^{N} \left(X_{g_j} - X_{l_{j-i+1}} \right) \left(Y_{\gamma_j} - Y_{\lambda_{j-i+1}} \right) - [X,Y]_T \right) \overset{st}{\rightsquigarrow} \mathbf{N}\left(0, \mathbf{AVAR}_{dis,sub}\right) \ ,$$

with asymptotic variance

$$\mathbf{AVAR}_{dis,sub} = \frac{2}{3} T \int_0^T G'(t) (\sigma_t^X \sigma_t^Y)^2 (1 + \rho_t^2) \, dt \ . \tag{A.5}$$

Proposition A.4. *On the Assumptions 1,2 and 4, the discretization error of the generalized multiscale estimator with the noise-optimal weights given in* (A.1) *converges with rate $\sqrt{N/M_N}$ stably in law to a centred mixed Gaussian limit as $M_N \to \infty, N \to \infty, M_N/N^\alpha \to 0$ for every $\alpha > 2/3$:*

$$\sqrt{\frac{N}{M_N}} \left(\sum_{i=1}^{M_N} \frac{\alpha_{i,M_N}^{opt}}{i} \sum_{j=i}^{N} \left(X_{g_j} - X_{l_{j-i+1}} \right) \left(Y_{\gamma_j} - Y_{\lambda_{j-i+1}} \right) - [X,Y]_T \right) \overset{st}{\rightsquigarrow} \mathbf{N}\left(0, \mathbf{AVAR}_{dis,multi}\right) \ ,$$

with asymptotic variance

$$\mathbf{AVAR}_{dis,multi} = \frac{26}{35} T \int_0^T G'(t) (\sigma_t^X \sigma_t^Y)^2 (1 + \rho_t^2) \, dt \ . \tag{A.6}$$

A.2.1. Time-change in the asymptotic quadratic variation of time

Proposition A.5. *In the proof of a central limit theorem for the discretization error of the closest synchronous approximation $T_k^{(N)}$, $k = 0, \ldots, N$, of our generalized multiscale estimator* (12) *on the Assumptions 2 and 5, we can additionally, without loss of further generality, assume that*

$$\sum_{k=1}^{N} \left(\Delta T_k^{(N)} - \frac{T}{N} \right)^2 = o\left(N^{-1} \right) \ . \tag{A.7}$$

Remark 2. From Assumptions 1 and 2, we can deduce directly that the sum above is at most of order N^{-1}. The stronger assertion, that the closest synchronous approximation defined by the times $T_k^{(N)}$, $k = 0, \ldots, N$ introduced in paragraph 3.2 is close to equidistant sampling in the sense that the sum above is of smaller asymptotic order than N^{-1}, is derived by the concept of a time-change in the asymptotic quadratic variation of time from Assumption 4. For the proof of (A.7) we refer to Lemma 1 from [30] where this concept has been presented for the univariate multiscale approach and it directly carries over to the synchronous multivariate case.
On the Assumption 4, a transformation g can be defined that maps the refresh times $T_k^{(N)}$ to values $g(T_k^{(N)})$, so that (A.7) holds true for the transformed synchronous observation scheme. Thanks to the fact that the corresponding time-changed processes $L_{g(t)}$ and $M_{g(t)}$ fulfill Assumption 1 again and the transformed observation scheme Assumption 2, we are able to prove a central limit theorem for the time-changed version of the discretization error if (A.7) did not hold.
Since the resulting asymptotic variance will be invariant under the transformation g, the central limit theorem will analogously hold true for the original sampling scheme. Hence, no further restriction has to be made when assuming (A.7).

A.2.2. Discretization error of the closest synchronous approximation

Note that it suffices to prove the foregoing limit theorems for the zero-drift case. Since our limit theorems are stable, asymptotic mixed normality is assured to hold for the general setting on Assumption 1. Denote $L_t = \int_0^t \sigma_s^X dW_s^X$ and $M_t = \int_0^t \sigma_s^Y dW_s^Y$ the continuous martingales that represent the efficient

processes under the equivalent martingale measure after a Girsanov transformation. The Novikov condition has been imposed in Assumption 1 to allow for this transformation.
The asymptotic mixed normality result is implied as marginal distribution at $t = T$ of a limiting time-changed Brownian motion which is proven to be the stable weak limit of the process corresponding to the discretization error (21) with the theory of [18].
We begin with the discretization error of the closest synchronous approximation of a one-scale subsampling estimator.

Proposition A.6. *On the same assumptions as in Proposition A.3, the continuous martingale*

$$\mathfrak{D}_t^N := \sqrt{\frac{N}{i_N T}} \left[\sum_{T_k \leq t} \left(\Delta L_{T_k} + \int_{T_k}^t \sigma_s^X dW_s^X \right) \left(\sum_{l=1}^{i \wedge k} \left(1 - \frac{l}{i} \right) \Delta M_{T_{k-l}} \right) \right.$$
$$\left. + \sum_{T_k \leq t} \left(\Delta M_{T_k} + \int_{T_k}^t \sigma_s^Y dW_s^Y \right) \left(\sum_{l=1}^{i \wedge k} \left(1 - \frac{l}{i} \right) \Delta L_{T_{k-l}} \right) \right]$$

for $t \in [0,T]$, where $\Delta \cdot_{T_k} = \cdot_{T_k} - \cdot_{T_{k-1}}$ is the backward difference operator, converges stably in law as $N \to \infty, i_N \to \infty, i_N/N \to 0$ to a limiting time-changed Brownian motion

$$\mathfrak{D}_t^N \overset{st}{\rightsquigarrow} \int_0^t \sqrt{v_{\mathfrak{D}_s}} d\mathfrak{W}_s^\perp \,,$$

where $\mathfrak{W}^\perp$ is independent of $\mathcal{F}$ and

$$v_{\mathfrak{D}_s} = \frac{2}{3} G'(s) (\sigma_s^X \sigma_s^Y)^2 (1 + \rho_s^2) \,.$$

Proof of Proposition A.6:
The subscript of the subsampling frequency is omitted in the following proof and C denotes a generic constant and $\delta_N = \sup_{i \in \{1,\ldots,N\}} (T_i - T_{i-1})$.
We apply a simplified martingale version of the stable central limit theorem 2–1 from [18]. For other applications and expositions of the theory from [18] we refer to [24], [12] and [7]. The above limit theorem is implied by the following three conditions:

$$[\mathfrak{D}]_t \xrightarrow{p} \int_0^t v_{\mathfrak{D}_s} \, ds \,, \tag{A.8a}$$

$$[\mathfrak{D}, L]_t \xrightarrow{p} 0 \,, \;\; [\mathfrak{D}, M]_t \xrightarrow{p} 0, \tag{A.8b}$$

$$\left[\mathfrak{D}, L^\perp\right]_t \xrightarrow{p} 0 \,, \;\; \left[\mathfrak{D}, M^\perp\right]_t \xrightarrow{p} 0, \tag{A.8c}$$

for all $t \in [0,T]$ and for all $M^\perp \in \mathcal{M}\perp$ and $L^\perp \in \mathcal{L}\perp$ that denote the set of $(\mathcal{F})_t$-adapted bounded martingales orthogonal to M and L in the sense that $\left[L, L^\perp\right] = 0$, respectively.
Calculating the quadratic variation of $\mathfrak{D}_t^N$ yields

$$\left[\mathfrak{D}^N\right]_t = \frac{N}{iT} \left[\sum_{T_k \leq t} \left(\Delta [L]_{T_k} \left(\sum_{l=1}^{i \wedge k} \left(1 - \frac{l}{i} \right) \Delta M_{T_{k-l}} \right)^2 + \Delta [M]_{T_k} \left(\sum_{l=1}^{i \wedge k} \left(1 - \frac{l}{i} \right) \Delta L_{T_{k-l}} \right)^2 \right) \right.$$
$$\left. + 2 \sum_{T_k \leq t} \Delta [L, M]_{T_k} \left(\sum_{l=1}^{i \wedge k} \left(1 - \frac{l}{i} \right) \Delta L_{T_{k-l}} \right) \left(\sum_{l=1}^{i \wedge k} \left(1 - \frac{l}{i} \right) \Delta M_{T_{k-l}} \right) \right] + \mathcal{O}_p(1)$$

