SFB 649 Discussion Paper 2014-005

Functional stable limit theorems for efficient spectral covolatility estimators

Randolf Altmeyer*
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

Functional stable limit theorems for efficient spectral covolatility estimators

Randolf Altmeyer & Markus Bibinger[1]

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

Abstract

We consider noisy non-synchronous discrete observations of a continuous semimartingale. Functional stable central limit theorems are established under high-frequency asymptotics in three setups: one-dimensional for the spectral estimator of integrated volatility, from two-dimensional asynchronous observations for a bivariate spectral covolatility estimator and multivariate for a local method of moments. The results demonstrate that local adaptivity and smoothing noise dilution in the Fourier domain facilitate substantial efficiency gains compared to previous approaches. In particular, the derived asymptotic variances coincide with the benchmarks of semiparametric Cramér-Rao lower bounds and the considered estimators are thus asymptotically efficient in idealized sub-experiments. Feasible central limit theorems allowing for confidence are provided.

Keywords: adaptive estimation, asymptotic efficiency, local parametric estimation, microstructure noise, integrated volatility, non-synchronous observations, spectral estimation, stable limit theorem

AMS 2000 subject classification: 62G05, 62G20, 62M10
JEL classes: C14, C32

1. Introduction

The estimation of integrated volatility and integrated covolatility (sometimes referred to as integrated variance and covariance) from high-frequency data is a vibrant current research topic in statistics for stochastic processes. Semimartingale log-price models are widely used in financial economics. Its pivotal role for portfolio optimization and risk management makes the integrated covolatility matrix a key quantity of interest in econometrics and finance. During the last decade the increasing availability of data from high-frequency recordings of trades or orders has provided the statistician with rich data sets. Yet, the consequence of this richness of data is double-edged. On the one hand, high-frequency observations are almost close to continuous-time observations which should allow for a very precise and highly efficient semiparametric estimation of integrated covolatilities. On the other hand, it has turned out that a traditional pure semimartingale model has several limitations in describing stylized facts of the high-frequency data and therefore does not afford a suitable risk estimation. It is nowadays well-known that effects ascribed to market microstructure frictions interfere at high frequencies with a latent price evolution that can be appropriately described by a semimartingale. Also, non-synchronous observations in a multi-dimensional setup require a thorough handling or can cause unwanted effects.

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

A prominent model that accounts for market microstructure in high-frequency observations is an additive noise model, in which a continuous semimartingale is observed with i.i.d. observation errors. The problem of estimating the integrated volatility in the one-dimensional model with noise, as well as multi-dimensional covolatility matrix estimation from *noisy and non-synchronous* observations, have attracted a lot of attention in recent years and stimulated numerous research contributions from different areas. While the importance for applications is one driving impulse, the model allures researchers from stochastic calculus and mathematical statistics foremost with its intriguing intrinsic properties and surprising new effects. In this work, we establish central limit theorems for estimators in a general setup, with noisy discrete observations of a continuous semimartingale in the one-dimensional case and noisy and non-synchronous observations multi-dimensional. The vital novelty is that the obtained asymptotic variances of the spectral estimators are smaller than for previous estimation approaches and coincide with the lower Cramér-Rao-type bounds in simplified sub-models. The estimators are thus *asymptotically efficient* in these sub-models in which the notion of efficiency is meaningful and explored in the literature. *Stability* of weak convergence and *feasible limit theorems* allow for confidence intervals.

There exist two major objectives in the strand of literature on volatility estimation:

1. Providing methods for feasible inference in general and flexible models.
2. Attaining the lowest possible asymptotic variance.

The one-dimensional parametric experiment in which the volatility σ is a constant parameter and without drift and with Gaussian i.i.d. noise has been well understood by a LAN (local asymptotic normality) result by Gloter and Jacod (2001). While it is commonly known that for n regularly spaced discrete observations $n^{1/2}$ is the optimal convergence rate in absence of noise and $2\sigma^4$ the variance lower bound, Gloter and Jacod (2001) showed that the optimal rate with noise declines to $n^{1/4}$ and the lower variance bound is $8\eta\sigma^3$, when η^2 is the variance of the noise. Recent years have witnessed the development and suggestion of various estimation methods in a nonparametric framework that can provide consistent rate-optimal estimators. *Stable central limit theorems* have been proved. Let us mention the prominent approaches by Zhang (2006), Barndorff-Nielsen et al. (2008), Jacod et al. (2009) and Xiu (2010) for the one-dimensional case and Aït-Sahalia et al. (2010), Barndorff-Nielsen et al. (2011), Bibinger (2011) and Christensen et al. (2013) for the general multi-dimensional setup. A major focus has been to attain a minimum (asymptotic) variance among all proposed estimators which at the slow optimal convergence rate could result in substantial finite-sample precision gains. For instance Barndorff-Nielsen et al. (2008) have put emphasis on the construction of a version of their kernel estimator which asymptotically attains the bound $8\eta\sigma^3$ in the parametric sub-experiment. Nonparametric efficiency is considered by Reiß (2011) in the one-dimensional setup and recently by Bibinger et al. (2013) in a multi-dimensional non-synchronous framework. Also for the nonparametric experiment without observational noise efficiency is subject of current research, see Renault et al. (2013) and Clément et al. (2013) for recent advances. Jacod and Mykland (2013) have proposed an adaptive version of their pre-average estimator which achieves an asymptotic variance of ca. $1.07 \cdot 8\eta \int_0^t \sigma_s^3\, ds$ and $8\eta \int_0^t \sigma_s^3\, ds$ is the nonparametric lower bound.

Reiß (2011) introduced a spectral approach motivated by an equivalence of the nonparametric and a locally parametric experiment using a local method of moments. In contrast to all previous estimation techniques, the spectral estimator in Reiß (2011) attains the Cramér-Rao efficiency lower bound for the asymptotic variance. The estimator has been extended to the multi-dimensional setting in Bibinger and Reiß (2013) and Bibinger et al. (2013). However, the notion of nonparametric efficiency and the construction of the estimators have been restricted to the simplified statistical experiment where a

continuous martingale without drift and with time-varying but deterministic volatility is observed with additive Gaussian noise. The main contribution of this work is to investigate the spectral approach under model misspecification in the general nonparametric standard setup, i.e. with drift, a general random volatility process and a more general error distribution. We show that the estimators significantly improve upon existing estimation methods also in more complex models which are of central interest in finance and econometrics. We pursue a high-frequency asymptotic distribution theory. The main results are functional stable limit theorems with optimal convergence rates and with variances that coincide with the lower bounds in the sub-experiments. The asymptotic analysis combines the theory to prove stable limit theorems by Jacod (1997), applied in similar context also in Fukasawa (2010) and Hayashi and Yoshida (2011), with Fourier analysis, matrix algebra and proofs of tightness results. Those pop up by the fact that the efficient spectral estimation employs smoothing in the Fourier domain, because the method of moments is based on multi-dimensional Fisher information matrix calculus and because the estimation is carried out in a two-stage approach where in a first step the local covolatility matrix is pre-estimated from the same data.
This article is structured into six sections. Following this introduction, Section 2 introduces the statistical model and outlines all main results in a concise overview. Section 3 revisits the elements of the spectral estimation approach and the multivariate local method of moments. In Section 4 we explain the main steps of the strategy of proofs of the functional central limit theorems. Mathematical details are given in Section 6. In Section 5 we present a Monte Carlo study.

2. Statistical model & Main results

Let us first introduce the statistical model, fix the notation and gather all assumptions for the one- and the multi-dimensional setup.

2.1. Theoretical setup and assumptions

First, consider a one-dimensional Itô semimartingale

$$X_t = X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s\,, \tag{1}$$

on a filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\big)$.

Assumption (H-1). *We pursue the asymptotic analysis under two structural hypotheses for the volatility process of which one must be satisfied:*

$(\sigma-1)$ *There exists a random variable L with at least four finite moments, i.e. with $\mathbb{E}\left[L^4\right]<\infty$, such that $t\mapsto\sigma_t$ is almost surely α-Hölder continuous on $[0,1]$ for $\alpha>1/2$ and Hölder constant L, i.e. $|\sigma_t-\sigma_s|\leq L\,|t-s|^{\alpha}$, $0\leq t,s\leq 1$, almost surely.*

$(\sigma-2)$ *The process σ is itself a continuous Itô semimartingale, i.e. there exist a random variable σ_0 and adapted càdlàg processes $\tilde b=\big(\tilde b_t\big)_{0\leq t\leq 1}$, $\tilde\sigma=(\tilde\sigma_t)_{0\leq t\leq 1}$ and $\tilde\eta=(\tilde\eta_t)_{0\leq t\leq 1}$ such that*

$$\sigma_t = \sigma_0 + \int_0^t \tilde b_s\,ds + \int_0^t \tilde\sigma_s\,dW_s + \int_0^t \tilde\eta_s\,dW'_s. \tag{2}$$

W' is an $(\mathcal{F}_t)_{0\leq t\leq 1}$-Brownian motion which is independent of W.

Furthermore, suppose there exists a constant $\kappa > 0$ such that $|\sigma_t| > \kappa$ uniformly for $0 \le t \le 1$. For the drift process, assume there exists a random variable L' with $\mathbb{E}\left[(L')^2\right] < \infty$ such that $t \mapsto b_t$ is almost surely ν-Hölder continuous on $[0,1]$ for $\nu > 0$ and Hölder constant L', i.e. $|b_t - b_s| \le L' |t-s|^\nu$, $0 \le t, s \le 1$, almost surely.

We work within the model where we have indirect observations of X diluted by noise.

Assumption (Obs-1). *The semimartingale X is observed at regular times $i/n, i = 0, \ldots, n$, with observational noise:*

$$Y_i = X_{i/n} + \epsilon_i, \ \ i = 0, \ldots, n\,. \tag{3}$$

The discrete-time noise process $(\epsilon_i)_{i=0,\ldots,n}$ is a sequence of i.i.d. real random variables with $\mathbb{E}[\epsilon_i] = 0$, variance η^2 and having finite eighth moments. We assume the noise process is independent of $\mathcal{F}$. Set $\mathcal{G}_t = \mathcal{F}_t \otimes \sigma(\epsilon_0, \ldots, \epsilon_l : l/n \le t, l \in \mathbb{N})$ for $0 \le t \le 1$, and the filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{0 \le t \le 1}, \mathbb{P})$ which accommodates the signal and the noise process.

For the multi-dimensional case, we focus on a d-dimensional continuous Itô semimartingale

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s \tag{4}$$

on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ with $(\mathcal{F}_t)$ being a right-continuous and complete filtration and W being here a d-dimensional $(\mathcal{F}_t)$-adapted standard Brownian motion. The *integrated covolatility matrix* is denoted $\int_0^t \Sigma_s\, ds$, $\Sigma_s = \sigma_s \sigma_s^\top$. It coincides with the quadratic covariation matrix $[X, X]_t$ of the semimartingale X. We denote the spectral norm by $\|\cdot\|$ and define $\|f\|_\infty := \sup_{t \in [0,1]} \|f(t)\|$. Consider Hölder balls of order $\alpha \in (0,1]$ and with radius $R > 0$:

$$C^\alpha(R) = \{f \in C^\alpha([0,1], \mathbb{R}^{d \times d'}) | \|f\|_{C^\alpha} \le R\}\,,\ \|f\|_{C^\alpha} := \|f\|_\infty + \sup_{x \ne y} \frac{\|f(x) - f(y)\|}{|x-y|^\alpha}\,.$$

We assume the following regularity conditions.

Assumption (H-d). *The drift b is a d-dimensional $(\mathcal{F}_t)$-adapted process with $b \in C^\nu(R)$ for some $R > 0, \nu > 0$ and the stochastic instantaneous volatility process σ is a $(d \times d')$-dimensional $(\mathcal{F}_t)$-adapted process satisfying one of the following regularity conditions:*

($\Sigma - 1$) *$\sigma \in C^\alpha(R)$ for some $R > 0$ and with Hölder exponent $\alpha > 1/2$.*

($\Sigma - 2$) *σ is a continuous Itô semimartingale of the form* (4) *whose characteristics are assumed to be càdlàg.*

Furthermore, we assume that the positive definite matrix Σ_s satisfies $\Sigma_s \ge \underline{\Sigma} E_d$, uniformly, in the sense of Löwner ordering of positive definite matrices. This is the analogue of "bounded from below" for $d = 1$.

We consider a very general framework with noise and in which observations come at non-synchronous sampling times.

Assumption (Obs-d). *The signal process X of the form* (4) *is discretely and non-synchronously observed on* $[0,1]$. *Observation times $t_i^{(l)}, 0 \le i \le n_l, l = 1, \dots, d$, are described by quantile transformations $t_i^{(l)} = F_l^{-1}(i/n_l)$, $0 \le i \le n_l$, $1 \le l \le d$, with differentiable distribution functions $F_l, 1 \le l \le d$, $F_l(0) = 0$, $F_l(1) = 1$ and $F_l' \in C^\alpha([0,1],[0,1])$ with $\|F'\|_{C^\alpha}$ bounded for some $\alpha > 1/2$ and F_l' strictly positive. The observation times are independent of X. We pursue asymptotics as there exists a sequence n such that $n_p/n \to c_p$ with constants $0 < c_p < \infty$ for all $p = 1, \dots, d$. Observations are subject to an additive noise:*

$$Y_i^{(l)} = X_{t_i^{(l)}}^{(l)} + \epsilon_i^{(l)} \, , i = 0, \dots, n_l \, .$$

The observation errors are assumed to be i.i.d. with expectation zero and $\eta_l^2 = \mathbb{V}\mathrm{ar}(\epsilon_i^{(l)})$ and independent of X and the observation times. Furthermore, the errors are mutually independent for all components and eighth moments exist. Again, we shall write $(\Omega, \mathcal{G}, (\mathcal{G}_t), \mathbb{P})$ for the probability space that accommodates Y.

Our analysis includes deterministic and random observation times which are independent of Y. Though Assumption (Obs-d) displays to some extent still an idealization of realistic market microstructure dynamics, our observation model constitutes the established setup in related literature and captures the main ingredients of realistic log-price models. We aim at exploring the semiparametric efficient methods by Reiß (2011) and Bibinger et al. (2013) in this benchmark model.

2.2. Mathematical concepts and notation

Denote $\Delta_i^n Y^{(l)} = Y_i^{(l)} - Y_{i-1}^{(l)}, 1 \le i \le n_l, l = 1, \dots, d$, the increments of $Y^{(l)}$ and analogously for X and other processes. In the one-dimensional case, we write $\Delta^n Y = (\Delta_i^n Y)_{1 \le i \le n} \in \mathbb{R}^n$ for the vector of increments. We shall write $Z_n = \mathcal{O}_{\mathbb{P}}(\delta_n)$ ($Z_n = \scriptstyle{\mathcal{O}}_{\mathbb{P}}(\delta_n)$) for real random variables, to express that the sequence $\delta_n^{-1} Z_n$ is bounded (tends to zero) in probability under $\mathbb{P}$. Analogously $\mathcal{O}$ and $\scriptstyle{\mathcal{O}}$ are used for deterministic sequences. To express that terms are of the same asymptotic order we write $Z_n \asymp_p Y_n$ if $Z_n = \mathcal{O}_{\mathbb{P}}(Y_n)$ and $Y_n = \mathcal{O}_{\mathbb{P}}(Z_n)$ and likewise $\asymp$ for deterministic terms. Also, we use the short notation $A_n \lesssim B_n$ for $A_n = \mathcal{O}(B_n)$. Convergence in probability and weak convergence are denoted by $\stackrel{\mathbb{P}}{\to}$ and $\stackrel{d}{\longrightarrow}$;$\stackrel{st}{\longrightarrow}$ refers to stable weak convergence with respect to $\mathcal{G}$ - if not further specified. We write $X^n \stackrel{ucp}{\longrightarrow} X$ for processes X^n, X to express shortly that $\sup_{t\in[0,1]} |X_t^n - X_t| \stackrel{\mathbb{P}}{\to} 0$. δ_{lm} is Kronecker's delta, i.e. $\delta_{lm} = 1$ for $l = m$, $\delta_{lm} = 0$ else. For functional convergence of processes we focus on the space $\mathcal{D}([0,1])$, the space of càdlàg functions (right-continuous with left limits).

Recall the definition of stable weak convergence which is an essential concept in asymptotic theory for volatility estimation. For a sub-σ-field $\mathcal{A} \subseteq \mathcal{F}$, a sequence of random variables (X_n) taking values in a Polish space $(E, \mathcal{E})$ converges $\mathcal{A}$-stably, if

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

with a random probability measure μ on $(\Omega \times E, \mathcal{A} \otimes \mathcal{E})$ and for all continuous and bounded f and $\mathcal{A}$-measurable bounded random variables Z. The definition is equivalent to joint weak convergence of (Z, X_n) for every $\mathcal{A}$-measurable random variable Z. Thus $\mathcal{F}$-stable weak convergence means $\lim_{n\to\infty} \mathbb{E}\left[f(X_n)Z\right] = \mathbb{E}'\left[f(X)Z\right]$ for all bounded continuous f and bounded measurable Z, where the limit X of (X_n) is defined on an extended probability space $(\Omega', \mathcal{F}', \mathbb{P}')$. In our setup, the extended space will be given by the orthogonal product of $(\Omega, \mathcal{F}, \mathbb{P})$ and an auxiliary space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$.

We refer to Podolskij and Vetter (2010) for more information on stable convergence.

For the multi-dimensional setting the vec-operator and Kronecker products of matrices will be important. For a matrix $A \in \mathbb{R}^{d\times d}$ we write the entries $A_{pq}, 1 \le p \le d, 1 \le q \le d$, and the vector of its entries obtained by stacking its columns one below each other

$$\operatorname{vec}(A) = \big(A_{11}, A_{21}, \ldots, A_{d1}, A_{12}, A_{22}, \ldots, A_{d2}, \ldots, A_{d(d-1)}, A_{dd}\big)^\top \in \mathbb{R}^{d^2}.$$

The transpose of a matrix A is denoted by $A^\top$. For matrix functions in time , for instance the covolatility matrix, we write the entries $A_t^{(pq)}$. The Kronecker product $A \otimes B \in \mathbb{R}^{d^2\times d^2}$ for $A, B \in \mathbb{R}^{d\times d}$ is defined as

$$(A \otimes B)_{d(p-1)+q, d(p'-1)+q'} = A_{pp'} B_{qq'}, \quad p, q, p', q' = 1, \ldots, d.$$

In the multivariate limit theorems, we account for effects by non-commutativity of matrix multiplication. It will be useful to standardize limit theorems such that the matrix

$$\mathcal{Z} = \mathbb{COV}(\operatorname{vec}(ZZ^\top)), \text{ for } Z \sim N(0, E_d) \text{ standard Gaussian,} \tag{5}$$

appears as variance-covariance matrix of the standardized form instead of the identity matrix. This matrix is the sum of the d^2-dimensional identity matrix E_{d^2} and the so-called commutation matrix $C_{d,d}$ that maps a $(d \times d)$ matrix to the vec of its transpose $C_{d,d}\operatorname{vec}(A) = \operatorname{vec}(A^\top)$. The matrix $\mathcal{Z}/2$ is idempotent and introduced in Abadir and Magnus (2009), Chapter 11, as the *symmetrizer matrix*. Note that in the multi-dimensional experiment under equidistant synchronous non-noisy observations of X, the sample realised covolatility matrix $\widehat{IC} = \sum_{i=1}^n (X_{i/n} - X_{(i-1)/n})(X_{i/n} - X_{(i-1)/n})^\top$ obeys the central limit theorem:

$$\operatorname{vec}\left(n^{1/2}\Big(\widehat{IC} - \int_0^1 \Sigma_s\, ds\Big)\right) \xrightarrow{st} N\Big(0, \int_0^1 \big(\Sigma_s \otimes \Sigma_s\big)\, \mathcal{Z}\, ds\Big), \tag{6}$$

where similarly as in our result below the matrix $\mathcal{Z}$ appears as one factor in the asymptotic variance and remains after standardization. For background information on matrix algebra, especially tensor calculus using the Kronecker product and vec-operator we refer interested readers to Abadir and Magnus (2009). Note the crucial relation between the Kronecker product and the vec-operator $\operatorname{vec}(ABC) = (C^\top \otimes A)\operatorname{vec}(B)$.

In the multi-dimensional setup we introduce a diagonal matrix function of noise levels $\mathcal{H}(t) = \operatorname{diag}\left(\eta_l\big(\nu_l (F_l^{-1})'(t)\big)^{1/2}\right)_{1\le l\le d}$ incorporating constants ν_l when $n/n_l \to \nu_l$ and by a locally constant approximation of the observation frequencies a bin-wise locally constant approximation of $\mathcal{H}$:

$$\mathbf{H}_k^n = \operatorname{diag}(\eta_l^2 \nu_l (F_l^{-1})'((k-1)h_n))_{1\le l\le d} = \operatorname{diag}(H_l^{kh_n})_{1\le l\le d}\,. \tag{7}$$

We employ the notion of empirical scalar products in the fashion of Bibinger and Reiß (2013), which is recalled in Definition 1 in the Appendix, along with some useful properties.

2.3. *Outline of the main results*

In the sequel, we present the three major results of this work in Theorem 1, Theorem 2 and Theorem 3 and concisely discuss the consequences. Theorems 1 and 2 establish functional stable central limit theorems in a general semimartingale setting for the spectral integrated volatility and covolatility estimator by Bibinger and Reiß (2013). Theorem 3 gives a multivariate limit theorem for the localized method of moment approach by Bibinger et al. (2013). These methods developed in Reiß (2011), Bibinger and Reiß (2013) and Bibinger et al. (2013) and briefly explained in a nutshell in Section 3.1 attain asymptotic efficiency lower variance bounds in the simplified model without drift, with independent volatility and covolatility processes and normally distributed noise.

