THE UNIVERSITY OF CHICAGO

HIGH-FREQUENCY FUNCTIONAL INFERENCE OF DYNAMIC DATA IN TIME AND FREQUENCY DOMAINS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF STATISTICS

BY
YONGRUI CHEN

CHICAGO, ILLINOIS
AUGUST 2020

TO MY LOVING PARENTS

Σὰ βγεῖς στὸν πηγαιμὸ γιὰ τὴν ᾿Ιθάκη,
νὰ εὔχεσαι νἆναι μακρὺς ὁ δρόμος,
γεμάτος περιπέτειες, γεμάτος γνώσεις.

Κωνσταντίνος Π. Καβάφης, ΙΘΑΚΗ [1911]

When you set out on the journey to Ithaca,
pray that the road be long,
full of adventures, full of knowledge.

Constantine P. Cavafy, ITHACA [1911]

ACKNOWLEDGMENTS

The author is deeply indebted to Professor Per A. Mykland. He kindly opened the door and introduced the author to his research field. The author has benefited tremendously from his generous support and intellectual encouragement. After the stay at the University of Chicago, there is still a lot to learn from his insights and approaches to research. Being blissfully spoiled by Professor Mykland's limitless patience, the author hope more work will soon come to fruition.

The author would like to express gratitude to Professor Mihai Anitescu and Professor Wei Biao Wu. Professor Anitescu exposed the author to the vigorous working of computational scientists in Argonne National Laboratory. The author is grateful to Professor Wu for his generosity and insightful conversations. Their down-to-earth pragmatism, wide-ranging expertise, and mathematical rigor always inspire the author's work.

It would not be a complete description of the author's graduate school education without mentioning Professor Dacheng Xiu's supervision when the author worked as a research assistant on financial data analysis and Doctor Mei Wang's offering of valuable opportunities to work with domain experts on stimulating statistical problems.

The author comes last to the camaraderie in graduate school which is not last in thoughts. The technical discussions with Yoann Potiron in the earlier days, the sharing of research and intellectual curiosity with Christopher McKennan, Peter Panov, Byol Kim and others have become an integral part of the communal journey over the years.

This research was funded by National Science Foundation grants DMS 14-07812, DMS 17-13129, and Stevanovich Fellowship 2019. Some conference travels were supported by SoFiE travel scholarships.

ABSTRACT

This dissertation is about inference problems for general nonlinear functionals of time-varying covariance matrices using high-frequency observations, under a class of nonparametric multivariate Itô semimartingale models. Two distinctive approaches are investigated, which employ time-domain and frequency-domain techniques.

The time-domain technique adopts the pre-averaging method to attenuate noise and utilizes truncation to deal with possible sample-path discontinuities. The frequency-domain technique computes the discrete Fourier transform of sample-path increments which allows for asynchronous observations and missing data, and then estimates the Fourier coefficients of covariance matrix by finite-order Bohr convolution.

Given instantaneous covariance matrix estimates, the functional estimation is carried out by evaluating the functional of interest at these estimates. In case the functional is nonlinear, second-order bias correction becomes necessary in order to achieve the optimal convergence rates in various settings. Closed-form expressions of bias, asymptotic variance and their estimators are available.

Applications are considered. Particularly, the attention is focus on factor models by principal component analysis. The author demonstrates how to combine the general results in this dissertation with matrix calculus to conduct statistical inference of realized principal component analysis for non-stationary noisy high-frequency data.

Main results are:

(1) nonparametric plug-in frameworks for functional estimation;

(2) theoretical guidance on the choice of tuning parameters;

(3) bias corrections in the light of higher-order derivatives;

(4) statistical theories on consistency, convergence rates, asymptotic normality, efficiency;

(5) statistical uncertainty quantification for non-stationary factor analysis;

(6) an empirical analysis of large high-frequency panel data spanning 16 years.

Besides principal component analysis, this dissertation provides a methodological foundation for inference of continuous-time regression models, Laplace transform, generalized method of moments and specification tests, as well as statistical uncertainty quantification. The work presented here extends theories and methodologies of previous literature to more empirically realistic settings by solving non-trivial statistical challenges posed by noisy data and asynchronous observations.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Problem statement

In this dissertation, the author studies several problems surrounding statistical inference of

$$S(g)_t = \int_0^t g(\Sigma_s)\,\mathrm{d}s, \tag{1.1}$$

where $g(\cdot)$ is a smooth linear or nonlinear transformation that is of statistical interest, Σ_s is time-varying covariance matrix at time s. Σ_s can be understood as a matrix-valued random function of time.

The author proposes estimators and relevant uncertainty quantification using non-stationary high-frequency data. Such data are accumulating rapidly, cf. Figure 1.2 and Figure 1.1. A pictorial summary of the sampling frequency is provided in Figure 1.3. Principled analyses with statistical guarantees are in pressing need.

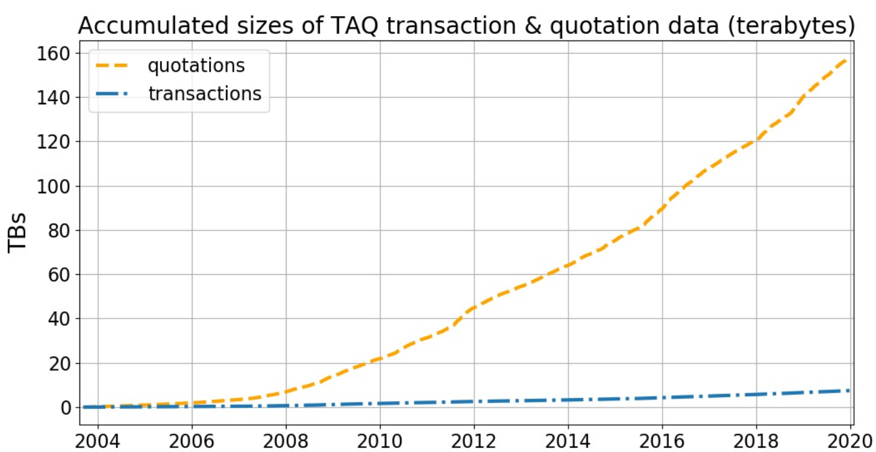


Figure 1.1: Accumulated sizes of TAQ quotation data

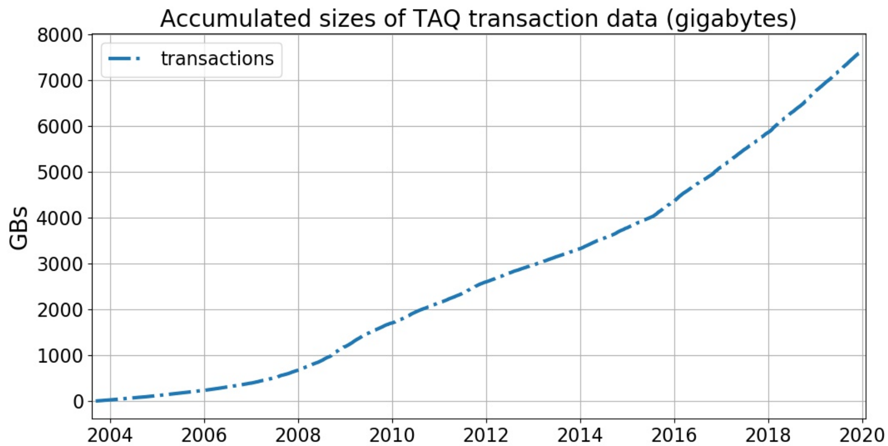


Figure 1.2: Accumulated sizes of TAQ transaction data

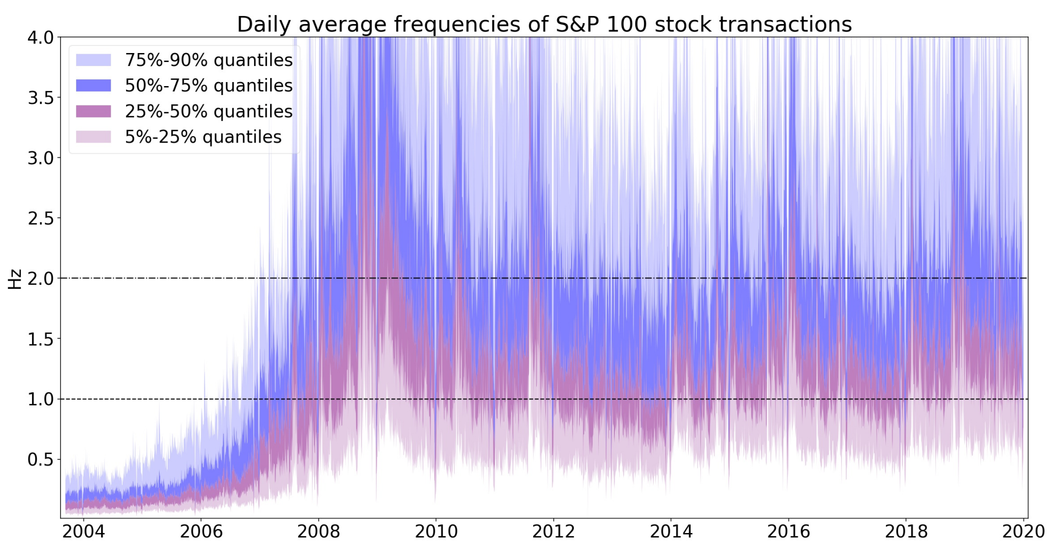


Figure 1.3: Daily average frequencies S&P 100 stock transactions

1.2 Motivations

Astronomical amount of high-frequency data have already attracted close attention from academic researchers, industry practitioners, economic policymakers and regulators. This leads to rapidly growing lines of research endeavors.

The availability of high frequency financial data brings both opportunities and challenges. With the increasing amount of data, according to properties of Itô semimartingales, we can estimate some important quantities of tremendous financial interest with unprecedented high accuracy, for instance powers of volatility and jump components. However, along opportunities come challenges.

- High-frequency data is often contaminated by microstructure noise. The presence of noise invalidates many methods which enjoy nice statistical properties like consistency and asymptotic mixed normality only when noise is absent.

- High-frequency data, especially tick-by-tick data, is not observed irregularly; in the multivariate setting, data in different dimensions are not recorded at the same time grid. This asynchronous/missing data issue causes conceptual and operational difficulties in estimating time-varying covariance.

High-frequency activities in the financial markets present numerous open statistical questions. How to model the observations of fleeting microscopic behaviors? Is there an identifiable and operational notion of cross section of non-stationary high-frequency data? How to design estimators such that they are consistent, rate-optimal, and efficient if achievable? What are the implications for market regulation and policy decisions? Inquiry and investigation into these questions have already become top priorities in the agenda of policymakers and researchers.

1.3 Literature review

Perhaps the most important research topic of all, and the very topic draws most of the attention and interest in high-frequency financial econometrics literature of the recent two decades, is the estimation of time-varying covariance matrix. Various applications concerned in financial econometrics involve volatility. Particularly, the main focus is nonparametric inference of stochastic volatility and related functionals. Various nonparametric estimation methods for volatility using noisy high-frequency financial data have already been established.

i. Zhang et al. [2005] found the first consistent estimator (two-time scale realized volatility, TSRV hereafter) using sub-sampling and averaging in the presence of additive independent noise and Zhang [2006] gave a multi-scale version with the optimal rate of convergence. Li and Mykland [2007] studied the robustness of TS with respect to assumptions on noise in general. Kalnina and Linton [2008] generalized the TS to the model with endogenous and diurnal noise and put forward a modified version of TSRV. Later, Aït-Sahalia et al. [2011] generalized the model to allow for correlated noise under stationary and strong-mixing conditions, a two-time-scale-based method to estimate covariance using asynchronous noisy data was introduced by Zhang [2011].

ii. Barndorff-Nielsen et al. [2008] constructed a kernel-based estimator under the model in which the noise process is temporarily dependent and possibly linearly correlated with the latent Itô process. The inference is robust to endogenous sampling times. Barndorff-Nielsen et al. [2009] investigated the performance of realized kernels in practice, and Barndorff-Nielsen et al. [2011] extended the realized kernel method to covariance matrix estimation.

iii. Jacod et al. [2009] designed a generalized version of the pre-averaging method [Podolskij and Vetter, 2009b], under a Markovian noise model which allows arbitrary fashion of noise but without noise autocorrelation. Jacod et al. [2010] provided an assortment of limit theorems relevant to the pre-averaging method for continuous Itô semimartingales and Itô semimartingales with jumps. Christensen et al. [2010] extended the pre-averaging method for multivariate observations and Hautsch and Podolskij [2013] applied the pre-averaging method in empirical analysis.

iv. Motivated by the likelihood method of Aït-Sahalia et al. [2005], Xiu [2010] established the quasi-maximum likelihood estimation (QMLE) for integrated volatility. Recently, Shephard and Xiu [2017] provided the econometric theory of multivariate QMLE.

v. Reiß [2011] studied the asymptotic properties of a common statistical model underlying the volatility estimation problem using noisy data, showed the statistical model is asymptotically equivalent to a continuously observable Gaussian shift model, and introduced a method utilizing spectral decomposition and local likelihood based on this asymptotic equivalence to estimate spot volatility. An efficient integrated volatility estimator can be formed by aggregating the spot estimates. To overcome non-convexity and estimate covariance matrices, Bibinger et al. [2014] extended this line of work and developed the local generalized method of moments (LMM) and showed this method achieves the efficiency bound. Later on, Altmeyer and Bibinger [2015] showed that the LMM is robust to drift and stochastic volatility. Recently, Bibinger et al. [2019] extended the LMM to accommodate autocorrelated noise and employed the method for spot covariation and correlation in an empirical application.

In recent years, research interests shifts from integrated volatility to spot volatility and the functionals. Various applications concerning high-frequency financial data require functionals of time-varying covariance matrices. To accommodate various needs of pressing importance, inferential theories of general nonlinear functionals of covariance matrices are developed in this thesis.

Previous literature already studied volatility functional estimation using noiseless and regularly sampled observations [Mykland and Zhang, 2009, Jacod and Rosenbaum, 2013, Li et al., 2019]. Following the framework of Jacod and Rosenbaum [2013], good spot volatility estimators are essential in the inference of volatility functionals. In the absence of noise, plugging in finite differences of the realized variances leads to efficient estimation [Jacod and Rosenbaum, 2013, Li et al., 2019]. In this noiseless setting, specialized methods for specific functionals include Laplace transform [Todorov and Tauchen, 2012], generalized method of moments [Li and Xiu, 2016], specification tests [Li et al., 2016], linear regressions [Li et al., 2017], and principal component analysis [Aït-Sahalia and Xiu, 2019].

However, in empirical applications high-frequency data is commonly noisy, and is observed irregularly and asynchronously. To utilize noisy observations, the author uses the pre-averaging method [Podolskij and Vetter, 2009b, Jacod et al., 2009, 2010]. For functional inference when data is asynchronous, the author adopts frequency-domain technique, particularly the Fourier-Malliavin method [Malliavin and Mancino, 2009, Clément and Gloter, 2011].

In developing the theoretical framework, the author tries to make the underlying data-generating process, noise mechanism, and the volatility functionals as general as possible. As a result, the methodologies developed in this thesis are applicable to a broad spectrum of statistical and econometric questions.

1.4 A panoramic review of applications

1.4.1 uncertainty quantification

In the univariate setting, the so-called quarticity which is a functional of volatility appears in the asymptotic variances of many extant volatility estimators. The multivariate counterpart involves more complicated volatility functionals in asymptotic variances. As a result, the volatility functional estimation facilitates uncertainty quantification of various volatility estimators.

1.4.2 Laplace transform

Todorov and Tauchen [2012] put forward an estimator of the realized Laplace transform of volatility defined as

$$\int_0^t e^{iw\Sigma_s}\,\mathrm{d}s.$$

This transform can be viewed as the characteristic function of volatility under the occupation measure. By matching the the moments of realized Laplace transform with those induced by a model, we can estimate model parameters or test the model. An open question noted by Todorov and Tauchen [2012] is the estimation of realized Laplace transform using noisy data.

1.4.3 generalized method of moments

Li and Xiu [2016] proposed the generalized method of integrated moments for financial high-frequency data. In estimating an option pricing model, one observes the process $Z_t = (t, X_t, r_t, d_t)$ where X_t is the price of the underlying observed without any noise, r_t is the short-term interest rate, d_t is the dividend yield. One model of the arbitrage-free option price under the risk-neutral probability measure is

$$\beta_t = f(Z_t, \Sigma_t; \theta^*),$$

where f is deterministic, θ^* is the true model parameter. The observed option price is often modeled as

$$Y_{i\Delta_n} = \beta_{i\Delta_n} + \epsilon_i,$$

where ϵ_i is pricing error and $\mathbb{E}(\epsilon_i) = 0$. Let $g(Z_t, \Sigma_t; \theta) = \mathbb{E}[Y_t - f(Z_t, \Sigma_t; \theta)]$, then we have the following integrated moment condition:

$$G(\theta^*) = 0,$$

where $G(\theta) = \int_0^t g(Z_s, \Sigma_s; \theta)\, \mathrm{d}s$.

1.4.4 linear regression

In the practice of linear factor models and financial hedging, one faces the tasks of computing the factor loadings and the hedge ratios. These tasks can be formulated as the estimation of coefficient β in the time-series linear regression model

$$Z_t^c = \beta^{\mathrm{T}} S_t^c + R_t,$$

where

$$\begin{cases} S_t \equiv & S_0 + \int_0^t b_u^S \,\mathrm{d}u + \int_0^t \sigma_u^S \,\mathrm{d}W_u^S + J_t^S \\ Z_t \equiv & Z_0 + \int_0^t b_u^Z \,\mathrm{d}u + \beta^{\mathrm{T}} \int_0^t \sigma_u^S \,\mathrm{d}W_u^S + \int_0^t \sigma_u^R \,\mathrm{d}W_u^R + J_t^Z, \end{cases}$$

$\langle W^S, W^R \rangle = 0$, $S_t \in \mathbb{R}^{d-1}$, $Z_t \in \mathbb{R}$, and S^c, Z^c are the continuous parts of the Itô semimartingales.

Let $X = (S^{\mathrm{T}}, Z)^{\mathrm{T}}$, we can write $X_t = X_0 + \int_0^t b_u \,\mathrm{d}u + \int_0^t \sigma_u \,\mathrm{d}W_u + J_t$ where $b = (b^{S,T}, b^Z)^{\mathrm{T}}$, $W = (W^{S,\mathrm{T}}, W^R)^{\mathrm{T}}$, $J = (J^{S,\mathrm{T}}, J^Z)^{\mathrm{T}}$ and

$$\sigma = \begin{bmatrix} \sigma^S & 0 \\ \beta^{\mathrm{T}} \sigma^S & \sigma^R \end{bmatrix},$$

so

$$\Sigma \coloneqq \sigma\sigma^{\mathrm{T}} = \begin{bmatrix} \sigma^S\sigma^{S,\mathrm{T}} & \sigma^S\sigma^{S,\mathrm{T}}\beta \\ \beta^{\mathrm{T}}\sigma^S\sigma^{S,\mathrm{T}} & \beta^{\mathrm{T}}\sigma^S\sigma^{S,\mathrm{T}}\beta + (\sigma^R)^2 \end{bmatrix} \coloneqq \begin{bmatrix} \Sigma^{SS} & \Sigma^{SZ} \\ \Sigma^{ZS} & \Sigma^{ZZ} \end{bmatrix},$$

hence by letting $g(\Sigma) = \Sigma^{SS,-1}\Sigma^{SZ}$, we have $\beta = t^{-1}S(g)_t$. Li et al. [2017] proposed this method for the situation in which the process X can be perfectly observed without noise.

1.4.5 principal component analysis

An interesting question about stochastic volatility is its spectral structure $\Sigma_s v_s = \lambda_s v_s$. Aït-Sahalia and Xiu [2019] applied principal component analysis to non-stationary financial data by conducting inference on the realized eigenvalue $\int_0^t \lambda_s \, \mathrm{d}s$, realized eigenvector $\int_0^t v_s \, \mathrm{d}s$, realized principal component $\int_0^t v_{s-} \, \mathrm{d}X_s$. In the basic setting where λ_s is a simple eigenvalue of Σ_s and v_s is the corresponding eigenvector, both $g(\Sigma_s) = \lambda_s$ and $g(\Sigma_s) = v_s$ are continuously differentiable. It turns out that the inferential results of $S(g)$ are applicable to principal component analysis. More recently, Chen et al. [2019] extends principal component analysis to asynchronous and noisy high-dimensional data.

To improve the convergence rate and estimation accuracy of principal component analysis, the author will introduce new estimators in Section 5.3.

CHAPTER 2

STATISTICS OF ITÔ SEMIMARTINGALES

2.1 Model

The following general *Itô semimartingale model* is used in this dissertation.

The *latent* process is a $\mathbb{R}^d$-valued Itô semimartingale X described by the following *Grigelionis decomposition form* on a filtered probability space

$$\big(\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}^{(0)}_t)_t, \mathbb{P}^{(0)}\big),$$

which is large enough to support the Itô semimartingale and whose information filtration $(\mathcal{F}^{(0)}_t)_t$ is rich enough so that X_t is $\mathcal{F}^{(0)}_t$-measurable $\forall t \geq 0$ and

$$\begin{aligned} X_t = X_0 + \underbrace{\int_0^t b_u \, \mathrm{d}u}_{\text{drift}} + \underbrace{\int_0^t \sigma_u \, \mathrm{d}W_u}_{\text{Gaussian martingale}} & \\ + \underbrace{\int_{(0,t]\times E} \delta(\cdot\,; u, x) \mathbb{1}_{\{\|\delta(\cdot\,;u,x)\| \leq 1\}} \, (\mathfrak{p} - \mathfrak{q}) \, (\cdot\,; \mathrm{d}u, \mathrm{d}x)}_{\text{pure jump martingale}} & \\ + \underbrace{\int_{(0,t]\times E} \delta(\cdot\,; u, x) \mathbb{1}_{\{\|\delta(\cdot\,;u,x)\| > 1\}} \, \mathfrak{p}(\cdot\,; \mathrm{d}u, \mathrm{d}x)}_{\text{large jumps}}, & \end{aligned} \tag{2.1}$$

where

(1) b is a $\mathbb{R}^d$-valued, optional, càdlàg process;

(2) W is a $\mathbb{R}^{d'}$-valued standard Brownian motion;

(3) δ is $\mathbb{R}^d$-valued, predictable function on $\Omega^{(0)} \times \mathbb{R}^+ \times E$ with E being a complete separable

metric space (Polish space) endowed with a σ-finite measure λ having no atom;

(4) $\mathfrak{p}$ is a Poisson random measure on $\mathbb{R}^+ \times E$, and $\mathfrak{q} = \mathrm{d}t \cdot \lambda(\mathrm{d}x)$ is the predictable compensator (unique up to $\mathbb{P}^{(0)}$-null set) of $\mathfrak{p}$ in the sense that $t\lambda(A) = E_{\mathbb{P}^{(0)}}[\mu(\cdot\,; (0,t], A)]$ and $(\mu - \nu)((0,t], A)$ is a local martingale, for $\forall A \in \mathcal{B}(E)$ and $\forall t > 0$;

(5) σ is a $\mathbb{R}^d \times \mathbb{R}^{d'}$-valued, progressively measurable, càdlàg process with $d' \geq d$, and it satisfies the property that

$$c_t \equiv \sigma_t \sigma_t^{\mathrm{T}} \in \mathcal{M}_d^+, \forall t \geq 0, \tag{2.2}$$

almost surly, where $\mathcal{M}_d^+$ is the space of $\mathbb{R}^{d\times d}$-valued positive semi-definite matrices. Oftentimes, we need to assume the following assumption on the dynamics of stochastic volatility.

Assumption A-c (stochastic volatility)**.** *The drift b has a Hölder$-1/2$ sample path almost surely, and*

$$\begin{aligned} c_t = c_0 + \int_0^t b_u^{(c)}\,\mathrm{d}u + \int_0^t \sigma_u^{(c)}\,\mathrm{d}W_u \\ + \underbrace{\int_{(0,t]\times E} \delta^{(c)}(\cdot\,; u, x)\mathbb{1}_{\{\|\delta^{(c)}(\cdot\,;u,x)\|\leq 1\}}\,(\mathfrak{p}-\mathfrak{q})\,(\cdot\,; \mathrm{d}u, \mathrm{d}x)}_{\textit{pure jump martingale component in volatility}} \\ + \underbrace{\int_{(0,t]\times E} \delta^{(c)}(\cdot\,; u, x)\mathbb{1}_{\{\|\delta^{(c)}(\cdot\,;u,x)\|> 1\}}\,\mathfrak{p}(\cdot\,; \mathrm{d}u, \mathrm{d}x)}_{\textit{large volatility jumps}}, \end{aligned} \tag{2.3}$$

where $b^{(c)}$ is $\mathbb{R}^{d\times d}$-valued, optional, càdlàg process; $\sigma^{(c)}$ is $\mathbb{R}^{d\times d\times d'}$-tensor-valued, adapted, càdlàg process; $\delta^{(c)}$ is $\mathbb{R}^{d\times d}$-valued, predictable function on $\Omega^{(0)} \times \mathbb{R}^+ \times E$.

c_t is called stochastic volatility and can be understood as instantaneous covariance matrix. The *integrated volatility* (aka the continuous part of *quadratic variation*) is defined as

$$C_t = \int_0^t c_s \,\mathrm{d}s. \tag{2.4}$$

The integrated volatility can be understood as a generalization of the notion of covariance matrix to non-stationary continuous-time processes.

To model noisy observations, let's introduce *hidden Itô semimartingale model*.

(1) The d-dimensional *observable* process, namely Y, can be described on another filtered probability space $\big(\Omega^{(1)}, \mathcal{F}^{(1)}, (\mathcal{F}_t^{(1)})_t, \mathbb{P}^{(1)}\big)$.

(2) Define the difference between the observable process Y and the latent process X to be microstructure noise process ε.

(3) To investigate the potential dependence structure between the market microstructure noise ε and the latent process (or efficient price in financial term) X, and to discuss robustness of any inferential theory with respect to different natures of market microstructure noise, it is beneficial to introduce a conditional probability measure on $\big(\Omega^{(1)}, \mathcal{F}^{(1)}\big)$:

$$Q_t(\cdot\,,\cdot) : \big(\Omega^{(0)}, \mathcal{F}^{(0)}\big) \times \big(\Omega^{(1)}, \mathcal{F}_t^{(1)}\big) \longmapsto [0,1]. \tag{2.5}$$

In particular, we have $\int_{\Omega^{(1)}} Q_t\big(A^{(0)}, \mathrm{d}\omega^{(1)}\big) = 1,$ for $\forall t \geq 0$ and $\forall A^{(0)} \in \mathcal{F}^{(0)}$. Then, we can define the conditional noise variance process:

$$\gamma_t \equiv \int_{\Omega^{(1)}} Y_t(\omega^{(1)}) Y_t(\omega^{(1)})^{\mathrm{T}} \, Q_t(\omega^{(0)}, \mathrm{d}\omega^{(1)}) - X_t X_t^{\mathrm{T}}. \tag{2.6}$$

(4) All the stochastic dynamics in the hidden semimartingale model can be described on the filtered extension $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\big)$, where

(i) $\Omega = \Omega^{(0)} \times \Omega^{(1)}$;

(ii) $\mathcal{F} = \mathcal{F}^{(0)} \otimes \mathcal{F}^{(1)}$;

(iii) $\mathcal{F}_t = \bigcap_{s>t} \big(\mathcal{F}_s^{(0)} \otimes \mathcal{F}_s^{(1)}\big)$, $\widetilde{\mathcal{F}}_t = \bigcap_{s>t} \big(\mathcal{F}^{(0)} \otimes \mathcal{F}_s^{(1)}\big)$;

(iv) $\mathbb{P}\big(A^{(0)} \times \mathrm{d}\omega^{(1)}\big) = \mathbb{P}^{(0)}(A^{(0)}) \cdot \otimes_{t\geq 0} Q_t(A^{(0)}, \mathrm{d}\omega^{(1)})$, $\forall A^{(0)} \in \mathcal{F}^{(0)}$.

To preserve the property on the enlarged probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ of any Itô semimartingale which is originally defined on $\big(\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)})_t, \mathbb{P}^{(0)}\big)$, it requires that $Q_t(\cdot\,, A^{(1)}) : (\Omega^{(0)}, \mathcal{F}^{(0)}) \longmapsto [0,1]$ is $\mathcal{F}_t^{(0)}$-measurable $\forall A^{(1)} \in \mathcal{F}_t^{(1)}$.

(6) We let $\mathbb{E}(\cdot)$ be the expectation operator, either on $(\Omega^{(0)}, \mathcal{F}^{(0)})$ or $(\Omega, \mathcal{F})$; we let $E(\cdot|\mathcal{H})$ be the conditional expectation operator with $\mathcal{H}$ being $\mathcal{F}_t^{(0)}$, $\mathcal{F}_t^{(1)}$, $\mathcal{F}_t$ or $\widetilde{\mathcal{F}}_t$.

To guarantee identifiability and consistent estimation under this hidden Itô semimartingale model, we need to assume the microstructure noise is mean-0 conditional on $\mathcal{F}^{(0)}$. Moreover, for deriving second-order asymptotics of functional estimation under the hidden Itô semimartingale model, it also requires that the time-varying noise variance itself is another Itô semimartingale. Below we summarize the necessary assumption on the noise.

Assumption A-γ (noise)**.** *For* $\forall t \in \mathbb{R}^+$

$$\int_{\Omega^{(1)}} Y_t(\omega^{(1)})\, Q_t(\omega^{(0)}, \mathrm{d}\omega^{(1)}) = X_t(\omega^{(0)}),$$

further, for $t \neq s$,

$$\int_{\Omega^{(1)}} \int_{\Omega^{(1)}} \big[Y_t(\omega^{(1)}) - X_t(\omega^{(0)})\big] \big[Y_s(\omega^{(1)}) - X_s(\omega^{(0)})\big]^{\mathrm{T}} Q_t(\omega^{(0)}, \mathrm{d}\omega^{(1)})\, Q_s(\omega^{(0)}, \mathrm{d}\omega^{(1)}) = 0,$$

furthermore,

$$\begin{aligned}\gamma_t &= \gamma_0 + \int_0^t b_u^{(r)}\,\mathrm{d}u + \int_0^t \sigma_u^{(r)}\,\mathrm{d}W_u \\ &\quad + \int_{(0,t]\times\mathbb{R}^{d\times d}} \delta^{(r)}(\cdot\,; u,x)\mathbb{1}_{\{\|\delta^{(r)}(\cdot\,;u,x)\|\le 1\}}\,(\mathfrak{p}-\mathfrak{q})\,(\cdot\,;\mathrm{d}u,\mathrm{d}x) \\ &\quad + \int_{(0,t]\times E} \delta^{(r)}(\cdot\,;u,x)\mathbb{1}_{\{\|\delta^{(r)}(\cdot\,;u,x)\|>1\}}\,\mathfrak{p}(\cdot\,;\mathrm{d}u,\mathrm{d}x).\end{aligned}$$

It turns out to be helpful sometimes to decompose the Itô semimartingale into continuous part and discontinuous part.

- When the small jumps are summable, i.e.

  $$\int_{(0,t]\times E} \|\delta(\cdot\,;u,x)\|\mathbb{1}_{\{\|\delta(\cdot\,;u,x)\|\le 1\}}\mathfrak{q}(\cdot\,;u,x) < \infty,$$

  the following decomposition often appears in the analysis of various estimators

  $$X = X_0 + X' + J, \tag{2.7}$$

  where

  $$\begin{aligned}X'_t &= \int_0^t \left[b_u - \int_E \delta(\cdot;u,x)\mathbb{1}_{\{\|\delta(\cdot;u,x)\|\le 1\}}\,\lambda(\mathrm{d}x)\right]\mathrm{d}u + \int_0^t \sigma_u\,\mathrm{d}W_u \\ J_t &= \int_{(0,t]\times E} \delta(\cdot;u,x)\,\mathfrak{p}(\cdot;\mathrm{d}u,\mathrm{d}x).\end{aligned}$$

  The drift in X', namely,

  $$b'_t \equiv b_t - \int_E \delta(\cdot;t,x)\mathbb{1}_{\{\|\delta(\cdot;t,x)\|\le 1\}}\,\lambda(\mathrm{d}x), \tag{2.8}$$

  is the “genuine” drift.

- Another decomposition representation is also handy:

$$X = X_0 + X'' + J', \tag{2.9}$$

where

$$\begin{aligned} X''_t &= \int_0^t \left[b_u + \int_E \delta(\cdot; u, x) \mathbb{1}_{\{\|\delta(\cdot;u,x)\|>1\}} \lambda(\mathrm{d}x) \right] \mathrm{d}u + \int_0^t \sigma_u \, \mathrm{d}W_u \\ J'_t &= \int_{(0,t]\times E} \delta(\cdot; u, x) \, (\mathfrak{p} - \mathfrak{q})(\cdot; \mathrm{d}u, \mathrm{d}x). \end{aligned}$$

To gain desired asymptotic properties of various estimators, we need some regularity conditions on the sample path. We assume the following assumption on the characteristics of the Itô semimartingale. It is called *local boundedness.*

Assumption A-ν (local boundedness)**.** *There are a localization sequence $(\tau_n)_{n\in\mathbb{N}}$ (in other words, a sequence of stopping times) increasing to ∞, and a sequence of nonnegative bounded λ-integrable functions $(\Gamma_n)_{n\geq 1}$ on E such that $\forall \omega^{(0)} \in \Omega^{(0)}$,*

$$t \leq \tau_n(\omega^{(0)}) \implies \begin{cases} \|b_t(\omega^{(0)})\| \leq n, & \|\sigma_t(\omega^{(0)})\| \leq n \\ \|b_t^{(c)}(\omega^{(0)})\| \leq n, & \|\sigma_t^{(c)}(\omega^{(0)})\| \leq n \\ \|b_t^{(r)}(\omega^{(0)})\| \leq n, & \|\sigma_t^{(r)}(\omega^{(0)})\| \leq n, \end{cases}$$

$$t < \tau_n(\omega^{(0)}) \implies \begin{cases} \|\delta(\omega^{(0)}, t, x)\|^\nu \wedge 1 \leq \Gamma_n(x), \ \nu \in [0,1), \forall x \in \mathbb{R}^d \\ \|\delta^{(c)}(\omega^{(0)}, t, x)\|^2 \wedge 1 \leq \Gamma_n(x), \ \forall x \in \mathbb{R}^d \\ \|\delta^{(r)}(\omega^{(0)}, t, x)\|^2 \wedge 1 \leq \Gamma_n(x), \ \forall x \in \mathbb{R}^d, \end{cases}$$

and b' is càdlàg.

Based on the localization lemma **Lemma 4.4.9** in Jacod and Protter [2012], without loss of generality, we can replace the local boundedness assumption **Assumption A-ν** to the strong boundedness assumption:

Assumption SA-ν. *boundedness There are a constant K, and a nonnegative bounded λ-integrable function Γ on E such that $\forall\omega^{(0)} \in \Omega^{(0)}$ and $\forall t > 0$,*

$$\begin{cases} \|b_t(\omega^{(0)})\| \le K, & \|\sigma_t(\omega^{(0)})\| \le K \\ \|b_t^{(c)}(\omega^{(0)})\| \le K, & \|\sigma_t^{(c)}(\omega^{(0)})\| \le K \\ \|b_t^{(r)}(\omega^{(0)})\| \le K, & \|\sigma_t^{(r)}(\omega^{(0)})\| \le K, \end{cases}$$

$$\begin{cases} \|\delta(\omega^{(0)}, t, x)\|^\nu \le \Gamma(x),\ \nu \in [0,1), \forall x \in \mathbb{R}^d \\ \|\delta^{(c)}(\omega^{(0)}, t, x)\|^2 \le \Gamma(x),\ \forall x \in \mathbb{R}^d \\ \|\delta^{(r)}(\omega^{(0)}, t, x)\|^2 \le \Gamma(x),\ \forall x \in \mathbb{R}^d, \end{cases}$$

and b' is càdlàg.

2.2 In-fill asymptotics and notation

First we need some notation for observation times.

- U^j (resp. U^{jk}) is the j-th (resp. (j,k)-th) component of U where U is a $\mathbb{R}^d$-valued (resp. $\mathbb{R}^{d\times d}$-valued) process;
- $\mathcal{T}^j = \{\tau_h^j,\ h = 0, \cdots, n_j\}$ is the set of observation times of X^j;
- $I_h^j = (\tau_{h-1}^j, \tau_h^j]$ is time interval between two consecutive observations of X^j;
- $\Delta_h^j = \tau_h^j - \tau_{h-1}^j$ is the length of I_h^j, $\Delta^j = \max_h \Delta_h^j$;
- $\underline{n} = \min_j n_j$, $n = \max_j n_j$, $\Delta(n) = \max_j \Delta^j$;
- δ_h^j is the first-order difference operator according to the observational times $\mathcal{T}^j$, i.e., given a generic scalar process U, $\delta_h^j(U) = U(\tau_h^j) - U(\tau_{h-1}^j)$ is the increment of the process U over the time interval I_h^j.

The statistical methods in this thesis will be analyzed in the *in-fill* asymptotic setting.[1] Speaking in detail, the estimators are constructed from high-frequency data contaminated with noises from a finite time interval with an asymptotically shrinking mesh of the observation grid, i.e., $\Delta(n) \xrightarrow{\mathbb{P}} 0$.

The following notation is used in this dissertation. For $r \in \mathbb{N}^+$, $\mathcal{C}^r(\mathcal{S})$ denotes the space of r-time continuously differentiable functions on the domain $\mathcal{S}$; $\mathcal{S}_d^+$ is the convex cone of $d \times d$ positive semidefinite matrices; $\mathbb{I}_d$ is the $d \times d$ identity matrix; $\|\cdot\|$ denotes a norm on vectors, matrices or tensors; given $a \in \mathbb{R}$, $\lfloor a \rfloor$ denotes the largest integer no more than a; $a \vee b = \max\{a, b\}$, $a \wedge b = \min\{a, b\}$; $a_n \asymp b_n$ means both a_n/b_n and b_n/a_n are bounded for large n; $\mathbf{A}^{\mathrm{T}}$ is the transpose of the vector or matrix $\mathbf{A}$; for a multidimensional array, the entry index is written in the superscript, e.g., $X_t = (X_t^1, \cdots, X_t^d)^{\mathrm{T}}$, $\mathbf{A}^{jk}$ denotes the (j,k)-th entry in the matrix $\mathbf{A}$; $\partial_{jk} g$ and $\partial^2_{jk,lm} g$ denote the gradient and Hessian of g with respect to the (j,k)-th and (l,m)-th entries in its matrix argument; $\xrightarrow{\mathcal{L}-s(f)}$ (resp. $\xrightarrow{\mathcal{L}-s}$) denotes stable convergence of processes (resp. variables) in law[2]; $\xrightarrow{u.c.p.}$ denotes uniform convergence in probability on compact sets; $\mathcal{MN}(0, V)$ is a mixed Gaussian distribution with mean 0 and conditional variance V.

2.3 Moments of Itô semimartingale increments

The most important and most commonly used result in the inferential theory of Itô semimartingales is the *Burkholder-Davis-Gundy inequality.* For the slightly confined case where $p > 1$, the inequality is called *Burkholder-Gundy inequality.*

1. It is also known as fixed-domain asymptotics in spatial statistics, high-frequency asymptotics in financial econometrics.

2. See section 2.2.1, 2.2.2 in Jacod and Protter [2012] for stable convergence. The sampling distributions of my estimators depends on the sample path of covariance matrix. I need a mode of convergence in which my estimators converges jointly with other random variables, so that I can estimate the asymptotic variances of sampling distributions consistently to compute confidence intervals.

Theorem 1 (Burkholder-Davis-Gundy inequality)**.** *$\forall$ local martingale M adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$ and $\forall$ finite stopping times S and T such that $S \leq T$, $\forall p \geq 1$, $\exists 0 < c_p < C_p < \infty$ such that*

$$c_p \left\| [M,M]_T - [M,M]_S \right\|_{p/2}^{1/2} \leq \left\| \sup_{t \in (S,T]} \left| M_t - M_S \right| \right\|_p \leq C_p \left\| [M,M]_T - [M,M]_S \right\|_{p/2}^{1/2},$$

where $\|Z\|_p := \mathbb{E}[\|Z\|^p]^{1/p}$ for every p-power integrable random variable Z.

Proof. Please refer to Protter [2005]. □

To establish the consistency and stable convergence to a conditional centered Gaussian martingale, we need more estimates and moment bounds for Itô semimartingales (some are classical, others are refined versions of well-known results). Most of these moment bounds for semimartingales come from Jacod and Protter [2012] and the papers cited wherein.

The following lemmas provide the bounds (possibly infinite) for moments of different components of an Itô semimartingales of the form (2.1). In those lemmas, $\|\cdot\|$ denotes the Euclidean norm (i.e., L^2 norm) for vectors and the Frobenius norm for matrices. In practice, we can use those results under different conditions with various choices of the power p. The K's appearing in different places are not necessarily the same which possibly depend on $\{X_t\}_{t\geq 0}$, d, and we use K_q if it depends on an additional parameter q.

Lemma 1 (drift)**.** *Suppose T is a finite stopping time. For any Itô semimartingale given by (2.1), $\forall p \geq 1$*

$$\sup_{u \in [0,s]} \left\| \int_T^{T+u} b_t \, \mathrm{d}t \right\|^p \leq \left(\int_T^{T+s} \|b_u\| \, \mathrm{d}u \right)^p \leq s^p \left(\frac{1}{s} \int_T^{T+s} \|b_u\|^p \, \mathrm{d}u \right).$$

Proof. The first inequality is trivial. For the second inequality, applying Hölder's inequality, we can get

$$\int_T^{T+s} \|b_u\| \, \mathrm{d}u \leq \left(\int_T^{T+s} 1^{\frac{p}{p-1}} \, \mathrm{d}u \right)^{\frac{p-1}{p}} \cdot \left(\int_T^{T+s} \|b_u\|^p \, \mathrm{d}u \right)^{\frac{1}{p}}.$$

Taking the power p on both sides of the inequality will reach the desired conclusion. $\square$

Lemma 2 (continuous Gaussian martingale). *Suppose T is a finite stopping time. For any Itô semimartingale given by (2.1), and $\forall p \geq 1$,*

$$E\left(\sup_{u\in[0,s]} \left\| \int_T^{T+u} \sigma_t \, \mathrm{d}W_t \right\|^p \middle| \mathcal{F}_T^{(0)} \right) \leq K_p s^{p/2} E\left[\left(\frac{1}{s} \int_T^{T+s} \|\sigma_u\|^2 \, \mathrm{d}u \right)^{p/2} \middle| \mathcal{F}_T^{(0)} \right].$$

Proof. This is an immediate consequence of **Theorem 1**. $\square$

To discuss *moment conditions* of the jump components in an Itô semimartingale, we need to define some quantities associated with the predictable function δ.

Definition 1 (quantities associated with jumps).

$$\begin{aligned}
\widehat{\delta}(p, \alpha)_{t,s} &:= \frac{1}{s} \int_t^{t+s} \int_{\mathbb{R}^d} \|\delta(\cdot\,; u, x)\|^p \mathbb{1}_{\{\|\delta(\cdot\,; u, x)\| \leq \alpha\}} \, \lambda(\mathrm{d}x) \, \mathrm{d}u \\
\widehat{\delta}(p)_{t,s} &:= \frac{1}{s} \int_t^{t+s} \int_{\mathbb{R}^d} \|\delta(\cdot\,; u, x)\|^p \, \lambda(\mathrm{d}x) \, \mathrm{d}u \\
\widehat{\delta}'(p)_{t,s} &:= \widehat{\delta}(p, 1)_{t,s} + \frac{1}{s} \int_t^{t+s} \int_{\mathbb{R}^d} \|\delta(\cdot\,; u, x)\| \mathbb{1}_{\{\|\delta(\cdot\,; u, x)\| > 1\}} \, \lambda(\mathrm{d}x) \, \mathrm{d}u \\
\widehat{\delta}''(p)_{t,s} &:= \widehat{\delta}(p, 1)_{t,s} + \frac{1}{s} \int_t^{t+s} \lambda\big(\{x \in \mathbb{R}^d : \|\delta(\cdot\,; u, x)\| > 1\}\big) \, \mathrm{d}u.
\end{aligned}$$

Lemma 3 (jump martingale). *For an Itô semimartingale given by (2.1), we have the following results:*

(i) suppose $\int_0^t \int_{\mathbb{R}^d} \|\delta(\cdot\,; u, x)\|^2 \, \lambda(\mathrm{d}x) \, \mathrm{d}u < \infty$, $\forall t$, then J_t' in (2.9) is a locally square integrable martingale;

(ii) $\forall$ finite stopping time T, $\forall s > 0$ we have

- *if* $p \in [1,2]$

$$E\Big(\sup_{u\in[0,s]} \|J'_{T+u} - J'_T\|^p \Big| \mathcal{F}_T^{(0)}\Big) \le K_p\, sE\big[\widehat{\delta}(p)_{T,s} | \mathcal{F}_T^{(0)}\big], \tag{2.10}$$

- *if* $p \ge 2$

$$E\Big(\sup_{u\in[0,s]} \|J'_{T+u} - J'_T\|^p \Big| \mathcal{F}_T^{(0)}\Big) \le K_p\Big(sE\big[\widehat{\delta}(p)_{T,s} | \mathcal{F}_T^{(0)}\big] + s^{p/2}E\big[\widehat{\delta}(2)_{T,s}^{p/2} | \mathcal{F}_T^{(0)}\big]\Big). \tag{2.11}$$

Remark 1. The variable $\widehat{\delta}(p)_{T,s}$ can be infinite, neither increasing nor decreasing in p in general.

Under the assumption of **Lemma 3**, $\int_0^t \int_{\mathbb{R}^d} \|\delta(\cdot\,; u, x)\|^2 \,\lambda(\mathrm{d}x)\, \mathrm{d}u < \infty$, $\forall t$, $\widehat{\delta}(p)_{T,s}$ is finite $\forall p \ge 2$, yet $\widehat{\delta}(p)_{T,s}$ might be infinite when $p \in [0, 2)$. For this reason, the right-hand side (hereafter RHS) of (2.11) is usually finite, whereas the RHS of (2.10) is infinite in many cases. (2.10) and (2.11) agree in the case $p = 2$. When $\widehat{\delta}(p)_{T,s} < \infty$, the typical magnitude of the RHS of (2.10) is $O_p(s)$ regardless which value of p is taken in $[1, 2]$.

Lemma 4 (normalized jump martingale)**.** *Suppose $p \in [1, 2]$, $\forall q \in [0, 1/p]$, $\forall s \in [0, 1]$, $\forall$ finite stopping time T, we have the following result for the process J' in (2.9):*

$$E\bigg[\sup_{u\in[0,s]} \Big(\frac{\|J'_{T+u} - J'_T\|}{s^q} \wedge 1\Big)^p \Big| \mathcal{F}_T^{(0)}\bigg] \le K\, s^{1-pq} E\Big[\widehat{\delta}(p, s^{q/2})_{T,s} + s^{(p-1)q/2}\widehat{\delta}'(p)_{T,s} \big| \mathcal{F}_T^{(0)}\Big].$$

Remark 2. Under the assumption $\forall t$ $\int_0^t \int_{\mathbb{R}^d} \|\delta(\cdot\,; u, x)\|^2 \,\lambda(\mathrm{d}x)\, \mathrm{d}u < \infty$, using **Lemma 4**, upon taking $p = 2$ and $q = 1/2$, we have

$$\frac{J'_{T+s} - J'_T}{\sqrt{s}} \xrightarrow{\mathbb{P}^{(0)}} 0.$$

Lemma 5 (jump)**.** *We have the following result for the pure jump process J in (2.7):*

(i) if $\int_0^t \lambda\big(\{x \in \mathbb{R}^d : \|\delta(\cdot\,; u, x)\| > 0\}\big)\, \mathrm{d}u < \infty$, $\forall t > 0$, then the process J has finitely

many jumps on any finite interval, almost surly;

(ii) if $\int_0^t \int_{\mathbb{R}^d} \|\delta(\cdot\,; u, x)\| \, \lambda(\mathrm{d}x) \, \mathrm{d}u < \infty$, $\forall t > 0$, *then the process* J *has locally integrable variation.*

Furthermore, for $s > 0$ *and* $\forall$ *finite stopping time* T, *we have*

- *for* $p \in (0, 1]$

$$E\Big(\sup_{u\in[0,s]} \|J_{T+u} - J_T\|^p \Big| \mathcal{F}_T^{(0)} \Big) \leq K_p \, sE\big[\widehat{\delta}(p)_{T,s} \big| \mathcal{F}_T^{(0)}\big],$$

- *for* $p \geq 1$

$$E\Big(\sup_{u\in[0,s]} \|J_{T+u} - J_T\|^p \Big| \mathcal{F}_T^{(0)} \Big) \leq K_p \Big(sE\big[\widehat{\delta}(p)_{T,s} \big| \mathcal{F}_T^{(0)}\big] + s^p E\big[\widehat{\delta}(1)_{T,s}^p \big| \mathcal{F}_T^{(0)}\big] \Big).$$

Lemma 6 (normalized jump)**.** *Set* $p \in (0, 1]$, $\forall q \in [0, 1/p]$, $\forall s \in [0, 1]$, $\forall$ *finite stopping time* T, *the following holds for the d-dimensional process* J *in (2.7):*

$$E\Big[\sup_{u\in[0,s]} \Big(\frac{\|J_{T+u} - J_T\|}{s^q} \wedge 1 \Big)^p \Big| \mathcal{F}_T^{(0)} \Big] \leq K \, s^{1-pq} E\big[\widehat{\delta}(p, s^{q/2})_{T,s} + s^{pq/2} \, \widehat{\delta}'(p)_{T,s} \big| \mathcal{F}_T^{(0)}\big].$$

Lemma 1, **2**, **3**, **5** give the "moments" of non-normalized quantities; on the other hand, **Lemma 4**, **6** are results for normalized quantities of discontinuous parts of Itô semimartingales. The "moments" of normalized quantities are particular handy when $p \leq 2$. In the case where the predictable function δ is bounded in some way, we have a simpler "moment" estimates. In view of the usefulness of "moments" of normalized discontinuous martingales and jump components in deriving the bounds of errors in various estimators, we state a corollary.

Corollary 1 (normalized jumps)**.** *Suppose the d-dimensional predictable function satisfies*

$$\|\delta(\cdot;t,x)\| \leq \Gamma(x),\, \forall t,$$

where Γ *is a measurable function on* $\mathbb{R}$*. Set* $p > 0$*,* $\nu \in (0,2]$ *and* $q \in [0, 1/\nu)$*.*

(i) if $\nu \in (1,2]$ *and* $\int_{\mathbb{R}} \Gamma^{\nu}(x) \wedge \Gamma(x)\, \lambda(\mathrm{d}x) < \infty$*, then* $\forall s \in (0,1]$ *and* $\forall$ *finite stopping time* T*, the process* J' *in (2.9) satisfies*

$$E\Big[\sup_{u\in[0,s]} \Big(\frac{\|J'_{T+u} - J'_T\|}{s^q} \wedge 1\Big)^p \Big| \mathcal{F}_T\Big] \leq \begin{cases} K\, s^{(1-q\nu)p/\nu} a(s) & \text{if } p \leq \nu \\ K\, s^{1-q\nu} a(s) & \text{if } p \geq \nu, \end{cases}$$

where K *and* a *depend on* p, q, ν, Γ *and* λ*; when* $q > 0$*,* $a(s) \to 0$ *as* $s \to 0$*; when* $q = 0$*,* $a < \infty$*;*

(ii) if $\nu \in (0,1]$ *and* $\int_{\mathbb{R}} \Gamma^{\nu}(x) \wedge \Gamma(x)\, \lambda(\mathrm{d}x) < \infty$*, and if* $p > 1$*,* $q < (p-1)(p-\nu)$*, then* $\forall s \in (0,1]$ *and* $\forall$ *finite stopping time* T*, and* K*,* ϕ *in (i), the process* J' *satisfies*

$$E\Big[\sup_{u\in[0,s]} \Big(\frac{\|J'_{T+u} - J'_T\|}{s^q} \wedge 1\Big)^p \Big| F_T\Big] \leq K\, s^{1-q\nu} a(s);$$

(iii) if $\nu \in (0,1]$ *and* $\int_{\mathbb{R}} \Gamma^{\nu}(x) \wedge 1\, \lambda(\mathrm{d}x) < \infty$*, then* $\forall s \in (0,1]$*,* $\forall$ *finite stopping time* T *and* K*,* a *in (i), the process* J *in (2.7) satisfies*

$$E\Big[\sup_{u\in[0,s]} \Big(\frac{\|J_{T+u} - J_T\|}{s^q} \wedge 1\Big)^p \Big| \mathcal{F}_T\Big] \leq \begin{cases} K\, s^{(1-q\nu)p/\nu} a(s) & \text{if } p \leq \nu \\ K\, s^{1-q\nu} a(s) & \text{if } p \geq \nu. \end{cases}$$

Combine the results from **Lemma 1**, **2**, **3**, **5**, we will have an estimate for general Itô semimartingale.

Corollary 2 (Itô semimartingale)**.** *For an Itô semimartingale of the form (2.1),* $\forall$ *finite* *stopping time T and $\forall s > 0$ and $\forall p \geq 2$, the following holds true:*

$$
\begin{aligned}
E\Big(\sup_{u\in[0,s]} \|X_{T+u} - X_T\|^p \Big| \mathcal{F}_T\Big) \;&\leq\; K_p E\bigg[\Big(\int_T^{T+s} \|b_u\|\,\mathrm{d}u\Big)^p + \Big(\int_T^{T+s} \|\sigma_u\|^2\,\mathrm{d}u\Big)^{p/2} \\
&+\; \int_T^{T+s}\int_{\mathbb{R}^d} \|\delta(\cdot\,; u,x)\|^p \lambda(\mathrm{d}x)\,\mathrm{d}u \\
&+\; \Big(\int_T^{T+s}\int_{\{x:\|\delta(\cdot\,;u,x)\|\leq 1\}} \|\delta(\cdot\,; u,x)\|^2 \lambda(\mathrm{d}x)\,\mathrm{d}u\Big)^{p/2} \\
&+\; \Big(\int_T^{T+s}\int_{\{x:\|\delta(\cdot\,;u,x)\|> 1\}} \|\delta(\cdot\,; u,x)\| \lambda(\mathrm{d}x)\,\mathrm{d}u\Big)^{p} \Big|\mathcal{F}_T\bigg].
\end{aligned}
$$

Under **Assumption SA-ν**, we have a stronger version of estimate for Itô semimartingales for $p \geq 2$:

$$
\begin{aligned}
E\Big(\sup_{u\in[0,s]} \|X_{T+u} - X_T\|^p | \mathcal{F}_T\Big) \;&\leq\; K_p\mathbb{E}\bigg[(s^p + s^{p/2}) + s\int_{\mathbb{R}^d} \Gamma^p(x)\lambda(\mathrm{d}x) \\
&+ s^{p/2}\Big(\int_{\mathbb{R}^d} \Gamma^2(x)\mathbb{1}_{\{\Gamma(x)\leq 1\}}\lambda(\mathrm{d}x)\Big)^{p/2} \\
&+ s^p\Big(\int_{\mathbb{R}^d} \Gamma(x)\mathbb{1}_{\{\Gamma(x)> 1\}}\lambda(\mathrm{d}x)\Big)^{p}\bigg].
\end{aligned}
$$

Corollary 3 (standard estimates for X and c)**.** *In the form (2.7) for Itô semimartingale, under **Assumption A**-c, **SA**-ν we have the following results:*

$$
\begin{cases}
\big\|E\big(X'_{T+s} - X'_T|\mathcal{F}_T^{(0)}\big)\big\| + \big\|E\big(c_{T+s} - c_T|\mathcal{F}_T^{(0)}\big)\big\| + \big\|E\big(\gamma_{T+s} - \gamma_T|F_T^{(0)}\big)\big\| & \leq Ks \\
E\left(\sup_{u\in[0,s]} \big\|X'_{T+u} - X'_T\big\|^p|\mathcal{F}_T^{(0)}\right) & \leq Ks^{p/2} \\
E\left(\sup_{u\in[0,s]} \|c_{T+u} - c_T\|^p + \|\gamma_{T+u} - \gamma_T\|^p\,|\mathcal{F}_T^{(0)}\right) & \leq Ks^{(p/2)\wedge 1},
\end{cases}
$$

where T is an almost surely finite stopping time and $p \geq 0$.

Proof. Under **Assumption SA-ν**, the results for the continuous Itô semimartingale X' follow from **Lemma 1**, **2**; the results for c follow from **Lemma 1**, **2**, **3**, **5**. □

The moment bounds introduced above are well known and is far-reaching for theoretical analysis. For the inferential purpose in high-frequency framework with noise, we need more refined "moment estimates". This is the task of section 3.2.

CHAPTER 3

VOLATILITY MATRIX ESTIMATION WITH NOISY DATA: PRE-AVERAGING METHOD

3.1 Pre-averaging method

When the observations are entangled with noise, statistical methods designed for perfectly observed signals would suffer from biases (finite or explosive). If one is interested in some quantities related to the sample path of Itô semimartingale, one general approach to utilize the high-frequency data and meanwhile mitigate the noise effect is the so-called *pre-averaging* method [Podolskij and Vetter, 2009b, Jacod et al., 2009, 2010].

The idea of the aforementioned pre-averaging method is actually quite intuitive – to take a moving window sliding over the time period within which one takes averages (simple or weighted) of the observations recorded in the moving window.[1] If one controls the window length in such way that the length of smoothing window is shrinking to 0 and the number of observations in the window goes to infinity in the meantime, then the averages would be able to approximate the uncontaminated processes satisfactorily provided the microstructure noise does not exhibit long memory.

In this thesis, I consider the pre-averaging method when noisy observations are regularly spaced among all the dimensions.

Assumption A-τ (synchronous observations). $n_1 = \cdots = n_d$ *and* $\min_j \tau_h^j = \max_j \tau_h^j$ *for*

1. We distinguish this method from the contiguity method developed for statistics of Itô semimartingales [Mykland and Zhang, 2009]. The premise of the latter is that the parameter, such as volatility, of an Itô semimartingale is almost constant subject to some controllable and analyzable errors of magnitude the square root of the length of the local window. This idea is prevalent in various methodologies in high-frequency econometrics, and the local-consistency approximation in in-fill asymptotics is valid for many inferential purposes and is a cornerstone for existing high-frequency statistical procedures. For an application of contiguity in the presence of microstructure noise, please see Mykland and Zhang [2016].

$h = 1, \cdots, n_1$.

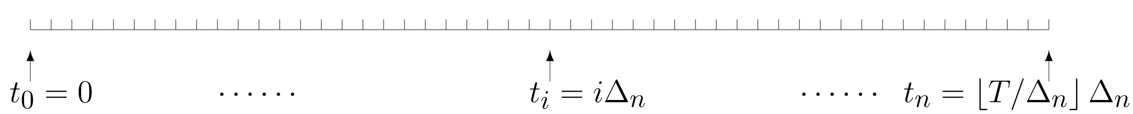
Assumption SA-τ (synchronous and regular observations)**.** *$n_1 = \cdots = n_d$, $\min_j \tau_h^j = \max_j \tau_h^j$ and $\min_h \Delta_h = \max_h \Delta_h$ for $h = 1, \cdots, n_1$.*

I assume **Assumption SA-τ** in chapter 3 and chapter 4. Under **Assumption SA-τ**, $\Delta_n := \Delta_h^j$, $\forall j = 1, \cdots, d$, $h = 1, \cdots, n$.