$$
\begin{aligned}
&= \frac{N}{iT}\Bigg[\sum_{T_k\le t}\Delta\left[L\right]_{T_k}\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\left(\Delta M_{T_{k-l}}\right)^2+\sum_{T_k\le t}\Delta\left[M\right]_{T_k}\\
&\quad\times\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\left(\Delta L_{T_{k-l}}\right)^2+2\sum_{T_k\le t}\Delta\left[L,M\right]_{T_k}\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\left(\Delta L_{T_{k-l}}\right)^2\Bigg]+\mathcal{O}_p(1)\\
&= \frac{N}{iT}\Bigg[\sum_{T_k\le t}\int_{T_{k-1}}^{T_k}(\sigma_s^X)^2ds\left(\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\int_{T_{k-l-1}}^{T_{k-l}}(\sigma_s^Y)^2ds\right)\\
&\qquad+\sum_{T_k\le t}\int_{T_{k-1}}^{T_k}(\sigma_s^Y)^2ds\left(\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\int_{T_{k-l-1}}^{T_{k-l}}(\sigma_s^X)^2ds\right)\\
&\qquad+2\sum_{T_k\le t}\int_{T_{k-1}}^{T_k}\rho_s\sigma_s^X\sigma_s^Yds\left(\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\int_{T_{k-l-1}}^{T_{k-l}}\rho_s\sigma_s^X\sigma_s^Yds\right)\Bigg]+\mathcal{O}_p(1)\\
&\underset{\text{Lemma A.7}}{=}\frac{N}{iT}\sum_{T_k\le t}2(1+\rho_{T_{k-1}}^2)(\sigma_{T_{k-1}}^X\sigma_{T_{k-1}}^Y)^2\left(\Delta T_k\right)^2\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2+\mathcal{O}_p(1)\\
&=\sum_{T_k\le t}\frac{2}{3}\frac{G^N(T_k)-G^N(T_{k-1})}{T_k-T_{k-1}}\left(\rho_{T_{k-1}}^2+1\right)\left(\sigma_{T_{k-1}}^X\sigma_{T_{k-1}}^Y\right)^2\Delta T_k+\mathcal{O}_p(1)\\
&\xrightarrow{p}\frac{2}{3}\int_0^t(1+\rho_s^2)(\sigma_s^X\sigma_s^Y)^2G'(s)\,ds\ .
\end{aligned}
$$

In the first step cross terms of the inner sums have been neglected since they are centred and by Itô isometry it can be shown that their second moments are bounded from above by $Ci^3\delta_N^3$. We frequently use estimates δ_N^{l-1} for sums of the type $\sum_i^N(\Delta T_i)^l, l>1$, by Hölder's inequality with the supremum norm to obtain upper bounds.
Subsequently, squared increments of L and M and the increments of the product $L\cdot M$ in these inner sums are substituted by the increments of the quadratic (co-)variation processes. The induced error terms are centred by Itô isometry and involving the Cauchy-Schwarz inequality it follows that $C\delta_N$ is an upper bound for their second moments.
The crucial non-standard approximation is that on each block $(T_{k-1},\dots,T_{k-i\vee0})$ the increments of the form $\int_{T_{k-l-1}}^{T_{k-l}}f(t)dt$ with continuous functions f for $l=1,\dots,k\wedge i$ are approximated by $\Delta T_k f\left(T_{k-1}\right)$. This blockwise approximation is treated in Lemma A.7 and makes use of the concept of a time-changed quadratic variation of times and particularly (A.7). Finally, $1/i\sum_{l=1}^i(1-(l/i))^2=1/3+\mathcal{O}(1)$ and the convergence in probability is ensured by Assumption 4 and the convergence of the Riemann sums to the integral.

Lemma A.7. *On the same assumptions as in Proposition A.3, it holds true that the term*

$$
\frac{N}{iT}\sum_{T_k\le t}\left(\int_{T_{k-1}}^{T_k}(\sigma_s^X)^2ds\left(\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\int_{T_{k-l-1}}^{T_{k-l}}(\sigma_s^Y)^2ds\right)-(\sigma_{T_{k-1}}^X\sigma_{T_{k-1}}^Y\Delta T_k)^2\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\right),
$$

and the analogous blockwise approximations of $\Delta[M]_{T_k}$ and $\Delta[M]_{T_k}$ by constant left-end points converge to zero in probability.

Proof. The approximation uses the concept of a time-change in the asymptotic quadratic variation of refresh times introduced in [30] which is expounded in Proposition A.5. By virtue of that concept we may suppose without loss of generality that the sampling design of the closest synchronous approximation satisfies (A.7).
The asymptotic orders of the three terms are deduced analogously and we restrict us to the proof of the

above given first term. An application of the mean value theorem yields

$$\frac{N}{iT}\sum_{T_k\le t}\int_{T_{k-1}}^{T_k}(\sigma_s^X)^2ds\left(\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\int_{T_{k-l-1}}^{T_{k-l}}(\sigma_s^Y)^2ds\right)$$
$$=\frac{N}{iT}\sum_{T_k\le t}(\sigma_{\zeta_k}^X)^2\Delta T_k\left(\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2(\sigma_{\zeta_{k-l}^*}^Y)^2\Delta T_{k-l}\right)$$

with $\zeta_k\in[T_{k-1},T_k]$, $\zeta_q^*\in[T_{q-1},T_q]$. Since the volatility processes σ^X,σ^Y are uniformly continuous on $[0,T]$ by Assumption 1

$$\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\left|(\sigma_{\zeta_{k-l}^*}^Y)^2-(\sigma_{T_{k-1}}^Y)^2\right|\Delta T_{k-l}\le i\delta_N\sup_{|t-s|\le i\delta_N}\left|(\sigma_t^Y)^2-(\sigma_s^Y)^2\right|=\mathcal{O}_{a.\,s.}(i\delta_N)\,,$$

$$\sum_{T_k\le t}\left|(\sigma_{\zeta_k}^X)^2-(\sigma_{T_{k-1}}^X)^2\right|(\sigma_{T_{k-1}}^Y)^2\Delta T_k\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\Delta T_{k-l}\le i\delta_N\sup_{|t-s|\le\delta_N}\left|(\sigma_t^Y)^2-(\sigma_s^Y)^2\right|=\mathcal{O}_{a.\,s.}(i\delta_N)$$

hold almost surely (denoted $a.\,s.$).
With the Cauchy-Schwarz inequality and (A.7), we obtain

$$\begin{aligned}&\frac{N}{iT}\sum_{T_k\le t}(\sigma_{T_{k-1}}^X\sigma_{T_{k-1}}^Y)^2\Delta T_k\left(\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\left|\Delta T_{k-l}-\frac{T}{N}\right|\right)\\&\quad\le\frac{N}{iT}\sup_{s\in[0,t]}\left(\sigma_t^X\sigma_t^Y\right)^2\sum_{l=1}^{i}\sum_{j=1}^{N-l}\left|\left(\Delta T_j-\frac{T}{N}\right)\Delta T_{j+l}\right|\\&\quad\le\frac{N}{iT}C\left(\sum_{j=1}^{N}\left(T_{(j+i)\vee N}-T_j\right)^2\sum_{j=1}^{N}\left(\Delta T_j-\frac{T}{N}\right)^2\right)^{1/2}=\mathcal{O}_{a.\,s.}(1)\,.\end{aligned}$$

Furthermore,

$$\frac{N}{iT}\sum_{T_k\le t}(\sigma_{T_{k-1}}^X\sigma_{T_{k-1}}^Y)^2\Delta T_k\left|\Delta T_k-\frac{T}{N}\right|\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)^2\le\frac{N}{T}C\left(\sum_{j=1}^{N}(\Delta T_j)^2\sum_{j=1}^{N}\left(\Delta T_j-\frac{T}{N}\right)^2\right)^{1/2}$$

holds, where the right-hand side converges to zero almost surely due to (A.7) and the Cauchy-Schwarz inequality. The preceding estimates imply the statement of the lemma. □

We proceed proving (A.8b) that the quadratic covariations $\left[\mathfrak{D}^N,L\right]_t$ and $\left[\mathfrak{D}^N,M\right]_t$ converge to zero in probability for all $t\in[0,T]$.

$$\left[\mathfrak{D}^N,L\right]_t=\sqrt{\frac{N}{iT}}\sum_{T_k\le t}\left(\Delta\left[L\right]_{T_k}\left(\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)\Delta M_{T_{k-l}}\right)+\Delta\left[L,M\right]_{T_k}\left(\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)\Delta L_{T_{k-l}}\right)\right)$$

has an expectation equal to zero for all $t\in[0,T]$ and the second moment is bounded above by $iN\delta_N^2$. The order follows from the evaluation of the second moment that is carried out analogously as for the calculation of $\left[\mathfrak{D}^N\right]_t$ before. For this reason $\left[\mathfrak{D}^N,L\right]$ converges to zero in probability on $[0,T]$. It can be directly deduced that $\left[\mathfrak{D}^N,M\right]_t=\mathcal{O}_p(1)$ as well. If $L^\perp$ is a bounded $(\mathcal{F}_t)$-martingale with $\left[L,L^\perp\right]\equiv0$, the quadratic covariation

$$\left[\mathfrak{D}^N,L^\perp\right]_t=\sqrt{\frac{N}{iT}}\sum_{T_k\le t}\Delta\left[L^\perp,M\right]_{T_k}\left(\sum_{l=1}^{i\wedge k}\left(1-\frac{l}{i}\right)\Delta L_{T_{k-l}}\right)$$

converges to zero in probability on $[0,T]$ what can be concluded following the same principles and also that $\left[\mathfrak{D}^N, M^{\perp}\right]_t = \mathcal{O}_p(1)\ \forall M^{\perp} \in \mathcal{M}^{\perp},\ t \in [0,T]$. An application of Jacod's Theorem from [18] completes the proof of Proposition A.6. □