Theorem 1. *In the one-dimensional experiment, on Assumption (H-1) and Assumption (Obs-1), for the adaptive spectral estimator of integrated squared volatility $\widehat{\mathbf{IV}}_{n,t}$, stated in* (22a) *below, the functional stable weak convergence*

$$n^{\frac{1}{4}}\left(\widehat{\mathbf{IV}}_{n,t}-\int_0^t \sigma_s^2\,ds\right)\xrightarrow{st}\int_0^t\sqrt{8\eta\,|\sigma_s^3|}\,dB_s \tag{8}$$

applies as $n\to\infty$ on $\mathcal{D}[0,1]$, where B is a Brownian motion defined on an extension of the original probability space $(\Omega,\mathcal{G},(\mathcal{G}_t)_{0\le t\le 1},\mathbb{P})$, independent of the original σ-algebra $\mathcal{G}$.
Moreover, the implicitly obtained variance estimator $\widehat{\mathcal{V}}_{n,t}^{\mathcal{IV}}$ in (22b) *provides for $0\le t\le 1$ the feasible central limit theorem:*

$$\left(\widehat{\mathcal{V}}_{n,t}^{\mathcal{IV}}\right)^{-1/2}\left(\widehat{\mathbf{IV}}_{n,t}-\int_0^t \sigma_s^2\,ds\right)\xrightarrow{d} N(0,1)\,. \tag{9}$$

Remark 1. The convergence rate in (8) and (9) is *optimal*, already in the parametric subexperiment, see Gloter and Jacod (2001). $\widehat{\mathbf{IV}}_{n,1}$ is asymptotically mixed normally distributed with random asymptotic variance $\int_0^1 8\eta\left|\sigma_s^3\right|\,ds$. This asymptotic variance coincides with the lower bound derived by Reiß (2011) in the subexperiment with time-varying but deterministic volatility, without drift and Gaussian error distribution. The spectral estimator of squared integrated volatility is hence *asymptotically efficient* in this setting. For the general semimartingale experiment the concept of asymptotic efficiency is not developed yet, it is conjectured that the lower bound has analogous structure, see Jacod and Mykland (2013). Theorem 1 establishes that the asymptotic variance of the estimator has the same form in the very general framework and stable convergence holds true. The *feasible limit theorem* (9) allows to provide confidence bands and is of pivotal importance for practical capability.

Theorem 2. *In the multi-dimensional experiment, on Assumption (H-d) and Assumption (Obs-d), for the adaptive spectral estimator of integrated covolatility $\widehat{\mathbf{ICV}}_{n,t}^{(p,q)}$, stated in* (25a) *below, the functional stable weak convergence*

$$n^{\frac{1}{4}}\left(\widehat{\mathbf{ICV}}_{n,t}^{(p,q)}-\int_0^t \Sigma_s^{(pq)}\,ds\right)\xrightarrow{st}\int_0^t v_s^{(p,q)}\,dB_s \tag{10}$$

applies for $n_p/n\to c_p$ and $n_q/n\to c_q$ with $0<c_p<\infty, 0<c_q<\infty$, as $n\to\infty$ on $\mathcal{D}[0,1]$, where B is a Brownian motion defined on an extension of the original probability space $(\Omega,\mathcal{G},(\mathcal{G}_t)_{0\le t\le 1},\mathbb{P})$, independent of the original σ-algebra $\mathcal{G}$. The asymptotic variance process is given by

$$\begin{aligned}\left(v_s^{(p,q)}\right)^2 &= 2\left((F_p^{-1})'(s)(F_q^{-1})'(s)c_p^{-1}c_q^{-1}(A_s^2-B_s)B_s\right)^{1/2}\\ &\quad\times\left(\sqrt{A_s+\sqrt{A_s^2-B_s}}-\operatorname{sgn}(A_s^2-B_s)\sqrt{A_s-\sqrt{A_s^2-B_s}}\right),\end{aligned} \tag{11}$$

with the terms

$$\begin{aligned}A_s &= \Sigma_s^{(pp)}\frac{(F_q^{-1})'(s)c_p}{(F_p^{-1})'(s)c_q}+\Sigma_s^{(qq)}\frac{(F_p^{-1})'(s)c_q}{(F_q^{-1})'(s)c_p}\,,\\ B_s &= 4\left(\Sigma_s^{(pp)}\Sigma_s^{(qq)}+\left(\Sigma_s^{(pq)}\right)^2\right)\,.\end{aligned}$$

Moreover, the implicitly obtained variance estimator $\widehat{\mathcal{V}}_{n,t}^{\mathcal{ICV}^{(p,q)}}$ in (25b) *provides for $0\le t\le 1$ the feasible central limit theorem:*

$$\left(\widehat{\mathcal{V}}_{n,t}^{\mathcal{ICV}^{(p,q)}}\right)^{-1/2}\left(\widehat{\mathbf{ICV}}_{n,t}^{(p,q)}-\int_0^t \Sigma_s^{(pq)}\,ds\right)\xrightarrow{d} N(0,1)\,. \tag{12}$$

Remark 2. The bivariate extension of the spectral method outperforms by its local adaptivity and Fourier domain smoothing previous approaches in most cases, see Bibinger and Reiß (2013) for a detailed discussion and survey on the different methods. Yet, it attains the multi-dimensional variance lower bound for estimating the integrated covolatility $\int_0^1 \Sigma_s^{(pq)}\,ds$ only in case that $[W^{(p)}, W^{(q)}] \equiv 0$. On the other hand the estimator already achieves a high efficiency and since it does not involve Fisher information weight matrices, it is less computationally costly than the efficient local method of moments approach. The general form of the asymptotic variance given in (11) is a bit complicated. In case that $[W^{(p)}, W^{(q)}] \equiv 0$ and for equal volatilities $\Sigma_t^{(11)} = \Sigma_t^{(22)} = \sigma_t$ it simplifies to $\int_0^t 4\eta|\sigma_s^3|\,ds$, being efficient for this setup. By its implicitly obtained rescaled version (12) allowing for confidence, the estimator is of high practical value.

Theorem 3. *In the multi-dimensional experiment, on Assumption (H-d) and Assumption (Obs-d), for the local method of moments estimator of the integrated covolatility matrix $\mathbf{LMM}_{n,t}$, stated in* (29a) *below, the functional stable weak convergence*

$$n^{\frac{1}{4}}\Big(\mathbf{LMM}_{n,t} - \mathrm{vec}\Big(\int_0^t \Sigma_s\,ds\Big)\Big) \xrightarrow{st} \int_0^t \big(\Sigma_s^{\frac{1}{2}} \otimes \big(\Sigma_s^{\mathcal{H}}\big)^{\frac{1}{4}}\big)\mathcal{Z}dB_s + \int_0^t \big(\big(\Sigma_s^{\mathcal{H}}\big)^{\frac{1}{4}} \otimes \Sigma_s^{\frac{1}{2}}\big)\mathcal{Z}dB_s^{\perp} \quad (13)$$

applies, with $\mathcal{H}(t) = \mathrm{diag}(\eta_p \nu_p^{1/2} F_p'(t)^{-1/2})_p \in \mathbb{R}^{d\times d}$ and $\big(\Sigma^{\mathcal{H}}\big)^{1/4}$ the matrix square root of $\big(\Sigma^{\mathcal{H}}\big)^{1/2} := \mathcal{H}(\mathcal{H}^{-1}\Sigma\mathcal{H}^{-1})^{1/2}\mathcal{H}$, as $n \to \infty$ and $n/n_p \to \nu_p$ for $p = 1, \ldots, d$, on $\mathcal{D}[0,1]$, where $\mathcal{Z}$ is the matrix defined in (5) *and B and $B^{\perp}$ are two independent d^2-dimensional Brownian motions, both defined on an extension of the original probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{0\le t\le 1}, \mathbb{P})$, independent of the original σ-algebra $\mathcal{G}$. The point-wise marginal central limit theorem reads*

$$n^{\frac{1}{4}}\left(\mathbf{LMM}_{n,1} - \mathrm{vec}\Big(\int_0^1 \Sigma_s\,ds\Big)\right) \xrightarrow{st} MN\left(0, \mathbf{I}^{-1}\mathcal{Z}\right), \quad (14)$$

where MN means mixed normal distribution, with the asymptotic variance-covariance matrix

$$\mathbf{I}^{-1} = 2\int_0^1 \big(\Sigma_s \otimes \big(\Sigma_s^{\mathcal{H}}\big)^{1/2} + \big(\Sigma_s^{\mathcal{H}}\big)^{1/2} \otimes \Sigma_s\big)\,ds\,. \quad (15)$$

Moreover, the implicitly obtained variance-covariance matrix estimator $\mathbf{I}_{n,t}^{-1}$ in (29b) *provides for $0 \le t \le 1$ the feasible central limit theorem:*

$$\mathbf{I}_{n,t}^{1/2}\left(\mathbf{LMM}_{n,t} - \mathrm{vec}\Big(\int_0^t \Sigma_s\,ds\Big)\right) \xrightarrow{d} N(0, \mathcal{Z})\,. \quad (16)$$

Remark 3. The local method of moments attains the lower asymptotic variance bound derived in Bibinger et al. (2013) for a nonparametric experiment with deterministic covolatility matrix, without drift and Gaussian error distribution. Thus, the local method of moments is *asymptotically efficient* in this subexperiment.

The asymptotic variance of estimating an integrated volatility decreases as the information inherent in the observation of correlated components can be exploited. In the multi-dimensional observation model the attained minimum asymptotic variance of estimating integrated squared volatility can become much smaller than the bound in (8) for $d = 1$. In an idealized parametric model with $\sigma > 0$, the variance can be reduced up to $(8/\sqrt{d})\eta\sigma^3$ in comparison to the one-dimensional lower bound $8\eta\sigma^3$, see Bibinger et al. (2013) for a deeper discussion of the lower bound.

3. Spectral estimation of integrated volatility and the integrated covolatility matrix

3.1. Elements of spectral estimation

We shall concisely recapitulate the building blocks of spectral estimation in the sequel. For simplicity we start with the one-dimensional framework, $d = 1$. We partition the time span $[0, 1]$ into equidistant bins $[(k-1)h_n, kh_n]$, $k = 1, \ldots, h_n^{-1} \in \mathbb{N}$, $h_n \to 0$ as $n \to \infty$. For the ease of exposition, suppose without loss of generality $nh_n \in \mathbb{N}$, the number of observations on each bin. Consider a statistical experiment in which we approximate σ_t by a locally constant function. Then, presume on each bin $[(k-1)h_n, kh_n]$ a locally constant squared volatility $\sigma^2_{(k-1)h_n}$. Consequently, in this experiment we estimate $\int_{(k-1)h_n}^{kh_n} \sigma_s^2\, ds$ by $h_n \hat{\sigma}^2_{(k-1)h_n}$, solving locally parametric estimation problems. For this purpose, Reiß (2011) motivated to smooth noisy data bin-wise in the Fourier domain and construct an efficient estimator by a linear combination of smoothed spectral statistics over different frequencies. Thereto, consider the $L^2([0,1])$-orthonormal systems of trigonometric functions:

$$\Phi_{jk}(t) = \sqrt{\frac{2}{h_n}} \sin\left(j\pi h_n^{-1}\left(t - (k-1)\,h_n\right)\right) \mathbf{1}_{[(k-1)h_n, kh_n]}(t)\,, \quad j \geq 1,\, 0 \leq t \leq 1\,, \tag{17a}$$

$$\varphi_{jk}(t) = 2n\sqrt{\frac{2}{h_n}} \sin\left(\frac{j\pi}{2nh_n}\right) \cos\left(j\pi h_n^{-1}\left(t - (k-1)\,h_n\right)\right) \mathbf{1}_{[(k-1)h_n, kh_n]}(t)\,. \tag{17b}$$

Those provide weight functions for the *spectral statistics*:

$$S_{jk} = \sum_{i=1}^{n} \Delta_i^n Y\, \Phi_{jk}(i/n)\,, j = 1, \ldots, nh_n - 1, k = 1, \ldots, h_n^{-1}. \tag{18}$$

The spectral statistics are the principal element of the considered estimation techniques. They are related to the pre-averages of Jacod et al. (2009) which have been designed for our one-dimensional estimation problem, as well. A main difference is that we keep the bins fixed which makes the construction of the spectral approach simple. Bin-wise the spectral estimation profits from an advanced smoothing method in the Fourier domain, i.e. using the weight function of a discrete sine transformation. The spectral statistics hence *de-correlate* the observations and form their bin-wise principal components. Reiß (2011) showed that this leads to a semiparametrically efficient estimation approach of squared integrated volatility in a nonparametric setup with deterministic volatility, without drift and normally distributed noise. The bin-width is chosen as $h_n \asymp n^{-1/2} \log n$ to attain the optimal convergence rates and for the results in Section 2.1. This becomes clear in the proofs in Section 6. The log-factor plays a role in the convergence of the sum of variances over different frequencies. The leading asymptotic order $n^{-1/2}$ for the bin-width is analogous to the pre-average and kernel bandwidths, cf. Jacod et al. (2009) and Barndorff-Nielsen et al. (2008), and balances the error by discretization which increases with increasing h_n and the error due to noise which decreases as h_n increases. Let us point out that the basis functions (17a) and (17b) are slightly scaled versions of the respective basis functions in Bibinger and Reiß (2013) and Bibinger et al. (2013) for a more convenient exposition, but these factors which equal the empirical norms of the φ_{jk} have to be considered when translating expressions.

3.2. The spectral estimator of integrated volatility

Locally parametric estimates for the squared volatility $\hat{\sigma}^2_{(k-1)h_n}$ are obtained by linear combinations with weights w_{jk} of bias-corrected squared spectral statistics:

$$\hat{\sigma}^2_{(k-1)h_n} = \sum_{j=1}^{nh_n - 1} w_{jk}\left(S_{jk}^2 - [\varphi_{jk}, \varphi_{jk}]_n \frac{\eta^2}{n}\right). \tag{19}$$

The correction for the bias due to noise incorporates the empirical norm from Definition 1 and the noise level η which is in general unknown – but can be consistently estimated from the data with $n^{1/2}$ convergence rate, e. g. by $\hat{\eta}^2 = (2n)^{-1}\sum_{i=1}^{n}(\Delta_i^n Y)^2$, see Zhang et al. (2005) for an asymptotic analysis of this estimator.
The estimator of the integrated squared volatility $\int_0^t \sigma_s^2\,ds$ is constructed as Riemann sum

$$\sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n\hat{\sigma}^2_{(k-1)h_n} = \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n \sum_{j=1}^{nh_n-1} w_{jk}\Big(S_{jk}^2 - [\varphi_{jk},\varphi_{jk}]_n\,\frac{\eta^2}{n}\Big)\,, \tag{20}$$

such that the estimator at $t = 1$ becomes simply the average of local estimates in the case of equi-spaced bins. Set $I_{jk} = \big(\mathbb{V}\text{ar}\big(S_{jk}^2\big)\big)^{-1}$ and $I_k = \sum_{j=1}^{nh_n-1} I_{jk}$. The variance of the above estimator becomes minimal and equal to $\sum_{k=1}^{\lfloor th_n^{-1}\rfloor} I_k^{-1}$ for the oracle weights

$$w_{jk} = I_k^{-1}I_{jk} = \frac{\Big(\sigma^2_{(k-1)h_n} + \frac{\eta^2}{n}\,[\varphi_{jk},\varphi_{jk}]_n\Big)^{-2}}{\sum_{m=1}^{nh_n-1}\Big(\sigma^2_{(k-1)h_n} + \frac{\eta^2}{n}\,[\varphi_{mk},\varphi_{mk}]_n\Big)^{-2}} \tag{21}$$

for $k = 1,\ldots,h_n^{-1}$ and $j = 1,\ldots,nh_n - 1$, when the noise is Gaussian. For general noise distribution it turns out that the first-order variance remains invariant. It is essential to develop an adaptive version of the estimator, for which we replace the oracle optimal weights by data-driven estimated optimal weights. Additionally to the estimated noise variance, a bin-wise consistent estimator of the local volatilities $\sigma^2_{(k-1)h_n}$ with some convergence rate suffices. Local pre-estimates of the volatilities $\sigma^2_{(k-1)h_n}$ can be constructed by using the same ansatz as in (19), but involving only a small number $J_n \ll nh_n - 1$ of frequencies and constant weights $w_{jk} = J_n^{-1}$ and then averaging over $K_n \asymp n^{1/4}$ bins in a neighborhood of $(k-1)h_n$. This estimator attains at least a $n^{1/8}$ rate of convergence, the latter in case of $\alpha \approx 1/2$ under $(\sigma - 1)$ or under $(\sigma - 2)$ in Assumption (H-1). For a smoother volatility, K_n is chosen larger leading to a faster convergence rate, see also Bibinger and Reiß (2013) for a discussion on the estimation of the instantaneous volatility. Since plugging in the pre-estimates of local squared volatilities and of the noise variance implicitly provides estimates $\hat{I}_{jk}, \hat{I}_k$ for I_{jk}, I_k and thus also $\hat{w}_{jk} = \hat{I}_k^{\,-1}\hat{I}_{jk}$ for w_{jk} as well, we can define the final adaptive spectral estimator of volatility and the estimator for its variance based on a two-stage approach:

$$\widehat{\mathbf{IV}}_{n,t} = \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n \sum_{j=1}^{nh_n-1} \hat{w}_{jk}\Big(S_{jk}^2 - [\varphi_{jk},\varphi_{jk}]_n\,\frac{\hat{\eta}^2}{n}\Big)\,, \tag{22a}$$

$$\widehat{\mathcal{V}}^{\mathcal{IV}}_{n,t} = \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n^2\,\hat{I}_k^{-1}\,. \tag{22b}$$

3.3. The spectral covolatility estimator

The spectral covolatility estimator from Bibinger and Reiß (2013) is the obvious extension of the one-dimensional estimator using cross-products of spectral statistics:

$$S_{jk}^{(p)} = \sum_{i=1}^{n_p}\Delta_i^n Y^{(p)}\Phi_{jk}\left(\frac{t_i^{(p)} + t_{i-1}^{(p)}}{2}\right)\ ,\ j \ge 1, p = 1,\ldots,d, k = 1,\ldots,h_n^{-1}. \tag{23}$$

The basis functions (Φ_{jk}) are defined as in (17a). Under non-synchronous observations we modify in (17b) the factor in front of the cosine to the simpler expression $\sqrt{2}\pi j h_n^{-3/2}$, such that $\Phi'_{jk} = \varphi_{jk}$. This meets the original idea by Reiß (2011) to use orthogonal systems of functions and their derivatives. While in the case of regular observations on the grid $i/n, i = 0, \ldots, n$, we can slightly profit by the discrete Fourier analysis, we use for non-synchronous observations from now on the continuous-time analogues which coincide with the first-order discrete expressions. Starting from an asymptotically equivalent continuous-time observation model, where we set $S_{jk}^{(p)} = \int \varphi_{jk}(t) dY^{(p)}(t)$ as in Bibinger et al. (2013), integration by parts leads to the more natural discrete-time approximation (23) here where the Φ_{jk} are evaluated at mid-times $(t_{i+1}^{(p)} - t_i^{(p)})/2$ (see Bibinger et al. (2013) for details). The standardization is $[\varphi_{jk}, \varphi_{jk}] = \int_0^1 \varphi_{jk}^2(t)\, dt = h_n^{-2}\pi^2 j^2$. The weights $w_{jk}^{p,q} = (I_k^{(p,q)})^{-1} I_{jk}^{(p,q)}$,

$$I_{j(k+1)}^{(p,q)} = \Big(\Sigma_{kh_n}^{(pp)}\Sigma_{kh_n}^{(qq)} + (\Sigma_{kh_n}^{(pq)})^2 + H_p^{kh_n} H_q^{kh_n}[\varphi_{jk}, \varphi_{jk}]^2 + \big(\Sigma_{kh_n}^{(pp)} H_q^{kh_n} + \Sigma_{kh_n}^{(qq)} H_p^{kh_n}\big)[\varphi_{jk}, \varphi_{jk}]\Big)^{-1}$$

depend on the volatilities, covolatility and noise levels of the considered components as defined in (7). The local noise level combines the global noise variance η_p^2 and local observation densities. It can be estimated with

$$\hat{H}_p^{kh_n} = \frac{\sum_{i=1}^{n_p} \big(\Delta_i Y^{(p)}\big)^2}{2h_n} \sum_{kh_n \le t_v^{(p)} \le (k+1)h_n} \big(t_v^{(p)} - t_{v-1}^{(p)}\big)^2, \tag{24}$$

see the asymptotic identity (43) below. The bivariate spectral covolatility estimator with adaptive weights for $p \neq q, p, q \in \{1, \ldots, d\}$ is

$$\widehat{\mathbf{ICV}}_{n,t}^{(p,q)} = \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n \sum_{j=1}^{nh_n - 1} \hat{w}_{jk}^{(p,q)} \Big(S_{jk}^{(p)} S_{jk}^{(q)}\Big), \tag{25a}$$

where we choose the sequence n such that $nh_n \in \mathbb{N}$, and the estimator for its variance:

$$\widehat{\mathcal{V}}_{n,t}^{\mathcal{ICV}(p,q)} = \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n^2 \big((\hat{I}_k)^{(p,q)}\big)^{-1}. \tag{25b}$$

A more general version of the spectral covolatility estimator for a model including cross-correlation of the noise (in a synchronous framework) can be found in Bibinger and Reiß (2013). For a simpler exposition and since this notion of cross-correlation is not adequate for the more important non-synchronous case, we restrict ourselves here to noise according to Assumption (Obs-d).