For notational simplicity, for a process U and a filtration $(\mathcal{G}_t)_t$ living on some underlying probability space, we define the following discretized objects:

$$\begin{aligned}
U_h^n &:= U_{h\Delta_n} \\
\Delta_h^n U &:= U_h^n - U_{h-1}^n \\
\mathcal{G}_h^n &:= \mathcal{G}_{h\Delta_n} \\
I(n, h, h') &:= (h\Delta_n, (h + h')\Delta_n] \\
I(n, h) &:= I(n, h, 1).
\end{aligned}$$

A prototypical regular grid would be the following.



To precisely guide the idea, let's mathematically describe the procedure. First we choose a local window of length k_n used to estimate the spot volatility matrix at time $i\Delta_n$, similarly we can also take a local sub-window of length $l_n = o(k_n)$ to conduct pre-averaging in the local window $I(n, i, l_n) = (i\Delta_n, (i + l_n)\Delta_n]$, then estimate the spot volatility matrix by pre-averaged data.

A visual illustration for the computing scheme for the incoming estimator $\widehat{c}_i^n$ of c_i^n for $i =$

$0, 1, \cdots, \lfloor t/\Delta_n \rfloor - k_n$ or more generally c_t, $t \in I(n, i-1)$ for $i = 1, \cdots, \lfloor t/\Delta_n \rfloor - k_n$ is in Figure 3.1.

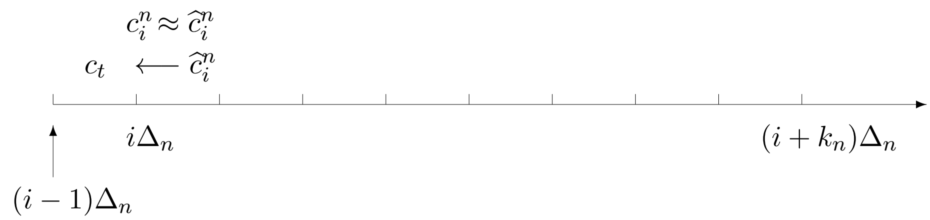


Figure 3.1: Computing scheme of instantaneous covariance

(1) Choose the bandwidths k_n and l_n satisfying

$$\begin{cases} k_n \to \infty, & k_n\Delta_n \to 0 \\ l_n \to \infty, & l_n/k_n \to 0; \end{cases} \tag{3.1}$$

(2) Choose the weight function $\varphi \in C([0,1])$ for the pre-averaging method such that

$$\begin{aligned} &\textstyle\int_0^1 \varphi^2(u)\,\mathrm{d}u > 0,\ \varphi(0) = \varphi(1) = 0 \\ &\varphi(\cdot) \text{ is piecewise } C^1; \end{aligned} \tag{3.2}$$

We also need the following quantities associated with the function φ:

$$\begin{aligned} \phi_0(s) &:= \int_s^1 \varphi(u)\varphi(u-s)\,\mathrm{d}u, & \psi_0 &:= \phi_0(0) \\ \phi_1(s) &:= \int_s^1 \varphi'(u)\varphi'(u-s)\,\mathrm{d}u, & \psi_1 &:= \phi_1(0) \\ \varphi_j^n &:= \varphi(j/l_n), \end{aligned} \tag{3.3}$$

and for any continuous-time processes U, define

$$\begin{aligned} \overline{U}_i^n(l_n;\varphi) &:= \frac{1}{(\psi_0 l_n)^{1/2}} \sum_{h=1}^{l_n-1} \varphi_h^n \Delta_{i+h}^n U \\ &= \frac{-1}{(\psi_0 l_n)^{1/2}} \sum_{h=0}^{l_n-1} (\varphi_{h+1}^n - \varphi_h^n) U_{i+h}^n, \end{aligned} \tag{3.4}$$

and

$$\widehat{U}_i^n(l_n;\varphi) := \frac{1}{2\psi_0 l_n} \sum_{h=0}^{l_n-1} (\varphi_{h+1}^n - \varphi_h^n)^2 \Delta_{i+h}^n U \cdot \Delta_{i+h}^n U^{\mathrm{T}}; \tag{3.5}$$

(3) Compute the pre-averaged values in the local window $I(n, i, k_n)$:

$$\overline{Y}_{i+j}^n(l_n;\varphi) = \frac{1}{(\psi_0 l_n)^{1/2}} \sum_{h=1}^{l_n-1} \varphi_h^n \Delta_{i+j+h}^n Y, \qquad j = 1, \cdots, k_n - l_n + 1. \tag{3.6}$$

Example 1 (pre-averaging via arithmetic average)**.** The pre-averaging method appears in Podolskij and Vetter [2009b], Wang and Mykland [2014] takes a special form corresponding to $\varphi(u) = (u \wedge (1-u))_+$ in (3.6). Note that $\psi_0 = 1/4$. For the ease of exposition, suppose l_n is a even integer, then the following gives us an intuitive understanding of $\overline{Y}_i^n(l_n)$.

$$\begin{aligned} \overline{Y}_i^n(l_n) &= 2 \sum_{k=1}^{l_n/2-1} \frac{k}{l_n} \Delta_{i+k}^n Y + 2 \sum_{k=l_n/2}^{l_n-1} \frac{l_n - k}{l_n} \Delta_{i+k}^n Y \\ &= \underbrace{\frac{2}{l_n} \sum_{k=l_n/2}^{l_n-1} Y_{i+k}^n}_{\widehat{X}_{i+l_n/2}^n} - \underbrace{\frac{2}{l_n} \sum_{k=0}^{l_n/2-1} Y_{i+k}^n}_{\widehat{X}_i^n}. \end{aligned}$$

3.2 Pre-averaging for continuous Itô semimartingales

Define pre-averaged spot volatility estimator for continuous Itô semimartingale model as

$$\widehat{c}_i^{*n}(k_n, l_n; \varphi) = \frac{1}{(k_n - l_n)\Delta_n} \sum_{h=1}^{k_n - l_n + 1} \left[\overline{Y}_{i+h}^{*n}(l_n; \varphi) \cdot \overline{Y}_{i+h}^{*n}(l_n; \varphi)^{\mathrm{T}} - \widehat{Y}_{i+h}^{*n}(l_n; \varphi) \right], \tag{3.7}$$

where $Y^* = X' + \epsilon$, X' is defined in (2.7), and in view of (3.4), (3.5),

$$\begin{aligned} \overline{Y}_i^{*n}(l_n; \varphi) &:= \frac{1}{(\psi_0 l_n)^{1/2}} \sum_{h=1}^{l_n - 1} \varphi_h^n \Delta_{i+h}^n Y^*, \\ \widehat{Y}_i^{*n}(l_n; \varphi) &:= \frac{1}{2\psi_0 l_n} \sum_{h=0}^{l_n - 1} (\varphi_{h+1}^n - \varphi_h^n)^2 \Delta_{i+h}^n Y^* \Delta_{i+h}^n Y^{*\mathrm{T}}. \end{aligned}$$

Though the estimator (3.7) is infeasible to implement if the stream of high-frequency data contains jumps, we need this intermediate quantity in our statistical analysis. Here we give some estimates of the various terms regarding the 1st-order and 2nd-order properties of (3.49). The following analysis before section 4.3 can be skipped in the first reading. These analyses will be called upon in section 4.3.3. In the analysis, for notational ease, we will write $\overline{U}_i^n$ in place of $\overline{U}_i^n(l_n; \varphi)$ for any process U in discussion.

3.2.1 variables to analyze

The Russian playwright Anton Chekhov said “If there is a rifle handing on the wall in act one, it must be fired in the next act. Otherwise it has no business being there”. First of all, define

$$\begin{aligned} \overline{C}_i^n &= \frac{1}{\psi_0 l_n} \sum_{h=1}^{l_n - 1} (\varphi_h^n)^2 \Delta_{i+h}^n C \\ D_i^n &= \overline{C}_i^n - c_i^n \Delta_n \end{aligned}$$

$$
\begin{aligned}
\Gamma_i^n &= \Gamma_{i,i}^n, \quad \Gamma_{h,h'}^n = \frac{1}{\psi_0 l_n} \sum_{v=h\vee h'}^{h\wedge h'+l_n-1} (\varphi_{v-h+1}^n - \varphi_{v-h}^n)(\varphi_{v-h'+1}^n - \varphi_{v-h'}^n)\, \gamma_v^n \\
R_i^n &= \widehat{Y}_i^{*n} - \Gamma_i^n \\
\zeta_i^n &= \overline{Y}_i^{*n}(l_n;\varphi) \cdot \overline{Y}_i^{*n}(l_n;\varphi)^{\mathrm{T}} - \overline{C}_i^n - \Gamma_i^n. \qquad (3.8)
\end{aligned}
$$

The magnitudes of these quantities are influential to the behavior of pre-averaging method. We can interpret $\overline{C}_j^n$ as discretization of volatility due to the kernel, and interpret $\Gamma_{j,k}^n$ as noise contribution to the pre-averaging estimator.

To get a stable central limit theorem, we proceed in the following two steps:

- express the 1st- and 2nd-order properties of $\overline{Y}_i^n$'s in terms of $\overline{X}_i^n$'s and $\Gamma_{j,k}^n$'s;
- analyze the sums of the form $\sum_{l=1}^{pl_n-1} \zeta_{i+j}^n$, i.e. the summation of ζ_i^n's over intervals of length pl_n where p is a fixed positive integer, and separate these sums by intervals of length l_n to ensure independence.[2]

Given $p \in \mathbb{N}^+$, we need to consider more variables. They are

$$
\begin{aligned}
\zeta(W,p)_i^n &= \sum_{h=i}^{i+pl_n-1} \left[(\sigma_i^n \overline{W}_h^n) \cdot (\sigma_i^n \overline{W}_h^n)^{\mathrm{T}} - \overline{C}_h^n \right] \\
\zeta(X,p)_i^n &= \sum_{h=i}^{i+pl_n-1} \left[\overline{X}_h^n \cdot {\overline{X}_h^n}^{\mathrm{T}} - \overline{C}_h^n \right] \\
\zeta(U,p)_i'^n &= \sum_{h=i}^{i+pl_n-2} \sum_{h'=h+1}^{i+pl_n-1} \overline{U}_h^n \cdot {\overline{U}_{h'}^n}^{\mathrm{T}} \phi_1\left(\frac{h'-h}{l_n}\right), \qquad U = W, \sigma_i^n W \text{ or } X
\end{aligned}
$$

2. This step is necessary since the pre-averaging estimator (3.7) involves overlapping intervals such that $\overline{Y}_j^{*n}$ and $\overline{Y}_k^{*n}$ are dependent if $|j-k| < l_n$.

$$\begin{aligned}
\zeta(Y,p)_i^n &= \sum_{h=i}^{i+pl_n-1} \zeta_h^n && (3.9)\\
\lambda(p)_i^n &= \sup_{s,u\in I(n,i,pl_n)} \left(\|b_s-b_u\|+\|c_s-c_u\|+\|\gamma_s-\gamma_u\|\right)\\
\overline{\lambda}(p)_i^n &= \sqrt{E\left[(\lambda(p)_i^n)^2|\mathcal{F}_i^n\right]}.
\end{aligned}$$

Given $p\in\mathbb{N}^+$, let $m(n,p)=\left\lfloor \frac{k_n}{(p+1)l_n}\right\rfloor$, $a(n,p,h)=1+h(p+1)l_n$, $b(n,p,h)=a(n,p,h)+pl_n$. According to (3.7), (3.8) and (3.9), we can decompose the estimation error as

$$\beta(p)_i^n = \widehat{c}_i^{*n} - c_i^n = \xi(0)_i^n + \xi(1)_i^n + \xi(2)_i^n + N(p)_i^n + M(p)_i^n, \tag{3.10}$$

where

$$\begin{aligned}
\xi(0)_i^n &= \frac{1}{k_n-l_n}\sum_{h=1}^{k_n-l_n+1} c_{i+h}^n - c_i^n\\
\xi(1)_i^n &= \frac{1}{(k_n-l_n)\Delta_n}\sum_{h=1}^{k_n-l_n+1} D_{i+h}^n\\
\xi(2)_i^n &= \frac{-1}{(k_n-l_n)\Delta_n}\sum_{h=1}^{k_n-l_n+1} R_{i+h}^n\\
N(p)_i^n &= \frac{1}{(k_n-l_n)\Delta_n}\left(\sum_{h=0}^{m(n,p)-1} \zeta(Y,1)_{i+b(n,p,h)}^n + \sum_{h=m(n,p)(p+1)l_n}^{k_n-l_n} \zeta_{i+1+h}^n\right)\\
M(p)_i^n &= \frac{1}{(k_n-l_n)\Delta_n}\sum_{h=0}^{m(n,p)-1} \zeta(Y,p)_{i+a(n,p,h)}^n. && (3.11)
\end{aligned}$$

Since the pre-averaging method is based on smoothing kernel φ, to express the asymptotic behavior we need to define

$$\Phi_{lm} \equiv \int_0^1 \phi_l(s)\phi_m(s)\,\mathrm{d}s, \quad \Psi_{lm} \equiv \int_0^1 s\,\phi_l(s)\phi_m(s)\,\mathrm{d}s \quad l,m=0,1, \tag{3.12}$$

where ϕ_l, $l=0,1$ are defined in (3.3).

3.2.2 estimates of discretization and noise offset effects

Assuming **Assumption A-c**, **Assumption A-ν**, based on **Corollary 3**, we can already estimate $\xi(0)_j^n$:

$$\left.\begin{array}{rcl}\left\|E\Big(\xi(0)_i^n|\mathcal{F}_i^{(0),n}\Big)\right\| & \leq & Kk_n\Delta_n \\ E\Big(\|\xi(0)_i^n\|^q|\mathcal{F}_i^{(0),n}\Big) & \leq & K_q(k_n\Delta_n)^{(q/2)\wedge 1}, \quad q\geq 0\end{array}\right\}. \tag{3.13}$$

Lemma 7. *Under* ***Assumption A-**c *and* ***Assumption A-**ν*,*

$$\begin{array}{rcl}\left\|E\left(\xi(1)_i^n|\mathcal{F}_i^{(0),n}\right)\right\| & \leq & K\left(l_n\Delta_n+l_n^{-1}\right) \\ E\left(\|\xi(1)_i^n\|^q|\mathcal{F}_i^{(0),n}\right) & \leq & K_q\left[(l_n\Delta_n)^{(q/2)\wedge 1}+l_n^{-q}\right],\ q\in\mathbb{N}^+.\end{array}$$

Proof. Note that

$$\begin{aligned}D_i^n &= \frac{1}{\psi_0 l_n}\sum_{h=1}^{l_n-1}(\varphi_h^n)^2\left(\Delta_{i+h}^n C-c_i^n\Delta_n\right)+\frac{c_i^n\Delta_n}{\psi_0}\left[\sum_{h=1}^{l_n-1}(\varphi_h^n)^2\frac{1}{l_n}-\psi_0\right] \\ &= \frac{\Delta_n}{\psi_0 l_n}\sum_{h=1}^{l_n-1}(\varphi_h^n)^2\left(c_{i+h-1}^n-c_i^n\right)+\frac{1}{\psi_0 l_n}\sum_{h=1}^{l_n-1}(\varphi_h^n)^2\left(\Delta_{i+h}^n C-c_{i+h-1}^n\Delta_n\right)+O_p\left(\frac{\Delta_n}{l_n}\right).\end{aligned}$$

(i) Since the c is locally bounded, by Fubini's theorem and **Corollary 3**,

$$\begin{array}{rl}\left\|E\left(c_{i+h-1}^n-c_i^n|\mathcal{F}_i^{(0),n}\right)\right\| & \leq Kh\Delta_n \\ \left\|E\left(\Delta_{i+h}^n C-c_{i+h-1}^n\Delta_n|\mathcal{F}_i^{(0),n}\right)\right\| & \leq \int_{I(n,i+h-1)}\left\|E\left(c_s-c_{i+h-1}^n|\mathcal{F}_i^{(0),n}\right)\right\|\mathrm{d}s\leq K\Delta_n^2.\end{array}$$

(ii) By **Corollary 3** again,

$$\begin{array}{rcl}E\left(\|c_{i+h-1}^n-c_i^n\|^q|\mathcal{F}_i^{(0),n}\right) & \leq & K_q\left(h\Delta_n\right)^{(q/2)\wedge 1} \\ E\left(\left\|\Delta_{i+h}^n C-c_{i+h-1}^n\Delta_n\right\|^q|\mathcal{F}_i^{(0),n}\right) & \leq & K_q\,\Delta_n^{(\frac{q}{2}\wedge 1)+q}.\end{array} \tag{3.14}$$

(iii) By Hölder's inequality, for $q \in \mathbb{N}^+$,

$$E\left(\left\|\prod_{l=1}^{q} D_{i+h_l}^n\right\| |\mathcal{F}_i^{(0),n}\right) \leq \prod_{l=1}^{q} E\left(\|D_{i+h_l}^n\|^q |\mathcal{F}_i^{(0),n}\right)^{1/q}.$$

Combine (i),(ii) with Hölder's inequality, we have

$$\left.\begin{array}{rcl} \left\|E\left(D_i^n|\mathcal{F}_i^{(0),n}\right)\right\| & \leq & K\left(l_n\Delta_n^2 + l_n^{-1}\Delta_n\right) \\ E\left(\|D_i^n\|^q|\mathcal{F}_i^{(0),n}\right) & \leq & K_q\Big(l_n^{(q/2)\wedge 1}\Delta_n^{[(q/2)\wedge 1]+q} + l_n^{-q}\Delta_n^q\Big) \end{array}\right\}, \tag{3.15}$$

use (iii) and consider $q \in \mathbb{N}^+$,

$$\|\xi(1)_i^n\|^q \leq \frac{1}{(k_n - l_n)^q \Delta_n^q} \sum_{h_1=1}^{k_n-l_n+1} \cdots \sum_{h_q=1}^{k_n-l_n+1} E\left(\left\|\prod_{l=1}^{q} D_{i+h_l}^n\right\| |\mathcal{F}_i^{(0),n}\right),$$

we prove this lemma. □

Lemma 8. *Under* ***Assumption A-***γ, ***Assumption A-***ν

$$\|E\left(\xi(2)_i^n|\mathcal{F}_i^n\right)\| \leq K l_n^{-2};$$

$$E\left(\|\xi(2)_i^n\|^q|\mathcal{F}_i^n\right) \leq \begin{cases} K\Big(k_n^{-1/2}l_n^{-2}\Delta_n^{-1} + k_n^{-1}l_n^{-1}\Delta_n^{-1/2}\Big), & q = 1; \\ K_q\, k_n^{-q}l_n^{-q}\Delta_n^{-q}\left(k_n l_n^{-q} + \Delta_n\right), & q \in \mathbb{N}^+, q \geq 2. \end{cases}$$

Proof. we can write

$$\xi(2)_i^n = \frac{-1}{2\psi_0(k_n - l_n)\Delta_n}(F_i^n - G_i^n),$$

where

$$\begin{array}{rl} F_i^n = & \sum_{h=1}^{k_n} \eta_{i+h}^n \times \psi_1^{n,h}, \quad \eta_i^n = \Delta_i^n Y \cdot \Delta_i^n Y^{\mathrm{T}} - \gamma_{i-1}^n - \gamma_i^n \\ G_i^n = & \sum_{h=1}^{k_n} \Delta_{i+h}^n \gamma \times \psi_1^{n,h}, \end{array}$$

and $\psi_1^{n,h} = \frac{1}{l_n}\sum_{h'=0\vee(h-k_n+l_n-1)}^{(h\wedge l_n)-1}(\varphi_{h'+1}^n - \varphi_{h'}^n)^2$. Note that

$$\left.\begin{array}{lcl}\|E(\eta_i^n|\mathcal{F}_{i-1}^n)\| & = & E\left(\|\Delta_i^n X\|^2|\mathcal{F}_{i-1}^n\right) \leq K\Delta_n \\ E\left(\|\eta_i^n\|^q|\mathcal{F}_{i-1}^n\right) & \leq & K_q \\ \mathbb{E}\left(\left\|\prod_{l=1}^q \eta_{h_l}^n\right\|\right) & = & \mathbb{E}\left(\prod_{l=1}^q \|\Delta_{h_l}^n X\|^2\right) \leq K_q\Delta_n^q, \quad \min_{l,l'}|h_l - h_{l'}| \geq 2\end{array}\right\}, \tag{3.16}$$

thus

$$\left.\begin{array}{ll}\left\|E\left(F_i^n|\mathcal{F}_i^n\right)\right\| & \leq K k_n l_n^{-2}\Delta_n \\ E\left(\|F_i^n\|^q|\mathcal{F}_i^n\right) & \leq K_q k_n l_n^{-2q},\ q\in\mathbb{N}^+\end{array}\right\}.$$

Besides

$$G_i^n = \sum_{h'=1}^{l_n-1}\Delta_{i+h'}^n\gamma\frac{1}{l_n}\sum_{h=0}^{h'-1}(\varphi_{h+1}^n-\varphi_h^n)^2 + \sum_{h'=k_n-l_n+2}^{k_n}\Delta_{i+h'}^n\gamma\frac{1}{l_n}\sum_{h=h'-k_n+l_n-1}^{l_n-1}(\varphi_{h+1}^n-\varphi_h^n)^2$$
$$+ (\gamma_{i+k_n-l_n+1}^n - \gamma_{i+l_n-1}^n)\frac{1}{l_n}\sum_{h=0}^{l_n-1}(\varphi_{h+1}^n-\varphi_h^n)^2,$$

by **Corollary 3**, $\|E(\gamma_{i+p}^n-\gamma_i^n|\mathcal{F}_i^n)\|\leq Kp\Delta_n$, $E\left(\|\gamma_{i+p}^n-\gamma_i^n\|^q|\mathcal{F}_i^n\right)\leq K(p\Delta_n)^{(q/2)\wedge 1}$, thus

$$\left.\begin{array}{ll}\left\|E\left(G_i^n|\mathcal{F}_i^n\right)\right\| & \leq K\,k_n l_n^{-2}\Delta_n \\ E\left(\|G_i^n\|^q|\mathcal{F}_i^n\right) & \leq K_q\left(k_n^{(q/2)\wedge 1}l_n^{-2q}+l_n^{-q}\right)\Delta_n^{(q/2)\wedge 1}\end{array}\right\}.$$

By these intermediate estimates of F_i^n and G_i^n with Hölder's inequality and Jensen's inequality, this lemma is proved. □

3.2.3 estimates of Brownian motion functionals

Given any function f, define $f_n(t) = \sum_{h=1}^{l_n-1} f_h^n \mathbf{1}_{((h-1)\Delta_n, h\Delta_n]}(t)$; for a generic process U, define

$$U_t^n(f,s) = \int_t^{t+s} f_n(u-t)\,\mathrm{d}U_u. \tag{3.17}$$

We have

$$\bar{X}_i^n = \frac{1}{\psi_0^{1/2} l_n^{1/2}} X_{i\Delta_n}^n(\varphi, (l_n - 1)\Delta_n), \qquad \bar{C}_i^n = \frac{1}{\psi_0 l_n} C_{i\Delta_n}^n(\varphi^2, (l_n - 1)\Delta_n), \tag{3.18}$$

hence by **Assumption SA-ν**,

$$\|\bar{C}_i^n\| \leq K\Delta_n. \tag{3.19}$$

We need a lemma to facilitate estimates of (3.9).

Lemma 9. *For $p \in \mathbb{N}^+$, and $l, m = 0, 1$,*

$$\sum_{h=i}^{i+pl_n-2} \sum_{h'=h+1}^{i+pl_n-1} \phi_l\left(\frac{h'-h}{l_n}\right) \phi_m\left(\frac{h'-h}{l_n}\right) = l_n^2\left(p\Phi_{lm} - \Psi_{lm}\right) + O(pl_n).$$

Proof.

$$\begin{aligned}
&\sum_{h=i}^{i+pl_n-2} \sum_{h'=h+1}^{i+pl_n-1} \phi_l\left(\frac{h'-h}{l_n}\right) \phi_m\left(\frac{h'-h}{l_n}\right) \\
&= \left(\sum_{h=i}^{i+(p-1)l_n} l_n \sum_{h'=h+1}^{h+l_n-1} + \sum_{h=i+(p-1)l_n+1}^{i+pl_n-2} l_n \sum_{h'=h+1}^{i+pl_n-1}\right) \phi_l\left(\frac{h'-h}{l_n}\right) \phi_m\left(\frac{h'-h}{l_n}\right) \frac{1}{l_n} \\
&\qquad = (p-1)l_n^2\,\Phi_{lm} + O(pl_n) + l_n^2 \sum_{h=1}^{l_n-1} \frac{l_n - h}{l_n} \phi_l\left(h/l_n\right) \phi_m\left(h/l_n\right) \frac{1}{l_n} \\
&\qquad\qquad = \left(p\Phi_{lm} - \Psi_{lm}\right) l_n^2 + O(pl_n).
\end{aligned}$$

□

Lemma 10. *Under **Assumption A-**c, **Assumption A-**ν,*

$$
\begin{aligned}
E\left(\|\zeta(W,p)_i^n\|^4|\mathcal{F}_i^{(0),n}\right) &\leq Kp^4 l_n^4\Delta_n^4 \\
E\left[\zeta(W,p)_i'^n|\mathcal{F}_i^{(0),n}\right] &= \frac{1}{\psi_0}(p\Phi_{01}-\Psi_{01})\,\mathbf{I}_d\, l_n^2\Delta_n + p\, O_p\left(l_n^{3/2}\Delta_n\right) \\
E\left[\zeta(W,p)_i^{n,jk}\zeta(W,p)_i^{n,lm}|\mathcal{F}_i^{(0),n}\right] &= \frac{2}{\psi_0^2}(p\Phi_{00}-\Psi_{00})(c_i^{n,jl}c_i^{n,km}+c_i^{n,jm}c_i^{n,kl})\, l_n^2\Delta_n^2 \\
&\quad + p^2\, O_p\left[l_n^{3/2}\Delta_n^2\left(1+l_n^{1/2}\overline{\lambda}(p)_i^n\right)\right].
\end{aligned}
$$

Proof. By the scaling property of Brownian motion,

$$
\bar{W}_h^n \overset{\mathcal{L}}{=} -\frac{\Delta_n^{1/2}}{\psi_0^{1/2}}\sum_{h'=0}^{l_n-1}(\varphi_{h'+1}^n-\varphi_{h'}^n)W_{(h+h')/l_n},
$$

so

$$
\bar{W}_h^n \overset{\mathcal{L}}{=} \frac{\Delta_n^{1/2}}{\psi_0^{1/2}}U_{h/l_n} + e_h^n, \tag{3.20}
$$

where $U_t = -\int_0^1 \varphi'(u)W_{t+u}\,\mathrm{d}u$ and $e_h^n = \frac{\Delta_n^{1/2}}{\psi_0^{1/2}}\sum_{h'=0}^{l_n-1}\int_{(h+h')/l_n}^{(h+h'+1)/l_n}\varphi'(u)(W_u - W_{(h+h')/l_n})\,\mathrm{d}u$. Note that $\mathbb{E}(\|e_h^n\|^q)\leq K_q(\Delta_n/l_n)^{q/2}$. Besides $U_t=\int_0^1\varphi(u)\,\mathrm{d}W_{u+t}$, so U_t is a Gaussian random vector with independent entries, so $\mathbb{E}(U_tU_{t+s}^{\mathrm{T}})=\phi_0(s)\mathbf{I}_d$ by Itô isometry, hence for $h,h'\geq i$,

$$
\begin{aligned}
E\left[\bar{W}_h^n{\bar{W}_{h'}^n}^{\mathrm{T}}|\mathcal{F}_i^{(0),n}\right] &= \frac{\Delta_n}{\psi_0}\phi_0\left(\frac{|h'-h|}{l_n}\right)\mathbf{I}_d + O_p(l_n^{-1/2}\Delta_n) \\
E\left[\|\bar{W}_h^n\|^{2m}|\mathcal{F}_i^{(0),n}\right] &= \Delta_n^m(2m-1)!! + O_p(l_n^{-1}\Delta_n^m),\ m\in\mathbb{N}^+.
\end{aligned} \tag{3.21}
$$

(1) Based on **Assumption SA-**ν, (3.19) and (3.21)

$$
\begin{aligned}
E\left(\|\zeta(W,p)_i^n\|^4|\mathcal{F}_i^{(0),n}\right) &= p^4 l_n^4\, O_p\left(\max_{i\leq h\leq i+pl_n}\left[\mathbb{E}(\|\bar{W}_h^n\|^8)+\mathbb{E}(\|\bar{W}_h^n\|^6\|\bar{C}_h^n\|)\right.\right. \\
&\quad \left.\left. + \mathbb{E}(\|\bar{W}_h^n\|^4\|\bar{C}_h^n\|^2) + \mathbb{E}(\|\bar{W}_h^n\|^2\|\bar{C}_h^n\|^3)+\mathbb{E}(\|\bar{C}_h^n\|^4)\right]\right) \leq Kp^4 l_n^4\Delta_n^4,
\end{aligned}
$$

hence the inequality is confirmed.

(2) By (3.21),

$$E\big[\zeta(W,p)_i'^n|\mathcal{F}_i^{(0),n}\big] = \frac{\Delta_n}{\psi_0}\sum_{h=i}^{i+pl_n-2}\sum_{h'=h+1}^{i+pl_n-1}\phi_0\left(\frac{h'-h}{l_n}\right)\phi_1\left(\frac{h'-h}{l_n}\right)\mathbf{I}_d + p\,O_p(l_n^{3/2}\Delta_n),$$

according to **Lemma 9**, the first equality is proved.

(3) Let $U_i^n(p) = \sum_{h=i}^{i+pl_n-1}(\sigma_i^n\bar{W}_h^n)(\sigma_i^n\bar{W}_h^n)^{\mathrm{T}}$, $S_i^n(p) = \sum_{h=i}^{i+pl_n-1}\bar{C}_h^n$, then

$$\begin{aligned}\zeta(W,p)_i^{n,jk}\zeta(W,p)_i^{n,lm} &= U_i^n(p)^{jk}U_i^n(p)^{lm} + S_i^n(p)^{jk}S_i^n(p)^{lm} \\ &\quad - U_i^n(p)^{jk}S_i^n(p)^{lm} - U_i^n(p)^{lm}S_i^n(p)^{jk}. \end{aligned} \tag{3.22}$$

(i) By the decomposition of D_i^n in the proof of **Lemma 7**, $\bar{C}_h^n = c_i^n\Delta_n + O_p(l_n^{-1}\Delta_n + \Delta_n\lambda(p)_i^n)$ for $i < h < i+pl_n$, hence

$$S_i^n(p) = c_i^n\, pl_n\Delta_n + p\,O_p\big[\Delta_n + l_n\Delta_n\lambda(p)_i^n\big]. \tag{3.23}$$

(ii) Moreover, by (3.20),

$$U_i^n(p) \stackrel{\mathcal{L}}{=} \frac{\Delta_n}{\psi_0}\sum_{h=i}^{i+pl_n-1}(\sigma_i^n U_{h/l_n})(\sigma_i^n U_{h/l_n})^{\mathrm{T}} + e_i^n(p),$$

where $\mathbb{E}(\|e_i^n(p)\|^q) \le K_q\, p^q\, l_n^{q/2}\Delta_n^q$. Hence by (3.21)

$$E\big[U_i^n(p)|\mathcal{F}_i^{(0),n}\big] = c_i^n\, pl_n\Delta_n + p\,O_p(l_n^{1/2}\Delta_n). \tag{3.24}$$

(iii) Furthermore, let $\xi_{h,h'}^{n,jk,lm} = (\sigma_i^{n,j\cdot}U_{h/l_n})(\sigma_i^{n,k\cdot}U_{h/l_n})^{\mathrm{T}}(\sigma_i^{n,l\cdot}U_{h'/l_n})(\sigma_i^{n,m\cdot}U_{h'/l_n})^{\mathrm{T}}$ where

$\sigma_i^{n,j\cdot}$ is the j-th row of σ_i^n, then

$$E\big[U_i^n(p)^{jk}U_i^n(p)^{lm}|\mathcal{F}_i^{(0),n}\big] = \frac{\Delta_n^2}{\psi_0^2}\sum_{h=i}^{i+pl_n-1}\sum_{h'=i}^{i+pl_n-1}\mathbb{E}\big(\xi_{h,h'}^{n,jk,lm}\big) + p^2 O_p(l_n^{3/2}\Delta_n^2).$$

Let $U_{h,h'}^n = U_{h/l_n} - U_{h'/l_n}\phi_0(|h'-h|/l_n)/\psi_0$, using the fact that $U_{h,h'}^n$ and U_{h'/l_n} are independent,

$$\begin{aligned}\mathbb{E}\big(\xi_{h,h'}^{n,jk,lm}\big) = \mathbb{E}\big[(\sigma_i^{n,j\cdot}U_{h,h'}^n)(\sigma_i^{n,k\cdot}U_{h,h'}^n)^{\mathrm{T}}\big]\cdot\mathbb{E}\big[(\sigma_i^{n,l\cdot}U_{h'/l_n})(\sigma_i^{n,m\cdot}U_{h'/l_n})^{\mathrm{T}}\big]& \\ + \frac{\phi_0\big(\frac{|h-h'|}{l_n}\big)^2}{\psi_0^2}\mathbb{E}\big(\xi_{h',h'}^{n,jk,lm}\big).&\end{aligned}$$

Note $\mathbb{E}\big(U_{h,h'}^n U_{h,h'}^n{}^{\mathrm{T}}\big) = (\psi_0 - \phi_0^2/\psi_0)\mathbf{I}_d$, $\mathbb{E}\big(U_{h'/l_n}U_{h'/l_n}{}^{\mathrm{T}}\big) = \psi_0\mathbf{I}_d$; based on Gaussian moments and the representation $\xi_{h',h'}^{n,jk,lm} = \prod_{v=j,k,l,m}\sum_{r=1}^{d'}\sigma_i^{n,vr}U_{h'/l_n}^r$,

$$\mathbb{E}\big(\xi_{h',h'}^{n,jk,lm}\big) = \psi_0^2\big(c_i^{n,jk}c_i^{n,lm} + c_i^{n,jl}c_i^{n,km} + c_i^{n,jm}c_i^{n,kl}\big),$$

so

$$\mathbb{E}\big(\xi_{h,h'}^{n,jk,lm}\big) = \psi_0^2 c_i^{n,jk}c_i^{n,lm} + \phi_0\left(\frac{|h-h'|}{l_n}\right)^2\big(c_i^{n,jl}c_i^{n,km} + c_i^{n,jm}c_i^{n,kl}\big).$$

Apply **Lemma 9** to these intermediate results in (iii),

$$\begin{aligned}E\big[U_i^n(p)^{jk}U_i^n(p)^{lm}|\mathcal{F}_i^{(0),n}\big] &= c_i^{n,jk}c_i^{n,lm}\,p^2 l_n^2\Delta_n^2 \\ &+ \frac{2}{\psi_0^2}\,(p\Phi_{00} - \Psi_{00})\,\big(c_i^{n,jl}c_i^{n,km} + c_i^{n,jm}c_i^{n,kl}\big)l_n^2\Delta_n^2 + p^2\,O_p(l_n^{3/2}\Delta_n^2). \quad (3.25)\end{aligned}$$

Combine (3.22), (3.23), (3.24) and (3.25), we can prove the second equality. □

Because φ is piecewise C^1, $|\varphi^n_{h+1} - \varphi^n_h| \le K/l_n$.

Since the smoothing kernel satisfies $\varphi(0) = \varphi(1) = 0$, $\sum_{h=0}^{l_n-1}(\varphi^n_{h+1} - \varphi^n_h) = 0$, hence $\bar{X}^n_i = -\sum_{h=0}^{l_n-1}(\varphi^n_{h+1} - \varphi^n_h)(X^n_{i+h} - X^n_i)/(\psi_0 l_n)^{1/2}$. By **Corollary 3**,

$$E\left(\|\bar{X}^n_i\|^q|\mathcal{F}^{(0),n}_i\right) \le K_q\Delta_n^{q/2}. \tag{3.26}$$

Lemma 11. *Under* ***Assumption A****-c,* ***Assumption A****-ν,*

$$\begin{aligned}
E\left(\|\zeta(X,p)^n_i\|^4|\mathcal{F}^{(0),n}_i\right) &\le Kp^4 l_n^4\Delta_n^4 \\
\left\|E\left[\zeta(X,p)^n_i|\mathcal{F}^{(0),n}_i\right]\right\| &\le Kp\left(l_n^2\Delta_n^2 + l_n^{3/2}\Delta_n^{3/2}\,\overline{\lambda}(p)^n_i\right) \\
E\left[\zeta(X,p)'^n_i|\mathcal{F}^{(0),n}_i\right] &= \frac{1}{\psi_0}(p\Phi_{01} - \Psi_{01})\,c^n_i\,l_n^2\Delta_n \\
&\quad + p^2\,O_p\left(l_n^2\Delta_n\left[l_n^{1/2}\Delta_n^{1/2} + \overline{\lambda}(p)^n_i\right]\right) \\
E\left[\zeta(X,p)^{n,jk}_i\zeta(X,p)^{n,lm}_i|\mathcal{F}^{(0),n}_i\right] &= \frac{2}{\psi_0^2}(p\Phi_{00} - \Psi_{00})\,(c^{n,jl}_i c^{n,km}_i + c^{n,jm}_i c^{n,kl}_i)\,l_n^2\Delta_n^2 \\
&\quad + p^2\,O_p\left(l_n^2\Delta_n^2\left[l_n^{-1/2} + l_n^{1/2}\Delta_n^{1/2} + \overline{\lambda}(p)^n_i\right]\right).
\end{aligned}$$

Proof. We prove these estimates case by case.

(1) Based on (3.19) and (3.26), a similar calculation as part (1) in the proof for **Lemma 10** yields the result.

(2) Recall (3.18), by **Corollary 3**,

$$\left\|E\left[X^n_t(\varphi,s)|\mathcal{F}^{(0)}_t\right]\right\| \le Ks \tag{3.27}$$

$$E\left(\sup_{t>0}\|X^n_t(\varphi,s)\|^q|\mathcal{F}^{(0)}_t\right) \le K_q\,s^{q/2}. \tag{3.28}$$

Let $Z_s \equiv \int_0^s \varphi_n(u-t)\,\mathrm{d}X_u$, we can see that $Z_s = 0$ if $s \le t$ and $X^n_t(\varphi,s) = Z_{t+s}$. According to Itô's formula, $Z^j_{t+s}Z^k_{t+s} = \int_t^{t+s} Z^j_u \, \mathrm{d}Z^k_u + \int_t^{t+s} Z^k_u \, \mathrm{d}Z^j_u + \int_t^{t+s} \varphi_n(u-t)^2 \, \mathrm{d}C^{jk}_u$, so

$$X^n_t(\varphi, s) \cdot X^n_t(\varphi, s)^{\mathrm{T}} = B(n,t,s) + M(n,t,s) + C^n_t(\varphi^2, s), \tag{3.29}$$

where

$$\begin{aligned} B(n,t,s)^{jk} &= \int_t^{t+s} \left[X^{n,j}_{u-s}(\varphi,s)\, b^k_u + X^{n,k}_{u-s}(\varphi,s)\, b^j_u\right] \varphi_n(u-t)\, \mathrm{d}u \\ M(n,t,s)^{jk} &= \int_t^{t+s} \left[X^{n,j}_{u-s}(\varphi,s)\, \sigma^{k\cdot}_u + X^{n,k}_{u-s}(\varphi,s)\, \sigma^{j\cdot}_u\right] \varphi_n(u-t)\, \mathrm{d}W_u. \end{aligned}$$

Since $E\big[M(n,t,s)|\mathcal{F}^{(0)}_t\big] = 0$, use (3.29) we get

$$E\left[\zeta(X,p)^n_i|\mathcal{F}^{(0),n}_i\right] = \frac{1}{\psi_0 l_n} \sum_{h=i}^{i+pl_n-1} E\left[B(n, h\Delta_n, (l_n-1)\Delta_n)|\mathcal{F}^n_i\right]. \tag{3.30}$$

Write $B(n, h\Delta_n, (l_n-1)\Delta_n) = B^n_h(1) + B^n_h(2)$, where

$$\begin{aligned} B^n_h(1)^{jk} &= \int_{I(n,h,l_n-1)} \big[X^{n,j}_{u-(l_n-1)\Delta_n}(\varphi,(l_n-1)\Delta_n) b^{n,k}_h \\ &\qquad + X^{n,k}_{u-(l_n-1)\Delta_n}(\varphi,(l_n-1)\Delta_n) b^{n,j}_h\big] \varphi_n(u-h\Delta_n)\, \mathrm{d}u \\ B^n_h(2)^{jk} &= \int_{I(n,h,l_n-1)} \big[X^{n,j}_{u-(l_n-1)\Delta_n}(\varphi,(l_n-1)\Delta_n)(b^k_u - b^{n,k}_h) \\ &\qquad + X^{n,k}_{u-(l_n-1)\Delta_n}(\varphi,(l_n-1)\Delta_n)(b^j_u - b^{n,j}_h)\big] \varphi_n(u-h\Delta_n)\, \mathrm{d}u. \end{aligned}$$

On one hand, by (3.27), $\big\|E\big[B^n_h(1)|\mathcal{F}^{(0),n}_h\big]\big\| \leq K l_n^2 \Delta_n^2$; on the other hand, by Cauchy-Schwarz inequality and (3.28), and $\max_{i \leq h \leq i+l_n-1} \overline{\lambda}(1)^n_h \leq \overline{\lambda}(p)^n_h$,

$$\begin{aligned} &\Big\|E\big[B^n_h(2)|\mathcal{F}^{(0),n}_h\big]\Big\| \\ &\quad \leq l_n \Delta_n \overline{\lambda}(1)^n_h \cdot E\left(\sup_{u \in I(n,h,l_n-1)} \|X^n_{u-(l_n-1)\Delta_n}(\varphi,(l_n-1)\Delta_n)\|^2 |\mathcal{F}^{(n),n}_h\right)^{1/2} \\ &\quad \leq K l_n^{3/2} \Delta_n^{3/2}\, \overline{\lambda}(p)^n_i, \end{aligned}$$

therefore this case is checked.

(3)&(4) These 2 cases can be justified in a similar fashion.

(i) Note for $u > v > t$, $(X_u - X_v) - \sigma_t(W_u - W_v) = \int_v^u b_s\,\mathrm{d}s + \int_v^u(\sigma_s - \sigma_t)\,\mathrm{d}W_s$, we have for $i\Delta_n \le t < t+s \le (i+pl_n)\Delta_n$,

$$E\left(\sup_{u,v\in(t,t+s]} \|(X_u - X_v) - \sigma_t(W_u - W_v)\|^q \Big| \mathcal{F}_i^{(0),n}\right) \le K_{p,q}\left[s^q + s^{q/2}(\bar{\lambda}(p)_i^n)^q\right],$$

hence for $i \le h < h' \le i + pl_n$

$$E\left(\|(X_{h'}^n - X_h^n) - \sigma_i^n(W_{h'}^n - W_h^n)\|^q|\mathcal{F}_i^{(0),n}\right) \le K_q\,p^q\left[(l_n\Delta_n)^q + (l_n\Delta_n)^{q/2}(\bar{\lambda}(p)_i^n)^q\right]. \tag{3.31}$$

(ii) For $m \in \mathbb{N}^+$, $a^m - b^m = (a-b)\sum_{l=0}^{m-1} a^l b^{m-1-l}$, so

$$\left\|(\bar{X}_h^{n,j})^m - (\sigma_i^{n,j\cdot}\bar{W}_h^n)^m\right\| \le Km\Delta_n^{(m-1)/2}\left\|\bar{X}_h^n - \sigma_i^n\bar{W}_h^n\right\|. \tag{3.32}$$

(iii) By (3.4) and $\sum_{h=0}^{l_n-1}(\varphi_{h+1}^n - \varphi_h^n) = 0$, $\bar{X}_h^n - \sigma_i^n\bar{W}_h^n = -(\psi_0 l_n)^{-1/2}\sum_{l=0}^{l_n-1}(\varphi_{l+1}^n - \varphi_l^n)[(X_{h+l}^n - X_h^n) - \sigma_i^n(W_{h+l}^n - W_h^n)]$, combined with (3.31), (3.32) and $|\varphi_{h+1}^n - \varphi^n| < K/l_n$,

$$E\left(\left\|(\bar{X}_h^{n,j})^m - (\sigma_i^{n,j\cdot}\bar{W}_h^n)^m\right\|^q|\mathcal{F}_i^{(0),n}\right) \le K_q\,p^q\,m^q\Delta_n^{qm/2}\left[(l_n\Delta_n)^{q/2} + (\bar{\lambda}(p)_i^n)^q\right]. \tag{3.33}$$

Now we are ready to check the last 2 cases. Let $\xi_{i,h,h'}^n = \bar{X}_h^n\bar{X}_{h'}^{n\,\mathrm{T}} - (\sigma_i^n\bar{W}_h^n)(\sigma_i^n\bar{W}_{h'}^n)^{\mathrm{T}}$.

- $\zeta(X,p)_i'^n - \zeta(\sigma_i^n W,p)_i'^n = \sum_{h=i}^{i+pl_n-2}\sum_{h'=h+1}^{i+pl_n-1}\xi_{i,h,h'}^n\phi_1\left(\frac{h'-h}{l_n}\right)$.

  Note $\xi_{i,h,h'}^n = \bar{X}_h^n(\bar{X}_{h'}^n - \sigma_i^n\bar{W}_{h'}^n)^{\mathrm{T}} + \sigma_i^n\bar{W}_{h'}^n(\bar{X}_h^n - \sigma_i^n\bar{W}_h^n)^{\mathrm{T}}$, by Cauchy-Schwarz inequality, **Assumption SA-ν**, (3.21), (3.26) and (3.33) with $m=1, q=2$,

$$
\begin{aligned}
E\left(\left\|\zeta(X,p)_i^{\prime n,jk} - c_i^{n,jk}\zeta(W,p)_i^{\prime n,jk}\right\| |\mathcal{F}_i^{(0),n}\right) \\
&\leq Kp^2 l_n^2 \sup_{i\leq h\leq i+pl_n} E\left(\|\bar{X}_h^n\|^2 \vee \|\bar{W}_h^n\|^2 |\mathcal{F}_i^{(0),n}\right)^{1/2} \\
&\quad \times \sup_{i\leq h\leq i+pl_n} E\left(\|\bar{X}_h^n - \sigma_i^n \bar{W}_h^n\|^2 |\mathcal{F}_i^{(0),n}\right)^{1/2} \\
&\leq Kp^2 l_n^2 \Delta_n \left[(l_n\Delta_n)^{1/2} + \overline{\lambda}(p)_i^n\right].
\end{aligned}
$$

- $\zeta(X,p)_i^{n,jk}\zeta(X,p)_i^{n,lm} - \zeta(W,p)_i^{n,jk}\zeta(W,p)_i^{n,lm} = [\zeta(X,p)_i^{n,jk} - \zeta(W,p)_i^{n,jk}]\zeta(W,p)_i^{n,lm} + \zeta(X,p)_i^{n,jk}[\zeta(X,p)_i^{n,lm} - \zeta(W,p)_i^{n,lm}]$. Note $\zeta(X,p)_i^n - \zeta(W,p)_i^n = \sum_{h=i}^{i+pl_n-1}\xi_{i,h,h}^n$, Cauchy-Schwarz inequality and (3.33) with $m=2, q=2$ yields

$$
E\left(\|\zeta(X,p)_i^n - \zeta(W,p)_i^n\|^2 |\mathcal{F}_i^{(0),n}\right) \leq Kp^2 l_n^2 \Delta_n^2 \left[(l_n\Delta_n) + (\overline{\lambda}(p)_i^n)^2\right],
$$

apply Jensen's inequality and Cauchy-Schwarz inequality,

$$
\begin{aligned}
E&\left[\zeta(X,p)_i^{n,jk}\zeta(X,p)_i^{n,lm} - \zeta(W,p)_i^{n,jk}\zeta(W,p)_i^{n,lm}|\mathcal{F}_i^{(0),n}\right] \\
&\leq E\left[\|\zeta(X,p)_i^n - \zeta(W,p)_i^n\|^2|\mathcal{F}_i^{(0),n}\right]^{1/2} \cdot E\left[\|\zeta(X,p)_i^n + \zeta(W,p)_i^n\|^2\|\mathcal{F}_i^{(0),n}\right]^{1/2} \\
&\leq Kpl_n\Delta_n\left[(l_n\Delta_n)^{1/2} + \overline{\lambda}(p)_i^n\right] \cdot E\left(\|\zeta(X,p)_i^n + \zeta(W,p)_i^n\|^4\right)^{1/4},
\end{aligned}
$$

based on estimates of $E\big[\|\zeta(W,p)_i^n\|^4|\mathcal{F}_i^{(0),n}\big]$ and $E\big[\|\zeta(X,p)_i^n\|^4|\mathcal{F}_i^{(0),n}\big]$, this case is checked.

$\square$

Lemma 12. *Under **Assumption A-**γ**, A-**ν,*

$$i \leq h, h' \leq i + pl_n - 1 \Rightarrow \quad \Gamma^n_{h,h'} = \frac{1}{\psi_0 l_n^2}\, \phi_1\Big(\frac{|h'-h|}{l_n}\Big)\gamma^n_i + O_p\big[l_n^{-3} + l_n^{-2}\lambda(p)^n_i\big],$$

and

$$\sum_{h=i}^{i+pl_n-2} \sum_{h'=h+1}^{i+pl_n-1} \Gamma^{n,jk}_{h,h'}\Gamma^{n,lm}_{h,h'} = \frac{1}{\psi_0^2 l_n^2}\,(p\Phi_{11} - \Psi_{11})\,\gamma^{n,jk}_i\gamma^{n,lm}_i + p^2\, O_p\big[l_n^{-3} + l_n^{-2}\lambda(p)^n_i\big].$$

Proof. Firstly, note that $|\varphi^n_{l+1} - \varphi^n_l| \leq K/l_n$ and $\|\gamma^n_l - \gamma^n_i\| \leq \lambda(p)^n_i$ for $i \leq l \leq i+(p+1)l_n$, so for $i \leq h \leq h' \leq i + (p+1)l_n$, by **Assumption SA-**$\nu$,

$$\begin{aligned}
\Gamma^n_{h,h'} &= \frac{1}{\psi_0 l_n} \sum_{v=h'}^{h+l_n-1} (\varphi^n_{v-h+1} - \varphi^n_{v-h})(\varphi^n_{v-h'+1} - \varphi^n_{v-h'})[\gamma^n_i + (\gamma^n_v - \gamma^n_i)] \\
&= \frac{1}{\psi_0 l_n^2}\phi_1\left(\frac{h'-h}{l_n}\right)\gamma^n_i + O_p\big[l_n^{-3} + l_n^{-2}\lambda(p)^n_i\big],
\end{aligned}$$

so $\Gamma^{n,jk}_{h,h'}\Gamma^{n,lm}_{h,h'} = \psi_0^{-2} l_n^{-4}\phi_1^2\left(\frac{h'-h}{l_n}\right)\gamma^{n,jk}_i\gamma^{n,lm}_i + O_p\left(l_n^{-5} + l_n^{-4}\lambda(p)^n_i\right)$, and

$$\begin{aligned}
\sum_{h=i}^{i+pl_n-2} \sum_{h'=h+1}^{i+pl_n-1} \Gamma^{n,jk}_{h,h'}\Gamma^{n,lm}_{h,h'} = \frac{\gamma^{n,jk}_i\gamma^{n,lm}_i}{\psi_0^2 l_n^4} \sum_{h=i}^{i+pl_n-2} \sum_{h'=h+1}^{i+pl_n-1} \phi_1^2\left(\frac{h'-h}{l_n}\right) \\
+ p^2\, O_p\big[l_n^{-3} + l_n^{-2}\lambda(p)^n_i\big],
\end{aligned}$$

by **Lemma 9**, we finish the proof. ☐

Lemma 13. *Assume* ***Assumption A-***γ *and* ***Assumption A-***ν, $\forall q,r\in\mathbb{N}^+$, *we have*

$$E\left[(\overline{Y}^{n,j}_h-\overline{X}^{n,j}_h)^q(\overline{Y}^{n,k}_{h'}-\overline{X}^{n,k}_{h'})^r|\mathcal{F}'^n_{h\wedge h'-1}\right]=\begin{cases}0 & q+r=1\\ \Gamma^{n,jk}_{h,h'} & q=r=1\\ O_p\big(l_n^{-7/2}\big)\mathbf{1}_{\{|h-h'|\le l_n\}} & q+r=3\\ O_p(l_n^{-8}) & q+r=8,\end{cases}$$

and

$$E\left[\big(\overline{Y}^{n,j}_h-\overline{X}^{n,j}_h\big)\big(\overline{Y}^{n,k}_h-\overline{X}^{n,k}_h\big)\big(\overline{Y}^{n,l}_{h'}-\overline{X}^{n,l}_{h'}\big)\big(\overline{Y}^{n,m}_{h'}-\overline{X}^{n,m}_{h'}\big)|\mathcal{F}'^n_{h\wedge h'}\right]$$
$$=\Gamma^{n,jk}_{h}\Gamma^{n,lm}_{h'}+\Gamma^{n,jl}_{h,h'}\Gamma^{n,km}_{h,h'}+\Gamma^{n,jm}_{h,h'}\Gamma^{n,kl}_{h,h'}+O_p(l_n^{-5}).$$

Proof. We prove these estimates case by case.

(1) $q+r=1$. Since $\bar{Y}^n_h-\bar{X}^n_h=(\psi_0 l_n)^{-1/2}\sum_{v=0}^{l_n-1}(\varphi^n_{v+1}-\varphi^n_v)(Y^n_{h+v}-X^n_{h+v})$, this case can be confirmed by **Assumption A-**γ, **Assumption A-**ν;

(2) $q=r=1$. $(\overline{Y}^{n,j}_h-\overline{X}^{n,j}_h)(\overline{Y}^{n,k}_{h'}-\overline{X}^{n,k}_{h'})=(\psi_0 l_n)^{-1}\sum_{v=0}^{l_n-2}\sum_{v'=0}^{l_n-1}(\varphi^n_{v+1}-\varphi^n_v)(\varphi^n_{v'+1}-\varphi^n_{v'})\varepsilon^{n,j}_{h+v}\varepsilon^{n,k}_{h'+v'}$. By **Assumption A-**$\gamma$, **Assumption A-**$\nu$, consider $\mathrm{supp}(\varphi)\subseteq[0,1]$,

$$E\big[(\overline{Y}^{n,j}_h-\overline{X}^{n,j}_h)(\overline{Y}^{n,k}_{h'}-\overline{X}^{n,k}_{h'})|\mathcal{F}'^n_{h\wedge h'-1}\big]$$
$$=\frac{1}{\psi_0 l_n}\sum_{v=h\vee h'}^{h\wedge h'+l_n-1}\left(\varphi^n_{v-h+1}-\varphi^n_{v-h}\right)\left(\varphi^n_{v-h'+1}-\varphi^n_{v-h'}\right)\gamma^{n,jk}_v,$$

hence this term is checked;

(3) $q+r=3$. According to **Assumption A-**γ, **Assumption A-**ν,

$$E\big[(\overline{Y}_h^{n,j}-\overline{X}_h^{n,j})^2(\overline{Y}_{h'}^{n,k}-\overline{X}_{h'}^{n,k})|\mathcal{F}'^n_{h\wedge h'-1}\big]=$$
$$(\psi_0 l_n)^{-3/2}\sum_{v=h\vee h'}^{h\wedge h'+l_n-1}(\varphi^n_{v-h+1}-\varphi^n_{v-h})^2(\varphi^n_{v-h'+1}-\varphi^n_{v-h'})E\big[(\varepsilon_v^{n,j})^2\,\varepsilon_v^{n,k}|\mathcal{F}'^n_{h\wedge h'-1}\big],$$

then this case is confirmed;

(4) $q+r=8$. This can be shown by viewing $(\overline{Y}_h^{n,j}-\overline{X}_h^{n,j})^q$ as the sum of

$$(\psi_0 l_n)^{-q/2}\prod_{r=1}^{q}(\varphi^n_{v_r+1}-\varphi^n_{v_r})(Y^{n,j}_{h+v_r}-X^{n,j}_{h+v_r})\text{'s;}$$

(5) Let

$$\xi_v^{n,jk,lm}=E\big(\varepsilon_v^{n,j}\varepsilon_v^{n,k}\varepsilon_v^{n,l}\varepsilon_v^{n,m}|\mathcal{F}^n_{v-1}\big)-E\big(\varepsilon_v^{n,j}\varepsilon_v^{n,k}|\mathcal{F}^n_{v-1}\big)E\big(\varepsilon_v^{n,l}\varepsilon_v^{n,m}|\mathcal{F}^n_{v-1}\big),$$

and

$$\eta_h^{n,jk}\;=\;\frac{1}{\psi_0 l_n}\sum_{v=0}^{h+l_n-2}(\varphi^n_{v-h+1}-\varphi^n_{v-h})\sum_{v'=v+1}^{h+l_n-1}(\varphi^n_{v'-h+1}-\varphi^n_{v'-h})\big(\varepsilon_v^{n,j}\varepsilon_{v'}^{n,k}+\varepsilon_v^{n,k}\varepsilon_{v'}^{n,j}\big).$$

Based on the case $q=r=1$,

$$\begin{aligned}E&\Big[\big(\overline{Y}_h^{n,j}-\overline{X}_h^{n,j}\big)\big(\overline{Y}_h^{n,k}-\overline{X}_h^{n,k}\big)\big(\overline{Y}_{h'}^{n,l}-\overline{X}_{h'}^{n,l}\big)\big(\overline{Y}_{h'}^{n,m}-\overline{X}_{h'}^{n,m}\big)|\mathcal{F}'^n_{h\wedge h'-1}\Big]\\ &=\Gamma_h^{n,jk}\Gamma_{h'}^{n,lm}+\frac{1}{\psi_0^2 l_n^2}\sum_{v=h\vee h'}^{h\wedge h'+l_n-1}(\varphi^n_{v-h+1}-\varphi^n_{v-h})^2(\varphi^n_{v-h'+1}-\varphi^n_{v-h'})^2\,\xi_v^{n,jk,lm}\\ &\qquad+E\Big(\eta_h^{n,jk}\eta_{h'}^{n,lm}|\mathcal{F}'^n_{h\wedge h'-1}\Big),\end{aligned}$$

since for $v' > v, h' > h$, $E(\varepsilon_v^{n,j}\varepsilon_{v'}^{n,k}\varepsilon_h^{n,l}\varepsilon_{h'}^{n,m}|\mathcal{F}'^n_{h\wedge v}) \neq 0$ if and only if $v = h, v' = h'$,

$$
\begin{aligned}
E\big(\eta_h^{n,jk}\eta_{h'}^{n,lm}|\mathcal{F}'^n_{h\wedge h'}\big) = \frac{1}{\psi_0^2 l_n^2}\sum_{v=h\vee h'}^{h\wedge h'+l_n-2}\ \sum_{v'=h\vee h', v'\neq v}^{h\wedge h'+l_n-1}(\varphi^n_{v-h+1}-\varphi^n_{v-h})(\varphi^n_{v-h'+1}-\varphi^n_{v-h'})\\
\times(\varphi^n_{v'-h+1}-\varphi^n_{v'-h})(\varphi^n_{v'-h'+1}-\varphi^n_{v'-h'})\big(\gamma_v^{n,jl}\gamma_{v'}^{n,km}+\gamma_v^{n,jm}\gamma_{v'}^{n,kl}\big)\\
= \Gamma^{n,jl}_{h,h'}\Gamma^{n,km}_{h,h'}+\Gamma^{n,jm}_{h,h'}\Gamma^{n,kl}_{h,h'}+O_p(l_n^{-5}),
\end{aligned}
$$

which validates this case. □

Lemma 14. *Assume **Assumption A-**γ, **SA-**ν,*

$$
\begin{aligned}
E(\zeta_h^n|\mathcal{F}'^n_h) &= \bar{X}_h^n{\bar{X}_h^n}^{\mathrm{T}} - \bar{C}_h^n\\
\|E\left(\zeta_h^n|\mathcal{F}_h^n\right)\| &\leq K\left(l_n\Delta_n^2 + l_n^{1/2}\Delta_n^{3/2}\,\bar{\lambda}(1)_h^n\right)\\
E\left(\zeta_h^{n,jk}\zeta_{h'}^{n,lm}|\mathcal{F}'^n_{h\wedge h'}\right) &= \big(\bar{X}_h^{n,j}\bar{X}_h^{n,k}-\bar{C}_h^{n,jk}\big)\big(\bar{X}_{h'}^{n,l}\bar{X}_{h'}^{n,m}-\bar{C}_{h'}^{n,lm}\big)\\
&\quad+\bar{X}_h^{n,j}\bar{X}_{h'}^{n,l}\Gamma^{n,km}_{h,h'}+\bar{X}_h^{n,j}\bar{X}_{h'}^{n,m}\Gamma^{n,kl}_{h,h'}\\
&\quad+\bar{X}_h^{n,k}\bar{X}_{h'}^{n,l}\Gamma^{n,jm}_{h,h'}+\bar{X}_h^{n,k}\bar{X}_{h'}^{n,m}\Gamma^{n,jl}_{h,h'}\\
&\quad+\Gamma^{n,jl}_{h,h'}\Gamma^{n,km}_{h,h'}+\Gamma^{n,jm}_{h,h'}\Gamma^{n,kl}_{h,h'}+O_p\big[l_n^{-5}+\big(\|\bar{X}_h^n\|+\|\bar{X}_{h'}^n\|\big)l_n^{-7/2}\big]\\
E\left(\|\zeta_h^n\|^4|\mathcal{F}_h^n\right) &\leq K\big(l_n^{-8}+\Delta_n^4\big).
\end{aligned}
$$

Proof. We prove these case by case.