Proposition A.8. *On the same assumptions as in Proposition A.4, the continuous martingale*

$$\mathfrak{M}_t^N := \sqrt{\frac{N}{M_N}} \sum_{i=1}^{M_N} \left[\sum_{T_k \le t} \left(\Delta L_{T_k} + \int_{T_k}^t \sigma_s^X dW_s^X \right) \left(\sum_{l=1}^{i \wedge k} \left(1 - \frac{l}{i} \right) \Delta M_{T_{k-l}} \right) \right.$$
$$\left. + \sum_{T_k \le t} \left(\Delta M_{T_k} + \int_{T_k}^t \sigma_s^Y dW_s^Y \right) \left(\sum_{l=1}^{i \wedge k} \left(1 - \frac{l}{i} \right) \Delta L_{T_{k-l}} \right) \right]$$

for $t \in [0,T]$ converges stably in law as $N \to \infty, M_N \to \infty, M_N/N^{\alpha} \to 0$ for every $\alpha > 2/3$ to a limiting time-changed Brownian motion

$$\mathfrak{M}_t^N \overset{st}{\rightsquigarrow} \int_0^t \sqrt{v_{\mathfrak{M}_s}} d\tilde{\mathfrak{W}}_s^{\perp} \,,$$

where $\tilde{\mathfrak{W}}^{\perp}$ is independent of $\mathcal{F}$ and with

$$v_{\mathfrak{M}_s} = \frac{26}{35} T G'(s) (\sigma_s^X \sigma_s^Y)^2 (1 + \rho_s^2) \,.$$

Proof of Proposition A.8:
The discretization error of the generalized multiscale estimator calculated with the closest synchronous approximation under the equivalent martingale measure where the drift terms equal zero

$$\sum_{i=1}^{M_N} \frac{\alpha_{i,M_N}^{opt}}{i} \sum_{j=i}^{N} \left(L_{T_j} - L_{T_{j-i}} \right) \left(M_{T_j} - M_{T_{j-i}} \right) - [X,Y]_T$$
$$= \sum_{i=1}^{M_N} \alpha_{i,M_N}^{opt} \left(\frac{1}{i} \sum_{j=i}^{N} \left(L_{T_j} - L_{T_{j-i}} \right) \left(M_{T_j} - M_{T_{j-i}} \right) - [X,Y]_T \right)$$

equals the weighted sum of $M_N \to \infty$ discretization errors of the type considered in Proposition A.3 because $\sum_{i=1}^{M_N} \alpha_{i,M_N}^{opt} = 1$. Note, that all approximation errors in the preceding proof of Proposition A.6 converge to zero in probability as long as $N \to \infty, i/N^{\alpha} \to 0$ for every $\alpha > 2/3$.
We begin with the proof of a multivariate stable central limit theorem for a finite-dimensional vector:

Lemma A.9. *Consider the sequence of K-dimensional vectors $\mathbb{D}^N = \left(\mathfrak{D}^{i_N^1}, \ldots, \mathfrak{D}^{i_N^K} \right)$ where the entries $\mathfrak{D}^{i_N^k}, k = 1, \ldots, K < \infty$ are the continuous martingales*

$$\mathfrak{D}_t^{i_N^k} = \left[\sum_{T_r \le t} \left(\Delta L_{T_r} + \int_{T_r}^t \sigma_s^X dW_s^X \right) \left(\sum_{l=1}^{i_N^k \wedge r} \left(1 - \frac{l}{i_N^k} \right) \Delta M_{T_{r-l}} \right) \right.$$
$$\left. + \sum_{T_r \le t} \left(\Delta M_{T_r} + \int_{T_r}^t \sigma_s^Y dW_s^Y \right) \left(\sum_{l=1}^{i_N^k \wedge r} \left(1 - \frac{l}{i_N^k} \right) \Delta L_{T_{r-l}} \right) \right]$$

with a sequence of integers $i_N^k, k = 1, \ldots, K$. On the Assumptions 1, 2 and 4 and if for every $k \in \{1, \ldots, K\}$ there exists a constant q_k with $i_N^k/M_N \to q_k$, the following stable convergence holds true as $N \to \infty, M_N \to \infty, M_N/N^{\alpha} \to 0$ for every $\alpha > 2/3$:

$$\sqrt{\frac{N}{M_N}} \mathbb{D}_t^N \overset{st}{\rightsquigarrow} \int_0^t w_s d\mathbb{W}_s \,, \tag{A.9}$$

with a K-dimensional Brownian motion $\mathbb{W}$ independent of $\mathcal{F}$ and a predictable process w_s with

$$(w_s w_s^*)_{mn} = \frac{T}{3}\min(q_m,q_n)\left(3-\frac{\min(q_m,q_n)}{\max(q_m,q_n)}\right)(1+\rho_s^2)\left(\sigma_s^X\sigma_s^Y\right)^2 G'(s) \tag{A.10}$$

with the convention that for $q_m = q_n = 0$ the ratio is one.
For $\mathbb{D}_T^N$ we obtain the following multivariate stable central limit theorem

$$\sqrt{\frac{N}{M_N}}\mathbb{D}_T^N \overset{st}{\rightsquigarrow} \mathbf{N}\left(0,\eta^2\Sigma\right)\ , \tag{A.11}$$

with $\eta^2 = 2T\int_0^T(1+\rho_t^2)(\sigma_t^X\sigma_t^Y)^2G'(t)dt$ and

$$\Sigma_{mn} = \frac{1}{6}\min(q_m,q_n)\left(3-\frac{\min(q_m,q_n)}{\max(q_m,q_n)}\right)\ .$$

Proof. Define for $k\in\{1,\ldots,K\}$ the continuous martingales

$$\mathfrak{M}_t^{i_N^k} = \sqrt{\frac{N}{M_N}}\mathfrak{D}_t^{i_N^k}\ .$$

By virtue of Proposition A.3, we already have that

$$\left[\mathfrak{M}^{i_N^k}\right]_t \xrightarrow{p} \frac{2}{3}Tq_k\int_0^t(1+\rho_s^2)(\sigma_s^X\sigma_s^Y)^2G'(s)\,ds\ .$$

The limit of the quadratic covariations $\left[\mathfrak{M}^{i_N^m},\mathfrak{M}^{i_N^k}\right]$ is derived using the same approximations as for the quadratic variation in the preceding proof:

$$\begin{aligned}
\left[\mathfrak{M}^{i_N^m},\mathfrak{M}^{i_N^k}\right]_t &= \frac{N}{M_N}\left[\sum_{T_r\le t}\Delta[L]_{T_r}\left(\sum_{l=1}^{\min(i_N^m,i_N^k,r)}\left(1-\frac{l}{i_N^m}\right)\left(1-\frac{l}{i_N^k}\right)\left(\Delta M_{T_{r-l}}\right)^2\right)\right.\\
&\quad+\sum_{T_r\le t}\Delta[M]_{T_r}\left(\sum_{l=1}^{\min(i_N^m,i_N^k,r)}\left(1-\frac{l}{i_N^m}\right)\left(1-\frac{l}{i_N^k}\right)\left(\Delta L_{T_{r-l}}\right)^2\right)\\
&\quad\left.+\sum_{T_r\le t}2\Delta[L,M]_{T_r}\left(\sum_{l=1}^{\min(i_N^m,i_N^k,r)}\left(1-\frac{l}{i_N^m}\right)\left(1-\frac{l}{i_N^k}\right)\Delta L_{T_{r-l}}\Delta M_{T_{r-l}}\right)\right]+\mathcal{O}_p(1)\\
&= N\sum_{T_r\le t}2\,\frac{G^{(N)}(T_r)-G^{(N)}(T_{r-1})}{\Delta T_r}(\rho_{T_{r-1}}^2+1)(\sigma_{T_{r-1}}^X\sigma_{T_{r-1}}^Y)^2\Delta T_r\\
&\qquad\times\left(\sum_{l=1}^{\min(i_N^m,i_N^k,r)}\left(1-\frac{l}{i_N^m}\right)\left(1-\frac{l}{i_N^k}\right)\right)+\mathcal{O}_p(1)\\
&\xrightarrow{p} 2T\int_0^t(\rho_s^2+1)(\sigma_s^X\sigma_s^Y)^2G'(s)\left(\frac{1}{6}\min(q_m,q_k)\left(3-\frac{\min(q_m,q_k)}{\max(q_m,q_k)}\right)\right)ds\ ,
\end{aligned}$$