3.4. Local method of moments

Consider the *vectors* of spectral statistics:

$$S_{jk} = \Big(\sum_{i=1}^{n_p} \Big(Y_i^{(p)} - Y_{i-1}^{(p)}\Big)\Phi_{jk}\Big(\frac{t_{i-1}^{(p)} + t_i^{(p)}}{2}\Big)\Big)_{1\le p\le d}, \tag{26}$$

for all $k = 1, \ldots, h_n^{-1}$ and $j \geq 1$. Averaging empirical covariances $S_{jk}S_{jk}^\top$ over different spectral frequencies $j = 1, \ldots, J_n$ and over a set of $(2K_n + 1)$ adjacent bins yields a consistent estimator of the instantaneous covolatility matrix:

$$\hat{\Sigma}_s^{pilot} = (2K_n + 1)^{-1} \sum_{k=\lfloor sh_n^{-1}\rfloor - K_n}^{\lfloor sh_n^{-1}\rfloor + K_n} J_n^{-1} \sum_{j=1}^{J_n} \Big(S_{jk}S_{jk}^\top - \hat{\mathbf{H}}_k^n\Big), \tag{27}$$

with $\hat{\mathbf{H}}_k^n$ the estimated noise levels matrix (7), not discussing end effects for $s < K_n h_n$ and $s > 1 - K_n h_n$ here.
The fundamental novelty of the local method of moments approach is to involve multivariate Fisher informations as optimal weight matrices which are $(d^2 \times d^2)$ matrices of the following form:

$$W_{jk} = I_k^{-1} I_{jk} = \Big(\sum_{u=1}^{nh_n-1} \big(\Sigma_{(k-1)h_n} + [\varphi_{uk}, \varphi_{uk}]\mathbf{H}_k^n\big)^{-\otimes 2}\Big)^{-1} \big(\Sigma_{(k-1)h_n} + [\varphi_{jk}, \varphi_{jk}]\mathbf{H}_k^n\big)^{-\otimes 2}, \quad (28)$$

$nh_n \in \mathbb{N}$, where $A^{\otimes 2} = A \otimes A$ denotes the Kronecker product of a matrix with itself and $A^{-\otimes 2} = (A^{\otimes 2})^{-1} = (A^{-1})^{\otimes 2}$. With the pilot estimates and estimators for the noise level at hand, we derive estimated optimal weight matrices for building a linear combination over spectral frequencies $j = 1, \ldots, nh_n - 1$, similar as above. The final estimator of the vectorization of the integrated covolatility matrix $\operatorname{vec}(\int_0^t \Sigma_s \, ds)$, becomes

$$\mathbf{LMM}_{n,t} = \sum_{k=1}^{\lfloor th_n^{-1} \rfloor} h_n \sum_{j=1}^{nh_n-1} \hat{W}_{jk} \operatorname{vec}\left(S_{jk} S_{jk}^{\top} - \hat{\mathbf{H}}_k^n \right) , \quad (29a)$$

and the implicitly derived estimator of its variance-covariance matrix:

$$\hat{\mathbf{I}}_{n,t}^{-1} = \sum_{k=0}^{\lfloor th_n^{-1} \rfloor} h_n^2 \Big(\sum_{j=1}^{nh_n-1} \hat{I}_{jk} \Big)^{-1} . \quad (29b)$$

4. Asymptotic theory

We start with the one-dimensional experiment first. We decompose X using a locally constant approximation of the volatility path and the approximation error:

$$X_t = X_0 + \tilde{X}_t + (X_t - X_0 - \tilde{X}_t) , \quad (30a)$$

where we define

$$\tilde{X}_t = \int_0^t \sigma_{\lfloor sh_n^{-1} \rfloor h_n} \, dW_s , \quad (30b)$$

as a simplified process without drift and with locally constant volatility. The asymptotic theory is conducted for oracle versions of the spectral estimators first with optimal oracle weights. Below the effect of a pre-estimation for the fully adaptive estimator is shown to be asymptotically negligible at first order. In the following, we distinguish between $\widehat{\mathbf{IV}}_{n,t}^{or}(Y)$, the oracle version of the spectral volatility estimator (22a) from noisy observations, and $\widehat{\mathbf{IV}}_{n,t}^{or}(\tilde{X} + \epsilon)$ for the oracle estimator in a simplified experiment in which $\tilde{X}$ instead of X is observed with noise. It turns out that both have the same asymptotic limiting distribution. In order to establish a functional limit theorem, we decompose the estimation error of the oracle version of (22a) in the following way:

$$\widehat{\mathbf{IV}}_{n,t}^{or}(Y) - \int_0^t \sigma_s^2 \, ds = \widehat{\mathbf{IV}}_{n,t}^{or}(\tilde{X} + \epsilon) - \int_0^t \sigma^2_{\lfloor sh_n^{-1} \rfloor h_n} \, ds \quad (31a)$$

$$+ \widehat{\mathbf{IV}}_{n,t}^{or}(Y) - \widehat{\mathbf{IV}}_{n,t}^{or}(\tilde{X} + \epsilon) - \int_0^t \big(\sigma_s^2 - \sigma^2_{\lfloor sh_n^{-1} \rfloor h_n} \big) \, ds . \quad (31b)$$

The proof of the functional central limit theorem (CLT) falls into three major parts. First, we prove the result of Theorem 1 for the right-hand side of (31a). In the second step the approximation error in (31b) is shown to be asymptotically negligible. Finally, we establish that the same functional stable CLT carries over to the adaptive estimators by proving that the error of the plug-in estimation of optimal weights is asymptotically negligible.

Proposition 4.1. *On the assumptions of Theorem 1, it holds true that*

$$n^{\frac{1}{4}}\left(\widehat{\mathbf{IV}}^{or}_{n,t}(\tilde{X}+\epsilon)-h_n\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\sigma^2_{(k-1)h_n}\right)\xrightarrow{st}\int_0^t\sqrt{8\eta\,|\sigma_s^3|}\,dB_s\,, \tag{32}$$

as $n\to\infty$ on $\mathcal{D}\,[0,1]$ where B is a Brownian motion defined on an extension of the original probability space $(\Omega,\mathcal{G},(\mathcal{G}_t)_{0\le t\le 1},\mathbb{P})$, independent of the original σ-algebra $\mathcal{G}$.

Proposition 4.2. *On the assumptions of Theorem 1, it holds true that:*

$$n^{\frac{1}{4}}\Big(\widehat{\mathbf{IV}}^{or}_{n,t}(Y)-\widehat{\mathbf{IV}}^{or}_{n,t}(\tilde{X}+\epsilon)-\int_0^t\left(\sigma_s^2-\sigma^2_{\lfloor sh_n^{-1}\rfloor h_n}\right)ds\Big)\xrightarrow{ucp}0 \text{ as } n\to\infty\,. \tag{33}$$

Theorem 1 is then an immediate consequence of the following proposition:

Proposition 4.3. *On the assumptions of Theorem 1:*

$$n^{\frac{1}{4}}\left|\widehat{\mathbf{IV}}_{n,t}-\widehat{\mathbf{IV}}^{or}_{n,t}(Y)\right|\xrightarrow{ucp}0\quad \text{as } n\to\infty\,. \tag{34}$$

Finally, by consistency of the variance estimators and Slutsky's Lemma the feasible limit theorems for the adaptive estimators are valid. The proof of the functional stable CLT is based on the asymptotic theory developed by Jacod (1997). In order to apply Theorem 3–1 of Jacod (1997) (or equivalently Theorem 2.6 of Podolskij and Vetter (2010), we illustrate the rescaled estimation error as a sum of increments:

$$n^{\frac{1}{4}}\left(\widehat{\mathbf{IV}}^{or}_{n,t}(\tilde{X}+\epsilon)-h_n\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\sigma^2_{(k-1)h_n}\right)=\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\zeta_k^n\,, \tag{35}$$

$$\zeta_k^n=n^{\frac{1}{4}}h_n\sum_{j=1}^{nh_n-1}w_{jk}\left(\tilde{S}^2_{jk}-\mathbb{E}\left[\tilde{S}^2_{jk}|\mathcal{G}_{(k-1)h_n}\right]\right),\ k=1,\ldots,h_n^{-1}. \tag{36}$$

with $\tilde{S}_{jk}$ being spectral statistics build from observations of $\tilde{X}+\epsilon$. For the proof of the functional stable CLT, we need to verify the following five conditions:

(J1) $$\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}\left[\zeta_k^n\middle|\mathcal{G}_{(k-1)h_n}\right]\xrightarrow{ucp}0\,.$$

Convergence of the sum of conditional variances

(J2) $$\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}\left[(\zeta_k^n)^2\left|\mathcal{G}_{(k-1)h_n}\right.\right]\xrightarrow{\mathbb{P}}\int_0^t v_s^2\,ds\,,$$

with a predictable process v_s, and a Lyapunov-type condition

$$\text{(J3)}\qquad \sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}\left[(\zeta_k^n)^4\,\middle|\mathcal{G}_{(k-1)h_n}\right]\stackrel{\mathbb{P}}{\rightarrow}0\,.$$

Finally, stability of weak convergence is ensured if

$$\text{(J4)}\qquad \sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}\left[\zeta_k^n(W_{kh_n}-W_{(k-1)h_n})\middle|\mathcal{G}_{(k-1)h_n}\right]\stackrel{\mathbb{P}}{\rightarrow}0\,,$$

where W is the Brownian motion driving the signal process X and

$$\text{(J5)}\qquad \sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}\left[\zeta_k^n(N_{kh_n}-N_{(k-1)h_n})\middle|\mathcal{G}_{(k-1)h_n}\right]\stackrel{\mathbb{P}}{\rightarrow}0\,,$$

for all bounded martingales N which are orthogonal to W. Next, we strive for a stable CLT for the estimation errors of the covolatility estimator (25a) and the local method of moments approach (29a). A non-degenerate asymptotic variance is obtained when $n/n_p\rightarrow\nu_p$ with $0<\nu_p<\infty$ as $n\rightarrow\infty$ for all $p\in\{1,\ldots,d\}$. We transform the non-synchronous observation model from Assumption (Obs-d) to a synchronous observation model and show that the first order asymptotics of the considered estimators remain invariant. Hence, the effect of non-synchronous sampling on the spectral estimators is shown to be asymptotically negligible. In the idealized martingale framework Bibinger et al. (2013) have found that non-synchronicity effects are asymptotically immaterial in terms of the information content of underlying experiments by a (strong) asymptotic equivalence in the sense of Le Cam of the discrete non-synchronous and a continuous-time observation model. This constitutes a fundamental difference to the non-noisy case where the asymptotic variance of the prominent Hayashi-Yoshida estimator in the functional CLT hinges on interpolation effects, see Hayashi and Yoshida (2011). In the presence of the dominant noise part, however, at the slower optimal convergence rate, the influence of sampling schemes boils down to local observation densities. These time-varying local observation densities are shifted to locally time-varying noise levels (indeed locally increased noise is equivalent to locally less frequent observations). Here, we shall explicitly prove that if we pass from a non-synchronous to a synchronous reference scheme the transformation errors of the estimators are asymptotically negligible.

Lemma 4.4. *Denote $\bar t_i^{(l)}=\big(t_i^{(l)}+t_{i-1}^{(l)}\big)/2, l=1,\ldots,d$. On Assumptions (H-d) and (Obs-d), we can work under synchronous sampling when considering the signal part X, i.e. for $l,m\in\{1,\ldots,d\}$ uniformly in t for both, $w_{jk}^{l,m}$ as in Section 3.3 or defined as entries of* (28)*:*

$$\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}h_n\sum_{j\ge1}w_{jk}^{l,m}\sum_{v=1}^{n_l}\Big(X_{t_v^{(l)}}^{(l)}-X_{t_{v-1}^{(l)}}^{(l)}\Big)\Phi_{jk}(\bar t_v^{(l)})\sum_{i=1}^{n_m}\Big(X_{t_i^{(m)}}^{(m)}-X_{t_{i-1}^{(m)}}^{(m)}\Big)\Phi_{jk}(\bar t_i^{(m)})+\mathcal{O}_{\mathbb{P}}(n^{-1/4})$$
$$=\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}h_n\sum_{j\ge1}w_{jk}^{l,m}\sum_{v=1}^{n_l}\Big(X_{t_v^{(l)}}^{(l)}-X_{t_{v-1}^{(l)}}^{(l)}\Big)\Phi_{jk}(\bar t_v^{(l)})\sum_{i=1}^{n_l}\Big(X_{t_i^{(l)}}^{(m)}-X_{t_{i-1}^{(l)}}^{(m)}\Big)\Phi_{jk}(\bar t_i^{(l)})\,.$$

Note that $(F_l^{-1})',(F_m^{-1})'$ affect the asymptotics of our estimators as can be seen in (11), but are treated as part of the summands due to noise.

Under a synchronous reference observation scheme the strategy of the asymptotic analysis is similar as for the one-dimensional setup. Analogous decompositions in leading terms from the simplified model without drift and with a locally constant covolatility matrix and remainders are considered for the multivariate method of moments estimator (29a) and the spectral covolatility estimator (25a). In order to prove Theorem 2 for instance, we apply Jacod's limit theorem to the sum of increments

$$\zeta_k^n = n^{\frac{1}{4}} h_n \Big(\sum_{j\ge 1} w_{jk}^{p,q} \big(\tilde S_{jk}^{(p)} \tilde S_{jk}^{(q)} \big) - \mathbb{E}\left[\tilde S_{jk}^{(p)} \tilde S_{jk}^{(q)} \Big| \mathcal{G}_{(k-1)h_n} \right] \Big), \tag{38}$$

for $k = 1, \ldots, h_n^{-1}$ with $\tilde S_{jk}^{(p)}$ as defined in (23), but based on observations of $\tilde X + \epsilon$. By including the case $p = q$ with a bias correction the one-dimensional result is generalized to non-equidistant sampling.

The asymptotic negligibility of the plug-in estimation in Proposition 4.3 is proven in Section 6 exploiting a uniform bound on the derivative of the weights as function of σ_t. In fact, it turns out that the weights are robust enough in misspecification of the pre-estimated local volatility to render the difference between oracle and adaptive estimator asymptotically negligible. This carries over to the multivariate methods.

5. Simulations

In the sequel, the one-dimensional spectral integrated volatility estimator's (22a) finite sample performance is investigated in a random volatility simulation scenario. We sample regular observations $Y_1, \ldots, Y_n$ as in (3) with $\epsilon_i \overset{iid}{\sim} N(0, \eta^2)$ and the simulated diffusion

$$X_t = \int_0^t b\, ds + \int_0^t \sigma_s\, dW_s\,.$$

In a first baseline scenario configuration we set $\sigma_s = 1$ constant, and then

$$\sigma_t^2 = \left(\int_0^t \tilde\sigma \cdot \lambda\, dW_s + \int_0^t \sqrt{1-\lambda^2} \cdot \tilde\sigma\, dW_s^{\perp} \right) \cdot f(t)\,, \tag{39}$$

with $W^{\perp}$ a standard Brownian motion independent of W and f a deterministic seasonality function

$$f(t) = 0.1(1 - t^{\frac{1}{3}} + 0.5 \cdot t^2)\,.$$

The drift is set $b = 0.1$ and $\tilde\sigma = 0.01$. The superposition of a continuous semimartingale as random component with a time-varying seasonality modeling volatility's typical U-shape mimics very general realistic volatility characteristics. We implement the oracle version of the estimator (22a) and the adaptive two-stage procedure with pre-estimated optimal weights. Table 1 presents Monte Carlo results for different scenario configurations. In particular, we consider different tuning parameters (bin-widths) and possible dependence of the finite-sample behavior on the leverage magnitude and the magnitude of the noise variance. We compute the estimators' root mean square errors (RMSE) at $t = 1$, for each configuration based on 1000 Monte Carlo iterations, and fix in each configuration one realization of a volatility path to compare the RMSEs to the theoretical asymptotic counterparts in the realized relative efficiency (RE):

$$\text{RE}(\widehat{\mathbf{IV}}_{n,1}) = \frac{\sqrt{\Big((\text{mean}(\widehat{\mathbf{IV}}_{n,1}) - \int_0^1 \sigma_s^2\, ds)^2 + \mathbb{V}\text{ar}(\widehat{\mathbf{IV}}_{n,1}) \Big) \cdot \sqrt{n}}}{\sqrt{8\eta \int_0^1 \sigma_s^3\, ds}}\,. \tag{40}$$

n	σ	h_n^{-1}	η	λ	RE($\widehat{\mathbf{IV}}_{n,1}^{or}$)	RE($\widehat{\mathbf{IV}}_{n,1}$)
30000	1	25	0.01	–	1.01	1.43
5000	1	25	0.01	–	1.02	1.47
30000	(39)	25	0.01	0.5	1.09	1.75
30000	(39)	25	0.01	0.2	1.06	1.77
30000	(39)	25	0.01	0.8	1.09	1.75
30000	(39)	25	0.001	0.5	1.62	1.88
30000	(39)	25	0.1	0.5	1.20	1.69
30000	(39)	50	0.01	0.5	1.09	1.84
30000	(39)	10	0.01	0.5	1.16	1.86
5000	(39)	25	0.01	0.5	1.13	1.92
5000	(39)	50	0.01	0.5	1.08	1.75
5000	(39)	10	0.01	0.5	1.09	1.87

Table 1: Relative Efficiencies (RE) of oracle and adaptive spectral integrated volatility estimator in finite-sample Monte Carlo study.

Our standard sample size is $n = 30000$, a realistic number of observations in usual high-frequency applications as number of ticks over one trading day for liquid assets at NASDAQ. We also focus on smaller samples, $n = 5000$.
Throughout all simulations we fix a maximum spectral cut-off $J_p = 100$ in the pre-estimation step and $J = 150$ for the final estimator, which is large enough to render the approximation error by neglecting higher frequencies negligible. In summary, the Monte Carlo study confirms that the estimator performs well in practice and the Monte Carlo variances come very close to the theoretical lower bound, even in the complex wiggly volatility setting. The fully adaptive approach performs less well than the oracle estimator which is in light of previous results on related estimation approaches not surprising, see e.g. Bibinger and Reiß (2013) for a study including an adaptive multi-scale estimator (global smoothing parameter, but chosen data-driven). Still the adaptive estimator's performance is remarkably well in almost all configurations. Under very small noise level, the relative efficiency is not as close to 1 any more. Apart from this case, the RE comes very close to 1 for the oracle estimator, not depending on the magnitude of leverage, also for small samples, and being very robust with respect to different bin-widths.
A simulation study of the multivariate method of moments estimator in a random volatility setup can be found in Bibinger et al. (2013).

6. Proofs

6.1. Preliminaries

- *Empirical scalar products:*

Definition 1. Let $f, g : [0,1] \to \mathbb{R}$ be functions and $z = (z_i)_{1\le i\le n} \in \mathbb{R}^n$. We call the quantities

$$\langle f, g\rangle_n = \frac{1}{n}\sum_{i=1}^{n} f\left(\frac{i}{n}\right) g\left(\frac{i}{n}\right) ,$$
$$\langle z, g\rangle_n = \frac{1}{n}\sum_{i=1}^{n} z_i\, g\left(\frac{i}{n}\right) ,$$

the *empirical scalar product* of f, g and of z, g, respectively. We further define the "shifted" empirical scalar products

$$[f, g]_n = \frac{1}{n}\sum_{i=1}^{n} f\left(\frac{i-\frac{1}{2}}{n}\right) g\left(\frac{i-\frac{1}{2}}{n}\right),$$
$$[z, g]_n = \frac{1}{n}\sum_{i=1}^{n} z_i\, g\left(\frac{i-\frac{1}{2}}{n}\right).$$

Recall the notation $\Delta^n Y = \left(\Delta_i^n Y\right)_{1\le i\le n} \in \mathbb{R}^n$, the vector of increments and analogously $\Delta^n X$ and let $\epsilon = (\epsilon_i)_{0\le i\le (n-1)}$.

Lemma 6.1. *It holds that*

$$\langle \Phi_{jk}, \Phi_{mk}\rangle_n = \delta_{jm} , \tag{41a}$$

$$[\varphi_{jk}, \varphi_{mk}]_n = \delta_{jm} 4n^2 \sin^2\left(\frac{j\pi}{2nh}\right) . \tag{41b}$$

Moreover, we have the summation by parts decomposition of spectral statistics:

$$\langle n\triangle^n Y, \Phi_{jk}\rangle_n = \langle n\triangle^n X, \Phi_{jk}\rangle_n - [\epsilon, \varphi_{jk}]_n . \tag{41c}$$

Proof. The proofs of the orthogonality relations (41a) and (41b) are similar and we restrict ourselves to prove (41b). In the following we use the shortcut $N = nh_n$ and without loss of generality we consider the first bin $k = 1$. We make use of the trigonometric addition formulas which yield for $N \ge j \ge r \ge 1$:

$$\cos(j\pi N^{-1}(l+\tfrac{1}{2}))\cos(r\pi N^{-1}(l+\tfrac{1}{2})) = \cos((j+r)\pi N^{-1}(l+\tfrac{1}{2}))+\cos((j-r)\pi N^{-1}(l+\tfrac{1}{2})) .$$

The empirical norm for $j = r$ readily follows by $\cos(0) = 1$ and the following. We show that $\sum_{i=0}^{N-1}\cos(m\pi N^{-1}(i+\frac{1}{2})) = 0$ for $m \in \mathbb{N}$. First, consider m odd:

$$\begin{aligned}\sum_{i=0}^{N-1}\cos(m\pi N^{-1}(i+\tfrac{1}{2})) &= \sum_{i=0}^{\lfloor (N-2)/2\rfloor}\cos(m\pi N^{-1}(i+\tfrac{1}{2})) + \sum_{i=\lceil N/2\rceil}^{N-1}\cos(m\pi N^{-1}(i+\tfrac{1}{2}))\\ &= \sum_{i=0}^{\lfloor (N-2)/2\rfloor}\Big(\cos(m\pi N^{-1}(i+\tfrac{1}{2})) + \cos(m\pi N^{-1}(N-(i+\tfrac{1}{2})))\Big)\\ &= 0,\end{aligned}$$

since $\cos(x+\pi m) = -\cos(x)$ for m odd. Note that for $i = (N-1)/2 \in \mathbb{N}$, we leave out one addend which equals $\cos(m\pi/2) = 0$, and also that for m even by $\cos(x) = \cos(x+m\pi)$ the

two sums are equal.
For $m \in \mathbb{N}$ with m even, we differentiate the cases $N = 4k, k \in \mathbb{N}$; $N = 4k+2, k \in \mathbb{N}$ and $N = 2k+1, k \in \mathbb{N}$. If $N = 4k+2$, we decompose the sum as follows:

$$\sum_{i=0}^{N-1} \cos(m\pi N^{-1}(i+\tfrac{1}{2})) = \sum_{i=0}^{2k} \cos(m\pi(4k+2)^{-1}(i+\tfrac{1}{2})) + \sum_{i=2k+1}^{4k+1} \cos(m\pi(4k+2)^{-1}(i+\tfrac{1}{2}))\,.$$

The addends of the left-hand sum are symmetric around the point $m\pi/4$ at $i = k$ and of the right-hand sum around $3m\pi/4$ at $i = 3k+1$. Thereby, both sums equal zero by symmetry. More precisely, for m being not a multiple of 4 the sums directly yield zero. If m is a multiple of 4, we can split the sum into two or more sums which then equal zero again.
This observation for the first sum readily implies $\sum_{i=0}^{N-1} \cos(m\pi N^{-1}(i+\frac{1}{2})) = 0$ for $N = 2k+1$, since in this case

$$\sum_{i=0}^{2k} \cos(m\pi N^{-1}(i+\tfrac{1}{2})) = \sum_{i=1}^{2k} \cos(2m\pi(4k+2)^{-1}(i+\tfrac{1}{2})) = 0\,.$$

For $N = 4k$, we may as well exploit symmetry relations of the cosine. Decompose the sum

$$\sum_{i=0}^{N-1} \cos(m\pi N^{-1}(i+\tfrac{1}{2})) = \sum_{i=0}^{2k-1} \cos(m\pi(4k)^{-1}(i+\tfrac{1}{2})) + \sum_{i=2k}^{4k-1} \cos(m\pi(4k)^{-1}(i+\tfrac{1}{2}))\,.$$

Symmetry around $m\pi/4$ and $3m\pi/4$ is similar as above, but these points lie off the discrete grid this time. Yet, analogous reasoning as above yields that both sums equal zero again, what completes the proof of (41b).
Applying summation by parts to $\langle n\triangle^n\epsilon, \Phi_{jk}\rangle_n$ and using $\Phi_{jk}(1) = \Phi_{jk}(0) = 0$ yields

$$\langle n\triangle^n\epsilon, \Phi_{jk}\rangle_n = \sum_{l=1}^{n} \triangle_l^n \epsilon \Phi_{jk}\left(\frac{l}{n}\right) = -\sum_{l=1}^{n} \epsilon_{l-1}\left(\Phi_{jk}\left(\frac{l}{n}\right) - \Phi_{jk}\left(\frac{l-1}{n}\right)\right)\,.$$

The equality $\sin(x+h) - \sin(x) = 2\sin\left(\frac{h}{2}\right)\cos\left(x+\frac{h}{2}\right)$ for $x, h \in \mathbb{R}$ gives

$$\Phi_{jk}\left(\frac{l}{n}\right) - \Phi_{jk}\left(\frac{l-1}{n}\right) = \frac{1}{n}\varphi_{jk}\left(\frac{l-\frac{1}{2}}{n}\right)$$

what yields the claim. □

- *Basic estimates for drift and Brownian terms:* For all $p \ge 1$ and $s, (s+t) \in [(k-1)h_n, kh_n]$ for some $k = 1, \dots, h_n^{-1}$:

$$\mathbb{E}\left[\|\tilde{X}_{s+t} - \tilde{X}_s\|^p \big| \mathcal{G}_s\right] \le K_p t^{p/2}\,, \tag{42a}$$

with $\tilde{X}_s = \sigma_{(k-1)h_n} \int_{(k-1)h_n}^{s} dW_t$, introduced in Section 4,

$$\mathbb{E}\left[\|X_{s+t} - \tilde{X}_{s+t} - X_s + \tilde{X}_s\|^p \big| \mathcal{G}_s\right] \le K_p \mathbb{E}\left[\left(\int_t^{s+t} \|\sigma_\tau - \sigma_s\|^2\, d\tau\right)^{\frac{p}{2}} \Big| \mathcal{G}_s\right] \le K_p t^p\,, \tag{42b}$$

$$\mathbb{E}\left[\Big\|\int_s^{s+t} b_u\,du\Big\|^p\Big|\mathcal{G}_s\right]\le K_p t^p, \tag{42c}$$

with generic constant K_p depending on p by Itô isometry, Cauchy-Schwarz and Burkholder-Davis Gundy inequalities with Assumption (H-1) and Assumption (H-d), respectively.

