INFERENCE FOR VOLATILITY FUNCTIONALS OF MULTIVARIATE ITÔ SEMIMARTINGALES OBSERVED WITH JUMP AND NOISE

By Richard Y. Chen*

University of Chicago

This paper presents the nonparametric inference for nonlinear volatility functionals of general multivariate Itô semimartingales, in high-frequency and noisy setting. Pre-averaging and truncation enable simultaneous handling of noise and jumps. Second-order expansion reveals explicit biases and a pathway to bias correction. Estimators based on this framework achieve the optimal convergence rate. A class of stable central limit theorems are attained with estimable asymptotic covariance matrices. This paper form a basis for infill asymptotic results of, for example, the realized Laplace transform, the realized principal component analysis, the continuous-time linear regression, and the generalized method of integrated moments, hence helps to extend the application scopes to more frequently sampled noisy data.

1. Introduction. This paper concerns statistical inference and applications for integrated volatility functionals from high-frequency data modeled by an Itô semimartingale observed with noise. The functionals have of the form

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

here t is finite, $g : \mathbb{R}^{d\times d} \mapsto \mathbb{R}^r$ is any three-times continuously differentiable function on some compact set, c_s is a positive-definite matrix that is the instantaneous covariance of the continuous part of the Itô semimartingale. In absence of noise, inferential frameworks of volatility functional estimation were established by [1, 2, 3]. Subsequently, specialized methodologies for various applications with novel empirical results blossomed in recent years, for example, [4, 5, 6].

To cope with noise, this paper embeds the *pre-averaging* method [7, 8, 9] into the general framework [1]. In this sense, this work extends the inferential

framework to accommodate noisy data, and generalizes the pre-averaging method to nonlinear transformations of volatility. On the road to a rate-optimal central limit theorem (CLT) with such generality, there are the following technicalities:

- *Stochastic volatility*: an nonparametric model (2.1) is used for robust-ness, yet, it becomes crucial to simultaneously control statistical error (due to noise) and discretization error (attributable to evolving parameters).
- *Noise & Jump*: there is an interplay between noise and jump, which necessitates truncating jumps on top of local moving averages, in order to recover volatility from noisy and jumpy observations.
- *Dependence*: because of overlapping windows in pre-averaging, the local moving averages are highly correlated to which standard CLTs does not apply. The "big block - small block" technique of [8] is used instead.
- *Bias*: generally there is an asymptotic bias due to nonlinearity of g in (1.1). In this paper, the bias is explicitly calculated and removed.
- *Unbounded derivatives*: some important applications, e.g., precision matrix estimation and linear regression, correspond to g's with singularities in derivatives near the origin, where the original framework [1] does not apply. A spatial localization argument by [6] is called upon in conjunction with a uniform convergence result.

It is the author's sincere hope, by solving these technicalities above, this paper will be able to offer a share of contribution to push the inferential framework to a new frontier of potentials and possibilities, and lend the effort to extend the corresponding applications to adopt noisy high-frequency data where exciting new stories await.

2. Setting.

2.1. *Model.* This paper assumes the data is generated from a process Y, and for any $t > 0$ there is a probability transition kernel Q_t linking another process X to Y where X is a solution to the stochastic differential equation

$$(2.1) \qquad X_t = X_0 + \int_0^t b_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}W_s + J_t$$

$b_s \in \mathbb{R}^d$, $\sigma_s \in \mathbb{R}^{d \times d'}$ with $d \leq d'$ and the volatility $c_s = \sigma_s \sigma_s^{\mathrm{T}}$ is positive semidefinite, W is a d'-dimensional standard Brownian motion, J is purely discontinuous process described by (A.1).

In this model, the noisy observations are samples from Y, and the underlying process before noise contamination is assumed as an Itô semimartingale.

An example of this model is

$$(2.2) \qquad Y_t = f(X_t, \varepsilon_t)$$

where $f : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}^d$ such that the conditional mean of Y_t is X_t and $\epsilon_t = Y_t - f(X_t, \varepsilon_t)$ is a white noise process. Generally, the noise model induced by (Q_t) incorporates additive white noise, rounding error, the combination thereof as special cases. Besides the probabilistic structure, the inferential framework also requires additional assumptions:

- the drift b has a smooth trajectory in certain sense (see appendix A);
- the volatility c is a locally spatially restricted Itô semimartingale[1] such that both c and c^{-1} is locally bounded;
- J may exhibit infinite activities but has finite variation, i.e., finite-length trajectory;
- the noise variance is an Itô semimartingale; conditioning on all the information on X, there is no autocorrelation in noise.[2]

These assumptions are necessary for the CLT and for applicability over functions of statistical interest. For readers interested in the precise description of the model specification and assumptions, please refer to appendix A.

2.2. *Observations.* This work treats regularly sampled observations and considers in-fill asymptotics[3]. Specifically, the samples are observed every Δ_n time units on a finite time interval $[0, t]$ where $n = \lfloor t/\Delta_n \rfloor$ is the sample size. As $n \to \infty$, $\Delta_n \to 0$ while t is fixed.

Throughout this paper, U_i^n is written for $U_{i\Delta_n}$ where U can be a process or filtration, for example, c_i^n denotes the value of volatility c at time $i\Delta_n$; for any process U, $\Delta_i^n U$ represents the increment $U_i^n - U_{i-1}^n$.

---

[1]However, it is important to accommodate long-memory volatility model. The volatility functional inference in long-memory and noisy setting is an open question under investigation.

[2]When the observations are mixed with colored noise, the statistical property of this methodology is unknown. Since it is empirically important, the author hopes this question can be illuminated by future research.

[3]aka fixed-domain asymptotics, high-frequency asymptotics, small-interval asymptotics

2.3. *Notations.* 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; $\| \cdot \|$ 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; for a multidimensional array, the entry index is written in the superscript, e.g., $X_t = (X_t^1, \cdots, X_t^d)^{\mathrm{T}}$, c^{jk} denotes the (j, k) entry in the matrix c; $\partial_{jk} g$ and $\partial^2_{jk,lm} g$ denote the gradient and Hessian with respect to the (j, k)-th and (l, m)-th entries; $\overset{\mathcal{L}-s(f)}{\longrightarrow}$ (resp. $\overset{\mathcal{L}-s}{\longrightarrow}$) denotes stable convergence of processes (resp. variables) in law[4]; $\overset{u.c.p.}{\longrightarrow}$ denotes uniform convergence in probability on compact sets; $\mathcal{MN}(0, V)$ is a mixed Gaussian distribution with mean 0 and conditional variance V.

3. Methods. The estimation methodology consists of 5 components:

   i. local moving averages of noisy data by a smoothing kernel φ, which act as proxies for X_i^n's;
   ii. jump truncation operated on local moving averages;
   iii. spot volatility estimator $\widehat{c}_i^n$'s for estimating c_i^n's;
   iv. Riemann sum of $g(\widehat{c}_i^n)$'s for approximating $\int g(c_s) \, \mathrm{d}s$;
   v. bias correction due to the nonlinearity, e.g., in case of $d = 1$ and constant volatility, by Taylor expansion, the estimation error of the plug-in estimator $g(\widehat{c})$ can be decomposed as

$$g(\widehat{c}) - g(c) = \underbrace{\partial g(c)(\widehat{c} - c)}_{\text{variance}} + \underbrace{\frac{1}{2}\partial^2 g(c)(\widehat{c} - c)^2}_{\text{bias}} + \underbrace{O_p(|\widehat{c} - c|^3)}_{\text{negligible}}$$

   the bias arises from the quadratic form of estimation error of $\widehat{c}$, provided g has a non-zero Hessian. This bias term does not affect the consistency, but one needs to explicitly correct the bias to get a CLT.

The moving-average idea is due to [8, 7]; the truncation is modified from (16.4.4) in [10]; the plug-in and bias correction are inspired by [1]. The specific recipe is given next.

---

[4]See section 2.2.1, 2.2.2 in [10] for stable convergence. The sampling variation of the estimator depends on the realization of the process c, hence we need a mode of convergence in which the estimator converges jointly with other variables, so that one can consistently estimate the asymptotic variance to compute confidence intervals.

3.1. *Building blocks.* For the local moving averages, we choose a smoothing kernel φ such that

(3.1) $\qquad \begin{aligned} &\operatorname{supp}(\varphi) \subset (0,1), \ \int_0^1 \varphi^2(s)\,\mathrm{d}s > 0 \\ &\varphi \in \mathcal{C} \text{ is piecewise } \mathcal{C}^1; \ \varphi' \text{ is piecewise Lipschitz} \end{aligned}$

Choose an integer l_n as the number of observations in each smoothing window, define $\varphi_h^n = \varphi(h/l_n)$ and $\psi_n = \sum_{h=1}^{l_n-1}(\varphi_h^n)^2$. Associate the following quantities with a generic process U:

(3.2) $\qquad \begin{aligned} \bar{U}_i^n &= (\psi_n)^{-1/2} \sum_{h=1}^{l_n-1} \varphi_h^n \Delta_{i+h-1}^n U \\ \widehat{U}_i^n &= (2\psi_n)^{-1} \sum_{h=0}^{l_n-1} (\varphi_{h+1}^n - \varphi_h^n)^2 \Delta_{i+h}^n U \cdot \Delta_{i+h}^n U^{\mathrm{T}} \end{aligned}$

$\bar{Y}_i^n$ is a local moving average of the noisy data Y_i^n's and is a proxy for $\Delta_i^n X$, $\widehat{Y}_i^n$ serves as noise correction to $\bar{Y}_i^n$. Based on these 2 ingredients, choose $k_n > l_n$, define the spot volatility estimator as

(3.3) $\qquad \widehat{c}_i^n \equiv \dfrac{1}{(k_n - l_n)\Delta_n} \sum_{h=1}^{k_n-l_n+1} \left(\bar{Y}_{i+h}^n \cdot \bar{Y}_{i+h}^{n,\mathrm{T}} \mathbb{1}_{\{\|\bar{Y}_{i+h}^n\| \leq \nu_n\}} - \widehat{Y}_{i+h}^n \right)$

where $\nu_n \asymp \Delta_n^\rho$ is a truncation threshold for jumps. The choice of ρ is stated in (3.6). A spot noise variance estimator is also needed:

(3.4) $\qquad \widehat{\gamma}_i^n \equiv \dfrac{1}{2m_n} \sum_{h=1}^{m_n} \Delta_{i+h}^n Y \cdot \Delta_{i+h}^n Y^{\mathrm{T}}$

where $m_n = \lfloor \theta' \Delta_n^{-1/2} \rfloor$, θ' positive finite.