(1) Because $\zeta_h^n = \big(\bar{X}_h^n{\bar{X}_h^n}^{\mathrm{T}} - \bar{C}_h^n\big) + \big[(\bar{Y}_h^n-\bar{X}_h^n)(\bar{Y}_h^n-\bar{X}_h^n)^{\mathrm{T}}-\Gamma_h^n\big] + \bar{X}_h^n(\bar{Y}_h^n-\bar{X}_h^n)^{\mathrm{T}} + (\bar{Y}_h^n-\bar{X}_h^n){\bar{X}_h^n}^{\mathrm{T}}$, the first case can be validated by **Assumption A-**γ and **Lemma 13**;

(2) In view of (3.17) and (3.29), we have $E(\zeta_h^n|\mathcal{F}_h^n) = (\psi_0 l_n)^{-1}E[B(n, h\Delta_n, (l_n-1)\Delta_n)|\mathcal{F}_i^n]$. By a similar calculation in part **(2)** of the proof for **Lemma 11**, we can confirm this case;

(3) Let $\xi^{n,j_1\cdots j_q}_{h_1\cdots h_q} = \prod_{v=1}^{q}\big(\overline{Y}^{n,j_v}_{h_v} - \overline{X}^{n,j_v}_{h_v}\big)$, by **Assumption A-**$\gamma$,

$$E\big(\zeta^{n,jk}_h \zeta^{n,lm}_{h'} | \mathcal{F}'^n_{h\wedge h'}\big) = \big(\overline{X}^{n,j}_h \overline{X}^{n,k}_h - \overline{C}^{n,jk}_h\big)\big(\overline{X}^{n,l}_{h'} \overline{X}^{n,m}_{h'} - \overline{C}^{n,lm}_{h'}\big) + \sum_{r=1}^{3} \alpha(r)^{n,jk,lm}_{h,h'},$$

where

$$\begin{aligned}
\alpha(1)^{n,jk,lm}_{h,h'} &= \overline{X}^{n,j}_h \overline{X}^{n,l}_{h'} \xi^{n,km}_{hh'} + \overline{X}^{n,j}_h \overline{X}^{n,m}_{h'} \xi^{n,kl}_{hh'} + \overline{X}^{n,k}_h \overline{X}^{n,l}_{h'} \xi^{n,jm}_{hh'} + \overline{X}^{n,k}_h \overline{X}^{n,m}_{h'} \xi^{n,jl}_{hh'} \\
\alpha(2)^{n,jk,lm}_{h,h'} &= \xi^{n,jklm}_{hhh'h'} - \xi^{n,jk}_{hh}\Gamma^{n,lm}_{h'} - \xi^{n,lm}_{h'h'}\Gamma^{n,jk}_{h} + \Gamma^{n,jk}_{h}\Gamma^{n,lm}_{h'} \\
\alpha(3)^{n,jk,lm}_{h,h'} &= \overline{X}^{n,j}_h \xi^{n,klm}_{hh'h'} + \overline{X}^{n,k}_h \xi^{n,jlm}_{hh'h'} + \overline{X}^{n,l}_{h'} \xi^{n,jkm}_{hhh'} + \overline{X}^{n,m}_{h'} \xi^{n,jkl}_{hhh'},
\end{aligned}$$

then this case can be justified by **Lemma 13** and (3.26);

(4) According to **(1)**, $\|\zeta^n_i\|^4 \le K\big(\|\overline{C}^n_i\|^4 + \|\Gamma^n_i\|^4 + \|\overline{X}^n_i\|^8 + \|\overline{Y}^n_i - \overline{X}^n_i\|^8\big)$. By (3.19), (3.26), the first claim in **Lemma 12** and the case $q+r=8$ in **Lemma 13**, this case can be checked. $\square$

Set $\theta = l_n\Delta_n^{1/2}$. Before stating the next results, for a given $p \in \mathbb{N}^+$, we need to define a $d\times d\times d\times d$-tensor-valued function over the domain $\mathcal{M}^+_d \times \mathcal{M}^+_d$:

$$\Xi(x,z;p) = \frac{2\theta}{\psi_0^2}\left[\frac{p\Phi_{00} - \Psi_{00}}{p+1}\Sigma(x) + \frac{p\Phi_{01} - \Psi_{01}}{\theta^2(p+1)}\Theta(x,z) + \frac{p\Phi_{11} - \Psi_{11}}{\theta^4(p+1)}\Sigma(z)\right], \tag{3.34}$$

where Σ and Θ are also tensor-valued functions such that

$$\begin{aligned}
\Sigma(x)^{jk,lm} &= x^{jl}x^{km} + x^{jm}x^{kl} \\
\Theta(x,z)^{jk,lm} &= x^{jl}z^{km} + x^{jm}z^{kl} + x^{km}z^{jl} + x^{kl}z^{jm}.
\end{aligned} \tag{3.35}$$

Similarly, define another tensor function

$$\Xi(x,z) = \frac{2\theta}{\psi_0^2}\left[\Phi_{00}\Sigma(x) + \frac{\Phi_{01}}{\theta^2}\Theta(x,z) + \frac{\Phi_{11}}{\theta^4}\Sigma(z)\right], \tag{3.36}$$

we have $\Xi(c,\gamma;p) \to \Xi(c,\gamma)$ pointwise as $p \to \infty$.

Lemma 15. *Under **Assumption A-**γ, **SA-**ν,*

$$\begin{aligned}
E[\zeta(Y,p)_i^n|\mathcal{F}_i'^n] &= \zeta(X,p)_i^n \\
\|E[\zeta(Y,p)_i^n|\mathcal{F}_i^n]\| &\leq Kp\left(l_n^2\Delta_n^2 + l_n^{3/2}\Delta_n^{3/2}\,\overline{\lambda}(p)_i^n\right) \\
E[\|\zeta(Y,p)_i^n\|^q|\mathcal{F}_i^n] &\leq K_q\, p^{\lfloor q/2\rfloor\vee 1}\left(l_n^{-q} + l_n^q\Delta_n^q\right), \qquad q = 1,2,3,4,
\end{aligned}$$

moreover

$$\begin{aligned}
&\left|E\left[\zeta(Y,p)_i^{n,jk}\zeta(Y,p)_i^{n,lm}|\mathcal{F}_i^n\right] - (p+1)l_n\Delta_n^{3/2}\,\Xi(c_i^n,\gamma_i^n;p)^{jk,lm}\right| \leq \\
&\qquad Kp^2\left[l_n^{-3/2}(l_n^{-3/2} + \Delta_n^{1/2} + l_n^2\Delta_n^{3/2} + l_n^4\Delta_n^{5/2}) + (l_n^{-2} + \Delta_n + l_n^2\Delta_n^2)\,\overline{\lambda}(p)_i^n\right].
\end{aligned} \tag{3.37}$$

Proof. We are going to check these 4 estimates case by case.

(1)&**(2)** These 2 cases are consequences of the second claim in **Lemma 11** and the first claim in **Lemma 14**.

(3) Note ζ_h^n and $\zeta_{h'}^n$ are independent if $|h'-h| > l_n$, this case can be checked by the second and last claims in **Lemma 14** with Jensen inequality.

(4) Let

$$\begin{aligned}
\alpha(1)_{h,h'}^{n,jk,lm} &= \overline{X}_h^{n,j}\overline{X}_{h'}^{n,l}\Gamma_{h,h'}^{n,km} + \overline{X}_h^{n,j}\overline{X}_{h'}^{n,m}\Gamma_{h,h'}^{n,kl} + \overline{X}_h^{n,k}\overline{X}_{h'}^{n,l}\Gamma_{h,h'}^{n,jm} + \overline{X}_h^{n,k}\overline{X}_{h'}^{n,m}\Gamma_{h,h'}^{n,jl} \\
\alpha(2)_{h,h'}^{n,jk,lm} &= \Gamma_{h,h'}^{n,jl}\Gamma_{h,h'}^{n,km} + \Gamma_{h,h'}^{n,jm}\Gamma_{h,h'}^{n,kl},
\end{aligned}$$

according to **Lemma 14**,

$$\zeta(Y,p)_i^{n,jk}\zeta(Y,p)_i^{n,lm} = \sum_{h=i}^{i+pl_n-1}\big(\overline{X}_h^{n,j}\overline{X}_h^{n,k}-\overline{C}_h^{n,jk}\big)\sum_{h'=i}^{i+pl_n-1}\big(\overline{X}_{h'}^{n,l}\overline{X}_{h'}^{n,m}-\overline{C}_{h'}^{n,lm}\big)$$
$$+\sum_{h=i}^{i+pl_n-2}\sum_{h'=h+1}^{i+pl_n-1}\Big[\alpha(1)_{h,h'}^{n,jk,lm}+\alpha(1)_{h,h'}^{n,lm,jk}+\alpha(2)_{h,h'}^{n,jk,lm}+\alpha(2)_{h,h'}^{n,lm,jk}\Big]$$
$$+\sum_{h=i}^{i+pl_n-1}\Big[\alpha(1)_{h,h}^{n,jk,lm}+\alpha(2)_{h,h}^{n,jk,lm}\Big]+p^2\,O_p\big(l_n^{-3}+l_n^{-3/2}\Delta_n^{1/2}\big),$$

by (3.9), (3.26), as well as **Lemma 12**,

$$E\Big[\zeta(Y,p)_i^{n,jk}\zeta(Y,p)_i^{n,lm}|\mathcal{F}_i'^n\Big]=\zeta(X,p)_i^{n,jk}\zeta(X,p)_i^{n,lm}$$
$$+\frac{2}{\psi_0 l_n^2}\Big[\zeta(X,p)_i'^{n,jl}\gamma_i^{n,km}+\zeta(X,p)_i'^{n,jm}\gamma_i^{n,kl}+\zeta(X,p)_i'^{n,kl}\gamma_i^{n,jm}+\zeta(X,p)_i'^{n,km}\gamma_i^{n,jl}\Big]$$
$$+\frac{2(\gamma_i^{n,jl}\gamma_i^{n,km}+\gamma_i^{n,jm}\gamma_i^{n,kl})}{\psi_0^2 l_n^2}(p\Phi_{11}-\Psi_{11})$$
$$+p^2\,O_p\big[l_n^{-3}+l_n^{-3/2}\Delta_n^{1/2}+(l_n^{-2}+\Delta_n)\overline{\lambda}(p)_i^n\big].$$

combined with **Lemma 11**, this case is checked. □

3.2.6 estimates of local estimation errors

Lemma 16. *Under **Assumption SA-ν**, $\forall p\in\mathbb{N}^+$,*

$$\Delta_n\sum_{i=0}^{N_t^n-1}\sum_{h=1}^{k_n-l_n+1}\overline{\lambda}(1)_{i+h}^n\xrightarrow{\mathbb{P}}0$$
$$\frac{k_n\Delta_n}{pl_n}\sum_{i=0}^{N_t^n-1}\sum_{h=m(n,p)(p+1)l_n}^{k_n-l_n}\overline{\lambda}(1)_{ik_n+1+h}^n\xrightarrow{\mathbb{P}}0$$
$$pl_n\Delta_n\sum_{i=0}^{N_t^n-1}\sum_{h=0}^{m(n,p)-1}\overline{\lambda}(p)_{ik_n+a(n,p,h)}^n\bigvee\overline{\lambda}(1)_{ik_n+b(n,p,h)}^n\xrightarrow{\mathbb{P}}0;$$

in addition, if ***Assumption A**-c,* ***A**-γ hold,*

$$\overline{\lambda}(p)_i^n \leq K p^{1/2} l_n^{1/2} \Delta_n^{1/2}.$$

Proof. **Corollary 3** implies the last claim. Let's focus on the first two claims. Set

$$\begin{aligned} N(\epsilon)_{u,t} &= \textstyle\sum_{s\in[u,u+t]} \left(\mathbf{1}_{\{\|\Delta b_s\|>\epsilon\}} + \mathbf{1}_{\{\|\Delta c_s\|>\epsilon\}} + \mathbf{1}_{\{\|\Delta \gamma_s\|>\epsilon\}}\right) \\ \rho(\epsilon,\delta)_t &= \sup_{u<s<u+\delta<t}(\|b_s - b_u\| + \|c_s - c_u\| + \|\gamma_s - \gamma_u\|)\mathbf{1}_{\{N(\epsilon)_{u,\delta}=0\}}, \end{aligned}$$

since $k_n - l_n - m(n,p)(p+1)l_n \leq p l_n$, we have

$$\left.\begin{aligned} \overline{\lambda}(p)_i^n &\leq K N(\epsilon)_{i\Delta_n, p l_n \Delta_n} + \rho(\epsilon, p l_n \Delta_n)_t \\ \textstyle\sum_{h=1}^{k_n-l_n+1} \overline{\lambda}(1)_{i+h}^n &\leq K l_n N(\epsilon)_{i\Delta_n, k_n\Delta_n} + k_n \rho(\epsilon, l_n\Delta_n)_t \\ \textstyle\sum_{h=m(n,p)(p+1)l_n}^{k_n-l_n} \overline{\lambda}(1)_{i+1+h}^n &\leq K l_n N(\epsilon)_{[i+1+m(n,p)(p+1)l_n]\Delta_n, p l_n \Delta_n} \\ &\qquad + p l_n \rho(\epsilon, l_n\Delta_n)_t \end{aligned}\right\}. \tag{3.38}$$

Note that $N_t^n \asymp t(k_n\Delta_n)^{-1}$ and $N_t^n m(n,p) \asymp t(p l_n \Delta_n)^{-1}$, by (3.38),

$$\begin{aligned} \Delta_n \sum_{i=0}^{N_t^n-1} \sum_{h=1}^{k_n-l_n+1} \overline{\lambda}(1)_{ik_n+h}^n &\leq K\left[l_n\Delta_n N(\epsilon)_{0,t} + t\rho(\epsilon, l_n\Delta_n)_t\right] \\ p l_n \Delta_n \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \overline{\lambda}(p)_{ik_n+a(n,p,h)}^n &\leq K\left[p l_n \Delta_n N(\epsilon)_{0,t} + t\,\rho(\epsilon, p l_n\Delta_n)_t\right] \\ p^{-1}k_n l_n^{-1}\Delta_n \sum_{i=0}^{N_t^n-1} \sum_{h=m(n,p)(p+1)l_n}^{k_n-l_n} \overline{\lambda}(1)_{ik_n+1+h}^n &\leq K\left[p^{-1}k_n\Delta_n N(\epsilon)_{0,t} + t\,\rho(\epsilon, l_n\Delta_n)_t\right], \end{aligned}$$

since $\mathbb{E}[N(\epsilon)_{0,t}] < K_\epsilon t$, and $\limsup_{\Delta_n\to 0} \rho(\epsilon, k l_n \Delta_n)_t \leq 3\epsilon$, $\forall k \in \mathbb{N}^+$, this lemma can be proved. $\square$

Lemma 17. *If we let $l_n\Delta_n^{1/2} \asymp \theta$ where θ is positive finite and $k_n\Delta_n^{3/4} \to 0$, $\forall p \in \mathbb{N}^+$,*

$$\begin{aligned}
\|E\left(M(p)_i^n|\mathcal{F}_i^n\right)\| &\leq K\Big[\Delta_n^{1/2} + pk_n^{-1}\Delta_n^{-1/4}\sum_{h=0}^{m(n,p)-1}\overline{\lambda}(p)_{i+a(n,p,h)}^n\Big] \\
\|E\left(N(p)_i^n|\mathcal{F}_i^n\right)\| &\leq K\Bigg[\big(p^{-1}\Delta_n^{1/2} + pk_n^{-1}\big)k_n^{-1}\Delta_n^{-1/4}\sum_{h=0}^{m(n,p)-1}\overline{\lambda}(1)_{i+b(n,p,h)}^n \\
&\qquad + k_n^{-1}\Delta_n^{1/4}\sum_{h=m(n,p)(p+1)l_n}^{k_n-l_n}\overline{\lambda}(1)_{i+1+h}^n\Bigg];
\end{aligned}$$

*additionally, if **Assumption A**-c, **A**-γ hold,*

$$\begin{aligned}
E\left(\|M(p)_i^n\|^q|\mathcal{F}_i^n\right) &\leq \begin{cases} K_q\big(k_n\Delta_n^{1/2}\big)^{-q/2}, & q = 1,2,4 \\ K\big(k_n\Delta_n^{1/2}\big)^{-2}, & q = 3 \end{cases} \\
E\left(\|N(p)_i^n\|^q|\mathcal{F}_i^n\right) &\leq \begin{cases} K_q\big[p^{-q/2} + p^{\lfloor q/2\rfloor\vee 1}(k_n\Delta_n^{1/2})^{-q/2}\big](k_n\Delta_n^{1/2})^{-q/2}, & q = 1,2,4 \\ K\big[p^{-1} + p(k_n\Delta_n^{1/2})^{-1}\big]\big(k_n\Delta_n^{1/2}\big)^{-2}, & q = 3 \end{cases}
\end{aligned}$$

Proof. We can write

$$N(p)_i^n = N(p,1)_i^n + N(p,2)_i^n, \tag{3.39}$$

where

$$\begin{aligned}
N(p,1)_i^n &= (k_n - l_n)^{-1}\Delta_n^{-1}\textstyle\sum_{h=0}^{m(n,p)-1}\zeta(Y,1)_{i+b(n,p,h)}^n \\
N(p,2)_i^n &= (k_n - l_n)^{-1}\Delta_n^{-1}\textstyle\sum_{h=m(n,p)(p+1)l_n}^{k_n-l_n}\zeta_{i+1+h}^n.
\end{aligned}$$

(1) By **Lemma 14**, **Lemma 15** and $m(n,p) \asymp p^{-1}k_nl_n^{-1}$, $k_n - l_n - m(n,p)(p+1)l_n \leq pl_n$,

$$\left.\begin{aligned}
\|E[M(p)_i^n|\mathcal{F}_i^n]\| &\leq K\Big[l_n\Delta_n + pk_n^{-1}l_n^{3/2}\Delta_n^{1/2}\textstyle\sum_{h=0}^{m(n,p)-1}\overline{\lambda}(p)_{i+a(n,p,h)}^n\Big] \\
\|E[N(p,1)_i^n|\mathcal{F}_i^n]\| &\leq K\Big[p^{-1}l_n\Delta_n + k_n^{-1}l_n^{3/2}\Delta_n^{1/2}\textstyle\sum_{h=0}^{m(n,p)-1}\overline{\lambda}(1)_{i+b(n,p,h)}^n\Big] \\
\|E[N(p,2)_i^n|\mathcal{F}_i^n]\| &\leq K\Big[pk_n^{-1}l_n^2\Delta_n + k_n^{-1}l_n^{1/2}\Delta_n^{1/2}\textstyle\sum_{h=m(n,p)(p+1)l_n}^{k_n-l_n}\overline{\lambda}(1)_{i+1+h}^n\Big]
\end{aligned}\right\}, \tag{3.40}$$

then the estimates of norms of expectations can be deduced from (3.38) and (3.40);

(2) for $q = 2,3,4$, by the independence among $\zeta(Y,p)^n_{i+a(n,p,h)}$'s and the independence among $\zeta(Y,1)^n_{i+b(n,p,h)}$'s,

$$E\big[\|M(p)^n_i\|^2|\mathcal{F}^n_i\big] \leq \frac{K}{k_n^2\Delta_n^2}\left(\sum_{h=0}^{m(n,p)-1} E\big[\|\zeta(Y,p)^n_{i+a(n,p,h)}\|^2|\mathcal{F}^n_i\big]\right.$$
$$\left. + \sum_{h=0}^{m(n,p)-1}\sum_{h'=0}^{m(n,p)-1} \big\|E\big[\zeta(Y,p)^n_{i+a(n,p,h)}|\mathcal{F}^n_i\big]\big\| \cdot \big\|E\big[\zeta(Y,p)^n_{i+a(n,p,h')}|\mathcal{F}^n_i\big]\big\|\right),$$

$$E\big[\|M(p)^n_i\|^3|\mathcal{F}^n_i\big] \leq \frac{K}{k_n^2\Delta_n^2}\left(\sum_{h=0}^{m(n,p)-1} E\big[\|\zeta(Y,p)^n_{i+a(n,p,h)}\|^3|\mathcal{F}^n_i\big]\right.$$
$$+ \sum_{h=0}^{m(n,p)-1}\sum_{h'=0}^{m(n,p)-1} E\big(\|\zeta(Y,p)^n_{i+a(n,p,h)}\|^2|\mathcal{F}^n_i\big)\big\|E\big[\zeta(Y,p)^n_{i+a(n,p,h')}|\mathcal{F}^n_i\big]\big\|$$
$$\left. + \prod_{r=1}^{3}\sum_{h_r=0}^{m(n,p)-1} \big\|E\big[\zeta(Y,p)^n_{i+a(n,p,h_r)}|\mathcal{F}^n_i\big]\big\|\right),$$

$$E(\|M(p)^n_i\|^4|\mathcal{F}^n_i) \leq \frac{K}{k_n^2\Delta_n^2}\left(\sum_{h=0}^{m(n,p)-1} E\big(\|\zeta(Y,p)^n_{i+a(n,p,h)}\|^4|\mathcal{F}^n_i\big)\right.$$
$$+ \sum_{h=0}^{m(n,p)-1} E\big(\|\zeta(Y,p)^n_{i+a(n,p,h)}\|^3|\mathcal{F}^n_i\big) \sum_{h=0}^{m(n,p)-1} \big\|E\big[\zeta(Y,p)^n_{i+a(n,p,h)}|\mathcal{F}^n_i\big]\big\|$$
$$+ \sum_{h=0}^{m(n,p)-1}\sum_{h'=0}^{m(n,p)-1} E\big(\|\zeta(Y,p)^n_{i+a(n,p,h)}\|^2|\mathcal{F}^n_i\big)E\big(\|\zeta(Y,p)^n_{i+a(n,p,h')}\|^2|\mathcal{F}^n_i\big)$$
$$+ \sum_{h=0}^{m(n,p)-1} E\big(\|\zeta(Y,p)^n_{i+a(n,p,h)}\|^2|\mathcal{F}^n_i\big)\prod_{r=1}^{2}\sum_{h_r=0}^{m(n,p)-1} \big\|E\big[\zeta(Y,p)^n_{i+a(n,p,h_r)}|\mathcal{F}^n_i\big]\big\|$$
$$\left. + \prod_{r=1}^{4}\sum_{h_r=0}^{m(n,p)-1} \big\|E\big[\zeta(Y,p)^n_{i+a(n,p,h_r)}|\mathcal{F}^n_i\big]\big\|\right).$$

We have similar expansions for $E(\|N(p,1)^n_i\|^2|\mathcal{F}^n_i)$. Similar calculation for the third claim of **Lemma 15** can give an estimate of $E(\|N(p,2)_i^n\|^q|\mathcal{F}_i^n)$. The estimates of expectations of norms are readily implied by **Lemma 14**, **15** and the last claim in **Lemma 16** with Jensen's inequality. $\square$

Lemma 18. *Under* ***Assumption A***-c, ***A***-γ, ***SA***-ν, *if we let* $l_n\Delta_n^{1/2} \asymp \theta$ *where* θ *is positive finite and* $k_n\Delta_n^{3/4} \to 0$, *given* $p \in \mathbb{N}^+$,

$$\begin{aligned}
\|E(\beta_i^n|\mathcal{F}_i^n)\| &\leq K\left[k_n\Delta_n + k_n^{-1}\Delta_n^{1/4}\sum_{h=1}^{k_n-l_n+1}\overline{\lambda}(1)_{i+h}^n\right] \\
E(\|\beta_i^n\|^q|\mathcal{F}_i^n) &\leq \begin{cases} K_q\left[(k_n\Delta_n)^{(q/2)\wedge 1} + (k_n\Delta_n^{1/2})^{-q/2}\right], & q=1,2,4 \\ K\left[k_n\Delta_n + (k_n\Delta_n^{1/2})^{-2}\right], & q=3 \end{cases}
\end{aligned}$$

additionally,

$$\begin{aligned}
&\left|E\big(\beta_i^{n,jk}\beta_i^{n,lm}|\mathcal{F}_i^n\big) - \frac{1}{k_n\Delta_n^{1/2}}\Xi(c_i^n,\gamma_i^n)^{jk,lm}\right| \\
&\qquad \leq K\left[k_n\Delta_n + p^{-1}(k_n\Delta_n^{1/2})^{-1} + p^2(k_n\Delta_n^{1/2})^{-2}\sum_{h=0}^{m(n,p)-1}\overline{\lambda}(p+1)_{i+a(n,p,h)}^n\right],
\end{aligned}$$

where $\Xi(c,\gamma)_i^n$ *is defined in (3.36).*

Proof. The estimates for $\|E(\beta_i^n|\mathcal{F}_i^n)\|$ and $E(\|\beta_i^n\|^q|\mathcal{F}_i^n)$ can be derived from (3.10), (3.13), **Lemma 7**, **8** and the independence between ζ_h^n, $\zeta_{h'}^n$ provided $|h'-h| > l_n$. Next, we will focus on the estimate of $E\big(\beta_i^{n,jk}\beta_i^{n,lm}|\mathcal{F}_i^n\big)$ and proceed as the following 5 steps.

(1) *Notation*

We need to define some variables. Recall the variables in (3.8), define

$$\begin{aligned}
R(p)_{i+v}^n &= \textstyle\sum_{h=v}^{v+pl_n-1} R_{i+h}^n \\
D(p)_{i+v}^n &= \textstyle\sum_{h=v}^{v+pl_n-1} D_{i+h}^n \\
A(p)_{i+v}^n &= \textstyle\sum_{h=v}^{v+pl_n-1}(c_{i+h}^n - c_i^n)\Delta_n,
\end{aligned}$$

and define

$$\begin{aligned}
\alpha^n_{i,h} &= -R^n_{i+h} + D^n_{i+h} + (c^n_{i+h} - c^n_i)\Delta_n + \zeta^n_{i+h} \\
\alpha(p)^n_{i,h} &= -R(p)^n_{i+h} + D(p)^n_{i+h} + A(p)^n_{i+h} + \zeta(Y,p)^n_{i+h} \\
\alpha(p)'^n_{i,h} &= -R(p)^n_{i+h} + D(p)^n_{i+h} + A(p)^n_{i+h},
\end{aligned}$$

then we have

$$\beta^n_i = \frac{1}{(k_n - l_n)\Delta_n}\left(\sum_{h=0}^{m(n,p)-1} \alpha(p+1)^n_{i,a(n,p,h)} + \sum_{h=m(n,p)(p+1)l_n}^{k_n-l_n} \alpha^n_{i,1+h}\right). \tag{3.41}$$

(2) *Ingredient Estimates*

Note

$$A(p)^n_{i+v} = \Delta_n \sum_{h=v+1}^{v+pl_n-1} \Delta^n_{i+h} c \times (v + pl_n - h) + pl_n\Delta_n(c^n_{i+v} - c^n_i),$$

we have

$$\left.\begin{aligned}
&E\big(\|A(p)^n_{i+v}\|^q|\mathcal{F}^n_i\big) &&\leq K_q(k_n l_n\Delta_n)^q\Delta_n^{(q/2)\wedge 1}\big(pl_n + v^{(q/2)\wedge 1}\big) \\
&\|E[A(p)^n_{i+v}|\mathcal{F}^n_i]\| &&\leq Kpl_n\Delta_n^2(pl_n + v)
\end{aligned}\right\}.$$

Use (3.15) and the fact that $D^n_i \perp\!\!\!\perp D^n_{i+h}$ if $|h| > l_n$ to estimate $E(\|D(p)^n_i\|^2|\mathcal{F}^n_i)$; multiply the bounds in **Lemma 7**, **8** with $k_n^q\Delta_n^q$ and then swap k_n with pl_n to estimate $E(\|R(p)^n_i\|^2|\mathcal{F}^n_i)$, $\|E(R(p)^n_i|\mathcal{F}^n_i)\|$, $\|E(D(p)^n_i|\mathcal{F}^n_i)\|$; use **Lemma 15** and let $l_n \asymp \Delta_n^{-1/2}$, we have the following scaling properties.

scaling properties	$E(\\| \cdot \\|^2 \| \mathcal{F}^n_i)$	$\\|E(\cdot \| \mathcal{F}^n_i)\\|$
$D(p)^n_h$	$p\Delta_n^{3/2}$	$p\Delta_n$
$R(p)^n_h$	$p\Delta_n^{3/2}$	$p\Delta_n^{3/2}$
$A(p)^n_{i+v}$	$p^2\Delta_n^2\big(p\Delta_n^{-1/2} + v\big)$	$p\Delta_n^{3/2}\big(p\Delta_n^{-1/2} + v\big)$
$\zeta(Y,p)^n_h$	$p\Delta_n$	$p\big[\Delta_n + \Delta_n^{3/4}\overline{\lambda}(p)^n_h\big]$

(3) *Main & Edge Terms*

By the ingredient estimates above about $E(\|\cdot\|^2|\mathcal{F}_i^n)$ and Cauchy-Schwarz inequality,

$$E\left(\left|\alpha(p+1)_{i,v}^{n,jk}\alpha(p+1)_{i,v}^{n,lm}-\zeta(Y,p+1)_{i+v}^{n,jk}\zeta(Y,p+1)_{i+v}^{n,lm}\right||\mathcal{F}_i^n\right)$$
$$\leq K\big(p^2\Delta_n^{5/4}+p^{3/2}\Delta_n^{3/2}v^{1/2}\big). \quad (3.42)$$

Let $\alpha(p)_i^{\prime\prime n}=\sum_{h=m(n,p)(p+1)l_n}^{k_n-l_n}\alpha_{i,1+h}^n$, according to the table above, by Cauchy-Schwarz inequality and independence between $\alpha(p)_i^{\prime\prime n}$ and $\alpha(p+1)_{i,a(n,p,h)}^n$ with $h\leq m(n,p)-2$,

$$\mathbb{E}\left(\left|\alpha(p)_i^{\prime\prime n,jk}\alpha(p)_i^{\prime\prime n,lm}\right|+\left|\alpha(p)_i^{\prime\prime n,jk}\alpha(p+1)_{i,a(n,p,m(n,p)-1)}^{n,lm}\right|\right)\leq Kp\Delta_n,$$
$$\mathbb{E}\left(\left|\alpha(p)_i^{\prime\prime n,jk}\sum_{h=0}^{m(n,p)-2}\alpha(p+1)_{i,a(n,p,h)}^{n,lm}\right|\right)$$
$$\leq Kp\Big[k_n^3\Delta_n^{7/2}+pk_n\Delta_n^{9/4}\sum_{h=0}^{m(n,p)-2}\overline{\lambda}(p+1)_{i+a(n,p,h)}^n\Big]. \quad (3.43)$$

(4) *Cross Terms*

By the ingredient estimates above and Cauchy-Schwarz inequality, for $v\neq v'$,

$$\left|E\left[\alpha(p)_{i,v}^{\prime n,jk}\alpha(p)_{i,v'}^{\prime n,lm}|\mathcal{F}_i^n\right]\right|\leq Kp^2\Delta_n^2\big[p\Delta_n^{-1/2}+p^{1/2}\Delta_n^{-1/4}(v^{1/2}+v'^{1/2})+v^{1/2}v'^{1/2}\big], \quad (3.44)$$

furthermore by exploiting independence,

$$\left|E\left[\zeta(Y,p)_{i+v}^{n,jk}\alpha(p)_{i,v'}^{\prime n,lm}|\mathcal{F}_i^n\right]\right|\leq$$
$$E\left(\left|\zeta(Y,p)_{i+v}^{n,jk}\right|^2|\mathcal{F}_i^n\right)^{1/2}E\left(\left|\alpha(p)_{i,v'}^{\prime n,lm}\right|^2|\mathcal{F}_i^n\right)^{1/2}\mathbb{1}_{\{|v'-v|=1\}}$$
$$+\left|E\left[\zeta(Y,p)_{i+v}^{n,jk}|\mathcal{F}_i^n\right]\right|\cdot\left|E\left[\alpha(p)_{i,v'}^{\prime n,lm}|\mathcal{F}_i^n\right]\right|+e(p)_{i,v,v'}^{n,jk,lm},$$

where $e(p)_{i,v,v'}^{n,jk,lm}=pl_n\Delta_n\big|E\big[\zeta(Y,p)_{i+v}^{n,jk}(c_{i+v'}^{n,lm}-c_{i+v}^{n,lm})|\mathcal{F}_i^n\big]\big|\mathbf{1}_{\{v'>v\}}$. Using **Lemma 15**, the ingredient estimates above and **Assumption SA-ν**, we get

$$\Big|E\Big[\zeta(Y,p)^{n,jk}_{i+v}\alpha(p)'^{n,lm}_{i,v'}|\mathcal{F}^n_i\Big]\Big| \leq K\Big[\big(p^2\Delta_n^{5/4}+p^{3/2}\Delta_n^{3/2}v'^{1/2}\big)\mathbb{1}_{\{|v'-v|=1\}} \\ +p^2\big(\Delta_n^{3/2}+\Delta_n^{5/2}v'+\Delta_n^{5/4}\overline{\lambda}(p)^n_{i+v}+\Delta_n^{9/4}v'\overline{\lambda}(p)^n_{i+v}\big)\Big]. \quad (3.45)$$

(5) *Decomposition*

Given $j,k,l,m=1,\cdots,d$, we have

$$\Big|E\big(\beta_i^{n,jk}\beta_i^{n,lm}|\mathcal{F}^n_i\big)-(k_n\Delta_n^{1/2})^{-1}\Xi(c^n_i,\gamma^n_i)^{jk,lm}\Big| = \xi_i^{n,1}+\xi_i^{n,2}+\xi_i^{n,3}+\xi_i^{n,4}+\xi_i^{n,5}+\epsilon(p)^n_i,$$

where

$$\begin{aligned}
\xi_i^{n,1} &= \frac{1}{k_n^2\Delta_n^2}\sum_{h=0}^{m(n,p)-1}E\Big(\Big|\alpha(p+1)^{n,jk}_{i,a(n,p,h)}\alpha(p+1)^{n,lm}_{i,a(n,p,h)} \\
&\qquad -\zeta(Y,p+1)^{n,jk}_{i+a(n,p,h)}\zeta(Y,p+1)^{n,lm}_{i+a(n,p,h)}\Big|\,|\mathcal{F}^n_i\Big) \\
\xi_i^{n,2} &= \frac{1}{k_n^2\Delta_n^2}\sum_{h=0}^{m(n,p)-2}\sum_{h'=h+1}^{m(n,p)-1}\Big|E\Big[\alpha(p+1)^{n,jk}_{i,a(n,p,h)}\alpha(p+1)^{n,lm}_{i,a(n,p,h')} \\
&\qquad +\alpha(p+1)^{n,lm}_{i,a(n,p,h)}\alpha(p+1)^{n,jk}_{i,a(n,p,h')}|\mathcal{F}^n_i\Big]\Big|,
\end{aligned}$$

and

$$\begin{aligned}
\xi_i^{n,3} &= \frac{1}{k_n^2\Delta_n^2}\sum_{h=0}^{m(n,p)-1}E\Big(\Big|E\Big[\zeta(Y,p+1)^{n,jk}_{i+a(n,p,h)}\zeta(Y,p+1)^{n,lm}_{i+a(n,p,h)}|\mathcal{F}^n_{i+a(n,p,h)}\Big] \\
&\qquad -(p+2)\theta\Delta_n\Xi\big(c^n_{i+a(n,p,h)},\gamma^n_{i+a(n,p,h)};p+1\big)^{jk,lm}\Big|\,|\mathcal{F}^n_i\Big) \\
\xi_i^{n,4} &= \frac{(p+2)\theta}{k_n^2\Delta_n}\sum_{h=0}^{m(n,p)-1}\Big|E\Big[\Xi\big(c^n_{i+a(n,p,h)},\gamma^n_{i+a(n,p,h)};p+1\big)^{jk,lm} \\
&\qquad -\Xi(c^n_i,\gamma^n_i;p+1)^{jk,lm}|\mathcal{F}^n_i\Big]\Big|
\end{aligned}$$

$$\xi_i^{n,5} = \frac{1}{k_n \Delta_n^{1/2}} \left| \frac{(p+2)\theta}{k_n \Delta_n^{1/2}} \left\lfloor \frac{k_n}{(p+1)l_n} \right\rfloor - 1 \right| \left| \Xi(c_i^n, \gamma_i^n; p+1)^{jk,lm} \right| + \frac{1}{k_n \Delta_n^{1/2}} \left| \Xi(c_i^n, \gamma_i^n; p+1)^{jk,lm} - \Xi(c_i^n, \gamma_i^n)^{jk,lm} \right|.$$

By (3.43), $\epsilon(p)_i^n \leq Kp\big[(k_n^2\Delta_n)^{-1} + pk_n^{-1}\Delta_n^{1/4} \sum_{h=0}^{m(n,p)-1} \overline{\lambda}(p+1)_{i+a(n,p,h)}^n\big]$ is the edge term.

- by (3.42), $\xi_i^{n,1} \leq Kp^{1/2}k_n^{-1/2}$;
- by (3.44) and (3.45), $\xi_i^{n,2} \leq K\Big[k_n\Delta_n + p^{1/2}k_n^{-1/2} + p\Delta_n^{3/4} \sum_{h=0}^{m(n,p)-1} \overline{\lambda}(p+1)_{i+a(n,p,h)}^n\Big]$;
- by (3.37), $\xi_i^{n,3} \leq K\big[p(k_n\Delta_n^{1/4})^{-1} + p^2(k_n^2\Delta_n)^{-1} \sum_{h=0}^{m(n,p)-1} \overline{\lambda}(p+1)_{i+a(n,p,h)}^n\big]$;
- by **Assumption A-**c, **A-**γ, **SA-**ν, $\xi_i^{n,4} \leq K\Delta_n^{1/2}$; by (3.34), $\xi^{n,5} \leq K(pk_n\Delta_n^{1/2})^{-1}$.

Combine these upper bounds on $\xi_i^{n,r}$, $r = 1,2,3,4,5$ and $\epsilon(p)_i^n$, this claim can be justified.

□

For the purpose of correcting bias and quantifying uncertainty, we need to estimate $\Xi(c_i^n, \gamma_i^n)$'s. I use an plug-in estimator - plug in the spot estimates of volatility and noise covariance into function Ξ.

Formally, we estimate $\Xi(c_i^n, \gamma_i^n)$ by $\Xi(\widehat{c}_i^{*n}, \widehat{\gamma}_i^n)$, where $\widehat{c}_i^{*n}$ is defined in (3.51) and

$$\widehat{\gamma}_i^n = \frac{1}{2m_n} \sum_{h=1}^{m_n} \Delta_{i+h}^n Y \cdot \Delta_{i+h}^n Y^{\mathrm{T}}, \tag{3.46}$$

with $m_n = \lfloor \theta' \Delta_n^{-1/2} \rfloor$, θ' positive finite. To study the accuracy of this estimator, we define

$$\kappa_i^n = \Xi(\widehat{c}_i^{*n}, \widehat{\gamma}_i^n) - \Xi(c_i^n, \gamma_i^n), \tag{3.47}$$

and we have the follow lemma regarding κ_i^n.

Lemma 19.

$$E(\|\kappa_i^n\|^q|\mathcal{F}_i^n) \le K_q\Big[(k_n\Delta_n)^{q/2} + (k_n\Delta_n^{1/2})^{-q/2}\Big], \quad q=1,2.$$

Proof. Let $\chi_i^n = \widehat{\gamma}_i^n - \gamma_i^n$, we have

$$\chi_i^n = \frac{1}{2m_n}\sum_{h=1}^{m_n}\eta_{i+h}^n + \frac{1}{m_n}\sum_{h=1}^{m_n}(\gamma_{i+h}^n - \gamma_i^n) + \frac{1}{2m_n}(\gamma_i^n - \gamma_{i+m_n}^n),$$

where $\eta_i^n = \Delta_i^n Y \Delta_i^n Y^{\mathrm{T}} - \gamma_i^n - \gamma_{i-1}^n$. Note

$$\Big\|\sum_{h=1}^{m_n}\eta_{i+h}^n\Big\|^2 = \sum_{h=1}^{m_n}\|\eta_{i+h}^n\|^2 + 2\sum_{h=1}^{m_n-1}\|\eta_{i+h}^n\eta_{i+h+1}^n\| + 2\sum_{h=1}^{m_n-2}\sum_{h'=h+2}^{m_n}\|\eta_{i+h}^n\eta_{i+h'}^n\|.$$

By (3.16),

$$E\Big(\Big\|\sum_{h=1}^{m_n}\eta_{i+h}^n\Big\|^2\Big|\mathcal{F}_i^n\Big) \le K\big(m_n + m_n^2\Delta_n^2\big),$$

by **Corollary 3**, the choice of m_n and Jensen's inequality

$$E(\|\chi_i^n\|^q|\mathcal{F}_i^n) \le K_q\Delta_n^{q/4},\ q=1,2, \tag{3.48}$$

by (3.35) and (3.36),

$$\|\kappa_i^n\| \le K\big(\|\beta_i^n\| + \|\chi_i^n\| + \|\beta_i^n\|^2 + \|\beta_i^n\|\|\chi_i^n\| + \|\chi_i^n\|^2\big),$$

hence this lemma can be checked by **Lemma 18** and (3.48). □

3.3 Pre-averaging for Itô semimartingales

If the Itô semimartingale is driven by a jump process of finite activity besides the Brownian motion, we can identify the intervals in which jump occurred[3]. For the purpose of estimating volatility, we can discard those intervals that contain jumps. A natural approach is to select a threshold value $\nu_n = \alpha\Delta_n^{\rho}$, where α is positive finite and possibly depends on the local volatility level, and $\rho < 1/2$ by classic estimates in section 2.3 to keep the Brownian increments (diffusion components).

I define pre-averaged spot volatility estimator[4] for general Itô semimartingale model as

$$\widehat{c}_i^n(k_n, l_n, \nu_n; \varphi) = \frac{1}{(k_n - l_n)\Delta_n} \sum_{h=1}^{k_n - l_n + 1} \Big[\overline{Y}_{i+h}^n(l_n; \varphi) \cdot \overline{Y}_{i+h}^n(l_n; \varphi)^{\mathrm{T}} \mathbb{1}_{\{\|\overline{Y}_{i+h}^n(l_n;\varphi)\| \leq \nu_n\}} - \widehat{Y}_{i+h}^n(l_n; \varphi)\Big], \quad (3.49)$$

where

$$\nu_n \asymp \Delta_n^{\rho} \quad (3.50)$$

with some suitable choice of $\rho < 1/2$.

The second term on the right hand side of (3.49) is for bias-correction. It is negligible in the first-order asymptotics due to noise, hence a "small order" correction. It does not matter to

3. However, the problem of positioning jumps are much more difficult, depending on the jump activity of the underlying process (Blumenthal-Getoor index). In the simplest case where there are finite jumps on a bounded interval, one can easily separate the jumps from the continuous evolution. If the jump activity is infinite, the matter is complicate to discuss and beyond the discussion of estimating spot volatility. Nonetheless, for the purpose of identifying the intervals containing jumps (single or multiple jumps) and estimating volatility and the functionals, the intuition described here would suffice to guide us to the right path.

4. There are a few slightly different versions of pre-averaging estimators in the extant literature. For the sake of generality, we adapt the convention from Jacod et al. [2009, 2010], Christensen et al. [2010]. A special usage example of pre-averaging can be found in Podolskij and Vetter [2009b], Wang and Mykland [2014].

the consistency, so one can also use

$$\widehat{c}_i^n(k_n, l_n; \varphi) \equiv \frac{1}{(k_n - l_n)\Delta_n} \sum_{h=1}^{k_n - l_n + 1} \bar{Y}_{i+h}^n(l_n;\varphi) \cdot \bar{Y}_{i+h}^n(l_n;\varphi)^{\mathrm{T}} \mathbb{1}_{\{\|\bar{Y}_{i+h}^n(l_n;\varphi)\| \le \nu_n\}}. \tag{3.51}$$

The advantage of the estimator (3.51) is that the estimator is guaranteed to be positive semi-definite in finite sample.

Lemma 20. *Under **Assumption A-**γ and **Assumption SA-**ν with $\nu \in [0,1)$, choose l_n such that $l_n\Delta_n^{1/2} = \theta + o(\Delta_n^{1/4})$, we have*

$$\begin{aligned} E\left(\|\bar{Y}_i^{*n}\|^q | \mathcal{F}_i^n\right) &\le K_q \Delta_n^{q/2} \\ E\left(\|\bar{Y}_i^n\|^q | \mathcal{F}_i^n\right) &\le K_q \Delta_n^{(q/2)\wedge(q/4+1/2)} \\ E\left(\|\widehat{Y}_i^{*n}\|^q \vee \|\widehat{Y}_i^n\|^q | \mathcal{F}_i^n\right) &\le K_q \Delta_n^q \\ E\left[\left(\frac{\|\bar{J}_i^n\|}{\Delta_n^w} \wedge 1\right)^q | \mathcal{F}_i^n\right] &\le K_q\, a_n\, \Delta_n^{[1/2-(w-1/4)\nu]\times[1\wedge(q/\nu)]}, \end{aligned}$$

for some $a_n \to 0$.

Proof. We show these inequalities case by case.

(1)&**(2)** We define $\bar{X}_i'^n$, $\bar{\varepsilon}_i^n$, $\bar{J}_i^n$ according to (3.4), which are associated with process X', ε, J respectively. By (3.26), $E\big(\|\bar{X}_i'^n\|^q|\mathcal{F}_i^{(0),n}\big) \le K_q\Delta_n^{q/2}$; by **Theorem 1**, $E\big(\|\bar{\varepsilon}_i^n\|^q|\widetilde{\mathcal{F}}_i^n\big) \le K_q l_n^{-q}$; we can write $\bar{J}_i^n = (\psi_0 l_n)^{-1/2} J_{i\Delta_n}^n(\varphi, (l_n-1)\Delta_n)$ according to (3.17), by **Lemma 5**, $E\big(\|\bar{J}_i^n\|^q|\mathcal{F}_i^{(0),n}\big) \le K_q l_n^{1-q/2}\Delta_n$. Notice $\bar{Y}_i^{*n} = \bar{X}_i'^n + \bar{\varepsilon}_i^n$, $\bar{Y}_i^n = \bar{Y}_i^{*n} + \bar{J}_i^n$, the first 2 cases can be verified.

(3) We define $\widehat{X}_i'^n$, $\widehat{\varepsilon}_i^n$, $\widehat{J}_i^n$ according to (3.5), which are associated with process X', ε, J respectively. By **Lemma 5**, **Corollary 3**, $E\big(\|\widehat{X}_i'^n\|^q|\mathcal{F}_i^{(0),n}\big) \le K_q l_n^{-2q}\Delta_n^q$, $E\big(\|\widehat{J}_i^n\|^q|\mathcal{F}_i^{(0),n}\big) \le K_q l_n^{-2q}\Delta_n$; besides, $E\big(\|\widehat{\varepsilon}_i^n\|^q|\widetilde{\mathcal{F}}_i^n\big) \le K_q l_n^{-2q}$. Notice $\|\widehat{Y}_i^{*n}\| \le K\big(\|\widehat{X}_i'^n\| + \|\widehat{\varepsilon}_i^n\|\big)$, $\|\widehat{Y}_i^n\| \le K\big(\|\widehat{X}_i'^n\| + \|\widehat{\varepsilon}_i^n\| + \|\widehat{J}_i^n\|\big)$, this case is checked.

(4) Note that

$$\frac{\|\overline{J}_i^n\|}{\Delta_n^w} = \frac{J_{i\Delta_n}^n(\varphi,(l_n-1)\Delta_n)}{(\psi_0\theta)^{1/2}\Delta_n^{w-1/4}} = \frac{J_{i\Delta_n}^n(\varphi,(l_n-1)\Delta_n)}{\psi_0^{1/2}\theta^{1-2w}(l_n\Delta_n)^{2w-1/2}}$$

then this case can be checked by **Corollary 1**. □

Lemma 21. *Under the definition (3.51) and (3.49), when $\nu < 1$, there is a sequence $a_n \to 0$ such that*

$$E\left(\|\widehat{c}_i^n - \widehat{c}_i^{*n}\|^q | \mathcal{F}_i^n\right) \le K_q\Big[a_n \Delta_n^{1/2-(\rho-1/4)\nu-(1-2\rho)q} + \Delta_n^{1/2}\Big].$$

Proof. We have

$$\left\|\left(\overline{Y}_i^n \cdot \overline{Y}_i^{n,\mathrm{T}} \mathbb{1}_{\{\|\overline{Y}_i^n\| \le \nu_n\}} - \widehat{Y}_i^n\right) - \left(\overline{Y}_i^{*n} \cdot \overline{Y}_i^{*n,\mathrm{T}} - \widehat{Y}_i^{*n}\right)\right\| \le \eta_i^n(1) + \eta_i^n(2) + \eta_i^n(3),$$

where

$$\begin{aligned}
\eta_i^n(1) &= \left\|\overline{Y}_i^n \cdot \overline{Y}_i^{n,\mathrm{T}} \mathbb{1}_{\{\|\overline{Y}_i^n\| \le \nu_n\}} - \overline{Y}_i^{*n} \cdot \overline{Y}_i^{*n,\mathrm{T}} \mathbb{1}_{\{\|\overline{Y}_i^{*n}\| \le \nu_n\}}\right\| \\
\eta_i^n(2) &= \left\|\widehat{Y}_i^n - \widehat{Y}_i^{*n}\right\| \\
\eta_i^n(3) &= \left\|\overline{Y}_i^{*n}\right\|^2 \mathbb{1}_{\{\|\overline{Y}_i^{*n}\| > \nu_n\}}.
\end{aligned}$$

(1) estimate of $\eta_i^n(1)$

Consider $z = (y+v)(y+v)^{\mathrm{T}} \mathbb{1}_{\{\|y+v\| \le u_n\}} - yy^{\mathrm{T}} \mathbb{1}_{\{\|y\| \le u_n\}}$, we have the following 4 cases:

$$\begin{aligned}
\|y+v\| \vee \|y\| \le u_n &\;\Rightarrow\; z \asymp (\|y\|\|v\| + \|v\|^2)\mathbb{1}_{\{\|v\| \le 2u_n\}} \\
\|y+v\| \le u_n < \|y\| &\;\Rightarrow\; z < K\|y\|^2 \left(\|y\|/u_n\right)^{2/(1-2\rho)} \\
u_n/2 < \|y\| \le u_n < \|y+v\| &\;\Rightarrow\; z \asymp \|y\|^2 \left(\|y\|/u_n\right)^{2/(1-2\rho)} \\
2\|y\| \le u_n < \|y+v\| &\;\Rightarrow\; z \le K u_n^2 \mathbb{1}_{\{u_n < 2\|v\|\}},
\end{aligned}$$

hence

$$\|z\| \le u_n^{-2/(1-2\rho)} \|y\|^{2+2/(1-2\rho)} + (1+\|y\|)\big(\|v\| \wedge 1 + \|v\|^2 \wedge u_n^2\big).$$

Let $u_n = \nu_n/\Delta_n^{1/2}$, $y = \bar{Y}_i^{*n}/\Delta_n^{1/2}$ and $v = \bar{J}_i^n/\Delta_n^{1/2}$, we have

$$\|z\| \le u_n^{-2/(1-2\rho)} (Z_i^n)^{2+2/(1-2\rho)} + (1+Z_i^n)\big[Q_i^n + u_n^2 (V_i^n)^2\big],$$

where $Z_i^n = \|\bar{Y}_i^{*n}\|/\Delta_n^{1/2}$, $Q_i^n = \big(\|\bar{J}_i^n\|/\Delta_n^{1/2}\big) \wedge 1$, $V_i^n = \big(\|\bar{J}_i^n\|/\Delta_n^{\rho}\big) \wedge 1$.

Let $\mathcal{H}$ be the σ-algebra generated by random variables $\mathfrak{p}(A)$ with $A \in \mathcal{B}(\mathbb{R}^+ \times \mathbb{R}^d)$ being Borel sets. Let $(\bar{\mathcal{F}}_t)$ be the smallest filtration such that $\widetilde{\mathcal{F}}_t \subset \bar{\mathcal{F}}_t$ and $\mathcal{H} \subset \bar{\mathcal{F}}_t$. By **Lemma 20**,

$$\begin{aligned}
E\big[(Z_i^n)^q | \bar{\mathcal{F}}_i^n\big] &\le K_q \\
E\big[(Q_i^n)^q | \mathcal{F}_i^n\big] &\le K_q\, \Delta_n^{(1/2-\nu/4)\times[1\wedge(q/\nu)]}\, a_n \\
E\big[(V_i^n)^q | \mathcal{F}_i^n\big] &\le K_q\, \Delta_n^{[1/2-(\rho-1/4)\nu]\times[1\wedge(q/\nu)]}\, a_n,
\end{aligned}$$

where $a_n \to 0$. In case $\nu < 1$, $q \ge 1$, note $\eta_i^n(1) = \|z\|\Delta_n$, by successive conditioning on $\bar{\mathcal{F}}_i^n$ and then $\mathcal{F}_i^n$,

$$\begin{aligned}
E\big[(\eta_i^n(1))^q | \mathcal{F}_i^n\big] \le K_q \Delta_n^q \Big[\Delta_n^q + \Big(\Delta_n^{1/2-\nu/4} + \Delta_n^{(2\rho-1)q+1/2-(\rho-1/4)\nu}\Big)\, a_n\Big] \\
\le K_q \Delta_n^{2\rho q+1/2-(\rho-1/4)\nu}\, a_n.
\end{aligned}$$

(2) estimate of $\eta_i^n(2)$

Note $\widehat{Y}_i^n - \widehat{Y}_i^{*n} = (\psi_0 l_n)^{-1} \sum_{h=0}^{l_n-1} (\varphi_{h+1}^n - \varphi_h^n)^2 e_{i+h}^n$, where $e_i^n = \Delta_i^n X' \Delta_i^n J^{\mathrm{T}} + \Delta_i^n J \Delta_i^n J^{\mathrm{T}}/2 + \Delta_i^n \varepsilon \Delta_i^n J^{\mathrm{T}}$. By Cauchy-Schwarz inequality, **Lemma 5**, **Corollary 3**, $E\big(\|e_i^n\|^q | \mathcal{F}_i^n\big) \le K_q \Delta_n^{1/2}$ for $q \ge 1$, therefore $E\big[(\eta_i^n(2))^q | \mathcal{F}_i^n\big] \le K_q \Delta_n^{q+1/2}$.

(3) estimate of $\eta_i^n(3)$

By Markov's inequality and **Lemma 20**,

$$P\big(\|\bar{Y}_i^{*n}\| > \nu_n|\mathcal{F}_i^n\big) \le E\big(\|\bar{Y}_i^{*n}\|^{2q'}|\mathcal{F}_i^n\big)(\alpha^{2q'}\Delta_n^{2\rho q'})^{-1} \asymp \Delta_n^{(1-2\rho)q'}. \tag{3.52}$$

By Cauchy-Schwarz inequality and (3.52),

$$E\big[(\eta_i^n(3))^q|\mathcal{F}_i^n\big] \le K_{q,q'}E\left(\|\bar{Y}_i^{*n}\|^{4q}|\mathcal{F}_i^n\right)^{1/2}\Delta_n^{(1-2\rho)q'/2} \le K_{q,q'}\Delta_n^{q+(1/2-\rho)q'}.$$

Combine estimates of $\eta_i^n(m), m = 1,2,3$ with sufficiently large q' for $\eta_i^n(3)$, we can deduce this lemma. □

Define $\overline{c}_i^n = (k_n - l_n)^{-1}\sum_{h=1}^{k_n-l_n+1} c_i^n$, we have an uniform convergence result for $\widehat{c}_i^n \xrightarrow{\mathbb{P}} \overline{c}_i^n$.

Lemma 22. *Assume* ***Assumption A-****c,* ***A-****γ,* ***A-****ν, $k_n \asymp \Delta_n^\kappa$ with $\kappa \in \left(\frac{2}{3} \vee \frac{2+\nu}{4}, \frac{3}{4}\right)$, and in (3.50) $\rho \in [1/4 + (1-\kappa)/(2-\nu), 1/2)$, then*

$$\sup_{i\in I_n}\|\widehat{c}_i^n - \overline{c}_i^n\| = o_p(1),$$

where $I_n = \{0, \cdots, N_t^n - 1\}$.

Proof. We prove this lemma in 2 steps.