- *Local quadratic variations of time:*

$$\sum_{(k-1)h_n\le t_i^{(l)}\le kh_n}\big(t_i^{(l)}-t_{i-1}^{(l)}\big)^2\asymp\sum_{(k-1)h_n\le t_i^{(l)}\le kh_n}H_l^{kh_n}\eta_l^{-2}n_l^{-1}\big(t_i^{(l)}-t_{i-1}^{(l)}\big)=H_l^{kh_n}\eta_l^{-2}n_l^{-1}h_n. \tag{43}$$

The left-hand side is a localized measure of variation in observation times in the vein of the quadratic variation of time by Zhang et al. (2005). It appears in the variance of the estimator and is used to estimate $(F_l^{-1})'((k-1)h_n)$. Under $F_l'\in C^\alpha$ with $\alpha>1/2$ the approximation error by $H_l^{kh_n}$ is $\mathcal{O}(n^{-1/4})$. The asymptotic identity applies to deterministic observation times in deterministic manner and to random exogenous sampling in terms of convergence in probability.

- *Extending local to uniform boundedness:*
On the compact time span $[0,1]$, we can strengthen the structural Assumption (H-1) and assume b_s and $\sigma_s,\tilde b_s,\tilde\sigma_s$ are uniformly bounded. This is based on the localization procedure given in Jacod (2012), Lemma 6. 6 in Section 6. 3.

- *Order of optimal weights:*
Recall the definition of the optimal weights (21). An upper bound for these weights is

$$\begin{aligned}w_{jk}\lesssim I_{jk}=\Big(\sigma^2_{(k-1)h_n}+\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{jk}\right]_n\Big)^{-2}&\lesssim\Big(1+\frac{j^2}{nh_n^2}\Big)^{-2}\\ &\lesssim\begin{cases}1 & \text{for } j\le\sqrt{n}h_n\\ j^{-4}n^2h_n^4 & \text{for } j>\sqrt{n}h_n\end{cases}\end{aligned} \tag{44}$$

what also gives

$$\begin{aligned}\sum_{j=1}^{nh_n-1}w_{jk}\Big(\sigma^2_{(k-1)h_n}+\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{jk}\right]_n\Big)&\lesssim\sum_{j=1}^{\lfloor\sqrt{n}h_n\rfloor}\Big(1+\frac{j^2}{h_n^2n}\Big)+\sum_{j=\lceil\sqrt{n}h_n\rceil}^{nh_n-1}\Big(1+\frac{j^2}{nh_n^2}\Big)j^{-4}n^2h_n^4\\ &\lesssim\sqrt{n}h_n+nh_n^2\,.\end{aligned} \tag{45}$$

6.2. **Proof of Proposition 4.1**

Recall the definition of spectral statistics (18) and denote for $j=1,\ldots,nh_n-1,k=1,\ldots,h_n^{-1}$:

$$\tilde S_{jk}=\Big\langle n(\triangle^n\tilde X+\triangle^n\epsilon),\Phi_{jk}\Big\rangle_n=\Big\langle n\Delta^n\tilde X,\Phi_{jk}\Big\rangle_n-\left[\epsilon,\varphi_{jk}\right]_n\,,$$

where $\tilde X$ is the signal process in the locally parametric experiment. It holds that

$$\begin{aligned}\mathbb{E}\left[\tilde S_{jk}^2\Big|\mathcal{G}_{(k-1)h_n}\right]&=\mathbb{E}\left[\Big(\Big\langle n\triangle^n\tilde X,\Phi_{jk}\Big\rangle_n-\left[\epsilon,\varphi_{jk}\right]_n\Big)^2\Big|\mathcal{G}_{(k-1)h_n}\right]\\ &=\mathbb{E}\left[\Big\langle n\triangle^n\tilde X,\Phi_{jk}\Big\rangle_n^2-2\Big\langle n\triangle^n\tilde X,\Phi_{jk}\Big\rangle_n\left[\epsilon,\varphi_{jk}\right]_n+\left[\epsilon,\varphi_{jk}\right]_n^2\Big|\mathcal{G}_{(k-1)h_n}\right]\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\left[\left\langle n\triangle^n\tilde{X},\Phi_{jk}\right\rangle_n^2\middle|\mathcal{G}_{(k-1)h_n}\right]+\mathbb{E}\left[\left[\epsilon,\varphi_{jk}\right]_n^2\middle|\mathcal{G}_{(k-1)h_n}\right]\\
&= \sigma^2_{(k-1)h_n}+\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{jk}\right]_n . \qquad (46)
\end{aligned}
$$

We have defined ζ_k^n above such that

$$
n^{\frac{1}{4}}\Big(\widetilde{\mathrm{IV}}_{n,t}-h_n\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\sigma^2_{(k-1)h_n}\Big)=n^{\frac{1}{4}}h_n\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\sum_{j=1}^{nh_n-1}w_{jk}\Big(\tilde{S}^2_{jk}-\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{jk}\right]_n-\sigma^2_{(k-1)h_n}\Big)=\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\zeta_k^n
$$

when we shortly express $\widetilde{\mathrm{IV}}_{n,t}=\widehat{\mathbf{IV}}^{or}_{n,t}(\tilde{X}+\epsilon)$. We have to verify (J1)-(J5). (J1) is trivial as the ζ_k^n are centered conditional on $\mathcal{G}_{(k-1)h_n}$. The proof of (J2) is done in two steps. In paragraph 6.2.1 we calculate explicitly the variance which is the left-hand side of (J2). For this we consider at first general weights $w_{jk}\ge 0$, $\sum_{j=1}^{nh_n-1}w_{jk}=1$ which satisfy $w_{jk}\in\mathcal{G}_{(k-1)h_n}$ for all $k=1,\dots,h_n^{-1}$, $j=1,\dots,nh_n-1$. After that we find optimal weights minimizing the variance. In paragraph 6.2.2 we let $n\to\infty$ and calculate the resulting limiting asymptotic variance. The proofs of (J3), (J4) and (J5) follow in paragraph 6.2.3.

6.2.1. **Computation of the variance**

$$
\begin{aligned}
\mathbb{E}\left[(\zeta_k^n)^2\middle|\mathcal{G}_{(k-1)h_n}\right] &= n^{\frac{1}{2}}h_n^2\sum_{j,m=1}^{nh_n-1}w_{jk}w_{mk}\,\mathbb{E}\Big[\Big(\tilde{S}^2_{jk}-\mathbb{E}\left[\tilde{S}^2_{jk}\middle|\mathcal{G}_{(k-1)h_n}\right]\Big)\\
&\qquad\cdot\Big(\tilde{S}^2_{mk}-\mathbb{E}\left[\tilde{S}^2_{mk}\middle|\mathcal{G}_{(k-1)h_n}\right]\Big)\Big|\mathcal{G}_{(k-1)h_n}\Big]\\
&= n^{\frac{1}{2}}h_n^2\sum_{j,m=1}^{nh_n-1}w_{jk}w_{mk}\big(T^n_{j,m,k}(1)+T^n_{j,m,k}(2)+T^n_{j,m,k}(3)\big)\,,
\end{aligned}
$$

with the following three addends:

$$
\begin{aligned}
T^n_{j,m,k}(1)&=\mathbb{E}\left[\left(\left\langle n\triangle^n\tilde{X},\Phi_{jk}\right\rangle_n^2-\sigma^2_{(k-1)h_n}\right)\left(\left\langle n\triangle^n\tilde{X},\Phi_{mk}\right\rangle_n^2-\sigma^2_{(k-1)h_n}\right)\middle|\mathcal{G}_{(k-1)h_n}\right],\\
T^n_{j,m,k}(2)&=\mathbb{E}\left[4\left\langle n\triangle^n\tilde{X},\Phi_{jk}\right\rangle_n\left[\epsilon,\varphi_{jk}\right]_n\left\langle n\triangle^n\tilde{X},\Phi_{mk}\right\rangle_n\left[\epsilon,\varphi_{mk}\right]_n\middle|\mathcal{G}_{(k-1)h_n}\right],\\
T^n_{j,m,k}(3)&=\mathbb{E}\left[\left(\left[\epsilon,\varphi_{jk}\right]_n^2-\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{jk}\right]_n\right)\left(\left[\epsilon,\varphi_{mk}\right]_n^2-\frac{\eta^2}{n}\left[\varphi_{mk},\varphi_{mk}\right]_n\right)\middle|\mathcal{G}_{(k-1)h_n}\right]
\end{aligned}
$$

for frequencies j,m. Independence of the noise and of the Brownian increments yield the identities

$$
\begin{aligned}
\mathbb{E}\left[\left[\epsilon,\varphi_{jk}\right]_n\left[\epsilon,\varphi_{mk}\right]_n\right]&=\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{mk}\right]_n\,,\\
\mathbb{E}\left[\left\langle n\triangle^n\tilde{X},\Phi_{jk}\right\rangle_n\left\langle n\triangle^n\tilde{X},\Phi_{mk}\right\rangle_n\middle|\mathcal{G}_{(k-1)h_n}\right]&=\delta_{jm}\sigma^2_{(k-1)h_n},
\end{aligned}
$$

which already imply

$$
T^n_{j,m,k}(2)=4\frac{\eta^2}{n}\delta_{jm}\left[\varphi_{jk},\varphi_{mk}\right]_n\sigma^2_{(k-1)h_n}\,,
$$

because noise and signal are independent. We further obtain by another polynomial expansion

$$\mathbb{E}\left[[\epsilon,\varphi_{jk}]_n^2\,[\epsilon,\varphi_{mk}]_n^2\right]$$
$$= n^{-4}\sum_{l,l',p,p'=1}^{n}\Big(\mathbb{E}\left[\epsilon_l\epsilon_{l'}\epsilon_p\epsilon_{p'}\right]\varphi_{jk}\Big(\frac{l-\frac12}{n}\Big)\varphi_{jk}\Big(\frac{l'-\frac12}{n}\Big)\varphi_{mk}\Big(\frac{p-\frac12}{n}\Big)\varphi_{mk}\Big(\frac{p'-\frac12}{n}\Big)\Big).$$

Only the cases $l=l'\neq p=p'$, $l=p\neq l'=p'$, $l=p'\neq l'=p$ or $l=l'=p=p'$ produce non-zero results in the expectation. Hence, denoting by $\eta'=\mathbb{E}[\epsilon_1^4]$ the fourth moment of the observation errors, we end up with

$$\mathbb{E}\left[[\epsilon,\varphi_{jk}]_n^2\,[\epsilon,\varphi_{mk}]_n^2\right]=\frac{1}{n^4}\sum_{l,l',p,p'}\left(\eta^4\left(\delta_{ll'}\delta_{pp'}+\delta_{lp}\delta_{l'p'}+\delta_{lp'}\delta_{l'p}\right)+\eta'\delta_{lp}\delta_{l'p'}\delta_{ll'}-3\eta^4\delta_{lp}\delta_{l'p'}\delta_{ll'}\right)$$
$$\cdot\Big(\varphi_{jk}\Big(\frac{l-\frac12}{n}\Big)\varphi_{jk}\Big(\frac{l'-\frac12}{n}\Big)\varphi_{mk}\Big(\frac{p-\frac12}{n}\Big)\varphi_{mk}\Big(\frac{p'-\frac12}{n}\Big)\Big)$$
$$=\frac{\eta^4}{n^2}\left([\varphi_{jk},\varphi_{jk}]_n\,[\varphi_{mk},\varphi_{mk}]_n+2\,[\varphi_{jk},\varphi_{mk}]_n^2\right)+\frac{\eta'-3\eta^4}{n^4}\sum_{l=1}^{n}\Big(\varphi_{jk}^2\Big(\frac{l-\frac12}{n}\Big)\varphi_{mk}^2\Big(\frac{l-\frac12}{n}\Big)\Big).$$

Arguing similarly and using that $\mathbb{E}[(\triangle_l^nW)^4]=3\,\mathbb{E}[(\triangle_l^nW)^2]$ for $l\in\mathbb{N}$, we obtain

$$\mathbb{E}\left[\left\langle n\triangle^n\tilde X,\Phi_{jk}\right\rangle_n^2\left\langle n\triangle^n\tilde X,\Phi_{mk}\right\rangle_n^2\middle|\,\mathcal{G}_{(k-1)h_n}\right]$$
$$=\sigma_{(k-1)h_n}^4\left(\langle\Phi_{jk},\Phi_{jk}\rangle_n\langle\Phi_{mk},\Phi_{mk}\rangle_n+2\langle\Phi_{jk},\Phi_{mk}\rangle_n^2\right)=\sigma_{(k-1)h_n}^4\left(1+2\delta_{jm}\right)\,.$$

From the identities so far we obtain

$$T_{j,m,k}^n(1)=\mathbb{E}\left[\left\langle n\triangle^n\tilde X,\Phi_{jk}\right\rangle_n^2\left\langle n\triangle^n\tilde X,\Phi_{mk}\right\rangle_n^2\middle|\,\mathcal{G}_{(k-1)h_n}\right]$$
$$-\mathbb{E}\left[\left\langle n\triangle^n\tilde X,\Phi_{jk}\right\rangle_n^2\middle|\,\mathcal{G}_{(k-1)h_n}\right]\mathbb{E}\left[\left\langle n\triangle^n\tilde X,\Phi_{mk}\right\rangle_n^2\middle|\,\mathcal{G}_{(k-1)h_n}\right]$$
$$=\sigma_{(k-1)h_n}^4\left(1+2\delta_{mj}\right)-\sigma_{(k-1)h_n}^4=2\delta_{jm}\sigma_{(k-1)h_n}^4,$$
$$T_{j,m,k}^n(3)=\mathbb{E}\left[\left([\epsilon,\varphi_{jk}]_n^2-\frac{\eta^2}{n}\,[\varphi_{jk},\varphi_{jk}]_n\right)\left([\epsilon,\varphi_{mk}]_n^2-\frac{\eta^2}{n}\,[\varphi_{mk},\varphi_{mk}]_n\right)\right]$$
$$=\mathbb{E}\left[[\epsilon,\varphi_{jk}]_n^2\,[\epsilon,\varphi_{mk}]_n^2\right]-\frac{\eta^4}{n^2}\,[\varphi_{jk},\varphi_{jk}]_n\,[\varphi_{mk},\varphi_{mk}]_n$$
$$=\frac{2\eta^4}{n^2}\,[\varphi_{jk},\varphi_{mk}]_n^2+\frac{\eta'-3\eta^4}{n^3}\,\left[\varphi_{jk}^2,\varphi_{mk}^2\right]_n\,.$$

In all, the conditional variance is given by

$$\mathbb{E}\left[(\zeta_k^n)^2\middle|\,\mathcal{G}_{(k-1)h_n}\right]$$
$$=\sqrt{n}h_n^2\left(\sum_{j=1}^{nh_n-1}w_{jk}^2\left(2\sigma_{(k-1)h_n}^4+4\frac{\eta^2}{n}\sigma_{(k-1)h_n}^2\,[\varphi_{jk},\varphi_{jk}]_n+\frac{2\eta^4}{n^2}\,[\varphi_{jk},\varphi_{jk}]_n^2\right)\right)+R_n$$
$$=\sqrt{n}h_n^2\sum_{j=1}^{nh_n-1}w_{jk}^2\,2\left(\sigma_{(k-1)h_n}^2+\frac{\eta^2}{n}\,[\varphi_{jk},\varphi_{jk}]_n\right)^2+R_n$$

with remainder $R_n = \frac{\eta'-3\eta^4}{n^3}[\varphi_{jk}^2,\varphi_{mk}^2]_n$. Observe that $R_n = 0$ for Gaussian noise. In this case, analogous to Bibinger and Reiß (2013), we find that the optimal weights minimizing the variance, under the constraint $\sum_{j=1}^{nh_n-1} w_{jk} = 1$, which assures unbiasedness of the estimator, are given by (21). The optimization can be done with Lagrange multipliers. R_n is then a remainder in case that $\eta' \neq 3\eta^4$. With the weights (21) and using (45), we can bound R_n by:

$$\begin{aligned}
R_n &= \sqrt{n}h_n^2\frac{\eta'-3\eta^4}{n^4}\left(\sum_{i=1}^{n}\left(\sum_{j=1}^{nh_n-1} w_{jk}\varphi_{jk}^2\left(\frac{i-\frac{1}{2}}{n}\right)\right)^2\right)\\
&\lesssim \sqrt{n}h_n^2\frac{1}{n^4}\left(n\left(\frac{n}{h_n}\right)^2\left(\sum_{j=1}^{nh_n-1} w_{jk}\left(\sigma_{(k-1)h_n}^2+\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{jk}\right]_n\right)\right)^2\right)\\
&\lesssim n^{-1/2}\left(\sqrt{n}h_n + nh_n^2\right)^2 = \mathcal{O}(1)\,.
\end{aligned}$$

We therefore obtain

$$\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}\left[(\zeta_k^n)^2\middle|\mathcal{G}_{(k-1)h_n}\right] = \sqrt{n}h_n^2\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\sum_{j=1}^{nh_n-1}(I_k^{-2}I_{jk}^2)I_{jk}^{-1}+\mathcal{O}(1) = \sqrt{n}h_n^2\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}I_k^{-1}+\mathcal{O}(1)$$

as variance of the estimator.

6.2.2. **The asymptotic variance of the estimator**

The key to the asymptotic variance is to recognize

$$(\sqrt{n}h_n)^{-1}I_k = \frac{1}{\sqrt{n}h_n}\sum_{j=1}^{nh_n-1}\frac{1}{2}\left(\sigma_{(k-1)h_n}^2+\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{jk}\right]_n\right)^{-2}$$

as a Riemann sum, ending up with the "double-Riemann-sum" $\sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n((\sqrt{n}h_n)^{-1}I_k)^{-1}$. The scaling factor $(\sqrt{n}h_n)^{-1}$ is the right choice for the first Riemann sum which becomes clear after two Taylor expansions. First, expanding the sine for each frequency j we find $0 \le \xi_j \le j\pi/(2nh_n)$ with

$$I_{jk} = \frac{1}{2}\left(\sigma_{(k-1)h_n}^2+4\eta^2 n\left(\frac{j\pi}{2nh_n}-\frac{\xi_j^3}{6}\right)^2\right)^{-2}.$$

Second, we expand $x \mapsto \frac{1}{2}\left(\sigma_{(k-1)h_n}^2+4\eta^2 nx^2\right)^{-2}$ which yields $\frac{j\pi}{2nh_n}-\frac{\xi_j^3}{6} \le \xi_j' \le \frac{j\pi}{2nh_n}$ such that

$$I_{jk} = \tilde{I}_{jk}+R_{jk} \quad\text{with}\quad R_{jk} = \frac{4\eta^2 n\xi_j'}{(\sigma_{(k-1)h_n}^2+4\eta^2 n\xi_j'^2)^3}\frac{\xi_j^3}{6} \tag{47}$$

where we define $\tilde{I}_{jk} = \frac{1}{2}(\sigma_{(k-1)h_n}^2+\eta^2(\frac{j\pi}{2\sqrt{n}h_n})^2)^{-2}$. Now it becomes clear that $\sqrt{n}h_n$ is indeed the right factor because

$$\left|\frac{1}{\sqrt{n}h_n}\sum_{j=1}^{nh_n-1}\tilde{I}_{jk}-\int_0^{\sqrt{n}-\frac{1}{\sqrt{n}h_n}}\frac{1}{2}\left(\sigma_{(k-1)h_n}^2+\eta^2\pi^2x^2\right)^{-2}dx\right|$$

$$
\begin{aligned}
&= \left| \sum_{j=1}^{nh_n-1} \int_{\frac{j-1}{\sqrt{n}h_n}}^{\frac{j}{\sqrt{n}h_n}} \left(\frac{1}{2}\big(\sigma^2_{(k-1)h_n} + \eta^2\pi^2 j^2 h_n^{-2} n^{-1}\big)^{-2} - \frac{1}{2}\big(\sigma^2_{(k-1)h_n} + \eta^2\pi^2 x^2\big)^{-2} \right) dx \right| \\
&\lesssim \sum_{j=1}^{nh_n-1} \int_{\frac{j-1}{\sqrt{n}h_n}}^{\frac{j}{\sqrt{n}h_n}} \left| x - \frac{j}{\sqrt{n}h_n} \right| dx \max_{\frac{j-1}{\sqrt{n}h_n} \le y \le \frac{j}{\sqrt{n}h_n}} \left(y\big(\sigma^2_{(k-1)h_n} + \eta^2\pi^2 y^2\big)^{-3} \right) \\
&\le \left(\frac{1}{\sqrt{n}h_n} \right)^2 \left(\sum_{j=1}^{nh_n-1} \left(\max_{\frac{j-1}{\sqrt{n}h_n} \le y \le \frac{j}{\sqrt{n}h_n}} \left(y\big(\sigma^2_{(k-1)h_n} + \eta^2\pi^2 y^2\big)^{-3} \right) \right) \right) \\
&= \left(\frac{1}{\sqrt{n}h_n} \right)^2 \left(\sum_{j=1}^{\lfloor \sqrt{n}h_n \rfloor} \frac{j}{\sqrt{n}h_n} + \sum_{j=\lceil \sqrt{n}h_n \rceil}^{nh_n-1} \left(\frac{\sqrt{n}h_n}{j-1} \right)^5 \right) \\
&\lesssim \left(\frac{1}{\sqrt{n}h_n} \right)^2 \left(\sqrt{n}h_n + \sum_{j=1}^{nh_n-1-\lceil \sqrt{n}h_n \rceil} \left(\frac{\sqrt{n}h_n}{j + \lceil \sqrt{n}h_n \rceil} \right)^5 \right) \le C \left(\frac{1}{\sqrt{n}h_n} \right)^2 .
\end{aligned}
$$

for some positive constant C which does neither depend on $\sigma^2_{(k-1)h_n}$ nor on n. We choose h_n such that $\sqrt{n}h_n \to \infty$. Though we consider all possible spectral frequencies $j = 1, \dots, nh_n - 1$, we shall see in the following that the I_{jk} for $j \ge \lceil n^\beta h_n \rceil$ become asymptotically negligible for a suitable $0 \le \beta < 1$. By virtue of monotonicity of the sine on $\left[0, \frac{\pi}{2}\right]$ and $\sin(x) \ge x/2$ for $0 \le x \le 1$, it follows that

$$
\begin{aligned}
\frac{1}{\sqrt{n}h_n} \sum_{j=\lceil n^\beta h_n \rceil}^{nh_n-1} I_{jk} &\lesssim \frac{1}{\sqrt{n}h_n} \sum_{j=\lceil n^\beta h_n \rceil}^{nh_n-1} \left(n \sin^2\left(\frac{n^\beta h_n \pi}{2nh_n} \right) \right)^{-2} \\
&\le \frac{1}{\sqrt{n}h_n} nh_n \left(n \sin^2\left(\frac{n^\beta h_n \pi}{2nh_n} \right) \right)^{-2} \\
&\le \sqrt{n} \left(n \left(\frac{n^{\beta-1}\pi}{4} \right)^2 \right)^{-2} \lesssim n^{\frac{1}{2}-4\beta+2} = n^{\frac{5}{2}-4\beta} .
\end{aligned}
$$