3.2. *The estimator.*

DEFINITION 1. Let $N_t^n = \lfloor t/(k_n \Delta_n) \rfloor$, the estimator of (1.1) is defined as

$$\widehat{S}(g)_t^n \equiv k_n \Delta_n \sum_{i=0}^{N_t^n-1} \left[g(\widehat{c}_{ik_n}^n) - B(g)_{ik_n}^n \right] \times a_t^n$$

where $B(g)_i^n$ is a de-biasing term of the form

$$B(g)_i^n = \dfrac{1}{2k_n \Delta_n^{1/2}} \sum_{j,k,l,m=1}^d \partial_{jk,lm}^2 g(\widehat{c}_i^n) \times \Xi(\widehat{c}_i^n, \widehat{\gamma}_i^n)^{jk,lm}$$

$\widehat{c}_i^n$, $\widehat{\gamma}_i^n$ are defined in (3.3), (3.4), Ξ is defined in (4.3), and $a_t^n = t/(N_t^n k_n \Delta_n)$ is a finite-sample adjustment.[5]

Besides φ, there are 3 tuning parameters in this estimator:

$a\Delta_n^b$	scale a	rate b	description
l_n	θ	$-1/2$	length of overlapping window for local moving averages
k_n	ϱ	$-\kappa$	length of disjoint window for spot volatility estimation
ν_n	α	ρ	truncation level for jumps

With suitable choices of l_n, k_n, ν_n in (3.3), this estimator is applicable to any function $g : \mathcal{S}_d^+ \mapsto \mathbb{R}^r$ that satisfies

$$(3.5) \qquad\qquad\qquad g \in \mathcal{C}^3(\mathcal{S})$$

where $\mathcal{S} \supset \cup_m \mathcal{S}_m^\epsilon$ for some $\epsilon > 0$, $\mathcal{S}_m^\epsilon = \left\{ A \in \mathcal{S}_d^+ : \inf_{M \in \mathcal{S}_m} \|A - M\| \leq \epsilon \right\}$ and $\mathcal{S}_m$ is identified in assumption A-ν.

3.3. *Choosing tuning parameters.* A proper combination of the tuning parameters is crucial for consistency, CLT, and optimal convergence rate. For these objectives, one needs

$$(3.6) \qquad \begin{cases} l_n & \asymp & \theta \Delta_n^{-1/2} \\ k_n & \asymp & \varrho \Delta_n^{-\kappa} \quad & \text{where } \kappa \in \left(\frac{2}{3} \vee \frac{2+\nu}{4}, \frac{3}{4} \right) \\ \nu_n & = & \alpha \Delta_n^\rho \quad & \text{where } \rho \in \left[\frac{1}{4} + \frac{1-\kappa}{2-\nu}, \frac{1}{2} \right) \end{cases}$$

$\theta, \varrho, \alpha > 0$ are positive finite, and $\nu \in [0, 1)$ is introduced in assumption A-ν which dictates the jump intensity.

The rest of this section offers an intuition for (3.6). The reader can skip this part without affecting understanding of the main result in section 4.

1. *l_n influences the convergence rate*
   In the example (2.2), according to (3.2),

   $$\overline{Y}_i^n = \overline{X}_i^n + \overline{\varepsilon}_i^n$$

   and we can write $\overline{\varepsilon}_i^n = -\psi_n^{-1/2} \sum_{h=0}^{l_n-1} (\varphi_{h+1}^n - \varphi_h^n) \varepsilon_{i+h}^n$. Under the conditional independence of ε_i^n's, $\overline{\varepsilon}_i^n = O_p(l_n^{-1})$; $\overline{X}_i^n = O_p(\Delta_n^{-1/2})$ by (B.12).

---

[5]Overlapping intervals are used to compute $\widehat{c}_i^n$'s, non-overlapping intervals are used to compute $\widehat{S}(g)_t^n$. The local moving averages computed over overlapping intervals in (3.3) are necessary to achieve the optimal convergence rate. By contrast, overlapping intervals in $\widehat{S}(g)_t'^n \equiv \Delta_n \sum_{i=0}^{\lfloor t/\Delta_n \rfloor - 1} [g(\widehat{c}_i^n) - B(g)_i^n]$ do not improve the convergence rate nor efficiency, though lead to robustness in finite sample. In fact, the overlapping-interval-based estimator has the same asymptotic result as that of $\widehat{S}(g)_t^n$ in section 4.

By taking $l_n \asymp \Delta_n^{-1/2}$ the orders of $\bar{X}_i^n$ and $\bar{\varepsilon}_i^n$ are equal, this choice of local smoothing window will deliver the optimal rate of convergence.

2. k_n *dictates bias-correction and the CLT form*

   Here let's focus on the case $d = 1$, X is continuous, then

   $$\widehat{c}_i^n - c_i^n = d_i^n + s_i^n$$

   - $d_i^n = \frac{1}{(k_n - l_n)\Delta_n} \int_{i\Delta_n}^{(i+k_n-l_n+1)\Delta_n} (c_s - c_i^n)\, \mathrm{d}s$ is the "*discretization error*", $d_i^n = O_p((k_n\Delta_n)^{1/2})$ by (B.2);

   - $s_i^n \approx \frac{1}{(k_n - l_n)\Delta_n} \Delta_n^{1/4} (\chi_{i+k_n-l_n+1}^n - \chi_i^n)$ is the "*statistical error*", where χ is a continuous Itô semimartingale, this result is due to (3.8) in [8], so $s_i^n = O_p((k_n\Delta_n^{1/2})^{-1/2})$.

   Balancing the orders of d_i^n and s_i^n by setting $\kappa = 3/4$ will result in the minimum order of total estimation error. However, in the case $\kappa \geq 3/4$ the bias involves volatility of volatility and volatility jump, which are difficult to estimate and de-bias in applications. Therefore, it is advisable to choose $\kappa < 3/4$, in which case the statistical error dominates in the bias, thereby the thorny terms are circumvented. Besides, to achieve successful de-biasing of statistical error and negligibility of higher-order Taylor-expansion terms, we need $\kappa > 2/3$.
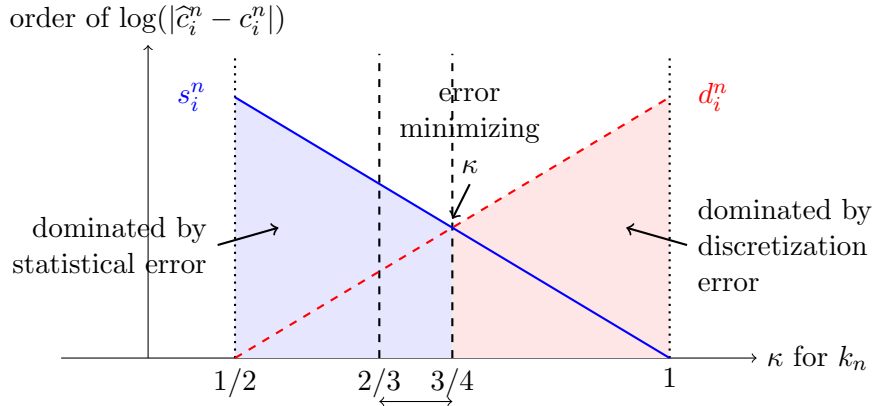
   Section 3.1, 3.2 of [1] give a similar discussion in absence of noise.

3. ν_n *disentangles volatility from jump variation*

   $\|\bar{Y}_i^n\| = O_p(\Delta_n^{1/2})$ if there is no jump over $[i\Delta_n, (i+l_n)\Delta_n]$, via (B.6). By choosing $\rho < 1/2$, the truncation level, which is $\nu_n > \Delta_n^{1/2}$, keeps the diffusion movements and discards jumps in a certain sense. To effectively filter out the jumps, the truncation level should be bounded above and the upper bounds depends on the jump activity index ν.

REMARK.   If the reader is interested to estimate (1.1) with g satisfying

(3.7) $\|\partial_h g(x)\| \leq K_h (1 + \|x\|^{r-h}), \ h = 0, 1, 2, 3, \ r \geq 3$

the requirements on k_n and ν_n can be loosened and become

(3.8)
$$
\begin{aligned}
k_n \Delta_n^\kappa &\asymp \varrho, \ \text{where } \kappa \in \left(\tfrac{2}{3}, \tfrac{3}{4} \right) \\
\nu_n &= \alpha \Delta_n^\rho, \ \text{where } \rho \in \left[\tfrac{1}{4} + \tfrac{1}{4(2-\nu)}, \tfrac{1}{2} \right)
\end{aligned}
$$

For wider applicability, we choose to accommodate the functional space (3.5) and retain the requirement (3.6).

4. Asymptotics.

4.1. *Elements.*   Before stating the asymptotic result, some elements appearing in the limit need to be defined. Associate the following quantities with the smoothing kernel φ for $l, m = 0, 1$:

(4.1)
$$
\begin{aligned}
\phi_0(s) &= \int_s^1 \varphi(u)\varphi(u-s)\,\mathrm{d}u, \quad \phi_1(s) = \int_s^1 \varphi'(u)\varphi'(u-s)\,\mathrm{d}u \\
\Phi_{lm} &= \int_0^1 \phi_l(s)\phi_m(s)\,\mathrm{d}s, \qquad \Psi_{lm} = \int_0^1 s\,\phi_l(s)\phi_m(s)\,\mathrm{d}s
\end{aligned}
$$

Define Σ, Θ, Υ as $\mathbb{R}^{d \times d \times d \times d}$-valued functions, such that for $x, z \in \mathbb{R}^{d \times d}$, $j, k, l, m = 1, \cdots, d$,

(4.2)
$$
\begin{aligned}
\Sigma(x)^{jk,lm} &= x^{jl}x^{km} + x^{jm}x^{kl} \\
\Theta(x, z)^{jk,lm} &= x^{jl}z^{km} + x^{jm}z^{kl} + x^{km}z^{jl} + x^{kl}z^{jm}
\end{aligned}
$$

and Ξ also as a tensor-valued function

(4.3) $\Xi(x, z) = \dfrac{2\theta}{\phi_0(0)^2} \left[\Phi_{00}\Sigma(x) + \dfrac{\Phi_{01}}{\theta^2}\Theta(x, z) + \dfrac{\Phi_{11}}{\theta^4}\Sigma(z) \right]$

where θ is introduced in (3.6).
Now we are ready to describe the limit process.

DEFINITION 2.   Given g satisfying (3.5) or (3.7), $Z(g)$ is a process defined on an extension of the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ specified in $(A.3)$, such that conditioning on $\mathcal{F}$, $Z(g)$ is a mean-0 Gaussian process with conditional variance

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

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

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

with $\partial_{jk}g(c)$ being the first-order partial derivative of g with respect to c^{jk}, γ being the conditional covariance process of noise defined in $(A.4)$.

4.2. *The formal results.*

THEOREM 1. *Assume assumptions A-ν, A-γ. Given g satisfying (3.5), we control the tuning parameters l_n, k_n, ν_n according to (3.6), then we have the following stable convergence in law of discretized process to a conditional continuous Itô semimartingale on compact subsets of $\mathbb{R}^+$:*

$$(4.5) \qquad \Delta_n^{-1/4}\left[\widehat{S}(g)^n - S(g)\right] \xrightarrow{\mathcal{L}-s(f)} Z(g)$$

where $S(g)$ is defined in (1.1), $\widehat{S}(g)^n$ is from definition 1, $Z(g)$ is identified in definition 2.

Theorem 1 is valid over the functional space (3.5), which is as general as the current literature can get. If applications require functionals whose derivatives satisfy the polynomial growth condition (3.7), we can put less restrictions on the tuning parameters.

THEOREM 2. *Assume assumptions A-ν, A-γ. Replace the functional space (3.5) with (3.7), replace the tuning conditions (3.6) on k_n, ν_n with (3.8), then (4.5) still holds true.*

However, theorem 2 rules out operations that involve matrix inversion, it is not applicable to, for instance, inference of linear regression models. In the rest of this paper, we focus on the results over the general functional space (3.5).