(1) Show $\sup_{i\in I_n}\|\widehat{c}_i^n - \widehat{c}_i^{*n}\| \xrightarrow{\mathbb{P}} 0$.

Note that $|I_n| \asymp (k_n\Delta_n)^{-1}$, according to **Lemma 21**, there is a sequence $a_n \to 0$ such that

$$E\left(\sup_{i\in I_n}\|\widehat{c}_i^n - \widehat{c}_i^{*n}\||\mathcal{F}_i^n\right) \le K\Big[a_n\Delta_n^{\kappa-1/2-(\rho-1/4)\nu-(1-2\rho)} + \Delta_n^{\kappa-1/2}\Big],$$

by the condition on κ, we get the result in this step.

(2) Show $\sup_{i\in I_n}\|\widehat{c}_i^{*n} - \overline{c}_i^n\| \xrightarrow{\mathbb{P}} 0$.

Note that $\widehat{c}_i^{*n} - \overline{c}_i^n = \xi(1)_i^n + \xi(2)_i^n + N(p)_i^n + M(p)_i^n$, by **Lemma 7**, **8**, **17** and $\kappa < 3/4$, $E\big(\|\widehat{c}_i^{*n} - \overline{c}_i^n\|^4 \|\mathcal{F}_i^n\big) \leq K\Delta_n^{2\kappa-1}$, so

$$E\left(\sup_{i\in I_n} \|\widehat{c}_i^{*n} - \overline{c}_i^n\| |\mathcal{F}_i^n\right) \leq K\Delta_n^{3\kappa-2},$$

since $\kappa > 2/3$, we accomplish this step. Combine these 2 steps, this lemma is verified. $\square$

CHAPTER 4

INFERENCE OF VOLATILITY MATRIX FUNCTIONALS BY PRE-AVERAGING METHOD

4.1 Functional estimation when noise matters

Now we are ready to study the functional inference problem of (1.1). We assume the following smoothness property of the matrix functional:

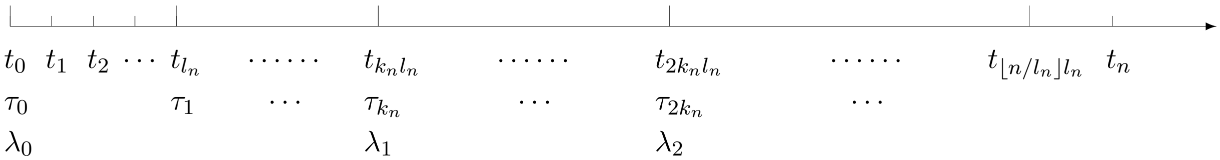
$$\|\partial^h g(c)\| \leq K(1 + \|c\|^{r-h}), \qquad h = 0, 1, 2, 3. \tag{4.1}$$

Jacod and Rosenbaum [2013] studied the same problem when the observations do not suffer any noise, so did Li et al. [2019]. What if our high-frequency observations are noisy? We can use pre-averaging! We compute the instantaneous covariance matrices by pre-averaging, and plug them in the functional of interest and apply necessary bias correction.

Overlapping intervals would result in more efficiency than the non-overlapping alternative when we use the pre-averaging method to estimate covariance matrices, while overlapping intervals will not lead to any asymptotic efficiency gain in estimating the integral by Riemann sums. Taking this fact into account, it is reasonable to use overlapping intervals to estimate instantaneous covariances and use non-overlapping intervals as building blocks for functional estimators.

If one uses non-overlapping smoothing windows[1], the corresponding sampling scheme can be depicted as the following.

1. Note that, we will use overlapping window of length l_n to improve efficiency and our choice of sampling scheme is not as the picture describes. This sampling scheme pictured above is only meant to demonstrate the double moving windows in my methodology.



l_n is the window length (in terms of observations) for kernel smoothing; k_n is the window length (in terms of local smoothing windows) for volatility estimation. $\{t_i = i\Delta_n\}_{i=0,\cdots,n}$ is the set of time points where noisy high-frequency data is observed; the local windows for de-noising are obtained via partition points τ_i, $i = 0, \cdots, \lfloor n/(l_nk_n)\rfloor\, k_n$; the time points λ_i, $i = 0, \cdots, \lfloor n/(l_nk_n)\rfloor$ are those on which spot volatilities are to be estimated.

Mentioned in section 3.1, the spot volatility estimate based on the data in the interval $I(n, ik_n, k_n) = (ik_n\Delta_n, (i+1)k_n\Delta_n]$ is:

$$\widehat{c}^n_{ik_n} = \frac{1}{(k_n - l_n)\Delta_n} \sum_{j=1}^{k_n-l_n+1} \left[\overline{Y}^n_{ik_n+j}(l_n;\varphi) \cdot \overline{Y}^n_{ik_n+j}(l_n;\varphi)^{\mathrm{T}} \mathbb{1}_{\{\|\overline{Y}^n_{ik_n+j}(l_n;\varphi)\| \le \nu_n\}} \right. \\ \left. - \widehat{Y}^n_{ik_n+j}(l_n;\varphi) \right],$$

where $i = 0, 1, 2, \cdots, \lfloor t/(k_n\Delta_n)\rfloor - 1$, and $\overline{Y}^n_i(l_n;\varphi)$, $\widehat{Y}^n_i(l_n;\varphi)$ are defined in (3.4), (3.5), $\nu_n = \alpha\Delta_n^\rho$, $\alpha > 0, \rho \in [0, 1/2)$. The estimator of integrated volatility functional g over $[0, t]$ is defined as

$$\widehat{S}(g)^n_t \equiv k_n\Delta_n \sum_{i=0}^{N^n_t-1} \left[g(\widehat{c}^n_{ik_n}) - \frac{1}{2k_n\Delta_n^{1/2}} \sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}\sum_{m=1}^{d} \partial^2_{jk,lm} g(\widehat{c}^n_{ik_n}) \times \Xi\big(\widehat{c}^n_{ik_n}, \widehat{\gamma}^n_{ik_n}\big)^{jk,lm} \right], \tag{4.2}$$

where $N^n_t = \lfloor t/(k_n\Delta_n)\rfloor$ and $\Xi(c, \gamma)$ is defined in (3.36).

There are three tuning parameters in the spot estimator $\widehat{c}^n_h$ and the functional estimator $\widehat{S}(g)^n_t$. They are summarized in Table 4.1.

tuning parameters	description
l_n	length of overlapping window for local moving averages
k_n	length of disjoint window for spot volatility estimation
ν_n	truncation level for jumps

Table 4.1: Summary and description of tuning parameters in the pre-averaging method

4.2 Jacod's theorem

A groundbreaking landmark and widely used probabilistic result for deriving central limit theorems associated with discretized stochastic processes is the following theorem applicable to discretized processes. Literature calls the result *stable convergence to a progressive conditional continuous process of independent increments*. It is adapted from Theorem IX.7.28 in Jacod and Shiryaev [2003].

Theorem 2. *Suppose M is a $\mathbb{R}^d$-valued continuous local martingale satisfying $E\|M_t\| \leq \infty$, $\forall t \geq 0$, and the components in the triangular array $(\chi_i^n)_{1\leq i\leq n,\, n\in\mathbb{N}}$ are square integrable, and χ_i^n is $\mathcal{F}_i^n$-measurable for each i.*

Assume C, G, B are continuous, adapted processes on the probability space $\mathcal{B} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$, and they take values $\mathcal{M}_+^p$, $\mathbb{R}^{p\times d}$, $\mathbb{R}^p$, respectively. Furthermore, $\forall t > 0$, $\forall \varepsilon > 0$, $1 \leq j, k \leq p$, $1 \leq l \leq d$, and $\forall$ bounded martingale N on the same probability space which is orthogonal to all components of the process M, assume

$$\sup_{t\geq 0} \left| \sum_{i=1}^{\lfloor nt/T \rfloor} E(\chi_i^n | \mathcal{F}_{i-1}^n) - B_t \right| \xrightarrow{\mathbb{P}} 0 \tag{4.3}$$

$$\sum_{i=1}^{\lfloor nt/T \rfloor} \left[E\big(\chi_i^{n,j} \chi_i^{n,k} | \mathcal{F}_{i-1}^n\big) - E\big(\chi_i^{n,j} | \mathcal{F}_{i-1}^n\big) E\big(\chi_i^{n,k} | F_{i-1}^n\big) \right] \xrightarrow{\mathbb{P}} C_t^{jk} \tag{4.4}$$

$$\sum_{i=1}^{\lfloor nt/T \rfloor} E\big(\chi_i^{n,j} \Delta_i^n M^l | \mathcal{F}_{i-1}^n\big) \xrightarrow{\mathbb{P}} G_t^{jl} \tag{4.5}$$

$$\sum_{i=1}^{\lfloor nt/T \rfloor} E\left(|\chi_i^n|^2 \mathbb{1}_{\{|\chi_i^n|>\varepsilon\}} | \mathcal{F}_{i-1}^n\right) \xrightarrow{\mathbb{P}} 0 \tag{4.6}$$

$$\sum_{i=1}^{\lfloor nt/T \rfloor} E\left(\chi_i^n \Delta_i^n N | \mathcal{F}_{i-1}^n\right) \xrightarrow{\mathbb{P}} 0, \tag{4.7}$$

then $\exists$ an extension $\widetilde{\mathcal{B}}$ of the probability space $\mathcal{B}$ and a M-biased $\mathcal{F}$-progressive conditional continuous martingale $\widetilde{X}$ with independent increments on this extension satisfying $\forall t > 0$

$$\begin{aligned} \langle \widetilde{X}, \widetilde{X} \rangle_t &= C_t \\ \langle \widetilde{X}, M \rangle_t &= G_t, \end{aligned}$$

and such that

$$\sum_{i=1}^{\lfloor nt/T \rfloor} \chi_i \xrightarrow{\mathcal{L}-s} B + \widetilde{X}_t.$$

4.3 Central limit theorems

4.3.1 decomposition

By Cramér-Wold theorem, we can suppose g is a $\mathbb{R}$-valued function without loss of generality.

Based on 2nd-order Taylor expansion, we can decompose the estimation error of volatility functional in the following way:

$$\Delta_n^{-1/4}\left[\widehat{S}(g)^n - S(g)\right] = \overline{V}^{n,0} + \overline{V}^{n,1} + \overline{V}^{n,2} + \overline{V}^{n,3} + \overline{V}(p)^{n,4}, \tag{4.8}$$

where

$$\overline{V}_t^{n,0} \;=\; \Delta_n^{-1/4}\left[\sum_{i=0}^{N_t^n-1}\int_{I(n,ik_n,k_n)} g(c_{ik_n}^n)-g(c_s)\,\mathrm{d}s-\int_{N_t^nk_n\Delta_n}^{t} g(c_s)\,\mathrm{d}s\right]$$

$$\begin{aligned}
\overline{V}_t^{n,1} \;&=\; k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\Bigg[g(\widehat{c}_{ik_n}^n)-g(\widehat{c}_{ik_n}^{*n})-\frac{1}{2k_n\Delta_n^{1/2}}\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}\sum_{m=1}^{d}\\
&\qquad\Big(\partial_{jk,lm}^2 g(\widehat{c}_{ik_n}^n)\times\Xi(\widehat{c}_{ik_n}^n,\widehat{\gamma}_{ik_n}^n)^{jk,lm}-\partial_{jk,lm}^2 g(\widehat{c}_{ik_n}^{*n})\times\Xi(\widehat{c}_{ik_n}^{*n},\widehat{\gamma}_{ik_n}^n)^{jk,lm}\Big)\Bigg]\\
\overline{V}_t^{n,2} \;&=\; k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\sum_{j=1}^{d}\sum_{k=1}^{d}\partial_{jk}g(c_{ik_n}^n)\left[\xi(0)_{ik_n}^{n,jk}+\xi(1)_{ik_n}^{n,jk}+\xi(2)_{ik_n}^{n,jk}\right]\\
\overline{V}_t^{n,3} \;&=\; k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\Bigg[g(\widehat{c}_{ik_n}^{*n})-g(c_{ik_n}^n)-\sum_{j=1}^{d}\sum_{k=1}^{d}\partial_{jk}g(c_{ik_n}^n)(\widehat{c}_{ik_n}^{*n,jk}-c_{ik_n}^{n,jk})\\
&\qquad-\frac{1}{2k_n\Delta_n^{1/2}}\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}\sum_{m=1}^{d}\partial_{jk,lm}^2 g(\widehat{c}_{ik_n}^{*n})\times\Xi(\widehat{c}_{ik_n}^{*n},\widehat{\gamma}_{ik_n}^n)^{jk,lm}\Bigg]
\end{aligned}$$

$$\overline{V}(p)_t^{n,4}=k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\sum_{j=1}^{d}\sum_{k=1}^{d}\partial_{jk}g(c_{ik_n}^n)\times\left[N(p)_{ik_n}^{n,jk}+M(p)_{ik_n}^{n,jk}\right].$$

4.3.2 asymptotic negligibility

In this subsection, we show $\overline{V}^{n,m}\xrightarrow{u.c.p.}0$, $m=0,1,2,3$. One recurring intermediate result I rely upon is the following maximal inequality.

Lemma 23. *Let $Z_i, i=1,\cdots,M$ be random variables and $\mathcal{H}_i=\sigma(Z_1,\cdots,Z_i)$ be the σ-algebra generated by $Z_1,\cdots,Z_i$, then*

$$\mathbb{E}\left(\sup_{m=1,\cdots,M}\left\|\sum_{i=1}^{m}\left[Z_i-E\left(Z_i|\mathcal{H}_i\right)\right]\right\|\right)\le K\Big(\sum_{i=1}^{M}\mathbb{E}\big(\|Z_i\|^2\big)\Big)^{1/2}.$$

Proof. Since $\{\mathcal{H}_i\}$ is the natural filtration of $\{Z_i\}$, $\{\sum_i^m [Z_i - E(Z_i|\mathcal{H}_i)]\}_m$ is a martingale. By Jensen's inequality and Doob's maximal inequality,

$$\mathbb{E}\left(\sup_{m=1,\cdots,M}\left\|\sum_{i=1}^m [Z_i - E(Z_i|\mathcal{H}_i)]\right\|\right) \leq \mathbb{E}\left(\sup_{m=1,\cdots,M}\left\|\sum_{i=1}^m [Z_i - E(Z_i|\mathcal{H}_i)]\right\|^2\right)^{1/2}$$
$$\leq \left(4\sup_{m=1,\cdots,M}\mathbb{E}\left(\left\|\sum_{i=1}^m [Z_i - E(Z_i|\mathcal{H}_i)]\right\|^2\right)\right)^{1/2} \leq K\Big(\sum_{i=1}^M \mathbb{E}\Big(\|Z_i\|^2\Big)\Big)^{1/2}.$$

$\square$

Lemma 24. *If we let $k_n\Delta_n^{3/4} \to 0$,*

$$\bar{V}^{n,0} \xrightarrow{u.c.p.} 0.$$

Proof. One can write

$$\bar{V}_t^{n,0} = A_t^n + \Delta_n^{-1/4}\int_{N_t^n k_n\Delta_n}^t g(c_s)\,\mathrm{d}t, \tag{4.9}$$

where $A_t^n = -\sum_{i=0}^{N_t^n-1}\left(\theta_{ik_n}^n + \theta_{ik_n}'^n\right)$ and

$$\begin{cases} \theta_{ik_n}^n &= \Delta_n^{-1/4}\int_{I(n,ik_n,k_n)}\left[g(c_u) - g(c_{ik_n}^n) - \sum_{j=1}^d\sum_{k=1}^d \partial_{jk}g(c_{ik_n}^n)(c_s^{jk} - c_{ik_n}^{n,jk})\right]\mathrm{d}s \\ \theta_{ikn}'^n &= \Delta_n^{-1/4}\sum_{j=1}^d\sum_{k=1}^d \partial_{jk}g(c_{ik_n}^n)\int_{I(n,ik_n,k_n)}(c_s^{jk} - c_{ik_n}^{n,jk})\,\mathrm{d}s. \end{cases}$$

Due to the fact that $k_n\Delta_n^{3/4} \to 0$ and $\Delta_n^{-1/4}\|\int_{N_t^n k_n\Delta_n}^t g(c_s)\,\mathrm{d}t\| \leq Kk_n\Delta_n^{3/4}$, it suffices to show $A^n \xrightarrow{u.c.p.} 0$. By assumption (4.1), $\|\theta_{ik_n}^n\| \leq K\Delta_n^{-1/4}\int_{I(n,ik_n,k_n)}\|c_s - c_i^n\|^2\,\mathrm{d}s$, by Fubini theorem $E\|\theta_{ik_n}^n\| \leq Kk_n^2\Delta_n^{7/4}$. By **Corollary 3**, $\left\|E(\theta_{ik_n}'^n|\mathcal{F}_{ik_n}^{(0),n})\right\| \leq Kk_n^2\Delta_n^{4/7}$. By

Cauchy-Schwarz inequality and **Corollary 3** again,

$$E\Big(\|\theta'^n_{ik_n}\|^2\Big|F^{(0),n}_{ik_n}\Big) \le K\Delta_n^{-1/2}\sup_{j,k=1,\cdots,d} E\left(\left|\int_{I(n,ik_n,k_n)} c_s^{jk} - c^{n,jk}_{ik_n}\,\mathrm{d}s\right|^2 \Big|F^{(0),n}_{ik_n}\right)$$

$$\le K\Delta_n^{-1/2}\sup_{j,k=1,\cdots,d} E\left(\Delta_n\int_{I(n,ik_n,k_n)} (c_s^{jk} - c^{n,jk}_{ik_n})^2\,\mathrm{d}s\Big|\mathcal{F}^{(0),n}_{ik_n}\right) \le Kk_n^2\Delta_n^{5/2}.$$

Note that

$$\sup_{s\in[0,t]}\|A^n_t\| \le \sum_{i=0}^{N^n_t-1}\|\theta^n_{ik_n}\| + \sum_{i=0}^{N^n_t-1}\big\|E\big(\theta'^n_{ik_n}|\mathcal{F}^{(0),n}_{ik_n}\big)\big\| + \sup_{s\in[0,t]}\left\|\sum_{i=0}^{N^n_s-1}\Big[\theta'^n_{ik_n} - E\big(\theta'^n_{ik_n}|\mathcal{F}^{(0),n}_{ik_n}\big)\Big]\right\|,$$

because $\Big\{\theta'^n_{ik_n} - E\Big(\theta'^n_{ik_n}|\mathcal{F}^{(0),n}_{ik_n}\Big)\Big\}_i$ is a martingale difference sequence w.r.t. $\big(\mathcal{F}^{(0),n}_{ik_n}\big)_i$, based on **Lemma 23**,

$$\sup_{s\in[0,t]}\left\|\sum_{i=0}^{N^n_s-1}\Big[\theta'^n_{ik_n} - E\Big(\theta'^n_{ik_n}|\mathcal{F}^{(0),n}_{ik_n}\Big)\Big]\right\| \le 2\left(\sum_{i=0}^{N^n_t-1}\mathbb{E}\big(\|\theta'^n_i\|^2\big)\right)^{1/2} \le K\sqrt{t}k_n^{1/2}\Delta_n^{3/4},$$

thus $\mathbb{E}\big(\sup_{s\in[0,t]}\|A^n_s\|\big) \le K\big(tk_n\Delta_n^{3/4} + \sqrt{t}k_n^{1/2}\Delta_n^{3/4}\big)$, thus $A^n \xrightarrow{u.c.p.} 0$ by Markov inequality.

Therefore, $\bar{V}^{n,0} \xrightarrow{u.c.p.} 0$. □

Lemma 25. *If $\nu_n \asymp \Delta_n^\rho$ in (3.49) satisfies $\rho \in [(3-\nu)/[4(2-\nu)], 1/2)$,*

$$\bar{V}^{n,1} \xrightarrow{u.c.p.} 0.$$

Proof. Define function g_n on $\mathcal{M}^+_d \times \mathcal{M}^+_d$ by

$$g_n(x,z) = g(x) - \frac{1}{2k_n\Delta_n^{1/2}}\sum_{j=1}^d\sum_{k=1}^d\sum_{l=1}^d\sum_{m=1}^d \partial^2_{jk,lm}g(x)\times\Xi(x,z)^{jk,lm},$$

we have

$$
\begin{aligned}
\|g_n(x,z)-g_n(y,z)\| \leq{}& \|g(x)-g(y)\| \\
&+\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}\sum_{m=1}^{d}\frac{1}{k_n\Delta_n^{1/2}}\left\|\partial^2_{jk,lm}g(x)\times\Xi(x,z)^{jk,lm}-\partial^2_{jk,lm}g(y)\times\Xi(y,z)^{jk,lm}\right\|,
\end{aligned}
$$

according to (4.1),

$$
\begin{aligned}
\|g_n(x,z)-g_n(y,z)\| \leq{}& K\left[1+(\|x\|\vee\|y\|)^{r-1}\right]\|x-y\| \\
&+\frac{K}{k_n\Delta_n^{1/2}}\left[1+(\|x\|\vee\|y\|)^{r-3}\right]\|x-y\|\left(\|x\|^2+\|z\|^2\right) \\
&+\frac{K}{k_n\Delta_n^{1/2}}\left(1+\|y\|^{r-2}\right)\|x-y\|\left(\|x-y\|^2+\|z\|\|x-y\|\right),
\end{aligned}
$$

so $\|g_n(x,z)-g_n(y,z)\|\leq K\left[1+(\|x\|\vee\|y\|)^{r-1}\right]\|x-y\|$ when n is sufficiently large. Note

$$
\begin{aligned}
\mathbb{E}\left(\sup_{s\in[0,t]}\left\|\overline{V}_s^{n,1}\right\|\right) &\leq k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\left\|g_n(\widehat{c}^n_{ik_n},\widehat{\gamma}^n_{ik_n})-g_n(\widehat{c}^{*n}_{ik_n},\widehat{\gamma}^n_{ik_n})\right\| \\
&\leq Kt\left(a_n\Delta_n^{1/4-(\rho-1/4)\nu-(1-2\rho)}+\Delta_n^{1/4}\right),
\end{aligned}
$$

by **Lemma 21**. Since $\rho>\frac{3-\nu}{4(2-\nu)}$, $1/4-(\rho-1/4)\nu-(1-2\rho)>0$, this lemma is verified. $\square$

Lemma 26. *If we let $l_n\asymp\Delta_n^{-1/2}$ and $k_n\Delta_n^{3/4}\to 0$,*

$$
\overline{V}^{n,2}\xrightarrow{u.c.p.}0.
$$

Proof. Consider the process

$$
\overline{V}_t^{\prime n}=k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\sum_{j=1}^{d}\sum_{k=1}^{d}\partial_{jk}g(c^n_{ik_n})\times\xi^{n,jk}_{ik_n},
$$

and suppose $\xi_i^n \in \mathbb{R}^{d\times d}$ satisfies

$$\left.\begin{array}{rcl} \left\|E\left(\xi_i^n|\mathcal{F}_i^n\right)\right\| & \leq & K\Delta_n^{1/4}a_n \\ E\left(\|\xi_i^n\|^2|\mathcal{F}_i^n\right) & \leq & K(k_n\Delta_n^{1/2})^{-1}b_n \end{array}\right\} \tag{4.10}$$

where $a_n, b_n \to 0$. Since ∂g is bounded,

$$\mathbb{E}\left(\sup_{s\in[0,t]}\left\|\bar{V}_s'^n\right\|\right) \leq Kk_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\mathbb{E}\left(\left\|E\left(\xi_{ik_n}^n|\mathcal{F}_{ik_n}^n\right)\right\|\right) + Kk_n\Delta_n^{3/4}\mathbb{E}\left(\sup_{s\in[0,t]}\left\|\sum_{i=0}^{N_s^n-1}\left[\xi_{ik_n}^n - E\left(\xi_{ik_n}^n|\mathcal{F}_{ik_n}^n\right)\right]\right\|\right),$$

by **Lemma 23**,

$$\mathbb{E}\left(\sup_{s\in[0,t]}\left\|\sum_{i=0}^{N_s^n-1}\left[\xi_{ik_n}^n - E\left(\xi_{ik_n}^n|\mathcal{F}_{ik_n}^n\right)\right]\right\|\right) \leq K\left(\sum_{i=0}^{N_t^n-1}\mathbb{E}\left(\|\xi_{ik_n}^n\|^2\right)\right)^{1/2},$$

note that $k_n\Delta_nN_t^n \asymp t$, we have

$$\mathbb{E}\left(\sup_{s\in[0,t]}\left\|\bar{V}_s'^n\right\|\right) \leq K\left(ta_n + \sqrt{tb_n}\right) \to 0,$$

thereby we show that $\bar{V}'^n$ is uniformly asymptotically negligible on compact intervals.

To show the asymptotic negligibility of $\bar{V}^{n,2}$, we need to show ξ_i satisfies (4.10) in each of the following 3 cases:

(1) $\xi_i^n = \xi(0)_i^n$
by (3.13), (4.10) is realized with $a_n = k_n\Delta_n^{3/4}$, $b_n = (k_n\Delta_n^{3/4})^2$;

(2) $\xi_i^n = \xi(1)_i^n$
by **Lemma 7**, (4.10) holds with $a_n = l_n\Delta_n^{3/4} + (l_n\Delta_n^{1/4})^{-1}$, $b_n = k_n\Delta_n^{3/4}\big[l_n\Delta_n^{3/4} +$

$(l_n\Delta_n^{1/8})^{-2}]$;

(3) $\xi_i^n = \xi(2)_i^n$

according to **Lemma 8**, (4.10) is satisfied with $a_n = (l_n\Delta_n^{1/8})^{-2}$, $b_n = (l_n\Delta_n^{3/8})^{-4}$. $\square$

Lemma 27. *If we let $l_n\Delta_n^{1/2} = \theta$ where θ is positive finite, and $k_n\Delta_n^{3/4} \to 0$, $k_n\Delta_n^{2/3} \to \infty$,*

$$\overline{V}^{n,3} \xrightarrow{u.c.p.} 0.$$

Proof. We can rewrite $\overline{V}^{n,3}$ as

$$\overline{V}^{n,3} = G^n + H^n,$$

where

$$\begin{aligned} G_t^n &\equiv k_n\Delta_n^{3/4} \sum_{i=0}^{N_t^n-1} \left[s_{ik_n}^n + u_{ik_n}^n + E\left(v_{ik_n}^n | \mathcal{F}_{ik_n}^n\right)\right] \\ H_t^n &\equiv k_n\Delta_n^{3/4} \sum_{i=0}^{N_t^n-1} \left[v_{ik_n}^n - E\left(v_{ik_n}^n | \mathcal{F}_{ik_n}^n\right)\right], \end{aligned}$$

and

$$\begin{aligned} s_i^n &= g(c_i^n + \beta_i^n) - g(c_i^n) - \sum_{j=1}^d \sum_{k=1}^d \partial_{jk} g(c_i^n)\beta_i^{n,jk} \\ &\qquad - \frac{1}{2}\sum_{j=1}^d\sum_{k=1}^d\sum_{l=1}^d\sum_{m=1}^d \partial^2_{jk,lm} g(c_i^n)\beta_i^{n,jk}\beta_i^{n,lm} \\ u_i^n &= \frac{1}{2k_n\Delta_n^{1/2}} \sum_{j=1}^d\sum_{k=1}^d\sum_{l=1}^d\sum_{m=1}^d \left[\partial^2_{jk,lm} g(c_i^n)\,\Xi(c_i^n,\gamma_i^n)^{jk,lm} - \partial^2_{jk,lm} g(\widehat{c}_i^{*n})\,\Xi(\widehat{c}_i^{*n},\widehat{\gamma}_i^n)^{jk,lm}\right] \\ v_i^n &= \frac{1}{2}\sum_{j=1}^d\sum_{k=1}^d\sum_{l=1}^d\sum_{m=1}^d \partial^2_{jk,lm} g(c_i^n)\left[\beta_i^{n,jk}\beta_i^{n,lm} - (k_n\Delta_n^{1/2})^{-1}\Xi(c_i^n,\gamma_i^n)^{jk,lm}\right]. \end{aligned}$$

By (3.47) and (4.1),

$$
\begin{aligned}
\|s_i^n\| &\leq K(1+\|c_i^n\|^{r-3})\|\beta_i^n\|^3 \\
\|u_i^n\| &\leq K(k_n\Delta_n^{1/2})^{-1}\big[\|\partial^2 g(c_i^n)\| \times \|\kappa_i^n\| \\
&\qquad\qquad +\|\partial^2 g(c_i^n+\beta_i^n) - \partial^2 g(c_i^n)\| \times (\|\Xi(c_i^n,\gamma_i^n)\| + \|\kappa_i^n\|)\big],
\end{aligned}
$$

hence by **Assumption SA-ν**,

$$
\|s_i^n\| + \|u_i^n\| \leq K\big[\|\beta_i^n\|^3 + (k_n\Delta_n^{1/2})^{-1}(\|\beta_i^n\| + \|\kappa_i^n\| + \|\beta_i^n\|\|\kappa_i^n\|)\big]. \tag{4.11}
$$

(1) Study of G^n

By **Lemma 19**, we have $E(\|u_i^n\||\mathcal{F}_i^n) \leq K(k_n\Delta_n^{1/2})^{-3/2}$. By (4.1), **Assumption SA-ν** and the last claim in **Lemma 18**, we have

$$
\|E(v_i^n|\mathcal{F}_i^n)\| \leq K\big[k_n\Delta_n + (pk_n\Delta_n^{1/2})^{-1} + p^2(k_n^2\Delta_n)^{-1}\sum_{h=0}^{m(n,p)-1}\overline{\lambda}(p+1)_{i+a(n,p,h)}^n\big].
$$

These yield

$$
\begin{aligned}
\mathbb{E}\left(\sup_{s\in[0,t]}\|G_s^n\|\right) &\leq k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\left[\mathbb{E}\left(\|s_{ik_n}^n\| + \|u_{ik_n}^n\|\right) + \mathbb{E}\big(\left\|E(v_{ik_n}^n|\mathcal{F}_{ik_n}^n)\right\|\big)\right] \\
&\leq Kt\left[k_n\Delta_n^{3/4} + (k_n\Delta_n^{2/3})^{-3/2} + (pk_n\Delta_n^{3/4})^{-1}\right] \\
&\quad +Kp(k_n\Delta_n^{3/4})^{-1}\cdot\mathbb{E}\left(pl_n\Delta_n\sum_{i=0}^{N_t^n-1}\sum_{h=0}^{m(n,p)-1}\overline{\lambda}(p+1)_{ik_n+a(n,p,h)}^n\right),
\end{aligned}
$$

by the last claim in **Lemma 16**, $G^n \stackrel{u.c.p.}{\longrightarrow} 0$.

(2) Study of H^n

First we have by (4.1), **Assumption SA-ν**, **Lemma 18**,

$$E(\|v_i^n\|^2|\mathcal{F}_i^n) \leq K\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}\sum_{m=1}^{d} E\left(\left|\beta_i^{n,jk}\beta_i^{n,lm} - \frac{1}{k_n\Delta_n^{1/2}}\Xi_i^{n,jk,lm}\right|^2 |\mathcal{F}_i^n\right)$$
$$\leq K\left[k_n\Delta_n + (k_n\Delta_n^{1/2})^{-2}\right].$$

Lemma 23 implies

$$\mathbb{E}\left(\sup_{s\in[0,t]}\|H_s^n\|\right) \leq Kk_n\Delta_n^{3/4}\left(\sum_{i=0}^{N_t^n-1}\mathbb{E}(\|v_{ik_n}^n\|^2)\right)^{1/2} \leq K\sqrt{t}\left[k_n\Delta_n^{3/4} + (k_n\Delta_n^{1/2})^{-1/2}\right],$$

hence $H^n \stackrel{u.c.p.}{\longrightarrow} 0$.

Therefore, $\overline{V}^{n,3} = G^n + H^n \stackrel{u.c.p.}{\longrightarrow} 0$. □

4.3.3 stable convergence

Define

$$\begin{aligned}
\overline{V}(p,0)_t^{n,4} &= k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\sum_{j=1}^{d}\sum_{k=1}^{d}\partial_{jk}g(c_{ik_n}^n)\times N(p)_{ik_n}^{n,jk}\\
\overline{V}(p,1)_t^{n,4} &= k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\sum_{j=1}^{d}\sum_{k=1}^{d}\partial_{jk}g(c_{ik_n}^n)\times M(p)_{ik_n}^{n,jk},
\end{aligned}$$

note that $\overline{V}^{n,4} = \overline{V}(p,0)^{n,4} + \overline{V}(p,1)^{n,4}$.

Lemma 28. *If we let $l_n\Delta_n^{1/2} \asymp \theta$ where θ is positive finite, $l_n/k_n \to 0$, $p \to \infty$, $p/k_n \to 0$,*

$$\overline{V}(p,0)^{n,4} \stackrel{u.c.p.}{\longrightarrow} 0.$$

Proof. Based on (3.39), we can write $\overline{V}(p,0)^{n,4} = k_n(k_n-l_n)^{-1}\left[\overline{V}(p,0,0)^{n,4}+\overline{V}(p,0,1)^{n,4}\right]$,

where

$$
\begin{aligned}
\bar{V}(p,0,0)^{n,4} &= \textstyle\sum_{j=1}^{d}\sum_{k=1}^{d}\Delta_n^{-1/4}\sum_{i=0}^{N_t^n-1}\sum_{h=0}^{m(n,p)-1}\partial_{jk}g(c_{ik_n}^n)\times\zeta(Y,1)_{ik_n+b(n,p,h)}^{n,jk}\\
\bar{V}(p,0,1)^{n,4} &= \textstyle\sum_{j=1}^{d}\sum_{k=1}^{d}\Delta_n^{-1/4}\sum_{i=0}^{N_t^n-1}\sum_{h=m(n,p)(p+1)l_n}^{k_n-l_n}\partial_{jk}g(c_{ik_n}^n)\times\zeta_{ik_n+1+h}^{n,jk}.
\end{aligned}
$$

By **Lemma 15**, **23**

$$
\begin{aligned}
\mathbb{E}\left(\sup_{s\in[0,t]}\left\|\bar{V}(p,0,0)_s^{n,4}\right\|\right) &\le K\Delta_n^{-1/4}\sum_{i=0}^{N_t^n-1}\sum_{h=0}^{m(n,p)-1}\left\|\mathbb{E}\left[\zeta(Y,1)_{ik_n+b(n,p,h)}^n\right]\right\|\\
&\quad + K\Delta_n^{-1/4}\left(\sum_{i=0}^{N_t^n-1}\sum_{h=0}^{m(n,p)-1}\mathbb{E}\left[\left\|\zeta(Y,1)_{ik_n+b(n,p,h)}^n\right\|^2\right]\right)^{1/2}\\
&\le K\Big[tp^{-1}l_n\Delta_n^{3/4}+\sqrt{t}\big(\theta^{-3/2}+\theta^{1/2}\big)p^{-1/2}\\
&\qquad + \theta^{1/2}l_n\Delta_n\sum_{i=0}^{N_t^n-1}\sum_{h=0}^{m(n,p)-1}\bar{\lambda}(1)_{ik_n+b(n,p,h)}^n\Big].
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\mathbb{E}\left(\sup_{s\in[0,t]}\left\|\bar{V}(p,0,1)_s^{n,4}\right\|\right) &\le K\Big[tp\theta^2(k_n\Delta_n^{1/4})^{-1}+\sqrt{t}\big(\theta^{-3/2}+\theta^{1/2}\big)(p/k_n)^{1/2}\\
&\qquad + l_n^{1/2}\Delta_n^{5/4}\sum_{i=0}^{N_t^n-1}\sum_{h=m(n,p)(p+1)l_n}^{k_n-l_n}\bar{\lambda}(1)_{ik_n+1+h}^n\Big],
\end{aligned}
$$

then this lemma follows from **Lemma 16**. □

Lemma 29. *If we let $l_n \asymp \theta\Delta_n^{-1/2}$, then for $\forall p\in\mathbb{N}^+$,*

$$
\bar{V}(p,1)^{n,4}\stackrel{\mathcal{L}-s(f)}{\longrightarrow} Z(p),
$$

where $Z(p)$ is a process defined on an extension of the space $\big(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge0},\mathbb{P}\big)$, such that *conditioning on $\mathcal{F}$ it is a continuous mean-0 Gaussian martingale with variance*

$$\widetilde{E}[Z(p)Z(p)^{\mathrm{T}}|\mathcal{F}] = \int_0^t \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d \sum_{m=1}^d \partial_{jk} g(c_s) \partial_{lm} g(c_s)^{\mathrm{T}} \, \Xi(c_s, \gamma_s; p)^{jk,lm} \, \mathrm{d}s,$$

where $\widetilde{E}$ is the conditional expectation operator on the extended probability space and $\Xi(x, z; p)$ is defined in (3.34).

Proof. We will prove this result using **Theorem 2**.

Firstly, we can write

$$\overline{V}(p,1)_t^{n,4} = \frac{k_n}{k_n - l_n} \sum_{j=1}^d \sum_{k=1}^d \Delta_n^{-1/4} \sum_{i=0}^{N_t^n - 1} \sum_{h=0}^{m(n,p)-1} \zeta(Y,p)_{ik_n + a(n,p,h)}^{n,jk} \times \partial_{jk} g(c_{ik_n}^n),$$

Secondly, let $\mathcal{H}(p)_{i,h}^n = \mathcal{F}_{i+a(n,p,h)}^n$, by **Lemma 15**, **16**,

$$\begin{aligned}
\Delta_n^{-1/2} \sum_{i=0}^{N_t^n - 1} \sum_{h=0}^{m(n,p)-1} & \left\| E\left[\zeta(Y,p)_{ik_n + a(n,p,h)}^n | \mathcal{H}(p)_{i,h}^n\right] \right\|^2 \\
& \leq K p l_n^3 \Delta_n^{5/2} \left[t + p \Delta_n^{1/2} \sum_{i=0}^{N_t^n - 1} \sum_{h=0}^{m(n,p)-1} \overline{\lambda}(p)_{ik_n + a(n,p,h)}^n \right] \xrightarrow{\mathbb{P}} 0,
\end{aligned}$$

and note that by Jensen's inequality

$$\begin{aligned}
\|\partial g(c_i^n)\|^4 E(\|\zeta(Y,p)_i^n\|^4 | \mathcal{F}_i^n) & \geq \|\partial g(c_i^n)\|^4 E\left(\|\zeta(Y,p)_i^n\|^4 \mathbb{1}_{\{\|\zeta(Y,p)_i^n\| > \epsilon / \|\partial g(c_i^n)\|\}} | \mathcal{F}_i^n \right) \\
& \geq \left(\|\partial g(c_i^n)\|^2 E\left(\|\zeta(Y,p)_i^n\|^2 \mathbb{1}_{\{\|\zeta(Y,p)_i^n\| > \epsilon / \|\partial g(c_i^n)\|\}} | \mathcal{F}_i^n \right) \right)^2.
\end{aligned}$$

Hence, the following 4 statements about convergence in probability for any indexes j, k, l, m can verify the conditions (4.3) - (4.7) of **Theorem 2**, which then will prove the stable convergence:

$$\Delta_n^{-1/4} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \|\partial g(c_{ik_n}^n)\| \cdot \left\| E\left[\zeta(Y,p)_{ik_n+a(n,p,h)}^n | \mathcal{H}(p)_{ik_n,h}^n\right] \right\| \xrightarrow{\mathbb{P}} 0 \quad (4.12)$$

$$\Delta_n^{-1} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \|\partial g(c_{ik_n}^n)\|^4 \cdot E\left[\left\|\zeta(Y,p)_{ik_n+a(n,p,h)}^n\right\|^4 | \mathcal{H}(p)_{ik_n,h}^n\right] \xrightarrow{\mathbb{P}} 0, \quad (4.13)$$

and

$$\begin{aligned} \Delta_n^{-1/2} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} & \partial_{jk} g(c_{ik_n}^n) \partial_{lm} g(c_{ik_n}^n)^{\mathrm{T}} \\ & \times E\Big[\zeta(Y,p)_{ik_n+a(n,p,h)}^{n,jk} \zeta(Y,p)_{ik_n+a(n,p,h)}^{n,lm} | \mathcal{H}(p)_{ik_n,h}^n\Big] \\ & \xrightarrow{\mathbb{P}} \int_0^t \partial_{jk} g(c_s) \partial_{lm} g(c_s)^{\mathrm{T}}\, \Xi(c_s, \gamma_s; p)^{jk,lm}\, \mathrm{d}s \end{aligned} \quad (4.14)$$

$$\Delta_n^{-1/4} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \|\partial g(c_{ik_n}^n)\| \cdot \left\| E\Big[\zeta(Y,p)_{ik_n+a(n,p,h)}^n \Delta N(p)_{ik_n,h}^n | \mathcal{H}(p)_{ik_n,h}^n\Big] \right\| \xrightarrow{\mathbb{P}} 0, \quad (4.15)$$

where N is a bounded martingale orthogonal to W or $N = W^l$ for some $l = 1, 2, \cdots, d'$, and $\Delta N(p)_{i,h}^n = N_{i+b(n,p,h)}^n - N_{i+a(n,p,h)}^n$.

(1) (4.12) can be shown by the second claim of **Lemma 15**,

$$\begin{aligned} \Delta_n^{-1/4} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} & \|\partial g(c_{ik_n}^n)\| \cdot \left\| E\left[\zeta(Y,p)_{ik_n+a(n,p,h)}^n | \mathcal{H}(p)_{ik_n,h}^n\right] \right\| \leq \\ & K\left(t l_n \Delta_n^{3/4} + p l_n^{3/2} \Delta_n^{5/4} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \overline{\lambda}(p)_{ik_n+a(n,p,h)}^n \right) \xrightarrow{\mathbb{P}} 0. \end{aligned}$$

(2) (4.13) can be verified by the third claim of **Lemma 15**,

$$\Delta_n^{-1} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \|\partial g(c_{ik_n}^n)\|^4 \cdot E\left[\left\|\zeta(Y,p)_{ik_n+a(n,p,h)}^n\right\|^4 |\mathcal{H}(p)_{ik_n,h}^n\right]$$

$$\leq K t\, p\left(l_n^{-5}\Delta_n^{-2} + l_n^3 \Delta_n^2\right) \to 0.$$

(3) (4.14) be established by the last claim of **Lemma 15**,

$$\begin{aligned}
&\Delta_n^{-1/2} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \partial_{jk} g(c_{ik_n}^n)\, \partial_{lm} g(c_{ik_n}^n)^{\mathrm{T}} \\
&\qquad\qquad \times E\big[\zeta(Y,p)_{ik_n+a(n,p,h)}^{n,jk} \zeta(Y,p)_{ik_n+a(n,p,h)}^{n,lm} |\mathcal{H}(p)_{ik_n,h}^n\big] \\
&= \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \partial_{jk} g(c_{ik_n}^n)\, \partial_{lm} g(c_{ik_n}^n)^{\mathrm{T}} \times \Xi\big(c_{ik_n+a(n,p,h)}^n, \gamma_{ik_n+a(n,p,h)}^n; p\big)^{jk,lm} (p+1) l_n \Delta_n \\
&\qquad\qquad + p\, O_p\left(t\Delta_n^{1/4} + p\Delta_n^{1/2} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \overline{\lambda}(p)_{ik_n+a(n,p,h)}^n\right),
\end{aligned}$$

then (4.14) can be proved by **Lemma 16** and Riemann summation.

(4) (4.15) follows from the same argument of Jacod et al. [2009] for (5.58) therein.

□

Theorem 3. *Assume **Assumption A**-c, **A**-γ, **A**-ν hold for $\nu \in [0,1)$. Suppose $g \in C^3(\mathcal{M}_d^+)$ satisfies (4.1) for some constants $K > 0$ and $r \geq 3$.*

Furthermore, assume either one of the following two conditions:

(i) X is continuous, and $\nu_n \Delta_n^{-\rho} \to \infty$ with $\rho \in [0, 1/2)$

(ii) $\nu_n \asymp \Delta_n^\rho$ with

$$\rho \in [1/4 + 1/(4(2-\nu)), 1/2). \tag{4.16}$$

Beside, given θ positive finite, we control the bandwidths in such way that

$$\left.\begin{array}{rcl} l_n\Delta_n^{1/2} & = & \theta \\ k_n\Delta_n^{2/3} & \to & \infty \\ k_n\Delta_n^{3/4} & \to & 0 \end{array}\right\}. \tag{4.17}$$

Then we have the following functional stable convergence in law:

$$\Delta_n^{-1/4}\left[\widehat{S}(g)^n - S(g)\right] \overset{\mathcal{L}-s(f)}{\longrightarrow} Z, \tag{4.18}$$

where Z is a process defined on an extension of the space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$, such that conditioning on $\mathcal{F}$ it is a continuous mean-0 Gaussian martingale with variance

$$\widetilde{E}[ZZ^{\mathrm{T}}|\mathcal{F}] = V(g),$$

where $\widetilde{E}$ is the conditional expectation operator on the extended probability space and

$$V(g)_t \equiv \int_0^t \sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}\sum_{m=1}^{d} \partial_{jk}g(c_s)\,\partial_{lm}g(c_s)^{\mathrm{T}}\,\Xi(c_s,\gamma_s)^{jk,lm}\,\mathrm{d}s, \tag{4.19}$$

with $\Xi(c,\gamma)$ being defined in (3.36).

Proof. This result can be obtained by **Lemma**, **24**, **25**, **26**, **27**, **28**, **29**, and that

$$\Xi(c_s,\gamma_s;p) \to \Xi(c_s,\gamma_s),$$

as $p \to \infty$. $\square$

We can estimate the asymptotic variance by plugging in spot estimates:

$$\widehat{V}(g)_t^n \equiv k_n \Delta_n \sum_{i=0}^{N_t^n-1} \sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}\sum_{m=1}^{d} \partial_{jk} g(\widehat{c}_{ik_n}^n)\, \partial_{lm} g(\widehat{c}_{ik_n}^n)^{\mathrm{T}}\, \Xi(\widehat{c}_{ik_n}^n, \widehat{\gamma}_{ik_n}^n)^{jk,lm}, \tag{4.20}$$

where $\widehat{c}_i^n$ is defined as (3.49); $\widehat{\gamma}_i^n$ is defined as (3.46).

Remark 3. We have for all finite t,

$$\left\|\widehat{V}(g)_t^n - V(g)_t\right\| = O_p\big((k_n \Delta_n^{1/2})^{-1}\big).$$

Since $E(\|\beta_i^n\|^q | \mathcal{F}_i^n) > E(\|\chi_i^n\|^q | \mathcal{F}_i^n)$, the estimation error of $\widehat{V}(g)$ is dominated by β_i^n, so it is dictated by the tuning parameter l_n for pre-averaging. We are satisfied with $\widehat{V}(g)$ being consistent and this order of error, and will not purse the central limit theorem of $\widehat{V}(g)$ in this dissertation.

4.3.4 positive semi-definite plug-ins

As in (3.51), we can modify the original definition of the pre-averaging estimator (3.49) for spot volatility by dispensing with the noise-correction term $\widehat{Y}_i^n$ in (3.49),

$$\widehat{c}_i^{n,\mathrm{psd}} \equiv \frac{1}{(k_n - l_n)\Delta_n} \sum_{h=1}^{k_n - l_n + 1} \overline{Y}_{i+h}^n \cdot \overline{Y}_{i+h}^{n,\mathrm{T}} \mathbb{1}_{\{\|\overline{Y}_{i+h}^n\| \leq \nu_n\}}. \tag{4.21}$$

This new estimator (4.21) is positive semidefinite by definition.

Correspondingly, we define the following conceptual "estimator" like (3.7),

$$\widehat{c}_i^{*n,\mathrm{psd}} = \frac{1}{(k_n - l_n)\Delta_n} \sum_{h=1}^{k_n - l_n + 1} \overline{Y}_{i+h}^{*n} \cdot \overline{Y}_{i+h}^{*n,\mathrm{T}} \tag{4.22}$$

Similarly to (3.10), we have the following decomposition:

$$\beta(p)_i'^n = \widehat{c}_i^{*n,\mathrm{psd}} - c_i^n = \xi(0)_i^n + \xi(1)_i^n + \xi(2)_i'^n + N(p)_i^n + M(p)_i^n, \tag{4.23}$$

where $\xi(0)_i^n, \xi(1)_i^n, N(p)_i^n, M(p)_i^n$ are defined in (3.11), and

$$\xi(2)_i'^n = \frac{1}{(k_n - l_n)\Delta_n} \sum_{h=1}^{k_n - l_n + 1} \Gamma_{i+h}^n,$$

with Γ_i^n being defined in (3.8).

Define the functional estimator with positive semidefinite covariance plug-ins as

$$\begin{aligned}\widehat{S}(g)_t^{n,\mathrm{psd}} \equiv k_n \Delta_n \sum_{i=0}^{N_t^n - 1} \Big[& g\big(\widehat{c}_{ik_n}^{n,\mathrm{psd}}\big) \\ & - \frac{\theta \Phi_{00}}{\psi_0^2 k_n \Delta_n^{1/2+\delta}} \sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}\sum_{m=1}^{d} \partial_{jk,lm}^2 g\big(\widehat{c}_{ik_n}^{n,\mathrm{psd}}\big) \times \Sigma\big(\widehat{c}_{ik_n}^{n,\mathrm{psd}}\big)^{jk,lm} \Big],\end{aligned} \tag{4.24}$$

where $\Sigma(c)$ is defined as (3.35).

Theorem 4. *Assume* ***Assumption A-****c,* ***A-****γ,* ***A-****ν hold for $\nu \in [0,1)$. Suppose $g \in C^3(\mathcal{M}_d^+)$ satisfies (4.1) for some constants $K > 0$ and $r \geq 3$.*

Given θ positive finite, we control the bandwidths l_n, k_n and the truncation threshold ν_n in such way that

$$\begin{cases} l_n & \asymp & \theta \Delta_n^{-1/2-\delta} & \delta \in \left(\frac{1}{10}, \frac{1}{2}\right) \\ k_n & \asymp & \varrho \Delta_n^{-\kappa} & \kappa \in \Big(\left(\frac{2}{3} + \frac{2\delta}{3}\right) \vee \left(\frac{2+\nu}{4} + \frac{(2-\nu)\delta}{2}\right), \frac{3}{4} + \frac{\delta}{2}\Big) \\ \nu_n & = & \alpha \Delta_n^{\rho} & \rho \in \Big[\frac{1}{4} + \frac{\delta}{2} + \frac{1-\kappa}{2-\nu}, \frac{1}{2}\Big), \end{cases} \tag{4.25}$$

then we have the following functional stable convergence in law

$$\Delta_n^{-1/4+\delta/2}\left[\widehat{S}(g)^{n,\mathrm{psd}} - S(g)\right] \stackrel{\mathcal{L}-s(f)}{\longrightarrow} Z', \tag{4.26}$$

where Z' is a process defined on an extension of the space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$, such that conditioning on $\mathcal{F}$ it is a continuous mean-0 Gaussian martingale with variance

$$\widetilde{E}[Z' Z'^{\mathrm{T}}|\mathcal{F}] = V(g)',$$

where $\widetilde{E}$ is the conditional expectation operator on the extended probability space and

$$V(g)'_t \equiv \frac{2\theta\Phi_{00}}{\psi_0^2}\int_0^t \sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}\sum_{m=1}^{d} \partial_{jk} g(c_s)\, \partial_{lm} g(c_s)^{\mathrm{T}}\, \Sigma(c_s)^{jk,lm}\, \mathrm{d}s, \tag{4.27}$$

with $\Sigma(c)$ being defined as (3.35).

4.4 Simulation

As a proof of concept, when $d = 1$ estimators corresponding to $g(c) = c^2$, $g(c) = c^{-1}$, $g(c) = \log(c)$ are calculated based on the simulation model

$$\begin{cases} Y_i^n &= X_i^n + \varepsilon_i^n \\ \mathrm{d}X_t &= .03\,\mathrm{d}t + \sqrt{c_t}\,\mathrm{d}W_t + J_t^X\,\mathrm{d}N_t^X \\ \mathrm{d}c_t &= 6(.16 - c_t)\,\mathrm{d}t + .5\sqrt{c_t}\,\mathrm{d}B_t + \sqrt{c_{t-}}J_t^c\,\mathrm{d}N_t^c, \end{cases}$$

where $\varepsilon_i^n \stackrel{\text{i.i.d.}}{\sim} N(0, .005^2)$, $\mathbb{E}[(W_{t+\Delta_n} - W_t)(B_{t+\Delta_n} - B_t)] = -.6\Delta_n$, $J_t^X \sim N(-.01, .02^2)$, $N_{t+\Delta_n}^X - N_t^X \sim \text{Poisson}(36\Delta_n)$, $\log(J_t^c) \sim N(-5, .8)$, $N_{t+\Delta_n}^c - N_t^c \sim \text{Poisson}(12\Delta_n)$.

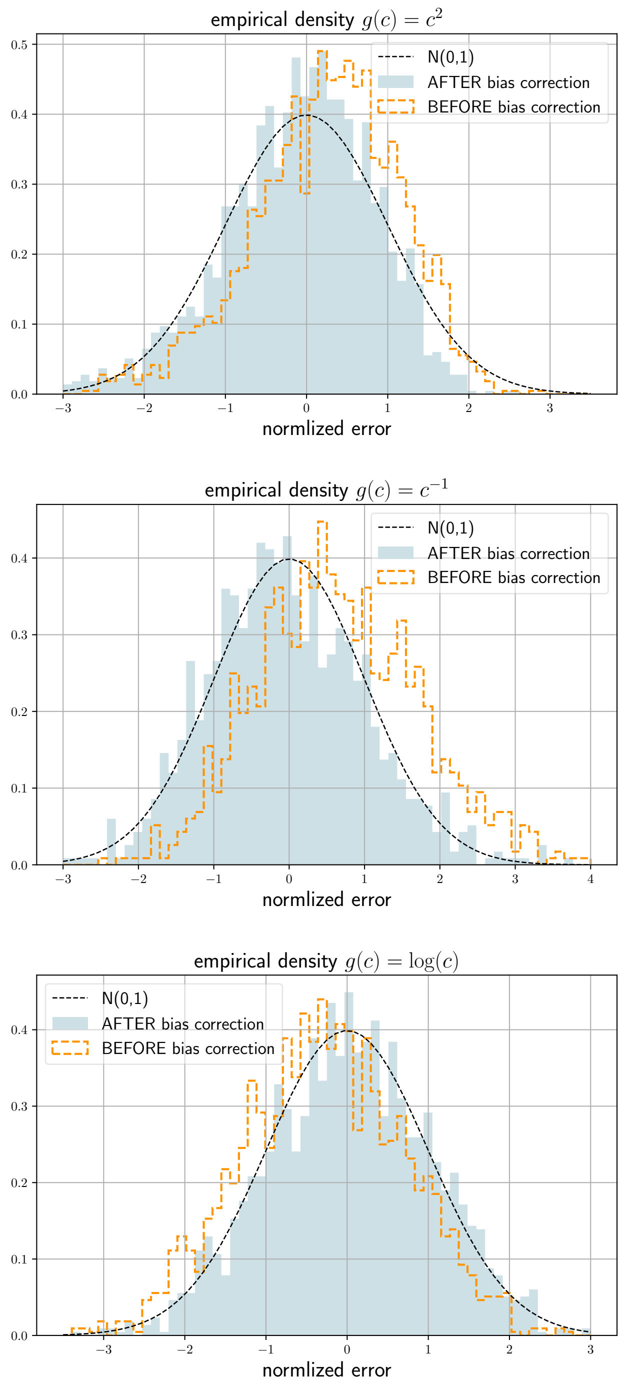
Each simulation employs 23400×21 data points with sampling interval $\Delta_n = 1s$. The author chooses the tuning parameters according to Table 4.2, where $\overline{\sigma}^2$ is an estimate of the average volatility by bipower variation [Podolskij and Vetter, 2009a].

functionals	l_n	k_n	ν_n
$g(c) = c^2$	$\lfloor \Delta_n^{-.5} \rfloor$	$\lfloor \Delta_n^{-.69} \rfloor$	$1.6\overline{\sigma}^2\Delta_n^{.47}$
$g(c) = c^{-1}$	$\lfloor \Delta_n^{-.5} \rfloor$	$\lfloor \Delta_n^{-.7} \rfloor$	$1.5\overline{\sigma}^2\Delta_n^{.47}$
$g(c) = \log(c)$	$\lfloor \Delta_n^{-.5} \rfloor$	$\lfloor \Delta_n^{-.7} \rfloor$	$1.5\overline{\sigma}^2\Delta_n^{.47}$

Table 4.2: Tuning parameters of the pre-averaging method in simulation

Simulation results are shown in Figure 4.1.

Figure 4.1: Simulation of volatility functional estimation by the pre-averaging method



CHAPTER 5

STATISTICAL UNCERTAINTY QUANTIFICATION FOR REALIZED PRINCIPAL COMPONENT ANALYSIS

Factor analysis is one of powerful statistical tools that help us obtain information and gain insights from huge and complex datasets. It reduces dimensionality, greatly facilitates data visualization, reveals common trends et cetera. One major approach to factor analysis is the principal component analysis (hereafter PCA) [Jolliffe, 2002], which can be done by estimation of spectral structure of covariance matrix. Along the same line, PCA of high-frequency data modeled by Itô semimartingales can be investigated through the lens of stochastic volatility matrix spectra, i.e., $\Sigma_t q_t = \lambda_t q_t$, where $\Sigma_t = c_t$ is the instantaneous covariance matrix at time t. In this chapter, I denote the instantaneous covariance matrix by Σ_t (notation widely used in factor model literature) rather than c_t (notation in the literature of limit theorems for stochastic processes).

The inferential theory of PCA of high-frequency data relies fundamentally on the results of volatility matrix functionals $\int_0^T g(\Sigma_t)\,\mathrm{d}t$. In the basic setting where λ_t is a simple eigenvalue of Σ_t and q_t is the corresponding eigenvector, the mapping $\Sigma_t \mapsto \lambda_t$ and $\Sigma_t \mapsto q_t$ are three-times continuously differentiable, cf. Magnus and Neudecker [2007]. Therefore the inferential results for $S(g)$ in section 4 are immediately applicable.

Based on Jacod and Rosenbaum [2013], Aït-Sahalia and Xiu [2019] developed the realized PCA method for high-frequency financial data, in absence of microstructure noise. The realized PCA method provides inference on the realized eigenvalue $\int_0^t \lambda_s\,\mathrm{d}s$, realized eigenvector $\int_0^t q_s\,\mathrm{d}s$, realized principal component $\int_0^t v_{s-}\,\mathrm{d}X_s$. Aït-Sahalia and Xiu [2017] applied this high-frequency PCA methodology under a high-dimensional approximate factor model to estimate large volatility matrices. Recently, Chen et al. [2019] extended the realized PCA to asynchronously observed noisy data.

5.1 Formulation of PCA

Suppose we have observations $X_0, X_1, \cdots, X_n$ of a random variable X taking values in $\mathbb{R}^d$. PCA seeks to reduce the "apparent" dimensionality of data and find an "intrinsic" dimensionality in which the original data structure/relation is preserved as much as possible.

Let $\Sigma = \mathrm{Var}(X)$ and suppose the following eigenvalue factorization

$$\Sigma = Q \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_d \end{bmatrix} Q^{\mathrm{T}}, \tag{5.1}$$

where $Q = [q_1, \cdots, q_d]$ is orthogonal and $\lambda_1 \geq \cdots \geq \lambda_d$.