We deduce that $\frac{1}{\sqrt{n}h_n} \sum_{j=\lceil n^\beta h_n \rceil}^{nh_n-1} I_{jk} = \mathcal{O}(1)$, for every $5/8 < \beta < 1$. Moreover, we obtain for the first $\lfloor n^\beta h_n \rfloor$ summands of the remainder term

$$
\begin{aligned}
\frac{1}{\sqrt{n}h_n} \sum_{j=1}^{\lfloor n^\beta h_n \rfloor} R_{jk} &= \frac{1}{\sqrt{n}h_n} \sum_{j=1}^{\lfloor n^\beta h_n \rfloor} \frac{4\eta^2 n \xi_j'}{\left(\sigma^2_{(k-1)h_n} + 4\eta^2 n \xi_j'^2 \right)^3} \frac{\xi_j^3}{6} \lesssim \frac{n}{\sqrt{n}h_n} \sum_{j=1}^{\lfloor n^\beta h_n \rfloor} \left(\xi_j^3 \xi_j' \right) \\
&\le \frac{n}{\sqrt{n}h_n} \sum_{j=1}^{\lfloor n^\beta h_n \rfloor} \left(\frac{j\pi}{nh_n} \right)^4 \lesssim \frac{1}{\sqrt{n}h_n} n^\beta h_n n^{4(\beta-1)+1} = n^{5\beta - \frac{7}{2}} .
\end{aligned}
$$

Hence $\frac{1}{\sqrt{n}h_n} \sum_{j=1}^{\lfloor n^\beta h_n \rfloor} R_{jk} = \mathcal{O}(1)$ for every $\beta < 7/10$. As the tails are asymptotic negligible we thus have $\frac{1}{\sqrt{n}h_n} \sum_{j=1}^{nh_n-1} R_{jk} = \mathcal{O}(1)$ and, in particular,

$$
\frac{1}{\sqrt{n}h_n} \sum_{j=1}^{nh_n-1} I_{jk} = \int_0^{\sqrt{n} - \frac{1}{\sqrt{n}h_n}} \frac{1}{2} \left(\sigma^2_{(k-1)h_n} + \eta^2\pi^2 x^2 \right)^{-2} dx + \mathcal{O}(1) .
$$

Substitution and an application of the recursion formula

$$\int_0^y (b^2+(x-a)^2)^{-k}\,dx = \left.\frac{x-a}{2(k-1)b^2\big(b^2+(x-a)^2\big)^{k-1}}\right|_0^y + \frac{2k-3}{2\,(k-1)\,b^2}\int_0^y \big(b^2+(x-a)^2\,\big)^{1-k}dx$$

for $y \geq 0$, $k=2, a=0$ and $b=1$, yields

$$\begin{aligned}
&\int_0^y \frac{1}{2}\left(\sigma^2_{(k-1)h_n} + \eta^2\pi^2 x^2\right)^{-2}\,dx \\
&= \int_0^y \frac{1}{2\sigma^4_{(k-1)h_n}}\left(1+\left(\frac{\eta\pi}{\sigma_{(k-1)h_n}}x\right)^2\right)^{-2}\,dx \\
&= \frac{1}{2\eta\pi\left|\sigma_{(k-1)h_n}\right|^3}\int_0^{\frac{\eta\pi}{\left|\sigma_{(k-1)h_n}\right|}y}\left(1+x^2\right)^{-2}\,dx \\
&= \frac{1}{2\eta\pi\left|\sigma_{(k-1)h_n}\right|^3}\left(\frac{\frac{\eta\pi}{\left|\sigma_{(k-1)h_n}\right|}y}{2\left(1+\left(\frac{\eta\pi}{\left|\sigma_{(k-1)h_n}\right|}y\right)^2\right)} + \frac{1}{2}\int_0^{\frac{\eta\pi}{\left|\sigma_{(k-1)h_n}\right|}y}\left(1+x^2\right)^{-1}\,dx\right) \\
&= \frac{y}{4\left|\sigma_{(k-1)h_n}\right|^4\left(1+\left(\frac{\eta\pi}{\left|\sigma_{(k-1)h_n}\right|}y\right)^2\right)} + \frac{1}{4\eta\pi\left|\sigma_{(k-1)h_n}\right|^3}\arctan\left(\frac{\eta\pi}{\left|\sigma_{(k-1)h_n}\right|}y\right).
\end{aligned}$$

As $\kappa < |\sigma_s| < C$ uniformly for all $0 \leq s \leq 1$ and because $\arctan(x) \to \pi/2$ as $x \to \infty$, as well as $\sqrt{n} - \frac{1}{\sqrt{n}h_n} \to \infty$ as $n \to \infty$, we have

$$\begin{aligned}
\frac{1}{\sqrt{n}h_n}\sum_{j=1}^{nh_n-1} I_{jk} &= \frac{\sqrt{n}-\frac{1}{\sqrt{n}h_n}}{4\left|\sigma_{(k-1)h_n}\right|^4\left(1+\left(\frac{\eta\pi}{\left|\sigma_{(k-1)h_n}\right|}\left(\sqrt{n}-\frac{1}{\sqrt{n}h_n}\right)\right)^2\right)} \\
&\quad + \frac{1}{4\eta\pi\left|\sigma_{(k-1)h_n}\right|^3}\arctan\left(\frac{\eta\pi}{\left|\sigma_{(k-1)h_n}\right|}\left(\sqrt{n}-\frac{1}{\sqrt{n}h_n}\right)\right) + \mathcal{O}(1) \\
&= \frac{1}{8\eta\left|\sigma_{(k-1)h_n}\right|^3} + \mathcal{O}(1).
\end{aligned}$$

The final step in the proof is another Taylor approximation:

$$\begin{aligned}
\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}\left[(\zeta_k^n)^2\middle|\mathcal{G}_{(k-1)h_n}\right] = \sqrt{n}h_n^2\sum_{k=1}^{\lfloor th_n^{-1}\rfloor} I_k^{-1} &= h_n\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\left(\frac{1}{\sqrt{n}h_n}\sum_{j=1}^{nh_n-1} I_{jk}\right)^{-1} \\
&= h_n\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\left(\frac{1}{8\eta\left|\sigma_{(k-1)h_n}\right|^3} + \mathcal{O}(1)\right)^{-1} \\
&= \left(h_n\sum_{k=1}^{\lfloor th_n^{-1}\rfloor} 8\eta\left|\sigma_{(k-1)h_n}\right|^3\right) + \mathcal{O}(1).
\end{aligned}$$

The last equality is true by Taylor and because σ is uniformly bounded. Because σ is continuous we obtain the claim by Riemann approximation, i.e.

$$\sum_{k=1}^{\lfloor th_n^{-1}\rfloor} \mathbb{E}\left[(\zeta_k^n)^2 \middle| \mathcal{G}_{(k-1)h_n} \right] \to 8\eta \int_0^t |\sigma_s|^3 \, ds$$

almost surely as $n \to \infty$ establishing (J2) with the asymptotic expression of Theorem 1.

6.2.3. Lyapunov's criterion and stability of convergence

So far, we have proved (J1) and (J2). Next, we shall prove that the Lyapunov condition (J3) is satisfied. For the sum of fourth moments, we obtain by Minkowski's inequality, Jensen's inequality and $w_{jk} \in \mathcal{G}_{(k-1)h_n}$ for all $k = 1, \ldots, h_n^{-1}$ and $j = 1, \ldots, nh_n - 1$:

$$\begin{aligned}
\mathbb{E}\left[(\zeta_k^n)^4 \middle| \mathcal{G}_{(k-1)h_n} \right] &= nh_n^4 \left(\mathbb{E}\left[\left(\sum_{j=1}^{nh_n-1} w_{jk} \left(\tilde{S}_{jk}^2 - \mathbb{E}\left[\tilde{S}_{jk}^2 \middle| \mathcal{G}_{(k-1)h_n} \right] \right) \right)^4 \middle| \mathcal{G}_{(k-1)h_n} \right] \right) \\
&\leq nh_n^4 \left(\sum_{j=1}^{nh_n-1} w_{jk} \left(\mathbb{E}\left[\left(\tilde{S}_{jk}^2 - \mathbb{E}\left[\tilde{S}_{jk}^2 \middle| \mathcal{G}_{(k-1)h_n} \right] \right)^4 \middle| \mathcal{G}_{(k-1)h_n} \right] \right)^{\frac{1}{4}} \right)^4 \\
&\lesssim nh_n^4 \left(\sum_{j=1}^{nh_n-1} w_{jk} \left(\sigma_{(k-1)h_n}^8 + \frac{\eta^8}{n^4} \left[\varphi_{jk}, \varphi_{jk} \right]_n^4 \right)^{\frac{1}{4}} \right)^4 \\
&\lesssim nh_n^4 \left(\sum_{j=1}^{nh_n-1} w_{jk} \left(\mathbb{E}\left[\tilde{S}_{jk}^8 \middle| \mathcal{G}_{(k-1)h_n} \right] \right)^{\frac{1}{4}} \right)^4
\end{aligned}$$

If we can show

$$\mathbb{E}\left[\left\langle n\triangle^n \tilde{X}, \Phi_{jk} \right\rangle_n^8 \middle| \mathcal{G}_{(k-1)h_n} \right] \lesssim \sigma_{(k-1)h_n}^8, \tag{48}$$

$$\mathbb{E}\left[\left[\epsilon, \varphi_{jk} \right]_n^8 \right] \lesssim \frac{\eta^8 \left[\varphi_{jk}, \varphi_{jk} \right]_n^4}{n^4}, \tag{49}$$

then we are able to conclude that

$$\begin{aligned}
\mathbb{E}\left[\tilde{S}_{jk}^8 \middle| \mathcal{G}_{(k-1)h_n} \right] &\lesssim \mathbb{E}\left[\left\langle n\Delta^n \tilde{X}, \Phi_{jk} \right\rangle_n^8 \middle| \mathcal{G}_{(k-1)h_n} \right] + \mathbb{E}\left[\left[\epsilon, \varphi_{jk} \right]_n^8 \right] \\
&\lesssim \sigma_{(k-1)h_n}^8 + \frac{\eta^8}{n^4} \left[\varphi_{jk}, \varphi_{jk} \right]_n^4 \leq \left(\sigma_{(k-1)h_n}^2 + \frac{\eta^2}{n} \left[\varphi_{jk}, \varphi_{jk} \right]_n \right)^4 .
\end{aligned} \tag{50}$$

Hence, we obtain by (45)

$$\sum_{k=1}^{h_n^{-1}} \mathbb{E}\left[(\zeta_k^n)^4 \middle| \mathcal{G}_{(k-1)h_n} \right] \lesssim \sum_{k=1}^{h_n^{-1}} nh_n^4 \left(\sum_{j=1}^{nh_n-1} w_{jk} \left(\sigma_{(k-1)h_n}^2 + \frac{\eta^2}{n} \left[\varphi_{jk}, \varphi_{jk} \right]_n \right) \right)^4 \lesssim n^2 h_n^6 = \mathcal{O}(1)$$

which proves (J3). We are therefore left with proving (48) and (49). The first inequality holds because $\langle n\triangle^n \tilde{X}, \Phi_{jk} \rangle_n$ is $N(0, \sigma_{(k-1)h_n}^2)$-distributed conditional on $\mathcal{G}_{(k-1)h_n}$. Higher moments of the noise term can be treated with techniques as in the proofs of method of moments, see for instance Tao (2012). In order to see why the second inequality is satisfied, let $g_l = \epsilon_{((k-1)nh_n+l)} \varphi_{jk}\left(\frac{(k-1)nh_n+l-\frac{1}{2}}{n}\right)$ for $l = 1, \ldots, nh_n$, such that Polynomial expansion yields

$$\mathbb{E}\left[[\epsilon, \varphi_{jk}]_n^8\right] = n^{-8} \sum_{1 \le l_1, \ldots, l_8 \le nh_n} \mathbb{E}\left[g_{l_1} \cdots g_{l_8}\right]. \tag{51}$$

As the first eight moments of the noise exist, we obtain for each summand the same bound

$$|\mathbb{E}\left[g_{l_1} \cdots g_{l_8}\right]| \quad \lesssim \quad \eta^8 \left|\varphi_{jk}\left(\frac{l_1 - \frac{1}{2}}{n}\right) \cdots \varphi_{jk}\left(\frac{l_8 - \frac{1}{2}}{n}\right)\right| \lesssim \frac{\eta^8 \left[\varphi_{jk}, \varphi_{jk}\right]_n^4}{h_n^4}.$$

Moreover, the g_l are centered and independent. Therefore, we only have to consider summands where each g_l appears at least twice. In particular, in every summand there are at most four distinct g_l. Thus, our problem simplifies to

$$\left|\mathbb{E}\left[[\epsilon, \varphi_{jk}]_n^8\right]\right| \lesssim n^{-8} \frac{\eta^8 \left[\varphi_{jk}, \varphi_{jk}\right]_n^4}{h_n^4} \left(\sum_{r=0}^{4} N_r\right), \tag{52}$$

where N_r is the number of ways one can assign integers $l_1, \ldots, l_8$ in $\{1, \ldots, nh_n\}$ such that each l_i appears at least twice, and such that exactly $(4-r)$ integers appear. A crude bound can be obtained by combinatorial considerations and Stirling's formula such that

$$N_r \le (enh_n)^{4-r}\, 4^{4+r} \lesssim n^4 h_n^4.$$

Inserting this into (52) yields the claim in (49).
It remains to verify (J4) and (J5). Consider the telescoping sum

$$W_{kh_n} - W_{(k-1)h_n} = \sum_{m=(k-1)nh_n+1}^{knh_n} \triangle_m^n W.$$

By linearity it is enough to consider only one summand $\triangle_m^n W$ for some $m = (k-1)\,nh_n + 1, \ldots, knh_n$:

$$\begin{aligned}
&\mathbb{E}\left[\zeta_k^n \triangle_m^n W \middle| \mathcal{G}_{(k-1)h_n}\right] \\
&= n^{1/4} h_n \sum_{j=1}^{nh_n-1} \left(w_{jk} \mathbb{E}\left[\tilde{S}_{jk}^2 \triangle_m^n W \middle| \mathcal{G}_{(k-1)h_n}\right] - \mathbb{E}\left[\tilde{S}_{jk}^2 \middle| \mathcal{G}_{(k-1)h_n}\right] \mathbb{E}\left[\triangle_m^n W \middle| \mathcal{G}_{(k-1)h_n}\right]\right) \\
&= n^{1/4} h_n \sum_{j=1}^{nh_n-1} \Bigg(w_{jk} \mathbb{E}\Bigg[\left\langle n\triangle^n \tilde{X}, \Phi_{jk}\right\rangle_n^2 \triangle_m^n W - 2\left\langle n\triangle^n \tilde{X}, \Phi_{jk}\right\rangle_n [\epsilon, \varphi_{jk}]_n \triangle_m^n W \\
&\qquad\qquad + [\epsilon, \varphi_{jk}]_n^2 \triangle_m^n W \Bigg| \mathcal{G}_{(k-1)h_n}\Bigg]\Bigg) \\
&= n^{1/4} h_n \sum_{j=1}^{nh_n-1} \left(w_{jk} \mathbb{E}\left[\left\langle n\triangle^n \tilde{X}, \Phi_{jk}\right\rangle_n^2 \triangle_m^n W \middle| \mathcal{G}_{(k-1)h_n}\right]\right).
\end{aligned}$$

The second equality holds because the $\triangle_m^n W$ are centered. The last one holds because the noise is centered and because noise and signal are independent. The expectation, however, vanishes for all frequencies j which follows by independence of the Brownian increments:

$$\mathbb{E}\left[\left\langle n\triangle^n \tilde{X}, \Phi_{jk}\right\rangle_n^2 \triangle_m^n W \middle| \mathcal{G}_{(k-1)h_n}\right]$$
$$= \sum_{l,p=1}^{n} \sigma_{(k-1)h_n}^2 \mathbb{E}\left[\triangle_l^n W \triangle_p^n W \triangle_m^n W\right|\right] \Phi_{jk}\left(\frac{l}{n}\right)\Phi_{jk}\left(\frac{p}{n}\right) = 0.$$

Therefore $\sum_{k=1}^{\lfloor t h_n^{-1}\rfloor} \mathbb{E}\left[\zeta_k^n \triangle_k^{h_n^{-1}} W \middle| \mathcal{G}_{(k-1)h_n}\right] = 0 \quad$ for all $0 \le t \le 1$.

Let N be a bounded $(\mathcal{G}_t)_{0\le t\le 1}$-martingale with $N_0 = 0$ and $\langle W, N\rangle \equiv 0$. For the telescoping sum

$$N_{kh_n} - N_{(k-1)h_n} = \sum_{m=(k-1)nh+1}^{knh} \triangle_m^n N$$

by linearity it is enough to consider $\mathbb{E}\left[\zeta_k^n \triangle_m^n N \right| \mathcal{G}_{(k-1)h_n}\right]$ for some $m = (k-1)\,nh_n + 1, \ldots, knh_n$. Just like above we end up with

$$\mathbb{E}\left[\zeta_k^n \triangle_m^n N \right| \mathcal{G}_{(k-1)h_n}\right] \;=\; n^{1/4} h_n \sum_{j=1}^{nh_n - 1} w_{jk} \mathbb{E}\left[\left\langle n\triangle^n \tilde{X}, \Phi_{jk}\right\rangle_n^2 \triangle_m^n N \middle| \mathcal{G}_{(k-1)h_n}\right].$$

The expectation

$$\mathbb{E}\left[\left\langle n\triangle^n \tilde{X}, \Phi_{jk}\right\rangle_n^2 \triangle_m^n N \middle| \mathcal{G}_{(k-1)h_n}\right] = \sum_{l,p=1}^{n} \sigma_{(k-1)h_n}^2 \mathbb{E}\left[\triangle_l^n W \triangle_p^n W \triangle_m^n N\right|\right] \Phi_{jk}\left(\frac{l}{n}\right)\Phi_{jk}\left(\frac{p}{n}\right).$$

vanishes, except for $l = p = m$, $p < l = m$ or $l < p = m$. The last two cases are handled similarly. For instance, for $l < p = m$, the process $(W_s N_s)_{0\le s\le 1}$ is a $(\mathcal{G}_s)_{0\le s\le 1}$-martingale because $\langle N, W\rangle \equiv 0$ such that

$$\mathbb{E}\left[\triangle_l^n W \triangle_p^n W \triangle_m^n N \right| \mathcal{G}_{(k-1)h_n}\right] = \mathbb{E}\left[\triangle_l^n W \mathbb{E}\left[\triangle_m^n W \triangle_m^n N \right| \mathcal{G}_{\frac{m-1}{n}}\right] \middle| \mathcal{G}_{(k-1)h_n}\right] = 0.$$

With respect to the first case we obtain by Itô's formula

$$\mathbb{E}\left[\left(\triangle_l^n W\right)^2 \triangle_l^n N \middle| \mathcal{G}_{(k-1)h_n}\right] = \mathbb{E}\left[\left(\left(\triangle_l^n W\right)^2 - \left(\frac{1}{n}\right)\right) \triangle_l^n N \middle| \mathcal{G}_{(k-1)h_n}\right]$$
$$+ \mathbb{E}\left[\left(\frac{1}{n}\right) \triangle_l^n N \middle| \mathcal{G}_{(k-1)h_n}\right] = \mathbb{E}\left[\left(\int_{\frac{l-1}{n}}^{\frac{l}{n}} W_s\, dW_s\right) \triangle_l^n N \middle| \mathcal{G}_{(k-1)h_n}\right].$$

However, $\left(\left(\int_0^t W_s\, dW_s\right)\cdot N_t\right)_{0\le t\le 1}$ is also a $(\mathcal{G}_t)_{0\le t\le 1}$-martingale because $\left\langle \int_0^\cdot W_s\, dW_s, N\right\rangle_t = \int_0^t W_s d\left\langle W, N\right\rangle_s = 0$. As for the last two cases above, this implies

$$\mathbb{E}\left[\left(\triangle_l^n W\right)^2 \triangle_l^n N \middle| \mathcal{G}_{(k-1)h_n}\right] = 0.$$

This completes the proof of Proposition 4.1.

