

# Nonparametric Estimation of the Leverage Effect: A Trade-off between Robustness and Efficiency\*

Ilze Kalnina<sup>†</sup>                      Dacheng Xiu<sup>‡</sup>  
Université de Montréal              Chicago Booth

This version: November 22, 2015

## Abstract

We consider two new approaches to nonparametric estimation of the leverage effect. The first approach uses stock prices alone. The second approach uses the data on stock prices as well as a certain volatility instrument, such as the CBOE volatility index (VIX) or the Black-Scholes implied volatility. The theoretical justification for the instrument-based estimator relies on a certain invariance property, which can be exploited when high frequency data is available. The price-only estimator is more robust since it is valid under weaker assumptions. However, in the presence of a valid volatility instrument, the price-only estimator is inefficient as the instrument-based estimator has a faster rate of convergence. We consider two empirical applications, in which we study the relationship between the leverage effect and the debt-to-equity ratio, credit risk, and illiquidity.

KEYWORDS: derivatives; VIX; implied volatility; high frequency data; spot correlation;

JEL CODES: G12, C22, C14.

---

\*We benefited much from discussions with Torben Andersen, Oleg Bondarenko, Stéphane Bonhomme, Francis Diebold, Kirill Evdokimov, Christian Hansen, Zhiguo He, Jean Jacod, Bryan Kelly, Jia Li, Yingying Li, Per Mykland, Benoit Perron, Mark Podolskij, Jeffrey Russell, Olivier Scaillet, Kevin Sheppard, George Tauchen, and Dan Wang. We thank seminar participants at Boston University, Duke University, Purdue University, Princeton University, Queen's University, University of Chicago, University of Pennsylvania, and University of Rochester, as well as participants of the 11th International Symposium on Econometric Theory and Applications, Frontiers in Financial Econometrics at Hitotsubashi University, Financial Econometrics Conference at the Toulouse School of Economics, Econometric Society European Meeting 2014, the SOFIE 2014 conference, the CIREQ Econometrics Conference 2014, the High-Frequency Data and High-Frequency Trading Conference at Chicago, the Econometric Study Group 2013 conference at Bristol, the 9th GNYEC, the First Conference in Econometric Theory at Universidad de San Andrés, the 5th Annual Modeling High-Frequency Data in Finance Conference, and the 11th World Congress of the Econometric Society.

<sup>†</sup>Département de sciences économiques and CIREQ, Université de Montréal. Address: C.P.6128, succ. centre-ville, Montréal, QC H3C 3J7, Canada. E-mail address: [ilze.kalnina@umontreal.ca](mailto:ilze.kalnina@umontreal.ca). Kalnina gratefully acknowledges financial support from *Institut de finance mathématique de Montréal*, SSHRC, and FQRSC.

<sup>‡</sup>Booth School of Business, University of Chicago. Address: 5807 S Woodlawn Avenue, Chicago, IL 60637, USA. E-mail address: [dacheng.xiu@chicagobooth.edu](mailto:dacheng.xiu@chicagobooth.edu). Xiu gratefully acknowledges financial support from the FMC Faculty Scholar Fund at the University of Chicago Booth School of Business.

# 1 Introduction

One of the most important empirical stylized facts about the volatility is the leverage effect, which refers to the generally negative correlation between an asset return and its volatility changes. The term “leverage” dates back to an early influential economic hypothesis of [Black \(1976\)](#) that explains this correlation using the debt-to-equity ratio, a common financial leverage measure. The estimation of the leverage effect is challenging because volatility is not observable.

We develop two qualitatively different approaches to nonparametric estimation of the leverage effect using high frequency data. We also study the empirical relationship between the leverage effect and the debt-to-equity ratio. Our results extend the large body of research that has used high frequency data fruitfully to estimate volatility measures of asset returns, see, e.g., [Andersen, Bollerslev, Diebold, and Labys \(2003\)](#). These methods are now commonly used in economics and finance, see, e.g., [Patton and Verardo \(2012\)](#), [Bandi and Renò \(2014\)](#), [Bollerslev, Li, and Todorov \(2015\)](#), and [Segal, Shaliastovich, and Yaron \(2015\)](#).