Consider the dimension-reducing linear transform $F = WX$ from $\mathbb{R}^d$ to $\mathbb{R}^r$, where $W = [w_1, \cdots, w_r]^{\mathrm{T}}$ and $w_1, \cdots, w_r$ are mutually orthogonal, i.e., $WW^{\mathrm{T}} = \mathbf{I}_r$. We see that $\widehat{X} := W^{\mathrm{T}}F = W^{\mathrm{T}}WX$ is the projection of X onto the subspace $\mathbb{R}^d$.

In the same spirit of least squares, one natural criterion of dimension reduction is to minimize the mean square error of the projection $\widehat{X}$ with respect to X, which is

$$\begin{aligned}
\mathbb{E}\big(\|\widehat{X} - X\|^2\big) &= \mathbb{E}\big(\mathrm{tr}\big[(\widehat{X} - X)(\widehat{X} - X)^T\big]\big) \\
&= \mathrm{tr}\big(\Sigma\big) - \mathrm{tr}\big(\Sigma W^{\mathrm{T}} W\big) - \mathrm{tr}\big(W^{\mathrm{T}} W \Sigma\big) + \mathrm{tr}\big(W^{\mathrm{T}} W \Sigma W^{\mathrm{T}} W\big) \\
&= \mathrm{tr}\big(\Sigma\big) - \mathrm{tr}\big(W \Sigma W^{\mathrm{T}}\big) \\
&= \mathrm{tr}\big(\Sigma\big) - \mathbb{E}(\|F\|^2).
\end{aligned}$$

Therefore, equivalent to minimizing the mean squared projection error, we would like to find weight vectors $w_1, \cdots, w_r \in \mathbb{R}^d$ such that

- $\|w_j\| = 1,\ j = 1, \cdots, r;$

- $w_j^{\mathrm{T}} X \perp\!\!\!\perp w_k^{\mathrm{T}} X,\ j \neq k\ (\because w_j^{\mathrm{T}} w_k = 0,\ j \neq k)$;

- $w_j = \arg\max_\alpha \alpha^{\mathrm{T}} \Sigma \alpha$.

We can derive the solution using Lagrange multipliers. Define

$$\mathcal{L}_1(\alpha, \beta_1) = \alpha^{\mathrm{T}} \Sigma \alpha - \beta_1(\alpha^{\mathrm{T}} \alpha - 1),$$

then

$$\frac{\partial}{\partial \alpha} \mathcal{L}_1 = 2\Sigma\alpha - 2\beta_1 \alpha = 0 \implies \Sigma\alpha = \beta_1 \alpha,$$

it also implies $\alpha^{\mathrm{T}} \Sigma \alpha = \beta_1$. Thus we should take $\beta_1 = \lambda_1$ and $w_1 = q_1$.

Next, for $j = 2, \cdots, r$, define

$$\mathcal{L}_j(\alpha, \beta_j, \theta_1, \cdots, \theta_{j-1}) = \alpha^{\mathrm{T}} \Sigma \alpha - \beta_j(\alpha^{\mathrm{T}} \alpha - 1) - \sum_{k=1}^{j-1} \theta_k \alpha^{\mathrm{T}} q_k,$$

the first-order condition is

$$\frac{\partial}{\partial \alpha} \mathcal{L}_j = 2\Sigma\alpha - 2\beta_j \alpha - \sum_{k=1}^{j-1} \theta_k q_k = 0, \tag{5.2}$$

multiply q_k^{T}, $k = 1, \cdots, j-1$ respectively on both sides of (5.2),

$$q_k^{\mathrm{T}} \frac{\partial}{\partial \alpha} \mathcal{L}_j = 2\lambda_k \alpha^{\mathrm{T}} q_k - 2\beta_j \alpha^{\mathrm{T}} q_k - \theta_k q_k^{\mathrm{T}} q_k = \theta_k = 0,$$

hence $\Sigma\alpha = \beta_j \alpha$, since we require $\alpha^{\mathrm{T}} q_k = 0,\ k = 1, \cdots, j-1$, we should let $\beta_k = \lambda_k$ and $\alpha = q_k$ in the optimization problem.

Therefore we shall use the linear transformation

$$F = [q_1, \cdots, q_r]^{\mathrm{T}} X.$$

The reconstruction $W^{\mathrm{T}}F$, is the most feature-preserving among all linear transforms, in the sense of minimizing the mean squared error. Such F is called "*principal components*" in PCA and "*factor*" in the parlance of factor analysis.

5.2 Matrix calculus relevant to PCA

The eigenvalues and eigenvectors can be considered as functions of the corresponding covariance matrix. Since Σ is real symmetric, $\exists\ d$ pairs $\big(\lambda_r(\Sigma),\, q_r(\Sigma)\big)$, $r = 1, \cdots, d$, where $\lambda_r(\Sigma) \in \mathbb{R}$, $q_r(\Sigma) \in \mathbb{R}^d$, such that

$$\Sigma\, q_r(\Sigma) = \lambda_r(\Sigma)\, q_r(\Sigma), \tag{5.3}$$

and we can recast (5.1) as

$$\Sigma = [q_1(\Sigma), \cdots, q_d(\Sigma)] \begin{bmatrix} \lambda_1(\Sigma) & & \\ & \ddots & \\ & & \lambda_d(\Sigma) \end{bmatrix} \begin{bmatrix} q_1(\Sigma)^{\mathrm{T}} \\ \vdots \\ q_d(\Sigma)^{\mathrm{T}} \end{bmatrix}. \tag{5.4}$$

By requiring the unit length and specifying the directions, eigenvectors are uniquely determined.

According to *Weyl's inequality*, $\exists K > 0$, $\forall$ "small" perturbation matrix δ,

$$|\lambda_r(\Sigma + \delta) - \lambda_r(\Sigma)| \leq K\|\delta\|,$$

consequently, the eigenvalues are Lipschitz continuous functions.

The derivatives of eigenvalues and eigenvectors with respect to the corresponding covariance matrix are indispensable to the functional inferential results presented in section 4. We will discuss the matrix calculus relevant to eigenvalues and eigenvectors in the rest of this section.

The existence of relevant derivatives hinges upon whether the eigenvalue is simple or repeated. If an eigenvalue is simple, both the gradients and Hessians of the eigenvalue and the associated eigenvector exist. However, if an eigenvalue is repeated, the eigenvalue is only first-order differentiable but is not twice differentiable, the eigenvector is not differentiable at all.

5.2.1 simple eigenvalues and corresponding eigenvectors

Let ∂_{jk}, $\partial^2_{jk,lm}$ be shorthand for differentiation operators $\frac{\partial}{\partial \Sigma_{jk}}$, $\frac{\partial^2}{\partial \Sigma_{jk} \partial \Sigma_{lm}}$ acting on the function $f(\Sigma)$. For example, $\partial_{jk}\Sigma$ is the matrix in which the (j,k)-th entry is 1 and all other entries is 0.

Before deriving the derivatives, we need to show that simple eigenvalues and the corresponding eigenvectors are differentiable functions. Suppose q_r is an eigenvector associated with a simple eigenvalue λ_r of Σ. Define the function

$$\Theta_\Sigma(q, \lambda) = \begin{bmatrix} \Sigma q - \lambda q \\ q^{\mathrm{T}} q - 1 \end{bmatrix},$$

we can see that $\Theta_\Sigma \in C^\infty(\mathbb{R}^d \times \mathbb{R})$ and $\Theta_\Sigma(q_r, \lambda_r) = 0$. Note that

$$|\nabla \Theta_\Sigma(q_r, \lambda_r)| = \begin{vmatrix} \Sigma - \lambda_r \mathbf{I}_d & q_r \\ 2q_r^{\mathrm{T}} & 0 \end{vmatrix} = -2 \prod_{v \neq r} (\lambda_v - \lambda_r) \neq 0,$$

by the *implicit function theorem*, $\lambda_r(\cdot)$ and $q_r(\cdot)$ are C^∞ in a neighborhood of Σ.

(1) Gradients

Based on the definition (5.3) and chain rule, for $r = 1, \cdots, d$,

$$\partial_{jk}\Sigma \times q_r(\Sigma) + \Sigma \times \partial_{jk}q_r(\Sigma) = \partial_{jk}\lambda_r(\Sigma) \times q_r(\Sigma) + \lambda_r(\Sigma) \times \partial_{jk}q_r(\Sigma), \tag{5.5}$$

Let $q_r(\Sigma)^{\mathrm{T}}$ multiply with both sides of (5.5),

$$\partial_{jk}\lambda_r(\Sigma) = q_{r,j} \times q_{r,k}. \tag{5.6}$$

By (5.5),

$$\big[\lambda_r(\Sigma)\,\mathbf{I}_d - \Sigma\big] \times \partial_{jk}q_r(\Sigma) = \partial_{jk}\Sigma \times q_r(\Sigma) - \partial_{jk}\lambda_r(\Sigma) \times q_r(\Sigma).$$

Assume λ_r is a simple eigenvalue of Σ, then

$$\big(\lambda_r\mathbf{I}_d - \Sigma\big)^{\dagger} = [q_1, \cdots, q_d] \begin{bmatrix} \frac{1}{\lambda_r-\lambda_1} & & & & & & \\ & \ddots & & & & & \\ & & \frac{1}{\lambda_r-\lambda_{r-1}} & & & & \\ & & & 0 & & & \\ & & & & \frac{1}{\lambda_r-\lambda_{r+1}} & & \\ & & & & & \ddots & \\ & & & & & & \frac{1}{\lambda_r-\lambda_d} \end{bmatrix} \begin{bmatrix} q_1^{\mathrm{T}} \\ \vdots \\ q_d^{\mathrm{T}} \end{bmatrix}, \tag{5.7}$$

so

$$\big(\lambda_r\mathbf{I}_d - \Sigma\big)^{\dagger} q_r = 0, \tag{5.8}$$

note $A = \big(\lambda_r\mathbf{I}_d - \Sigma\big)^{\dagger}\big(\lambda_r\mathbf{I}_d - \Sigma\big) + q_r q_r^{\mathrm{T}}$ is symmetric and idempotent,

$$\mathrm{rk}(A) = \mathrm{tr}(A) = \mathrm{tr}\big[\big(\lambda_r\mathbf{I}_d - \Sigma\big)^{\dagger}\big(\lambda_r\mathbf{I}_d - \Sigma\big)\big] + \mathrm{tr}(q_r q_r^{\mathrm{T}}) = d,$$

thus $A = \mathbf{I}_d$; since $q_r^{\mathrm{T}} q_r = 1$, $q_r^{\mathrm{T}}\,\partial_{jk}q_r = 0$, hence $\big(\lambda_r\mathbf{I}_d - \Sigma\big)^{\dagger}\big(\lambda_r\mathbf{I}_d - \Sigma\big)\,\partial_{jk}q_r = \partial_{jk}q_r$.

Therefore, when λ_r is a simple eigenvalue,

$$\partial_{jk} q_r(\Sigma) = \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger} \times \partial_{jk}\Sigma \times q_r,$$

i.e.,

$$\partial_{jk} q_r(\Sigma) = \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger}_{\cdot j} \times q_{r,k} \tag{5.9}$$

$$\partial_{jk} q_{r,s}(\Sigma) = \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger}_{sj} \times q_{r,k},$$

equivalently,

$$\partial_{jk} q_r(\Sigma) = \sum_{v \neq r} \frac{1}{\lambda_r - \lambda_v} \, q_{v,j} \times q_{r,k} \times q_v \tag{5.10}$$

$$\partial_{jk} q_{r,s}(\Sigma) = \sum_{v \neq r} \frac{1}{\lambda_r - \lambda_v} \, q_{v,s} \times q_{v,j} \times q_{r,k}. \tag{5.11}$$

It follows from (5.8) and (5.9) that

$$\begin{aligned}
& \sum_{j,k,l,m=1}^{d} \partial_{jk} q_r \, \partial_{lm} q_r^{\mathrm{T}} \big(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\big) \\
= \; & \sum_{j,l=1}^{d} \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger}_{\cdot j} \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger}_{l \cdot} \times \sum_{k,m=1}^{d} q_{r,k} \, q_{r,m} \big(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\big) \\
= \; & \sum_{j,l=1}^{d} \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger}_{\cdot j} \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger}_{l \cdot} \times \sum_{k=1}^{d} q_{r,k} \big(\lambda_r \Sigma_{jl} \, q_{r,k} + \lambda_r \Sigma_{kl} \, q_{r,j}\big) \\
= \; & \lambda_r \sum_{j,l=1}^{d} \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger}_{\cdot j} \times \Sigma_{jl} \times \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger}_{l \cdot} \\
& \qquad + \lambda_r^2 \sum_{j=1}^{d} \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger}_{\cdot j} \, q_{r,j} \times \sum_{l=1}^{d} \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger}_{l \cdot} \, q_{r,l} \\
= \; & \lambda_r \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger} \, \Sigma \, \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger} + \lambda_r^2 \big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger} q_r \big[\big(\lambda_r \, \mathbf{I}_d - \Sigma\big)^{\dagger} q_r\big]^{\mathrm{T}},
\end{aligned}$$

so

$$\sum_{j,k,l,m=1}^{d} \partial_{jk} q_r \, \partial_{lm} q_r^{\mathrm{T}} \left(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\right)$$

$$= \lambda_r \left(\lambda_r \, \mathbf{I}_d - \Sigma\right)^{\dagger} \Sigma \left(\lambda_r \, \mathbf{I}_d - \Sigma\right)^{\dagger} = \sum_{v \neq r} \frac{\lambda_r \lambda_v}{(\lambda_r - \lambda_v)^2} q_v q_v^{\mathrm{T}}. \quad (5.12)$$

(2) Hessians

In view of (5.6) and (5.11), when λ_r is a simple eigenvalue,

$$\begin{aligned}
\partial^2_{jk,lm} \lambda_r(\Sigma) &= \partial_{lm} q_{r,j} \times q_{r,k} + q_{r,j} \times \partial_{lm} q_{r,k} \\
&= \left[\lambda_r \, \mathbf{I}_d - \Sigma\right]^{\dagger}_{jl} \times q_{r,k} \, q_{r,m} + \left[\lambda_r \, \mathbf{I}_d - \Sigma\right]^{\dagger}_{kl} \times q_{r,j} \, q_{r,m} \\
&= \sum_{v \neq r} \frac{1}{\lambda_r - \lambda_v} \times \left(q_{v,j} \, q_{v,l} \, q_{r,k} \, q_{r,m} + q_{v,k} \, q_{v,l} \, q_{r,j} \, q_{r,m}\right). \qquad (5.13)
\end{aligned}$$

For the eigenvector associated with the simple eigenvalue λ_r, by (5.10) we have

$$\begin{aligned}
\partial^2_{jk,lm} q_r = &\sum_{v \neq r} \frac{\partial_{lm}\lambda_v - \partial_{lm}\lambda_r}{(\lambda_r - \lambda_v)^2} \times q_{v,j} \, q_{r,k} \, q_v \\
&+ \sum_{v \neq r} \frac{1}{\lambda_r - \lambda_v} \times \left(q_{v,j} \, q_{r,k} \times \partial_{lm} q_v + \partial_{lm} q_{v,j} \times q_{r,k} \, q_v + \partial_{lm} q_{r,k} \times q_{v,j} \, q_v\right),
\end{aligned}$$

by (5.6),

$$\begin{aligned}
\partial^2_{jk,lm} q_r &= \sum_{v \neq r} \frac{1}{(\lambda_r - \lambda_v)^2} \times \left(q_{v,l} \, q_{v,m} \, q_{v,j} \, q_{r,k} - q_{v,j} \, q_{r,k} \, q_{r,l} \, q_{r,m}\right) q_v \\
&+ \sum_{v \neq r} \sum_{b \neq v} \frac{1}{(\lambda_r - \lambda_v)(\lambda_v - \lambda_b)} \times q_{b,l} \, q_{v,m} \, q_{v,j} \, q_{r,k} \, q_b \\
&+ \sum_{v \neq r} \sum_{b \neq v} \frac{1}{(\lambda_r - \lambda_v)(\lambda_v - \lambda_b)} \times q_{b,j} \, q_{b,l} \, q_{v,m} \, q_{r,k} \, q_v \\
&+ \sum_{v \neq r} \sum_{b \neq r} \frac{1}{(\lambda_r - \lambda_v)(\lambda_r - \lambda_b)} \times q_{b,k} \, q_{b,l} \, q_{r,m} \, q_{v,j} \, q_v, \qquad (5.14)
\end{aligned}$$

note that

$$\begin{aligned}\sum_{j,k,l,m=1}^{d} \big(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\big)\, q_{v,l}\, q_{v,m}\, q_{v,j}\, q_{r,k} &= 2\lambda_v^2 \sum_{j,k=1}^{d} q_{v,j}^2\, q_{v,k}\, q_{r,k} = 2\lambda_v^2 \mathbb{1}_{\{v=r\}} \\ \sum_{j,k,l,m=1}^{d} \big(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\big)\, q_{v,j}\, q_{r,k}\, q_{r,l}\, q_{r,m} &= 2\lambda_r^2 \sum_{j,k=1}^{d} q_{r,k}^2\, q_{v,j}\, q_{r,j} = 2\lambda_r^2 \mathbb{1}_{\{v=r\}},\end{aligned}$$

and

$$\sum_{j,k,l,m=1}^{d} \big(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\big)\, q_{b,l}\, q_{v,m}\, q_{v,j}\, q_{r,k} =$$
$$\lambda_b\lambda_v \sum_{j,k=1}^{d} \big(q_{b,j}\, q_{v,k}\, q_{v,j}\, q_{r,k} + q_{v,j}^2\, q_{b,k}\, q_{r,k}\big) = \lambda_b\lambda_v\big(\mathbb{1}_{\{b=v=r\}} + \mathbb{1}_{\{b=r\}}\big)$$

$$\sum_{j,k,l,m=1}^{d} \big(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\big)\, q_{b,j}\, q_{b,l}\, q_{v,m}\, q_{r,k} =$$
$$\lambda_b\lambda_v \sum_{j,k=1}^{d} \big(q_{b,j}^2\, q_{v,k}\, q_{r,k} + q_{b,j}\, q_{b,k}\, q_{v,j}\, q_{r,k}\big) = \lambda_b\lambda_v\big(\mathbb{1}_{\{v=r\}} + \mathbb{1}_{\{b=v=r\}}\big)$$

$$\sum_{j,k,l,m=1}^{d} \big(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\big)\, q_{b,k}\, q_{b,l}\, q_{r,m}\, q_{v,j} =$$
$$\lambda_b\lambda_r \sum_{j,k=1}^{d} \big(q_{b,k}\, q_{b,j}\, q_{r,k}\, q_{v,j} + q_{b,k}^2\, q_{r,j}\, q_{v,j}\big) = \lambda_b\lambda_r\big(\mathbb{1}_{\{b=r=v\}} + \mathbb{1}_{\{r=v\}}\big),$$

then it follows from (5.14) that

$$\sum_{j,k,l,m=1}^{d} \partial^2_{jk,lm}\, q_r\, \big(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\big) = -\sum_{v\neq r} \frac{\lambda_r\lambda_v}{(\lambda_r-\lambda_v)^2}\, q_r. \tag{5.15}$$

5.2.2 spectral functions of repeated eigenvalues

Based on (5.6), (5.10), (5.13), (5.14), if λ_r is a simple eigenvalue of Σ, we know both $\lambda_r(\cdot)$ and $q_r(\cdot)$ are twice differentiable at Σ. However, if λ_r is a repeated eigenvalue, $\lambda_r(\cdot)$ is only first-order differentiable but is not second-order differentiable, and $q_r(\cdot)$ is not differentiable at Σ.

For this reason, the general functional results in **Theorem 3** and **Theorem 4** can be applied to simple eigenvalues and the associated eigenvectors, but they can not be directly applied to repeated eigenvalues and the associated eigenvectors. This is bad news. The good news is we can still have a central limit theorem for repeated eigenvalues formulated as a spectral function.

Let F be a function defined on the space $\mathcal{S}_d^+$ of positive semi-definite matrices, we call F is a *spectral function* if the value of F only depends on the eigenvalues of its argument, i.e., $\forall \Sigma \in \mathcal{S}_d^+$, $\forall d \times d$ orthogonal matrix O,

$$F(\Sigma) = F(O^{\mathrm{T}} \Sigma O).$$

Let f be a function defined on $\mathbb{R}^d$, we call f is a symmetric function if for any $d \times d$ permutation matrix P, and $\forall x \in \mathbb{R}^d$,

$$f(x) = f(Px).$$

Define λ to be a function on $\mathcal{S}_d^+$ such that for $\Sigma \in \mathcal{S}_d^+$,

$$\lambda(\Sigma) = \left[\lambda_1(\Sigma), \cdots, \lambda_d(\Sigma)\right]^{\mathrm{T}},$$

with $\lambda_1(\Sigma) \geq \cdots \geq \lambda_d(\Sigma)$. Note by definition, a spectral functional is a symmetric function of eigenvalues. For a spectral function F, we can write $F = f \circ \lambda$ with some symmetric function f and $\circ$ meaning function composition.

Here we restate a proposition based on Theorem 1.1 of Lewis [1996] and Theorem 3.3 of Lewis and Sendov [2001].

Proposition 1. *Suppose* $\Sigma = Q\,\mathrm{Diag}(\lambda(\Sigma))\,Q^{\mathrm{T}}$, $F = f \circ \lambda$, *then*

$$
\begin{aligned}
\partial_{jk} F(\Sigma) &= \sum_{r=1}^{d} \partial_r f(\lambda(\Sigma))\, Q_{jr} Q_{kr} \\
\partial^2_{jk,lm} F(\Sigma) &= \sum_{r,v=1}^{d} \partial^2_{rv} f(\lambda(\Sigma))\, Q_{jr} Q_{kr} Q_{lv} Q_{mv} + \sum_{r,v=1}^{d} \mathcal{H}^f_{r,v}(\lambda(\Sigma))\, Q_{jr} Q_{lr} Q_{kv} Q_{mv},
\end{aligned}
$$

where

$$
\mathcal{H}^f_{r,v}(x) = \begin{cases} 0, & r = v \\ \partial^2_{rr} f(x) - \partial^2_{vv} f(x), & r \neq v,\ x_r = x_v \\ \frac{\partial_r f(x) - \partial_v f(x)}{x_r - x_v}, & r \neq v,\ x_r \neq x_v. \end{cases}
$$

According to **Proposition 1**, spectral function of the form $F = f \circ \lambda$ is twice continuously differentiable at the point $\Sigma \in \mathcal{S}_d^+$ if and only if the symmetric function f is twice continuously differential at the point $\lambda(\Sigma)$.

Suppose the eigenvalues of Σ form K clusters so that

$$
\lambda_{r_0+1} \geq \cdots \geq \lambda_{r_1} > \lambda_{r_1+1} \geq \cdots \geq \lambda_{r_2} > \cdots\cdots > \lambda_{r_{K-1}+1} \geq \cdots \geq \lambda_{r_K}, \tag{5.16}
$$

where $r_0 = 0$, $r_K = d$. We define the following symmetric functions on $\mathbb{R}^d$

$$
f_h(x) = \frac{1}{r_h - r_{h-1}} \sum_{j=r_{h-1}+1}^{r_h} x_j.
$$

Define spectral functions $F_h = f_h \circ \lambda$ for $h = 1, \cdots, K$. These functions return simple averages of repeated eigenvalues. Note that $\partial_r f_h = (r_h - r_{h-1})^{-1} \sum_{j=r_{h-1}+1}^{r_h} \mathbb{1}_{\{j=r\}}$ and $\partial^2_{rr} f_h = 0$. According to **Lemma 1**,

$$\partial_{jk} F_h(\Sigma) = \sum_{r=1}^{d} \partial_r f_h(\lambda(\Sigma))\, Q_{jr} Q_{kr} = \frac{1}{r_h - r_{h-1}} \sum_{r=r_{h-1}+1}^{r_h} Q_{jr} Q_{kr}, \tag{5.17}$$

and

$$\begin{aligned}
\partial^2_{jk,lm} F_h(\Sigma) &= \sum_{r=1}^{d} \sum_{v \neq r} \mathcal{H}^{f_h}_{r,v}(\lambda(\Sigma))\, Q_{jr} Q_{lr} Q_{kv} Q_{mv} \\
&= \frac{1}{r_h - r_{h-1}} \sum_{b=r_{h-1}+1}^{r_h} \sum_{r=1}^{d} \sum_{\lambda_v \neq \lambda_r} \frac{\mathbf{1}_{\{b=r\}} - \mathbf{1}_{\{b=v\}}}{\lambda_r - \lambda_v}\, Q_{jr} Q_{lr} Q_{kv} Q_{mv} \\
&= \frac{1}{r_h - r_{h-1}} \left(\sum_{r=r_{h-1}+1}^{r_h} \sum_{\lambda_v \neq \lambda_r} \frac{1}{\lambda_r - \lambda_v} + \sum_{v=r_{h-1}+1}^{r_h} \sum_{\lambda_r \neq \lambda_v} \frac{1}{\lambda_v - \lambda_r} \right) Q_{jr} Q_{lr} Q_{kv} Q_{mv} \\
&= \frac{1}{r_h - r_{h-1}} \sum_{r=r_{h-1}+1}^{r_h} \sum_{\lambda_v \neq \lambda_r} \frac{1}{\lambda_r - \lambda_v} \big(Q_{jr} Q_{lr} Q_{kv} Q_{mv} + Q_{jv} Q_{lv} Q_{kr} Q_{mr} \big).
\end{aligned}$$

By (5.7),

$$(\lambda_r \mathbf{I}_d - \Sigma)^{\dagger}_{jl} = \sum_{\lambda_v \neq \lambda_r} \frac{1}{\lambda_r - \lambda_v} Q_{jv} Q_{lv},$$

hence

$$\partial^2_{jk,lm} F_h(\Sigma) = \frac{1}{r_h - r_{h-1}} \sum_{r=r_{h-1}+1}^{r_h} \Big[\big(\lambda_r \mathbf{I}_d - \Sigma\big)^{\dagger}_{km} Q_{jr} Q_{lr} + \big(\lambda_r \mathbf{I}_d - \Sigma\big)^{\dagger}_{jl} Q_{kr} Q_{mr} \Big]. \tag{5.18}$$

According to (5.17),

$$
\begin{aligned}
&\sum_{j,k,l,m=1}^{d} \partial_{jk} F_h(\Sigma)\, \partial_{lm} F_h(\Sigma) \left(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\right) \\
= \;& \frac{1}{(r_h - r_{h-1})^2} \sum_{r,v=r_{h-1}+1}^{r_h} \sum_{j,k,l,m=1}^{d} Q_{jr}Q_{kr}Q_{lv}Q_{mv}\big(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\big) \\
= \;& \frac{2}{(r_h - r_{h-1})^2} \sum_{r,v=r_{h-1}+1}^{r_h} \lambda_v^2 \sum_{j=1}^{d} Q_{jr}Q_{jv} \sum_{k=1}^{d} Q_{kr}Q_{kv} \\
= \;& \frac{2}{(r_h - r_{h-1})^2} \sum_{r,v=r_{h-1}+1}^{r_h} \lambda_v^2\, \mathbb{1}_{\{r=v\}},
\end{aligned}
$$

so

$$
\sum_{j,k,l,m=1}^{d} \partial_{jk} F_h(\Sigma)\, \partial_{lm} F_h(\Sigma) \left(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\right) = \frac{2}{r_h - r_{h-1}} \lambda_{r_h}^2. \tag{5.19}
$$

According to (5.18),

$$
\sum_{j,k,l,m=1}^{d} \partial^2_{jk,lm} F_h(\Sigma) \left(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\right) = \frac{2}{r_h - r_{h-1}} \sum_{r=r_{h-1}+1}^{r_h} \Gamma_r,
$$

where

$$
\begin{aligned}
\Gamma_r \;&=\; \sum_{j,k,l,m=1}^{d} \left(\lambda_r \mathbf{I}_d - \Sigma\right)^{\dagger}_{km} Q_{jr}Q_{lr}\big(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\big) \\
&=\; \sum_{k,m=1}^{d} \left(\lambda_r \mathbf{I}_d - \Sigma\right)^{\dagger}_{km} \sum_{j=1}^{d} \big(\lambda_r \Sigma_{km} Q_{jr}^2 + \lambda_r \Sigma_{jm} Q_{jr} Q_{kr}\big) \\
&=\; \lambda_r \sum_{k,m=1}^{d} \left(\lambda_r \mathbf{I}_d - \Sigma\right)^{\dagger}_{km} \times \Sigma_{km} + \lambda_r^2 \sum_{k=1}^{d} Q_{kr} \sum_{m=1}^{d} \left(\lambda_r \mathbf{I}_d - \Sigma\right)^{\dagger}_{km} Q_{mr} \\
&=\; \lambda_r \mathrm{tr}\big[\left(\lambda_r \mathbf{I}_d - \Sigma\right)^{\dagger} \times \Sigma\big].
\end{aligned}
$$

In consideration of (5.4), (5.7),

$$\begin{aligned}\operatorname{tr}\big[\big(\lambda_r \mathbf{I}_d - \Sigma\big)^\dagger \times \Sigma\big] = \operatorname{tr}\Big(\sum_{\lambda_v \neq \lambda_r} \frac{\lambda_v}{\lambda_r - \lambda_v} q_v q_v^{\mathrm{T}}\Big) &= \sum_{\lambda_v \neq \lambda_r} \frac{\lambda_v}{\lambda_r - \lambda_v} \operatorname{tr}\big(q_v^{\mathrm{T}} q_v\big) \\ &= \sum_{\lambda_v \neq \lambda_r} \frac{\lambda_v}{\lambda_r - \lambda_v} \|q_v\|^2 = \sum_{\lambda_v \neq \lambda_r} \frac{\lambda_v}{\lambda_r - \lambda_v},\end{aligned}$$

thus

$$\sum_{j,k,l,m=1}^{d} \partial^2_{jk,lm} F_h(\Sigma) \left(\Sigma_{jl}\Sigma_{km} + \Sigma_{jm}\Sigma_{kl}\right) = 2 \sum_{\lambda_v \neq \lambda_{r_h}} \frac{\lambda_{r_h}\lambda_v}{\lambda_{r_h} - \lambda_v}. \tag{5.20}$$

5.3 Asymptotic normality of realized eigenvalues and eigenvectors

In this section, we provide a central limit theorem for realized (simple and repeated) eigenvalues of $\Sigma_t = \sigma_t \sigma_t^{\mathrm{T}}$, and a central limit theorem for realized eigenvectors associated with simple eigenvalues, based on **Theorem 4**.

Suppose (5.16), and let $\mathcal{K}_q = \{r_{q-1}+1, \cdots, r_q\}$ for $q = 1, \cdots, K$. Let

$$\lambda(\Sigma_t) = \left[\frac{1}{r_1 - r_0} \sum_{r \in \mathcal{K}_1} \lambda_r, \cdots, \frac{1}{r_K - r_{K-1}} \sum_{r \in \mathcal{K}_K} \lambda_r\right]^{\mathrm{T}}.$$

Note that $\lambda(\cdot)$ is a spectral function. The objects to estimate are

$$\begin{aligned} S(\lambda)_t &= \int_0^t \lambda(\Sigma_s)\,\mathrm{d}s \\ S(q_r)_t &= \int_0^t q_r(\Sigma_s)\,\mathrm{d}s,\ r = 1 \cdots, d. \end{aligned}$$

We estimate these two realized quantities by Riemann sums of the spot estimates $\lambda\big(\widehat{\Sigma}_h^{n,\mathrm{psd}}\big)$

and $q_r\big(\widehat{\Sigma}_h^{n,\mathrm{psd}}\big)$ at times $h = 0, k_n, 2k_n, \cdots$.

Define

$$
\begin{aligned}
\widehat{\lambda}_{r,h}^n &= \lambda_r\big(\widehat{\Sigma}_h^{n,\mathrm{psd}}\big) \\
\overline{\lambda}_{q,h}^n &= \frac{1}{r_q - r_{q-1}} \sum_{r\in\mathcal{K}_q} \widehat{\lambda}_{r,h}^n,
\end{aligned}
$$

then we can write

$$
\lambda(\widehat{\Sigma}_h^{n,\mathrm{psd}}) = \left[\overline{\lambda}_{1,h}^n, \cdots, \overline{\lambda}_{K,h}^n\right]^{\mathrm{T}}.
$$

We define the estimator of realized eigenvalues to be

$$
\begin{aligned}
\widehat{S}(\lambda)_t^n &= \left[\widehat{S}(\lambda)_{1,t}^n, \cdots, \widehat{S}(\lambda)_{K,t}^n\right]^{\mathrm{T}} \\
\widehat{S}(\lambda)_{r,t}^n &\equiv k_n\Delta_n \sum_{h=0}^{N_t^n-1} \left[1 - \frac{2\theta\Phi_{00}}{\psi_0^2 k_n \Delta_n^{1/2+\delta}} \sum_{v\notin\mathcal{K}_r} \frac{\widehat{\lambda}_{v,hk_n}}{\overline{\lambda}_{r,hk_n}^n - \widehat{\lambda}_{v,hk_n}^n}\right] \overline{\lambda}_{r,hk_n}^n, \qquad (5.21)
\end{aligned}
$$

and define the estimator of realized eigenvectors associated with simple eigenvalues to be

$$
\widehat{S}(q_r)_t^n \equiv k_n\Delta_n \sum_{h=0}^{N_t^n-1} \left[1 + \frac{\theta\Phi_{00}}{\psi_0^2 k_n \Delta_n^{1/2+\delta}} \sum_{v\neq r} \frac{\widehat{\lambda}_{r,hk_n}^n \widehat{\lambda}_{v,hk_n}^n}{\big(\widehat{\lambda}_{r,hk_n}^n - \widehat{\lambda}_{v,hk_n}^n\big)^2}\right] q_r(\widehat{\Sigma}_{hk_n}^{n,\mathrm{psd}}), \qquad (5.22)
$$

where θ, Φ_{00} and ψ_0 are defined in (4.17), (3.3), (3.12).

Remark 4. Suppose Σ is the covariance matrix of X, and we have $\Sigma q_r = \lambda_r q_r$ where λ_r is the r-th largest eigenvalue. When we permute the elements of X by a permutation matrix P, the covariance becomes $P\Sigma P^{\mathrm{T}}$. There is some reason to permute X in empirical finance, i.e., sort assets before covariance matrix estimation, cf. Fan et al. [2016].

Suppose $P\Sigma P^{\mathrm{T}}\widetilde{q}_r = \widetilde{\lambda}_r\widetilde{q}_r$, what is the relationship between $\widetilde{\lambda}_r$ and λ_r, the relationship between $\widetilde{q}_r$ and q_r?

Since P is an orthogonal matrix, $PP^{\mathrm{T}} = I$,

$$P\big(\Sigma P^{\mathrm{T}}\widetilde{q}_r - \widetilde{\lambda}_r P^{\mathrm{T}}\widetilde{q}_r\big) = 0,$$

because P is of full rank,

$$\Sigma P^{\mathrm{T}}\,\widetilde{q}_r = \widetilde{\lambda}_r\, P^{\mathrm{T}}\widetilde{q}_r,$$

hence

$$\begin{aligned}\widetilde{\lambda}_r &= \lambda_r \\ \widetilde{q}_r &= Pq_r.\end{aligned}$$

Proposition 2. *Assume* ***Assumption A-****c,* ***A-****γ,* ***A-****ν hold for $\nu \in [0,1)$. Given θ positive finite, we control the bandwidths l_n, k_n and the truncation threshold ν_n according to (4.25), then we have the following functional stable convergence in law:*

$$\Delta_n^{-1/4+\delta/2}\left[\widehat{S}(\lambda)^n - S(\lambda)\right] \overset{\mathcal{L}-s(f)}{\longrightarrow} Z(\lambda) \tag{5.23}$$

where $Z(\lambda)$ is a process defined on an extension of the space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\big)$, such that conditioning on $\mathcal{F}$ it is a continuous mean-0 Gaussian martingale with variance

$$\widetilde{E}[Z(\lambda)Z(\lambda)^{\mathrm{T}}|\mathcal{F}] = V(\lambda)$$

where $\widetilde{E}$ is the conditional expectation operator on the extended probability space and

$$V(\lambda)_t = \frac{4\theta\Phi_{00}}{\psi_0^2}\begin{bmatrix} \frac{1}{r_1-r_0}\int_0^t \lambda_{r_1}(\Sigma_s)^2\,\mathrm{d}s & & \\ & \ddots & \\ & & \frac{1}{r_K-r_{K-1}}\int_0^t \lambda_{r_K}(\Sigma_s)^2\,\mathrm{d}s \end{bmatrix} \tag{5.24}$$

Proof. This is the consequence of **Theorem 4**, (5.19), (5.20). $\square$

Proposition 3. *Assume* ***Assumption A****-c,* ***A****-γ,* ***A****-ν hold for $\nu \in [0,1)$ and $\lambda_{r,t}$ is a simple eigenvalue of Σ_t, $\forall t > 0$. Given θ positive finite, we control the bandwidths l_n, k_n and the truncation threshold ν_n in accordance with (4.25), then we have the following functional stable convergence in law:*

$$\Delta_n^{-1/4+\delta/2}\left[\widehat{S}(q_r)^n - S(q_r)\right] \xrightarrow{\mathcal{L}-s(f)} Z(q_r) \tag{5.25}$$

where $Z(q_r)$ is a process defined on an extension of the space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\big)$, such that conditioning on $\mathcal{F}$ it is a continuous mean-0 Gaussian martingale with variance

$$\widetilde{E}[Z(q_r)Z(q_r)^{\mathrm{T}}|\mathcal{F}] = V(q_r)$$

where $\widetilde{E}$ is the conditional expectation operator on the extended probability space and

$$V(q_r)_t = \frac{2\theta\Phi_{00}}{\psi_0^2}\int_0^t \sum_{v\neq r}\frac{\lambda_{r,s}\lambda_{v,s}}{(\lambda_{r,s}-\lambda_{v,s})^2}\, q_v(\Sigma_s)\, q_v(\Sigma_s)^{\mathrm{T}}\,\mathrm{d}s \tag{5.26}$$

Proof. This is the consequence of **Theorem 4**, (5.12), (5.15). $\square$

5.4 PCA of TAQ millisecond data: 2003-2019

So far, the author have demonstrated how to utilize the theoretical results on volatility matrix functionals in Chapter 4 to develop a useful statistical theory and method of realized PCA for non-stationary time-dependent noisy high-frequency data.

In this section, this statistical method is applied to the `TAQ` dataset. Snippets of this dataset are plotted in Figure 5.1 and Figure 5.2. Realized PCA is instrumental to reduce the apparent dimensionality and facilitate a better understanding of both temporal evolution and cross-section structures.

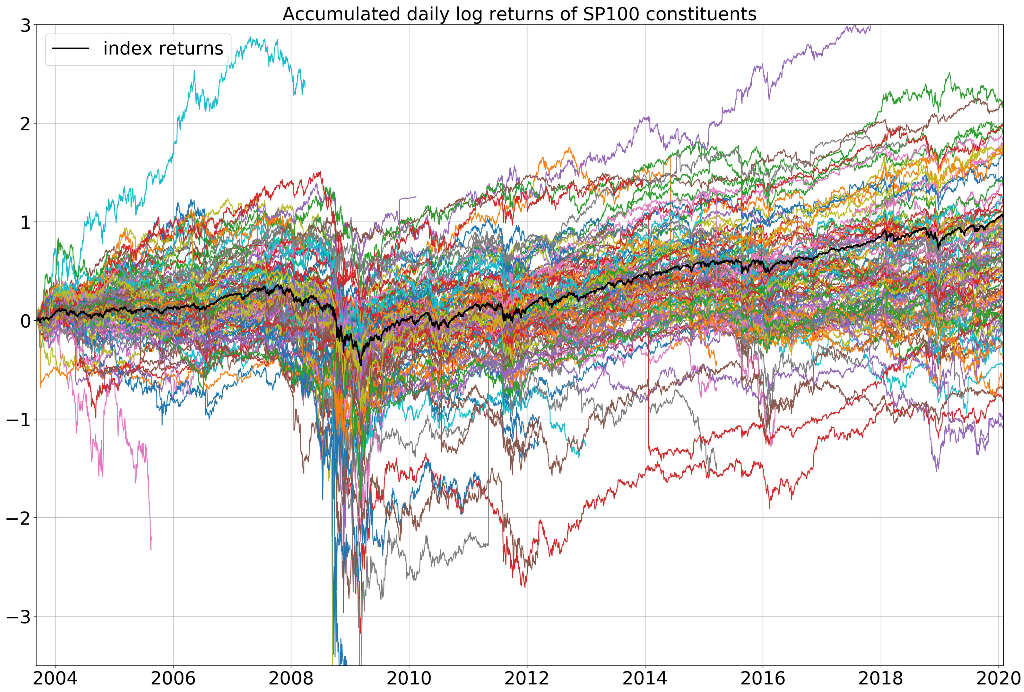


Figure 5.1: Accumulated daily log returns of S&P 100 stocks

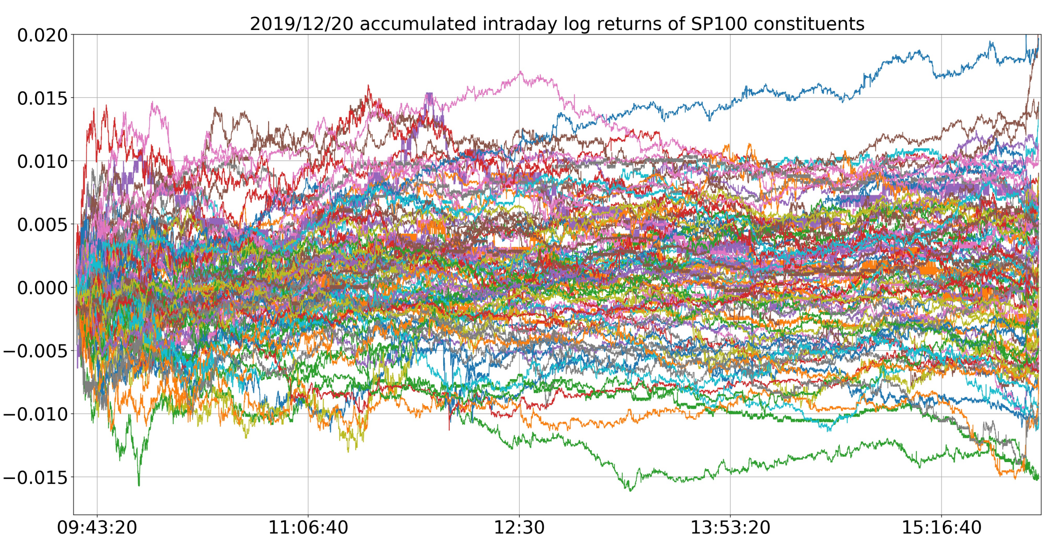


Figure 5.2: Accumulated intraday log returns of S&P 100 stocks on 2019/12/20

The data to be analyzed is transaction data of 90 S&P 100 stocks (the dimensionality is 90), sampled every second during 9:35 - 15:55 EST on each business day, from 2003/09/10 to 2019/12/31. There were 4106 business days in 852 weeks. By focusing on the time window 9:35 - 15:55 EST, we can remove overnight returns which contain big jumps and returns incurred in the first and last 5 minutes of trading hours which are relatively volatile. 10 illiquid assets out of the S&P 100 constituents were excluded from the statistical analysis.

The realized PCA ought to be applicable to financial data on the logarithmic scale, because asset prices are modeled by exponential jump-diffusions such as geometric Brownian motion.

Instantaneous volatility matrix and its eigenvalues and eigenvectors were estimated for every business day. The realized eigenvalues and realized eigenvectors were computed weekly by aggregating the instantaneous eigenvalues and instantaneous eigenvectors within each week.

The temporal evolution of four leading eigenvalue estimates along with statistical uncertainty are plotted in Figure 5.3. It shows $\lambda_q / \sum_j \lambda_j$ rather than the original scale, hence it shows the proportions of total variation that can be explained by principal components. The first eigenvalues are conspicuously separated from the others and indicate the first principal component explains more than 60% of the cross-section variation for the majority of time. Moreover, it is evident from this dataset that the first four leading eigenvalues were simple rather than repeated.

It is reasonable to interpret the first principal component as a market indicator on how the economy performs in general. As all the major corporations and their stock values are moving with the overall economy, the eigenvector corresponding to the first principal component is expected to comprise only positive entries.

An interesting question is how the first principal component is compared with S&P 100 index.

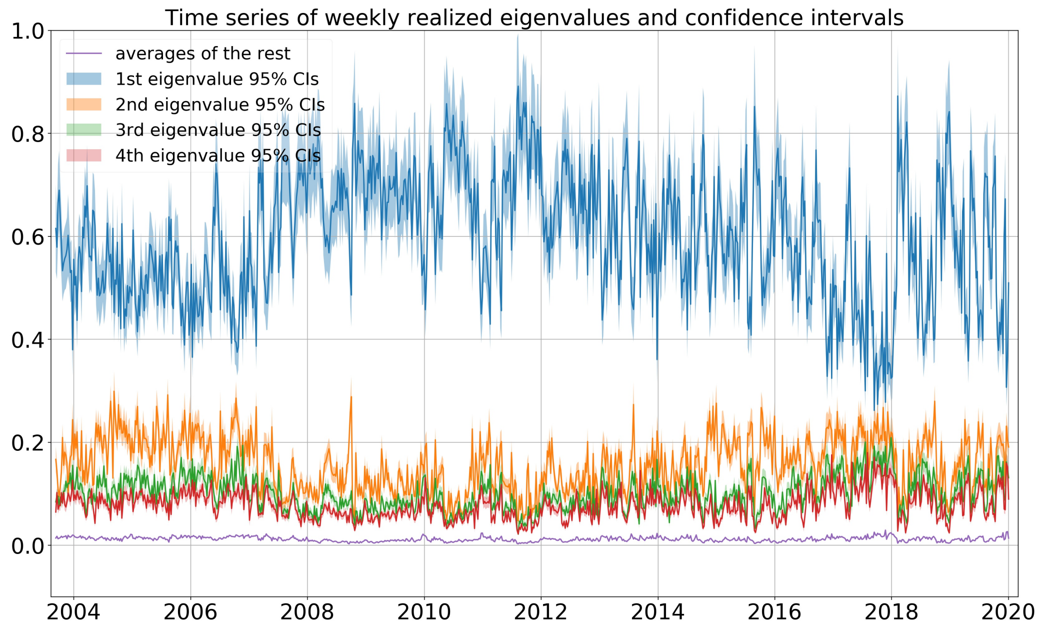


Figure 5.3: Weekly integrated eigenvalues and the confidence intervals

Chen et al. [2019] calculated the accumulated returns from 2007 to 2017 of the portfolio based on the eigenvector q_1 associated with the first principal component. Chen et al. [2019] normalized the eigenvector so that $\sum_j^d q_{1,j} = 1$ and used $q_{1,j}$'s as portfolio weights so that the resultant portfolio is self-financing. A curious finding of Chen et al. [2019] is that this portfolio outperformed the market index hence beat the passive investment. Here the author would like to test this portfolio using pre-averaging-based PCA, over a longer time span.

The principal components are $\int_0^t q_r(\Sigma_{s-})\, \mathrm{d}X_s$, $r = 1, 2, \cdots$, where $X_t^j = \log(S_t^j)$ and S_t is the vector of asset prices at time t. However, one obviously can not transact in logarithmic currency, so a pending question is how to find a portfolio that is analogous to the first principal component and requires trading on the currency's original scale.

Chen et al. [2019] proposed the following construction. Let

$$R_{t,\Delta} = \left[(S_{1,t+\Delta} - S_{1,t})/S_{1,t}, \cdots, (S_{d,t+\Delta} - S_{d,t})/S_{d,t}\right]^{\mathrm{T}}$$

be the vector of returns in percentage. When $\|R_{t,\Delta}\|$ is small,

$$\log\left(1 + q_1^{\mathrm{T}} R_{t,\Delta}\right) \approx q_1^{\mathrm{T}} R_{t,\Delta},$$

and

$$q_{1,j} \frac{S_{j,t+\Delta} - S_{j,t}}{S_{j,t}} \approx q_{1,j} \log\left(1 + \frac{S_{j,t+\Delta} - S_{j,t}}{S_{j,t}}\right) = q_{1,j} \log\left(\frac{S_{j,t+\Delta}}{S_{j,t}}\right),$$

so

$$q_1^{\mathrm{T}}(X_{t+\Delta} - X_t) \approx \log\left(1 + q_1^{\mathrm{T}} R_{t,\Delta}\right).$$

Therefore, it is sensible to regard $\sum_t \log(1+q_1^{\mathrm{T}} R_{t,\Delta})$ as an approximation to the first realized principal component, which is also the accumulated log return if there is no transaction cost and $\sum_j^d q_{1,j} = 1$. In this case, $\exp\left\{\sum_t \log(1 + q_1^{\mathrm{T}} R_{t,\Delta})\right\} - 1 = \prod_t (1 + q_1^{\mathrm{T}} R_{t,\Delta}) - 1$ is the accumulated return of this self-financing portfolio in the absence of transaction cost.

Since the volatility matrix and the eigenvector q_1 are updated on every business day, the portfolio can be re-weighted daily. One caveat is that the numerical values of eigenvector estimates are relative less stable. Roughly speaking, the eigenvector viewed as a function of the corresponding positive semidefinite matrix is less smooth than the eigenvalue. Due to this reason, heuristically, an error in the covariance matrix estimation tends to result in a larger error in the eigenvector estimate, as compared to the eigenvalue counterpart. To stabilize eigenvector estimates, it is advisable to compute moving averages of the initial eigenvector estimates over a window of six business days, hence the author re-weighted the portfolio on all the business days from 2003/09/17 to 2020/12/31.

Before the portfolio re-allocation according to the new weights on the next business day, one can either let positions stand overnight or clear positions at the end of trading hours. The former allows overnight returns to be absorbed into the portfolio returns.

The accumulated returns of the portfolio that mimics the first principal component are plotted in Figure 5.4. The portfolio returns were computed using daily prices of S&P 100 constituents which can be obtained from the `Compustat` database. As we can see, morphologies of the portfolios that mimics the first principal component are similar to that of S&P 100 index. The portfolio that was cleared at the end of trading hours slightly underperformed the index. Interestingly enough, the portfolio that stood overnight manifestly outperformed the index.

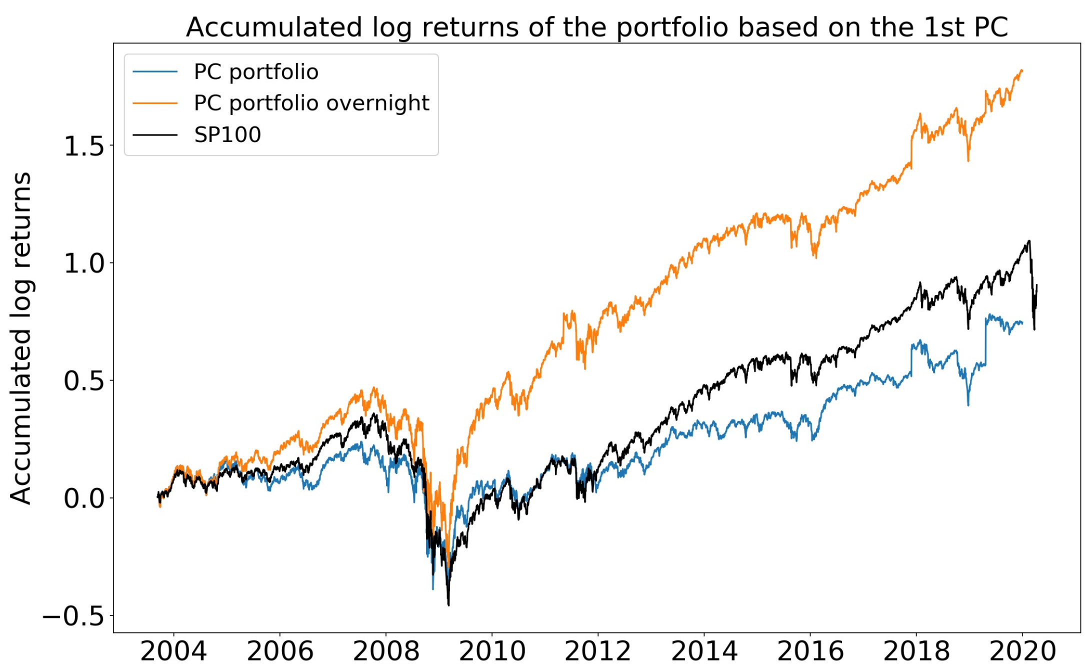


Figure 5.4: Accumulated log returns of the portfolio based on the first principal component

CHAPTER 6

VOLATILITY MATRIX ESTIMATION WITH MISSING DATA: FOURIER-MALLIAVIN METHOD

6.1 The Fourier-Malliavin method

In this chapter, the author considers a different setting: there is no observational noise present in the data, but the data is sampled irregularly in each dimension and asynchronously across different dimensions. The absence of noise means we can directly observe the values of X at discrete time points. However, the pre-averaging method in this dissertation is designed for regularly and synchronously sampled data, hence an new volatility estimator is needed.

This new instantaneous volatility matrix estimator is inspired by the premise of Fourier analysis to avoid temporal alignment or data interpolation altogether. Given a function φ on $[0, T]$, for $q \in \mathbb{N}^+$, define *Fourier transform* and *Fourier-Stieltjes transform* as

$$\begin{aligned} F(\varphi)_q &= \textstyle\int_0^T \varphi(u)\, e^{-i2\pi qu/T}\, \mathrm{d}u \\ F(\mathrm{d}\varphi)_q &= \textstyle\int_0^T e^{-i2\pi qu/T}\, \mathrm{d}\varphi(u). \end{aligned}$$

If φ is periodic, it can be expanded into Fourier series:

$$\varphi(t) = \frac{1}{T} \sum_{q=-\infty}^{\infty} F(\varphi)_q\, e^{i2\pi qt/T}. \tag{6.1}$$

Define

$$\widehat{\varphi}^M(t) = \frac{1}{T} \sum_{q=-M+1}^{M-1} \left(1 - \frac{|q|}{M}\right) F(\varphi)_q\, e^{i2\pi qt/T}, \tag{6.2}$$

by a classic result of approximation theory,

$$\sup_{t\in[0,T]} \|\widehat{\varphi}^M(t) - \varphi(t)\| \leq \lambda_\varphi(4/M), \tag{6.3}$$

where λ_φ is the modulus of continuity defined by

$$\lambda_f(\varphi) = \sup_{\|x-y\|\leq\Delta} \|\varphi(x) - \varphi(y)\|. \tag{6.4}$$

In this and the next chapters, the notation introduced in section 2.2 is still in force. To avoid notational clutter, a slight change is adopted: the component index is written as subscript instead of superscript, i.e., for a $\mathbb{R}^d$-valued (resp. $\mathbb{R}^{d\times d}$-valued) process U, U_j (resp. U_{jk}) is the j-th (resp. (j,k)-th) component of U.

In statistical applications, we can not evaluate the exact Fourier transform nor Fourier-Stieltjes transform because of discrete sampling, i.e., our signal is digital rather than analog. However, according to Malliavin and Mancino [2002, 2009], we can approximate the Fourier-Stieltjes transform of X and the Fourier transform of c over $[0,T]$ by the following quantities

$$\begin{aligned}
\widehat{F}(\mathrm{d}X_j)^n_s &\equiv \sum_{h=1}^{n_j} e^{-i2\pi s\tau^j_h/T}\,\delta^j_h(X_j), && (6.5)\\
\widehat{F}(c_{jk})^{n,N}_q &\equiv \frac{1}{2N+1}\sum_{|s|\leq N} \widehat{F}(\mathrm{d}X_j)^n_{q-s} \times \widehat{F}(\mathrm{d}X_k)^n_s. && (6.6)
\end{aligned}$$

The available frequency coordinates for $\widehat{F}(\mathrm{d}X_j)^n_s$ are $0, \pm1, \cdots, \pm\lfloor n_j/2\rfloor$, and given $N \leq \lfloor n_k/2\rfloor$ the available frequency coordinates for $\widehat{F}(c_{jk})^{n,N}_q$ are $0, \pm1, \cdots, \pm(\lfloor n_j/2\rfloor - N)$.

Based on the Fourier coefficient estimates $\widehat{F}(c_{jk})^{n,N}_q$'s, the spot volatility can be estimated by *Fourier-Fejér inversion*

$$\widehat{c}_{jk}^{n,N,M}(t) \equiv \frac{1}{T} \sum_{|q|<M} \Big(1 - \frac{|q|}{M}\Big) \widehat{F}(c_{jk})_q^{n,N} \, e^{i2\pi qt/T}, \tag{6.7}$$

where $M \leq \lfloor n_j/2 \rfloor - N + 1$.

Remark 5. The estimator $\widehat{F}(\mathrm{d}X_j)_s^n$ is the discrete Fourier transform (hereafter DFT) of the increments of X_j; the volatility spectrum estimator $\widehat{F}(c_{jk})_q^{n,N}$ is based on the idea akin to that of *Bohr convolution*, i.e., a scaled discrete convolution of the finite sequences $\widehat{F}(\mathrm{d}X_j)_s^n$'s and $\widehat{F}(\mathrm{d}X_k)_s^n$'s; the spot volatility estimator $\widehat{c}_{jk}^{n,N,M}(t)$ is the M-order Cesàro summation of inverse discrete Fourier transforms (hereafter IDFT) of the Fourier coefficient estimates, in the other word, IDFT with Fejér kernel.

Remark 6. It is worthwhile to understand the fundamental constraints on the tuning parameters N and M in the estimators $\widehat{F}(c_{jk})_q^{n,N}$ and $\widehat{c}_{jk}^{n,N,M}(t)$, namely,

$$\begin{cases} N & \leq & \lfloor n_k/2 \rfloor \wedge \big(\lfloor n_j/2 \rfloor - M + 1\big) \\ M & \leq & \lfloor n_j/2 \rfloor - N + 1. \end{cases}$$

Firstly, we start from the Fourier-Stieltjes transform (6.5), in order to avoid aliasing due to discrete sampling of the continuous-time signal, we only compute $\widehat{F}(\mathrm{d}X_j)_s^n$ for $|s| \leq \lfloor n_j/2 \rfloor$[1].

Secondly, due to the form of convolution in the definition (6.6) and the constraint by the Nyquist frequency, it follows that $N \leq \lfloor n_k/2 \rfloor$.

Thirdly, according to the definition (6.7) and the Nyquist frequency, $M \leq \lfloor n_j/2 \rfloor - N + 1$.[2]

1. It is the so-called *Nyquist frequency (folding frequency)*, which is the highest frequency coordinate without aliasing.

2. An interesting modification is to use different N's for different frequency coordinate q's. For instance, when estimating a low frequency we can use a relative large N to increase accuracy.

6.1.1 its relation with kernel methods

Here we provide some intuition of the Fourier-Malliavin method of Malliavin and Mancino [2009], and compare it with other nonparametric volatility estimators.