The asymptotic result is stated with a probabilistic flavor, which is necessary to express the strongest convergence[6] by appendix B. There is an alternative formulation which is more relevant for statistical applications:

$$(4.6) \qquad n^{1/4}\left[\widehat{S}(g)^n_t - S(g)_t\right] \xrightarrow{\mathcal{L}-s} \mathcal{MN}\left(0, \sqrt{t}V(g)_t\right)$$

this is true under the same conditions and t is finite.

4.3. *Confidence intervals.* The asymptotic variance in (4.6) can be estimated by plugging in spot estimates of volatility (3.3) and noise covariance matrix (3.4):

$$(4.7) \quad \widehat{V}(g)^n_t \equiv k_n\Delta_n \sum_{i=0}^{N^n_t-1} \sum_{j,k,l,m=1}^{d} \partial_{jk}g(\widehat{c}^n_{ik_n})\,\partial_{lm}g(\widehat{c}^n_{ik_n})^{\mathrm{T}}\,\Xi(\widehat{c}^n_{ik_n},\widehat{\gamma}^n_{ik_n})^{jk,lm}$$

---

[6]It is functional stable convergence (or stable convergence of processes) in law.

PROPOSITION 1. $\widehat{V}(g)_t^n$ *is consistent under* (3.6) *and assumptions* A-ν, A-γ. *Specifically, for all finite* t,

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

where κ *is specified in* (3.6).

PROOF. The asymptotic variance (4.4) is a smooth functional of spot volatility and instantaneous noise covariance, so the consistency of (4.7) follows from the consistence of the spot volatility estimator (3.3) and the noise covariance estimator (3.4). According to lemma 5 and (B.26), the error rate of (4.7) is determined by the estimation error of spot volatility. Therefore, the error rate of (4.7) is the same as the error rate of the volatility functional estimator without bias correction, which is $(k_n\Delta_n^{1/2})^{-1}$, then the proposition follows from (3.6). □

Based on theorem 1, proposition 1 and the property of stable convergence, we have the following feasible central limit theorem:

COROLLARY 1. *Under* (3.6) *and assumptions* A-ν, A-γ, *we have*

$$(4.8) \qquad \left[\Delta_n^{1/2}\,\widehat{V}(g)_t^n\right]^{-1/2}\left[\widehat{S}(g)_t^n - S(g)_t\right] \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \mathbb{I}\right)$$

in restriction to the event $\{\omega \in \Omega, \widehat{V}(g)_t^n$ *is positive definite*$\}$, *where* Ω *is defined in* (A.3).

5. Applications.

5.1. *Quarticity estimation.* In the univariate setting, the so-called quarticity $\int_0^t c_s^2\,\mathrm{d}s$ appears in the asymptotic variances of many extant volatility estimators. The multivariate counterpart involves $\int_0^t c_s^{jl}c_s^{km} + c_s^{jm}c_s^{kl}\,\mathrm{d}s$, e.g., $\Xi(c_s, \gamma_s)^{jk,lm}$ in (4.4). Since the quarticity is an integrated functional of volatility, the volatility functional estimator facilitates uncertainty quantification for various volatility estimators.

5.2. *Realized Laplace transform.* [11] put forward an estimator of the realized Laplace transform of volatility defined as

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

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 parameter(s) or test the model. An open question noted by [11] is the estimation of realized Laplace transform using noisy data. By the nonparametric estimation of volatility path in the first stage and the bias-corrected Riemann summation of functional plug-ins in the second stage, this paper contributes a rate-optimal solution to the open question.

5.3. *Generalized method of moments (GMM).*   [4] 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

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

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 $g(Z_t, c_t; \theta) = \mathbb{E}[Y_t - f(Z_t, c_t; \theta)]$, then we have the following integrated moment condition:

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

where $G(\theta) = \int_0^t g(Z_s, c_s; \theta)\, \mathrm{d}s$. Utilizing noisy observations of X at higher frequencies, $\widehat{S}(g)_t^n$ of this paper provides a means to compute a bias-corrected sample moment function of GMM.

5.4. *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 the 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 $b = (b^{S,\mathrm{T}}, b^Z)^{\mathrm{T}}$, $W = (W^{S,\mathrm{T}}, W^R)^{\mathrm{T}}$, $J = (J^{S,\mathrm{T}}, J^Z)^{\mathrm{T}}$ and

$$\sigma = \left[\begin{array}{cc} \sigma^S & 0 \\ \beta^{\mathrm{T}} \sigma^S & \sigma^R \end{array} \right]$$

so

$$c = \sigma \sigma^{\mathrm{T}} = \left[\begin{array}{cc} \sigma^S \sigma^{S,\mathrm{T}} & \sigma^S \sigma^{S,\mathrm{T}} \beta \\ \beta^{\mathrm{T}} \sigma^S \sigma^{S,\mathrm{T}} & \beta^{\mathrm{T}} \sigma^S \sigma^{S,\mathrm{T}} \beta + (\sigma^R)^2 \end{array} \right] \equiv \left[\begin{array}{cc} c^{SS} & c^{SZ} \\ c^{ZS} & c^{ZZ} \end{array} \right]$$

hence by letting $g(c) = s^{SS,-1} c^{SZ}$, we have $\beta = t^{-1} S(g)_t$. [6] proposed this method for the situation in which the process X can be perfectly observed. When the observations contain noise, the methodology of this paper can extend the estimator of [6] to wider applicability.

5.5. *Principal component analysis (PCA).* An interesting question about stochastic volatility is its spectral structure $c_s v_s = \lambda_s v_s$. [5] applied PCA to nonstationary 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 c_s and v_s is the corresponding eigenvector, $g(c_s) = \lambda_s$ and $g(c_s) = v_s$ are three-times continuously differentiable, therefore the inferential results of $S(g)$ are applicable. More recently, [12] extends the realized PCA to asynchronously observed high-dimensional noisy data, while this paper extends the realized PCA methodology to be both noise-robust and rate-optimal.

6. Simulation. As a proof of concept, estimators corresponding to $g(c) = c^2$, $g(c) = \log(c)$ when $d = 1$ 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 $\varepsilon_i^n \overset{\text{i.i.d.}}{\sim} N(0, .005^2)$, $\mathbb{E}[(W_{t+\Delta} - W_t)(B_{t+\Delta} - B_t)] = -.6\Delta$, $J_t^X \sim N(-.01, .02^2)$, $N_{t+\Delta}^X - N_t^X \sim \text{Poisson}(36\Delta)$, $\log(J_t^c) \sim N(-5, .8)$, $N_{t+\Delta}^c - N_t^c \sim \text{Poisson}(12\Delta)$.

Each simulation employs 23400×21 data points with $\Delta_n = 1s$. We choose the following tuning parameters:

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

where $\overline{\sigma}^2$ is an estimate of the average volatility by bipower variation [13]. The results are shown in figure 1.

FIG 1. *Simulation of volatility functional estimators*
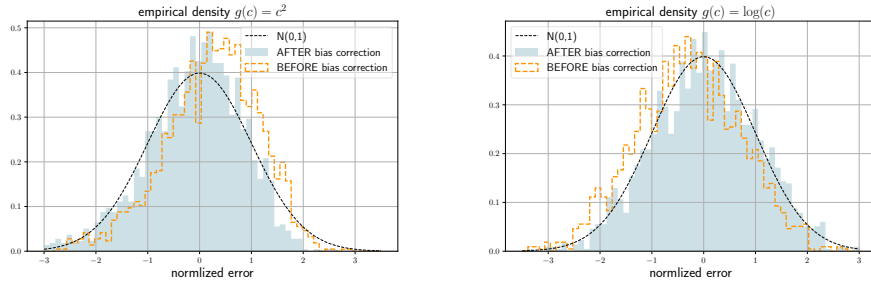
7. Discussions.

7.1. *Jump truncation.* It is worthwhile to take account of the dimension d and volatility levels in jump truncation. One possibility is to use the truncation indicator $\prod_{r=1}^{d} \mathbb{1}_{\{|\overline{Y}_i^{r,n}| \leq \alpha_r \Delta_n^\rho\}}$ where α_r is related to the volatility of the r-th component.

7.2. *Semi-efficiency.* The efficient bound of volatility estimation is studied by [14]. Nonetheless, pre-averaging hasn't been able to attain the efficiency bound. According to [15], by taking the moving average adaptively in the time domain, the asymptotic variance of the pre-averaging method can be within 7% of the efficiency bound. Using adaptive pre-averaging in volatility functional estimation is beyond the scope of this paper and is currently under investigation.

An efficient alternative is the spectral approach [16, 17]. The multi-scale approach [18], realized kernels [19], quasi-likelihood [20] are equally capable to rate-optimally handle noise. In the univariate case realized kernels and quasi-likelihood could be improved to be nearly efficient, cf. [21]. The pre-averaging are adopted in this paper to simultaneously handle price jumps and microstructure noise.

7.3. *Finite-sample consideration.* First, we consider effective jump truncations. It is worthwhile to take account of the dimension d and volatility levels in jump truncation. One possibility is to use the truncation indicator $\prod_{r=1}^{d} \mathbb{1}_{\{|\overline{Y}_i^{r,n}| \leq \alpha_r \Delta_n^\rho\}}$ where α_r is related to the volatility of the r-th component.

Next, in finite sample, the spot volatility estimator (3.3) might not be positive semidefinite due to the noise-correction term $\widehat{Y}_i^n$. [22] suggests to increase l_n to attenuate noise in $\overline{Y}_i^n$ and dispense with $\widehat{Y}_i^n$:

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

and let $l_n \asymp \theta \Delta_n^{-1/2 - \delta}$ where $\delta \in (.1, .5)$. According to [23], if one plugs in $\widetilde{c}_i^n$ with the tunning parameters satisfying

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

another central limit theorem holds after a different bias correction. However, doing so sacrifices the convergence rate, which drops from $n^{1/4}$ down to $n^{1/4 - \delta/2}$ (strictly less than $n^{1/5}$). Moreover, the choice of tunning parameters becomes less robust compared to (3.6).

To preserve the optimal convergence rate in the unfortunate event where the spot volatility estimator is not positive semidefinite, it is advisable to project $\widehat{c}_i^n$ onto the convex cone $\mathcal{S}_d^+$ with respect to the Frobenius norm. Suppose $\widehat{c}_i^n = Q\Lambda Q^{\mathrm{T}}$ is the eigenvalue factorization, the positive semidefinite projection is $\widetilde{c}_i^n = Q\Lambda_+ Q^{\mathrm{T}}$ where $\Lambda_+^{jj} = \Lambda^{jj} \vee 0$. By the convex geometry of $\mathcal{S}_d^+$, we have $\|\widetilde{c}_i^n - c_i^n\| \leq \|\widehat{c}_i^n - c_i^n\|$, hence the convergence rate is retained.