Our first approach to the leverage effect estimation only uses observations on the asset prices. It is an analog of the low frequency approach that has been common since [Black \(1976\)](#): one first conducts preliminary estimation of the volatility over small windows, and then computes the correlation between returns and the increments of the estimated volatility. However, in general there is an errors-in-variables bias associated with the preliminary estimation of volatility. We propose an estimator that corrects the biases due to the pre-estimation of volatility, and is valid for a very general class of semimartingales. We call this estimator the Price-only Realized Leverage (*PRL*). The rate of convergence of the *PRL* estimator is  $n^{1/4}$ .

Our second approach is to replace the preliminary estimation of volatility with high frequency observations of certain financial derivatives. It is well known that financial derivatives contain useful information about volatility. However, the implied volatilities from these derivatives are at best a proxy for the actual volatility as they are contaminated by risk premia. To purge the impact of risk premia, additional assumptions are necessary to model this contamination, which link the risk-neutral dynamics with the objective dynamics. For this purpose, we provide one such condition, that is, [Assumption 2](#) below, which allows the estimation of the leverage effect using high-frequency data on certain volatility instruments, such as the VIX or Black-Scholes implied volatility.<sup>1</sup> We call

---

<sup>1</sup>Since September 22, 2003, the VIX (the volatility index of CBOE) has been constructed by the CBOE using a weighted portfolio of options:  $(VIX_t/100)^2 = \frac{2e^{r_t, \tau}}{\tau} \left( \int_0^{e^{f_t, \tau}} \frac{P(\tau, x)}{x^2} dx + \int_{e^{f_t, \tau}}^{\infty} \frac{C(\tau, x)}{x^2} dx \right)$ , where  $P(\tau, x)$  and  $C(\tau, x)$  are put and call options with time-to-maturity  $\tau$  and strike  $x$ , and  $f_{t, \tau}$  is the log price of forward contracts, see, e.g., [Carr and Wu \(2009\)](#).

the resulting estimator the Instrument-based Realized Leverage (*IRL*). We develop the asymptotic theory for this estimator, and show that it has a faster ( $n^{1/2}$ ) rate of convergence, which in practice allows estimation of the leverage effect on relatively short windows such as one month or one quarter. Estimation of the leverage effect at such short time periods is important as it allows linking the time series of the estimated leverage effect with the time series of economic and financial variables.

The two estimators we develop are complementary to each other and have their own advantages and disadvantages. The *IRL* is more efficient in the sense that it has a faster rate of convergence, but it requires (i) the availability of the data on a volatility instrument, and (ii) that Assumption 2 adequately describes the data at hand. In Section 4.1, we argue that this assumption holds in a reasonably broad class of models. We stress that in particular it holds for various popular parametric models used in the derivative pricing literature. We provide a simple high-frequency Durbin-Wu-Hausman test that allows one to evaluate the validity of Assumption 2, see Durbin (1954), Wu (1973), and Hausman (1978). We also discuss settings where this assumption does not hold, and illustrate the impact of that on the *IRL*. The advantage of the *PRL* is that it does not rely on the availability of the additional data or Assumption 2. The cost of this robustness is a substantially lower precision. The precision can of course be increased by using longer time intervals, and we find that the *PRL* becomes practical if used over time periods of multiple years of data.<sup>2</sup> Over 11 years, the standard error for the leverage effect of the S&P 500 is 0.073.

We illustrate the finite sample performance of the estimators in Monte Carlo simulations and two empirical applications. In the Monte Carlo section, we consider the estimation of the leverage effect in models where Assumption 2 holds, and in a model where this assumption fails. Higher precision of the *IRL* makes it possible to investigate how the leverage effect changes over time, which is important for financial applications. We provide a time series of monthly leverage effects of the S&P 500 index using the VIX as a volatility instrument.<sup>3</sup> We also conduct a time series regression with the estimated leverage effect and important financial indicators, such as the credit risk, illiquidity, and the debt-to-equity ratio. Overall, we find that the leverage effect of the S&P 500 index tends to be stronger in bad times. This finding is consistent with, e.g., Bandi and Renò (2012) who document that the leverage effect of the S&P 500 is stronger when volatility is higher. In the second empirical application, we use the *PRL* estimator to document a significant relationship between the

---

<sup>2</sup>The finding that the price-only estimator requires multiple years of data for precise estimation appears to be in line with earlier empirical analysis by Aït-Sahalia, Fan, and Li (2013) who use 4 years of 1-minute data on the S&P 500 futures.