Suppose $d = 1$ and we observe the univariate process at times $\{\tau_0, \tau_1, \cdots, \tau_n\}$, and let $\delta_h = \delta_h^1$ be the first-order difference operator. Based on (6.6) and (6.7), one has

$$\begin{aligned}
\widehat{F}(c)_q^{n,N} &= \frac{1}{2N+1} \sum_{|s|\le N} \sum_{h=1}^{n} \sum_{v=1}^{n} e^{-i2\pi q \tau_h/T} e^{i2\pi s(\tau_h-\tau_v)/T} \delta_h(X)\delta_v(X) \\
&= \sum_{h=1}^{n} \sum_{v=1}^{n} \Big(\frac{1}{2N+1} \sum_{|s|\le N} e^{i2\pi s(\tau_h-\tau_v)/T}\Big) e^{-i2\pi q \tau_h/T} \delta_h(X)\delta_v(X) \\
&= \sum_{h=1}^{n} e^{-i2\pi q \tau_h/T} \delta_h(X)^2 + \sum_{h\ne v} e^{-i2\pi q \tau_h/T} \Big(\sum_{|s|\le N} e^{i2\pi s(\tau_h-\tau_v)/T}\Big) \frac{\delta_h(X)\delta_v(X)}{2N+1} \\
&= \sum_{h=1}^{n} e^{-i2\pi q \tau_h/T} \delta_h(X)^2 + \sum_{h\ne v} e^{-i2\pi q \tau_h/T} D^N\Big(\frac{\tau_h-\tau_v}{T}\Big) \frac{\delta_h(X)\delta_v(X)}{2N+1}, \qquad (6.8)
\end{aligned}$$

where $D^N(\cdot)$ is a kernel function defined later in (6.16). Furthermore, based on (6.7) and (6.8), one has

$$\begin{aligned}
\widehat{c}^{n,N,M}(t) &= \frac{1}{T} \sum_{h=1}^{n} \sum_{|q|<M} \Big(1 - \frac{|q|}{M}\Big) e^{i2\pi q(t-\tau_h)/T} \delta_h(X)^2 \\
&\quad + \frac{1}{T} \sum_{h\ne v} \sum_{|q|<M} \Big(1 - \frac{|q|}{M}\Big) e^{i2\pi q(t-\tau_h)/T}\, D^N\Big(\frac{\tau_h-\tau_v}{T}\Big) \frac{\delta_h(X)\delta_v(X)}{2N+1} \\
&= \frac{1}{T} \sum_{h=1}^{n} F^M\Big(\frac{t-\tau_h}{T}\Big) \delta_h(X)^2 + \frac{1}{T} \sum_{h\ne v} F^M\Big(\frac{t-\tau_h}{T}\Big) D^N\Big(\frac{\tau_h-\tau_v}{T}\Big) \frac{\delta_h(X)\delta_v(X)}{2N+1}, \quad (6.9)
\end{aligned}$$

where $F^M(\cdot)$ is another kernel function defined later in (6.18).

Figure 6.1 shows some examples of the kernel functions $F^M(\cdot)$ and $D^N(\cdot)$. There are some wiggles away from the origin due to the fact that they are trigonometric polynomials. As N

and M become large, the kernel put more weight around the origin. Some of their analytical properties are given in Section 6.3.

If one interprets $\delta_h(X)^2$ as a proxy of c_{τ_h}, then $\widehat{F}(c)_q^{n,N}$ as an estimator of the Fourier coefficient of c is a combination of the DFT of the proxies $\delta_j(X)^2$'s and cross terms involving the sample auto-covariances weighted by the kernel $D^N(\cdot)$:

$$\widehat{F}(c)_q^{n,N} = \text{DFT of volatility proxies} + \text{weighted sum of sample auto-covariance;}$$

similarly, $\widehat{c}^{n,N,M}(t)$ can be interpreted as a kernel estimator plus cross terms, the kernel is $F^M(\cdot)$ and the cross terms are sample auto-covariances weighted by both $F^M(\cdot)$ and $D^N(\cdot)$:

$$\widehat{c}^{n,N,M}(t) = \text{a kernel estimator} + \text{weighted sum of sample auto-covariance.}$$

Naturally, we shall ask, given the possible variations due to the cross terms, why not just use the DFT of volatility proxies $\delta_j(X)$'s to estimate the Fourier coefficients? Why not just use the kernel estimator to estimate the spot volatility?

There are a few considerations, here are two significant merits of the Fourier-based estimators.

- In multivariate settings, (6.6) and (6.7) can estimate the spot co-volatility and its Fourier coefficients when different processes are observed asynchronously, because the Bohr convolution is computed in the frequency domain and one does not need to worried about data misalignment in the time domain; however many other estimators require data alignment as a prerequisite.

- When the sampling frequency is high enough so that microstructure noise ε is present, given an appropriate choice of N and M, the estimators (6.6) and (6.7) are still consistent; whereas $\delta^j(X+\varepsilon)^2$ can no longer be a good proxy for volatility.

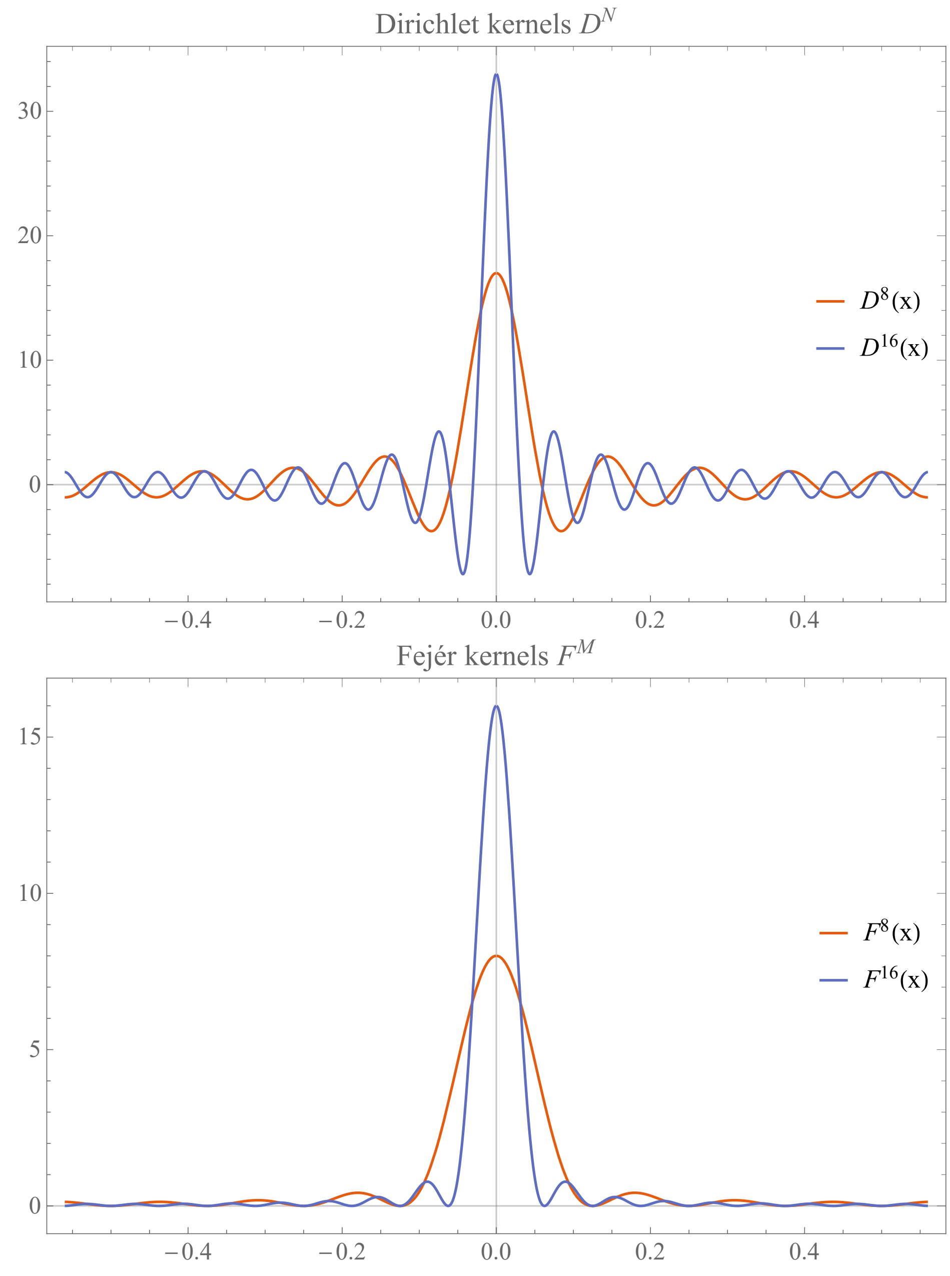


Figure 6.1: Dirichlet kernels and Fejér kernels

6.1.2 additional assumptions

Signal reconstruction by the Fourier transform suffers from the so-called *Gibbs phenomenon* if the signal exhibits any jump discontinuity. To avoid the Gibbs phenomenon, I need to impose the continuity assumption on the data generating process X.

Assumption A-X (continuity)**.** $\delta(t,x)=0$ *in (2.1) and* $\delta^c(t,x)=0$ *in (2.3) almost surely for* $t\in[0,T]$.

In order to establish an inferential theory for the Fourier-Malliavin method, it is necessary for us to put constraints on the smoothness of volatility sample paths.

Assumption A-α (volatility regularity)**.** $c(0)=c(T)$ *almost surely. The modulus of continuity of* c *satisfies*

$$\omega_c(\Delta)\leq\Delta^\alpha,$$

where $\alpha>0$, ω_c *is defined in (6.4).*

Remark 7. Assumption A-X and **A-ν** are needed for consistency; **Assumption A-α** with $\alpha>1/2$ is further needed for central limit theorems. **Assumption A-ν**, **A-α** can be rephrased that the volatility as a function of time belong to the Hölder ball

$$H^\alpha(K):=\left\{f\in C([0,T])\,\middle|\,\sup_{t\in[0,T]}\|f(t)\|+\sup_{t\neq u}\frac{\|f(t)-f(u)\|}{|t-u|^\alpha}\leq K\right\}$$

for some $K>0$.

6.2 Estimating the Fourier coefficients of volatility

In this section, we study the asymptotic theory for volatility spectrum inference, i.e., the convergence of $\widehat{F}(c)_q^{n,N}$.

First, we introduce a short-hand representation of the data generating process X under

Assumption A-X:

$$X = X(0) + A + M,$$

where $A(t) = \int_0^t b(u)\,\mathrm{d}u,\; M(t) = \int_0^t \sigma(u)\,\mathrm{d}W(u)$.

Note that

$$\widehat{F}(c_{jk})_q^{n,N} - F(c_{jk})_q = R(0)_{jk,q}^{n,N} + R(1)_{jk,q}^{n,N} + R(2)_{jk,q}^{N}, \tag{6.10}$$

where

$$\begin{aligned}
R(0)_{jk,q}^{n,N} &= \frac{1}{2N+1}\sum_{|s|\le N}\left[\widehat{F}(\mathrm{d}X_j)_{q-s}^n \widehat{F}(\mathrm{d}X_k)_s^n - \widehat{F}(\mathrm{d}M_j)_{q-s}^n \widehat{F}(\mathrm{d}M_k)_s^n\right]\\
R(1)_{jk,q}^{n,N} &= \frac{1}{2N+1}\sum_{|s|\le N}\left[\widehat{F}(\mathrm{d}M_j)_{q-s}^n \widehat{F}(\mathrm{d}M_k)_s^n - F(\mathrm{d}M_j)_{q-s} F(\mathrm{d}M_k)_s\right]\\
R(2)_{jk,q}^{N} &= \frac{1}{2N+1}\sum_{|s|\le N} F(\mathrm{d}M_j)_{q-s} F(\mathrm{d}M_k)_s - F(c_{jk})_q.
\end{aligned}$$

$R(0)_{jk,q}^{n,N}$ is the effect of the drift term; $R(1)_{jk,q}^{n,N}$ is the effect of discrete observations of the continuous-time model; $R(2)_{jk,q}^{N}$ is the error due to finite Bohr convolution.

Study of $R(1)^{n,N}$

We have

$$\widehat{F}(\mathrm{d}M_j)_q^n - F(\mathrm{d}M_j)_q = \int_0^T \lambda_{j,q}^n(t)\,\mathrm{d}M_j(t),$$

where

$$\lambda_{j,q}^n(t) = \sum_{h=1}^{n^j} e^{-i2\pi q\tau_h^j/T}\left[1 - e^{-i2\pi q(t-\tau_h^j)/T}\right]\mathbb{1}_{I_h^j}(t). \tag{6.11}$$

Note $|\lambda^n_{j,q}(t)| \le KT^{-1}q\Delta(n)$, by *Burkholder-Gundy inequality,*

$$\begin{aligned}\mathbb{E}\big(|\widehat{F}(\mathrm{d}M_j)^n_{q-s}|^4\big) &\le KT^2\\ \mathbb{E}\big(|F(\mathrm{d}M_j)_q|^4\big) &\le KT^2\\ \mathbb{E}\big(|\widehat{F}(\mathrm{d}M_j)^n_q - F(\mathrm{d}M_j)_q|^4\big) &\le KT^{-2}q^4\Delta(n)^4,\end{aligned}$$

by *Cauchy-Schwarz inequality,*

$$\begin{aligned}&\mathbb{E}\big(|\widehat{F}(\mathrm{d}M_j)^n_{q-s}\widehat{F}(\mathrm{d}M_k)^n_s - F(\mathrm{d}M_j)_{q-s}F(\mathrm{d}M_k)_s|^2\big)\\ &\qquad\le 2\Big[\mathbb{E}\big(|\widehat{F}(\mathrm{d}M_j)^n_{q-s}|^4\big)^{1/2}\cdot\mathbb{E}\big(|\widehat{F}(\mathrm{d}M_k)^n_s - F(\mathrm{d}M_k)_s|^4\big)^{1/2}\\ &\qquad\qquad + \mathbb{E}\big(|F(\mathrm{d}M_k)_s|^4\big)^{1/2}\cdot\mathbb{E}\big(|\widehat{F}(\mathrm{d}M_j)^n_{q-s} - F(\mathrm{d}M_j)^n_{q-s}|^4\big)^{1/2}\Big],\end{aligned}$$

so

$$\mathbb{E}\big(|R(1)^{n,N}_{jk,q}|^2\big) \le KN^2\Delta(n)^2. \tag{6.12}$$

Study of $R(2)^N$

Define a $\mathbb{C}$-valued martingale $\Gamma^j_q(t) = \int_0^t e^{-i2\pi qu/T}\,\mathrm{d}M_j(u)$ for $j = 1,\cdots,d$. We see that $\Gamma^j_q(T) = F(\mathrm{d}M_j)_q$. By Itô's formula,

$$\Gamma^j_{q-s}(T)\cdot\Gamma^k_s(T) = F(c_{jk})_q + \int_0^T \Gamma^j_{q-s}(t)\,\mathrm{d}\Gamma^k_s(t) + \int_0^T \Gamma^k_s(t)\,\mathrm{d}\Gamma^j_{q-s}(t),$$

hence

$$R(2)^N_{jk,q} = \Lambda(1)^N_{jk,q} + \Lambda(2)^N_{jk,q},$$

where

$$\begin{aligned}\Lambda(1)^N_{jk,q} &= \tfrac{1}{2N+1}\textstyle\sum_{|s|\le N}\int_0^T \Gamma^j_{q-s}(t)\,\mathrm{d}\Gamma^k_s(t)\\ \Lambda(2)^N_{jk,q} &= \tfrac{1}{2N+1}\textstyle\sum_{|s|\le N}\int_0^T \Gamma^k_s(t)\,\mathrm{d}\Gamma^j_{q-s}(t).\end{aligned}$$

By (6.16), we have

$$\begin{aligned}
\Lambda(1)^N_{jk,q} &= \int_0^T \sigma_{k,\cdot}(t)\,\mathrm{d}W(t) \int_0^t e^{-i2\pi qu/T}\frac{1}{2N+1}D^N\Big(\frac{u-t}{T}\Big)\sigma_{j,\cdot}(t)\,\mathrm{d}W(u)\\
\Lambda(2)^N_{jk,q} &= \int_0^T e^{-i2\pi qt/T}\sigma_{j,\cdot}(t)\,\mathrm{d}W(t) \int_0^t \frac{1}{2N+1}D^N\Big(\frac{u-t}{T}\Big)\sigma_{k,\cdot}(t)\,\mathrm{d}W(u),
\end{aligned}$$

by Itô isometry[3], (6.27) and (6.21),

$$\begin{aligned}
\mathbb{E}\Big(|\Lambda(1)^N_{jk,q}|^2\Big) = \mathbb{E}&\left[\int_0^T \Big(\int_0^t e^{-i2\pi qu/T}\frac{1}{2N+1}D^N\Big(\frac{u-t}{T}\Big)\,\mathrm{d}X_j(u)\Big)^2 c_{kk}(t)\,\mathrm{d}t\right]\\
&\asymp \frac{1}{2N+1}\int_0^T\int_0^t F^{2N+1}\Big(\frac{u-t}{T}\Big)\,\mathrm{d}u\,\mathrm{d}t \asymp \frac{T^2}{N},
\end{aligned}$$

the term $\Lambda(2)^N_{jk,q}$ can be bounded by a similar argument, so

$$\mathbb{E}\big(\big|R(2)^N_{jk,q}\big|\big) \asymp TN^{-1/2}. \tag{6.13}$$

Study of $R(0)^{n,N}$

For a generic scalar process, we can write $\widehat{F}(\mathrm{d}U)^n_q = \int_0^T \beta^n_{j,q}(t)\,\mathrm{d}U(t)$ for $j=1,\cdots,d$, where $\beta^n_{j,q}(t) = \sum_{h=1}^{n^j} e^{-i2\pi q\tau^j_h/T}\mathbb{1}_{I^j_h}(t)$.

By linearity of discrete Fourier transform, $\widehat{F}(\mathrm{d}X_j)^n_q = \widehat{F}(\mathrm{d}A^j)^n_q + \widehat{F}(\mathrm{d}M_j)^n_q$, so

$$\begin{aligned}
&\widehat{F}(\mathrm{d}X_j)^n_{q-s}\widehat{F}(\mathrm{d}X_k)^n_s - \widehat{F}(\mathrm{d}M_j)^n_{q-s}\widehat{F}(\mathrm{d}M_k)^n_s =\\
&\qquad \widehat{F}(\mathrm{d}A_j)^n_{q-s}\widehat{F}(\mathrm{d}A_k)^n_s + \widehat{F}(\mathrm{d}A_j)^n_{q-s}\widehat{F}(\mathrm{d}M_k)^n_s + \widehat{F}(\mathrm{d}A_k)^n_s\widehat{F}(\mathrm{d}M_j)^n_{q-s},
\end{aligned}$$

3. It is also called as Itô energy identity.

By *Parseval's identity*,

$$\sum_{s=-\infty}^{\infty} |F(\mathrm{d}A_j)_s|^2 = \int_0^T |b_j(t)|^2 \, \mathrm{d}t < \infty,$$

note

$$\widehat{F}(\mathrm{d}A_j)_q^n - F(\mathrm{d}A_j)_q = \int_0^T \lambda_{j,q}^n(t) \, b_j(t) \, \mathrm{d}t,$$

where $\lambda_{j,q}^n(t)$ is defined in (6.11). By Cauchy-Schwarz inequality,

$$|R(0)_{jk,q}^{n,N}| \leq \frac{K}{2N+1} \int_0^T \|b(t)\|^2 \, \mathrm{d}t + \Big(\frac{K}{2N+1} \int_0^T \|b(t)\|^2 \, \mathrm{d}t \Big)^{1/2} \times \\ \Bigg[\Big(\frac{1}{2N+1} \sum_{|s| \leq N} |F(\mathrm{d}M_k)_s|^2 \Big)^{1/2} + \Big(\frac{1}{2N+1} \sum_{|s| \leq N} |F(\mathrm{d}M_j)_{q-s}|^2 \Big)^{1/2} \Bigg].$$

From the study of the term $R(2)_{jk,q}^N$, we know

$$\sum_{|s| \leq N} F(\mathrm{d}M_k)_s^2 = \int_0^T D^N \Big(\frac{2t}{T} \Big) c_{kk}(t) \, \mathrm{d}t \\ + 2 \int_0^T \sigma_{k,\cdot}(t) \, \mathrm{d}W(t) \int_0^t D^N \Big(\frac{t+u}{T} \Big) \sigma_{k,\cdot}(u) \, \mathrm{d}W(u),$$

by Cauchy-Schwarz inequality, (6.27), (6.21),

$$\frac{1}{2N+1} \int_0^T D^N \Big(\frac{2t}{T} \Big) c_{kk}(t) \, \mathrm{d}t \\ \leq \frac{1}{\sqrt{2N+1}} \Big(\int_0^T F^{2N+1} \Big(\frac{2t}{T} \Big) \, \mathrm{d}t \Big)^{1/2} \Big(\int_0^T c_{kk}(t)^2 \, \mathrm{d}t \Big)^{1/2} \leq KTN^{-1/2},$$

by Jensen's inequality, Burkholder-Gundy inequality, (6.27), (6.21),

$$
\begin{aligned}
&\mathbb{E}\Big[\frac{1}{2N+1}\int_0^T \sigma_{k,\cdot}(t)\,\mathrm{d}W(t)\int_0^t D^N\Big(\frac{t+u}{T}\Big)\sigma_{k,\cdot}(u)\,\mathrm{d}W(u)\Big] \\
&\qquad\qquad \leq \frac{1}{2N+1}\mathbb{E}\Big[\Big(\int_0^T \sigma_{k,\cdot}(t)\,\mathrm{d}W(t)\int_0^t D^N\Big(\frac{t+u}{T}\Big)\sigma_{k,\cdot}(u)\,\mathrm{d}W(u)\Big)^2\Big]^{1/2} \\
&\qquad\qquad\qquad\qquad\qquad \leq \frac{K}{\sqrt{2N+1}}\Big[\int_0^T\int_0^t F^{2N+1}\Big(\frac{t+u}{T}\Big)\,\mathrm{d}u\,\mathrm{d}t\Big]^{1/2} \leq KTN^{-1/2},
\end{aligned}
$$

hence $\frac{1}{2N+1}\sum_{|s|\leq N}|F(\mathrm{d}M_k)_s|^2 \leq KTN^{-1/2}$.

Similarly, $\frac{1}{2N+1}\sum_{|s|\leq N}|F(\mathrm{d}M_k)_{q-s}|^2 \leq KTN^{-1/2}$. Thus

$$
\mathbb{E}\big(\big|R(0)^{n,N}_{jk,q}\big|\big) \leq KTN^{-3/4}. \tag{6.14}
$$

According to (6.12), (6.13) and (6.14), we have the following lemma.

Proposition 4. *Under **Assumption A**-ν, **A**-X,*

$$
\mathbb{E}\Big(\Big|\widehat{F}(c_{jk})^{n,N}_q - F(c_{jk})_q\Big|\Big) \leq K\Big[N\Delta(n) + TN^{-1/2}\Big].
$$

Remark 8. For volatility spectrum estimation on a finite time horizon, depending on the tuning parameter N, we have the following magnitudes for various error terms in Table 6.1.

Table 6.1: Estimation errors in volatility spectrum estimation

error sources	drift effect	discretization & asynchronicity effects	statistical error
magnitudes	$\leq KN^{-3/4}$	$\leq KN\Delta(n)$	$\asymp N^{-1/2}$

Remark 9. A good piece of news is that the size of the drift effect is dominated by other terms regardless of the choice N. In subsequent asymptotic analysis of the volatility spectrum

estimator, we can safely assume, without loss of generality,

$$X(t) = X(0) + \int_0^t \sigma(u)\,\mathrm{d}W(u). \tag{6.15}$$

6.3 Dirichlet kernel and Fejér kernel

For asymptotic analysis, some trigonometric polynomials pervasive in Fourier analysis will be introduced. Define the q-order *Dirichlet kernel* as

$$D^q(x) = \sum_{|s|\le q} e^{i2\pi sx}, \tag{6.16}$$

we have, for $x \notin \mathbb{N}$,

$$\begin{aligned} D^q(x) &= e^{-i2\pi qx}\frac{1-e^{i2\pi(2q+1)x}}{1-e^{i2\pi x}} = \frac{e^{i2\pi(2q+1)x/2}-e^{-i2\pi(2q+1)x/2}}{e^{i2\pi x/2}-e^{-i2\pi x/2}} \\ &= \frac{\sin[\pi(2q+1)x]}{\sin(\pi x)}, \end{aligned}$$

hence

$$D^q(x) = \begin{cases} \frac{\sin[\pi(2q+1)x]}{\sin(\pi x)}, & x \notin \mathbb{N} \\ 2q+1, & x \in \mathbb{N}. \end{cases} \tag{6.17}$$

Based on Dirichlet kernels, we define *Fejér kernel* of order M as

$$F^M(x) = \frac{1}{M}\sum_{q=0}^{M-1} D^q(x), \tag{6.18}$$

we have

$$F^M(x) = \sum_{|q|<M}\Big(1-\frac{|q|}{M}\Big)e^{i2\pi qx}. \tag{6.19}$$

Note for $x \notin \mathbb{N}$,

$$
\begin{aligned}
MF^M(x) = \sum_{q=0}^{M-1} \frac{\sin[\pi(2q+1)x]}{\sin(\pi x)} &= \frac{1}{\sin(\pi x)} \mathrm{Im}\Big(\sum_{q=0}^{M-1} e^{i2\pi(q+1/2)x} \Big) \\
&= \frac{1}{\sin(\pi x)} \mathrm{Im}\Big(\frac{e^{i2\pi Mx} - 1}{e^{i\pi x} - e^{-i\pi x}} \Big) = \frac{1 - \cos(2\pi Mx)}{2\sin(\pi x)^2},
\end{aligned}
$$

hence we have

$$
F^M(x) = \begin{cases} \frac{\sin(\pi Mx)^2}{M\sin(\pi x)^2}, & x \notin \mathbb{N} \\ M, & x \in \mathbb{N}, \end{cases} \tag{6.20}
$$

and

$$
F^{2M+1}(x) = \frac{1}{2M+1} D^M(x)^2. \tag{6.21}
$$

Furthermore, we have the following result,

$$
\frac{1}{M} \int_0^T F^M\Big(\frac{t}{T}\Big)^2 \mathrm{d}t \to \frac{2T}{3}, \tag{6.22}
$$

cf. remark 5.2 in Cuchiero and Teichmann [2015].

6.3.1 properties of the Dirichlet kernels

In this section we study the asymptotic effects of discrete and asynchronous observations through the lens of the Dirichlet kernel and the Fejér kernel.

Define step functions of time, for $j = 1, \cdots, d$,

$$
\begin{aligned}
\underline{\theta}_j^n(t) &= \sup\big\{\tau_h^j,\, \tau_h^j \le t\big\} \\
\overline{\theta}_j^n(t) &= \inf\big\{\tau_h^j,\, t \le \tau_h^j\big\} \wedge \tau_{n_j}^j.
\end{aligned} \tag{6.23}
$$

Define the shifted and scaled Dirichlet kernel

$$d_{jk}^{n,N}(t,u) = \frac{1}{2N+1} D^N\Big(\frac{\theta_j^n(t) - \theta_k^n(u)}{T}\Big), \;\; j,k = 1, \cdots, d. \tag{6.24}$$

The function $d_{jk}^{n,N}(t,u)$ was introduced in Clément and Gloter [2011] and it is indispensable in the upcoming asymptotic analysis.

Note $\big|\big[\theta_j^n(t) - \theta_k^n(t)\big]/T\big| < 1$. By (6.17)

$$d_{jk}^{n,N}(t,t) = \begin{cases} \frac{\sin\big((2N+1)\pi[\theta_j^n(t)-\theta_k^n(t)]/T\big)}{(2N+1)\sin\big(\pi[\theta_j^n(t)-\theta_k^n(t)]/T\big)} & \text{if } \theta_j^n(t) - \theta_k^n(t) \neq 0 \\ 1 & \text{if } \theta_j^n(t) - \theta_k^n(t) = 0, \end{cases}$$

hence

- if the j-th and k-th components are observed synchronously, $d_{jk}^{n,N}(t,t) = 1$;
- if the j-th and k-th components are observed asynchronously but N is chosen in a way such that $N\Delta(n) \to 0$, $d_{jk}^{n,N}(t,t) \xrightarrow{\mathbb{P}} 1$.

The following lemma is a modified adaptation of the Lemma 3 in Clément and Gloter [2011]. It investigates the L^p norm of the shift and scaled Dirichlet kernel.

Lemma 30. *For $p > 1$, $N \leq \lfloor (n_j \wedge n_k)/2 \rfloor$, $\exists K_p < \infty$,*

$$\sup_{j,k} \sup_{t\in[0,T]} \int_0^T \big|d_{jk}^{n,N}(t,u)\big|^p \,\mathrm{d}u \leq K_p N^{-1}.$$

Proof. By the definitions (6.16), (6.24) and the periodicity of the Dirichlet kernel, it suffices to study

$$\sup_j \sup_{a\in[0,T]} \int_{a-T/2}^{a+T/2} \left|\frac{1}{2N+1} D^N\left(\frac{\theta_j^n(t) - a}{T}\right)\right|^p \mathrm{d}t \leq K_p N^{-1}.$$

Note that

$$\Big|\frac{1}{2N+1}D^N(x/T)\Big| \leq 1 \wedge \frac{2T}{(2N+1)|x|},$$

and $\forall a \in [0,T]$, $\forall j = 1\cdots,d$, $|t-a| > |\theta_j^n(t)-a| - |\theta_j^n(t)-t|$, thus

$$\begin{aligned}
&\int_{a-T/2}^{a+T/2} \Big|\frac{1}{2N+1}D^N\Big(\frac{\theta_j^n(t)-a}{T}\Big)\Big|^p \,\mathrm{d}t \\
&\leq \Big(\int_{a-T/2}^{a-\frac{2T}{2N+1}-\Delta(n)} + \int_{a+\frac{2T}{2N+1}+\Delta(n)}^{a+T/2}\Big)\Big|\frac{1}{2N+1}D^N\Big(\frac{\theta_j^n(t)-a}{T}\Big)\Big|^p \,\mathrm{d}t + \frac{4T}{2N+1} + 2\Delta(n) \\
&\qquad \leq \Big(\int_{a-T/2}^{a-\frac{2T}{2N+1}-\Delta(n)} + \int_{a+\frac{2T}{2N+1}+\Delta(n)}^{a+T/2}\Big)\Big|\frac{2T}{(2N+1)(t-a)}\Big|^p \,\mathrm{d}t + \frac{4T}{2N+1} + 2\Delta(n),
\end{aligned}$$

by a change of variable,

$$\Big(\int_{a-T/2}^{a-\frac{2T}{2N+1}-\Delta(n)} + \int_{a+\frac{2T}{2N+1}+\Delta(n)}^{a+T/2}\Big)\Big|\frac{2T}{(2N+1)(t-a)}\Big|^p \,\mathrm{d}t \leq \frac{4T}{2N+1}\int_1^\infty x^{-p}\,\mathrm{d}x,$$

thereby this lemma is proved. $\square$

For $j,k,l,m = 1\cdots,d$, by Fubini's theorem and Hölder's inequality,

$$\begin{aligned}
&N^2 \int_0^T \int_0^T \mathrm{d}t\,\mathrm{d}u \Big[\int_0^{t\wedge u} d_{jk}^{n,N}(t,v)\, d_{jk}^{n,N}(u,v)\,\mathrm{d}v \int_0^v d_{lm}^{n,N}(v,\vartheta)^2\,\mathrm{d}\vartheta\Big] \\
&\qquad \leq N^2 \int_0^T \mathrm{d}v \Big[\int_v^T d_{jk}^{n,N}(t,v)\,\mathrm{d}t \int_v^T d_{jk}^{n,N}(u,v)\,\mathrm{d}u \int_0^v d_{lm}^{n,N}(v,\vartheta)^2\,\mathrm{d}\vartheta\Big] \\
&\qquad\quad \leq N^2 \int_0^T \mathrm{d}v \Big(\int_v^T \big|d_{jk}^{n,N}(t,v)\big|\,\mathrm{d}t\Big)^2 \Big(\int_0^v \big|d_{lm}^{n,N}(v,\vartheta)\big|^2\,\mathrm{d}\vartheta\Big) \\
&\qquad\qquad \leq T^{\frac{3p-2}{p}} N\Big(\int_0^T \big|d_{jk}^{n,N}(t,v)\big|^p\,\mathrm{d}t\Big)^{2/p}\Big(N\int_0^T \big|d_{lm}^{n,N}(v,\vartheta)\big|^2\,\mathrm{d}\vartheta\Big).
\end{aligned}$$

according to **Lemma 30**, we have for $p > 1$,

$$\sup_{j,k,l,m} N^2 \int_0^T \int_0^T \mathrm{d}t \, \mathrm{d}u \Big[\int_0^{t\wedge u} d_{jk}^{n,N}(t,v)\, d_{jk}^{n,N}(u,v)\, \mathrm{d}v \int_0^v d_{lm}^{n,N}(v,\vartheta)^2 \, \mathrm{d}\vartheta \Big] \\ \leq T^{\frac{3p-2}{p}} N^{1-2/p}. \quad (6.25)$$

As Clément and Gloter [2011], we formulate the assumption on the irregular and asynchronous observation times through the shifted and scaled Dirichlet kernel.

Assumption A-d^2. *For $j,k,l,m = 1,\cdots,d$, the quadratic integrals of $d_{jk}^{n,N}$ and $d_{lm}^{n,N}$ converge as $n, N \to \infty$. Specifically, $\exists$ L^1 functions $\tilde{\theta}_{jk,lm}$, $\acute{\theta}_{jk,lm}$, $\check{\theta}_{jk,lm}$, $\grave{\theta}_{jk,lm}$, such that $\forall t \in [0,T]$,*

$$\begin{aligned}
\int_0^t N \int_0^u d_{jk}^{n,N}(u,v)\, d_{lm}^{n,N}(u,v)\, \mathrm{d}v \, \mathrm{d}u &\xrightarrow{\mathbb{P}} \int_0^t \tilde{\theta}_{jk,lm}(u)\, \mathrm{d}u \\
\int_0^t N \int_0^u d_{jk}^{n,N}(u,v)\, d_{lm}^{n,N}(v,u)\, \mathrm{d}v \, \mathrm{d}u &\xrightarrow{\mathbb{P}} \int_0^t \acute{\theta}_{jk,lm}(u)\, \mathrm{d}u \\
\int_0^t N \int_0^u d_{jk}^{n,N}(v,u)\, d_{lm}^{n,N}(u,v)\, \mathrm{d}v \, \mathrm{d}u &\xrightarrow{\mathbb{P}} \int_0^t \check{\theta}_{jk,lm}(u)\, \mathrm{d}u \\
\int_0^t N \int_0^u d_{jk}^{n,N}(v,u)\, d_{lm}^{n,N}(v,u)\, \mathrm{d}v \, \mathrm{d}u &\xrightarrow{\mathbb{P}} \int_0^t \grave{\theta}_{jk,lm}(u)\, \mathrm{d}u.
\end{aligned}$$

The following lemma follows from **Assumption A-d^2** and **Lemma 30**.

Lemma 31. *For any $f_0, f_1 \in C([0,T])$, we have $\forall t \in [0,T]$,*

$$\begin{aligned}
\int_0^t f_0(u)\, \mathrm{d}u \, N \int_0^u d_{jk}^{n,N}(u,v)\, d_{lm}^{n,N}(u,v)\, f_1(v)\, \mathrm{d}v &\xrightarrow{\mathbb{P}} \int_0^t \tilde{\theta}_{jk,lm}(u)\, f_0(u) f_1(u)\, \mathrm{d}u \\
\int_0^t f_0(u)\, \mathrm{d}u \, N \int_0^u d_{jk}^{n,N}(u,v)\, d_{lm}^{n,N}(v,u)\, f_1(v)\, \mathrm{d}v &\xrightarrow{\mathbb{P}} \int_0^t \acute{\theta}_{jk,lm}(u)\, f_0(u) f_1(u)\, \mathrm{d}u \\
\int_0^t f_0(u)\, \mathrm{d}u \, N \int_0^u d_{jk}^{n,N}(v,u)\, d_{lm}^{n,N}(u,v)\, f_1(v)\, \mathrm{d}v &\xrightarrow{\mathbb{P}} \int_0^t \check{\theta}_{jk,lm}(u)\, f_0(u) f_1(u)\, \mathrm{d}u \\
\int_0^t f_0(u)\, \mathrm{d}u \, N \int_0^u d_{jk}^{n,N}(v,u)\, d_{lm}^{n,N}(v,u)\, f_1(v)\, \mathrm{d}v &\xrightarrow{\mathbb{P}} \int_0^t \grave{\theta}_{jk,lm}(u)\, f_0(u) f_1(u)\, \mathrm{d}u.
\end{aligned}$$

Proof. Let's focus on the first convergence. $\forall\epsilon > 0$, $|\theta^n_j(u) - \theta^n_k(v)| \wedge |\theta^n_l(u) - \theta^n_m(v)| > \epsilon$ if $|u - v| > \epsilon$ and n is sufficiently large. By the property of Dirichlet kernel and (6.24), when both n and N are sufficiently large,

$$\int_0^{u-\epsilon} d^{n,N}_{jk}(u,v)\, d^{n,N}_{lm}(u,v)\, \mathrm{d}v \le KN^{-2},$$

then combined with **Assumption A-**d^2, it implies

$$\begin{aligned} N\int_{u-\epsilon}^{u} d^{n,N}_{jk}(u,v)\, d^{n,N}_{lm}(u,v)\, \mathrm{d}v &\xrightarrow{\mathbb{P}} \tilde{\theta}_{jk,lm}(u) \\ N\int_0^{u-\epsilon} d^{n,N}_{jk}(u,v)\, d^{n,N}_{lm}(u,v)\, \mathrm{d}v &\xrightarrow{\mathbb{P}} 0. \end{aligned}$$

Since $f_1 \in C([0,T])$,

$$N\int_0^{u} d^{n,N}_{jk}(u,v)\, d^{n,N}_{lm}(u,v)\, f_1(v)\, \mathrm{d}v = N\int_{u-\epsilon}^{u} d^{n,N}_{jk}(u,v)\, d^{n,N}_{lm}(u,v)\, f_1(v)\, \mathrm{d}v + o_p(1).$$

Because ϵ can be chosen arbitrarily small and f_1 is continuous, the first claim is shown. The other 3 claims in this lemma follow from similar arguments. □

The following lemma reveals the limiting behavior of shifted and scaled Dirichlet kernels when temporal spacings are synchronous (but possibly irregular) across different dimensions.

Lemma 32. *Assume* ***Assumption A-***τ*, then* $\forall t \in [0,T]$, $\forall f_0, f_1 \in C([0,T])$,

$$\begin{aligned} \int_0^T f_0(u)\, \mathrm{d}u\, N\int_0^{u} d^{n,N}_{jk}(u,v)\, d^{n,N}_{lm}(u,v)\, f_1(v)\, \mathrm{d}v &\xrightarrow{\mathbb{P}} \frac{T}{4}\int_0^T f_0(u) f_1(u)\, \mathrm{d}u \\ \int_0^T f_0(u)\, \mathrm{d}u\, N\int_0^{u} d^{n,N}_{jk}(u,v)\, d^{n,N}_{lm}(v,u)\, f_1(v)\, \mathrm{d}v &\xrightarrow{\mathbb{P}} \frac{T}{4}\int_0^T f_0(u) f_1(u)\, \mathrm{d}u \\ \int_0^T f_0(u)\, \mathrm{d}u\, N\int_0^{u} d^{n,N}_{jk}(v,u)\, d^{n,N}_{lm}(u,v)\, f_1(v)\, \mathrm{d}v &\xrightarrow{\mathbb{P}} \frac{T}{4}\int_0^T f_0(u) f_1(u)\, \mathrm{d}u \\ \int_0^T f_0(u)\, \mathrm{d}u\, N\int_0^{u} d^{n,N}_{jk}(v,u)\, d^{n,N}_{lm}(v,u)\, f_1(v)\, \mathrm{d}v &\xrightarrow{\mathbb{P}} \frac{T}{4}\int_0^T f_0(u) f_1(u)\, \mathrm{d}u. \end{aligned}$$

Proof. Because of synchronous observations, $d^{n,N}_{jk}(u,v)\,d^{n,N}_{lm}(u,v) = d^{n,N}_{11}(u,v)^2$, then by (6.24) and (6.20),

$$\int_0^T f_0(u)\,\mathrm{d}u\, N\int_0^u d^{n,N}_{11}(u,v)^2\, f_1(v)\,\mathrm{d}v$$
$$= \frac{N}{2N+1}\int_0^T f_0(u)\,\mathrm{d}u \int_0^u F^{2N+1}\Big(\frac{\theta^n_j(u)-\theta^n_k(v)}{T}\Big) f_1(v)\,\mathrm{d}v,$$

by Riemann summation,

$$\int_0^T f_0(u)\,\mathrm{d}u \int_0^u F^{2N+1}\Big(\frac{\theta^n_j(u)-\theta^n_k(v)}{T}\Big) f_1(v)\,\mathrm{d}v$$
$$= \int_0^T f_0(u)\,\mathrm{d}u \int_0^u F^{2N+1}\Big(\frac{u-v}{T}\Big) f_1(v)\,\mathrm{d}v + O_p(n^{-1}),$$

via changes of variables,

$$\int_0^T f_0(u)\,\mathrm{d}u \int_0^u F^{2N+1}\Big(\frac{u-v}{T}\Big) f_1(v)\,\mathrm{d}v = T^2\int_0^1 f_0(Tu)\,\mathrm{d}u\int_0^u F^{2N+1}(u-v) f_1(Tv)\,\mathrm{d}v,$$

note that F^{2N+1} is a delta sequence, as $N\to\infty$, $\int_0^u F^{2N+1}(u-v)f_1(Tv)\,\mathrm{d}v \to f_1(Tu)/2$, then this lemma follows from a change of variable. □

6.3.2 classical results on the Fejér kernel

Given a function φ on $[0,T]$, define its *truncated Fourier inversion* as

$$\overline{\varphi}^q(x) = T^{-1}\sum_{|s|\le q} F(\varphi)_s\, e^{i2\pi s x/T}.$$

One can express $\overline{\varphi}^q$ as the convolution between φ and a Dirichlet kernel,

$$\begin{aligned}\overline{\varphi}^q(x) &= \frac{1}{T}\sum_{|s|\le q} e^{i2\pi sx/T}\int_0^T \varphi(u)e^{-i2\pi su/T}\,\mathrm{d}u \\ &= \frac{1}{T}\int_0^T \varphi(u)\sum_{|s|\le q} e^{i2\pi s(x-u)/T}\,\mathrm{d}u = \frac{1}{T}\int_0^T \varphi(u)D^q[(x-u)/T]\,\mathrm{d}u. \end{aligned} \tag{6.26}$$

Given an interval I with $|I|=1$, $\forall M\in\mathbb{N}^+$,

$$\int_I F^M(x)\,\mathrm{d}x = \frac{1}{M}\sum_{q=0}^{M-1}\int_I\sum_{|s|\le q} e^{i2\pi sx}\,\mathrm{d}x = \frac{1}{M}\sum_{q=0}^{M-1}\int_I \mathrm{d}x,$$

as a consequence, for example,

$$\int_{-1/2}^{1/2} F^M(x)\,\mathrm{d}x = 1, \qquad \int_0^1 F^M(x)\,\mathrm{d}x = 1. \tag{6.27}$$

Based on (6.20), we know $F^M(x)\le \sin(\pi\delta)^{-2}M^{-1}$ for $\delta\le|x|\le 1/2$. Hence $\forall\delta\in(0,1/2)$,

$$\int_{\delta\le|x|\le 1/2} F^M(x)\,\mathrm{d}x \le K(\delta^2 M)^{-1}. \tag{6.28}$$

Moreover, we have

$$F^M(x)\le\frac{M}{1+M^2x^2}, \qquad x\in[-1/2,1/2]. \tag{6.29}$$

Given a generic function φ on $[0,T]$, define

$$\widehat{\varphi}^M(t) = \frac{1}{T}\sum_{|q|<M}\Big(1-\frac{|q|}{M}\Big)F(\varphi)_q\, e^{i2\pi qt/T}, \tag{6.30}$$

note that

$$\widehat{\varphi}^M(t) = \frac{1}{TM}\sum_{q=0}^{M-1}\sum_{|s|\le q} F(\varphi)_s\, e^{i2\pi st/T} = \frac{1}{M}\sum_{q=0}^{M-1}\overline{\varphi}^q(t),$$

and by (6.26), (6.18),

$$\frac{1}{M}\sum_{q=0}^{M-1}\overline{\varphi}^q(t) = \frac{1}{T}\int_0^T \varphi(u)\frac{1}{M}\sum_{q=0}^{M-1} D^q[(t-u)/T]\,\mathrm{d}u = \frac{1}{T}\int_0^T \varphi(u)F^M[(t-u)/T]\,\mathrm{d}u,$$

so we have the representation

$$\widehat{\varphi}^M(t) = \frac{1}{T}\int_0^T \varphi(u)F^M[(t-u)/T]\,\mathrm{d}u. \tag{6.31}$$

We have the following lemma which implies the uniform convergence of $\widehat{\varphi}^M$:

$$\sup_{t\in[0,T]}\left|\widehat{\varphi}^M(t)-\varphi(t)\right| \le K\omega_\varphi(1/M), \qquad \text{if } \varphi(0)=\varphi(T)$$
$$\sup_{t\in[1/M,T-1/M]}\left|\widehat{\varphi}^M(t)-\varphi(t)\right| \le K\omega_\varphi(1/M), \qquad \text{if } \varphi(0)\neq\varphi(T).$$

Lemma 33. *If the function φ is continuous, then*

$$\sup_{t\in[0,T]}\left|\frac{1}{T}\int_0^T F^M\Big(\frac{t-u}{T}\Big)\varphi(u)\,\mathrm{d}u-\varphi(t)\right| \le K\omega_\varphi(1/M), \qquad \textit{if } \varphi(0)=\varphi(T)$$
$$\sup_{t\in[1/M,T-1/M]}\left|\frac{1}{T}\int_0^T F^M\Big(\frac{t-u}{T}\Big)\varphi(u)\,\mathrm{d}u-\varphi(t)\right| \le K\omega_\varphi(1/M), \qquad \textit{if } \varphi(0)\neq\varphi(T).$$

Proof. By (6.27),

$$\delta^M(t) \coloneqq T\left|\frac{1}{T}\int_0^T F^M\Big(\frac{t-u}{T}\Big)\varphi(u)\,\mathrm{d}u-\varphi(t)\right| = \left|\int_0^T F^M\Big(\frac{t-u}{T}\Big)\big[\varphi(t)-\varphi(u)\big]\,\mathrm{d}u\right|.$$

(1) If $\varphi(0)=\varphi(T)$, by periodization, we can extend the definition of φ to the real line and retain its modulus of continuity. By a change of variable and the periodicity of F^M and φ,

$$\begin{aligned}\delta^M(t) = \left|\int_{t-T}^{t} F^M(u/T)\big[\varphi(t)-\varphi(t-u)\big]\,\mathrm{d}u\right| &\le \int_{-T/2}^{T/2} F^M(u/T)\cdot|\varphi(t)-\varphi(t-u)|\,\mathrm{d}u \\ &= \Big(\int_{-1/M}^{1/M} + \int_{1/M\le|u|\le T/2}\Big) F^M(u/T)\cdot|\varphi(t)-\varphi(t-u)|\,\mathrm{d}u.\end{aligned}$$

since

$$F^M(x) \le \frac{M}{1+M^2x^2}, \qquad x\in[-1/2,1/2],$$

we have

$$\int_{-1/M}^{1/M} F^M(u/T)\cdot|\varphi(t)-\varphi(t-u)|\,\mathrm{d}u \le K\omega_\varphi(M^{-1})\int_{-1/M}^{1/M} F^M(u/T)\,\mathrm{d}u \le K\omega_\varphi(M^{-1}),$$

and

$$\begin{aligned}&\int_{1/M\le|u|\le T/2} F^M(u/T)\cdot|\varphi(t)-\varphi(t-u)|\,\mathrm{d}u \\ &\qquad\le \frac{K}{M}\int_{1/M\le|u|\le T/2} |u|^{\alpha-2}\,\mathrm{d}u \le K\big[\omega_\varphi(M^{-1})+M^{-1}\big],\end{aligned}$$

then this lemma in the case $\varphi(0)=\varphi(T)$ is proved.

(2) If $\varphi(0)\ne\varphi(T)$, then for $t\in[1/M, T-1/M]$,

$$\delta^M(t) \le \Big(\int_0^{t-1/M} + \int_{t-1/M}^{t+1/M} + \int_{t+1/M}^{T}\Big) F^M\Big(\frac{t-u}{T}\Big)\cdot|\varphi(t)-\varphi(u)|\,\mathrm{d}u,$$

then by a similar argument, the lemma in the case $\varphi(0)\ne\varphi(T)$ can also be proved. □

6.3.3 more results on the Fejér kernels

Lemma 34. *If $M/n \to 0$, for $\theta_j^n(t)$ defined in (6.23), $\forall f \in C([0,T])$, $\forall t \in [0,T]$,*

$$\left| \int_0^T F^M\Big(\frac{t-\theta_j^n(u)}{T}\Big)\, f(u)\,\mathrm{d}u - \int_0^T F^M\Big(\frac{t-u}{T}\Big)\, f(u)\,\mathrm{d}u \right| \le KT\frac{M}{n}.$$

Proof. Denote the L.H.S. by $D(t)_{j,T}^{n,M}$, note

$$\begin{aligned} D(t)_{j,T}^{n,M} &\le K \int_0^T \Big| F^M\Big(\frac{t-\theta_j^n(u)}{T}\Big) - F^M\Big(\frac{t-u}{T}\Big)\Big| \,\mathrm{d}u \\ = K \sum_{h=1}^{n_j} \int_{I_h^j} \Big| F^M\Big(\frac{t-\tau_h^j}{T}\Big) - F^M\Big(\frac{t-u}{T}\Big)\Big| \,\mathrm{d}u &= K \sum_{h=1}^{n_j} \Delta_h^j \Big| F^M\Big(\frac{t-\tau_h^j}{T}\Big) - F^M\Big(\frac{t-u_h^j}{T}\Big)\Big|, \end{aligned}$$

where $u_h^j \in I_h^j$ for each h by mean value theorem.

Let $J_b = \Big(\big(t+\frac{bT}{M}\big)\vee 0, \big(t+\frac{(b+1)T}{M}\big)\wedge T\Big]$, $B_0 = \inf\big\{b, t+\frac{(b+1)T}{M} > 0\big\}$, $B_1 = \sup\big\{b, t+\frac{bT}{M} < T\big\}$, then

$$D(t)_{j,T}^{n,M} \le K \sum_{b=B_0}^{B_1} \sum_{\tau_h^j \in J_b} \Delta_h^j \Big| F^M\Big(\frac{t-\tau_h^j}{T}\Big) - F^M\Big(\frac{t-u_h^j}{T}\Big)\Big|.$$

Based on mean value theorem, $\exists v_h^j \in [u_h^j, \tau_h^j]$ for each h such that $F^M\Big(\frac{t-\tau_h^j}{T}\Big) - F^M\Big(\frac{t-u_h^j}{T}\Big) = (\tau_h^j - v_h^j)\,\partial F^M\Big(\frac{t-v_h^j}{T}\Big)$, so

$$\begin{aligned} \sum_{\tau_h^j \in J_b} \Delta_h^j \Big| F^M\Big(\frac{t-\tau_h^j}{T}\Big) - F^M\Big(\frac{t-u_h^j}{T}\Big)\Big| &\le \Delta(n)^2 \sum_{\tau_h^j \in J_b} \Big| \partial F^M\Big(\frac{t-v_h^j}{T}\Big)\Big| \\ &\le K\frac{n\Delta(n)^2}{M} \sup_{v \in J_b} \Big| \partial F^M\Big(\frac{t-v}{T}\Big)\Big|. \end{aligned}$$

Based on (6.20),

$$\sup_{v\in J_b}\Big|\partial F^M\Big(\frac{t-v}{T}\Big)\Big| \le KM \sup_{v\in J_b} F^M\Big(\frac{t-v}{T}\Big) \le KM^2\int_{J_b} F^M\Big(\frac{t-u}{T}\Big)\,\mathrm{d}u,$$

thus we have

$$D(t)^{n,M}_{j,T} \le KMn\Delta(n)^2\sum_{b=B_0}^{B_1}\int_{J_b} F^M\Big(\frac{t-u}{T}\Big)\,\mathrm{d}u \asymp KT\frac{M}{n},$$

from which this lemma follows. □

6.4 Univariate asymptotic analysis

First of all, let's consider the estimation of univariate volatility. In this section, we let $X = X_1$, $c = c_{11}$, $\tau_h = \tau^1_h$, $\delta_h = \delta^1_h$, $\underline{\theta}^n(t) = \underline{\theta}^n_1(t)$, $\theta^n(t) = \theta^n_1(t)$. We are interested in the asymptotics of

$$N^{1/2}\int_0^T \rho(t)\big[\overline{\widehat{c}^{n,N,M}(t)} - c(t)\big]\,\mathrm{d}t,$$

for a α-Hölder continuous function ρ where α is specified in **Assumption A-**α.

Because $T-\tau_n \le \Delta(n)$, w.l.o.g., we can assume $\tau_n = T$. By the definition (6.7),

$$\begin{aligned}
&\int_0^T \rho(t)\big[\overline{\widehat{c}^{n,N,M}(t)} - c(t)\big]\,\mathrm{d}t \\
&\qquad = \frac{1}{T}\sum_{|q|<M}\Big(1-\frac{|q|}{M}\Big)\overline{\widehat{F}(c)^{n,N}_q}\int_0^T \rho(t)e^{-i2\pi qt/T}\,\mathrm{d}t - \int_0^T \rho(t)\,c(t)\,\mathrm{d}t \\
&\qquad = \frac{1}{T}\sum_{|q|<M}\Big(1-\frac{|q|}{M}\Big)F(\rho)_q\,\overline{\widehat{F}(c)^{n,N}_q} - \int_0^T \rho(t)\,c(t)\,\mathrm{d}t.
\end{aligned}$$

6.4.1 decomposition

Let

$$A^{n,N,M} = \frac{1}{T} \sum_{|q|<M} \Big(1 - \frac{|q|}{M}\Big) F(\rho)_q \, \overline{\widehat{F}(c)_q^{n,N}}.$$

Note by (6.8),

$$\overline{\widehat{F}(c)_q^{n,N}} = \sum_{h=1}^{n} e^{i2\pi q \tau_h / T} \delta_h(X)^2 + \alpha_{q,0}^{n,N} + \alpha_{q,1}^{n,N},$$

where

$$\begin{aligned} \alpha_{q,0}^{n,N} &= \sum_{h=2}^{n} e^{i2\pi q \tau_h / T} \delta_h(X) \sum_{v=1}^{h-1} D^N \Big(\frac{\tau_h - \tau_v}{T}\Big) \frac{\delta_v(X)}{2N+1} \\ \alpha_{q,1}^{n,N} &= \sum_{v=2}^{n} \delta_v(X) \sum_{h=1}^{v-1} e^{i2\pi q \tau_h / T} D^N \Big(\frac{\tau_v - \tau_h}{T}\Big) \frac{\delta_h(X)}{2N+1}, \end{aligned} \tag{6.32}$$

hence

$$\begin{aligned} A^{n,N,M} &= \frac{1}{T} \sum_{|q|<M} \Big(1 - \frac{|q|}{M}\Big) F(\rho)_q \times \Big[\sum_{h=1}^{n} e^{i2\pi q \tau_h / T} \delta_h(X)^2 + \alpha_{q,0}^{n,N} + \alpha_{q,1}^{n,N} \Big] \\ &= \sum_{h=1}^{n} \widehat{\rho}^M(\tau_h) \, \delta_h(X)^2 + \frac{1}{T} \sum_{|q|<M} \Big(1 - \frac{|q|}{M}\Big) F(\rho)_q \left(\alpha_{q,0}^{n,N} + \alpha_{q,1}^{n,N}\right). \end{aligned}$$

Therefore, we have the following decomposition:

$$N^{1/2} \int_0^T \rho(t) \big[\overline{\widehat{c}^{n,N,M}(t)} - c(t)\big] \, \mathrm{d}t = o_T^{n,M} + e_T^{n,N,M}, \tag{6.33}$$

where

$$
\begin{aligned}
o_T^{n,M} &= N^{1/2} \sum_{h=1}^{n} \int_{I_h} \big[\widehat{\rho}^M(\tau_h) - \rho(t)\big] c(t)\,\mathrm{d}t \\
e_T^{n,N,M} &= \frac{N^{1/2}}{T} \sum_{|q|<M} \Big(1 - \frac{|q|}{M}\Big) F(\rho)_q \left(\alpha_{q,0}^{n,N} + \alpha_{q,1}^{n,N}\right) \\
&\qquad\qquad + N^{1/2} \sum_{h=1}^{n} \widehat{\rho}^M(\tau_h) \Big[\delta_h(X)^2 - \int_{I_h} c(t)\,\mathrm{d}t\Big],
\end{aligned}
$$

and the function $\widehat{\rho}^M$ is defined by (6.2).

Based on (6.27), (6.31), (6.23), Fubini's theorem, and **Lemma 34**,

$$
\begin{aligned}
o_T^{n,M} = N^{1/2} \int_0^T c(t)\,\mathrm{d}t\, \frac{1}{T} \int_0^T F^M\Big(\frac{u-\theta^n(t)}{T}\Big) \big[\rho(u) - \rho(t)\big]\,\mathrm{d}u \\
= \frac{N^{1/2}}{T} J_T^M + O_p\big(N^{1/2} M/n\big),
\end{aligned}
$$

where

$$
J_T^M = \int_0^T \int_0^T F^M\Big(\frac{u-t}{T}\Big) \big[\rho(u) - \rho(t)\big] c(t)\,\mathrm{d}t\,\mathrm{d}u.
$$

By symmetry of variables, $J_T^M = \int_0^T \int_0^T F^M[(u-t)/T] \big[\rho(t) - \rho(u)\big] c(u)\,\mathrm{d}u\,\mathrm{d}t$, hence

$$
J_T^M = -\frac{1}{2} \int_0^T \mathrm{d}u \int_0^T F^M\Big(\frac{u-t}{T}\Big) \big[\rho(u) - \rho(t)\big] \big[c(u) - c(t)\big]\,\mathrm{d}t.
$$

The modulus of continuity of ρ is determined by that of c, let

$$
L_T^M(u) := \int_0^T F^M[(u-t)/T] \big[\rho(u) - \rho(t)\big] \big[c(u) - c(t)\big]\,\mathrm{d}t,
$$

by periodicity of c and ρ, $L_T^M(u) = \int_{u-T/2}^{u+T/2} F^M[(u-t)/T] \big[\rho(u) - \rho(t)\big] \big[c(u) - c(t)\big]\,\mathrm{d}t$.