since $\sum_{l=1}^m(1-(l/m))(1-(l/M)) = (1/2)m-(m^2/6M)-1/8+1/(12M)$ for $m,M\in\mathbb{Z}$. The multi-dimensional version of Jacod's stable central limit Theorem 2–1 from [18] enables us to prove the result of stable weak convergence of the vector provided we can verify the conditions

$$\left[\mathbb{D}^N,\mathbb{L}\right]_t \xrightarrow{p} 0\ \ ,\quad \left[\mathbb{D}^N,\mathbb{M}\right]_t \xrightarrow{p} 0,\ \ \forall t\in[0,T]\ ,$$

where $\mathbb{L}$ denotes the vector with entries $\mathbb{L}^j = L, j=1,\ldots,K$ and $\mathbb{M}$ with $\mathbb{M}^j = M, j=1,\ldots,K$, respectively, and

$$\left[\mathbb{D}^N,\mathbb{L}^\perp\right]_t \xrightarrow{p} 0\ \ ,\quad \left[\mathbb{D}^N,\mathbb{M}^\perp\right]_t \xrightarrow{p} 0,\ \ \forall t\in[0,T]$$

where $\mathbb{L}^\perp$ and $\mathbb{M}^\perp$ are bounded $(\mathcal{F}_t)$-adapted martingales orthogonal to $\mathbb{L}$ and $\mathbb{M}$, respectively. That is because the reference continuous martingales for all entries of the vector $\mathbb{D}^N$ are L and M. The componentwise proof of the conditions above is yet analogous as for the univariate case in the preceding proof. We conclude that the asymptotic distribution of the vector is described by a limiting time-changed Brownian motion on $[0,T]$, and the marginal distribution at time T by a mixed Gaussian limit, where the normal distribution is defined as well as for all componentwise marginals on an orthogonal extension of the original underlying probability space. □

From the preceding multivariate limit theorem the Cramér-Wold device allows to conclude the weak convergence of all one-dimensional linear combinations of the transformed discretization errors of a finite collection of one-scale subsampling estimators. For an asymptotically $\mathbf{N}(0,\Sigma)$-distributed random vector the sum of all components is asymptotically normally distributed with variance $\sum_{i,j}(\Sigma_{ij})$ by the Cramér-Wold device and the normality of any linear sum of components of a multivariate normal distribution (see e. g. pp. 516-517 in [25]).
The asymptotic variance in Proposition A.4 is deduced from the multivariate limit and

$$\sum_{k,l}(\Sigma_{k,l}) = 2\sum_{k=1}^{M_N}\sum_{l=1}^{k}\frac{l}{6M_N}\left(3-\frac{l}{k}\right)\alpha_{k,M_N}^{opt}\alpha_{l,M_N}^{opt} + \mathcal{O}(1) = \frac{13}{35} + \mathcal{O}(1)$$

with the weights (A.1) inserted.
For the completion of the proof of Propositions A.8 and hence A.4, it remains to extend the result for asymptotically infinitely many addends. This part of the proof can be adopted from [30] where a stable central limit theorem for a multiscale estimator for the integrated volatility in the univariate setting is proved. □

A.2.3. Discretization error due to the lack of synchronicity

Proposition A.10. *On the Assumptions 1 and 2, it holds true that*

$$\mathcal{A}_T^N = \frac{1}{i}\sum_{j=i}^{N-1}\left[\left(L_{g_j}-L_{T_j}\right)\left(\sum_{k=j-i+1}^{j}\Delta M_{T_k}\right) + \left(M_{\gamma_j}-M_{T_j}\right)\left(\sum_{k=j-i+1}^{j}\Delta L_{T_k}\right)\right.$$
$$\left. + \Delta L_{T_{j+1}}\left(\sum_{k=j-i+1}^{j}\left(M_{T_k}-M_{\lambda_{k+1}}\right)\right) + \Delta M_{T_{j+1}}\left(\sum_{k=j-i+1}^{j}\left(L_{T_k}-L_{l_{k+1}}\right)\right)\right] = \mathcal{O}_p\left(\sqrt{i}N\right).$$

for the error associated with interpolation errors $\mathcal{A}_T^N$ for a one-scale subsampling estimator.

Proof. $\mathcal{A}_T^N$ is the endpoint of a $\mathcal{F}_{j,N} = \mathcal{F}_{T_{j+1}^{(N)}}$-measurable discrete martingale with conditional expectation zero, since the addends incorporate products of Brownian increments over disjoint time intervals. The conditional variance yields

$$\frac{1}{i^2}\sum_{j=i}^{N-1}\mathbb{E}\left[\left(\left(L_{g_j}-L_{T_j}\right)\left(\sum_{k=j-i+1}^{j}\Delta M_{T_k}\right) + \left(M_{\gamma_j}-M_{T_j}\right)\left(\sum_{k=j-i+1}^{j}\Delta L_{T_k}\right)\right.\right.$$
$$\left.\left. + \Delta L_{T_{j+1}}\left(\sum_{k=j-i+1}^{j}\left(M_{T_k}-M_{\lambda_{k+1}}\right)\right) + \Delta M_{T_{j+1}}\left(\sum_{k=j-i+1}^{j}\left(L_{T_k}-L_{l_{k+1}}\right)\right)\right)^2 \Big| \mathcal{F}_{T_j^{(N)}}\right]$$
$$= \frac{1}{i^2}\sum_{j=i}^{N-1}\left(\mathbb{E}\left[(L_{g_j}-L_{T_j})^2\right]\left(\sum_{k=j-i+1}^{j}\Delta M_{T_k}\right)^2 + \mathbb{E}\left[(M_{\gamma_j}-M_{T_j})^2\right]\left(\sum_{k=j-i+1}^{j}\Delta L_{T_k}\right)^2\right.$$
$$\left. + \mathbb{E}\left[(\Delta L_{T_{j+1}})^2\right]\left(\sum_{k=j-i+1}^{j}\left(M_{T_k}-M_{\lambda_{k+1}}\right)\right)^2 + \mathbb{E}\left[(\Delta M_{T_{j+1}})^2\right]\left(\sum_{k=j-i+1}^{j}\left(L_{T_k}-L_{l_{k+1}}\right)\right)^2\right.$$

$$
\begin{aligned}
&+ \mathbb{E}\left[\int_{T_j}^{g_j} (\sigma_t^X)^2 dt\right] \left(\sum_{k=j-i+1}^{j} \left(M_{T_k} - M_{\lambda_{k+1}}\right)\right) \left(\sum_{k=j-i+1}^{j} \Delta M_{T_k}\right) \\
&+ \mathbb{E}\left[\int_{T_j}^{\gamma_j} (\sigma_t^Y)^2 dt\right] \left(\sum_{k=j-i+1}^{j} \left(L_{T_k} - L_{l_{k+1}}\right)\right) \left(\sum_{k=j-i+1}^{j} \Delta L_{T_k}\right) \\
&+ \mathbb{E}\left[\int_{T_j}^{g_j} \rho_t \sigma_t^X \sigma_t^Y dt\right] \left(\sum_{k=j-i+1}^{j} \left(L_{T_k} - L_{l_{k+1}}\right)\right) \left(\sum_{k=j-i+1}^{j} \Delta M_{T_k}\right) \\
&+ \mathbb{E}\left[\int_{T_j}^{\gamma_j} \rho_t \sigma_t^X \sigma_t^Y dt\right] \left(\sum_{k=j-i+1}^{j} \left(M_{T_k} - M_{\lambda_{k+1}}\right)\right) \left(\sum_{k=j-i+1}^{j} \Delta L_{T_k}\right)\Bigg) \\
&= \mathcal{O}_p\left(i^{-1}N^{-1}\right) .
\end{aligned}
$$

The variance of the term is of order $(iN)^{-1}$ which can be proved by taking the expectation of the above given conditional variance and an upper bound of the second moment. The asymptotic orders of the addends follow from taking the expectations using Itô isometry and analyzing the differences of the addends minus their expectations, that converge to zero at a faster rate. That part is similar to the proofs above and we forgo a more detailed computation here. □

Denote $A_T^{N,i}$ the error due to non-synchronicity and interpolations for a fixed subsampling frequency $i = 1, \ldots, M_N$ in the following. The error due to asynchronicity of the generalized multiscale estimator (12) equals the weighted sum $\sum_{i=1}^{M_N} \alpha_{i,M_N}^{opt} A_T^{N,i}$. It has expectation zero and the variance is of order

$$
\begin{aligned}
\mathbb{V}\mathrm{ar}\left(\sum_{i=1}^{M_N} \alpha_{i,M_N}^{opt} \mathcal{A}_T^{N,i}\right) &= \sum_{i,k} \alpha_{i,M_N}^{opt} \alpha_{k,M_N}^{opt} \mathbb{C}\mathrm{ov}\left(\mathcal{A}_T^{N,i}, \mathcal{A}_T^{N,k}\right) \\
&= \underbrace{\sum_{i=1}^{M_N} \left(\alpha_{i,M_N}^{opt}\right)^2 \mathbb{E}\left[\left(\mathcal{A}_T^{N,i}\right)^2\right]}_{=\mathcal{O}\left(M_N^{-2}N^{-1}\right)} + \underbrace{\sum_{i \neq k} \alpha_{i,M_N}^{opt} \alpha_{k,M_N}^{opt} \mathbb{E}\left[\mathcal{A}_T^{N,i} \mathcal{A}_T^{k,N}\right]}_{=\mathcal{O}\left(M_N^{-1}N^{-1}\right)} = o\left(\frac{M_N}{N}\right) .
\end{aligned}
$$

Thus, the error due to interpolations is of smaller asymptotic order than the discretization error of the closest synchronous approximation and asymptotically negligible.