<sup>3</sup>We also use the Black-Scholes implied volatility constructed from intraday S&P 500 options in an earlier draft. The two time series of estimates share a similar pattern.

cross-section of the leverage effect of individual firms and the debt-to-equity ratio. These empirical analyses illustrate the strengths of both estimators and the trade-offs between them. Both analyses support the leverage hypothesis of [Black \(1976\)](#), while also suggesting that the debt-to-equity ratio is likely not the only determinant of the financial leverage.

For more than two decades, parametric models have been used to capture the leverage effect of daily stock returns. For example, the popular EGARCH model of [Nelson \(1991\)](#) is motivated by the inability of the standard GARCH models to capture the leverage effect. Many papers have also estimated the leverage effect in parametric stochastic volatility models. Such models assume a constant leverage effect, and their estimation involves either the Markov Chain Monte Carlo algorithm or particle filters, see, e.g., [Jacquier, Polson, and Rossi \(2004\)](#), [Yu \(2005\)](#), [Pitt, Malik, and Doucet \(2014\)](#), and [Eraker \(2004\)](#), or moment- or likelihood-based approaches, see, e.g., [Harvey and Shephard \(1994\)](#), [Pan \(2002\)](#), and [Aït-Sahalia and Kimmel \(2007\)](#). These parametric leverage effect estimates depend on the specified volatility dynamics. In contrast, our framework is more agnostic about the dynamics of the volatility.

The use of the derivative information to estimate the leverage effect has been previously considered by a few papers. Most closely related paper is an empirical study by [Andersen, Bondarenko, and Gonzalez-Perez \(2015\)](#). Although authors mainly focus on developing an alternative volatility proxy “Corridor VIX”, they also use the *IRL* estimator in the empirical study. [Bollerslev, Sizova, and Tauchen \(2012\)](#) calculate the leverage effect as a correlation between returns and change of the VIX. Their estimates using the VIX are substantially more stable than the estimates using absolute returns as a volatility proxy in [Bollerslev, Litvinova, and Tauchen \(2006\)](#). Among the papers that propose fully parametric estimators, [Aït-Sahalia and Kimmel \(2007\)](#) also use the VIX, whereas [Bakshi, Cao, and Chen \(1997\)](#), [Pan \(2002\)](#), [Eraker \(2004\)](#), and [Broadie, Chernov, and Johannes \(2007\)](#) use S&P 500 options.

Several papers are related to the *PRL* estimator. [Aït-Sahalia, Fan, and Li \(2013\)](#) note that empirically the correlation between returns and changes of the estimated volatility from high frequency data is close to zero. They call this phenomenon “the leverage effect puzzle.” In the parametric framework of the Heston model, they show that naive correlation estimator is biased. [Wang and Mykland \(2014\)](#), [Vetter \(2012b\)](#), and [Vetter \(2012a\)](#) provide estimators for the integrated covariation between the returns and their volatilities as well as for the integrated volatility of volatility. A correlation-type combination of their estimators provides a nonparametric equivalent to the bias-correction of [Aït-Sahalia, Fan, and Li \(2013\)](#), and converges to a certain volatility-weighted leverage

effect measure.<sup>4</sup> However, empirically this estimator produces insignificant estimates of the leverage effect with 11 years of 5 minute S&P 500 returns. In contrast, our *PRL* gives significant estimates in the same setting. This increase in precision of the *PRL* estimator is not surprising because it is constructed similarly to the efficient quarticity estimator in [Jacod and Rosenbaum \(2013\)](#), by aggregating a sequence of local estimates. To derive the asymptotic distribution of the *PRL*, we prove a general central limit theorem that significantly extends the results by [Jacod and Rosenbaum \(2013\)](#) and [Aït-Sahalia and Jacod \(2014\)](#). This theoretical result is of own interest. [Bandi and Renò \(2012\)](#), [Bandi and Renò \(2014\)](#), and [Aït-Sahalia, Fan, Laeven, Wang, and Yang \(2015\)](#) estimate related quantities, which are however different from the integrated leverage effect.