Note

$$\left|L_T^M(u)\right| \le \Big(\int_{|u-t|\le 1/M} + \int_{|u-t|>1/M}\Big) F^M\Big(\frac{u-t}{T}\Big)\left|\rho_(u)-\rho(t)\right|\left|c(u)-c(t)\right| \mathrm{d}t,$$

through an argument similar to the proof of **Lemma 33**, we have $\mathbb{E}\left|L_T^M(u)\right| \le K\big[M^{-2\alpha} + M^{-(1+\alpha)}\big]$, thus

$$\mathbb{E}\left|o_T^{n,M}\right| \le K\Big[\Big(\frac{N}{M^{4\alpha}}\Big)^{1/2} + \Big(\frac{N}{n}\Big)^{1/2}\frac{M}{n^{1/2}}\Big],$$

by Markov's inequality, we have shown the asymptotic negligibility in probability of $o_T^{n,M}$, i.e.

$$o_T^{n,M} \xrightarrow{\mathbb{P}} 0, \tag{6.34}$$

thus the asymptotics solely relies on $e_T^{n,N,M}$.

Plug in the definition (6.32), recall the definition (6.2), use (6.15) and Itô's formula, we have

$$e_T^{n,N,M} = \breve{e}(0)_T^{n,N,M} + \breve{e}(1)_T^{n,N,M} + \breve{e}(2)_T^{n,N,M},$$

where

$$\begin{aligned}
\breve{e}(0)_T^{n,N,M} &= N^{1/2}\sum_{h=2}^{n}\widehat{\rho}^M(\tau_h)\delta_h(X)\sum_{v=1}^{h-1} D^N\Big(\frac{\tau_h-\tau_v}{T}\Big)\frac{\delta_v(X)}{2N+1} \\
\breve{e}(1)_T^{n,N,M} &= N^{1/2}\sum_{v=2}^{n}\delta_v(X)\sum_{h=1}^{v-1}\widehat{\rho}^M(\tau_h)D^N\Big(\frac{\tau_v-\tau_h}{T}\Big)\frac{\delta_h(X)}{2N+1} \\
\breve{e}(2)_T^{n,N,M} &= 2N^{1/2}\sum_{h=1}^{n}\widehat{\rho}^M(\tau_h)\int_{\tau_{h-1}}^{\tau_h}\mathrm{d}X(t)\int_{\tau_{h-1}}^{t}\mathrm{d}X(u).
\end{aligned}$$

By (6.23) and (6.24), we have the following representation

$$
\begin{aligned}
\breve{e}(0)_T^{n,N,M} &= N^{1/2}\int_{\tau_1}^{\tau_n} \widehat{\rho}^M(\theta^n(t))\,\mathrm{d}X(t)\int_0^{\underline{\theta}^n(t)} d^{n,N}(t,u)\,\mathrm{d}X(u)\\
\breve{e}(1)_T^{n,N,M} &= N^{1/2}\int_{\tau_1}^{\tau_n} \mathrm{d}X(t)\int_0^{\underline{\theta}^n(t)} \widehat{\rho}^M(\theta^n(u))\,d^{n,N}(t,u)\,\mathrm{d}X(u)\\
\breve{e}(2)_T^{n,N,M} &= 2N^{1/2}\int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))\,\mathrm{d}X(t)\int_{\underline{\theta}^n(t)}^{t} \mathrm{d}X(u).
\end{aligned}
$$

Note the for $u \in [\underline{\theta}^n(t), t]$, $\theta^n(t) = \theta^n(u)$, thus

$$
e_T^{n,N,M} = e(0)_T^{n,N,M} + e(1)_T^{n,N,M}, \tag{6.35}
$$

where

$$
\begin{aligned}
e(0)_T^{n,N,M} &= N^{1/2}\int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))\,\mathrm{d}X(t)\int_0^{t} d^{n,N}(t,u)\,\mathrm{d}X(u)\\
e(1)_T^{n,N,M} &= N^{1/2}\int_0^{\tau_n} \mathrm{d}X(t)\int_0^{t} \widehat{\rho}^M(\theta^n(u))\,d^{n,N}(t,u)\,\mathrm{d}X(u).
\end{aligned}
$$

6.4.2 preparing some martingales

Now let's define some martingales useful in the incoming asymptotic analysis. They are

$$
\begin{aligned}
U^{n,N}(t) &= \int_0^t d^{n,N}(t,u)\,\sigma(u)\,\mathrm{d}W(u)\\
\widetilde{U}^{n,N}(t) &= \int_0^t d^{n,N}(t,u)\,\widehat{\rho}^M(\theta^n(u))\,\sigma(u)\,\mathrm{d}W(u)\\
\widehat{U}^{n,N}(t) &= \int_0^u F^M\Big(\frac{t-\theta^n(v)}{T}\Big)\,d^{n,N}(v,u)\,\sigma(v)\,\mathrm{d}W(v),
\end{aligned}
\tag{6.36}
$$

and Itô martingales

$$\begin{aligned}
Z^{n,N}(t) &= \int_0^t d^{n,N}(t,u)\,\sigma(u)\,U^{n,N}(u)\,\mathrm{d}W(u) \\
\check{Z}^{n,N}(t) &= \int_0^t d^{n,N}(t,u)\,\sigma(u)\,\widetilde{U}^{n,N}(u)\,\mathrm{d}W(u) \\
\mathring{Z}^{n,N}(t) &= \int_0^t d^{n,N}(t,u)\,\widehat{\rho}^M(\theta^n(u))\,\sigma(u)\,U^{n,N}(u)\,\mathrm{d}W(u) \\
\widetilde{Z}^{n,N}(t) &= \int_0^t d^{n,N}(t,u)\,\widehat{\rho}^M(\theta^n(u))\,\sigma(u)\,\widetilde{U}^{n,N}(u)\,\mathrm{d}W(u).
\end{aligned} \tag{6.37}$$

By Itô's formula,

$$\begin{aligned}
U^{n,N}(t)^2 &= \int_0^t d^{n,N}(t,u)^2\,c(u)\,\mathrm{d}u + 2Z^{n,N}(t) \\
U^{n,N}(t)\,\widetilde{U}^{n,N}(t) &= \int_0^t d^{n,N}(t,u)^2\,\widehat{\rho}^M(\theta^n(u))\,c(u)\,\mathrm{d}u + \check{Z}^{n,N}(t) + \mathring{Z}^{n,N}(t) \\
\widetilde{U}^{n,N}(t)^2 &= \int_0^t d^{n,N}(t,u)^2\,\widehat{\rho}^M(\theta^n(u))^2\,c(u)\,\mathrm{d}u + 2\widetilde{Z}^{n,N}(t).
\end{aligned} \tag{6.38}$$

Lemma 35. *There exist some finite positive constant K such that*

$$\begin{aligned}
&\mathbb{E}\big[U^{n,N}(t)U^{n,N}(u)\big] \bigvee \mathbb{E}\big[U^{n,N}(t)\widetilde{U}^{n,N}(u)\big] \bigvee \mathbb{E}\big[\widetilde{U}^{n,N}(t)\widetilde{U}^{n,N}(u)\big] \\
&\qquad\qquad \le K\int_0^{t\wedge u} d^{n,N}(t,v)\,d^{n,N}(u,v)\,\mathrm{d}v,
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\big[Z^{n,N}(t)Z^{n,N}(u)\big] \bigvee \mathbb{E}\big[\check{Z}^{n,N}(t)\check{Z}^{n,N}(u)\big] \bigvee \mathbb{E}\big[\mathring{Z}^{n,N}(t)\mathring{Z}^{n,N}(u)\big] \bigvee \mathbb{E}\big[\widetilde{Z}^{n,N}(t)\widetilde{Z}^{n,N}(u)\big] \\
&\qquad\qquad \le K\int_0^{t\wedge u} d^{n,N}(t,v)\,d^{n,N}(u,v)\,\mathrm{d}v\,\Big(\int_0^v d^{n,N}(v,s)^2\,\mathrm{d}s\Big).
\end{aligned}$$

Proof. By Itô's formula,

$$\begin{aligned}
\mathbb{E}\big[U^{n,N}(t)U^{n,N}(u)\big] &= \int_0^{t\wedge u} d^{n,N}(t,v)\, d^{n,N}(u,v)\, \mathbb{E}[c(v)]\, \mathrm{d}v \\
\mathbb{E}\big[U^{n,N}(t)\widetilde{U}^{n,N}(u)\big] &= \int_0^{t\wedge u} d^{n,N}(t,v)\, d^{n,N}(u,v)\, \mathbb{E}\big[\widehat{\rho}^M(\theta^n(v))\, c(v)\big]\, \mathrm{d}v \\
\mathbb{E}\big[\widetilde{U}^{n,N}(t)\widetilde{U}^{n,N}(u)\big] &= \int_0^{t\wedge u} d^{n,N}(t,v)\, d^{n,N}(u,v)\, \mathbb{E}\big[\widehat{\rho}^M(\theta^n(v))^2\, c(v)\big]\, \mathrm{d}v,
\end{aligned}$$

thereby the first claim is shown.

According to Itô's formula, (6.37), (6.38),

$$\begin{aligned}
\mathbb{E}\big[Z^{n,N}(t)Z^{n,N}(u)\big] = \mathbb{E}\int_0^{t\wedge u} & d^{n,N}(t,v)\, d^{n,N}(u,v)\, U^{n,N}(v)^2\, c(v)\, \mathrm{d}v \\
&\le K\int_0^{t\wedge u} d^{n,N}(t,v)\, d^{n,N}(u,v)\, \mathbb{E}\big[U^{n,N}(v)^2\big]\, \mathrm{d}v,
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\big[\mathring{Z}^{n,N}(t)\mathring{Z}^{n,N}(u)\big] = \mathbb{E}\int_0^{t\wedge u} & d^{n,N}(t,v)\, d^{n,N}(u,v)\, U^{n,N}(v)^2\, \widehat{\rho}^M(\theta^n(v))\, c(v)\, \mathrm{d}v \\
&\le K\int_0^{t\wedge u} d^{n,N}(t,v)\, d^{n,N}(u,v)\, \mathbb{E}\big[U^{n,N}(v)^2\big]\, \mathrm{d}v,
\end{aligned}$$

similarly,

$$\begin{aligned}
\mathbb{E}\big[\check{Z}^{n,N}(t)\check{Z}^{n,N}(u)\big] &\le K\int_0^{t\wedge u} d^{n,N}(t,v)\, d^{n,N}(u,v)\, \mathbb{E}\big[\widetilde{U}^{n,N}(v)^2\big]\, \mathrm{d}v \\
\mathbb{E}\big[\widetilde{Z}^{n,N}(t)\widetilde{Z}^{n,N}(u)\big] &\le K\int_0^{t\wedge u} d^{n,N}(t,v)\, d^{n,N}(u,v)\, \mathbb{E}\big[\widetilde{U}^{n,N}(v)^2\big]\, \mathrm{d}v,
\end{aligned}$$

then the second claim follows from the first claim proved earlier. $\square$

6.4.3 stable convergence

By (6.36), we can write

$$
\begin{aligned}
e(0)_T^{n,N,M} &= N^{1/2} \int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))\, \sigma(t)\, U^{n,N}(t)\, \mathrm{d}W(t) \\
e(1)_T^{n,N,M} &= N^{1/2} \int_0^{\tau_n} \sigma(t)\, \widetilde{U}^{n,N}(t)\, \mathrm{d}W(t).
\end{aligned} \tag{6.39}
$$

First, let's consider, for $r = 1 \cdots, d'$,

$$
\begin{aligned}
\langle e(0)^{n,N,M}, W_r \rangle_T &= N^{1/2} \int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))\, \sigma_{1,r}(t)\, U^{n,N}(t)\, \mathrm{d}t \\
\langle e(1)^{n,N,M}, W_r \rangle_T &= N^{1/2} \int_0^{\tau_n} \sigma_{1,r}(t)\, \widetilde{U}^{n,N}(t)\, \mathrm{d}t.
\end{aligned}
$$

Notice that

$$
\begin{aligned}
\langle e(0)^{n,N,M}, W_r \rangle_T^2 = N \int_0^{\tau_n} \int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t)) \widehat{\rho}^M(\theta^n(u))\, \sigma_{1,r}(t) \sigma_{1,r}(u) \\
\times U^{n,N}(t) U^{n,N}(u)\, \mathrm{d}t\, \mathrm{d}u,
\end{aligned}
$$

according to **Lemma 35**, and by Fubini's theorem and Hölder's inequality,

$$
\begin{aligned}
\mathbb{E}\big[\langle e(0)^{n,N,M}, W_r \rangle_T^2\big] &\le KN \int_0^{\tau_n} \int_0^{\tau_n} \mathrm{d}t\, \mathrm{d}u \Big(\int_0^{t \wedge u} d^{n,N}(t,v)\, d^{n,N}(u,v)\, \mathrm{d}v \Big) \\
&\le KN \int_0^{\tau_n} \mathrm{d}v \Big(\int_v^{\tau_n} \big|d^{n,N}(t,v)\big|\, \mathrm{d}t \Big) \Big(\int_v^{\tau_n} \big|d^{n,N}(u,v)\big|\, \mathrm{d}u \Big) \\
&\le KT^{\frac{3p-2}{p}} N \Big(\int_0^T \big|d^{n,N}(t,v)\big|^p\, \mathrm{d}t \Big)^{2/p};
\end{aligned}
$$

similarly,

$$
\langle e(1)^{n,N,M}, W_r \rangle_T^2 = N \int_0^{\tau_n} \int_0^{\tau_n} \sigma_{1,r}(t) \sigma_{1,r}(u) \times \widetilde{U}^{n,N}(t) \widetilde{U}^{n,N}(u)\, \mathrm{d}t\, \mathrm{d}u,
$$

through a similar argument applied to $\mathbb{E}[\langle e(0)^{n,N,M}, W_r\rangle_T^2]$,

$$\mathbb{E}\big[\big\langle e(1)^{n,N,M}, W_r\big\rangle_T^2\big] \leq KT^{\frac{3p-2}{p}} N\Big(\int_0^T \big|d^{n,N}(t,v)\big|^p \,\mathrm{d}t\Big)^{2/p}.$$

By **Lemma 30**, Jensen's inequality and Markov's inequality,

$$\big\langle e^{n,N,M}, W_r\big\rangle_T = \big\langle e(0)^{n,N,M}, W_r\big\rangle_T + \big\langle e(1)^{n,N,M}, W_r\big\rangle_T \xrightarrow{\mathbb{P}} 0. \tag{6.40}$$

Second, let's consider

$$\begin{aligned}\big\langle e^{n,N,M}, e^{n,N,M}\big\rangle_T &= \big\langle e(0)^{n,N,M}, e(0)^{n,N,M}\big\rangle_T \\ &\quad + 2\big\langle e(0)^{n,N,M}, e(1)^{n,N,M}\big\rangle_T + \big\langle e(1)^{n,N,M}, e(1)^{n,N,M}\big\rangle_T,\end{aligned} \tag{6.41}$$

note, by (6.39),

$$\begin{aligned}\big\langle e(0)^{n,N,M}, e(0)^{n,N,M}\big\rangle_T &= N\int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))^2\, c(t)\, U^{n,N}(t)^2 \,\mathrm{d}t \\ \big\langle e(0)^{n,N,M}, e(1)^{n,N,M}\big\rangle_T &= N\int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))\, c(t)\, U^{n,N}(t)\, \widetilde{U}^{n,N}(t) \,\mathrm{d}t \\ \big\langle e(1)^{n,N,M}, e(1)^{n,N,M}\big\rangle_T &= N\int_0^{\tau_n} c(t)\, \widetilde{U}^{n,N}(t)^2 \,\mathrm{d}t,\end{aligned}$$

in view of (6.38),

$$\begin{aligned}\big\langle e(0)^{n,N,M}, e(0)^{n,N,M}\big\rangle_T &= 2O(0)_T^{n,N,M} + V(0)_T^{n,N,M} \\ \big\langle e(0)^{n,N,M}, e(1)^{n,N,M}\big\rangle_T &= O(1)_T^{n,N,M} + O(2)_T^{n,N,M} + V(1)_T^{n,N,M} \\ \big\langle e(1)^{n,N,M}, e(1)^{n,N,M}\big\rangle_T &= 2O(3)_T^{n,N,M} + V(2)_T^{n,N,M},\end{aligned}$$

where

$$
\begin{aligned}
O(0)^{n,N,M}_T &= N \int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))^2\, c(t)\, Z^{n,N}(t)\,\mathrm{d}t \\
O(1)^{n,N,M}_T &= N \int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))\, c(t)\, \check{Z}^{n,N}(t)\,\mathrm{d}t \\
O(2)^{n,N,M}_T &= N \int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))\, c(t)\, \mathring{Z}^{n,N}(t)\,\mathrm{d}t \\
O(3)^{n,N,M}_T &= N \int_0^{\tau_n} c(t)\, \widetilde{Z}^{n,N}(t)\,\mathrm{d}t,
\end{aligned}
$$

and

$$
\begin{aligned}
V(0)^{n,N,M}_T &= \int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))^2\, c(t) \Big[N \int_0^t d^{n,N}(t,u)^2\, c(u)\,\mathrm{d}u \Big]\,\mathrm{d}t \\
V(1)^{n,N,M}_T &= \int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))\, c(t) \Big[N \int_0^t d^{n,N}(t,u)^2\, \widehat{\rho}^M(\theta^n(u))\, c(u)\,\mathrm{d}u \Big]\,\mathrm{d}t \\
V(2)^{n,N,M}_T &= \int_0^{\tau_n} c(t) \Big[N \int_0^t d^{n,N}(t,u)^2\, \widehat{\rho}^M(\theta^n(u))^2\, c(u)\,\mathrm{d}u \Big]\,\mathrm{d}t.
\end{aligned}
$$

Let consider the asymptotically negligible terms,

$$
\begin{aligned}
\big|O(0)^{n,N,M}_T\big|^2 &= N^2 \int_0^{\tau_n}\int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t))^2 \widehat{\rho}^M(\theta^n(u))^2\, c(t)c(u)\, Z^{n,N}(t) Z^{n,N}(u)\,\mathrm{d}t\,\mathrm{d}u \\
\big|O(1)^{n,N,M}_T\big|^2 &= N^2 \int_0^{\tau_n}\int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t)) \widehat{\rho}^M(\theta^n(u))\, c(t)c(u)\, \check{Z}^{n,N}(t) \check{Z}^{n,N}(u)\,\mathrm{d}t\,\mathrm{d}u \\
\big|O(2)^{n,N,M}_T\big|^2 &= N^2 \int_0^{\tau_n}\int_0^{\tau_n} \widehat{\rho}^M(\theta^n(t)) \widehat{\rho}^M(\theta^n(u))\, c(t)c(u)\, \mathring{Z}^{n,N}(t) \mathring{Z}^{n,N}(u)\,\mathrm{d}t\,\mathrm{d}u \\
\big|O(3)^{n,N,M}_T\big|^2 &= N^2 \int_0^{\tau_n}\int_0^{\tau_n} c(t)c(u)\, \widetilde{Z}^{n,N}(t) \widetilde{Z}^{n,N}(u)\,\mathrm{d}t\,\mathrm{d}u,
\end{aligned}
$$

by **Lemma 35**,

$$
\begin{aligned}
&\mathbb{E}\big(\big|O(0)^{n,N,M}_T\big|^2\big) \\
&\qquad \le K N^2 \int_0^{\tau_n}\int_0^{\tau_n} \mathrm{d}t\,\mathrm{d}u \Big[\int_0^{t\wedge u} d^{n,N}(t,v)\, d^{n,N}(u,v) \Big(\int_0^v d^{n,N}(v,s)^2\,\mathrm{d}s \Big)\,\mathrm{d}v \Big],
\end{aligned}
$$

by (6.25),

$$\mathbb{E}\big(\big|O(0)_T^{n,N,M}\big|^2\big) \leq KT^{\frac{3p-2}{p}}N^{1-2/p}. \tag{6.42}$$

By similar arguments, we can show the same upper bound applies to $\mathbb{E}\big(\big|O(1)_T^{n,N,M}\big|^2\big)$, $\mathbb{E}\big(\big|O(2)_T^{n,N,M}\big|^2\big)$, $\mathbb{E}\big(\big|O(3)_T^{n,N,M}\big|^2\big)$.

Now, let's consider the terms which contribute to the asymptotic variance. By **Lemma 33**, **31**,

$$\begin{aligned}
V(0)_T^{n,N,M} &\xrightarrow{\mathbb{P}} \frac{T}{2}\int_0^T \rho(t)^2\, c(t)^2\, \mathrm{d}t \\
V(1)_T^{n,N,M} &\xrightarrow{\mathbb{P}} \frac{T}{2}\int_0^T \rho(t)^2\, c(t)^2\, \mathrm{d}t \\
V(2)_T^{n,N,M} &\xrightarrow{\mathbb{P}} \frac{T}{2}\int_0^T \rho(t)^2\, c(t)^2\, \mathrm{d}t,
\end{aligned}$$

thus we have the following lemma

Lemma 36. *Assume* ***Assumption A**-ν, **A**-X, **A**-α. Let $N\Delta(n)^{1/2} \to \infty$, $N \leq \lfloor n/2 \rfloor$, $M \to \infty$, $M < N$,*

$$N^{1/2}\int_0^T \rho(t)\big[\overline{\widehat{c}^{n,N,M}(t)} - c(t)\big]\,\mathrm{d}t \xrightarrow{\mathcal{L}-s} \mathcal{MN}\Big(0, T\int_0^T \rho(t)^2 \times c(t)^2\,\mathrm{d}t\Big).$$

6.5 Bivariate asymptotic analysis

Here we study the asymptotics of

$$N^{1/2}\int_0^T \rho_{jk}(t)\big[\overline{\widehat{c}_{jk}^{n,N,M}(t)} - c_{jk}(t)\big]\,\mathrm{d}t,$$

for a α-Hölder continuous function ρ where α is equal to that in **Assumption A-**α.

Given a smooth function ρ on $\mathbb{R}^+$, by the definition (6.7), we have

$$\begin{aligned}\int_0^T \rho(t)\big[\overline{\widehat{c}_{jk}^{\,n,N,M}(t)} - c_{jk}(t)\big]\,\mathrm{d}t \\ = \frac{1}{T}\sum_{|q|<M}\Big(1-\frac{|q|}{M}\Big)\overline{\widehat{F}(c_{jk})_q^{n,N}}\int_0^T \rho_{jk}(t)e^{-i2\pi qt/T}\,\mathrm{d}t - \int_0^T \rho_{jk}(t)\,c_{jk}(t)\,\mathrm{d}t \\ = \frac{1}{T}\sum_{|q|<M}\Big(1-\frac{|q|}{M}\Big)F(\rho_{jk})_q\,\overline{\widehat{F}(c_{jk})_q^{n,N}} - \int_0^T \rho_{jk}(t)\,c_{jk}(t)\,\mathrm{d}t.\end{aligned}$$

6.5.1 decomposition

Recall the definitions (6.5) and (6.6), based on (6.23) and (6.24), we have the expression

$$\widehat{F}(\mathrm{d}X_j)_{q-s}^n \times \widehat{F}(\mathrm{d}X_k)_s^n = \int_0^T e^{-i2\pi(q-s)\theta_j^n(t)/T}\,\mathrm{d}X_j(t)\int_0^T e^{-i2\pi s\theta_k^n(u)/T}\,\mathrm{d}X_k(u),$$

by (6.15) and Itô's formula,

$$\widehat{F}(\mathrm{d}X_j)_{q-s}^n \times \widehat{F}(\mathrm{d}X_k)_s^n = y_{q,s,jk}^n + \chi(0)_{q,s,jk}^n + \chi(1)_{q,s,jk}^n,$$

where

$$\begin{aligned} y_{q,s,jk}^n &= \int_0^T e^{-i2\pi q\theta_j^n(t)/T} e^{i2\pi s\left[\theta_j^n(t)-\theta_k^n(t)\right]/T} c_{jk}(t)\,\mathrm{d}t \\ \chi(0)_{q,s,jk}^n &= \int_0^T e^{-i2\pi q\theta_j^n(t)/T}\,\mathrm{d}X_j(t)\int_0^t e^{i2\pi s\left[\theta_j^n(t)-\theta_k^n(u)\right]/T}\,\mathrm{d}X_k(u) \\ \chi(1)_{q,s,jk}^n &= \int_0^T \mathrm{d}X_k(t)\int_0^t e^{-i2\pi q\theta_j^n(u)/T} e^{i2\pi s\left[\theta_j^n(u)-\theta_k^n(t)\right]/T}\,\mathrm{d}X_j(u), \end{aligned}$$

then by (6.6), (6.16), (6.24),

$$\widehat{F}(c_{jk})_q^{n,N} = Y_{q,jk}^{n,N} + \Gamma(0)_{q,jk}^{n,N} + \Gamma(1)_{q,jk}^{n,N}, \tag{6.43}$$

where

$$
\begin{aligned}
Y_{q,jk}^{n,N} &= \int_0^T e^{-i2\pi q\theta_j^n(t)/T} d_{jk}^{n,N}(t,t)\, c_{jk}(t)\, \mathrm{d}t \\
\Gamma(0)_{q,jk}^{n,N} &= \int_0^T e^{-i2\pi q\theta_j^n(t)/T}\, \mathrm{d}X_j(t) \int_0^t d_{jk}^{n,N}(t,u)\, \mathrm{d}X_k(u) \\
\Gamma(1)_{q,jk}^{n,N} &= \int_0^T \mathrm{d}X_k(t) \int_0^t e^{-i2\pi q\theta_j^n(u)/T} d_{jk}^{n,N}(u,t)\, \mathrm{d}X_j(u).
\end{aligned}
$$

Let

$$
\begin{aligned}
A_{jk}^{n,N,M} &:= \frac{1}{T}\sum_{|q|<M}\Big(1-\frac{|q|}{M}\Big)F(\rho_{jk})_q\,\overline{\widehat{F}(c_{jk})_q^{n,N}} \\
&= \frac{1}{T}\sum_{|q|<M}\Big(1-\frac{|q|}{M}\Big)F(\rho_{jk})_q\Big[\overline{Y_{q,jk}^{n,N}}+\overline{\Gamma(0)_{q,jk}^{n,N}}+\overline{\Gamma(1)_{q,jk}^{n,N}}\Big],
\end{aligned}
$$

then by (6.2) and (6.43),

$$
N^{1/2}A_{jk}^{n,N,M} = N^{1/2}\int_0^T \widehat{\rho}_{jk}^M(\theta_j^n(t))\, d_{jk}^{n,N}(t,t)\, c_{jk}(t)\, \mathrm{d}t + e(0)_{jk,T}^{n,N,M} + e(1)_{jk,T}^{n,N,M},
$$

where

$$
\begin{aligned}
e(0)_{jk,T}^{n,N,M} &= N^{1/2}\int_0^T \widehat{\rho}_{jk}^M(\theta_j^n(t))\, \sigma_{j\cdot}(t)\, \mathrm{d}W(t) \int_0^t d_{jk}^{n,N}(t,u)\, \sigma_{k\cdot}(u)\, \mathrm{d}W(u) \\
e(1)_{jk,T}^{n,N,M} &= N^{1/2}\int_0^T \sigma_{k\cdot}(t)\, \mathrm{d}W(t) \int_0^t \widehat{\rho}_{jk}^M(\theta_j^n(u))\, d_{jk}^{n,N}(u,t)\, \sigma_{j\cdot}(u)\, \mathrm{d}W(u). \quad (6.44)
\end{aligned}
$$

Therefore, we have the following decomposition:

$$
N^{1/2}\int_0^T \rho_{jk}(t)\big[\overline{\widehat{c}_{jk}^{n,N,M}(t)}-c_{jk}(t)\big]\, \mathrm{d}t = o(0)_{jk,T}^{n,M}+o(1)_{jk,T}^{n,M}+e(0)_{jk,T}^{n,N,M}+e(1)_{jk,T}^{n,N,M}, \quad (6.45)
$$

where

$$
\begin{aligned}
o(0)^{n,M}_{jk,T} &= N^{1/2} \sum_{h=1}^{n_j} \int_{I^j_h} \big[\widehat{\rho}^M_{jk}(\tau^j_h) - \rho_{jk}(t)\big] c_{jk}(t)\,\mathrm{d}t \\
o(1)^{n,N,M}_{jk,T} &= N^{1/2} \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, c_{jk}(t)\, \big[d^{n,N}_{jk}(t,t) - 1\big]\,\mathrm{d}t.
\end{aligned}
\tag{6.46}
$$

On one hand, by (6.27), (6.31), Fubini's theorem, and **Lemma 34**,

$$
\begin{aligned}
o(0)^{n,M}_{jk,T} = N^{1/2} \int_0^T c_{jk}(t)\,\mathrm{d}t\, \frac{1}{T} \int_0^T F^M\Big(\frac{u-\theta^n_j(t)}{T}\Big)\big[\rho_{jk}(u) - \rho_{jk}(t)\big]\,\mathrm{d}u \\
= \frac{N^{1/2}}{T} J^M_{jk,T} + O_p\big(N^{1/2}M/n\big),
\end{aligned}
$$

where

$$
J^M_{jk,T} = \int_0^T \int_0^T F^M\Big(\frac{u-t}{T}\Big)\big[\rho_{jk}(u) - \rho_{jk}(t)\big] c_{jk}(t)\,\mathrm{d}t\,\mathrm{d}u.
$$

By symmetry of variables, $J^M_{jk,T} = \int_0^T \int_0^T F^M[(u-t)/T]\big[\rho_{jk}(t) - \rho_{jk}(u)\big] c_{jk}(u)\,\mathrm{d}u\,\mathrm{d}t$, hence

$$
J^M_{jk,T} = -\frac{1}{2} \int_0^T \mathrm{d}u \int_0^T F^M\Big(\frac{u-t}{T}\Big)\big[\rho_{jk}(u) - \rho_{jk}(t)\big]\big[c_{jk}(u) - c_{jk}(t)\big]\,\mathrm{d}t.
$$

The modulus of continuity of ρ is determined by that of c, let

$$
L^M_{jk,T}(u) \coloneqq \int_0^T F^M[(u-t)/T]\big[\rho_{jk}(u) - \rho_{jk}(t)\big]\big[c_{jk}(u) - c_{jk}(t)\big]\,\mathrm{d}t,
$$

by periodicity of c and ρ, $L^M_{jk,T}(u) = \int_{u-T/2}^{u+T/2} F^M[(u-t)/T]\big[\rho_{jk}(u) - \rho_{jk}(t)\big]\big[c_{jk}(u) - c_{jk}(t)\big]\,\mathrm{d}t$. Note

$$
\big|L^M_{jk,T}(u)\big| \le \Big(\int_{|u-t|\le 1/M} + \int_{|u-t|>1/M}\Big) F^M\Big(\frac{u-t}{T}\Big)\big|\rho_{jk}(u) - \rho_{jk}(t)\big|\big|c_{jk}(u) - c_{jk}(t)\big|\,\mathrm{d}t
$$

through an argument similar to the proof of **Lemma 33** , we have $\mathbb{E}\big|L^M_{jk,T}(u)\big| \le K\big[M^{-2\alpha} +$

$M^{-(1+\alpha)}\big]$, thus

$$\mathbb{E}\big|o(0)^{n,M}_{jk,T}\big| \le K\Big[\Big(\frac{N}{M^{4\alpha}}\Big)^{1/2} + \Big(\frac{N}{n}\Big)^{1/2}\frac{M}{n^{1/2}}\Big],$$

by Markov's inequality, we have shown the asymptotic negligibility in probability of $o(0)^{n,M}_{jk,T}$, i.e.,

$$o(0)^{n,M}_{jk,T} \xrightarrow{\mathbb{P}} 0. \tag{6.47}$$

On the other hand, by the Taylor series of the sine function and the definition (6.24), $\big|d^{n,N}_{jk}(t,t)-1\big| \le KN\Delta(n)$, so $\mathbb{E}\big(\big|o(1)^{n,N,M}_{jk,T}\big|\big) \le KN^{3/2}\Delta(n)$, hence

$$o(1)^{n,N,M}_{jk,T} \xrightarrow{\mathbb{P}} 0. \tag{6.48}$$

Thus the asymptotics is dictated by $e(0)^{n,N,M}_{jk,T} + e(1)^{n,N,M}_{jk,T}$.

6.5.2 preparing some martingales

Now let's define some Itô martingales useful in the incoming asymptotic analysis. They are

$$\begin{aligned}
U^{n,N}_{jk}(t) &= \int_0^t d^{n,N}_{jk}(t,u)\,\sigma_{k\cdot}(u)\,\mathrm{d}W(u)\\
\widetilde{U}^{n,N}_{jk}(t) &= \int_0^t d^{n,N}_{jk}(u,t)\,\widehat{\rho}^{M}_{jk}(\theta^n_j(u))\,\sigma_{j\cdot}(u)\,\mathrm{d}W(u)\\
\widehat{U}^{n,N}_{jk}(t) &= \int_0^u F^M\Big(\frac{t-\theta^n_j(v)}{T}\Big)\,d^{n,N}_{jk}(v,u)\,\sigma_{j\cdot}(v)\,\mathrm{d}W(v),
\end{aligned} \tag{6.49}$$

and

$$
\begin{aligned}
Z_{jk}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(t,u)\,\sigma_{k\cdot}(u)\,U_{jk}^{n,N}(u)\,\mathrm{d}W(u) \\
\breve{Z}_{jk}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(t,u)\,\sigma_{k\cdot}(u)\,\widetilde{U}_{jk}^{n,N}(u)\,\mathrm{d}W(u) \\
\mathring{Z}_{jk}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(u,t)\,\widehat{\rho}_{jk}^{M}(\theta_j^n(u))\,\sigma_{j\cdot}(u)\,U_{jk}^{n,N}(u)\,\mathrm{d}W(u) \\
\widetilde{Z}_{jk}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(u,t)\,\widehat{\rho}_{jk}^{M}(\theta_j^n(u))\,\sigma_{j\cdot}(u)\,\widetilde{U}_{jk}^{n,N}(u)\,\mathrm{d}W(u).
\end{aligned} \tag{6.50}
$$

By Itô's formula,

$$
\begin{aligned}
U_{jk}^{n,N}(t)^2 &= \int_0^t d_{jk}^{n,N}(t,u)^2\,c_{kk}(u)\,\mathrm{d}u + 2Z_{jk}^{n,N}(t) \\
U_{jk}^{n,N}(t)\,\widetilde{U}_{jk}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(t,u)\,d_{jk}^{n,N}(u,t)\,\widehat{\rho}_{jk}^{M}(\theta_j^n(u))\,c_{jk}(u)\,\mathrm{d}u + \breve{Z}_{jk}^{n,N}(t) + \mathring{Z}_{jk}^{n,N}(t) \\
\widetilde{U}_{jk}^{n,N}(t)^2 &= \int_0^t d_{jk}^{n,N}(u,t)^2\,\widehat{\rho}_{jk}^{M}(\theta_j^n(u))^2\,c_{jj}(u)\,\mathrm{d}u + 2\widetilde{Z}_{jk}^{n,N}(t).
\end{aligned} \tag{6.51}
$$

We have the following lemma about the magnitudes of quadratics.

Lemma 37. *Under* ***Lemma A-ν****, there exist some finite positive constant K such that*

$$
\begin{aligned}
\mathbb{E}\big[U_{jk}^{n,N}(t)U_{jk}^{n,N}(u)\big] &\le K\int_0^{t\wedge u} d_{jk}^{n,N}(t,v)\,d_{jk}^{n,N}(u,v)\,\mathrm{d}v \\
\mathbb{E}\big[U_{jk}^{n,N}(t)\widetilde{U}_{jk}^{n,N}(u)\big] &\le K\int_0^{t\wedge u} d_{jk}^{n,N}(t,v)\,d_{jk}^{n,N}(v,u)\,\mathrm{d}v \\
\mathbb{E}\big[\widetilde{U}_{jk}^{n,N}(t)\widetilde{U}_{jk}^{n,N}(u)\big] &\le K\int_0^{t\wedge u} d_{jk}^{n,N}(v,t)\,d_{jk}^{n,N}(v,u)\,\mathrm{d}v,
\end{aligned} \tag{6.52}
$$

and

$$
\begin{aligned}
\mathbb{E}\big[Z^{n,N}_{jk}(t)Z^{n,N}_{jk}(u)\big] &\leq K\int_0^{t\wedge u} d^{n,N}_{jk}(t,v)\,d^{n,N}_{jk}(u,v)\,\mathrm{d}v\int_0^{v} d^{n,N}_{jk}(v,\vartheta)^2\,\mathrm{d}\vartheta\\
\mathbb{E}\big[\breve{Z}^{n,N}_{jk}(t)\breve{Z}^{n,N}_{jk}(u)\big] &\leq K\int_0^{t\wedge u} d^{n,N}_{jk}(t,v)\,d^{n,N}_{jk}(u,v)\,\mathrm{d}v\int_0^{v} d^{n,N}_{jk}(\vartheta,v)^2\,\mathrm{d}\vartheta\\
\mathbb{E}\big[\mathring{Z}^{n,N}_{jk}(t)\mathring{Z}^{n,N}_{jk}(u)\big] &\leq K\int_0^{t\wedge u} d^{n,N}_{jk}(v,t)\,d^{n,N}_{jk}(v,u)\,\mathrm{d}v\int_0^{v} d^{n,N}_{jk}(v,\vartheta)^2\,\mathrm{d}\vartheta\\
\mathbb{E}\big[\widetilde{Z}^{n,N}_{jk}(t)\widetilde{Z}^{n,N}_{jk}(u)\big] &\leq K\int_0^{t\wedge u} d^{n,N}_{jk}(v,t)\,d^{n,N}_{jk}(v,u)\,\mathrm{d}v\int_0^{v} d^{n,N}_{jk}(\vartheta,v)^2\,\mathrm{d}\vartheta.
\end{aligned}
$$

Proof. By Itô's formula and (6.49),

$$
\begin{aligned}
\mathbb{E}\big[U^{n,N}_{jk}(t)U^{n,N}_{jk}(u)\big] &= \int_0^{t\wedge u} d^{n,N}_{jk}(t,v)\,d^{n,N}_{jk}(u,v)\,\mathbb{E}[c_{jk}(v)]\,\mathrm{d}v\\
\mathbb{E}\big[U^{n,N}_{jk}(t)\widetilde{U}^{n,N}_{jk}(u)\big] &= \int_0^{t\wedge u} d^{n,N}_{jk}(t,v)\,d^{n,N}_{jk}(v,u)\,\mathbb{E}\big[\widehat{\rho}^{M}_{jk}(\theta^n_j(v))\,c_{jk}(v)\big]\,\mathrm{d}v\\
\mathbb{E}\big[\widetilde{U}^{n,N}_{jk}(t)\widetilde{U}^{n,N}_{jk}(u)\big] &= \int_0^{t\wedge u} d^{n,N}_{jk}(v,t)\,d^{n,N}_{jk}(v,u)\,\mathbb{E}\big[\widehat{\rho}^{M}_{jk}(\theta^n_j(v))^2\,c_{jk}(v)\big]\,\mathrm{d}v,
\end{aligned}
$$

thereby the first claim follows.

According to Itô's formula, (6.58), (6.59),

$$
\begin{aligned}
\mathbb{E}\big[Z^{n,N}_{jk}(t)Z^{n,N}_{jk}(u)\big] = \mathbb{E}\int_0^{t\wedge u} d^{n,N}_{jk}(t,v)\,d^{n,N}_{jk}(u,v)\,c_{kk}(v)\,U^{n,N}_{jk}(v)^2\,\mathrm{d}v&\\
\leq K\int_0^{t\wedge u} d^{n,N}_{jk}(t,v)\,d^{n,N}_{jk}(u,v)\,\mathbb{E}\big[U^{n,N}_{jk}(v)^2\big]\,\mathrm{d}v,&
\end{aligned}
$$

similarly,

$$
\begin{aligned}
\mathbb{E}\big[\breve{Z}^{n,N}_{jk}(t)\breve{Z}^{n,N}_{jk}(u)\big] &\leq K\int_0^{t\wedge u} d^{n,N}_{jk}(t,v)\,d^{n,N}_{jk}(u,v)\,\mathbb{E}\big[\widetilde{U}^{n,N}_{jk}(v)^2\big]\,\mathrm{d}v\\
\mathbb{E}\big[\mathring{Z}^{n,N}_{jk}(t)\mathring{Z}^{n,N}_{jk}(u)\big] &\leq K\int_0^{t\wedge u} d^{n,N}_{jk}(v,t)\,d^{n,N}_{jk}(v,u)\,\mathbb{E}\big[U^{n,N}_{jk}(v)^2\big]\,\mathrm{d}v\\
\mathbb{E}\big[\widetilde{Z}^{n,N}_{jk}(t)\widetilde{Z}^{n,N}_{jk}(u)\big] &\leq K\int_0^{t\wedge u} d^{n,N}_{jk}(v,t)\,d^{n,N}_{jk}(v,u)\,\mathbb{E}\big[\widetilde{U}^{n,N}_{jk}(v)^2\big]\,\mathrm{d}v,
\end{aligned}
$$

then the second claim follows from the first claim proved earlier. $\square$

6.5.3 local mean square rate

According to (6.6), (6.43) and (6.49), we can write

$$\widehat{c}_{jk}^{n,N,M}(t) = \frac{1}{T}\int_0^T F^M\Big(\frac{t-\theta_j^n(u)}{T}\Big)\, d_{jk}^{n,N}(u,u)\, c_{jk}(u)\,\mathrm{d}u \\ + \frac{1}{T}\int_0^T F^M\Big(\frac{t-\theta_j^n(u)}{T}\Big)\, U_{jk}^{n,N}(u)\,\sigma_{j\cdot}(u)\,\mathrm{d}W(u) + \frac{1}{T}\int_0^T \widehat{U}_{jk,T}^{n,N}(t,u)\,\sigma_{k\cdot}(u)\,\mathrm{d}W(u),$$

therefore

$$\widehat{c}_{jk}^{n,N,M}(t) - c_{jk}(t) = Q(t,0)_{jk,T}^{n,N,M} + Q(t,1)_{jk,T}^{n,N,M} + \Omega(t)_{jk,T}^{n,N,M}, \tag{6.53}$$

where

$$\begin{aligned} Q(t,0)_{jk,T}^{n,N,M} &= \frac{1}{T}\int_0^T F^M\Big(\frac{t-\theta_j^n(u)}{T}\Big)\, U_{jk}^{n,N}(u)\,\sigma_{j\cdot}(u)\,\mathrm{d}W(u) \\ Q(t,1)_{jk,T}^{n,N,M} &= \frac{1}{T}\int_0^T \widehat{U}_{jk,T}^{n,N,M}(t,u)\,\sigma_{k\cdot}(u)\,\mathrm{d}W(u) \\ \Omega(t)_{jk,T}^{n,N,M} &= \frac{1}{T}\int_0^T F^M\Big(\frac{t-\theta_j^n(u)}{T}\Big)\, d_{jk}^{n,N}(u,u)\, c_{jk}(u)\,\mathrm{d}u - c_{jk}(t). \end{aligned}$$

By Itô's formula and Fubini's theorem,

$$\begin{aligned} \mathbb{E}\big[\big|Q(t,0)_{jk,T}^{n,N,M}\big|^2\big] &= \frac{1}{T^2}\int_0^T F^M\Big(\frac{t-\theta_j^n(u)}{T}\Big)^2\, \mathbb{E}\big[U_{jk}^{n,N}(u)^2\, c_{jj}(u)\big]\,\mathrm{d}u \\ \mathbb{E}\big[\big|Q(t,1)_{jk,T}^{n,N,M}\big|^2\big] &= \frac{1}{T^2}\int_0^T \mathbb{E}\big[\widehat{U}_{jk,T}^{n,N,M}(t,u)^2\, c_{kk}(u)\big]\,\mathrm{d}u, \end{aligned}$$

because of **Assumption A-ν** and **Lemma 37**,

$$\begin{aligned}\mathbb{E}\big[\big|Q(t,0)^{n,N,M}_{jk,T}\big|^2\big] &\leq K\int_0^T F^M\Big(\frac{t-\theta^n_j(u)}{T}\Big)^2\,\mathrm{d}u\int_0^u d^{n,N}_{jk}(u,v)^2\,\mathrm{d}v\\ &\leq K\Big[\sup_{u\in[0,T]}\int_0^T d^{n,N}_{jk}(u,v)^2\,\mathrm{d}v\Big]\cdot\Big[\int_0^T F^M\Big(\frac{t-\theta^n_j(u)}{T}\Big)^2\,\mathrm{d}u\Big],\end{aligned}$$

and by Fubini's theorem,

$$\begin{aligned}\mathbb{E}\big[\big|Q(t,1)^{n,N,M}_{jk,T}\big|^2\big] &\leq K\int_0^T \mathrm{d}u\int_0^u F^M\Big(\frac{t-\theta^n_j(v)}{T}\Big)^2 d^{n,N}_{jk}(v,u)^2\,\mathrm{d}v\\ &= K\int_0^T F^M\Big(\frac{t-\theta^n_j(v)}{T}\Big)^2\,\mathrm{d}v\int_v^T d^{n,N}_{jk}(v,u)^2\,\mathrm{d}u\\ &\leq K\Big[\sup_{v\in[0,T]}\int_0^T d^{n,N}_{jk}(v,u)^2\,\mathrm{d}u\Big]\cdot\Big[\int_0^T F^M\Big(\frac{t-\theta^n_j(u)}{T}\Big)^2\,\mathrm{d}u\Big],\end{aligned}$$

according to (6.22) and **Lemma 30**,

$$\big|Q(t,0)^{n,N,M}_{jk,T}\big|^2+\big|Q(t,1)^{n,N,M}_{jk,T}\big|^2\leq K\frac{M}{N}. \tag{6.54}$$

Notice

$$\Omega(t)^{n,N,M}_{jk,T}=\Omega(t,0)^{n,N,M}_{jk,T}+\Omega(t,1)^{n,M}_{jk,T}+\Omega(t,2)^{M}_{jk,T}, \tag{6.55}$$

where

$$\begin{aligned}\Omega(t,0)^{n,N,M}_{jk,T} &= \frac{1}{T}\int_0^T F^M\Big(\frac{t-\theta^n_j(u)}{T}\Big)\,c_{jk}(u)\,\big[d^{n,N}_{jk}(u,u)-1\big]\,\mathrm{d}u\\ \Omega(t,1)^{n,M}_{jk,T} &= \frac{1}{T}\int_0^T F^M\Big(\frac{t-\theta^n_j(u)}{T}\Big)\,c_{jk}(u)\,\mathrm{d}u-\frac{1}{T}\int_0^T F^M\Big(\frac{t-u}{T}\Big)\,c_{jk}(u)\,\mathrm{d}u\\ \Omega(t,2)^{M}_{jk,T} &= \frac{1}{T}\int_0^T F^M\Big(\frac{t-u}{T}\Big)\,c_{jk}(u)\,\mathrm{d}u-c_{jk}(t).\end{aligned}$$

Based on the Taylor series of sine function, **Assumption A-**ν, (6.27),

$$\mathbb{E}\Big(\sup_{t\in[0,T]}\big|\Omega(t,0)^{n,N,M}_{jk,T}\big|^2\Big)\le K\frac{N^4}{\underline{n}^4}\mathbb{1}_{\{j\neq k\}};$$

according to **Assumption A-**ν, **34**, we know

$$\mathbb{E}\Big(\sup_{t\in[0,T]}\big|\Omega(t,1)^{n,M}_{jk,T}\big|^2\Big)\le K\frac{M^2}{\underline{n}^2};$$

by **Lemma 33**,

$$\begin{array}{ll}\mathbb{E}\Big(\sup_{t\in[1/M,T-1/M]}\big|\Omega(t,2)^{M}_{jk,T}\big|^2\Big)\le KM^{-2\alpha}, & \text{if } c(0)\neq c(T)\\ \mathbb{E}\Big(\sup_{t\in[0,T]}\big|\Omega(t,2)^{M}_{jk,T}\big|^2\Big)\le KM^{-2\alpha}, & \text{if } c(0)=c(T),\end{array}$$

then based on (6.53), (6.54), (6.55), (7.2), we have the following proposition.

Proposition 5. *Under* ***Assumption A-***ν*,* ***A-***X*,* ***A-***α*, there exists a finite positive constant* K *such that* $\forall j,k=1,\cdots,d$*,*

$$\sup_{t\in[M^{-1},T-M^{-1}]}\mathbb{E}\big|\widehat{c}^{n,N,M}_{jk}(t)-c_{jk}(t)\big|^2\le K\Big(\frac{N^4}{\underline{n}^4}\mathbb{1}_{\{j\neq k\}}+M^{-2\alpha}+\frac{M}{N}\Big);$$

additionally, if $c(0)=c(T)$*,*

$$\sup_{t\in[0,T]}\mathbb{E}\big|\widehat{c}^{n,N,M}_{jk}(t)-c_{jk}(t)\big|^2\le K\Big(\frac{N^4}{\underline{n}^4}\mathbb{1}_{\{j\neq k\}}+M^{-2\alpha}+\frac{M}{N}\Big).$$

Remark 10. The various terms in the upper bound in proposition 5 arise from estimation errors of different natures, cf. (6.53) and (6.55). The sources of these estimation errors are:

- asynchronous observations;
- approximation by convolution with the Fejér kernel (one type of delta sequences in Fourier analysis);

- statistical error in the form of stochastic integrals of the Fejér kernel with respect to Brownian motion.

The magnitude of these estimation errors are summarized in table 6.2.

Table 6.2: Estimation errors of the Fourier-Malliavin method for spot volatility

error sources	asynchronicity error	delta sequence approximation	statistical error
magnitudes	$\asymp N^2\Delta(n)^2$	$\asymp M^{-\alpha}$	$\asymp \sqrt{M/N}$

6.5.4 stable convergence

By (6.44) and (6.49), we can write

$$\begin{aligned} e(0)^{n,N,M}_{jk,T} &= N^{1/2}\int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\,\sigma_{j\cdot}(t)\,U^{n,N}_{jk}(t)\,\mathrm{d}W(t) \\ e(1)^{n,N,M}_{jk,T} &= N^{1/2}\int_0^T \sigma_{k\cdot}(t)\,\widetilde{U}^{n,N}_{jk}(t)\,\mathrm{d}W(t). \end{aligned} \tag{6.56}$$

First, let's consider, for $r = 1\cdots, d'$,

$$\begin{aligned} \langle e(0)^{n,N,M}_{jk}, W_r\rangle_T &= N^{1/2}\int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\,\sigma_{jr}(t)\,U^{n,N}_{jk}(t)\,\mathrm{d}t \\ \langle e(1)^{n,N,M}_{jk}, W_r\rangle_T &= N^{1/2}\int_0^T \sigma_{kr}(t)\,\widetilde{U}^{n,N}_{jk}(t)\,\mathrm{d}t. \end{aligned}$$

Notice that

$$\langle e(0)^{n,N,M}_{jk}, W_r\rangle^2_T = N\int_0^T\int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\widehat{\rho}^M_{jk}(\theta^n_j(u))\,\sigma_{jr}(t)\sigma_{jr}(u) \times U^{n,N}_{jk}(t)U^{n,N}_{jk}(u)\,\mathrm{d}t\,\mathrm{d}u,$$

according to **Lemma 37**, Fubini's theorem, Hölder's inequality,

$$
\begin{aligned}
\mathbb{E}\big[\langle e(0)^{n,N,M}_{jk}, W_r\rangle^2_T\big] &\le KN \int_0^T \int_0^T \mathrm{d}t\,\mathrm{d}u \Big(\int_0^{t\wedge u} d^{n,N}_{jk}(t,v)\, d^{n,N}_{jk}(u,v)\,\mathrm{d}v \Big) \\
&\le KN \int_0^T \mathrm{d}v \Big(\int_v^T \big|d^{n,N}_{jk}(t,v)\big|\,\mathrm{d}t \Big)\Big(\int_v^T \big|d^{n,N}_{jk}(u,v)\big|\,\mathrm{d}u \Big) \\
&\le KN \int_0^T \mathrm{d}v \Big(\int_v^T \big|d^{n,N}_{jk}(t,v)\big|\,\mathrm{d}t \Big)^2 \le KT^{\frac{3p-2}{p}} N \Big(\int_0^T \big|d^{n,N}_{jk}(t,v)\big|^p\,\mathrm{d}t \Big)^{2/p};
\end{aligned}
$$

similarly,

$$
\langle e(1)^{n,N,M}_{jk}, W_r\rangle^2_T = N \int_0^{\tau_n} \int_0^{\tau_n} \sigma_{kr}(t)\sigma_{kr}(u) \times \widetilde{U}^{n,N}_{jk}(t)\, \widetilde{U}^{n,N}_{jk}(u)\,\mathrm{d}t\,\mathrm{d}u,
$$

by a similar argument applied to $\mathbb{E}[\langle e(0)^{n,N,M}_{jk}, W_r\rangle^2_T]$,

$$
\mathbb{E}\big[\langle e(1)^{n,N,M}_{jk}, W_r\rangle^2_T\big] \le KT^{\frac{3p-2}{p}} N \Big(\int_0^T \big|d^{n,N}_{jk}(v,t)\big|^p\,\mathrm{d}t \Big)^{2/p}.
$$

By **Lemma 30**, Jensen's inequality and Markov's inequality, we have the following lemma.

Lemma 38. *Under* ***Lemma A-ν***,

$$
\langle e(0)^{n,N,M}_{jk} + e(1)^{n,N,M}_{jk}, W_r\rangle_T \xrightarrow{\mathbb{P}} 0.
$$

Second, let's consider

$$
\begin{aligned}
\langle e(0)^{n,N,M}_{jk}, e(0)^{n,N,M}_{jk}\rangle_T &= N \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))^2\, c_{jj}(t)\, U^{n,N}_{jk}(t)^2\,\mathrm{d}t \\
\langle e(0)^{n,N,M}_{jk}, e(1)^{n,N,M}_{jk}\rangle_T &= N \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, c_{jk}(t)\, U^{n,N}_{jk}(t)\, \widetilde{U}^{n,N}_{jk}(t)\,\mathrm{d}t \\
\langle e(1)^{n,N,M}_{jk}, e(1)^{n,N,M}_{jk}\rangle_T &= N \int_0^T c_{kk}(t)\, \widetilde{U}^{n,N}_{jk}(t)^2\,\mathrm{d}t,
\end{aligned}
$$

in view of (6.59),

$$
\begin{aligned}
\big\langle e(0)^{n,N,M}_{jk}, e(0)^{n,N,M}_{jk}\big\rangle_T &= 2O(0)^{n,N,M}_{jk,T} + V(0)^{n,N,M}_{jk,T}\\
\big\langle e(0)^{n,N,M}_{jk}, e(1)^{n,N,M}_{jk}\big\rangle_T &= O(1)^{n,N,M}_{jk,T} + O(2)^{n,N,M}_{jk,T} + V(1)^{n,N,M}_{jk,T}\\
\big\langle e(1)^{n,N,M}_{jk}, e(1)^{n,N,M}_{jk}\big\rangle_T &= 2O(3)^{n,N,M}_{jk,T} + V(2)^{n,N,M}_{jk,T},
\end{aligned}
$$

where

$$
\begin{aligned}
O(0)^{n,N,M}_{jk,T} &= N\int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))^2\, c_{jj}(t)\, Z^{n,N}_{jk}(t)\,\mathrm{d}t\\
O(1)^{n,N,M}_{jk,T} &= N\int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, c_{jk}(t)\, \breve{Z}^{n,N}_{jk}(t)\,\mathrm{d}t\\
O(2)^{n,N,M}_{jk,T} &= N\int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, c_{jk}(t)\, \mathring{Z}^{n,N}_{jk}(t)\,\mathrm{d}t\\
O(3)^{n,N,M}_{jk,T} &= N\int_0^T c_{kk}(t)\, \widetilde{Z}^{n,N}_{jk}(t)\,\mathrm{d}t,
\end{aligned}
$$

and

$$
\begin{aligned}
V(0)^{n,N,M}_{jk,T} &= \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))^2\, c_{jj}(t)\,\mathrm{d}t\left[N\int_0^t d^{n,N}_{jk}(t,u)^2\, c_{kk}(u)\,\mathrm{d}u\right]\\
V(1)^{n,N,M}_{jk,T} &= \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, c_{jk}(t)\,\mathrm{d}t\left[N\int_0^t d^{n,N}_{jk}(t,u)\, d^{n,N}_{jk}(u,t)\, \widehat{\rho}^M_{jk}(\theta^n_j(u))\, c_{jk}(u)\,\mathrm{d}u\right]\\
V(2)^{n,N,M}_{jk,T} &= \int_0^T c_{kk}(t)\,\mathrm{d}t\left[N\int_0^t d^{n,N}_{jk}(u,t)^2\, \widehat{\rho}^M_{jk}(\theta^n_j(u))^2\, c_{jj}(u)\,\mathrm{d}u\right].
\end{aligned}
$$

Let consider the asymptotically negligible terms,

$$\begin{aligned}
\big|O(0)^{n,N,M}_{jk,T}\big|^2 &= N^2 \int_0^T \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))^2\, \widehat{\rho}^M_{jk}(\theta^n_j(u))^2\, c_{jj}(t)\, c_{jj}(u)\, Z^{n,N}_{jk}(t)\, Z^{n,N}_{jk}(u)\,\mathrm{d}t\,\mathrm{d}u \\
\big|O(1)^{n,N,M}_{jk,T}\big|^2 &= N^2 \int_0^T \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, \widehat{\rho}^M_{jk}(\theta^n_j(u))\, c_{jk}(t)\, c_{jk}(u)\, \breve{Z}^{n,N}_{jk}(t)\, \breve{Z}^{n,N}_{jk}(u)\,\mathrm{d}t\,\mathrm{d}u \\
\big|O(2)^{n,N,M}_{jk,T}\big|^2 &= N^2 \int_0^T \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, \widehat{\rho}^M_{jk}(\theta^n_j(u))\, c_{jk}(t)\, c_{jk}(u)\, \mathring{Z}^{n,N}_{jk}(t)\, \mathring{Z}^{n,N}_{jk}(u)\,\mathrm{d}t\,\mathrm{d}u \\
\big|O(3)^{n,N,M}_{jk,T}\big|^2 &= N^2 \int_0^T \int_0^T c_{kk}(t)\, c_{kk}(u)\, \widetilde{Z}^{n,N}_{jk}(t)\, \widetilde{Z}^{n,N}_{jk}(u)\,\mathrm{d}t\,\mathrm{d}u,
\end{aligned}$$

by **Lemma 37**,

$$\begin{aligned}
&\mathbb{E}\big(\big|O(0)^{n,N,M}_{jk,T}\big|^2\big) \\
&\qquad\qquad \le KN^2 \int_0^T \int_0^T \mathrm{d}t\,\mathrm{d}u \Big[\int_0^{t\wedge u} d^{n,N}_{jk}(t,v)\, d^{n,N}_{jk}(u,v)\,\mathrm{d}v \int_0^v d^{n,N}_{jk}(v,\vartheta)^2\,\mathrm{d}\vartheta\Big],
\end{aligned}$$

then by (6.25),

$$\mathbb{E}\big(\big|O(0)^{n,N,M}_{jk,T}\big|^2\big) \le KT^{\frac{3p-2}{p}} N^{1-2/p}. \tag{6.57}$$

By similar arguments, we can show the same upper bound applies to $\mathbb{E}\big(\big|O(1)^{n,N,M}_{jk,T}\big|^2\big)$, $\mathbb{E}\big(\big|O(2)^{n,N,M}_{jk,T}\big|^2\big)$, $\mathbb{E}\big(\big|O(3)^{n,N,M}_{jk,T}\big|^2\big)$.