A.3. Asymptotics of the cross term

For a one-scale subsampling estimator cross terms are asymptotically negligible and hence the stable central limit theorem in Theorem 4.1 is implied by Theorem A.3. For the proof of the stable central limit theorem in Theorem 3 for the multiscale approach, we cope with the asymptotics of the cross terms in this subsection.

Proposition A.11. *On the Assumptions 1, 2, 3 and 5, the cross terms of the generalized multiscale estimator* (12) *with noise-optimal weights* (A.1) *weakly converge to a mixed normal limit as* $M_N \to \infty$, $N \to \infty$, $M_N \delta_N \to 0$*:*

$$
\sqrt{M_N} \sum_{i=1}^{M_N} \frac{\alpha_{i,M_N}^{opt}}{i} \sum_{j=i}^{N} \left((X_{g_j} - X_{l_{j-i+1}})(\epsilon_{\gamma_j}^Y - \epsilon_{\lambda_{j-i+1}}^Y) + (Y_{\gamma_j} - Y_{\lambda_{j-i+1}})(\epsilon_{g_j}^X - \epsilon_{l_{j-i+1}}^X)\right)
$$

$$
\rightsquigarrow \mathbf{N}\left(0, \mathbf{AVAR}_{cross}\right) , \quad \text{(A.12)}
$$

with asymptotic variance

$$
\mathbf{AVAR}_{cross} = \frac{12}{5}\left(\eta_Y^2 \int_0^T (1 + T\, I_Y'(t))(\sigma_t^X)^2\, dt + \eta_X^2 \int_0^T (1 + T\, I_X'(t))(\sigma_t^Y)^2\, dt\right) . \quad \text{(A.13)}
$$

The convergence holds conditionally given the paths of the efficient processes.

Proof. This proof affiliates to the discussion in Section 4, where degrees of regularity of non-synchronous sampling schemes have been defined in Definition 2 that are assumed to converge to continuously differentiable functions.
On the Assumption 3 of independent observation noise of X and Y, the two different cross terms are uncorrelated and we prove a central limit theorem for the first one:

$$\sqrt{M_N}\sum_{i=1}^{M_N}\frac{\alpha^{opt}_{i,M_N}}{i}\sum_{j=i}^{N}(X_{g_j}-X_{l_{j-i+1}})(\epsilon^Y_{\gamma_j}-\epsilon^Y_{\lambda_{j-i+1}})\rightsquigarrow \mathbf{N}\left(0,\frac{12}{5}\eta_Y^2\int_0^T(1+T\,I'_Y(t))(\sigma^X_t)^2\,dt\right)\,.$$

The parallel result for the other term can be proved analogously.
For the purpose of a shorter notation we have left out superscripts of the observation times, and write α_i, $i=1,\dots,M_N$ for the weights although we are interested in the specific weights (A.1). Denote $\delta_N=\sup_{i\in\{1,\dots,N\}}\Delta T_i$ and $\gamma_{j,+}=\min\left(\tau_k\in\mathcal{T}^Y|\tau_k\in\mathcal{G}^{j+1}\right), g_{j,+}=\min\left(t_k\in\mathcal{T}^X|t_k\in\mathcal{H}^{j+1}\right)$ and C a generic constant as before. From

$$\begin{aligned}
&\mathbb{E}\left[\left(\sqrt{M_N}\sum_{i=1}^{M_N}\frac{\alpha_i}{i}\sum_{j=i}^{N}(X_{g_j}-X_{T_j})(\epsilon^Y_{\gamma_j}-\epsilon^Y_{\lambda_{j-i+1}})+(X_{T_{j-i}}-X_{l_{j-i+1}})(\epsilon^Y_{\gamma_j}-\epsilon^Y_{\lambda_{j-i+1}})\right)^2\right]\\
&\leq M_N\sum_{i,k\in\{1,\dots,M_N\}}\frac{\alpha_i\alpha_k}{ik}2\eta_Y^2\left(\sum_{j=i\vee k}^{N}\mathbb{E}(X_{g_j}-X_{T_j})^2+\sum_{j=0}^{N-(i\vee k)}\mathbb{E}(X_{T_j}-X_{l_{j+1}})^2\right)\\
&\leq M_N\,C4\eta_Y^2\sum_{i,k\in\{1,\dots,M_N\}}\frac{\alpha_i\alpha_k}{ik}=\mathcal{O}\left(M_N^{-1}\right)\,,
\end{aligned}$$

for the errors due to interpolations and

$$\begin{aligned}
&\mathbb{E}\left[\left(\sqrt{M_N}\sum_{i=1}^{M_N}\frac{\alpha_i}{i}\left(\sum_{k=N-i+1}^{N}\epsilon^Y_{\gamma_k}(X_{T_k}-X_{T_{k-i}})-\sum_{k=1}^{i}\epsilon^Y_{\lambda_k}(X_{T_{k+i}}-X_{T_k})\right)\right)^2\right]\\
&=M_N\sum_{i,k\in\{1,\dots,M_N\}}\frac{\alpha_i\alpha_k}{ik}\eta_Y^2\left(\sum_{r=N-(i\wedge k)+1}^{N}\mathbb{E}(X_{T_r}-X_{T_{r-i}})^2+\sum_{r=1}^{i\wedge k}\mathbb{E}(X_{T_{r+i}}-X_{T_r})^2\right)\\
&=\mathcal{O}\left(M_N\delta_N\right)
\end{aligned}$$

for boundary terms, we conclude that

$$\begin{aligned}
&\sqrt{M_N}\sum_{i=1}^{M_N}\frac{\alpha_i}{i}\sum_{j=i}^{N}(X_{g_j}-X_{l_{j-i+1}})(\epsilon^Y_{\gamma_j}-\epsilon^Y_{\lambda_{j-i+1}})\\
&=\sqrt{M_N}\sum_{i=1}^{M_N}\frac{\alpha_i}{i}\left(\sum_{j=i}^{N-i}\epsilon^Y_{\gamma_j}(X_{T_j}-X_{T_{j-i}})-\epsilon^Y_{\lambda_{j+1}}(X_{T_{j+i}}-X_{T_j})\right)+\mathcal{O}_p(1)\\
&=\sqrt{M_N}\sum_{j=2}^{N-2}\left(\epsilon^Y_{\gamma_j}\sum_{i=1}^{M^*_N(j)}\frac{\alpha_i}{i}(X_{T_j}-X_{T_{j-i}})-\epsilon^Y_{\lambda_{j+1}}\sum_{i=1}^{M^*_N(j)}(X_{T_{j+i}}-X_{T_j})\right)+\mathcal{O}_p(1)\\
&=\sqrt{M_N}\left(\sum_{j\in\mathcal{Y}_1}\epsilon^Y_{\gamma_j}\sum_{i=1}^{M^*_N(j)}\frac{\alpha_i}{i}\zeta^1_{i,j}+\sum_{j\in\mathcal{Y}_2}\epsilon^Y_{\gamma_j}\sum_{i=1}^{M^*_N(j)}\frac{\alpha_i}{i}\zeta^2_{i,j}+\sum_{j\in\mathcal{Y}_3}\epsilon^Y_{\gamma_j}\sum_{i=1}^{M^*_N(j)}\frac{\alpha_i}{i}\zeta^3_{i,j}\right.\\
&\qquad\left.+\sum_{j\in\mathcal{Y}_4}\epsilon^Y_{\gamma_j}\sum_{i=1}^{M^*_N(j)}\frac{\alpha_i}{i}\zeta^{4a}_{i,j}-\sum_{j\in\mathcal{Y}_4}\epsilon^Y_{\gamma_{j,+}}\sum_{i=1}^{M^*_N(j)}\frac{\alpha_i}{i}\zeta^{4b}_{i,j}\right)+\mathcal{O}_p(1)\,.
\end{aligned}$$