Both *PRL* and *IRL* are related to the literature on statistical inference based on preliminary estimation of spot variances and covariances. This literature dates back to [Comte and Renault \(1998\)](#) and [Kristensen \(2010\)](#), with [Jacod and Rosenbaum \(2013\)](#) being a more recent contribution. The error from this preliminary estimation is asymptotically negligible if the time span diverges sufficiently fast, see, e.g., [Bandi and Renò \(2011\)](#) and [Li and Patton \(2015\)](#). Our asymptotic approximation does not use a diverging time span and therefore requires taking into account the bias due to the preliminary estimation of variances and covariances. Empirically, we find that the effect of the bias correction for the *PRL* is sizeable even when the time span is 11 years.

The rest of the paper is organized as follows. Section 2 describes the model and the quantities of interest. Section 3 presents the *PRL* estimator and the associated asymptotic theory. Section 4 presents the *IRL* estimator and its asymptotic properties. Section 5 presents a Durbin-Wu-Hausman specification test. Section 6 provides Monte Carlo evidence and Section 7 presents empirical findings. Section 8 concludes. The Appendix and the Online Appendix contain the proofs.

## 2 Leverage Effect in Continuous Time

We work in a general nonparametric framework that allows for potential jumps in prices and volatility. This framework is commonly used in high frequency econometrics, see, e.g., [Aït-Sahalia and Jacod \(2014\)](#).

**Assumption 1.** *Suppose that  $X$ , the logarithm of the underlying asset price, follows an Itô semi-*

---

<sup>4</sup>In an earlier draft, we have provided the joint central limit theorem results for these components in a more general setting with jumps, and developed the asymptotic distribution of the estimator of this volatility-weighted leverage effect measure, *VWIL* (see equation (5)).

martingale,

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{E}} \delta(s, z) \mu(ds, dz), \quad (1)$$

and its spot variance, denoted by  $c_t = \sigma_t^2$ , follows

$$c_t = c_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s d\tilde{W}_s + \int_0^t \int_{\mathbb{E}} \tilde{\delta}(s, z) \mu(ds, dz), \quad (2)$$

where  $\mu(ds, dz)$  is a Poisson random measure on  $\mathbb{R}_+ \times \mathbb{E}$ .  $\mathbb{E}$  is an arbitrary Polish space with a  $\sigma$ -finite and infinite measure  $\lambda$  having no atoms, with an intensity measure  $\nu(ds, dz) = ds \otimes \lambda(dz)$ .  $W$  and  $\tilde{W}$  are standard Brownian motions. The correlation between  $W_s$  and  $\tilde{W}_s$  is  $\rho_s$ . Moreover,  $|\delta(\omega, t \wedge \tau_m(\omega), z)| \wedge 1 \leq \Gamma_m(z)$  and  $|\tilde{\delta}(\omega, t \wedge \tau_m(\omega), z)| \wedge 1 \leq \Gamma_m(z)$ , for all  $(\omega, t, z)$ , where  $(\tau_m)$  is a localizing sequence of stopping times and, for some  $r \in [0, 1]$ , the function  $\Gamma_m$  on  $\mathbb{E}$  satisfies  $\int_{\mathbb{E}} \Gamma_m(z)^r \lambda(dz) < \infty$ , i.e., the jumps are summable. In addition,  $b_s$  and  $\tilde{b}_s$  are locally bounded, and  $c_s^{(c)} = \tilde{\sigma}_s^2$  is càdlàg and locally bounded. For any  $s \in [0, t]$ ,  $c_s$ ,  $c_{s-}$ ,  $c_s^{(c)}$ , and  $c_{s-}^{(c)}$  are almost surely positive. Also,  $[X, X]_s^c \neq 0$  and  $[c, c]_s^c \neq 0$  hold almost surely, for each  $s$ .<sup>5</sup> Finally, we assume that  $\rho_s$  is càdlàg, and that  $|\rho_s|$  and  $|\rho_{s-}|$  are almost surely smaller than 1.