APPENDIX A: ASSUMPTIONS

This section presents details of model specification and assumptions.
The pure jump process is

$$(A.1) \qquad J_t = \int_{(0,t] \times E} \delta(s, x) \, \mathfrak{p}(\mathrm{d}s, \, \mathrm{d}x)$$

where δ is a $\mathbb{R}^d$-valued predictable function on $\mathbb{R}^+ \times E$, E is a Polish space, $\mathfrak{p}$ is a Poisson random measure with compensator $\mathfrak{q}(\mathrm{d}u, \mathrm{d}x) = \mathrm{d}u \otimes \lambda(\mathrm{d}x)$, λ is a σ-finite measure on E and has no atom. The volatility process is assumed to be an Itô semimartingale

$$(A.2) \quad c_t = c_0 + \int_0^t b_s^{(c)} \, \mathrm{d}s + \int_0^t \sigma_s^{(c)} \, \mathrm{d}W_s + \int_{(0,t] \times E} \delta^{(c)}(s, x) \, (\mathfrak{p} - \mathfrak{q})(\mathrm{d}s, \, \mathrm{d}x)$$

where $b^{(c)}$ is $\mathbb{R}^{d \times d}$-valued, optional, càdlàg; $\sigma^{(c)}$ is $\mathbb{R}^{d \times d \times d'}$-valued, adapted, càdlàg; $\delta^{(c)}$ is a $\mathbb{R}^{d \times d}$-valued predictable function on $\mathbb{R}^+ \times E$.[7]

Let $\left(\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)}), \mathbb{P}^{(0)}\right)$ be a filtered probability space with respect to which X, c are $(\mathcal{F}_t^{(0)})$-adapted; let $\left(\Omega^{(1)}, \mathcal{F}^{(1)}, (\mathcal{F}_t^{(1)}), \mathbb{P}^{(1)}\right)$ be another filtered probability space accommodating Y; $\forall t \geq 0$, $\forall A \in \mathcal{F}^{(0)}$, let $Q_t(A, \cdot)$ be a conditional probability measure on $\left(\Omega^{(1)}, \mathcal{F}^{(1)}\right)$.

All the stochastic dynamics above can be described on the filtered extension $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, where

(A.3)
$$\begin{cases} \Omega = \Omega^{(0)} \times \Omega^{(1)} \\ \mathcal{F} = \mathcal{F}^{(0)} \otimes \mathcal{F}^{(1)} \\ \mathcal{F}_t = \bigcap_{s>t} \left(\mathcal{F}_s^{(0)} \otimes \mathcal{F}_s^{(1)}\right), \quad \widetilde{\mathcal{F}}_t = \bigcap_{s>t} \left(\mathcal{F}^{(0)} \otimes \mathcal{F}_s^{(1)}\right) \\ \mathbb{P}\left(A \times \mathrm{d}\omega\right) = \mathbb{P}^{(0)}(A) \cdot \otimes_{t \geq 0} Q_t(A, \mathrm{d}\omega), \forall A \in \mathcal{F}^{(0)} \end{cases}$$

In the sequel, $\mathbb{E}(\cdot)$ denotes the expectation operator on $(\Omega^{(0)}, \mathcal{F}^{(0)})$ or $(\Omega, \mathcal{F})$; $E(\cdot|\mathcal{H})$ denotes the conditional expectation operator, with $\mathcal{H}$ being $\mathcal{F}_t^{(0)}$, $\mathcal{F}_t^{(1)}$, $\mathcal{F}_t$, $\widetilde{\mathcal{F}}_t$; $E_i^n(\cdot)$ denotes $E(\cdot|\mathcal{F}_i^n)$. Assumptions are collected below.

ASSUMPTION A-ν (regularity). b has $\frac{1}{2}$-Hölder sample path, i.e., $\forall t, s \geq 0$,

$$E\left(\sup_{u \in [0,s]} \|b_{t+u} - b_t\| \Big| \mathcal{F}_t^{(0)}\right) \leq K s^{1/2}, \; a.s.;$$

c is of the form (A.2), there is a sequence of triples $(\tau_m, \mathcal{S}_m, \Gamma_m)$, where τ_m is a stopping time and $\tau_m \nearrow \infty$; $\mathcal{S}_m \subset \mathcal{S}_d^+$ is convex, compact such that

$$t \in [0, \tau_m] \Rightarrow c_t \in \mathcal{S}_m;$$

Γ_m is a sequence of bounded λ-integrable functions on E, such that

$$t \in [0, \tau_m] \Longrightarrow \begin{cases} \|b_t\| + \|\sigma_t\| + \|b_t^{(c)}\| + \|\sigma_t^{(c)}\| \leq m \\ \|\delta(t, x)\|^{\nu} \wedge 1 \leq \Gamma_m(x), \; \nu \in [0, 1) \\ \|\delta^{(c)}(t, x)\|^2 \wedge 1 \leq \Gamma_m(x) \end{cases}$$

Let

$$\epsilon_t(\omega) = Y_t(\omega) - X_t$$

and the conditional covariance process of noise can be written as

(A.4)
$$\gamma_t = \int_{\Omega^{(1)}} \epsilon_t(\omega) \, \epsilon_t(\omega)^{\mathrm{T}} \, Q_t(\cdot, \mathrm{d}\omega)$$

---

[7]To guarantee positive semidefiniteness in concrete applications, one needs to impose parametric restrictions on the nonparametric model (A.2). The result in this paper holds true for any process c which is positive semidefinite and satisfies (A.2).

ASSUMPTION A-γ (noise). $\forall t \in \mathbb{R}^+$,

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

$\forall t \neq s, \; \forall A \in \mathcal{F}_{s\wedge t}^{(0)}$

$$\int_{\Omega^{(1)}\times\Omega^{(1)}} \epsilon_t(\omega) \, \epsilon_s(\omega)^{\mathrm{T}} \, Q_t(A, \, \mathrm{d}\omega) \, Q_s(A, \, \mathrm{d}\omega) = 0$$

furthermore,

$$\gamma_t = \gamma_0 + \int_0^t b_s^{(r)} \, \mathrm{d}s + \int_0^t \sigma_s^{(r)} \, \mathrm{d}W_s + \int_{(0,t]\times E} \delta^{(r)}(s, x) \, \mathfrak{p}(\mathrm{d}s, \, \mathrm{d}x)$$

for the same τ_m, Γ_m in assumption A-ν and $\forall q \leq 8$

$$t \in [0, \tau_m] \Longrightarrow \left\{ \begin{array}{l} \|b_t^{(r)}\| + \|\sigma_t^{(r)}\| + \mathbb{E}_{\Omega^{(1)}} \|\epsilon_t(\omega)\|^q \leq m \\ \|\delta^{(r)}(t, x)\|^2 \wedge 1 \leq \Gamma_m(x) \end{array} \right.$$

APPENDIX B: DERIVATION

B.1. Preliminaries. In the sequel, the constant K changes across lines but remains finite; K_q is a constant depending on q; $a_n = O_p(b_n)$ means $\forall \epsilon > 0$, $\exists M > 0$ such that $\sup_n \mathbb{P}(a_n/b_n > M) < \epsilon$; $a_n \asymp b_n$ means both a_n/b_n and b_n/a_n are bounded for large n. Six useful results are stated below. I. By a localization argument from section 4.4.1 in [10], without loss of generality we can assume $\exists$ a constant K, a bounded λ-integrable function Γ on E, a convex compact subset $\mathcal{S} \in \mathcal{S}_d^+$ and $\epsilon > 0$, $g \in \mathcal{C}^3(\mathcal{S}^\epsilon)$ where $\mathcal{S}^\epsilon$ denotes the ϵ-enlargement of $\mathcal{S}$ (see (3.5)), such that
(B.1)
$$\left\{ \begin{array}{l} \|b\| + \|\sigma\| + \|b^{(c)}\| + \|\sigma^{(c)}\| + \|b_t^{(r)}\| + \|\sigma_t^{(r)}\| + \mathbb{E}\|\epsilon_t\|^q \leq K, \; \forall q \leq 8 \\ \|\delta(t, x)\|^\nu \wedge 1 \leq \Gamma(x), \; \nu \in [0, 1) \\ \|\delta^{(c)}(t, x)\|^2 \wedge 1 + \|\delta^{(r)}(t, x)\|^2 \wedge 1 \leq \Gamma(x) \\ c \in \mathcal{S} \end{array} \right.$$

II. Define a continuous Itô semimartingale with parameters corresponding to those in (2.1),

$$X_t' = X_0 + \int_0^t b_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}W_s$$

Let $Y^* = Y - X + X'$. Based on (3.2), define

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

In the upcoming derivation, $\|\widehat{c}_i^n - \widehat{c}_i^{*n}\|$ is tightly bounded provided ν_n is properly chosen, the focus then will be shifted from $\widehat{c}_i^n$ to $\widehat{c}_i^{*n}$.

III. By estimates of Itô semimartingale increments, $\forall$ finite stopping time τ

$$(\text{B.2}) \quad \begin{cases} \left\|E\left(X_{\tau+s}' - X_\tau'|\mathcal{F}_\tau^{(0)}\right)\right\| + \left\|E\left(c_{\tau+s} - c_\tau|\mathcal{F}_\tau^{(0)}\right)\right\| \\ \qquad\qquad\qquad + \left\|E\left(\gamma_{\tau+s} - \gamma_\tau|\mathcal{F}_\tau^{(0)}\right)\right\| & \leq Ks \\ E\left(\sup_{u\in[0,s]}\left\|X_{\tau+u}' - X_\tau'\right\|^q |\mathcal{F}_\tau^{(0)}\right) & \leq Ks^{q/2} \\ E\left(\sup_{u\in[0,s]}\left\|c_{\tau+u} - c_\tau\right\|^q + \left\|\gamma_{\tau+u} - \gamma_\tau\right\|^q |\mathcal{F}_\tau^{(0)}\right) & \leq Ks^{(q/2)\wedge 1} \end{cases}$$

by Lemma 2.1.7, Corollary 2.1.9 in [10]

$$(\text{B.3}) \quad \begin{cases} E\left(\sup_{u\in[0,s]}\|J_{\tau+u} - J_\tau\|^q|\mathcal{F}_\tau^{(0)}\right) \leq K_q\, sE\left[\widehat{\delta}(q)_{\tau,s}|\mathcal{F}_\tau^{(0)}\right] \\ E\left[\sup_{u\in[0,s\wedge 1]}\left(\frac{\|J_{\tau+u}-J_\tau\|}{s^w} \wedge 1\right)^q|\mathcal{F}_\tau^{(0)}\right] \leq K\,s^{(1-w\nu)(q/\nu\wedge 1)}a(s) \end{cases}$$

where $\widehat{\delta}(q)_{t,s} \equiv s^{-1}\int_t^{t+s}\int_E \|\delta(u,x)\|^q\,\lambda(\mathrm{d}x)\,\mathrm{d}u$ and $a(s) \to 0$ as $s \to 0$.

IV. Let $\varphi_n(t) = \sum_{h=1}^{l_n-1}\varphi_h^n \mathbb{1}_{((h-1)\Delta_n, h\Delta_n]}(t)$. For a generic process U, define

$$(\text{B.4}) \qquad\qquad\qquad U_{t,s}^n = \int_t^{t+s}\varphi_n(u - t)\,\mathrm{d}U_u$$

this quantity is useful in analyzing $\bar{U}_i^n$.