Here, we aggregate the observation times $\gamma_j, \lambda_j,\ j = 2, \dots, N-2$ in disjoint sets conforming to the four cases discussed in Section 4. Denote thereto

$$\begin{aligned}
\mathcal{Y}_1 &= \{j \in \{2, \dots, N_2\} | \gamma_j \neq \gamma_{j-1}\,,\, \gamma_j \leq g_j\}\,,\\
\mathcal{Y}_2 &= \{j \in \{2, \dots, N_2\} | \gamma_j > g_j\,,\, \gamma_j \geq g_{j,+}\}\,,\\
\mathcal{Y}_3 &= \{j \in \{2, \dots, N_2\} | \gamma_j > g_j\,,\, \gamma_j < g_{j,+}\,,\, \gamma_{j,+} > g_{j,+}\}\,,\\
\mathcal{Y}_4 &= \{j \in \{2, \dots, N_2\} | \gamma_j > g_j\,,\, \gamma_j < g_{j,+}\,,\, \gamma_{j,+} \leq g_{j,+}\}\,,
\end{aligned}$$

and $M_N^*(j) = \min(j, N-j, M_N)$. The increments of X that are multiplied with each observation error differ according to the set $\mathcal{Y}_k,\ 1 \leq k \leq 4$ to which γ_j belongs. We use the notation

$$\begin{aligned}
\zeta_{i,j}^1 &= (X_{T_j} - X_{T_{j-i}}) - (X_{T_{j+i}} - X_{T_j})\,,\\
\zeta_{i,j}^2 &= (X_{T_j} - X_{T_{j-i}}) + (X_{T_{j+1}} - X_{T_{j-i+1}}) - (X_{T_{j+i+1}} - X_{T_{j+1}})\,,\\
\zeta_{i,j}^3 &= (X_{T_j} - X_{T_{j-i}}) - (X_{T_{j+i+1}} - X_{T_{j+1}})\,,\\
\zeta_{i,j}^{4a} &= (X_{T_j} - X_{T_{j-i}})\,,\ \zeta_{i,j}^{4b} = (X_{T_{j+i+1}} - X_{T_{j+1}})\,.
\end{aligned}$$

The resulting aggregated leading term above of the cross term is the endpoint of a discrete martingale with respect to the filtration $\mathcal{F}_{j,N} := \sigma\big(\epsilon^Y_{\tau_k} | \tau_k < \gamma_{j+1}\,,\, X, Y\big)$. Since if $j \in \mathcal{Y}_4 \Rightarrow \gamma_{j,+} < \gamma_{j+1}$, the martingale property with respect to the filtration $\mathcal{F}_{j,N}$ is assured by Assumption 3.
An application of the non-stable version of the central limit theorem for martingale triangular arrays from [14] will proof the asymptotic normality of the cross term conditionally on the paths of the efficient processes. The conditional Lindeberg condition can be verified (using Chebyshev's inequality or directly verifying the conditional Lyapunov condition) in the same way as before and we omit it here. The sum of conditional variances yields

$$\begin{aligned}
&\sum_{l \in \{1,2,3,4a\}} \left(\sum_{j \in \mathcal{Y}_l} \mathbb{E}\Big[\Big(\sqrt{M_N} \epsilon^Y_{\gamma_j} \sum_{i=1}^{M_N^*(j)} \frac{\alpha_i}{i} \zeta_{i,j}^l \Big)^2 \Big| \mathcal{F}_{j-1,N} \Big] + \sum_{j \in \mathcal{Y}_4} \mathbb{E}\Big[\Big(-\sqrt{M_N} \epsilon^Y_{\gamma_{j,+}} \sum_{i=1}^{M_N^*(j)} \frac{\alpha_i}{i} \zeta_{i,j}^{4b} \Big)^2 \Big| \mathcal{F}_{j-1,N} \Big] \right)\\
&= M_N \eta_Y^2 \left(\sum_{j \in \mathcal{Y}_1 \cup \mathcal{Y}_3 \cup \mathcal{Y}_4} \left(\sum_{i=1}^{M_N^*(j)} \frac{\alpha_i}{i} \zeta_{i,j}^1 \right)^2 + \sum_{j \in \mathcal{Y}_2} \left(\sum_{i=1}^{M_N^*(j)} \frac{\alpha_i}{i} \zeta_{i,j}^2 \right)^2 \right) + \mathcal{O}_p(1)\\
&= M_N \eta_Y^2 \left(\sum_{j \in \mathcal{Y}_1 \cup \mathcal{Y}_3 \cup \mathcal{Y}_4} \sum_{i,k \in \{1, \dots, M_N^*(j)\}} \frac{\alpha_i \alpha_k}{ik} \left(\zeta_{i \wedge k, j}^1 \right)^2 + \sum_{j \in \mathcal{Y}_2} \left(\sum_{i,k \in \{1, \dots, M_N^*(j)\}} \frac{\alpha_i \alpha_k}{ik} \left(\zeta_{i \wedge k, j}^2 \right)^2 \right) \right) + \mathcal{O}_p(1)\\
&= M_N \eta_Y^2 \left(\sum_{j \in \mathcal{Y}_1 \cup \mathcal{Y}_3 \cup \mathcal{Y}_4} \sum_{i,k \in \{1, \dots, M_N^*(j)\}} \frac{\alpha_i \alpha_k}{ik} \left((X_{T_j} - X_{T_{j-(i \wedge k)}})^2 + (X_{T_{j+(i \wedge k)}} - X_{T_j})^2 \right) \right.\\
&\qquad \left. + \sum_{j \in \mathcal{Y}_2} \sum_{i,k \in \{1, \dots, M_N^*(j)\}} \frac{\alpha_i \alpha_k}{ik} \left(4(X_{T_j} - X_{T_{j-(i \wedge k)}})^2 + (X_{T_{j+(i \wedge k)}} - X_{T_j})^2 \right) \right) + \mathcal{O}_p(1)\\
&= M_N \eta_Y^2 \sum_{j=2}^{N-2} \sum_{i,k \in \{1, \dots, M_N^*(j)\}} \frac{\alpha_i \alpha_k}{ik} (2 + \mathbb{1}_{\{j \in \mathcal{Y}_2\}}) (X_{T_j} - X_{T_{j-(i \wedge k)}})^2 + \mathcal{O}_p(1)\\
&= M_N \eta_Y^2 \left(\sum_{i,k \in \{1, \dots, M_N^*(j)\}} \frac{\alpha_i \alpha_k}{ik} \left(2(i \wedge k) \widehat{[X]}_T^{sub, i \wedge k} + \sum_{j=i \wedge k}^{N} \mathbb{1}_{\{j \in \mathcal{Y}_2\}} (X_{T_j} - X_{T_{j-(i \wedge k)}})^2 \right) \right) + \mathcal{O}_p(1)\\
&\xrightarrow{p} \frac{12}{5} \eta_Y^2 \left([X]_T + \int_0^T T\, I_Y'(t) (\sigma_t^X)^2 \, dt \right)\,.
\end{aligned}$$

Since for the shifted increments

$$(X_{T_{j+i+1}} - X_{T_{j+1}}) = (X_{T_{j+i}} - X_{T_j}) + \mathcal{O}_p\left(N^{-1/2} \right)$$

holds, where the order is for time instants of average length N^{-1}, the variances of the sums over all $j \in \mathcal{Y}_1$ and $j \in \mathcal{Y}_3$ are asymptotically equal. The variance of both uncorrelated sums over maxima γ_j and minima $\gamma_{j,+}$ distributed according to the fourth case is also asymptotically equal to the variances of those two addends. Only the asymptotic variance of the sum over all $j \in \mathcal{Y}_2$ is bigger. For this reason the total asymptotic variance hinges on the asymptotic degree of regularity of the non-synchronous sampling scheme $(\mathcal{T}^X, \mathcal{T}^Y)$ defined in Definition 2.
In the calculation of the asymptotic variance we have used that

$$\zeta^1_{i,j}\zeta^1_{i,k} = \left(\zeta^1_{i\wedge k,j}\right)^2 + \zeta^1_{i\wedge k,j}\left(\sum_{l=j-(i\vee k)+1}^{j-(i\wedge k)} \Delta X_{T_l} + \sum_{l=j+(i\wedge k)+1}^{j+(i\vee k)} \Delta X_{T_l}\right) ,$$

where the second remainder addend has an expectation equal to zero, and analogous formulae for $\zeta^2_{i,j}$, for all $1 \leq i \leq M_N\,, 1 \leq k \leq M_N\,,\, k \vee i \leq j \leq N - (i \vee k)$.
Furthermore, an application of the mean value theorem, Itô isometry and approximations in the same spirit as in the calculation of the asymptotic variance in the proof of the central limit theorem for the discretization errors of the estimators, lead to the Riemann sum in the calculation of the asymptotic variance above. The cross terms in $(\zeta^l_{i,j})^2$, $l = 1, 2$ are asymptotically negligible. Since in $\mathcal{Y}_4$ repeating maxima $\gamma_i = \gamma_{i+1}$ are considered only once, it holds true that $|\mathcal{Y}_1| + |\mathcal{Y}_3| + |\mathcal{Y}_4| + 2|\mathcal{Y}_2| = N - 3 \pm 1$ (the last addend can appear due to boundary term effects). In the last step we have used that