What is an appropriate measure of the leverage effect in this general framework? To motivate, note that in the special case of the popular Heston model, the leverage effect is usually associated with a parameter  $\rho$ , which equals

$$\rho = \lim_{\Delta \rightarrow 0} \text{Corr}(c_{s+\Delta} - c_s, X_{s+\Delta} - X_s), \quad \forall s \in [0, t]. \quad (3)$$

It also coincides with the correlation between the two Brownian motions of the Heston model, which is assumed constant over time. In general, this correlation  $\rho_s$  varies over time. We hence call it the spot leverage effect. Then, a natural measure of the leverage effect over the interval  $[0, t]$  is a scaled integral of the spot leverage over time, namely, the integrated leverage effect,

$$IL_t = \frac{1}{t} \int_0^t \rho_s ds. \quad (4)$$

Moreover, we can represent  $\rho_s$  alternatively as

$$\rho_s \equiv \frac{[X, c]_s^{c, \prime}}{\sqrt{[X, X]_s^{c, \prime}} \sqrt{[c, c]_s^{c, \prime}}}$$

where  $\prime$  denotes a derivative with respect to time.

Our  $IL$  measure is invariant to nonlinear transformations. In other words, for any smooth and

---

<sup>5</sup>The superscript  $c$  denotes the continuous part of the process, and  $[X, X]_s$  is the quadratic variation of  $X$  over  $[0, s]$ .

monotone increasing functions  $f$  and  $g$  with non-vanishing derivatives, by Itô's lemma, we have

$$\frac{[X, c]_s^{c'}}{\sqrt{[X, X]_s^{c'}}\sqrt{[c, c]_s^{c'}}} = \frac{[f(X), g(c)]_s^{c'}}{\sqrt{[f(X), f(X)]_s^{c'}}\sqrt{[g(c), g(c)]_s^{c'}}}.$$

As a result, we have

$$\int_0^t \frac{[X, c]_s^{c'}}{\sqrt{[X, X]_s^{c'}}\sqrt{[c, c]_s^{c'}}} ds = \int_0^t \frac{[f(X), g(c)]_s^{c'}}{\sqrt{[f(X), f(X)]_s^{c'}}\sqrt{[g(c), g(c)]_s^{c'}}} ds.$$

An important consequence is that variance-based leverage effect coincides with volatility-based leverage effect.

An alternative leverage effect measure can be defined as the scaled continuous part of the quadratic covariation between  $c$  and  $X$ ,

$$VWIL_t = \frac{[c, X]_t^c}{\sqrt{[X, X]_t^c}\sqrt{[c, c]_t^c}} = \frac{\int_0^t \rho_s \sigma_s \tilde{\sigma}_s ds}{\sqrt{\int_0^t \sigma_s^2 ds} \sqrt{\int_0^t \tilde{\sigma}_s^2 ds}}. \quad (5)$$

This leverage effect measure depends also on the path of the spot volatility and the volatility of volatility, hence we name it the Volatility-Weighted Integrated Leverage effect ( $VWIL$ ). This measure is only invariant to linear transformations.

While we do allow for general price and volatility jumps, they do not contribute to our definition of leverage. We choose not to include them for the following reasons. First, the  $IL$  is based on the spot correlation, which is an intuitive generalization of the  $\rho$  parameter in the Heston model, and which endows the  $IL$  measure with invariance to smooth transformations. However, the spot correlation is not well defined without the exclusion of jumps.<sup>6</sup> Second, Andersen, Bondarenko, and Gonzalez-Perez (2015) point out that the VIX contains artificial jumps due to its implementation by the CBOE. The  $IRL$  using the VIX is robust to the artificial jumps, because it only depends on the continuous part. Third, we find in our empirical study that truncating off the jump component makes the estimates more stable. Finally, Bollerslev, Kreschmer, Pigorsch, and Tauchen (2009) find that the leverage effect works primarily through the continuous components. We point out that the use of our method does not require taking a stand on the importance of jumps, and can be viewed as the estimation of the continuous part of the total leverage effect.<sup>7</sup>

<sup>6</sup>Only the continuous part of the quadratic variation is absolutely continuous and differentiable almost everywhere.