V. For $p \in \mathbb{N}^+$, $l, m = 0, 1$, by (4.1) and Riemann summation,

$$(\text{B.5}) \quad \sum_{h=i}^{i+pl_n-2}\sum_{h'=h+1}^{i+pl_n-1}\phi_l\left(\frac{h'-h}{l_n}\right)\phi_m\left(\frac{h'-h}{l_n}\right) = l_n^2\left(p\Phi_{lm} - \Psi_{lm}\right) + O(pl_n)$$

VI. By Jensen's inequality and Doob's maximal inequality, we have the following lemma:

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

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

B.2. Properties of spot estimator: I. jumps. By assumption A-ν, A-γ, (3.6), (B.1), (B.2), (B.3), and $\bar{J}_i^n = \psi_n^{-1/2} \bar{J}_{i\Delta_n,(l_n-1)\Delta_n}^n$ from (B.4),

(B.6)

$$
\begin{cases}
E_i^n\left(\|\bar{Y}_i^{*n}\|^q\right) & \leq & K_q\Delta_n^{q/2}, \ \forall q \leq 8 \\
E_i^n\left(\|\bar{Y}_i^n\|^q|\right) & \leq & K_q\Delta_n^{(q/2)\wedge(q/4+1/2)}, \ \forall q \leq 8 \\
E_i^n\left(\|\widehat{Y}_i^{*n}\|^q \vee \|\widehat{Y}_i^n\|^q\right) & \leq & K_q\Delta_n^q, \ \forall q \leq 4 \\
E_i^n\left[\left(\frac{\|\bar{J}_i^n\|}{\Delta_n^w} \wedge 1\right)^q\right] & \leq & K_q\Delta_n^{[1/2-(w-1/4)\nu]\times[1\wedge(q/\nu)]} a_n, \ \forall q \leq 8
\end{cases}
$$

for some $a_n \to 0$. We can write

$$
\left\|\left(\bar{Y}_i^n \cdot \bar{Y}_i^{n,\mathrm{T}}\mathbb{1}_{\{\|\bar{Y}_i^n\|\leq \nu_n\}} - \widehat{Y}_i^n\right) - \left(\bar{Y}_i^{*n} \cdot \bar{Y}_i^{*n,\mathrm{T}} - \widehat{Y}_i^{*n}\right)\right\| \leq \sum_{r=1}^3 \eta_i^{n,r}
$$

where

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

Let $u_n = \nu_n/\Delta_n^{1/2}$, $Z_i^n = \|\bar{Y}_i^{*n}\|/\Delta_n^{1/2}$, $Q_i^n = (\|\bar{J}_i^n\|/\Delta_n^{1/2}) \wedge 1$, $V_i^n = (\|\bar{J}_i^n\|/\Delta_n^{\rho}) \wedge 1$, we have

$$
\Delta_n\eta_i^{n,1} \leq u_n^{-2/(1-2\rho)}(Z_i^n)^{2+2/(1-2\rho)} + (1 + Z_i^n)\left[Q_i^n + u_n^2(V_i^n)^2\right]
$$

By successive conditioning and (B.3), there is a sequence $a_n \to 0$ such that

$$
E_i^n\left[(\eta_i^{n,1})^q\right] \leq K_q\Delta_n^{2\rho q+1/2-(\rho-1/4)\nu}a_n
$$

Analyzing $\eta_i^{n,2}$ with (B.2), (B.3), analyzing $\eta_i^{n,3}$ with Cauchy-Schwarz inequality, Markov's inequality, (B.6), we can get the following lemma:

LEMMA 2. *Assume (3.6), (B.1), assumption A-ν, A-γ, then $\exists\, a_n \to 0$ such that $\forall q \leq 4$,*

$$
E_i^n\left(\|\widehat{c}_i^n - \widehat{c}_i^{*n}\|^q\right) \leq K_q\left(a_n\Delta_n^{1/2-(\rho-1/4)\nu-(1-2\rho)q} + \Delta_n^{1/2}\right)
$$

B.3. Properties of spot estimator: II. continuous part.

B.3.1. *variables.* "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", said the Russian playwright Anton Chekhov. Define

$$\bar{C}_i^n = \frac{1}{\psi_n} \sum_{h=1}^{l_n-1} (\varphi_h^n)^2 \Delta_{i+h}^n C, \ C_t = \int_0^t c_s \, \mathrm{d}s$$

$$D_i^n = \bar{C}_i^n - c_i^n \Delta_n$$

$$\Gamma_h^n = \Gamma_{h,h}^n, \ \Gamma_{h,h'}^n = \frac{1}{\psi_n} \sum_{v=h\vee h'}^{h\wedge h'+l_n-1} (\varphi_{v-h+1}^n - \varphi_{n-h}^n)(\varphi_{v-h'+1}^n - \varphi_{v-h'}^n)\gamma_v^n$$

$$R_i^n = \widehat{Y}_i^{*n} - \Gamma_i^n$$

$$\zeta_i^n = \bar{Y}_i^{*n} \cdot \bar{Y}_i^{*n,\mathrm{T}} - \bar{C}_i^n - \Gamma_i^n$$

given $p \in \mathbb{N}^+$, define

$$\zeta(W,p)_i^n = \sum_{h=i}^{i+pl_n-1} \left[(\sigma_i^n \bar{W}_h^n) \cdot (\sigma_i^n \bar{W}_h^n)^\mathrm{T} - \bar{C}_h^n \right]$$

$$\zeta(X,p)_i^n = \sum_{h=i}^{i+pl_n-1} \left(\bar{X}_h^n \cdot \bar{X}_h^{n,\mathrm{T}} - \bar{C}_h^n \right)$$

$$\zeta(X,p)_i'^n = \sum_{h=i}^{i+pl_n-2} \sum_{h'=h+1}^{i+pl_n-1} \bar{X}_h^n \cdot \bar{X}_{h'}^{n,\mathrm{T}} \phi_1\left(\frac{h'-h}{l_n}\right)$$

$$\zeta(p)_i^n = \sum_{h=i}^{i+pl_n-1} \zeta_h^n$$

let $m(n,p) = \left\lfloor \frac{k_n}{(p+1)l_n} \right\rfloor$, $a(n,p,h) = 1 + h(p+1)l_n$, $b(n,p,h) = a(n,p,h)+pl_n$, then the estimation error of $\widehat{c}_i^{*n}$ can be decomposed as

$$(\text{B.7}) \qquad \beta_i^n \equiv \widehat{c}_i^{*n} - c_i^n = \xi_i^{n,0} + \xi_i^{n,1} + \xi_i^{n,2} + N(p)_i^n + M(p)_i^n$$

where

$$\xi_i^{n,0} \;\; = \;\; \frac{1}{k_n - l_n} \sum_{h=1}^{k_n - l_n + 1} c_{i+h}^n - c_i^n$$

$$\xi_i^{n,1} \;\; = \;\; \frac{1}{(k_n - l_n)\Delta_n} \sum_{h=1}^{k_n - l_n + 1} D_{i+h}^n$$

$$\xi_i^{n,2} \;\; = \;\; \frac{-1}{(k_n - l_n)\Delta_n} \sum_{h=1}^{k_n - l_n + 1} R_{i+h}^n$$

$$N(p)_i^n \;\; = \;\; \frac{1}{(k_n - l_n)\Delta_n} \Big(\sum_{h=0}^{m(n,p)-1} \zeta(1)_{i+b(n,p,h)}^n + \sum_{h=m(n,p)(p+1)l_n}^{k_n - l_n} \zeta_{i+1+h}^n \Big)$$

$$M(p)_i^n \;\; = \;\; \frac{1}{(k_n - l_n)\Delta_n} \sum_{h=0}^{m(n,p)-1} \zeta(p)_{i+a(n,p,h)}^n$$

B.3.2. *bounds on* $\|\xi_i^{n,r}\|$. By assumption A-ν, (B.1), (B.2)

(B.8)
$$\left\{ \begin{array}{rcl} \left\| E\left(\xi_i^{n,0} | \mathcal{F}_i^{(0),n} \right) \right\| & \leq & K k_n \Delta_n \\ E\left(\|\xi_i^{n,0}\|^q | \mathcal{F}_i^{(0),n} \right) & \leq & K_q (k_n \Delta_n)^{(q/2)\wedge 1}, \, q \geq 0 \end{array} \right.$$

combined with (3.6),

(B.9)
$$\left\{ \begin{array}{rcl} \left\| E\left(\xi_i^{n,1} | \mathcal{F}_i^{(0),n} \right) \right\| & \leq & K \Delta_n^{1/2} \\ E\left(\|\xi_i^{n,1}\|^q | \mathcal{F}_i^{(0),n} \right) & \leq & K_q \Delta_n^{[(q/2)\wedge 1]/2}, \, q \in \mathbb{N}^+ \end{array} \right.$$

By assumption A-γ,

(B.10)
$$\left\| E_i^n(\xi_i^{n,2}|) \right\| \leq K \Delta_n^{-1}$$
$$E_i^n\left(\|\xi_i^{n,2}\|^q \right) \leq \left\{ \begin{array}{l} K \, k_n^{-1/2}, \, q = 1; \\ K_q \left(k_n^{-q+1} + k_n^{-q} \Delta_n^{-q/2+1} \right), \, q = 2,3,4 \end{array} \right.$$

B.3.3. *estimates of* $\zeta(X,p)_i^n$ & $\zeta(X,p)_i'^n$. $\bar{C}_i^n \;\; = \;\; (\psi^n)^{-1} C_{i\Delta_n,(l_n-1)\Delta_n}^n$ in view of (B.4), hence by (B.1)

(B.11)
$$\|\bar{C}_i^n\| \leq K\Delta_n$$

According to (3.1) we have $\bar{X}_i^n = -\psi_n^{-1/2} \sum_{h=0}^{l_n-1} (\varphi_{h+1}^n - \varphi_h^n)(X_{i+h}^n - X_i^n)$, then by (B.2)

(B.12)
$$E\left(\|\bar{X}_i^n\|^q | \mathcal{F}_i^{(0),n} \right) \leq K_q \Delta_n^{q/2}$$

Adopt the argument for (5.21) in [8] in the multivariate setting, we have

$$
\text{(B.13)} \quad
\begin{aligned}
E\big(\bar{W}_h^n \bar{W}_{h'}^{n,\mathrm{T}}|\mathcal{F}_i^{(0),n}\big) &= \frac{l_n \Delta_n}{\psi_n}\phi_0\big(\tfrac{|h'-h|}{l_n}\big)\mathbf{I}_d + O_p(l_n^{-1/2}\Delta_n) \\
E\big(\|\bar{W}_h^n\|^{2m}|\mathcal{F}_i^{(0),n}\big) &= \Delta_n^m(2m-1)!! + O_p(l_n^{-1}\Delta_n^m), \ m \in \mathbb{N}^+
\end{aligned}
$$

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

$$
\text{(B.14)} \quad
\begin{aligned}
\zeta(W,p)_i^{n,jk}\zeta(W,p)_i^{n,lm} = &\; U_i^n(p)^{jk}U_i^n(p)^{lm} + S_i^n(p)^{jk}S_i^n(p)^{lm} \\
&- U_i^n(p)^{jk}S_i^n(p)^{lm} - U_i^n(p)^{lm}S_i^n(p)^{jk}
\end{aligned}
$$

By (B.5), (B.11), (B.13), (B.14), and through similar arguments in section 5.3 of [8] with a modification for multi-dimension, and exploit the connection between $\zeta(W,p)_i^n$ and $\zeta(X,p)_i^n$, in view of (4.2), we have the following lemma:

LEMMA 3. *Assume assumption A-ν, (B.1), l_n satisfies (3.6), then*

$$
\begin{aligned}
E\left(\|\zeta(X,p)_i^n\|^4|\mathcal{F}_i^{(0),n}\right) &\leq Kp^4\Delta_n^2 \\
\left\|E\left[\zeta(X,p)_i^n|\mathcal{F}_i^{(0),n}\right]\right\| &\leq Kp\Delta_n \\
E\left[\zeta(X,p)_i'^n|\mathcal{F}_i^{(0),n}\right] &= \frac{\theta^2 l_n}{\psi_n}(p\Phi_{01} - \Psi_{01})c_i^n + p^2\,O_p(\Delta_n^{1/4}) \\
E\left[\zeta(X,p)_i^{n,jk}\zeta(X,p)_i^{n,lm}|\mathcal{F}_i^{(0),n}\right] &= \frac{2\theta^4}{\psi_n^2}(p\Phi_{00} - \Psi_{00})\Sigma(c_i^n)^{jk,lm} \\
&\qquad + p^2\,O_p(\Delta_n^{5/4})
\end{aligned}
$$

B.3.4. *estimates of $\zeta(p)_i^n$.* For $i \leq h, h' \leq i + pl_n - 1$, by (B.5)

$$
\text{(B.15)} \qquad
\Gamma_{h,h'}^n = \frac{1}{\psi_n l_n}\phi_1\big(\frac{|h'-h|}{l_n}\big)\gamma_i + O_p(\Delta_n^{5/4})
$$

$$
\sum_{h=i}^{i+pl_n-2}\sum_{h'=h+1}^{i+pl_n-1}\Gamma_{h,h'}^{n,jk}\Gamma_{h,h'}^{n,lm} = \frac{1}{\psi_n^2}(p\Phi_{11} - \Psi_{11})\gamma_i^{n,jk}\gamma_i^{n,lm} + p^2\,O_p(\Delta_n^{5/4})
$$

Let $\xi_{h_1\cdots h_q}^{n,j_1\cdots j_q} = \prod_{v=1}^q\big(\bar{Y}_{h_v}^{n,j_v} - \bar{X}_{h_v}^{n,j_v}\big)$. By assumption A-$\gamma$,

$$
\text{(B.16)} \quad E\left[\big(\bar{Y}_h^{n,j} - \bar{X}_h^{n,j}\big)^q\big(\bar{Y}_{h'}^{n,k} - \bar{X}_{h'}^{n,k}\big)^r|\widetilde{\mathcal{F}}_{h\wedge h'-1}^n\right] =
$$

$$
\begin{cases}
0 & q+r=1 \\
\Gamma_{h,h'}^{n,jk} & q=r=1 \\
O_p\big(l_n^{-7/2}\big)\mathbb{1}_{\{|h-h'|\leq l_n\}} & q+r=3 \\
O_p(l_n^{-8}) & q+r=8
\end{cases}
$$

$$
E\left[\xi_{hhh'h'}^{n,jklm}|\widetilde{\mathcal{F}}_{h\wedge h'}^n\right] = \Gamma_h^{n,jk}\Gamma_{h'}^{n,lm} + \Gamma_{h,h'}^{n,jl}\Gamma_{h,h'}^{n,km} + \Gamma_{h,h'}^{n,jm}\Gamma_{h,h'}^{n,kl} + O_p\big(l_n^{-5}\big)
$$

then by (B.12), (B.16)

$$E(\zeta_h^n|\widetilde{\mathcal{F}}_h^n) = \bar{X}_h^n \bar{X}_h^{n\,\mathrm{T}} - \bar{C}_h^n, \ \ \|E_h^n(\zeta_h^n)\| \leq K\Delta_n^{3/2}, \ \ E_h^n(\|\zeta_h^n\|^4) \leq K\Delta_n^4$$

$$E\big(\zeta_h^{n,jk}\zeta_{h'}^{n,lm}|\widetilde{\mathcal{F}}_{h\wedge h'}^n\big) = \sum_{r=1}^3 \vartheta(r)_{h,h'}^{n,jk,lm}$$
$$+ \big(\bar{X}_h^{n,j}\bar{X}_h^{n,k} - \bar{C}_h^{n,jk}\big)\big(\bar{X}_{h'}^{n,l}\bar{X}_{h'}^{n,m} - \bar{C}_{h'}^{n,lm}\big)$$

where

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

Let $\Upsilon_{h,h'}^{n,jk,lm} = \Theta(\bar{X}_h^n\bar{X}_{h'}^{n,\mathrm{T}}, \Gamma_{h,h'}^n)^{jk,lm}$ in light of (4.2), then

$$E\left(\zeta_h^{n,jk}\zeta_{h'}^{n,lm}|\widetilde{\mathcal{F}}_{h\wedge h'}^n\right) = \big(\bar{X}_h^{n,j}\bar{X}_h^{n,k} - \bar{C}_h^{n,jk}\big)\big(\bar{X}_{h'}^{n,l}\bar{X}_{h'}^{n,m} - \bar{C}_{h'}^{n,lm}\big)$$
$$+ \Upsilon_{h,h'}^{n,jk,lm} + \Sigma(\Gamma_{h,h'}^n)^{jk,lm} + O_p\big(l_n^{-5} + (\|\bar{X}_h^n\| + \|\bar{X}_{h'}^n\|)l_n^{-7/2}\big)$$

hence

$$\zeta(p)_i^{n,jk}\zeta(p)_i^{n,lm} = \zeta(X,p)_i^{n,jk}\zeta(X,p)_i^{n,lm}$$
$$+ \sum_{h=i}^{i+pl_n-2}\sum_{h'=h+1}^{i+pl_n-1}\left[\Upsilon_{h,h'}^{n,jk,lm} + \Upsilon_{h,h'}^{n,lm,jk} + \Sigma(\Gamma_{h,h'}^n)^{jk,lm} + \Sigma(\Gamma_{h,h'}^n)^{lm,jk}\right]$$
$$+ \sum_{h=i}^{i+pl_n-1}\left[\Upsilon_{h,h'}^{n,jk,lm} + \Sigma(\Gamma_{h,h'}^n)^{jk,lm}\right] + p^2\,O_p(\Delta_n^{5/4})$$

then by (B.12), (B.15)

$$E\left[\zeta(p)_i^{n,jk}\zeta(p)_i^{n,lm}|\widetilde{\mathcal{F}}_i^n\right] = \zeta(X,p)_i^{n,jk}\zeta(X,p)_i^{n,lm} +$$
$$\frac{2}{\psi_n l_n}\Theta(\zeta(X,p)_i'^n, \gamma_i^n)^{jk,lm} + \frac{2}{\psi_n^2}(p\Phi_{11} - \Psi_{11})\,\Sigma(\gamma_i^n)^{jk,lm} + p^2\,O_p(\Delta_n^{5/4})$$

According to these results and lemma 3, one can get the following lemma

LEMMA 4.   *Assume assumption A-ν, A-γ, (B.1), l_n satisfies (3.6), then*

$$
\begin{aligned}
E[\zeta(p)_i^n | \widetilde{\mathcal{F}}_i^n] &= \zeta(X, p)_i^n \\
\| E_i^n [\zeta(p)_i^n] \| &\leq K p \Delta_n \\
E_i^n (\| \zeta(p)_i^n \|^q) &\leq K_q p^{\lfloor q/2 \rfloor \vee 1} \Delta_n^{q/2}, \quad q = 1, 2, 3, 4
\end{aligned}
$$

moreover

$$
\left| E_i^n \left[\zeta(p)_i^{n,jk} \zeta(p)_i^{n,lm} \right] - (p+1)\theta \Delta_n \, \Xi(c_i^n, \gamma_i^n; p)^{jk,lm} \right| \leq K \Delta_n^{5/4}
$$

where

$$
\text{(B.17)} \quad \Xi(x, z; p) = \frac{2\theta}{\phi_0(0)^2} \left[\frac{p\Phi_{00} - \Psi_{00}}{p+1} \Sigma(x) \right.
$$
$$
\left. + \frac{p\Phi_{01} - \Psi_{01}}{\theta^2(p+1)} \Theta(x, z) + \frac{p\Phi_{11} - \Psi_{11}}{\theta^4(p+1)} \Sigma(z) \right]
$$

Let $p \asymp \Delta_n^{-1/12}$, based on lemma 4,

$$
\begin{aligned}
\| E_i^n (M(p)_i^n) \| &\leq K \Delta_n^{1/2} \\
\| E_i^n (N(p)_i^n) \| &\leq K p^{-1} \Delta_n^{1/2}
\end{aligned}
$$

$$
\text{(B.18)} \quad E_i^n (\| M(p)_i^n \|^q) \leq \begin{cases} K_q \left(k_n \Delta_n^{1/2} \right)^{-q/2}, & q = 1, 2, 4 \\ K \left(k_n \Delta_n^{1/2} \right)^{-2}, & q = 3 \end{cases}
$$

$$
E_i^n (\| N(p)_i^n \|^q) \leq \begin{cases} K p (k_n \Delta_n^{1/2})^{-1} & q = 1 \\ K_q p^{-q/2} (k_n \Delta_n^{1/2})^{-q/2} & q = 2, 4 \\ K p^{-1} (k_n \Delta_n^{1/2})^{-2} & q = 3 \end{cases}
$$

B.3.5. *estimates of β_i^n.*   We need to define more variables:

$$
\text{(B.19)} \quad \begin{array}{ll} \zeta(p)_{i,h}^n = \zeta(p)_{i+a(n,p,h)}^n & A(p)_{i+v}^n = \sum_{h=v}^{v+pl_n-1} (c_{i+h}^n - c_i^n) \Delta_n \\ D(p)_{i+v}^n = \sum_{h=v}^{v+pl_n-1} D_{i+h}^n & R(p)_{i+v}^n = \sum_{h=v}^{v+pl_n-1} R_{i+h}^n \end{array}
$$

we have

TABLE 1
Estimates of ingredients

scaling properties	$E(\|\cdot\|^2 | \mathcal{F}_i^n)$	$\|E(\cdot | \mathcal{F}_i^n)\|$
$R(p)_h^n$	$p\Delta_n^{3/2}$	$p\Delta_n^{3/2}$
$D(p)_h^n$	$p\Delta_n^{3/2}$	$p\Delta_n$
$A(p)_{i+v}^n$	$p^2\Delta_n^2 \left(p\Delta_n^{-1/2} + v \right)$	$p\Delta_n^{3/2} \left(p\Delta_n^{-1/2} + v \right)$
$\zeta(p)_{i,h}^n$	$p\Delta_n$	$p\Delta_n$

Define

$$\alpha(p)_{i,h}^n = -R(p+1)_{i+a(n,p,h)}^n + D(p+1)_{i+a(n,p,h)}^n \\ + A(p+1)_{i+a(n,p,h)}^n + \zeta(p+1)_{i,h}^n$$

By table 1 and Cauchy-Schwarz inequality,

$$(\text{B.20}) \quad E_i^n\left(\left|\alpha(p)_{i,h}^{n,jk}\alpha(p)_{i,h}^{n,lm} - \zeta(p+1)_{i,h}^{n,jk}\zeta(p+1)_{i,h}^{n,lm}\right|\right) \\ \leq K\left(p^2\Delta_n^{5/4} + p^{3/2}\Delta_n^{3/2}v^{1/2}\right)$$