$$M_N \sum_{i,k\in\{1,\ldots,M_N\}} \frac{\alpha^{opt}_{i,M_N}\alpha^{opt}_{k,M_N}}{ik}(i \wedge k) = 6/5 + o(1)$$

when inserting the weights (A.1).
From the analysis for the asymptotic discretization error of a one-scale subsampling estimator, we know that

$$\widehat{[X]}^{sub,i\wedge k}_T = \frac{1}{i \wedge k}\sum_{j=i\wedge k}^{N} (X_{T_j} - X_{T_{j-(i\wedge k)}})^2 = [X]_T + \mathcal{O}_p\left(\sqrt{\frac{(i \wedge k)}{N}}\right)$$

holds true. Similarly, it can be deduced that

$$\begin{aligned}\frac{1}{i}\sum_{j=i}^{N} \mathbb{1}_{\{j\in\mathcal{Y}_2\}}(X_{T_j} - X_{T_{j-i}})^2 &= \frac{1}{i}\sum_{l=1}^{N} (\Delta X_{T_l})^2\left(\sum_{k=1}^{i} \mathbb{1}_{\{(k+l-1)\in\mathcal{Y}_2\}}\right) + \mathcal{O}_p\left(\sqrt{i/N}\right)\\ &= \int_0^T T\, I'_Y(t)(\sigma^X_t)^2\, dt + \mathcal{O}_p\left(\sqrt{i/N}\right) ,\end{aligned}$$

on Assumption 5. □

B. Proof of Proposition 5.1

Let R^N_j, R^n_j, R^m_j, $S^{N,X}_j$ and $S^{N,Y}_j$ denote the number of times $T^{(N)}_k, 0 \leq k \leq N$, $t^{(n)}_i$ and $\tau^{(m)}_j$ in the bins $[G^N_j, G^N_{j+1})$, $[(I_Y)^N_j, (I_Y)^N_{j+1})$, $[(I_X)^N_j, (I_X)^N_{j+1}), 0 \leq j \leq K_N - 1$,. Define the generalized multiscale estimator in the fashion of (12)

$$\widehat{\Delta\,[X,Y]}_{G^N_{j+1}} = \sum_{i=1}^{M_N(j)} \frac{\alpha^{opt}_{i,M_N(j)}}{i}\sum_{r=i}^{R^N_j}\left(\tilde{X}_{g_r} - \tilde{X}_{l_{r-i+1}}\right)\left(\tilde{Y}_{\gamma_r} - \tilde{Y}_{\lambda_{r-i+1}}\right)$$

for the increase of the quadratic covariation $\Delta\,[X,Y]_{G^N_{j+1}}$ and the univariate multiscale estimators

$$\widehat{\Delta\,[X]}_{G^N_{j+1}} = \sum_{i=1}^{M_n(j)} \frac{\alpha^{opt}_{i,M_n(j)}}{i}\sum_{r=i}^{R^n_j}\left(\tilde{X}_{t_r} - \tilde{X}_{t_{r-i}}\right)^2,\ \widehat{\Delta\,[Y]}_{G^N_{j+1}} = \sum_{i=1}^{M_m(j)} \frac{\alpha^{opt}_{i,M_m(j)}}{i}\sum_{r=i}^{R^m_j}\left(\tilde{Y}_{\tau_r} - \tilde{Y}_{\tau_{r-i}}\right)^2 ,$$

and $\widehat{\Delta\,[X]}_{(I_Y)_{j+1}^N}, \widehat{\Delta\,[Y]}_{(I_X)_{j+1}^N}$ analogously, where all binwise multiscale frequencies are of order $\sqrt{NK_N}$ with possibly differing constants. Essential when considering the multiscale estimators on bins is that on Assumption 4 the distances between sampling times are of order $N^{-1} \sim n^{-1} \sim m^{-1}$, whereas the numbers of observations $R_j^{\cdot}, S_j^{\cdot}, 1 \leq j \leq K_N$ in the specific bin are at most of order NK_N^{-1}. Following the analysis for the four uncorrelated parts of the estimation error in sections Appendix A, orders of the discretization variances yield

$$\sum_{i,k\in\{1,\ldots,M_N(j)\}} \frac{\alpha_{i,M_N(j)}^{opt}\alpha_{k,M_N(j)}^{opt}}{ik} \cdot i \cdot R_j^N \frac{i^2}{N^2} \sim M_N(j)\frac{R_j^N}{N^2} \sim \frac{M_N(j)}{K_N N} \,,$$

$M_n(j)/(nK_N)$, $M_m(j)/(mK_N)$, $M_n(j)\frac{S_j^{N,X}}{N^2}$ and $M_n(j)\frac{S_j^{N,Y}}{N^2}$, respectively. Cross terms are of order $R_j^N/(NM_N(j)) \sim (M_N(j)K_N)^{-1}$ and analogous orders for the univariate estimators. The errors due to noise instead depend only on the number of observations in the considered interval. Therefore, the addends are of orders $R_j^N/M_N^3(j) \sim N/\left(K_N M_N^3(j)\right)$ and $M_N^{-1}(j)$ and analogous for the univariate estimators.
Choosing all multiscale frequencies $M_\cdot(j) \sim N^{1/2}K_N^{1/2}$ for every j, so that $M_N(\cdot)N^{1/2} \to \infty$, the error due to end-effects in the noise part and the discretization error dominate asymptotically the two other addends and are of order $N^{-1/4}K_N^{-1/4}$. This holds as long as $K_N N^{-1/3} \to 0$, such that $M_N(j)(N/K_N)^{-1} \to 0$ as $N \to \infty$.
The estimators (22a)-(22d) are consistent as $K_N \to \infty$ with $K_N N^{-1/3} \to 0$ as $N \to \infty$, since

$$\begin{aligned}
\hat{I}_2 &= \sum_{j=1}^{K_N} \left(\frac{\widehat{\Delta\,[X]}_{G_j^N}\widehat{\Delta\,[Y]}_{G_j^N}}{\left(\Delta G_j^N\right)^2} \right) \frac{G^N(T)}{K_N} \\
&= \sum_{j=1}^{K_N} \left(\frac{\int_{G_{j-1}^N}^{G_j^N} \left(\sigma_t^X\right)^2 \, dt \int_{G_{j-1}^N}^{G_j^N} \left(\sigma_t^Y\right)^2 \, dt + \mathcal{O}_p\left(N^{-1/4}K_N^{-1/4}\right)}{\Delta G_j^N} \right)^2 \frac{G^N(T)}{K_N} \\
&= \sum_{j=1}^{K_N} \left(\sigma^X\right)^2_{\overline{G_j^N}} \left(\sigma^Y\right)^2_{\widetilde{G_j^N}} \frac{G^N(T)}{K_N} + \mathcal{O}_p\left(K_N^{1/4}N^{-1/4}\right) = \sum_{j=1}^{K_N} \left(\sigma^X\sigma^Y\right)^2_{G_{j-1}^N} \frac{G^N(T)}{K_N} + \mathcal{O}_p\left(K_N^{1/4}N^{-1/4}\right)
\end{aligned}$$

and similar conclusions for the other three estimators. We have used that $\Delta G_j^N \sim N^{-1}$ and apply the mean value theorem. $\overline{G_j^N}$ is some value $G_{j-1}^N \leq \overline{G_j^N} \leq G_j^N$. Finally, elementary inequalities $\left|\left(\sigma^X\right)^2_{\overline{G_j^N}}\left(\sigma^Y\right)^2_{\widetilde{G_j^N}} - \left(\sigma^X\sigma^Y\right)^2_{G_{j-1}^N}\right| \leq \left(\sigma^X\right)^2_{\overline{G_j^N}}\left|\left(\sigma^Y\right)^2_{\widetilde{G_j^N}} - \left(\sigma^Y\right)^2_{G_{j-1}^N}\right| + \left(\sigma^Y\right)^2_{G_{j-1}^N}\left|\left(\sigma^X\right)^2_{\overline{G_j^N}} - \left(\sigma^X\right)^2_{G_{j-1}^N}\right|$,
$\sum_{j=1}^{K_N}\left|\left(\rho\sigma^X\sigma^Y\right)^2_{\overline{G_j^N}} - \left(\rho\sigma^X\sigma^Y\right)^2_{G_{j-1}^N}\right| \frac{G^N(T)}{K_N} \leq \sup_{|t-s|\leq\Delta\sup_j G_j^n}\left|\rho_t\sigma_t^X\sigma_t^Y - \rho_s\sigma_s^X\sigma_s^Y\right| G^N(T) = \mathcal{O}_{a.\,s.}(1)$
are involved in the approximations of the type above. Considering (22c) and (22d), note that bin-widths chosen accordingly to I_Y^N are asymptotically of order K_N^{-1} in any interval of $[0,T]$ on that the corresponding part of the integral $\int I_Y'(t)(\sigma_t^X)^2\,dt$ is strictly positive.
Denote R_N^k, $k = 1,\ldots,4$, the orders of the approximation errors of the four above given integrals and their Riemann sums evaluated on the partition given K_N bins. The variance of the estimators $\widehat{\eta_X^2}$ and $\widehat{\eta_Y^2}$ for the noise variances are known to be $\mathbb{E}\left[\left(\epsilon_{t_1}^X\right)^4\right] N^{-1}$ and $\mathbb{E}\left[\left(\epsilon_{\tau_1}^Y\right)^4\right] N^{-1}$ and hence $\mathcal{O}\left(N^{-1}\right)$ on Assumption 3 from [31]. From