6.3. **Proof of Proposition 4.2**

We first give a general outline of the proof, deferring some technical details to the end of this section. By Taylor we have for all $k = 1, \ldots, h_n^{-1}$ and $j = 1, \ldots, nh_n - 1$, the existence of random variables ξ_{jk} such that $S_{jk}^2 - \tilde{S}_{jk}^2 = 2\tilde{S}_{jk}(S_{jk} - \tilde{S}_{jk}) + 2(\xi_{jk} - \tilde{S}_{jk})(S_{jk} - \tilde{S}_{jk})$ and $|\xi_{jk} - \tilde{S}_{jk}| \leq |S_{jk} - \tilde{S}_{jk}|$. This yields

$$
\begin{aligned}
n^{\frac{1}{4}}\left(\widehat{\mathbf{IV}}^{or}_{n,t}(Y) - \widehat{\mathbf{IV}}^{or}_{n,t}(\tilde{X} + \epsilon)\right) &= n^{\frac{1}{4}}\left(h_n \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} \sum_{j=1}^{nh_n-1} w_{jk}\left(S_{jk}^2 - \tilde{S}_{jk}^2\right)\right) \\
&= \left(n^{\frac{1}{4}} h_n \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} \sum_{j=1}^{nh_n-1} w_{jk}\left(2\tilde{S}_{jk}\left(S_{jk} - \tilde{S}_{jk}\right)\right)\right) \\
&\quad + \left(n^{\frac{1}{4}} h_n \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} \sum_{j=1}^{nh_n-1} w_{jk}\left(\xi_{jk} - \tilde{S}_{jk}\right)\left(S_{jk} - \tilde{S}_{jk}\right)\right).
\end{aligned}
$$

For the second sum above, which we denote by Z_t^n, we obtain by the Markov inequality and Step 1 below for any $\varepsilon > 0$

$$
\begin{aligned}
\mathbb{P}\left(\sup_{0\leq t\leq 1} |Z_t^n| > \varepsilon\right) &\leq \mathbb{P}\left(\left(n^{\frac{1}{4}} h_n \sum_{k=1}^{h_n^{-1}} \sum_{j=1}^{nh_n-1} w_{jk}\left|S_{jk} - \tilde{S}_{jk}\right|^2\right) > \varepsilon\right) \\
&\leq \varepsilon^{-1} n^{\frac{1}{4}} h_n \sum_{k=1}^{h_n^{-1}} \sum_{j=1}^{nh_n-1} w_{jk}\mathbb{E}\left[\left(S_{jk} - \tilde{S}_{jk}\right)^2\right] \\
&\lesssim \varepsilon^{-1} n^{\frac{1}{4}} h_n \to 0.
\end{aligned}
$$

Let $T_{jk}^n = \sum_{j=1}^{nh_n-1} w_{jk}\left(2\tilde{S}_{jk}\left(S_{jk} - \tilde{S}_{jk}\right)\right)$ and write the first sum above as $M_t^n + R_t^n$ with

$$
\begin{aligned}
M_t^n &= n^{\frac{1}{4}} h_n \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} \left(T_{jk}^n - \mathbb{E}\left[T_{jk}^n \middle| \mathcal{G}_{(k-1)h_n}\right]\right), \\
R_t^n &= n^{\frac{1}{4}} h_n \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} \mathbb{E}\left[T_{jk}^n \middle| \mathcal{G}_{(k-1)h_n}\right].
\end{aligned}
$$

In Step 2 we show that

$$
\sum_{k=1}^{\lfloor th_n^{-1}\rfloor} \mathbb{E}\left[\left(n^{\frac{1}{4}} h_n T_{jk}^n\right)^2\right] \to 0, \quad n \to \infty.
$$

A well known result thereby yields $M_t^n \xrightarrow{ucp} 0$. Finally, observe that

$$
\begin{aligned}
&\mathbb{E}\left[\left(2\tilde{S}_{jk}\left(S_{jk} - \tilde{S}_{jk}\right)\right)\middle| \mathcal{G}_{(k-1)h_n}\right] \\
&= \mathbb{E}\left[2\left(\left\langle n\triangle^n \tilde{X}, \Phi_{jk}\right\rangle_n - [\epsilon, \varphi_{jk}]_n\right)\left\langle n\triangle^n\left(X - \tilde{X}\right), \Phi_{jk}\right\rangle_n \middle| \mathcal{G}_{(k-1)h_n}\right] \\
&= \mathbb{E}\left[2\left\langle n\triangle^n \tilde{X}, \Phi_{jk}\right\rangle_n \left\langle n\triangle^n\left(X - \tilde{X}\right), \Phi_{jk}\right\rangle_n \middle| \mathcal{G}_{(k-1)h_n}\right],
\end{aligned}
$$

i.e. the noise terms vanish, thereby simplifying the following calculations.
Write $\mathbb{E}[(2\tilde{S}_{jk}(S_{jk}-\tilde{S}_{jk}))|\mathcal{G}_{(k-1)h_n}]$ as the sum $D^n_{jk}+V^n_{jk}$ with

$$\begin{aligned} D^n_{jk} &= \mathbb{E}\left[2\left\langle n\triangle^n\tilde{X},\Phi_{jk}\right\rangle_n\left(\sum_{l=1}^n\left(\int_{\frac{l-1}{n}}^{\frac{l}{n}} b_s\,ds\right)\Phi_{jk}\left(\frac{l}{n}\right)\right)\middle|\,\mathcal{G}_{(k-1)h_n}\right],\\ V^n_{jk} &= \mathbb{E}\left[2\left\langle n\triangle^n\tilde{X},\Phi_{jk}\right\rangle_n\left(\sum_{l=1}^n\left(\int_{\frac{l-1}{n}}^{\frac{l}{n}}\left(\sigma_s-\sigma_{(k-1)h_n}\right)\,dW_s\right)\Phi_{jk}\left(\frac{l}{n}\right)\right)\middle|\,\mathcal{G}_{(k-1)h_n}\right]. \end{aligned}$$

In Step 3 we show that $|D_{jk}+V_{jk}|\lesssim h_n^\beta$ for some $\beta>1/2$. This yields immediately

$$\sup_{0\le t\le 1}|R^n_t| \le n^{\frac14}h_n\sum_{k=1}^{h_n^{-1}}\sum_{j=1}^{nh_n-1} w_{jk}\left|D^n_{jk}+V^n_{jk}\right|\lesssim n^{\frac14}h_n^\beta=\mathcal{O}\left(1\right),$$

implying ucp-convergence. We therefore conclude that

$$n^{\frac14}\left(\widehat{\mathbf{IV}}^{or}_{n,t}(Y)-\widehat{\mathbf{IV}}^{or}_{n,t}(\tilde{X}+\epsilon)\right)\xrightarrow{ucp}0,\;\; n\to\infty.$$

The second claim

$$\int_0^t\left(\sigma_s^2-\sigma^2_{\lfloor sh_n^{-1}\rfloor h_n}\right)ds\xrightarrow{ucp}0,\;\; n\to\infty$$

follows from Step 4 and

$$\begin{aligned} &\mathbb{P}\left(\sup_{0\le t\le 1}\left|n^{\frac14}\left(\int_0^t\sigma_s^2\,ds-h_n\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\sigma^2_{(k-1)h_n}\right)\right|>\varepsilon\right)\\ &\quad\le \mathbb{P}\left(\sup_{0\le t\le 1}\left|n^{\frac14}\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\int_{(k-1)h_n}^{kh_n}\left(\sigma_s^2-\sigma^2_{(k-1)h_n}\right)\,ds\right|>\frac{\varepsilon}{2}\right)\\ &\quad\quad+\mathbb{P}\left(\sup_{0\le t\le 1}n^{\frac14}\int_{\lfloor th_n^{-1}\rfloor h_n}^t\sigma_s^2\,ds>\frac{\varepsilon}{2}\right)\\ &\quad\lesssim \varepsilon^{-1}n^{\frac14}\sum_{k=1}^{h_n^{-1}}\left|\mathbb{E}\left[\int_{(k-1)h_n}^{kh_n}\left(\sigma_s^2-\sigma^2_{(k-1)h_n}\right)\,ds\right]\right|+\varepsilon^{-1}\sup_{0\le t\le 1}n^{\frac14}\left(t-\lfloor th_n^{-1}\rfloor h_n\right)\\ &\quad\lesssim \varepsilon^{-1}n^{\frac14}h_n^\gamma \end{aligned}$$

for any $\varepsilon>0$ and some $\gamma>1/2$ what proves Proposition 4.2. We end this section with detailed proofs of Steps 1 – 4.

Step 1: $\mathbb{E}[(S_{jk}-\tilde{S}_{jk})^4]\lesssim h_n^2$
Using the decomposition

$$\begin{aligned} S_{jk}-\tilde{S}_{jk} &= \left\langle n\triangle^n(X-\tilde{X}),\Phi_{jk}\right\rangle_n \qquad (53)\\ &= \sum_{l=1}^n\left(\int_{\frac{l-1}{n}}^{\frac{l}{n}} b_s\,ds\right)\Phi_{jk}\left(\frac{l}{n}\right)+\sum_{l=1}^n\left(\int_{\frac{l-1}{n}}^{\frac{l}{n}}\left(\sigma_s-\sigma_{(k-1)h_n}\right)\,dW_s\right)\Phi_{jk}\left(\frac{l}{n}\right) \end{aligned}$$

into drift and volatility terms we obtain

$$\mathbb{E}\left[\left(S_{jk}-\tilde{S}_{jk}\right)^4\right] \lesssim \mathbb{E}\left[\left(\sum_{l=1}^{n}\left(\int_{\frac{l-1}{n}}^{\frac{l}{n}} b_s\,ds\right)\Phi_{jk}\left(\frac{l}{n}\right)\right)^4\right] + \mathbb{E}\left[\left(\sum_{l=1}^{n}\left(\int_{\frac{l-1}{n}}^{\frac{l}{n}} \left(\sigma_s-\sigma_{(k-1)h_n}\right)\,dW_s\right)\Phi_{jk}\left(\frac{l}{n}\right)\right)^4\right].$$

The first addend is bounded by h_n^2. For the second let $\kappa_l = \int_{\frac{l-1}{n}}^{\frac{l}{n}} \left(\sigma_s-\sigma_{(k-1)h_n}\right)\,dW_s$, such that

$$\mathbb{E}\left[\left(\sum_{l=1}^{n}\left(\int_{\frac{l-1}{n}}^{\frac{l}{n}} \left(\sigma_s-\sigma_{(k-1)h_n}\right)\,dW_s\right)\Phi_{jk}\left(\frac{l}{n}\right)\right)^4\right] = \sum_{l,l',p,p'} \mathbb{E}\left[\kappa_l\kappa_{l'}\kappa_p\kappa_{p'}\right]\Phi_{jk}\left(\frac{l}{n}\right)\Phi_{jk}\left(\frac{l'}{n}\right)\Phi_{jk}\left(\frac{p}{n}\right)\Phi_{jk}\left(\frac{p'}{n}\right).$$

Properties of the conditional expectation show that the only choices for l, l', p, p' with non-trivial results are $l, l' < p = p'$, $l < l' = p = p'$ and $l = l' = p = p'$. In all three cases we can conclude by the Burkholder inequality and (42b) that

$$\left|\mathbb{E}\left[\kappa_l\kappa_{l'}\kappa_p\kappa_{p'}\right]\Phi_{jk}\left(\frac{l}{n}\right)\Phi_{jk}\left(\frac{l'}{n}\right)\Phi_{jk}\left(\frac{p}{n}\right)\Phi_{jk}\left(\frac{p'}{n}\right)\right| \lesssim n^{-4}h_n^{-2}.$$

Observe that in any of the three mentioned cases we find at least two identical integers l, l', p or p'. In all, there are $nh_n \cdot \binom{nh_n-1}{2}\cdot 4!$ possibilities to choose such indices. Hence, we obtain

$$\mathbb{E}\left[\left(\sum_{l=1}^{n}\left(\int_{\frac{l-1}{n}}^{\frac{l}{n}} \left(\sigma_s-\sigma_{(k-1)h_n}\right)\,dW_s\right)\Phi_{jk}\left(\frac{l}{n}\right)\right)^4\right] \lesssim (nh_n)^3\, n^{-4}h_n^{-2} = n^{-1}h_n \lesssim h_n^2$$

and therefore the claim holds.

Step 2: $\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}[(n^{\frac{1}{4}}h_nT_{jk}^n)^2] \to 0, \quad n\to\infty$

Applying the Minkowski and Cauchy-Schwarz inequalities, we obtain

$$\begin{aligned}&\left\|n^{\frac{1}{4}}h_n\sum_{j=1}^{nh_n-1} w_{jk}\left(2\tilde{S}_{jk}\left(S_{jk}-\tilde{S}_{jk}\right)\right)\right\|_{L^2(\mathbb{P})}^2\\ &\le n^{\frac{1}{2}}h_n^2\left(\sum_{j=1}^{nh_n-1}\left\|w_{jk}\left(2\tilde{S}_{jk}\left(S_{jk}-\tilde{S}_{jk}\right)\right)\right\|_{L^2(\mathbb{P})}\right)^2\\ &\le n^{\frac{1}{2}}h_n^2\left(\sum_{j=1}^{nh_n-1} w_{jk}\left(\mathbb{E}\left[\tilde{S}_{jk}^4\right]\right)^{\frac{1}{4}}\left(\mathbb{E}\left[\left(S_{jk}-\tilde{S}_{jk}\right)^4\right]\right)^{\frac{1}{4}}\right)^2.\end{aligned}$$

By Step 1 we already know that $\mathbb{E}\big[\big(S_{jk}-\tilde{S}_{jk}\big)^4\big] \lesssim h_n^2$. Because σ is bounded, we obtain by (50) the bound

$$\mathbb{E}\left[\tilde{S}_{jk}^4\right] \le \mathbb{E}^{\frac{1}{2}}\left[\mathbb{E}\left[\tilde{S}_{jk}^8\middle|\mathcal{G}_{(k-1)h_n}\right]\right] \lesssim \mathbb{E}^{\frac{1}{2}}\left[\left(\sigma_{(k-1)h_n}^2+\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{jk}\right]_n\right)^4\right]$$

$$\lesssim \left(1+\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{jk}\right]_n\right)^2 \leq \left(1+\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{jk}\right]_n\right)^4.$$

Together with (45) it follows that

$$\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}\left[\left(n^{\frac{1}{4}}h_nT_{jk}^n\right)^2\right] \lesssim n^{\frac{1}{2}}h_n^3\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\left(\sum_{j=1}^{nh_n-1}w_{jk}\left(1+\frac{\eta^2}{n}\left[\varphi_{jk},\varphi_{jk}\right]_n\right)\right)^2$$
$$\lesssim n^{\frac{1}{2}}h_n^2\cdot n^2h_n^4=\mathcal{O}(1).$$

Step 3: $|D_{jk}+V_{jk}|\lesssim h_n^{\beta}$ for some $\beta>1/2$

Expanding the sums in V_{jk} and Itô isometry yield

$$\begin{aligned}|V_{jk}| &= \left|\sum_{l,m=1}^{n}\left(\mathbb{E}\left[\triangle_l^n\tilde{X}\left(\int_{\frac{m-1}{n}}^{\frac{m}{n}}\left(\sigma_s-\sigma_{(k-1)h_n}\right)\,dW_s\right)\middle|\mathcal{G}_{(k-1)h_n}\right]\Phi_{jk}\left(\frac{l}{n}\right)\Phi_{jk}\left(\frac{m}{n}\right)\right)\right|\\ &= \left|\sum_{l=1}^{n}\mathbb{E}\left[\int_{\frac{l-1}{n}}^{\frac{l}{n}}\left(\sigma_{(k-1)h_n}\left(\sigma_s-\sigma_{(k-1)h_n}\right)\right)\,ds\right]\Phi_{jk}^2\left(\frac{l}{n}\right)\right|.\end{aligned}$$

Observe that we have in the semimartingale case for σ that

$$\begin{aligned}\left|\mathbb{E}\left[\left(\sigma_{(k-1)h_n}\left(\sigma_s-\sigma_{(k-1)h_n}\right)\right)\,ds\right]\right| &= \left|\mathbb{E}\left[\sigma_{(k-1)h_n}\mathbb{E}\left[\sigma_s-\sigma_{(k-1)h_n}\middle|\mathcal{G}_{(k-1)h_n}\right]\right]\,ds\right|\\ &= \left|\mathbb{E}\left[\sigma_{(k-1)h_n}\int_{(k-1)h_n}^{s}\tilde{b}_r\,dr\right]\right|\lesssim h_n,\end{aligned} \tag{54}$$

because σ and $\tilde{b}$ are bounded. In the Hölder case, on the other hand, it holds similarly that

$$\left|\mathbb{E}\left[\left(\sigma_{(k-1)h_n}\left(\sigma_s-\sigma_{(k-1)h_n}\right)\right)\,ds\right]\right|\lesssim h_n^{\alpha}.$$

Hence, we can conclude in any case by Fubini that

$$\left|\mathbb{E}\left[\left(\sigma_{(k-1)h_n}\left(\sigma_s-\sigma_{(k-1)h_n}\right)\right)\,ds\right]\right|\lesssim h_n^{\beta'}$$

for some $\beta'>1/2$ such that $|V_{jk}|\lesssim h_n^{\beta'}$, as well. With respect to D_{jk}^n, we need an additional approximation. By Assumption (H-1) and the boundedness of $\mathbb{E}[|\langle n\triangle^n\tilde{X},\Phi_{jk}\rangle_n|]$, see (48):

$$\begin{aligned}&\left|\mathbb{E}\left[\left\langle n\triangle^n\tilde{X},\Phi_{jk}\right\rangle_n\int_{\frac{l-1}{n}}^{\frac{l}{n}}b_s\,ds\middle|\mathcal{G}_{(k-1)h_n}\right]\right|\\ &\lesssim\left|\mathbb{E}\left[\left\langle n\triangle^n\tilde{X},\Phi_{jk}\right\rangle_n\int_{\frac{l-1}{n}}^{\frac{l}{n}}\left(b_s-b_{(k-1)h_n}\right)\,ds\middle|\mathcal{G}_{(k-1)h_n}\right]\right|\\ &+\left|\frac{b_{(k-1)h_n}}{n}\mathbb{E}\left[\left\langle n\triangle^n\tilde{X},\Phi_{jk}\right\rangle_n\middle|\mathcal{G}_{(k-1)h_n}\right]\right|\lesssim h_n^{\nu}n^{-1}.\end{aligned}$$

Using this bound, we find with (b_s) being ν-Hölder that

$$\left|D_{jk}^n\right| \leq \sum_{l=1}^{n}\left|\mathbb{E}\left[2\left\langle n\triangle^n\tilde{X},\Phi_{jk}\right\rangle_n\int_{\frac{l-1}{n}}^{\frac{l}{n}}\left(b_s-b_{(k-1)h_n}\right)\,ds\middle|\mathcal{G}_{(k-1)h_n}\right]\right|\left|\Phi_{jk}\left(\frac{l}{n}\right)\right|$$

$$\lesssim \quad h_n^{\nu}\left(\frac{1}{n}\sum_{l=1}^{n}\left|\Phi_{jk}\left(\frac{l}{n}\right)\right|\right)\lesssim h_n^{\nu+\frac{1}{2}}\,.$$

We obtain the claim with $\beta = \min\left\{\nu + \frac{1}{2}, \beta'\right\}$. This is the only time we need the Hölder smoothness of the drift in Assumption (H-1).
Step 4: $\left|\mathbb{E}\left[\int_{(k-1)h_n}^{kh_n}\left(\sigma_s^2-\sigma_{(k-1)h_n}^2\right)\,ds\right]\right| \lesssim h_n^{1+\gamma}$ for some $\gamma > 1/2$.
For each block $k = 1,\ldots,h_n^{-1}$, and all $0 \le s \le 1$, we can find random variables $\xi_{k,s}$ with

$$\sigma_s^2-\sigma_{(k-1)h_n}^2 = 2\sigma_{(k-1)h}\left(\sigma_s-\sigma_{(k-1)h_n}\right)+2\left(\xi_{k,s}-\sigma_{(k-1)h_n}\right)\left(\sigma_s-\sigma_{(k-1)h_n}\right)$$

and $\left|\xi_{k,s}-\sigma_{(k-1)h_n}\right| \le \left|\sigma_s-\sigma_{(k-1)h_n}\right|$. Note that this implies in the semimartingale case for σ by (54) that

$$\begin{aligned}\left|\mathbb{E}\left[\sigma_s^2-\sigma_{(k-1)h_n}^2\right]\right| &\lesssim \left|\mathbb{E}\left[\sigma_{(k-1)h_n}\mathbb{E}\left[\left(\sigma_s-\sigma_{(k-1)h_n}\right)\middle|\,\mathcal{G}_{(k-1)h_n}\right]\right]\right|+\mathbb{E}\left[\left(\sigma_s-\sigma_{(k-1)h_n}\right)^2\right]\\ &\lesssim h_n+\mathbb{E}\left[\left(\sigma_s-\sigma_{(k-1)h_n}\right)^2\right].\end{aligned}$$

Hence, we see by Fubini and (42b) that

$$\left|\mathbb{E}\left[\int_{(k-1)h_n}^{kh_n}\left(\sigma_s^2-\sigma_{(k-1)h_n}^2\right)\,ds\right]\right| \quad\lesssim\quad h_n^2+\mathbb{E}\left[\int_{(k-1)h_n}^{kh_n}\left(\sigma_s-\sigma_{(k-1)h_n}\right)^2\,ds\right]\lesssim h_n^2.$$

In the Hölder case we obtain directly by the boundedness of σ:

$$\mathbb{E}\left[\int_{(k-1)h_n}^{kh_n}\left(\sigma_s^2-\sigma_{(k-1)h_n}^2\right)\,ds\right]\lesssim\quad \mathbb{E}\left[\int_{(k-1)h_n}^{kh_n}\left|\sigma_s-\sigma_{(k-1)h_n}\right|\,ds\right]\lesssim h_n^{1+\alpha}.$$

Because $\alpha > 1/2$, we obtain the claim with $\gamma = \min(\alpha, 1)$.

6.4. Proofs of Theorem 2 and Theorem 3 for oracle estimation

We decompose X similarly as in the proof of Theorem 1:

$$X_t = X_0+\bar{B}_t+\tilde{B}_t+\bar{C}_t+\tilde{C}_t\,, \tag{55}$$

where we denote

$$\bar{B}_t=\int_0^t b_{\lfloor sh_n^{-1}\rfloor h_n}\,ds\;,\;\tilde{B}_t=\int_0^t(b_s-b_{\lfloor sh_n^{-1}\rfloor h_n})\,ds\,, \tag{56a}$$

$$\bar{C}_t=\int_0^t \sigma_{\lfloor sh_n^{-1}\rfloor h_n}\,dW_s\;,\;\tilde{C}_t=\int_0^t(\sigma_s-\sigma_{\lfloor sh_n^{-1}\rfloor h_n})\,dW_s\,. \tag{56b}$$

In order to establish a functional CLT, we decompose the estimation errors of (29a) (and likewise (25a)) in the following way:

$$\mathbf{LMM}_{n,t}^{or}(Y)-\mathrm{vec}\Big(\int_0^t\Sigma_s\,ds\Big)=\mathbf{LMM}_{n,t}^{or}(\bar{C}+\epsilon)-\mathrm{vec}\left(\int_0^t\Sigma_{\lfloor sh_n^{-1}\rfloor h_n}\,ds\right) \tag{57a}$$

$$+\mathbf{LMM}^{or}_{n,t}(Y)-\mathbf{LMM}^{or}_{n,t}(\bar{C}+\epsilon)-\operatorname{vec}\left(\int_0^t\left(\Sigma_s-\Sigma_{\lfloor sh_n^{-1}\rfloor h_n}\right)ds\right). \tag{57b}$$

One crucial step to cope with multi-dimensional non-synchronous data is Lemma 4.4 which is proved next. Below, we give a concise proof of the functional CLTs for the estimators (29a) and (25a), where after restricting to a synchronous reference scheme many steps develop as direct extensions of the one-dimensional case. The stable CLTs for the *leading terms*, namely the right-hand side of (57a) and the analogue for estimator (25a), are established in paragraph 6.4.2. The *remainder terms* (57b) and their analogues are handled in paragraph 6.4.3.