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

$$\left|E_i^n\left(\beta_i^{n,jk}\beta_i^{n,lm}\right) - (k_n\Delta_n^{1/2})^{-1}\Xi(c_i^n, \gamma_i)^{jk,lm}\right| = \sum_{r=1}^5 \mu_i^{n,r} + pO_p((k_n^2\Delta_n)^{-1})$$

where

$$\mu_i^{n,1} = \frac{1}{k_n^2\Delta_n^2} \sum_{h=0}^{m(n,p)-1} E_i^n\left(\left|\alpha(p)_{i,h}^{n,jk}\alpha(p)_{i,h}^{n,lm} - \zeta(p+1)_{i,h}^{n,jk}\zeta(p+1)_{i,h}^{n,lm}\right|\right)$$

$$\mu_i^{n,2} = \frac{1}{k_n^2\Delta_n^2} \sum_{h=0}^{m(n,p)-2} \sum_{h'=h+1}^{m(n,p)-1} \left|E_i^n\left[\alpha(p)_{i,h}^{n,jk}\alpha(p)_{i,h'}^{n,lm} + \alpha(p)_{i,h}^{n,lm}\alpha(p)_{i,h'}^{n,jk}\right]\right|$$

$$\mu_i^{n,3} = \frac{1}{k_n^2\Delta_n^2} \sum_{h=0}^{m(n,p)-1} E_i^n\left(\left|\zeta(p+1)_{i,h}^{n,jk}\zeta(p+1)_{i,h}^{n,lm}\right.\right. \\ \left.\left. -(p+2)\theta\Delta_n\,\Xi\left(c_{i+a(n,p,h)}^n, \gamma_{i+a(n,p,h)}^n; p+1\right)^{jk,lm}\right|\right)$$

$$\mu_i^{n,4} = \frac{(p+2)\theta}{k_n^2\Delta_n} \sum_{h=0}^{m(n,p)-1} \left|E_i^n\left[\Xi\left(c_{i+a(n,p,h)}^n, \gamma_{i+a(n,p,h)}^n; p+1\right)^{jk,lm}\right.\right. \\ \left.\left. -\Xi(c_i^n, \gamma_i^n; p+1)^{jk,lm}\right]\right|$$

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

Use table 1 and (B.20) to get bounds on $\mu_i^{n,r}$, $r = 1, 2, 3, 4, 5$; combine (B.7), (B.8), (B.9), (B.10), (B.18), we get the following lemma:

LEMMA 5.   *Assume (3.6), (B.1) and assumption A-ν, A-γ, given $p \in \mathbb{N}^+$,*

$$
\begin{aligned}
\|E_i^n(\beta_i^n)\| &\leq K k_n \Delta_n \\
E_i^n(\|\beta_i^n\|^q) &\leq \begin{cases} K_q\left[(k_n\Delta_n)^{(q/2)\wedge 1} + (k_n\Delta_n^{1/2})^{-q/2}\right], & q = 1, 2, 4 \\ K k_n \Delta_n, & q = 3 \end{cases}
\end{aligned}
$$

additionally

$$
\left| E_i^n\left(\beta_i^{n,jk}\beta_i^{n,lm}\right) - (k_n\Delta_n)^{-1/2}\Xi(c_i^n, \gamma_i^n)^{jk,lm} \right|
$$
$$
\leq K\left[k_n\Delta_n + p^{-1}(k_n\Delta_n^{1/2})^{-1}\right]
$$

B.4. Structure. Define

$$
\begin{aligned}
\lambda(x, z) &= \sum_{j,k,l,m=1}^d \partial^2_{jk,lm} g(x) \times \Xi(x, z)^{jk,lm} \\
\eta_i^n &= \lambda(\widehat{c}_i^n, \widehat{\gamma}_i^n) - \lambda(\widehat{c}_i^{*n}, \widehat{\gamma}_i^n)
\end{aligned}
$$

As $\Delta_n^{-1/4}|a_t^n - 1| < k_n\Delta_n^{3/4} \to 0$, letting $a_t^n = 1$ doesn't not affect the asymptotic analysis. By Cramér-Wold theorem, we can suppose g is $\mathbb{R}$-valued. We have

(B.21) $\Delta_n^{-1/4}\left[\widehat{S}(g)^n - S(g)\right] = \bar{S}^{n,0} + \bar{S}^{n,1} + \bar{S}(p)^{n,2} + \bar{S}^{n,3} + \bar{S}(p)^{n,4}$

where

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

$$
\bar{S}_t^{n,1} = k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\left[g(\widehat{c}_{ik_n}^n) - g(\widehat{c}_{ik_n}^{*n}) - (2k_n\Delta_n^{1/2})^{-1}\eta_{ik_n}^n\right]
$$

$$
\bar{S}(p)_t^{n,2} = k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\sum_{j,k=1}^d \partial_{jk}g(c_{ik_n}^n)\left[\sum_{r=0}^2 \xi(r)_{ik_n}^{n,jk} + N(p)_{ik_n}^{n,jk}\right]
$$

$$
\bar{S}_t^{n,3} = k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\left[g(\widehat{c}_{ik_n}^{*n}) - g(c_{ik_n}^n) - \sum_{j,k=1}^d \partial_{jk}g(c_{ik_n}^n)\beta_{ik_n}^{n,jk}\right.
$$
$$
\left. -(2k_n\Delta_n^{1/2})^{-1}\lambda(\widehat{c}_i^{*n}, \widehat{\gamma}_i^n)\right]
$$

$$
\bar{S}(p)_t^{n,4} = k_n\Delta_n^{3/4}\sum_{i=0}^{N_t^n-1}\sum_{j,k=1}^d \partial_{jk}g(c_{ik_n}^n) \times M(p)_{ik_n}^{n,jk}
$$

B.5. Asymptotic negligibility. First of all, we need to get bounds on $\partial^r g(\widehat{c}_i^n)$, $r \leq 3$, where $\partial^r g$ denotes the r-th order partial derivatives. Let $\overline{c}_i^n = (k_n - l_n)^{-1} \sum_{h=1}^{k_n - l_n + 1} c_i^n$ and $I_t^n = \{0, 1, \cdots, N_t^n - 1\}$, note that $|I_t^n| \asymp (k_n \Delta_n)^{-1}$, according to lemma 2, there is a sequence $a_n \to 0$ such that

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

Note $\widehat{c}_i^{*n} - \overline{c}_i^n = \xi_i^{n,1} + \xi_i^{n,2} + N(p)_i^n + M(p)_i^n$, by (B.9), (B.10), (B.18) and $\kappa < 3/4$, $E_i^n\big(\|\widehat{c}_i^{*n} - \overline{c}_i^n\|^4 \big) \leq K\Delta_n^{2\kappa - 1}$, so

$$\mathbb{E}\Big(\sup_{i \in I_n} \|\widehat{c}_i^{*n} - \overline{c}_i^n\| \Big) \leq K\Delta_n^{3\kappa - 2}$$

hence by (3.6) and Markov's inequality

$$(\text{B.22}) \qquad\qquad \sup_{i \in I_n} \|\widehat{c}_i^n - \overline{c}_i^n\| = o_p(1)$$

According to (B.1) and convexity, $\overline{c}_i^n \in \mathcal{S}$. By (B.22) $\widehat{c}_i^n \in \mathcal{S}^\epsilon$ if n is sufficiently large. Therefore by (3.5), in asymptotic analysis we can assume

$$(\text{B.23}) \qquad\qquad \|\partial^r g(\widehat{c}_i^n)\| \leq K, \quad \forall r = 0, 1, 2, 3, \forall i \in I_t^n$$

Through an almost identical argument for lemma 4.4 in [1],

$$(\text{B.24}) \qquad\qquad \overline{S}^{n,0} \xrightarrow{u.c.p.} 0$$

Define function g_n on $\mathcal{S}_d^+ \times \mathcal{S}_d^+$ as $g_n(x, z) = g(x) - (k_n \Delta_n^{1/2})^{-1}\xi(x, z)$ according to (B.1),

$$\|g_n(x, z) - g_n(y, z)\| \leq K\|x - y\|$$
$$+ K(k_n \Delta_n^{1/2})^{-1}\|x - y\| \left(\|x\|^2 + \|z\|^2 + \|x - y\|^2 + \|z\|\|x - y\| \right)$$

so $\|g_n(x, z) - g_n(y, z)\| \leq K\|x - y\|$ when n is sufficiently large. By lemma 2

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

Since $\rho > \frac{3 - \nu}{4(2 - \nu)}$, $1/4 - (\rho - 1/4)\nu - (1 - 2\rho) > 0$, we have the following lemma:

LEMMA 6.   *Assume assumption A-ν, A-γ, (B.23), (3.6) then*

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

Given $e_i^n \in \mathbb{R}^{d \times d}$, consider the process

$$\widetilde{S}_t^n = k_n \Delta_n^{3/4} \sum_{i=0}^{N_t^n-1} \sum_{j,k=1}^{d} \partial_{jk} g(c_{ik_n}^n) \times e_{ik_n}^{n,jk}$$

suppose e_i^n satisfies

(B.25)
$$\begin{cases} \|E_i^n(e_i^n)\| & \leq & K\Delta_n^{1/4} a_n \\ E_i^n(\|e_i^n\|^2) & \leq & K(k_n \Delta_n^{1/2})^{-1} b_n \end{cases}$$

where $a_n, b_n \to 0$. Since ∂g is bounded by (B.1),

$$\mathbb{E}\left(\sup_{s\in[0,t]} \|\widetilde{S}_s^n\|\right) \leq Kk_n\Delta_n^{3/4} \sum_{i=0}^{N_t^n-1} \mathbb{E}\left(\left\|E_{ik_n}^n(e_{ik_n}^n)\right\|\right)$$

$$+ Kk_n\Delta_n^{3/4}\mathbb{E}\left(\sup_{s\in[0,t]} \left\|\sum_{i=0}^{N_s^n-1}\left[e_{ik_n}^n - E_{ik_n}^n(e_{ik_n}^n)\right]\right\|\right)$$

by lemma 1,

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

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

$$\mathbb{E}\left(\sup_{s\in[0,t]} \|\widetilde{S}_s^n\|\right) \leq K\left(ta_n + \sqrt{tb_n}\right) \to 0$$

To show the asymptotic negligibility of $\bar{S}^{n,2}$, we need to show ξ_i satisfies (B.25) in each of the following 4 cases:

  (i) when $e_i^n = \xi_i^{n,0}$, by (B.8), $a_n = k_n\Delta_n^{3/4}$, $b_n = (k_n\Delta_n^{3/4})^2$;
  (ii) when $e_i^n = \xi_i^{n,1}$, by (B.9), $a_n = \Delta_n^{1/4}$, $b_n = k_n\Delta_n$;
  (iii) when $e_i^n = \xi_i^{n,2}$, by (B.10), $a_n = \Delta_n^{3/4}$, $b_n = \Delta_n^{1/2}$;
  (iv) when $e_i^n = N(p)_i^n$, by (B.18), $a_n = p^{-1}\Delta_n^{1/4}$, $b_n = p^{-1}$.