<sup>7</sup>This interpretation is analogous to the estimation of the continuous part of beta in, e.g., Reiß, Todorov, and Tauchen (2015) and Aït-Sahalia, Kalnina, and Xiu (2014), and the principal component analysis of the continuous spot covariance matrix in Aït-Sahalia and Xiu (2014).

### 3 Estimation using Price Observations alone

The current section introduces the Price-only Realized Leverage (*PRL*), a nonparametric leverage effect estimator that uses price observations alone, and presents its asymptotic distribution. The *PRL* estimator uses preliminary estimation of volatility, and it is based on an aggregation of spot correlations between the returns and the estimated volatility increments.

To obtain preliminary estimates of spot volatility, we use an additional set of smaller and overlapping blocks, each containing  $l_n$  observations, within each larger non-overlapping block of size  $k_n$ . We estimate the unobservable spot variance with the truncated realized variance over each smaller block,

$$\hat{\sigma}_i^2 = \frac{1}{l_n \Delta_n} \sum_{j=0}^{l_n-1} (\Delta_{i+j}^n X)^2 1_{\{|\Delta_{i+j}^n X| \leq u_n\}},$$

with the threshold  $u_n$  satisfying conditions of Theorem 1 below. The use of overlapping blocks of size  $l_n$  is motivated by the fact that the non-overlapping alternative would be less efficient in the sense of having a larger asymptotic variance, as shown by Aït-Sahalia and Jacod (2014) for the estimation of the *VWIL*. Meanwhile, the use of overlapping blocks of size  $k_n$  would not improve the efficiency of the *PRL*, similar in spirit to the estimator of Jacod and Rosenbaum (2013).

To simplify the exposition, we denote

$$C_s = \begin{pmatrix} c_s & c_s^{(X,c)} \\ c_s^{(X,c)} & c_s^{(c)} \end{pmatrix} \equiv \begin{pmatrix} [X, X]_s^{c'} & [X, \sigma^2]_s^{c'} \\ [X, \sigma^2]_s^{c'} & [\sigma^2, \sigma^2]_s^{c'} \end{pmatrix}, \quad (6)$$

and notice that it implies  $c_s^{(X,c)} = \rho_s c_s^{1/2} (c_s^{(c)})^{1/2}$ . We can approximate each element of this matrix at  $s = ik_n \Delta_n$  as follows:<sup>8</sup>

$$\begin{aligned} \hat{c}_{ik_n+1} &= \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n-1} (\Delta_{ik_n+1+j}^n X)^2 1_{\{|\Delta_{ik_n+1+j}^n X| \leq u_n\}}, \\ \hat{c}_{ik_n+1}^{(X,c)} &= \frac{1}{l_n k_n \Delta_n} \sum_{j=0}^{k_n-1} \left\{ (\hat{\sigma}_{ik_n+1+j+l_n}^2 - \hat{\sigma}_{ik_n+1+j}^2) \sum_{l=0}^{2l_n-1} (\Delta_{ik_n+1+j+l}^n X) 1_{\{|\Delta_{ik_n+1+j+l}^n X| \leq u_n\}} \right\}, \\ \hat{c}_{ik_n+1}^{(c)} &= \frac{3}{2l_n k_n \Delta_n} \sum_{j=0}^{k_n-1} \left\{ (\hat{\sigma}_{ik_n+1+j+l_n}^2 - \hat{\sigma}_{ik_n+1+j}^2)^2 - \frac{4}{l_n} (\hat{\sigma}_{ik_n+1+j}^2)^2 \right\}, \end{aligned}$$

therefore the estimator of the spot leverage effect is:

$$\hat{\rho}_{ik_n+1} = \frac{\hat{c}_{ik_n+1}^{(X,c)}}{(\hat{c}_{ik_n+1})^{1/2} (\hat{c}_{ik_n+1}^{(c)})^{1/2}}.$$

<sup>8</sup>We use  $\hat{c}$  to denote the estimator of  $\sigma^2$  based on a block of size  $k_n$ , and use  $\hat{\sigma}^2$  when a block of size  $l_n$  is used.