6.4.1. Proof of Lemma 4.4

Consider for $(l,m)\in\{1,\ldots,d\}^2$, observation times $t_i^{(l)}=F_l^{-1}(i/n_l)$ and $t_i^{(m)}=F_m^{-1}(i/n_m)$ and suppose without loss of generality $n_m\le n_l$. Define a next-tick interpolation function by

$$t_+^{(l)}(s)=\min\left(t_v^{(l)},0\le v\le n_l|t_v^{(l)}\ge s\right),l=1,\ldots,d,$$

and analogously a previous-tick interpolation function by

$$t_-^{(l)}(s)=\max\left(t_v^{(l)},0\le v\le n_l|t_v^{(l)}\le s\right),l=1,\ldots,d.$$

We decompose increments of $X^{(l)}$ between adjacent observation times $t_{v-1}^{(l)},t_v^{(l)},v=1,\ldots,n_l$, in the sum of increments of $X^{(l)}$ over all time intervals $[t_{i-1}^{(m)},t_i^{(m)}]$ contained in $[t_{v-1}^{(l)},t_v^{(l)}]$ and the remaining time intervals at the left $\left[t_{v-1}^{(l)},t_+^{(m)}(t_{v-1}^{(l)})\right]$ and the right border $\left[t_-^{(m)}(t_v^{(l)}),t_v^{(l)}\right]$:

$$X^{(l)}_{t_v^{(l)}}-X^{(l)}_{t_{v-1}^{(l)}}=\left(X^{(l)}_{t_v^{(l)}}-X^{(l)}_{t_-^{(m)}(t_v^{(l)})}\right)+\sum_{\Delta_it^{(m)}\subset\Delta_vt^{(l)}}\left(X^{(l)}_{t_i^{(m)}}-X^{(l)}_{t_{i-1}^{(m)}}\right)+\left(X^{(l)}_{t_+^{(m)}(t_{v-1}^{(l)})}-X^{(l)}_{t_{v-1}^{(l)}}\right).$$

If there is only one observation of $X^{(m)}$ in $[t_{v-1}^{(l)},t_v^{(l)}]$, set $\sum_{\Delta_it^{(m)}\subset\Delta_vt^{(l)}}\left(X^{(l)}_{t_i^{(m)}}-X^{(l)}_{t_{i-1}^{(m)}}\right)=0$.

If there is no observation of $X^{(m)}$ in $[t_{v-1}^{(l)},t_v^{(l)}]$ we take the union of a set of intervals $\bigcup_{v\in V}[t_{v-1}^{(l)},t_v^{(l)}]$ which contains at least one observation time of $X^{(m)}$. We use an expansion of $\left(\Phi_{jk}(t)-\Phi_{jk}(s)\right)$. By virtue of $\sin(t)-\sin(s)=2\cos((t+s)/2)\sin((t-s)/2)$ and the sine expansion, we obtain for $s,t\in[kh_n,(k+1)h_n)$:

$$\left(\Phi_{jk}(t)-\Phi_{jk}(s)\right)\asymp\sqrt{2}h_n^{-3/2}j\pi\cos\left(j\pi h_n^{-1}(\tfrac{t+s}{2}-kh_n)\right)(t-s). \tag{58}$$

In particular, for $(t-s)=\mathcal{O}(n^{-1})$ we have that $\left(\Phi_{jk}(t)-\Phi_{jk}(s)\right)=\mathcal{O}\left(\varphi_{jk}(\tfrac{t+s}{2})n^{-1}\right)$.
With $u_v^{(m)}=(1/2)(t_+^{(m)}(t_v^{(l)})-t_-^{(m)}(t_v^{(l)}))$ and $\tilde{u}_v^{(m)}=(1/2)(t_+^{(m)}(t_{v-1}^{(l)})-t_-^{(m)}(t_{v-1}^{(l)}))$, we infer

$$\begin{aligned}&\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}h_n\sum_{j\ge1}w_{jk}^{l,m}\sum_{i=1}^{n_l}\Delta_iX^{(l)}\Phi_{jk}(\bar{t}_i^{(l)})\sum_{v=1}^{n_m}\Delta_vX^{(m)}\Phi_{jk}(\bar{t}_v^{(m)})\\
&=\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}h_n\sum_{j\ge1}w_{jk}^{l,m}\sum_{i=1}^{n_l}\left(X^{(l)}_{t_i^{(l)}}-X^{(l)}_{t_{i-1}^{(l)}}\right)\Phi_{jk}(\bar{t}_i^{(l)})\sum_{v=1}^{n_l}\left(X^{(m)}_{t_v^{(l)}}-X^{(m)}_{t_{v-1}^{(l)}}\right)\Phi_{jk}(\bar{t}_v^{(l)})\\
&+\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}h_n\sum_{j\ge1}w_{jk}^{l,m}\sum_{v=1}^{n_l}\left(X^{(l)}_{t_v^{(l)}}-X^{(l)}_{t_{v-1}^{(l)}}\right)\Phi_{jk}(\bar{t}_v^{(l)})\end{aligned}$$

$$\times\Big(\sum_{\Delta_i t^{(m)}\subset\Delta_v t^{(l)}}\Big(X^{(m)}_{t_i^{(m)}}-X^{(m)}_{t_{i-1}^{(m)}}\Big)(\Phi_{jk}(\bar t_i^{(m)})-\Phi_{jk}(\bar t_v^{(l)}))+$$
$$\Big(X^{(m)}_{t_+^{(m)}(t_{v-1}^{(l)})}-X^{(m)}_{t_{v-1}^{(l)}}\Big)\big(\Phi_{jk}(\tilde u_v^{(m)})-\Phi_{jk}(\bar t_v^{(l)})\big)+\Big(X^{(m)}_{t_v^{(l)}}-X^{(m)}_{t_-^{(m)}(t_v^{(l)})}\Big)\big(\Phi_{jk}(u_v^{(m)})-\Phi_{jk}(\bar t_v^{(l)}))\Big).$$

Applying the bound (58), we find that the order of the last summand is $\sum_k h_n\sum_j w_{jk}^{l,m} j/(nh_n)$ and since for all weights the bound (44) holds we conclude that the approximation error is uniformly of order $\mathcal{O}_{\mathbb{P}}(h_n)=\scriptstyle\mathcal{O}_{\mathbb{P}}(n^{-1/4})$.

6.4.2. **Leading terms**

This paragraph develops the asymptotics for the right-hand side of (57a) and the sum of the increments in (38). Observe that

$$\sum_{i=1}^{n_l-1}\varphi_{jk}^2\big(t_i^{(l)}\big)\Big(\tfrac{t_{i+1}^{(l)}-t_{i-1}^{(l)}}{2}\Big)^2\asymp\sum_{i=1}^{n_l-1}\varphi_{jk}^2\big(t_i^{(l)}\big)\tfrac{t_{i+1}^{(l)}-t_{i-1}^{(l)}}{2}\tfrac{H_l^{kh_n}}{\eta_l^2 n_l}\asymp\Big(\int_0^1\varphi_{jk}^2(t)\,dt\Big)\tfrac{H_l^{kh_n}}{\eta_l^2 n_l}\,. \tag{59}$$

The left approximation uses $(t_{i+1}^{(l)}-t_i^{(l)})/2=(H_l^{kh_n}+\mathcal{O}(h_n^\alpha))/(\eta_l^2 n_l)$ as in (43) with $\alpha>1/2$ by Assumption (Obs-d). Writing the integral on the right-hand side as sum over the subintervals and using mean value theorem, the differences when passing to the arguments $(t_i^{(l)})_i$ induce approximation errors of order $jh_n^{-1}n^{-1}$. Thus, the total approximation errors are of order $(h_n^\alpha+j(nh_n)^{-1})j^2(nh_n^2)^{-1}$. We focus on the oracle versions of (29a) and (25a) with their deterministic optimal weights. The proof follows the same methodology as the proof of Proposition 4.1 after restricting to a synchronous reference observation scheme. We concisely go through the details for cross terms and the proof for the bivariate spectral covolatility estimator.

We apply Theorem 3-1 of Jacod (1997) (or equivalently Theorem 2.6 in Podolskij and Vetter (2010)) again. For the spectral estimator (25a), consider

$$\zeta_k^n=n^{1/4}\,h_n\Big(\sum_{j\ge1}w_{jk}^{p,q}\zeta_{jk}^{(pq)}-\Sigma^{(pq)}_{(k-1)h_n}\Big)\,, \tag{60}$$

with the random variables

$$\zeta_{jk}^{(pq)}=\Bigg(\sum_{i=1}^{n_p}\Delta_i^n\bar C^{(p)}\Phi_{jk}\big(\bar t_i^{(p)}\big)-\sum_{i=1}^{n_p-1}\epsilon_i^{(p)}\varphi_{jk}\big(t_i^{(p)}\big)\frac{t_{i+1}^{(p)}-t_{i-1}^{(p)}}{2} \tag{61}$$
$$\times\sum_{v=1}^{n_q}\Delta_v^n\bar C^{(q)}\Phi_{jk}\big(\bar t_v^{(q)}\big)-\sum_{v=1}^{n_q-1}\epsilon_v^{(q)}\varphi_{jk}\big(t_v^{(q)}\big)\frac{t_{v+1}^{(q)}-t_{v-1}^{(q)}}{2}-\pi^2j^2h_n^{-2}\delta_{p,q}\hat H_p^{kh_n}n_p^{-1}\Bigg)\,.$$

The accordance with (38) follows from a generalization of the summation by parts identity(41c):

$$S_{jk}^{(p)}\asymp_p-\sum_{v=1}^{n_p-1}Y_v^{(p)}\Big(\Phi_{jk}\big(\bar t_{v+1}^{(p)}\big)-\Phi_{jk}\big(\bar t_v^{(p)}\big)\Big)$$
$$\asymp_p-\sum_{v=1}^{n_p-1}Y_v^{(p)}\varphi_{jk}(t_v^{(p)})\frac{t_{v+1}^{(p)}-t_{v-1}^{(p)}}{2}\,,$$

where the first remainder, which is only due to end-effects when $t_0^{(p)}\neq0$ or $t_{n_p}^{(p)}\neq1$, and the second remainder by application of mean value theorem and passing to arguments $t_v^{(p)}$ are asymptotically

negligible. This is obvious for the first remainder and the second is treated analogously as for the approximation between discrete and continuous-time norm of the (φ_{jk}) in the following.
By Lemma 4.4 we may without loss of generality work under synchronous observations $t_i, i = 0, \dots, n$, when considering the signal part X. Set $\bar{t}_i = (t_{i+1} - t_i)/2$. We shall write in the sequel terms of the signal part as coming from observations on a synchronous grid (t_i), while keeping to the actual grids for the noise terms. For the expectation we have

$$\begin{aligned}\mathbb{E}\left[\zeta_{jk}^{(pq)}\right] &= \sum_{i=1}^{n} \Phi_{jk}^2(\bar{t}_i)\mathbb{E}\big[\Delta_i^n \bar{C}^{(p)} \Delta_i^n \bar{C}^{(q)}\big] \\ &+ \sum_{i,v=1}^{(n_p \vee n_q)-1} \mathbb{E}\left[\epsilon_i^{(p)}\epsilon_i^{(q)}\right]\varphi_{jk}\big(t_i^{(p)}\big)\Big(\tfrac{t_{i+1}^{(p)}-t_{i-1}^{(p)}}{2}\Big)\varphi_{jk}\big(t_v^{(q)}\big)\Big(\tfrac{t_{v+1}^{(q)}-t_{v-1}^{(q)}}{2}\Big) - \pi^2 j^2 h_n^{-2}\tfrac{\delta_{p,q}\mathbb{E}\left[\hat{H}_p^{kh_n}\right]}{n_p} \\ &= \sum_{i=1}^{n} \Phi_{jk}^2(\bar{t}_i)(t_{i+1}-t_i)\Sigma_{(k-1)h_n}^{(pq)} + \delta_{p,q}\left(\eta_p^2 \sum_{i=1}^{n_p}\varphi_{jk}^2\big(t_i^{(p)}\big)\Big(\tfrac{t_{i+1}^{(p)}-t_{i-1}^{(p)}}{2}\Big)^2 - \pi^2 j^2 h_n^{-2}\tfrac{H_p^{kh_n}}{n_p}\right) \\ &= \Sigma_{(k-1)h_n}^{(pq)} + R_{n,k}\end{aligned}$$

by Itô isometry. The remainders due to the approximation (59) satisfy with (44) uniformly

$$R_{n,k} \lesssim \sum_{j=1}^{\lfloor\sqrt{n}h_n\rfloor} j^2 n^{-1}h_n^{-2}\big(h_n^{\alpha} + jn^{-1}h_n^{-1}\big) + \sum_{\lceil\sqrt{n}h_n\rceil}^{nh_n-1}\big(j^{-1}h_n + j^{-2}h_n^2 n h_n^{\alpha}\big) = \mathcal{O}\big(n^{-1/4}\big)\,.$$

Since $\sum_{j\ge1} w_{jk}^{p,q} = 1$, asymptotic unbiasedness is ensured:

$$\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}\left[\zeta_k^n\middle|\mathcal{G}_{(k-1)h_n}\right] = \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} n^{1/4}h_n\left(\sum_{j\ge1} w_{jk}^{p,q}\mathbb{E}[\zeta_{jk}^{(pq)}] - \Sigma_{(k-1)h_n}^{(pq)}\right) \xrightarrow{ucp} 0\,.$$

We now determine the asymptotic variance expression in (11):

$$\begin{aligned}\mathbb{V}\mathrm{ar}\big(\zeta_{jk}^{(pq)}\big) &= \Big(\sum_{i=1}^{n}\Phi_{jk}^2(\bar{t}_i)(t_{i+1}-t_i)\Big)^2\Big(\big(\Sigma_{(k-1)h_n}^{(pq)}\big)^2 + \Sigma_{(k-1)h_n}^{(pp)}\Sigma_{(k-1)h_n}^{(qq)}\Big) \\ &\quad + \eta_p^2\eta_q^2\Big(\sum_{i=1}^{n_p-1}\varphi_{jk}^2\big(t_i^{(p)}\big)\Big(\tfrac{t_{i+1}^{(p)}-t_{i-1}^{(p)}}{2}\Big)^2\Big)\Big(\sum_{i=1}^{n_q-1}\varphi_{jk}^2\big(t_i^{(q)}\big)\Big(\tfrac{t_{i+1}^{(q)}-t_{i-1}^{(q)}}{2}\Big)^2\Big) \\ &\quad + \left(\sum_{i=1}^{n}\Phi_{jk}^2(\bar{t}_i)(t_{i+1}-t_i)\Big(\eta_p^2\Sigma_{(k-1)h_n}^{(qq)}\sum_{i=1}^{n_p-1}\varphi_{jk}^2\big(t_i^{(p)}\big)\Big(\tfrac{t_{i+1}^{(p)}-t_{i-1}^{(p)}}{2}\Big)^2\right. \\ &\quad \left. + \eta_m^2\Sigma_{(k-1)h_n}^{(pp)}\sum_{i=1}^{n_q-1}\varphi_{jk}^2\big(t_i^{(q)}\big)\Big(\tfrac{t_{i+1}^{(q)}-t_{i-1}^{(q)}}{2}\Big)^2\Big)\right) \\ &\asymp \big(\Sigma_{(k-1)h_n}^{(pq)}\big)^2 + \Sigma_{(k-1)h_n}^{(pp)}\Sigma_{(k-1)h_n}^{(qq)} + \pi^2 j^2 h_n^{-2}\big(H_p^{kh_n}n_p^{-1}\Sigma_{(k-1)h_n}^{(qq)} + H_q^{kh_n}n_q^{-1}\Sigma_{(k-1)h_n}^{(pp)}\big) \\ &\quad + \pi^4 j^4 h_n^{-4} n_p^{-1} n_q^{-1} H_p^{kh_n} H_q^{kh_n},\end{aligned}$$

where the remainder is negligible by the same bounds as for the bias above. The sum of conditional variances with $w_{jk}^{p,q} = I_k^{-1}I_{jk}$, $I_k = \sum_{j\ge1} I_{jk}$, thus yields

$$\sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}\left[(\zeta_k^n)^2\,\middle|\mathcal{G}_{(k-1)h_n}\right] + \mathcal{O}(1) = \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n^2 n^{1/2}\sum_{j\ge1}\big(w_{jk}^{(pq)}\big)^2\mathbb{V}\mathrm{ar}\big(\zeta_{jk}^{(pq)}\big)$$

$$= \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n^2 n^{1/2} \sum_{j\geq 1} I_{jk} I_k^{-2} = \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n^2 n^{1/2} I_k^{-1}.$$

As $h_n\sqrt{n}\to\infty$, we obtain an asymptotic expression as the solution of an integral

$$\sum_{k=1}^{\lfloor th_n^{-1}\rfloor} \mathbb{E}\left[(\zeta_k^n)^2\,|\mathcal{G}_{(k-1)h_n}\right] = \sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n(\sqrt{n}h_n) I_k^{-1} \to \int_0^t \left(\int_0^\infty (f(\Sigma,\mathcal{H}(t),c_p,c_q;z))^{-1}dz\right)^{-1} ds$$

with a continuous limit function f which is the same as in Bibinger and Reiß (2013). Computing the solution of the integral using the explicit form of I_k and f yields the variance $\int_0^t \left(v_s^{(p,q)}\right)^2 ds$ with

$$\left(v_s^{(p,q)}\right)^2 = 2\left((F_p^{-1})'(s)(F_q^{-1})'(s)c_p^{-1}c_q^{-1}(A_s^2-B_s)B_s\right)^{1/2}$$
$$\times\left(\sqrt{A_s+\sqrt{A_s^2-B_s}} - \operatorname{sgn}(A_s^2-B_s)\sqrt{A_s-\sqrt{A_s^2-B_s}}\right),$$

and the terms

$$A_s = \Sigma_s^{(pp)}\frac{(F_q^{-1})'(s)c_p}{(F_p^{-1})'(s)c_q} + \Sigma_s^{(qq)}\frac{(F_p^{-1})'(s)c_q}{(F_q^{-1})'(s)c_p},$$
$$B_s = 4\left(\Sigma_s^{(pp)}\Sigma_s^{(qq)} + \left(\Sigma_s^{(pq)}\right)^2\right).$$

The detailed computation is carried out in Bibinger and Reiß (2013) and we omit it here. sgn denotes the sign taking values in $\{-1,+1\}$ and ensuring that the value of $\left(v_s^{(p,q)}\right)^2$ is always a positive real number. Contrarily to the one-dimensional case, in the cross term there is no effect of non-Gaussian noise on the variance because fourth noise moments do not occur and component-wise independence. The Lyapunov criterion follows from

$$\mathbb{E}\left[\left(\zeta_{jk}^{(pq)}\right)^4|\mathcal{G}_{(k-1)h_n}\right] \asymp 3\sum_{j\geq 1}\left(w_{jk}^{p,q}\right)^4 I_{jk}^{-2} \asymp 3\, I_k^{-4}\sum_{j\geq 1} I_{jk}^2 = \mathcal{O}(1)$$
$$\Rightarrow \sum_{k=1}^{\lfloor th_n^{-1}\rfloor}\mathbb{E}\left[\left(\zeta_k^n\right)^4|\mathcal{G}_{(k-1)h_n}\right] = \mathcal{O}\Big(n\sum_{k=1}^{\lfloor th_n^{-1}\rfloor} h_n^4\Big) = \mathcal{O}\left(n^{-1/4}\right).$$

By Cauchy-Schwarz and Burkholder-Davis-Gundy inequalities, we deduce

$$\mathbb{E}\left[h_n\sum_{j\geq 1} w_{jk}^{p,q}\sum_{i=1}^n \Delta_i^n\bar{C}^{(p)}\Delta_i^n\bar{C}^{(q)}\Phi_{jk}^2(\bar{t}_i)\sum_{i=1}^n\Delta_i^n W^{(p)}\right]$$
$$= h_n\sum_{j\geq 1} w_{jk}^{p,q}\sum_{i=1}^n\mathbb{E}\left[\Delta_i^n\bar{C}^{(p)}\Delta_i^n\bar{C}^{(q)}\Delta_i^n W^{(p)}\right]\Phi_{jk}^2(\bar{t}_i)$$
$$\leq h_n\sum_{j\geq 1} w_{jk}^{p,q}\sum_{i=1}^n (t_i-t_{i-1})^{3/2}\Phi_{jk}^2(\bar{t}_i) = \mathcal{O}\big(n^{-1/4}\big).$$

By the analogous estimate with $\Delta_i^n W^{(q)}$ the stability conditions are valid. This proves stable convergence of the leading term to the limit given in Theorem 2.

The heart of the proof of Theorem 3 is the asymptotic theory for the leading term (57a), namely the analysis of the asymptotic variance-covariance structure. This is carried out in detail in Bibinger et al.

(2013) for the idealized locally parametric experiment using bin-wise orthogonal transformation to a diagonal covariance structure. The only difference between our main term and the setup considered in Bibinger et al. (2013) is the Gaussianity of the noise component. Yet, in the deduction of the variance this only effects the terms with fourth noise moments where $\mathbb{E}[\epsilon_i^4] \neq 3\mathbb{E}[\epsilon_i^2]$ in general. Above we explicitly proved that the resulting remainder converges to zero for the one-dimensional estimator and this directly extends to the diagonal elements here. An intuitive heuristic reason why this holds is that the smoothed statistics are asymptotically still close to a normal distribution, though the normality which could have been used in Bibinger et al. (2013) does not hold here for fixed n in general. Based on the expressions of variances for cross products and squared spectral statistics above, coinciding their counterparts in the normal noise model when separating the remainder induced for the squares, we can pursue the asymptotics along the same lines as the proof of Corollary 4.3 in Bibinger et al. (2013). At this stage, we restrict to shed light on the connection between the expressions in (13) and the asymptotic variance-covariance matrix. Observe that $(A \otimes B)^\top = A^\top \otimes B^\top$ for matrices A, B, $\mathcal{Z}\mathcal{Z} = 2\mathcal{Z}$ and that $(A \otimes B)(C \otimes D) = (AC \otimes BD)$ for matrices A, B, C, D, such that

$$
\begin{aligned}
&\left(\Sigma_s^{\frac{1}{2}} \otimes \left(\Sigma_s^{\mathcal{H}}\right)^{\frac{1}{4}}\right)\mathcal{Z}\left(\left(\Sigma_s^{\frac{1}{2}} \otimes \left(\Sigma_s^{\mathcal{H}}\right)^{\frac{1}{4}}\right)\mathcal{Z}\right)^\top \\
&\quad = \left(\Sigma_s^{\frac{1}{2}} \otimes \left(\Sigma_s^{\mathcal{H}}\right)^{\frac{1}{4}}\right)2\mathcal{Z}\left(\Sigma_s^{\frac{1}{2}} \otimes \left(\Sigma_s^{\mathcal{H}}\right)^{\frac{1}{4}}\right)^\top \\
&\quad = 2\left(\Sigma_s \otimes \left(\Sigma_s^{\mathcal{H}}\right)^{\frac{1}{2}}\right)\mathcal{Z}\,,
\end{aligned}
$$

since $\mathcal{Z}$ commutes with $\left(\Sigma_s^{\frac{1}{2}} \otimes \left(\Sigma_s^{\mathcal{H}}\right)^{\frac{1}{4}}\right)$. Therefore, the expression in (13) is natural for the matrix square root of the asymptotic variance-covariance, where we use two independent terms because of non-commutativity of matrix multiplication. Conditions (J1) and (J3) and the stability conditions (J4) and (J5) can be analogously showed by element-wise adopting the results for squared and cross products of spectral statistics from above. Since any component of the estimator is a weighted sum of the entries of $S_{jk}S_{jk}^\top$, bias-corrected on the diagonal, the convergences to zero in probability follow likewise.