Now, let's consider the terms which contribute to the asymptotic variance. By **Lemma 31** and **Lemma 33**,

$$\begin{aligned}
V(0)^{n,N,M}_{jk,T} &\xrightarrow{\mathbb{P}} \int_0^T \rho_{jk}(t)^2\, \tilde{\theta}_{jk,jk}(t)\, c_{jj}(t)\, c_{kk}(t)\,\mathrm{d}t \\
V(1)^{n,N,M}_{jk,T} &\xrightarrow{\mathbb{P}} \int_0^T \rho_{jk}(t)^2\, \check{\theta}_{jk,jk}(t)\, c_{jk}(t)^2\,\mathrm{d}t \\
V(2)^{n,N,M}_{jk,T} &\xrightarrow{\mathbb{P}} \int_0^T \rho_{jk}(t)^2\, \grave{\theta}_{jk,jk}(t)\, c_{jj}(t)\, c_{kk}(t)\,\mathrm{d}t.
\end{aligned}$$

Based on the derivation above, we have the following lemma.

Lemma 39. *Assume* ***Assumption A**-ν, **A**-X, **A**-α, **A**-d^2. Let $N\Delta(n)^{1/2} \to \infty$, $N\Delta(n)^{2/3} \to 0$, $M \to \infty$, $M < N$,*

$$N^{1/2} \int_0^T \rho(t) \big[\overline{\widehat{c}_{jk}^{n,N,M}(t)} - c_{jk}(t)\big] \,\mathrm{d}t \xrightarrow{\mathcal{L}-s} \mathcal{MN}\big(0, V_{jk,T}\big),$$

where

$$V_{jk,T} = \int_0^T \rho(t)^2 \times \Big\{ \big[\tilde{\theta}_{jk,jk}(t) + \grave{\theta}_{jk,jk}(t)\big] c_{jj}(t)\, c_{kk}(t) + 2\check{\theta}_{jk,jk}(t)\, c_{jk}(t)^2 \Big\} \,\mathrm{d}t.$$

6.6 Multivariate asymptotic analysis

Finally, we are ready to tackle the general multivariate asymptotics

$$N^{1/2} \sum_{j,k=1}^{d} \int_0^T \rho_{jk}(t) \big[\overline{\widehat{c}_{jk}^{n,N,M}(t)} - c_{jk}(t)\big] \,\mathrm{d}t.$$

for a α-Hölder continuous function ρ where α is specified in **Assumption A-**α.

6.6.1 prepare more martingales

Define the following Itô martingales:

$$\begin{aligned}
Z_{jk,lm}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(t,u)\, \sigma_{k\cdot}(u)\, U_{lm}^{n,N}(u) \,\mathrm{d}W(u) \\
\breve{Z}_{jk,lm}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(t,u)\, \sigma_{k\cdot}(u)\, \widetilde{U}_{lm}^{n,N}(u) \,\mathrm{d}W(u) \\
\mathring{Z}_{jk,lm}^{n,N}(t) &= \int_0^t d_{lm}^{n,N}(u,t)\, \widehat{\rho}_{lm}^{M}(\theta_l^n(u))\, \sigma_{l\cdot}(u)\, U_{jk}^{n,N}(u) \,\mathrm{d}W(u) \\
\widetilde{Z}_{jk,lm}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(u,t)\, \widehat{\rho}_{jk}^{M}(\theta_j^n(u))\, \sigma_{j\cdot}(u)\, \widetilde{U}_{lm}^{n,N}(u) \,\mathrm{d}W(u).
\end{aligned} \tag{6.58}$$

Based the definition (6.49) and by Itô's formula,

$$
\begin{aligned}
U_{jk}^{n,N}(t)\,U_{lm}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(t,u)\,d_{lm}^{n,N}(t,u)\,c_{km}(u)\,\mathrm{d}u + Z_{jk,lm}^{n,N}(t) + Z_{lm,jk}^{n,N}(t) \\
U_{jk}^{n,N}(t)\,\widetilde{U}_{lm}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(t,u)\,d_{lm}^{n,N}(u,t)\,\widehat{\rho}_{lm}^{M}(\theta_l^n(u))\,c_{kl}(u)\,\mathrm{d}u + \breve{Z}_{jk,lm}^{n,N}(t) + \mathring{Z}_{jk,lm}^{n,N}(t) \\
\widetilde{U}_{jk}^{n,N}(t)\,U_{lm}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(u,t)\,d_{lm}^{n,N}(t,u)\,\widehat{\rho}_{jk}^{M}(\theta_j^n(u))\,c_{jm}(u)\,\mathrm{d}u + \breve{Z}_{lm,jk}^{n,N}(t) + \mathring{Z}_{lm,jk}^{n,N}(t) \\
\widetilde{U}_{jk}^{n,N}(t)\,\widetilde{U}_{lm}^{n,N}(t) &= \int_0^t d_{jk}^{n,N}(u,t)\,d_{lm}^{n,N}(u,t)\,\widehat{\rho}_{jk}^{M}(\theta_j^n(u))\,\widehat{\rho}_{lm}^{M}(\theta_l^n(u))\,c_{jl}(u)\,\mathrm{d}u \\
&\qquad + \widetilde{Z}_{jk,lm}^{n,N}(t) + \widetilde{Z}_{lm,jk}^{n,N}(t). \qquad (6.59)
\end{aligned}
$$

We have the following lemma about the magnitudes of quadratics.

Lemma 40. *There exist some finite positive constant K such that*

$$
\begin{aligned}
\mathbb{E}\big[Z_{jk,lm}^{n,N}(t)Z_{jk,lm}^{n,N}(u)\big] &\leq K\int_0^{t\wedge u} d_{jk}^{n,N}(t,v)\,d_{jk}^{n,N}(u,v)\,\mathrm{d}v \int_0^v d_{lm}^{n,N}(v,\vartheta)^2\,\mathrm{d}\vartheta \\
\mathbb{E}\big[\breve{Z}_{jk,lm}^{n,N}(t)\breve{Z}_{jk,lm}^{n,N}(u)\big] &\leq K\int_0^{t\wedge u} d_{jk}^{n,N}(t,v)\,d_{jk}^{n,N}(u,v)\,\mathrm{d}v \int_0^v d_{lm}^{n,N}(\vartheta,v)^2\,\mathrm{d}\vartheta \\
\mathbb{E}\big[\mathring{Z}_{jk,lm}^{n,N}(t)\mathring{Z}_{jk,lm}^{n,N}(u)\big] &\leq K\int_0^{t\wedge u} d_{lm}^{n,N}(v,t)\,d_{lm}^{n,N}(v,u)\,\mathrm{d}v \int_0^v d_{jk}^{n,N}(v,\vartheta)^2\,\mathrm{d}\vartheta \\
\mathbb{E}\big[\widetilde{Z}_{jk,lm}^{n,N}(t)\widetilde{Z}_{jk,lm}^{n,N}(u)\big] &\leq K\int_0^{t\wedge u} d_{jk}^{n,N}(v,t)\,d_{jk}^{n,N}(v,u)\,\mathrm{d}v \int_0^v d_{lm}^{n,N}(\vartheta,v)^2\,\mathrm{d}\vartheta.
\end{aligned}
$$

Proof. According to Itô's formula, (6.58), (6.59),

$$
\begin{aligned}
\mathbb{E}\big[Z_{jk,lm}^{n,N}(t)Z_{jk,lm}^{n,N}(u)\big] &= \mathbb{E}\int_0^{t\wedge u} d_{jk}^{n,N}(t,v)\,d_{jk}^{n,N}(u,v)\,c_{kk}(v)\,U_{lm}^{n,N}(v)^2\,\mathrm{d}v \\
&\leq K\int_0^{t\wedge u} d_{jk}^{n,N}(t,v)\,d_{jk}^{n,N}(u,v)\,\mathbb{E}\big[U_{lm}^{n,N}(v)^2\big]\,\mathrm{d}v,
\end{aligned}
$$

similarly,

$$
\begin{aligned}
\mathbb{E}\big[\breve{Z}^{n,N}_{jk,lm}(t)\breve{Z}^{n,N}_{jk,lm}(u)\big] &\leq K\int_0^{t\wedge u} d^{n,N}_{jk}(t,v)\, d^{n,N}_{jk}(u,v)\,\mathbb{E}\big[\widetilde{U}^{n,N}_{lm}(v)^2\big]\,\mathrm{d}v \\
\mathbb{E}\big[\mathring{Z}^{n,N}_{jk,lm}(t)\mathring{Z}^{n,N}_{jk,lm}(u)\big] &\leq K\int_0^{t\wedge u} d^{n,N}_{lm}(v,t)\, d^{n,N}_{lm}(v,u)\,\mathbb{E}\big[U^{n,N}_{jk}(v)^2\big]\,\mathrm{d}v \\
\mathbb{E}\big[\widetilde{Z}^{n,N}_{jk,lm}(t)\widetilde{Z}^{n,N}_{jk,lm}(u)\big] &\leq K\int_0^{t\wedge u} d^{n,N}_{jk}(v,t)\, d^{n,N}_{jk}(v,u)\,\mathbb{E}\big[\widetilde{U}^{n,N}_{lm}(v)^2\big]\,\mathrm{d}v,
\end{aligned}
$$

then this lemma follows from (6.52). □

6.6.2 stable convergence

By (6.47), (6.48), it suffices to study

$$
\Psi^{n,N,M}_T := \sum_{j,k=1}^{d}\Big[e(0)^{n,N,M}_{jk,T} + e(1)^{n,N,M}_{jk,T}\Big],
$$

because of **Lemma 38**, it remains to study the limit of the angle bracket $\big\langle \Psi^{n,N,M}, \Psi^{n,N,M}\big\rangle_T$ in probability.

$$
\begin{aligned}
\big\langle \Psi^{n,N,M}, \Psi^{n,N,M}\big\rangle_T = \sum_{j,k,l,m=1}^{d}\Big[&\big\langle e(0)^{n,N,M}_{jk}, e(0)^{n,N,M}_{lm}\big\rangle_T + \big\langle e(0)^{n,N,M}_{jk}, e(1)^{n,N,M}_{lm}\big\rangle_T \\
&+ \big\langle e(1)^{n,N,M}_{jk}, e(0)^{n,N,M}_{lm}\big\rangle_T + \big\langle e(1)^{n,N,M}_{jk}, e(0)^{n,N,M}_{lm}\big\rangle_T\Big],
\end{aligned}
$$

and by (6.56),

$$\begin{aligned}
\big\langle e(0)^{n,N,M}_{jk}, e(0)^{n,N,M}_{lm}\big\rangle_T &= N\int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\,\widehat{\rho}^M_{lm}(\theta^n_l(t))\,c_{jl}(t)\,U^{n,N}_{jk}(t)\,U^{n,N}_{lm}(t)\,\mathrm{d}t\\
\big\langle e(0)^{n,N,M}_{jk}, e(1)^{n,N,M}_{lm}\big\rangle_T &= N\int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\,c_{jm}(t)\,U^{n,N}_{jk}(t)\,\widetilde{U}^{n,N}_{lm}(t)\,\mathrm{d}t\\
\big\langle e(1)^{n,N,M}_{jk}, e(0)^{n,N,M}_{lm}\big\rangle_T &= N\int_0^T \widehat{\rho}^M_{lm}(\theta^n_l(t))\,c_{kl}(t)\,\widetilde{U}^{n,N}_{jk}(t)\,U^{n,N}_{lm}(t)\,\mathrm{d}t\\
\big\langle e(1)^{n,N,M}_{jk}, e(1)^{n,N,M}_{lm}\big\rangle_T &= N\int_0^T c_{km}(t)\,\widetilde{U}^{n,N}_{jk}(t)\,\widetilde{U}^{n,N}_{lm}(t)\,\mathrm{d}t,
\end{aligned}$$

so by (6.59),

$$\big\langle \Psi^{n,N,M}, \Psi^{n,N,M}\big\rangle_T = \sum_{j,k,l,m=1}^{d}\left[\sum_{r=0}^{3} O(r)^{n,N,M}_{jk,lm,T} + \sum_{r=0}^{3} V(r)^{n,N,M}_{jk,lm,T}\right], \tag{6.60}$$

where

$$\begin{aligned}
O(0)^{n,N,M}_{jk,lm,T} &= N\int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\,\widehat{\rho}^M_{lm}(\theta^n_l(t))\,c_{jl}(t)\,\big[Z^{n,M}_{jk,lm}(t) + Z^{n,M}_{lm,jk}(t)\big]\,\mathrm{d}t\\
O(1)^{n,N,M}_{jk,lm,T} &= N\int_0^T \Big[\widehat{\rho}^M_{jk}(\theta^n_j(t))\,c_{jm}(t)\,\breve{Z}^{n,N}_{jk,lm}(t) + \widehat{\rho}^M_{lm}(\theta^n_l(t))\,c_{kl}(t)\,\breve{Z}^{n,N}_{lm,jk}(t)\Big]\,\mathrm{d}t\\
O(2)^{n,N,M}_{jk,lm,T} &= N\int_0^T \Big[\widehat{\rho}^M_{jk}(\theta^n_j(t))\,c_{jm}(t)\,\mathring{Z}^{n,N}_{jk,lm}(t) + \widehat{\rho}^M_{lm}(\theta^n_l(t))\,c_{kl}(t)\,\mathring{Z}^{n,N}_{lm,jk}(t)\Big]\,\mathrm{d}t\\
O(3)^{n,N,M}_{jk,lm,T} &= N\int_0^T c_{km}(t)\,\big[\widetilde{Z}^{n,M}_{jk,lm}(t) + \widetilde{Z}^{n,M}_{lm,jk}(t)\big]\,\mathrm{d}t,
\end{aligned}$$

and

$$
\begin{aligned}
V(0)^{n,N,M}_{jk,lm,T} &= \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, \widehat{\rho}^M_{lm}(\theta^n_l(t))\, c_{jl}(t)\, \mathrm{d}t \Big[N \int_0^t d^{n,N}_{jk}(t,u)\, d^{n,N}_{lm}(t,u)\, c_{km}(u)\, \mathrm{d}u\Big] \\
V(1)^{n,N,M}_{jk,lm,T} &= \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, c_{jm}(t)\, \mathrm{d}t \Big[N \int_0^t d^{n,N}_{jk}(t,u)\, d^{n,N}_{lm}(u,t)\, \widehat{\rho}^M_{lm}(\theta^n_l(u))\, c_{kl}(u)\, \mathrm{d}u\Big] \\
V(2)^{n,N,M}_{jk,lm,T} &= \int_0^T \widehat{\rho}^M_{lm}(\theta^n_l(t))\, c_{kl}(t)\, \mathrm{d}t \Big[N \int_0^t d^{n,N}_{jk}(u,t)\, d^{n,N}_{lm}(t,u)\, \widehat{\rho}^M_{jk}(\theta^n_j(u))\, c_{jm}(u)\, \mathrm{d}u\Big] \\
V(3)^{n,N,M}_{jk,lm,T} &= \int_0^T c_{km}(t)\, \mathrm{d}t \Big[N \int_0^t d^{n,N}_{jk}(u,t)\, d^{n,N}_{lm}(u,t)\, \widehat{\rho}^M_{jk}(\theta^n_j(u))\, \widehat{\rho}^M_{lm}(\theta^n_l(u))\, c_{jl}(u)\, \mathrm{d}u\Big].
\end{aligned}
$$

To show the asymptotic negligibility of $O(r)^{n,N,M}_{jk,lm,T}$ for $r = 0, \cdots, 3$, by symmetry, it suffices to study the following terms:

$$
\begin{aligned}
\phi(0)^{n,N,M}_{jk,lm,T} &\coloneqq N \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, \widehat{\rho}^M_{lm}(\theta^n_l(t))\, c_{jl}(t)\, Z^{n,N}_{jk,lm}(t)\, \mathrm{d}t \\
\phi(1)^{n,N,M}_{jk,lm,T} &\coloneqq N \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, c_{jm}(t)\, \breve{Z}^{n,N}_{jk,lm}(t)\, \mathrm{d}t \\
\phi(2)^{n,N,M}_{jk,lm,T} &\coloneqq N \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, c_{jm}(t)\, \mathring{Z}^{n,N}_{jk,lm}(t)\, \mathrm{d}t \\
\phi(3)^{n,N,M}_{jk,lm,T} &\coloneqq N \int_0^T c_{km}(t)\, \widetilde{Z}^{n,N}_{jk,lm}(t)\, \mathrm{d}t.
\end{aligned}
$$

Note

$$
\begin{aligned}
\big|\phi(0)^{n,N,M}_{jk,lm,T}\big|^2 &= N^2 \int_0^T \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, \widehat{\rho}^M_{lm}(\theta^n_l(t))\, \widehat{\rho}^M_{jk}(\theta^n_j(u))\, \widehat{\rho}^M_{lm}(\theta^n_l(u)) \\
&\qquad\qquad c_{jl}(t)\, c_{jl}(u)\, Z^{n,N}_{jk,lm}(t)\, Z^{n,N}_{jk,lm}(u)\, \mathrm{d}t\, \mathrm{d}u \\
\big|\phi(1)^{n,N,M}_{jk,lm,T}\big|^2 &= N^2 \int_0^T \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, \widehat{\rho}^M_{jk}(\theta^n_j(u))\, c_{jm}(t)\, c_{jm}(u)\, \breve{Z}^{n,N}_{jk,lm}(t)\, \breve{Z}^{n,N}_{jk,lm}(u)\, \mathrm{d}t\, \mathrm{d}u \\
\big|\phi(2)^{n,N,M}_{jk,lm,T}\big|^2 &= N^2 \int_0^T \int_0^T \widehat{\rho}^M_{jk}(\theta^n_j(t))\, \widehat{\rho}^M_{jk}(\theta^n_j(u))\, c_{jm}(t)\, c_{jm}(u)\, \mathring{Z}^{n,N}_{jk,lm}(t)\, \mathring{Z}^{n,N}_{jk,lm}(u)\, \mathrm{d}t\, \mathrm{d}u \\
\big|\phi(3)^{n,N,M}_{jk,lm,T}\big|^2 &= N^2 \int_0^T \int_0^T c_{km}(t)\, c_{km}(u)\, \widetilde{Z}^{n,N}_{jk,lm}(t)\, \widetilde{Z}^{n,N}_{jk,lm}(u)\, \mathrm{d}t\, \mathrm{d}u,
\end{aligned}
$$

by **Lemma 40**,

$$\mathbb{E}\big(\big|\phi(0)^{n,N,M}_{jk,lm,T}\big|^2\big) \\ \leq KN^2 \int_0^T \int_0^T \mathrm{d}t\,\mathrm{d}u \Big[\int_0^{t\wedge u} d^{n,N}_{jk}(t,v)\, d^{n,N}_{jk}(u,v)\,\mathrm{d}v \int_0^v d^{n,N}_{lm}(v,\vartheta)^2\,\mathrm{d}\vartheta\Big],$$

by (6.25),

$$\mathbb{E}\big(\big|O(0)^{n,N,M}_{jk,lm,T}\big|^2\big) \leq KT^{\frac{3p-2}{p}} N^{1-2/p}. \tag{6.61}$$

By similar arguments, we can show the same upper bound applies to $\mathbb{E}\big(\big|\phi(1)^{n,N,M}_{jk,lm,T}\big|^2\big)$, $\mathbb{E}\big(\big|\phi(2)^{n,N,M}_{jk,lm,T}\big|^2\big)$, $\mathbb{E}\big(\big|\phi(3)^{n,N,M}_{jk,lm,T}\big|^2\big)$. Thus by Jensen's inequality and Markov's inequality, we can prove that $\phi(r)^{n,N,M}_{jk,lm,T}$, $r = 0,1,2,3$ all converge to 0 in probability.

By **Lemma 31** and **Lemma 33**,

$$\begin{aligned}
V(0)^{n,N,M}_{jk,lm,T} &\xrightarrow{\mathbb{P}} \int_0^T \rho_{jk}(t)\,\rho_{lm}(t)\,\tilde{\theta}_{jk,lm}(t)\,c_{jl}(t)\,c_{km}(t)\,\mathrm{d}t \\
V(1)^{n,N,M}_{jk,lm,T} &\xrightarrow{\mathbb{P}} \int_0^T \rho_{jk}(t)\,\rho_{lm}(t)\,\acute{\theta}_{jk,lm}(t)\,c_{jm}(t)\,c_{kl}(t)\,\mathrm{d}t \\
V(2)^{n,N,M}_{jk,lm,T} &\xrightarrow{\mathbb{P}} \int_0^T \rho_{jk}(t)\,\rho_{lm}(t)\,\check{\theta}_{jk,lm}(t)\,c_{jm}(t)\,c_{kl}(t)\,\mathrm{d}t \\
V(3)^{n,N,M}_{jk,lm,T} &\xrightarrow{\mathbb{P}} \int_0^T \rho_{jk}(t)\,\rho_{lm}(t)\,\grave{\theta}_{jk,lm}(t)\,c_{jl}(t)\,c_{km}(t)\,\mathrm{d}t.
\end{aligned}$$

Therefore, we have the following proposition.

Proposition 6. *Assume* ***Assumption A**-ν, **A**-X, **A**-α. Let $N \to \infty$, $N\Delta(n)^{2/3} \to 0$, $M \to \infty$, $M < N$,*

$$N^{1/2} \sum_{j,k=1}^{d} \int_0^T \rho_{jk}(t)\big[\overline{\widehat{c}^{n,N,M}_{jk}(t)} - c_{jk}(t)\big]\,\mathrm{d}t \xrightarrow{\mathcal{L}-s} \mathcal{MN}\big(0, V_T\big),$$

where

$$V_T = \sum_{j,k,l,m=1}^{d} \int_0^T \rho_{jk}(t)\,\rho_{lm}(t) \times$$
$$\Big\{ \big[\tilde{\theta}_{jk,lm}(t) + \grave{\theta}_{jk,lm}(t)\big]\, c_{jl}(t)\, c_{km}(t) + \big[\acute{\theta}_{jk,lm}(t) + \check{\theta}_{jk,lm}(t)\big]\, c_{jm}(t)\, c_{kl}(t) \Big\}\, \mathrm{d}t.$$

CHAPTER 7

INFERENCE OF VOLATILITY MATRIX FUNCTIONALS BY FOURIER-MALLIAVIN METHOD

7.1 Functional estimation given missing data

Given a functional of statistical interest and a nonparametric estimator of spot volatility, we can construct a functional estimator via the plug-in framework of Jacod and Rosenbaum [2013]. In this framework, computing a functional estimator entails (i) computing the nonparametric estimates of spot volatility at various time points; (ii) plugging the nonparametric estimates into the functional and computing the Riemann sum.

To cope with asynchronicity and generalize the framework of Jacod and Rosenbaum [2013], the author chooses the Fourier-Malliavin method to compute nonparametric estimates of spot volatility. The Fourier-Malliavin method for volatility functional estimation comprises of three steps:

1. estimate the Fourier coefficients of volatility by DFT and Bohr convolution;
2. estimate the spot volatility from the estimates of its Fourier coefficients by IDFT and Fejér kernel;
3. plug in the estimates of spot volatility and evaluate the functionals.

The first two steps have been described in Chapter 6. For the third step, the volatility functional estimator is defined as

$$\widehat{S}(g)^n_T \equiv \sum_{h=1+L}^{B-L} g\Big(\widehat{c}^{n,N,M}\left(\frac{hT}{B}\right)\Big)\frac{T}{B}. \tag{7.1}$$

There are four tuning parameters, namely N, M, B, L. The tuning parameters N and M are inherited from the spot estimator $\widehat{c}^{n,N,M}$. The tuning parameters B and L dictate how to construct the functional estimators:

- B is the number of plug-ins in the Riemann sum; a higher B results in a more accurate approximation to the integral, with the cost of higher computational load;
- L is the bandwidth at the boundaries of the time window, in which no spot estimate will be taken in the Riemann sum.

The boundary values of a volatility sample path $c(0)$ and $c(T)$ are different in general. However, the spot estimator (6.7) is based on trigonometric series and is periodic by construction, hence $\widehat{c}^{n,N,M}(0) = \widehat{c}^{n,N,M}(T)$. Because of this artifact, no spot estimate near the boundaries will be used in the functional estimator (7.1).

It is required that the tuning parameters N and M satisfy

$$\begin{cases} N & \to \infty \\ N & \leq \lfloor \underline{n}/2 \rfloor - M + 1 \\ M^{4\alpha}/N & \to \infty \\ M^2/N & \to 0, \end{cases} \tag{7.2}$$

where α is specified in assumption A-α, and B and L satisfy

$$\begin{cases} B/N^{1/2} \to \infty & \\ L = 0 & \text{if } c(0) = c(T) \\ L \asymp B/M & \text{if } c(0) \neq c(T). \end{cases} \tag{7.3}$$

The spot volatility is a crucial element in volatility functional estimation. For a given functional, the large sample properties of the functional estimator largely relies on the asymptotics of the nonparametric estimator of spot volatility.

The author have shown the analyses of Fourier coefficient estimation and volatility estimation in Chapter 6. In this chapter, the author is going to show the analysis of volatility functional estimation based on the Fourier-Malliavin method.

7.2 Central limit theorems

7.2.1 decomposition and asymptotic negligibility

We can write

$$N^{1/2}\big[\widehat{S}(g)_T^n - S(g)_T\big] = \overline{S}(0)_T^{n,N,M,B} + \overline{S}(1)_T^{n,N,M} + \overline{S}(2)_T^{n,N,M}, \tag{7.4}$$

where

$$\begin{aligned}
\overline{S}(0)_T^{n,N,M,B} &:= N^{1/2}\left[\sum_{h=1}^{B} g\big(\widehat{c}^{n,N,M}(hT/B)\big)\,T/B - \int_0^T g\big(\widehat{c}^{n,N,M}(t)\big)\,\mathrm{d}t\right] \\
\overline{S}(1)_T^{n,N,M} &:= N^{1/2}\int_0^T \Big\{g\big(\widehat{c}^{n,N,M}(t)\big) - g(c(t)) \\
&\qquad\qquad - \sum_{j,k=1}^{d} \partial_{jk} g(c(t))\Big[\widehat{c}_{jk}^{n,N,M}(t) - c_{jk}(t)\Big]\Big\}\,\mathrm{d}t \\
\overline{S}(2)_T^{n,N,M} &:= \sum_{j,k=1}^{d} N^{1/2}\int_0^T \partial_{jk} g(c(t))\Big[\widehat{c}_{jk}^{n,N,M}(t) - c_{jk}(t)\Big]\,\mathrm{d}t.
\end{aligned}$$

By Riemann summation,

$$\big\|\overline{S}(0)_T^{n,N,M,B}\big\| = O_p(N^{1/2}/B),$$

in view of (7.3),

$$\left\|\overline{S}(0)_T^{n,N,M,B}\right\| \xrightarrow{\mathbb{P}} 0. \tag{7.5}$$

By (4.1),

$$\left\|g\big(\widehat{c}^{n,N,M}(t)\big) - g(c(t)) - \sum_{j,k} \partial_{jk} g(c(t))\big[\widehat{c}_{jk}^{n,N,M}(t) - c_{jk}(t)\big]\right\| = O_p\big(\|\widehat{c}^{n,N,M}(t) - c(t)\|^2\big),$$

therefore

$$\left\|\overline{S}(1)_T^{n,N,M}\right\| \leq KTN^{1/2} \sup_{t\in[0,T]} \left\|\widehat{c}^{n,N,M}(t) - c(t)\right\|^2.$$

According to **Proposition 5**,

$$\mathbb{E}\left\|\overline{S}(1)_T^{n,N,M}\right\| \leq KT\frac{M}{N^{1/2}}.$$

Thus by (7.2) and Markov's inequality,

$$\left\|\overline{S}(1)_T^{n,N,M}\right\| \xrightarrow{\mathbb{P}} 0. \tag{7.6}$$

7.2.2 stable convergence

Let $\rho_{jk}(t) = \partial_{jk} g(c(t))$, we get

$$\overline{S}(2)_T^{n,N,M} = \sum_{j,k=1}^{d} N^{1/2} \int_0^T \rho_{jk}(t)\left[\overline{\widehat{c}_{jk}^{n,N,M}(t)} - c_{jk}(t)\right] \mathrm{d}t.$$

In view of (7.4), (7.5), (7.6), and **Proposition 6**, the author derives the following theorem.

Theorem 5. *Assume* ***Assumption A-***ν*,* ***A-***X*,* ***A-***d^2*, and* ***Assumption A-***α *with* $\alpha >$

$1/2$, $c(0) = c(T)$ and (4.1), and choose the tuning parameters according to (7.2), (7.3) with

- *$N \leq \lfloor \underline{n}/2 \rfloor - M + 1$ if **Assumption A-**τ holds,*
- *$N = o(\underline{n}^{4/5})$ if **Assumption A-**τ does not hold,*

then we have

$$N^{1/2}\big[\widehat{S}(g)^n_T - S(g)_T\big] \xrightarrow{\mathcal{L}-s} \mathcal{MN}\big(0, V(g)_T\big),$$

where

$$V(g)_T = \sum_{j,k,l,m=1}^{d} \int_0^T \partial_{jk} g(c(t))\, \partial_{lm} g(c(t)) \times \\ \Big\{\big[\tilde{\theta}_{jk,lm}(t) + \grave{\theta}_{jk,lm}(t)\big] c_{jl}(t)\, c_{km}(t) + \big[\acute{\theta}_{jk,lm}(t) + \check{\theta}_{jk,lm}(t)\big] c_{jm}(t)\, c_{kl}(t)\Big\}\, \mathrm{d}t. \quad (7.7)$$

We have the following corollary which immediately follows from **Theorem 5** and **Lemma 32**. It states that when different time series are observed synchronously, the functional estimator based on the Fourier-Malliavin method can be rate optimal and efficient.

Corollary 4. *Assume **Assumption A-**τ, **A-**ν, **A-**X, and **Assumption A-**α with $\alpha > 1/2$, $c(0) = c(T)$ and (4.1). Choose the tuning parameters according to (7.2), (7.3) with $N = \lfloor n/2 \rfloor - M + 1$, we have*

$$\Delta(n)^{-1/2}\big[\widehat{S}(g)^n_T - S(g)_T\big] \xrightarrow{\mathcal{L}-s} \mathcal{MN}\big(0, V(g)^*_T\big),$$

where

$$V(g)^*_T = \sum_{j,k,l,m=1}^{d} \int_0^T \partial_{jk} g(c(t))\, \partial_{lm} g(c(t)) \times \big[c_{jl}(t)\, c_{km}(t) + c_{jm}(t)\, c_{kl}(t)\big]\, \mathrm{d}t.$$

The convergence rate $\Delta(n)^{-1/2} \asymp n^{1/2}$ is optimal and the asymptotic variance $V(g)^*_T$

achieves the efficiency bound, cf. Jacod and Rosenbaum [2013] and Clément et al. [2013].

The author provides an estimator of the asymptotic variance (7.7), which is defined as

$$\widehat{V}(g)^n_T = \sum_{j,k,l,m=1}^{d} \left[\widehat{V}(0)^{n,N,M,B}_{jk,lm,T} + \widehat{V}(1)^{n,N,M,B}_{jk,lm,T}\right], \tag{7.8}$$

where

$$\begin{aligned}
\widehat{V}(0)^{n,N,M,B}_{jk,lm,T} &= \frac{T}{B}\sum_{h=1}^{B} \partial_{jk}g(\widehat{c}^{n,N,M}(t_h))\,\partial_{lm}g(\widehat{c}^{n,N,M}(t_h))\,\widehat{c}^{n,N,M}_{jl}(t_h)\,\widehat{c}^{n,N,M}_{km}(t_h) \\
&\quad\times N\delta(n)\sum_{v=1}^{\lfloor t_h/\delta(n)\rfloor}\left[d^{n,N}_{jk}(t_h,\vartheta_v)\,d^{n,N}_{lm}(t_h,\vartheta_v) + d^{n,N}_{jk}(\vartheta_v,t_h)\,d^{n,N}_{lm}(\vartheta_v,t_h)\right] \\
\widehat{V}(1)^{n,N,M,B}_{jk,lm,T} &= \frac{T}{B}\sum_{h=1}^{B} \partial_{jk}g(\widehat{c}^{n,N,M}(t_h))\,\partial_{lm}g(\widehat{c}^{n,N,M}(t_h))\,\widehat{c}^{n,N,M}_{jm}(t_h)\,\widehat{c}^{n,N,M}_{kl}(t_h) \\
&\quad\times N\delta(n)\sum_{v=1}^{\lfloor t_h/\delta(n)\rfloor}\left[d^{n,N}_{jk}(t_h,\vartheta_v)\,d^{n,N}_{lm}(\vartheta_v,t_h) + d^{n,N}_{jk}(\vartheta_v,t_h)\,d^{n,N}_{lm}(t_h,\vartheta_v)\right],
\end{aligned}$$

and $t_h = hT/B$ with B satisfying (7.3), $\vartheta_v = v\delta(n)$, $\delta(n) = \min_j \min_h \Delta^j_h$.

According to (4.1), **Proposition 5**, and the choices of t_h and ϑ_v, it immediately follows that under the conditions of **Theorem 5**,

$$\widehat{V}(g)^n_T \xrightarrow{\mathbb{P}} V(g)_T.$$

7.2.3 interference phenomenon

In the presence of asynchronous observations, the condition $N = o(\underline{n}^{4/5})$ in **Theorem 5** means the convergence rate is strictly less than $\underline{n}^{2/5}$. If we allow the limit distribution to be non-centered, we can improve the convergence rate to be exact $\underline{n}^{2/5}$.

To formulate this non-centered asymptotic result, we define "*cubic variation of time*" as $P^n_{jk}(t) := \underline{n}^2 \int_0^t \left[\theta^n_j(u) - \theta^n_k(u)\right]^2 \mathrm{d}u$, note

$$\begin{aligned}
P^n_{jk}(t) = \underline{n}^2 \sum_{I^j_h \cap I^k_v \neq \emptyset} \Big[& \big(\tau^j_h \wedge \tau^k_v - \tau^j_{h-1} \vee \tau^k_{v-1}\big)^2 \, |\tau^j_{h-1} - \tau^k_{v-1}| \, \mathbb{1}_{\{\tau^j_{h-1} \wedge \tau^k_{v-1} \leq t;\, I^j_h \not\subseteq I^k_v \text{ and } I^k_v \not\subseteq I^j_h\}} \\
& + \big(\tau^j_h \vee \tau^k_v - \tau^j_{h-1} \vee \tau^k_{v-1}\big)^2 \, |\tau^j_{h-1} - \tau^k_{v-1}| \, \mathbb{1}_{\{\tau^j_{h-1} \wedge \tau^k_{v-1} \leq t;\, I^j_h \subseteq I^k_v \text{ or } I^k_v \subseteq I^j_h\}} \\
& \qquad + \big(\tau^j_h \wedge \tau^k_v - \tau^j_{h-1} \vee \tau^k_{v-1}\big) \, |\tau^j_h - \tau^k_v|^2 \, \mathbb{1}_{\{\tau^j_{h-1} \vee \tau^k_{v-1} \leq t\}} \Big],
\end{aligned}$$

and under **Assumption A-**τ, $P^n_{jk}(t) = 0$ uniformly.

Assumption A-θ (cubic variation of time)**.** *$\forall j, k = 1, \cdots, d$, $\exists$ an integrable function ϱ_{jk}, such that $\forall t \in [0, T]$, as $n \to \infty$*

$$P^n_{jk}(t) \xrightarrow{\mathbb{P}} \int_0^t \varrho_{jk}(u) \, \mathrm{d}u.$$

The next proposition states a limit result with exact rate $\underline{n}^{2/5}$, and the cubic variation of time emerges as the bias in the asymptotic distribution.

Proposition 7. *Assume **Assumption A-**X, **A-**ν, **A-**τ, **A-**d^2, **A-**θ, **A-**α with $\alpha > 1/2$, $c(0) = c(T)$ and (4.1). Choose the tuning parameters according to (7.2), (7.3) with $N = \lfloor \kappa \underline{n}^{4/5} \rfloor \wedge (\lfloor \underline{n}/2 \rfloor - M + 1)$, we have*

$$\underline{n}^{2/5} \big[\widehat{S}(g)^n_T - S(g)_T\big] \xrightarrow{\mathcal{L}-s} \mathcal{MN}\big(\mu(g)_T, V(g)_T\big),$$

where $V(g)_T$ is defined as (7.7) and

$$\mu(g)_T = -\frac{2\pi^2 \kappa^{5/2}}{3T^2} \sum_{j,k=1}^d \int_0^T \partial_{jk} g\big(c(t)\big) \, c_{jk}(t) \, \varrho_{jk}(t) \, \mathrm{d}t,$$

*with ϱ_{jk} being defined in **Assumption A-**θ.*

The bias in the second order can be estimated by

$$\widehat{\mu}(g)_T = -\frac{2\pi^2\kappa^{5/2}}{3T^2}\sum_{j,k=1}^{d}\sum_{h=1}^{B}\partial_{jk}g\big(\widehat{c}^{n,N,M}(t_h)\big)\,\widehat{c}_{jk}^{n,N,M}(t_h)\,\big[P_{jk}^n(t_h)-P_{jk}^n(t_{h-1})\big].$$

In the asynchronous scenario, if $N=\lfloor \underline{n}/2\rfloor - M+1$, the functional estimator (7.1) generally is no longer consistent. However, there is still an asymptotic result with optimal convergence rate and a new limit. Define

$$\begin{aligned}
\underline{c}^{n,N}(t) &:= \big[d_{jk}^{n,N}(t,t)\,c_{jk}(t)\big]_{jk}\\
\underline{S}(g)_T^{n,N} &:= \int_0^T g\big(\underline{c}^{n,N}(t)\big)\,\mathrm{d}t,
\end{aligned}$$

we can write

$$N^{1/2}\big[\widehat{S}(g)_T^n - \underline{S}(g)_T^{n,N}\big] = \overline{S}(0)_T^{n,N,M,B} + \overline{\underline{S}}(1)_T^{n,N,M} + \overline{\underline{S}}(2)_T^{n,N,M}, \tag{7.9}$$

where $\overline{S}(0)_T^{n,N,M,B}$ is defined in (7.4) and

$$\begin{aligned}
\overline{\underline{S}}(1)_T^{n,N,M} &:= N^{1/2}\int_0^T \Big\{g\big(\widehat{c}^{n,N,M}(t)\big) - g\big(\underline{c}^{n,N}(t)\big)\\
&\qquad\qquad - \sum_{j,k=1}^{d}\partial_{jk}g\big(\underline{c}^{n,N}(t)\big)\Big[\widehat{c}_{jk}^{n,N,M}(t) - \underline{c}_{jk}^{n,N}(t)\Big]\Big\}\,\mathrm{d}t\\
\overline{\underline{S}}(2)_T^{n,N,M} &:= \sum_{j,k=1}^{d} N^{1/2}\int_0^T \partial_{jk}g\big(\underline{c}^{n,N}(t)\big)\Big[\widehat{c}_{jk}^{n,N,M}(t) - d_{jk}^{n,N}(t,t)\,c_{jk}(t)\Big]\,\mathrm{d}t.
\end{aligned}$$

By (4.1),

$$\Big\| g\big(\widehat{c}^{n,N,M}(t)\big) - g\big(\underline{c}(t)\big) - \sum_{j,k} \partial_{jk} g\big(\underline{c}^{n,N}(t)\big)\Big[\widehat{c}_{jk}^{n,N,M}(t) - \underline{c}_{jk}^{n,N}(t)\Big]\Big\| \\ = O_p\big(\big\|\widehat{c}^{n,N,M}(t) - \underline{c}^{n,N}(t)\big\|^2\big),$$

therefore

$$\big\|\underline{\overline{S}}(1)_T^{n,N,M}\big\| \leq KTN^{1/2} \sup_{t\in[0,T]} \big\|\widehat{c}^{n,N,M}(t) - \underline{c}^{n,N}(t)\big\|^2.$$

Notice that

$$\widehat{c}_{jk}^{n,N,M}(t) - \underline{c}_{jk}(t) = Q(t,0)_{jk,T}^{n,N,M} + Q(t,1)_{jk,T}^{n,N,M} + \underline{\Omega}(t,1)_{jk,T}^{n,N,M} + \underline{\Omega}(t,2)_{jk,T}^{N,M},$$

where $Q(t,0)_{jk,T}^{n,N,M}$ and $Q(t,1)_{jk,T}^{n,N,M}$ are defined by (6.53) and

$$\begin{aligned}
\underline{\Omega}(t,1)_{jk,T}^{n,N,M} &= \frac{1}{T}\int_0^T \Big[F^M\Big(\frac{t-\theta_j^n(u)}{T}\Big) - F^M\Big(\frac{t-u}{T}\Big)\Big]\, d_{jk}^{n,N}(u,u)\, c_{jk}(u)\, \mathrm{d}u \\
\underline{\Omega}(t,2)_{jk,T}^{N,M} &= \frac{1}{T}\int_0^T F^M\Big(\frac{t-u}{T}\Big)\, d_{jk}^{n,N}(u,u)\, c_{jk}(u)\, \mathrm{d}u - d_{jk}^{n,N}(t,t)\, c_{jk}(t),
\end{aligned}$$

by a similar argument preceding **Proposition 5**,

$$\mathbb{E}\big\|\underline{\overline{S}}(1)_T^{n,N,M}\big\| \leq KT\frac{M}{N^{1/2}}.$$

Thus by (7.2) and Markov's inequality,

$$\big\|\underline{\overline{S}}(1)_T^{n,N,M}\big\| \xrightarrow{\mathbb{P}} 0. \tag{7.10}$$

Let $\underline{\rho}_{jk}(t) = \partial_{jk} g\big(\underline{c}^{n,N}(t)\big)$, then

$$\overline{\underline{S}}(2)_T^{n,N,M} = \sum_{j,k=1}^{d} N^{1/2} \int_0^T \underline{\rho}_{jk}(t) \Big[\overline{\widehat{c}_{jk}^{n,N,M}(t)} - d_{jk}^{n,N}(t,t)\, c_{jk}(t)\Big]\, \mathrm{d}t. \tag{7.11}$$

The analysis of $\overline{\underline{S}}(2)_T^{n,N,M}$ can be done analogously as that of **Proposition 6**. Before state the result, we need an additional assumption on the almost everywhere convergence of the Dirichlet kernel of time gaps.

Assumption A-d (Dirichlet kernels of time)**.** *$\forall j, k = 1, \cdots, d$, $\exists$ an integrable function r_{jk}, such that $\forall t \in [0, T]$, as $n \to \infty$*

$$\int_0^t d_{jk}^{n,\lfloor \underline{n}/2 \rfloor}\big(u, \theta_j^n(u)\big)\, \mathrm{d}u \xrightarrow{\mathbb{P}} \int_0^t r_{jk}(u)\, \mathrm{d}u.$$

If $N = \lfloor \underline{n}/2 \rfloor - M + 1$, under other conditions, $\widehat{S}(g)_T^n - \underline{S}(g)_T^{n,N}$ converges rate-optimally to a mixed normal distribution.

Proposition 8. *Assume **Assumption A**-X, **A**-ν, **A**-τ, **A**-d^2, **A**-d, **A**-α with $\alpha > 1/2$, $c(0) = c(T)$ and (4.1), and choose the tuning parameters according to (7.2), (7.3) with $N = \lfloor \underline{n}/2 \rfloor - M + 1$, we have*

$$\underline{n}^{1/2} \big[\widehat{S}(g)_T^n - \underline{S}(g)_T^{n,N}\big] \xrightarrow{\mathcal{L}-s} \mathcal{MN}\big(0, \underline{V}(g)_T\big),$$

where

$$\begin{aligned}\underline{V}(g)_T = 2 \sum_{j,k,l,m=1}^{d} \int_0^T & \partial_{jk} g\big(r \odot c(t)\big)\, \partial_{lm} g\big(r \odot c(t)\big) \times \\ & \Big\{ \big[\tilde{\theta}_{jk,lm}(t) + \grave{\theta}_{jk,lm}(t)\big] c_{jl}(t)\, c_{km}(t) + \big[\acute{\theta}_{jk,lm}(t) + \check{\theta}_{jk,lm}(t)\big] c_{jm}(t)\, c_{kl}(t) \Big\}\, \mathrm{d}t,\end{aligned}$$

$r(t) = [r_{jk}(t)]_{jk}$, and $\odot$ denotes the Hadamard product.

Remark 11. The biases in **Proposition 7**, **8** arise from asynchronicity. Note $\widehat{F}(\mathrm{d}X_j)^n_s$ defined in (6.5) can be regarded as a wave function, the multiplication term $\widehat{F}(\mathrm{d}X_j)^n_{q-s} \times \widehat{F}(\mathrm{d}X_k)^n_s$ in the spectrum estimator (6.6) can be interpreted as a "superposition" of two waves. When the observation times of the j-th and k-th components are asynchronous, the waves $\widehat{F}(\mathrm{d}X_j)^n_{q-s}$ and $\widehat{F}(\mathrm{d}X_k)^n_s$ are out of phase. This results in the scaled and shifted Dirichlet kernel $d^{n,N}_{jk}(t,t)$ and is the source of asynchronicity biases.

7.3 Simulation

The author adopts the following simulation model to illustrate the functional estimator based on Fourier-Malliavin method

$$\begin{cases} \mathrm{d}X(t) &= .03\,\mathrm{d}t + \sqrt{c(t)}\,\mathrm{d}W(t) \\ c(t) &= \widetilde{c}(t) - [\widetilde{c}(T) - \widetilde{c}(0)]\,t/T \\ \mathrm{d}\widetilde{c}(t) &= 6(.16 - \widetilde{c}(t))\,\mathrm{d}t + .5\sqrt{\widetilde{c}(t)}\,\mathrm{d}B(t), \end{cases}$$

where $\mathbb{E}[(W_{t+\Delta} - W_t)(B_{t+\Delta} - B_t)] = -.6\Delta$. Each simulation employs 23400×21 data points with $\Delta(n) = 1s$.

In the first simulation experiment, we simulate synchronous observations and compute estimators for functionals $g(c) = c^2$, $g(c) = c^{-1}$, $g(c) = \log(c)$ based on the realized variance and the Fourier methods. For these three functionals, the tuning parameters are given in Table 7.1.

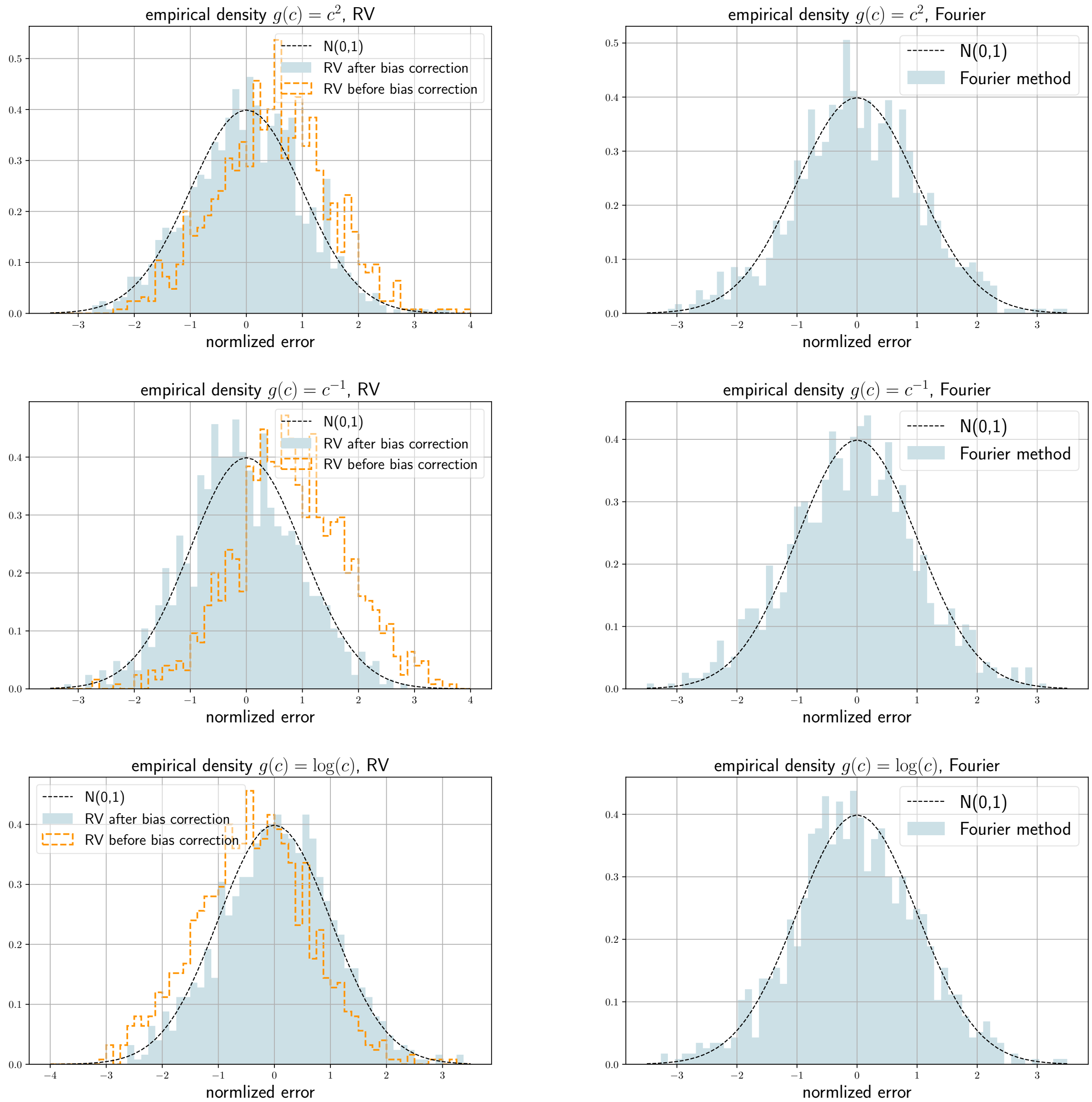
tuning parameters	k_n	N	M	B
values	$\lfloor \Delta_n^{-.45} \rfloor$	$\lfloor \underline{n}^{.75} \rfloor$	$\lfloor \underline{n}^{.3} \rfloor$	$\lfloor \underline{n}^{.8} \rfloor$

Table 7.1: Tuning parameters of the RV and Fourier-Malliavin method in simulation

The tuning parameter k_n for realized variance is chosen according to (3.6) in Jacod and Rosenbaum [2013].

The empirical densities of studentized estimators are shown in Figure 7.1.

Figure 7.1: Simulation of volatility functional estimation by realized variances & the Fourier-Malliavin method



CHAPTER 8

SUMMARY

The author presented, in the preceding chapters, some statistical theories and methodologies for covariance matrix functional estimation under nonparametric Itô semimartingale models. The models allow for non-stationarity and time-varying covariance matrix. The theories and methodologies are applicable to a wide range of nonlinear smooth functionals using high-frequency observations. Their applications include principal component analysis, generalized method of moments and specification tests, time series linear regression, statistical uncertainty quantification et cetera.

The novelties in this dissertation are the dealings with microstructure noise and asynchronous observations (missing/incomplete data), while the previous literature has only tackled this estimation problem in the absence of noise and observational asynchronicity. For the first time in the study of volatility functional estimation, the author demonstrated that microstructure noise can be handle by the pre-averaging method, and asynchronous observations can be utilized by the Fourier-Malliavin method.

New statistical results in this dissertation include convergence rates, correction of bias caused by functional non-linearity, bias as a manifestation of wave interference due to asynchronicity, asymptotic mixed normality and optimality. The asymptotic bias can be derived by second-order expansion. After bias correction, the pre-averaging method can achieve the optimal convergence rate of volatility functional estimation, and the asymptotic variance can be estimated. The Fourier-Malliavin method inherits wave properties from Fourier transform. The asynchronicity effect goes into the estimator through wave interference, and when the multivariate data is observe synchronously, the Fourier-Malliavin method is not only rate-optimal but also achieves the efficiency bound.

Based on the pre-averaging method and its statistical guarantee, a large-scale empirical analysis is implemented on the millisecond transaction data from the `TAQ` database. As a proof of concept, this demonstrates how to apply the results in this dissertation to complex large time-dependent data.

Nonetheless, there are theoretical limitations to the current work. One major limitation is that the methodologies proposed here can handle noisy or asynchronous data, but are not sufficient when noise and asynchronicity are both present. A good candidate to overcome this difficulty is the Fourier-Malliavin method. An advantage of the Fourier-Malliavin method is that it can cope with noise in a straightforward manner. For white noise, its energy is persistent across the whole spectrum. A working idea is to apply low-pass filters to sieve out noise at high frequencies. How to design low-pass filters for accurate covariance matrix functional estimation is still an open problem. A shortcoming of the pre-averaging method is statistical efficiency. Modifying the smoothing kernel in the light of Karhunen-Loève expansion is a potential approach to boost the efficiency of covariance matrix estimation in the presence of noise. Interestingly, it is spectrally adaptive rather than temporally adaptive, but it has not been investigated in the context of volatility functional estimation.

The author has implemented principal component analysis based on the pre-averaging method on noisy high-frequency data (e.g. transaction data sampled every second). Principal component analysis based on the Fourier-Malliavin method can be applied to noiseless synchronous high-frequency data (e.g. data sampled every minute), and can also be used for analysis of asynchronous data (e.g. liquid stocks are sampled at higher frequencies, less liquid stocks at lower frequencies). However, canonical Fast Fourier Transform (FFT) can not be directly applied to data that is asynchronously sampled. Typical sample sizes of high frequency financial datasets are hundred thousands or millions, hence computing the Fourier transforms without FFT is unfortunately not scalable. The author hopes to embed the non-uniform FFT (NFFT) algorithms into the Fourier-Malliavin method in the near future.

BIBLIOGRAPHY

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *The Review of Financial Studies*, 18(2):351–416.

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160:160–175.

Aït-Sahalia, Y. and Xiu, D. (2017). Using principal component analysis to estimate a high dimensional factor model with high-frequency data. *Journal of Econometrics*, 201:384–399.

Aït-Sahalia, Y. and Xiu, D. (2019). Principal component analysis of high frequency data. *Journal of American Statistical Association*, 114(525):287–303.

Altmeyer, R. and Bibinger, M. (2015). Functional stable limit theorems for quasi-efficient spectral covolatility estimators. *Stochastic Processes and their applications*, 125.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing realized 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., and Shephard, N. (2009). Realized kernels in practice: trades and quotes. *Econometrics Journal*, 12:C1–C32.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and 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*, 162:149–169.

Bibinger, M., Hautsch, N., Malec, P., and Reiß, M. (2014). Estimating the quadratic covariation matrix from noisy observations: local method of moments and efficiency. *The Annals of Statistics*, 42(4):1312–1346.

Bibinger, M., Hautsch, N., Malec, P., and Reiß, M. (2019). Estimating the spot covariation of asset prices - statistical theory and empirical evidence. *Journal of Business & Economic Statistics*, 37(3):419–435.

Chen, D., Mykland, P. A., and Zhang, L. (2019). The five trolls under the bridge: principal component analysis with asynchronous and noisy high frequency data. *Journal of American Statistical Association, forthcoming*.

Christensen, K., Kinnebrock, S., and Podolskij, M. (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *Journal of Econometrics*.

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

Clément, E. and Gloter, A. (2011). Limit theorems in the fourier transform method for the estimation of multivariate volatility. *Stochastic Processes and their Applications*, 121:1097–1124.

Cuchiero, C. and Teichmann, J. (2015). Fourier transform methods for pathwise covariance estimation in the presence of jumps. *Stochastic Processes and their Applications*, 125:116–160.

Fan, J., Furger, A., and Xiu, D. (2016). Incorporating global industrial classification standard into portfolio allocation: a simple factor-based large covariance matrix estimator with high-frequency data. *Journal of Business & Economic Statistics*, 34(4):489–503.

Hautsch, N. and Podolskij, M. (2013). Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: theory, implementation, and empirical evidence. *Journal of Business & Economic Statistics*, 31(2):165–183.

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

Jacod, J., Podolskij, M., and Vetter, M. (2010). Limit theorems for moving averages of discretized processes plus noise. *The Annals of Statistics*, 38(3):1478–1545.

Jacod, J. and Protter, P. (2012). *Discretization of Processes.* Springer-Verlag, Berlin Heidelberg.

Jacod, J. and Rosenbaum, M. (2013). Quarticity and other functionals of volatility: efficient estimation. *The Annals of Statistics*, 41(3):1462–1484.

Jacod, J. and Shiryaev, A. (2003). *Limit Theorems for Stochastic Processes (2ed.).* Springer-Verlag, Berlin Heidelberg.

Jolliffe, I. T. (2002). *Principal Component Analysis (2ed.).* Springer-Verlag, New York.

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

Lewis, A. S. (1996). Derivatives of spectral functions. *Mathematics of Operations Research*, 21(3):576–588.

Lewis, A. S. and Sendov, H. S. (2001). Twice differentiable spectral functions. *SIAM Journal on Matrix Analysis and Applications*, 23(2):368–386.

Li, J., Liu, Y., and Xiu, D. (2019). Efficient estimation of integrated volatility functionals via multiscale jackknife. *The Annals of Statistics*, 47(1):156–176.

Li, J., Todorov, V., and Tauchen, G. (2016). Inference theory for volatility functional dependencies. *Journal of Econometrics*, 193:17–34.

Li, J., Todorov, V., and Tauchen, G. (2017). Adaptive estimation of continuous-time regression models using high-frequency data. *Journal of Econometrics*, 200:36–47.

Li, J. and Xiu, D. (2016). Generalized method of integrated moments for high-frequency data. *Econometrica*, 84(4):1613–1633.

Li, Y. and Mykland, P. A. (2007). Are volatility estimators robust with respect to modeling assumption? *Bernoulli*, 13(3):601–622.

Magnus, J. and Neudecker, H. (2007). *Matrix Differential Calculus with Applications in Statistics and Econometrics (3ed.)*. John Wiley & Sons Ltd, West Sussex.

Malliavin, P. and Mancino, M. E. (2002). Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics*, 6:49–61.

Malliavin, P. and Mancino, M. E. (2009). A fourier transform method for nonparametric estimation of multivariate volatility. *The Annals of Statistics*, 37(4):1983–2010.

Mykland, P. A. and Zhang, L. (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica*, 77(5):1403–1445.

Mykland, P. A. and Zhang, L. (2016). Between data cleaning and inference: pre-averaging and robust estimators of the efficient price. *Journal of Econometrics*, 194:242–262.

Podolskij, M. and Vetter, M. (2009a). Bipower-type estimation in a noisy diffusion setting. *Stochastic Processes and their Applications*, 119:2803–2831.

Podolskij, M. and Vetter, M. (2009b). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, 15(3):634–658.

Protter, P. (2005). *Stochastic Integration and Differential Equations (2ed.)*. Springer-Verlag, Berlin Heidelberg.

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

Shephard, N. and Xiu, D. (2017). Econometric analysis of multivariate realised qml: estimation of the covariation of equity prices under asynchronous trading. *Journal of Econometrics*, 201:19–42.

Todorov, V. and Tauchen, G. (2012). The realized laplace transform of volatility. *Econometrica*, 80(3):1105–1127.

Wang, C. D. and Mykland, P. A. (2014). The estimation of leverage effect with high-frequency data. *Journal of the American Statistical Association*, 109(505):197–215.

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. (2011). Estimating covariation: Epps effect, microstructure noise. *Journal of Econometrics*, 160:33–47.

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