Our  $PRL$  can be constructed as follows:

$$PRL_t = \frac{k_n \Delta_n}{t} \sum_{i=0}^{N_t^n} \left( \hat{\rho}_{ik_n+1} - \frac{1}{2k_n \sqrt{\Delta_n}} \left( -\frac{5}{\beta} \frac{\hat{c}_{ik_n+1}^{3/2} \hat{c}_{ik_n+1}^{(X,c)}}{\left(\hat{c}_{ik_n+1}^{(c)}\right)^{3/2}} - \frac{151\beta}{120} \frac{\hat{c}_{ik_n+1}^{(X,c)}}{\hat{c}_{ik_n+1}^{1/2} \left(\hat{c}_{ik_n+1}^{(c)}\right)^{1/2}} + \frac{36}{\beta} \frac{\hat{c}_{ik_n+1}^{(X,c)} \hat{c}_{ik_n+1}^{7/2}}{\left(\hat{c}_{ik_n+1}^{(c)}\right)^{3/2}} + \frac{9}{\beta} \frac{\hat{c}_{ik_n+1}^{(X,c)} \hat{c}_{ik_n+1}^{3/2}}{\left(\hat{c}_{ik_n+1}^{(c)}\right)^{1/2}} + \frac{453\beta}{280} \frac{\hat{c}_{ik_n+1}^{(X,c)} \left(\hat{c}_{ik_n+1}^{(c)}\right)^{1/2}}{\hat{c}_{ik_n+1}^{1/2}} \right) \right), \quad (7)$$

where  $N_t^n = \lfloor (t/\Delta_n - 2l_n + 2)/k_n \rfloor - 1$  and  $\beta = l_n \sqrt{\Delta_n}$ . The component after  $\hat{\rho}_{ik_n+1}$  corrects an asymptotic bias. The next theorem presents the asymptotic distribution of the  $PRL$  estimator.

**Theorem 1.** *Suppose Assumption 1 holds. In addition,  $\sigma_t$  is a continuous Itô semimartingale. Suppose  $\beta = l_n \sqrt{\Delta_n} \in (0, \infty)$ ,  $k_n^3 \Delta_n^2 \rightarrow \infty$ , and  $k_n^4 \Delta_n^3 \rightarrow 0$ . Moreover, suppose  $u_n \asymp \Delta_n^\varpi$  with  $0 \leq r < 1/3$  and  $5/(12 - 6r) \leq \varpi < 1/2$ . Then,*

$$\Delta_n^{-1/4} (PRL_t - IL_t) \xrightarrow{L-\varpi} \mathcal{Z} \sqrt{V_t^{PRL}},$$

where  $\mathcal{Z}$  is a standard normal random variable defined on the extension of the original probability space, and the variance  $V_t^{PRL}$  is given by

$$V_t^{PRL} = \frac{1}{t^2} \int_0^t \left( \frac{23\beta}{15} + \frac{8}{3\beta} \frac{c_s^2}{c_s^{(c)}} - \frac{2}{\beta} \frac{c_s (c_s^{(X,c)})^2}{(c_s^{(c)})^2} + \frac{12}{\beta^3} \frac{c_s^3 (c_s^{(X,c)})^2}{(c_s^{(c)})^3} + \frac{\beta}{21} \frac{(c_s^{(X,c)})^2}{c_s c_s^{(c)}} \right) ds.$$