6.4.3. **Remainder terms**

After applying triangular inequality to (57b), it suffices to prove that

$$
n^{1/4}\|\,\mathbf{LMM}_{n,t}^{or}(Y) - \mathbf{LMM}_{n,t}^{or}(\bar{C} + \epsilon)\| \xrightarrow{ucp} 0\,, \tag{62}
$$

$$
n^{1/4}\left\|\int_0^t \operatorname{vec}\left(\Sigma_s - \Sigma_{\lfloor sh_n^{-1}\rfloor h_n}\right)ds\right\| \xrightarrow{ucp} 0\,. \tag{63}
$$

For $A, B \in \mathbb{R}^d$, we use in the following several times the elementary bound:

$$
\left\|AA^\top - BB^\top\right\| = \left\|B(A^\top - B^\top) + (A - B)A^\top\right\| \leq \left(\|A\| + \|B\|\right)\|A - B\|\,. \tag{64}
$$

Define analogously as above $\tilde{S}_{jk} = \left(\sum_{i=1}^{n_p} \Delta_i^n \bar{C}^{(p)}\Phi_{jk}\left(\bar{t}_i^{(p)}\right)\right)_{1\leq p\leq d}$, the spectral statistics in the locally constant volatility experiment. Then we can bound uniformly for all t:

$$
\|\,\mathbf{LMM}_{n,t}^{or}(Y) - \mathbf{LMM}_{n,t}^{or}(\bar{C} + \epsilon)\|
$$

$$
\begin{aligned}
&\leq \sum_{k=1}^{h_n^{-1}} h_n \left\| \sum_{j=1}^{nh_n-1} W_{jk} \operatorname{vec}\left(S_{jk}S_{jk}^{\top} - \tilde{S}_{jk}\tilde{S}_{jk}^{\top}\right) \right\| \\
&\leq \sum_{k=1}^{h_n^{-1}} h_n \sum_{j=1}^{nh_n-1} \|W_{jk}\| \big(\|S_{jk}\| + \|\tilde{S}_{jk}\|\big) \|S_{jk} - \tilde{S}_{jk}\| \\
&\lesssim \sum_{k=1}^{h_n^{-1}} h_n \sum_{j=1}^{nh_n-1} \Big(1 + \frac{j^2}{nh_n^2}\Big)^{-2} \|S_{jk} - \tilde{S}_{jk}\| = \mathcal{O}_{\mathbb{P}}\big(h_n\big) = \scriptstyle{\mathcal{O}}_{\mathbb{P}}(n^{-1/4})\,,
\end{aligned}
$$

what yields (62). We have used Lemma C.1 from Bibinger et al. (2013) for the magnitude of $\|W_{jk}\|$, the bound (64) and a bound for the sum over j, for which holds

$$
\frac{1}{\sqrt{n}h_n} \sum_{j=1}^{nh_n-1} \Big(1 + \frac{j^2}{nh_n^2}\Big)^{-2} \to \frac{\pi}{2}
$$

by an analogous integral approximation as used in the limiting variance before. Drift terms and cross terms including the drift are asymptotically negligible and are handled similarly as before. Directly neglecting drift terms, we deduce $\|S_{jk} - \tilde{S}_{jk}\| = \mathcal{O}_{\mathbb{P}}(h_n)$ uniformly from $(S_{jk} - \tilde{S}_{jk})^{(p)} \asymp_p \sum_{i=1}^{n_p} \Delta_i^n \tilde{C}^{(p)} \Phi_{jk}(\bar{t}_i)$ with (42b). (63) is equivalent to

$$
n^{1/4} \left\| \sum_{k=1}^{h_n^{-1}} \int_{(k-1)h_n}^{kh_n} \left(\Sigma_s - \Sigma_{(k-1)h_n}\right) ds \right\| \xrightarrow{ucp} 0\,. \tag{65}
$$

For $(\Sigma - 1)$ on Assumption (H-d), we have

$$
\left\| \sum_{k=1}^{h_n^{-1}} \int_{(k-1)h_n}^{kh_n} \left(\Sigma_s - \Sigma_{(k-1)h_n}\right) ds \right\| \leq \sum_{k=1}^{h_n^{-1}} h_n \sup_{(k-1)h_n \leq s \leq kh_n} \|\sigma_s - \sigma_{(k-1)h_n}\| = h_n^{\alpha} = \scriptstyle{\mathcal{O}}(n^{-1/4})
$$

uniformly, since $\alpha > 1/2$, by (64) and boundedness of σ on $[0, 1]$.
For $(\Sigma - 2)$ on Assumption (H-d), consider the decomposition:

$$
\begin{aligned}
\Sigma_s - \Sigma_{(k-1)h_n} &= \sigma_s\sigma_s^{\top} - \sigma_{(k-1)h_n}\sigma_{(k-1)h_n}^{\top} \\
&= (\sigma_s - \sigma_{(k-1)h_n})\sigma_{(k-1)h_n}^{\top} + \sigma_{(k-1)h_n}(\sigma_s^{\top} - \sigma_{(k-1)h_n}^{\top}) \\
&\quad + (\sigma_s - \sigma_{(k-1)h_n})(\sigma_s^{\top} - \sigma_{(k-1)h_n}^{\top})
\end{aligned}
$$

for $s \in [(k-1)h_n, kh_n]$. Here we derive uniformly the magnitude $h_n = \scriptstyle{\mathcal{O}}_{\mathbb{P}}(n^{-1/4})$ for all addends after application of the triangular inequality, since we can use that $(\sigma_s - \sigma_{(k-1)h_n})$ is a semimartingale starting in zero w.r.t the filtration restricted to the time subinterval and analogously for the second addend. Standard bounds for their first two moments then readily render the magnitudes and we conclude (63).
For the spectral covolatility estimator (25a) we may conduct an analysis of the remainder similarly as in the proof of Proposition 4.2. One can as well employ integration by parts of Itô integrals after supposing again a synchronous observation design $t_i, i = 0, \ldots, n$, possible according to Lemma 4.4:

$$
\Delta_i^n \tilde{C}^{(p)} \Delta_i^n \tilde{C}^{(q)} - \int_{t_{i-1}}^{t_i} \big(\Sigma_s^{(pq)} - \Sigma_{\lfloor sh_n^{-1} \rfloor h_n}^{(pq)}\big) ds
$$

$$= \int_{t_{i-1}}^{t_i} \big(\tilde C_s^{(p)} - \tilde C_{t_{i-1}}^{(p)}\big) d\tilde C_s^{(q)} + \int_{t_{i-1}}^{t_i} \big(\tilde C_s^{(q)} - \tilde C_{t_{i-1}}^{(q)}\big) d\tilde C_s^{(p)} \,. \tag{66}$$

with $\tilde C$ approximation errors as in (56b). Consider the random variables

$$\tilde\zeta_{jk}^{(pq)} = \sum_{i=1}^{n} \Delta_i \tilde C^{(p)} \Phi_{jk}(\bar t_i) \sum_{v=1}^{n} \Delta_v \tilde C^{(q)} \Phi_{jk}(\bar t_v)\,,$$

$$\tilde\zeta_k^n = h_n \sum_{j\ge 1} w_{jk}^{p,q} \tilde\zeta_{jk}^{(pq)} - \int_{kh_n}^{(k+1)h_n} \big(\Sigma_s^{(pq)} - \Sigma_{\lfloor s h_n^{-1}\rfloor h_n}^{(pq)}\big) ds\,.$$

Inserting (66) for $\Delta_i^n \tilde C^{(p)} \Delta_i^n \tilde C^{(q)}$, using $[\int Z\,dX, \int Z\,dX] = \int Z^2\,d[X,X]$ for Itô integrals and applying Burkholder-Davis-Gundy inequalities and using the bound (42b) for $\mathbb{E}\big[\big(\Delta_i^n \tilde C^{(p)}\big)^2\big]$, $\mathbb{E}\big[\big(\Delta_i^n \tilde C^{(q)}\big)^2\big]$, it follows that $\mathbb{E}\big[\big(\tilde\zeta_k^n\big)^2\big] = \mathcal{O}(n^{-1})$. Bounds for cross terms with $\tilde C$ and $\bar C$ readily follow by standard estimates and we conclude our claim.

6.5. **Proofs for adaptive estimation**

We carry out the proof of Proposition 4.3 in the case $d = 1$ explicitly. We need to show that

$$n^{1/4} \left| \widehat{\mathbf{IV}}_{n,t} - \widehat{\mathbf{IV}}_{n,t}^{or}(Y) \right| \xrightarrow{ucp} 0 \;\text{ as } n \to \infty\,. \tag{67}$$

Since the noise constitutes the dominant component in observed increments $\Delta_i^n Y = \Delta_i^n X + (\epsilon_i - \epsilon_{i-1})$, it is a simple task to show that the effect of estimating the noise level η is asymptotically negligible. We shall concentrate on the harder problem of analyzing the plug-in estimation of the estimated instantaneous squared volatility process σ_t^2 in the weights. We have to bound

$$\left| \widehat{\mathbf{IV}}_{n,t} - \widehat{\mathbf{IV}}_{n,t}^{or}(Y) \right| = \left| \sum_{k=1}^{\lfloor t h_n^{-1}\rfloor} h_n \sum_{j=1}^{n h_n - 1} \big(\hat w_{jk} - w_{jk}\big)\Big(S_{jk}^2 - [\varphi_{jk},\varphi_{jk}]_n \frac{\eta^2}{n}\Big) \right|,$$

uniformly with w_{jk} the optimal oracle weights (21) and $\hat w_{jk}$ their adaptive estimates. We introduce a coarse grid of blocks of lengths r_n such that $r_n h_n^{-1} \to \infty$ as $n \to \infty$. We analyze the above absolute value of the difference in this double asymptotic framework, where the plug-in estimators are evaluated on the coarse grid first:

$$\left| \widehat{\mathbf{IV}}_{n,t} - \widehat{\mathbf{IV}}_{n,t}^{or}(Y) \right| = \left| \sum_{m=1}^{\lfloor t r_n^{-1}\rfloor} h_n \sum_{k=(m-1)r_n h_n^{-1}+1}^{m r_n h_n^{-1}} \sum_{j=1}^{n h_n - 1} \big(w_j(\hat\sigma_{(m-1)r_n}^2) - w_j(\sigma_{(m-1)r_n}^2)\big) Z_{jk} \right|$$

with $Z_{jk} = S_{jk}^2 - [\varphi_{jk},\varphi_{jk}]_n \eta^2/n - \sigma_{(k-1)h_n}^2$, because $\sum w_{jk} = \sum \hat w_{jk} = 1$ and where we write the weights as function of σ^2:

$$w_j(\sigma^2) = \frac{\big(\sigma^2 + \frac{\eta^2}{n}[\varphi_{jk},\varphi_{jk}]_n\big)^{-2}}{\sum_{l=1}^{n h_n - 1} \big(\sigma^2 + \frac{\eta^2}{n}[\varphi_{lk},\varphi_{lk}]_n\big)^{-2}}\,.$$

We conclude the uniform upper bound

$$\left| \widehat{\mathbf{IV}}_{n,t} - \widehat{\mathbf{IV}}_{n,t}^{or}(Y) \right| \le \sum_{m=1}^{r_n^{-1}} h_n \sum_{j=1}^{n h_n - 1} \left| w_j(\hat\sigma_{(m-1)r_n}^2) - w_j(\sigma_{(m-1)r_n}^2) \right| \left| \sum_{k=(m-1)r_n h_n^{-1}+1}^{m r_n h_n^{-1}} Z_{jk} \right|,$$

since the weights depend here on the same block of the coarse grid not on k. We prove that the right-hand side is $\mathcal{O}_{\mathbb{P}}(n^{-1/4})$ in two steps: The above terms is $\mathcal{O}_{\mathbb{P}}(n^{-1/4})$ and also the remainder induced by the difference of a plug-in estimator on the coarse and finer grid is $\mathcal{O}_{\mathbb{P}}(n^{-1/4})$.
Covariances of the Z_{jk} are asymptotically negligible. We may directly consider $\tilde{Z}_{jk} = \tilde{S}_{jk}^2 - [\varphi_{jk}, \varphi_{jk}]_n \eta^2/n - \sigma^2_{(k-1)h_n}$, the statistics under locally parametric volatility and without drift by the asymptotic negligibility of the remainder proved above in Proposition 4.2. Then, the observation

$$\begin{aligned}\mathbb{E}[\tilde{Z}_{jk}\tilde{Z}_{j(k-l)}] &= \mathbb{E}[\mathbb{E}[\tilde{Z}_{jk}\tilde{Z}_{j(k-l)}|\mathcal{G}_{(k-1)h_n}]] \\ &= \mathbb{E}[\tilde{Z}_{j(k-l)}\mathbb{E}[\tilde{Z}_{jk}|\mathcal{G}_{(k-1)h_n}]] = 0\,,\end{aligned}$$

for all $k = 1, \ldots, h_n^{-1}, l = 1, \ldots, (k-1)$, by (46) shows that $\mathbb{V}\mathrm{ar}(\sum_k \tilde{Z}_{jk}) = \sum_k \mathbb{V}\mathrm{ar}(\tilde{Z}_{jk})$ and thus $\mathbb{V}\mathrm{ar}(\sum_k Z_{jk}) = \sum_k \mathbb{V}\mathrm{ar}(Z_{jk}) + \mathcal{O}(1)$.
The crucial property to ensure tightness of the adaptive approach is a uniform bound on the first derivatives of the weight functions: $w_j(\sigma^2)$ is continuously differentiable with the derivatives satisfying:

$$\left|w_j'\left(\sigma^2\right)\right| \lesssim w_j\left(\sigma^2\right)\log^2(n)\ . \tag{68}$$

In fact, $w_j(\sigma^2)$ is well-defined on $(-\frac{\eta^2}{n}\left[\varphi_{1k}, \varphi_{1k}\right]_n, \infty)$ and the differentiability is clear. To keep notation simple set $c_j = \frac{\eta^2}{n}\left[\varphi_{jk}, \varphi_{jk}\right]_n$. Then, we have

$$\begin{aligned}\left|w_j'(x)\right| &= \left|\frac{-2\left(x+c_j\right)^{-3}\sum_{m=1}^{nh-1}\left(x+c_m\right)^{-2} - \left(x+c_j\right)^{-2}\sum_{m=1}^{nh-1}\left((-2)\left(x+c_m\right)^{-3}\right)}{\left(\sum_{m=1}^{nh-1}\left(x+c_m\right)^{-2}\right)^2}\right| \\ &\leq 2w_j(x)\frac{\sum_{m=1}^{nh-1}\left(x+c_m\right)^{-2}\left|\left(x+c_j\right)^{-1} - \left(x+c_m\right)^{-1}\right|}{\sum_{m=1}^{nh-1}\left(x+c_m\right)^{-2}} \lesssim w_j(x)\log^2(n)\end{aligned}$$

for n sufficiently large. The last inequality follows from

$$\left|\left(x+c_j\right)^{-1} - \left(x+c_m\right)^{-1}\right| \leq \frac{1}{c_j} + \frac{1}{c_m} \lesssim \frac{1}{c_1} = \mathcal{O}\left(\log^2(n)\right).$$

The plug-in estimator established in Bibinger and Reiß (2013) satisfies $\|\hat{\sigma}^2 - \sigma^2\|_{L^1} = \mathcal{O}_{\mathbb{P}}(\delta_n)$ for the L^1-norm $\|\cdot\|_{L^1}$. Observe that with (44):

$$\sum_{j=1}^{nh_n-1} w_{jk}(\mathbb{V}\mathrm{ar}(\tilde{Z}_{jk}))^{1/2} \lesssim \sum_{j=1}^{nh_n-1}\left(j^{-4}(\sqrt{n}h_n)^4 \wedge 1\right)\left(1 \vee j^2(\sqrt{n}h_n)^{-2}\right) \lesssim \log^2(n)\,. \tag{69}$$

Therefore, using Δ-method we obtain that uniformly

$$\left|\widehat{\mathbf{IV}}_{n,t} - \widehat{\mathbf{IV}}^{or}_{n,t}(Y)\right| = \mathcal{O}_{\mathbb{P}}\left(\sqrt{\frac{h_n}{r_n}}(\log n)^4\delta_n\right),$$

with the order $|w_j(\hat{\sigma}^2_{(m-1)r_n} - w_j(\sigma^2_{(m-1)r_n})| = \mathcal{O}_{\mathbb{P}}(w_j(\sigma^2_{(m-1)r_n})\delta_n\log^2(n))$ using (68), δ_n^{-1} being the convergence rate of the plug-in estimator which is $n^{1/8}$ or faster, the negligibility of correlations of the $(Z_{jk})_k$ and (69). Here, we require that $r_n \to 0$ not too fast, i.e. $r_n^{-1} \lesssim n^{1/4}(\log n)^{-9}$.

Consider the remainder by the difference of coarse and finer grid. Since both versions are unbiased it is enough to bound the variance of the difference by:

$$\sum_{m=1}^{r_n^{-1}} \sum_{k=(m-1)r_n h_n^{-1}+1}^{m r_n h_n^{-1}} h_n^2 \left(\sum_{j=1}^{n h_n - 1} \left(\mathbb{E}\left[\left(w_j(\sigma^2_{(k-1)h_n}) - w_j(\sigma^2_{(m-1)r_n}) \right)^2 \tilde{Z}_{jk}^2 \right] \right)^{1/2} \right)^2 \lesssim h_n r_n \log^6(n)\,,$$

using (68) and (69), such that we require $r_n \to 0$ fast enough that
$r_n \log^6(n) \to 0$. In fact, we can easily find some $r_n \to 0$ to render the above errors negligible.
The proofs that Theorem 2 and Theorem 3 extend from the oracle to the adaptive versions of the estimators (25a) and (29a) can be conducted in an analogous way. For covariation matrix estimation, the key ingredient is the uniform bound on the norm of the matrix derivative of the weight matrix function $W_j(\Sigma)$ w.r.t. Σ, which is a matrix with d^6 entries and requires a notion of matrix derivatives, see Lemma C.2 in Bibinger et al. (2013). The proof is then almost along the same lines as the proof of Theorem 4.4 in Bibinger et al. (2013), with the only difference in the construction being that the Z_{jk} are not independent, but still have negligible correlations. The adaptivity in the proof of Theorem 4.4 of Bibinger et al. (2013) is proved under more delicate asymptotics of asymptotically separating sample sizes. For this reason, but at the same time not having the remainders, the restrictions on r_n are different there.

References

Abadir, K. M., Magnus, J. R., 2009. Matrix algebra, Vol. 1, Econometric Exercises. Cambridge University Press.

Aït-Sahalia, Y., Fan, J., Xiu, D., 2010. High-frequency estimates with noisy and asynchronous financial data. Journal of the American Statistical Association 105 (492), 1504–1516.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2008. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. Econometrica 76 (6), 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2011. Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Journal of Econometrics 162 (2), 149–169.

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

Bibinger, M., Hautsch, N., Malec, P., Reiß, M., 2013. Estimating the quadratic covariation matrix from noisy observations: Local method of moments and efficiency. preprint, HU Berlin.

Bibinger, M., Reiß, M., 2013. Spectral estimation of covolatility from noisy observations using local weights. Scandinavian Journal of Statistics, forthcoming.

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

Clément, E., Delattre, S., Gloter, A., 2013. An infinite dimensional convolution theorem with applications to the efficient estimation of the integrated volatility. Stochastic Processes and their Applications 123, 2500–2521.

Fukasawa, M., 2010. Realized volatility with stochastic sampling. Stochastic Processeses and their Applications 120, 209–233.

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

Hayashi, T., Yoshida, N., 2011. Nonsynchronous covariation process and limit theorems. Stochastic Processes and their Applications 121, 2416–2454.

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

Jacod, J., 2012. Statistics and high frequency data. Proceedings of the 7th Séminaire Européen de Statistique, La Manga, 2007: Statistical methods for stochastic differential equations, edited by M. Kessler, A. Lindner and M. Sørensen.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., Vetter, M., 2009. Microstructure noise in the continous case: the pre-averaging approach. Stochastic Processes and their Applications 119, 2803–2831.

Jacod, J., Mykland, P. A., 2013. Microstructure noise in the continuous case: Efficiency and the adaptive pre-averging method. working paper.

Podolskij, M., Vetter, M., 2010. Understanding limit theorems for semimartingales: a short survey. Statistica Neerlandica 64 (3), 329–351.

Reiß, M., 2011. Asymptotic equivalence for inference on the volatility from noisy observations. Annals of Statistics 39 (2), 772–802.

Renault, E., Sarisoy, C., Werker, B. J. M., 2013. Efficient estimation of integrated volatility and related processes. SSRN: `http://ssrn.com/abstract=2293570`.

Tao, T., 2012. Topics in random matrix theory. Vol. 132 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI.

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

Zhang, L., 2006. Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. Bernoulli 12 (6), 1019–1043.

Zhang, L., Mykland, P. A., Aït-Sahalia, Y., 2005. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association 100 (472), 1394–1411.

SFB 649 Discussion Paper Series 2014

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

001 "Principal Component Analysis in an Asymmetric Norm" by Ngoc Mai Tran, Maria Osipenko and Wolfgang Karl Härdle, January 2014.

002 "A Simultaneous Confidence Corridor for Varying Coefficient Regression with Sparse Functional Data" by Lijie Gu, Li Wang, Wolfgang Karl Härdle and Lijian Yang, January 2014.

003 "An Extended Single Index Model with Missing Response at Random" by Qihua Wang, Tao Zhang, Wolfgang Karl Härdle, January 2014.

004 "Structural Vector Autoregressive Analysis in a Data Rich Environment: A Survey" by Helmut Lütkepohl, January 2014.

005 "Functional stable limit theorems for efficient spectral covolatility estimators" by Randolf Altmeyer and Markus Bibinger, January 2014.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

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