Hence we have the following lemma:

LEMMA 7.   *Assume assumption A-ν, A-γ, (B.23), (3.6), and let $p \asymp \Delta_n^{-1/12}$, then*

$$\overline{S}(p)^{n,2} \xrightarrow{u.c.p.} 0$$

Let $\chi_i^n = \widehat{\gamma}_i^n - \gamma_i^n$, by (B.2), the choice of m_n and Jensen's inequality

(B.26) $$E_i^n(\|\chi_i^n\|^q) \le K_q \Delta_n^{q/4}, \, q = 1, 2$$

Let $\eta_i'^n = \lambda(\widehat{c}_i^{*n}, \widehat{\gamma}_i^n) - \lambda(c_i^n, \gamma_i^n)$, then by (4.2), (4.3)

$$\|\eta_i'^n\| \le K \left(\|\beta_i^n\| + \|\chi_i^n\| + \|\beta_i^n\|^2 + \|\beta_i^n\|\|\chi_i^n\| + \|\chi_i^n\|^2 \right)$$

hence by lemma 5, (B.26), (3.6)

(B.27) $$E_i^n(\|\eta_i'^n\|^q) \le K_q (k_n \Delta_n^{1/2})^{-q/2}, \quad q = 1, 2$$

We can rewrite $\overline{S}^{n,3}$ as

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

where

$$G_t^n = k_n \Delta_n^{3/4} \sum_{i=0}^{N_t^n - 1} \left[s_{ik_n}^n + u_{ik_n}^n + E_{ik_n}^n \left(v_{ik_n}^n \right) \right]$$

$$H_t^n = k_n \Delta_n^{3/4} \sum_{i=0}^{N_t^n - 1} \left[v_{ik_n}^n - E_{ik_n}^n \left(v_{ik_n}^n \right) \right]$$

$$s_i^n = g(c_i^n + \beta_i^n) - g(c_i^n) - \sum_{j,k=1}^{d} \partial_{jk} g(c_i^n) \beta_i^{n,jk}$$

$$-\frac{1}{2} \sum_{j,k,l,m=1}^{d} \partial_{jk,lm}^2 g(c_i^n) \beta_i^{n,jk} \beta_i^{n,lm}$$

$$u_i^n = \frac{1}{2k_n \Delta_n^{1/2}} \left[\xi(c_i^n, \gamma_i^n) - \xi(\widehat{c}_i^{*n}, \widehat{\gamma}_i^n) \right]$$

$$v_i^n = \frac{1}{2} \sum_{j,k,l,m=1}^{d} \partial_{jk,lm}^2 g(c_i^n) \left[\beta_i^{n,jk} \beta_i^{n,lm} - (k_n \Delta_n^{1/2})^{-1} \Xi(c_i^n, \gamma_i^n)^{jk,lm} \right]$$

By lemma 5, (B.23), (B.27), if we let $p \asymp \Delta_n^{-12}$,

$$(\text{B.28}) \qquad \mathbb{E}\left(\sup_{s \in [0,t]} \|G_s^n\|\right) \leq Kt\left[k_n\Delta_n^{3/4} + (k_n\Delta_n^{2/3})^{-1}\right]$$

and

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

then lemma 1 implies

$$(\text{B.29}) \quad \mathbb{E}\left(\sup_{s \in [0,t]} \|H_s^n\|\right) \leq Kk_n\Delta_n^{3/4}\left(\sum_{i=0}^{N_t^n-1} \mathbb{E}(\|v_{ik_n}^n\|^2)\right)^{1/2}$$
$$\leq K\sqrt{t}\left[k_n\Delta_n^{3/4} + (k_n\Delta_n^{1/2})^{-1/2}\right]$$

According to (3.6), (B.28), (B.29), we have the following lemma:

LEMMA 8. *Assume assumption A-ν, A-γ, (B.23), (3.6) then*

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

B.6. Stable convergence in law to a continuous Itô semimartingale.
Recall (B.19), we can write

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

Let $\mathcal{H}(p)_{i,h}^n = \mathcal{F}_{ik_n+a(n,p,h)}^n$, by lemma 4,

$$\Delta_n^{-1/2} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \left\|E\left[\zeta(p)_{ik_n,h}^n | \mathcal{H}(p)_{i,h}^n\right]\right\|^2 \leq Ktp\Delta_n$$

Let $\Lambda(p)_{i,h}^n = \partial g(c_{ik_n}^n)\zeta(p)_{ik_n,h}^n$, N is a bounded martingale orthogonal to W or $N = W^l$ for some $l = 1, \cdots, d'$, and $\Delta N(p)_{i,h}^n = N_{ik_n+b(n,p,h)}^n - $

$N^n_{ik_n+a(n,p,h)}$. The following 4 statements about convergence in probability for any indices j, k, l, m can verify the conditions of theorem IX.7.28 in [24]:

$$\text{(B.30)} \qquad \Delta_n^{-1/4} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \left\| E\left[\Lambda(p)^n_{i,h} | \mathcal{H}(p)^n_{i,h} \right] \right\| \xrightarrow{\mathbb{P}} 0$$

$$\text{(B.31)} \qquad \Delta_n^{-1} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} E\left[\left\| \Lambda(p)^n_{i,h} \right\|^4 | \mathcal{H}(p)^n_{i,h} \right] \xrightarrow{\mathbb{P}} 0$$

$$\text{(B.32)} \quad \Delta_n^{-1/4} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \left\| E\left[\Lambda(p)^n_{i,h} \Delta N(p)^n_{i,h} | \mathcal{H}(p)^n_{i,h} \right] \right\| \xrightarrow{\mathbb{P}} 0$$

(B.33)

$$\Delta_n^{-1/2} \sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \partial_{jk}g(c^n_{ik_n}) \partial_{lm}g(c^n_{ik_n})^{\mathrm{T}} E\left[\zeta(p)^{n,jk}_{ik_n,h} \zeta(p)^{n,lm}_{ik_n,h} | \mathcal{H}(p)^n_{i,h} \right]$$
$$\xrightarrow{\mathbb{P}} \int_0^t \partial_{jk}g(c_s) \partial_{lm}g(c_s)^{\mathrm{T}} \Xi(c_s, \gamma_s; p)^{jk,lm}\, \mathrm{d}s$$

Under (B.23), one can verify (B.30), (B.31) by the second and third claims of lemma 4, respectively. The same argument as that for (5.58) in [8] leads to (B.32). By the last claim of lemma 4, the left-hand side of (B.33) equals

$$\sum_{i=0}^{N_t^n-1} \sum_{h=0}^{m(n,p)-1} \partial_{jk}g(c^n_{ik_n}) \partial_{lm}g(c^n_{ik_n})^{\mathrm{T}} \times$$
$$\Xi\left(c^n_{ik_n+a(n,p,h)}, \gamma^n_{ik_n+a(n,p,h)}; p \right)^{jk,lm} (p+1)l_n\Delta_n + tp\, O_p(\Delta_n^{1/4})$$

then (B.32) is verified by Riemann summation. By theorem IX.7.28 in [24] we have the following lemma:

LEMMA 9. *Assume assumption A-ν, A-γ, (B.23), (3.6), then for $\forall p \in \mathbb{N}^+$,*

$$\overline{S}(p)^{n,4} \xrightarrow{\mathcal{L}-s(f)} Z(p)$$

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

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

where $\widetilde{E}$ is the conditional expectation operator on the extended probability space and $\Xi(x, z; p)$ is defined in (B.17).

By (B.24), lemma 6, 7, 8, 9, and $\Xi(x, z; p) \to \Xi(x, z)$ as $p \to \infty$, we arrive at the asymptotic result in section 4.

REFERENCES

[1] JACOD, J., ROSENBAUM, M. (2013). Quarticity and other functionals of volatility: efficient estimation. *The Annals of Statistics* 41 (3), 1462-1484.

[2] LI, J., LIU, Y., XIU, D. (2019). Efficient estimation of integrated volatility functionals via multiscale Jackknife. *The Annals of Statistics* 47 (1), 156-176.

[3] MYKLAND, P., ZHANG, L. (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica* 77 (5), 1403-1445.

[4] LI, J., XIU, D. (2016). Generalized method of integrated moments for high-frequency data. *Econometrica* 84 (4), 1613-1633.

[5] AÏT-SAHALIA, Y., XIU, D. (2019). Principal component analysis of high-frequency data. *Journal of the American Statistical Association* 114 (525), 287-303.

[6] LI, J., TODOROV, V., TAUCHEN, G. (2017). Adaptive estimation of continuous-time regression models using high-frequency data. *Journal of Econometrics* 200, 36-47.

[7] JACOD, J., PODOLSKIJ, M., VETTER, M. (2010). Limit theorems for moving averages of discretized processes plus noise. *The Annals of Statistics* 38 (3), 1478-1545.

[8] JACOD, J., LI, Y., MYKLAND, P., PODOLSKIJ, M., VETTER, M. (2009). Microstructure noise in the continuous case: the preaveraging approach. *Stochastic processes and their applications* 119, 2249-2276.

[9] PODOLSKIJ, M., VETTER, M. (2009). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli* 15 (3), 634-658.

[10] JACOD, J., PROTTER, P. (2012). *Discretization of Processes*. Springer-Verlag Berlin Heidelberg.

[11] TODOROV, V., TAUCHEN, G. (2012). The realized Laplace transform of volatility. *Econometrica* 80 (3), 1105-1127.

[12] CHEN, D., MYKLAND, P., ZHANG, L. (2018). The five trolls under the bridge: principal component analysis with asynchronous and noisy high frequency data. *Working paper*.

[13] PODOLSKIJ, M., VETTER, M. (2009). Bipower-type estimation in a noisy diffusion setting. *Stochastic Processes and their Applications* 119, 2803-2831.

[14] REIß, M. (2011). Asymptotic equivalence for inference on the volatility from noisy observations. *The Annals of Statistics* 39 (2), 772-802.

[15] JACOD, J., MYKLAND, P. (2015). Microstructure noise in the continuous case: approximate efficiency of the adaptive pre-averaging method. *Stochastic Processes and their Applications* 125, 2910-2936.

[16] BIBINGER, M., HAUTSCH, N., MALEC, P., 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.

[17] ALTMEYER, R., BIBINGER, M. (2015). Functional stable limit theorems for quasi-efficient spectral covolatility estimators. *Stochastic Processes and their applications* 125, 4556-4600.

[18] ZHANG, L. (2006). Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. *Bernoulli* 12 (6), 1019-1043.

32 CHEN, R. Y.

[19] BARNDORFF-NIELSEN, O.E., HANSEN, P.R., LUNDE, A., 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.

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

[21] CLINET, S., POTIRON, Y. (2018). Efficient asymptotic variance reduction when estimating volatility in high frequency data. *Journal of Econometrics* 206, 103-142.

[22] CHRISTENSEN, K., KINNEBROCK, S., PODOLSKIJ, M. (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *Journal of Econometrics* 159, 116-133.

[23] CHEN, Y. (2019). Inference for volatility functionals of Itô semimartingales observed with jump and noise: time domain and adaptiveness. *Working paper*.

[24] JACOD, J., SHIRYAEV, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2ed. Springer-Verlag Berlin Heidelberg.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
CHICAGO, IL 60637
E-MAIL: yrchen@uchicago.edu