$$\hat{I}_k = I_k + \mathcal{O}_p\left(R_N^k + K_N^{1/2}N^{-1/2}\right) \,,\ k = 1,\ldots,4\,,$$

we derive that

$$\widehat{\mathbf{AVAR}}_{multi} = \mathbf{AVAR}_{multi} + \mathcal{O}_p\left(\max_k R_N^k + K_N^{1/2}N^{-1/2}\right) \,,$$

and the same result for the one-scale estimator. □

References

[1] Y. Aït-Sahalia, J. Fan, D. Xiu, High-frequency estimates with noisy and asynchronous financial data, Journal of the American Statistical Association 105 (2010) 1504–1516.

[2] Y. Aït-Sahalia, L. Zhang, P.A. Mykland, Ultra high frequency volatility estimation with dependent microstructure noise, Journal of Econometrics, 160 (2011) 160–165.

[3] B. Awartani, V. Corradi, W. Distasor, Assessing market microstructure effects via realized volatility measures with an application to the dow jones industrial average stocks, Journal of Business and Economic Statistics 27 (2009) 251–265.

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

[5] 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).

[6] 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.

[7] M. Bibinger, Asymptotics of asynchronicity, Technical Report, Humboldt-Universität zu Berlin (2011)
URL=`http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2011-033.pdf`.

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

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

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

[11] J. Fan, Y. Wang, Multi-scale jump and volatility analysis for high-frequency data, Journal of the American Statistical Association 102 (2007) 1349–1362.

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

[13] A. Gloter, J. Jacod, Diffusions with measurement errors 1 and 2, ESAIM Probability and Statistics 5 (2001) 225–242.

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

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

[16] 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.

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

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

[19] J. Jacod, Y. Li, P.A. Mykland, M. Podolskij, M. Vetter, Microstructure noise in the continous case: the pre-averaging approach, Stochastic Processes and their Applications 119 (2009) 2249–2276.

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

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

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

[23] M. Podolskij, M. Vetter, Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps, Bernoulli 15 (2009) 634–658.

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

[25] C.R. Rao, Linear Statistical Inference and its Applications, Wiley, New York, 2 edition, 2001.

[26] M. Reiß, Asymptotic equivalence for inference on the volatility from noisy observations, Annals of Statistics, Forthcoming (2011).

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

[28] S. Utev, Central limit theorem for dependent random variables, In Grigelionis et al., Probability theory and mathematical statistics 2 (1990) 519–528.

[29] D. Xiu, Quasi-maximum likelihood estimation of volatility with high frequency data, Journal of Econometrics 159 (2010) 235–250.

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

[31] L. Zhang, P.A. Mykland, Y. Aït-Sahalia, A tale of two time scales: Determining integrated volatility with noisy high-frequency data, Journal of the American Statistical Association 100 (2005) 1394–1411.

SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Localising temperature risk" by Wolfgang Karl Härdle, Brenda López Cabrera, Ostap Okhrin and Weining Wang, January 2011.
002 "A Confidence Corridor for Sparse Longitudinal Data Curves" by Shuzhuan Zheng, Lijian Yang and Wolfgang Karl Härdle, January 2011.
003 "Mean Volatility Regressions" by Lu Lin, Feng Li, Lixing Zhu and Wolfgang Karl Härdle, January 2011.
004 "A Confidence Corridor for Expectile Functions" by Esra Akdeniz Duran, Mengmeng Guo and Wolfgang Karl Härdle, January 2011.
005 "Local Quantile Regression" by Wolfgang Karl Härdle, Vladimir Spokoiny and Weining Wang, January 2011.
006 "Sticky Information and Determinacy" by Alexander Meyer-Gohde, January 2011.
007 "Mean-Variance Cointegration and the Expectations Hypothesis" by Till Strohsal and Enzo Weber, February 2011.
008 "Monetary Policy, Trend Inflation and Inflation Persistence" by Fang Yao, February 2011.
009 "Exclusion in the All-Pay Auction: An Experimental Investigation" by Dietmar Fehr and Julia Schmid, February 2011.
010 "Unwillingness to Pay for Privacy: A Field Experiment" by Alastair R. Beresford, Dorothea Kübler and Sören Preibusch, February 2011.
011 "Human Capital Formation on Skill-Specific Labor Markets" by Runli Xie, February 2011.
012 "A strategic mediator who is biased into the same direction as the expert can improve information transmission" by Lydia Mechtenberg and Johannes Münster, March 2011.
013 "Spatial Risk Premium on Weather Derivatives and Hedging Weather Exposure in Electricity" by Wolfgang Karl Härdle and Maria Osipenko, March 2011.
014 "Difference based Ridge and Liu type Estimators in Semiparametric Regression Models" by Esra Akdeniz Duran, Wolfgang Karl Härdle and Maria Osipenko, March 2011.
015 "Short-Term Herding of Institutional Traders: New Evidence from the German Stock Market" by Stephanie Kremer and Dieter Nautz, March 2011.
016 "Oracally Efficient Two-Step Estimation of Generalized Additive Model" by Rong Liu, Lijian Yang and Wolfgang Karl Härdle, March 2011.
017 "The Law of Attraction: Bilateral Search and Horizontal Heterogeneity" by Dirk Hofmann and Salmai Qari, March 2011.
018 "Can crop yield risk be globally diversified?" by Xiaoliang Liu, Wei Xu and Martin Odening, March 2011.
019 "What Drives the Relationship Between Inflation and Price Dispersion? Market Power vs. Price Rigidity" by Sascha Becker, March 2011.
020 "How Computational Statistics Became the Backbone of Modern Data Science" by James E. Gentle, Wolfgang Härdle and Yuichi Mori, May 2011.
021 "Customer Reactions in Out-of-Stock Situations – Do promotion-induced phantom positions alleviate the similarity substitution hypothesis?" by Jana Luisa Diels and Nicole Wiebach, May 2011.

SFB 649, Ziegelstraße 13a, D-10117 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

022 "Extreme value models in a conditional duration intensity framework" by Rodrigo Herrera and Bernhard Schipp, May 2011.
023 "Forecasting Corporate Distress in the Asian and Pacific Region" by Russ Moro, Wolfgang Härdle, Saeideh Aliakbari and Linda Hoffmann, May 2011.
024 "Identifying the Effect of Temporal Work Flexibility on Parental Time with Children" by Juliane Scheffel, May 2011.
025 "How do Unusual Working Schedules Affect Social Life?" by Juliane Scheffel, May 2011.
026 "Compensation of Unusual Working Schedules" by Juliane Scheffel, May 2011.
027 "Estimation of the characteristics of a Lévy process observed at arbitrary frequency" by Johanna Kappus and Markus Reiß, May 2011.
028 "Asymptotic equivalence and sufficiency for volatility estimation under microstructure noise" by Markus Reiß, May 2011.
029 "Pointwise adaptive estimation for quantile regression" by Markus Reiß, Yves Rozenholc and Charles A. Cuenod, May 2011.
030 "Developing web-based tools for the teaching of statistics: Our Wikis and the German Wikipedia" by Sigbert Klinke, May 2011.
031 "What Explains the German Labor Market Miracle in the Great Recession?" by Michael C. Burda and Jennifer Hunt, June 2011.
032 "The information content of central bank interest rate projections: Evidence from New Zealand" by Gunda-Alexandra Detmers and Dieter Nautz, June 2011.
033 "Asymptotics of Asynchronicity" by Markus Bibinger, June 2011.
034 "An estimator for the quadratic covariation of asynchronously observed Itô processes with noise: Asymptotic distribution theory" by Markus Bibinger, June 2011.

SFB 649, Ziegelstraße 13a, D-10117 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".