Similarly, we can construct a consistent estimator of  $V_t^{PRL}$  using local estimators of  $c_{ik_n+1}$ ,  $c_{ik_n+1}^{(X,c)}$ , and  $c_{ik_n+1}^{(c)}$  as follows:

$$\hat{V}_t^{PRL} = \frac{k_n \Delta_n}{t^2} \sum_{i=0}^{N_t^n} \left( \frac{23\beta}{15} + \frac{8}{3\beta} \frac{\hat{c}_{ik_n+1}^2}{\hat{c}_{ik_n+1}^{(c)}} - \frac{2}{\beta} \frac{\hat{c}_{ik_n+1} \left(\hat{c}_{ik_n+1}^{(X,c)}\right)^2}{\left(\hat{c}_{ik_n+1}^{(c)}\right)^2} + \frac{12}{\beta^3} \frac{\hat{c}_{ik_n+1}^3 \left(\hat{c}_{ik_n+1}^{(X,c)}\right)^2}{\left(\hat{c}_{ik_n+1}^{(c)}\right)^3} + \frac{\beta}{21} \frac{\left(\hat{c}_{ik_n+1}^{(X,c)}\right)^2}{\hat{c}_{ik_n+1} \hat{c}_{ik_n+1}^{(c)}} \right). \quad (8)$$

Appendix A contains the proof of Theorem 1. Note that due to the preliminary estimation of volatility, the rate of convergence of the  $PRL$  estimator is slower than the usual  $\sqrt{n}$  rate of convergence.

To prove Theorem 1, we first prove a more general result about estimating  $\int_0^t g(C_s) ds$  for any smooth function  $g(\cdot)$ ; see Theorem A1 in Appendix A. This theorem extends Jacod and Rosenbaum (2013) who estimate  $\int_0^t g(c_s) ds$ . The spot matrix  $C_s$  in (6) not only contains the spot volatility  $c_s$  of the observable process  $X$ , but also volatilities of unobservable processes. Therefore, estimating integrals of functionals of  $C_s$  requires substantially more involved methods and first-order bias corrections. Theorem 1 is related to Wang and Mykland (2014), Vetter (2012b), and Chapter 8 of Aït-Sahalia and Jacod (2014), in that our  $\hat{c}_{ik_n+1}^{(X,c)}$  and  $\hat{c}_{ik_n+1}^{(c)}$ , for each block  $[(ik_n+1)\Delta_n, (i+1)k_n\Delta_n]$ , are similar to their estimators on  $[0, t]$ .

The above theorem assumes that there are no volatility jumps. Volatility jumps do not affect the estimation of the covariation between the continuous components of returns and volatilities,  $c_{ik_n+1}^{(X,c)}$ , because these jumps do not co-vary with the continuous component of returns, see, e.g., [Aït-Sahalia, Fan, Laeven, Wang, and Yang \(2015\)](#). However, volatility jumps affect the estimation of  $c_{ik_n+1}^{(c)}$ , because its estimator would converge to the total quadratic variation of the volatility process. To ensure a fair comparison between the *IRL* and *PRL* estimates in the empirical work, we suggest identifying volatility jumps using the jumps of the VIX, a strategy previously adopted by [Todorov and Tauchen \(2011\)](#) for the estimation of the volatility jump activity index. We thus modify the above  $\hat{c}_{ik_n+1}^{(c)}$  as follows,

$$\tilde{c}_{ik_n+1}^{(c)} = \frac{3}{2l_n k_n \Delta_n} \sum_{j=0}^{k_n-1} \left\{ (\hat{\sigma}_{ik_n+1+j+l_n}^2 - \hat{\sigma}_{ik_n+1+j}^2)^2 - \frac{4}{l_n} (\hat{\sigma}_{ik_n+1+j}^2)^2 \right\} 1_{\{|Z_{ik_n+1+j+l_n} - Z_{ik_n+1+j}| \leq v_n\}},$$

where  $Z$  is the VIX and  $v_n \asymp (l_n \Delta_n)^\varpi$ . In the empirical application, we find  $\tilde{c}^{(c)}$  and  $\hat{c}^{(c)}$  to be very close.<sup>9</sup> We also recommend a small-sample adjustment, which is to divide the above sum by the number of non-zero summands instead of  $k_n$ .

## 4 Estimation with a Volatility Instrument

### 4.1 Volatility Instrument

Financial derivatives are known to contain valuable information about the volatility. However, to make use of this information, certain assumptions that link the volatility dynamics under the risk-neutral measure  $\mathbb{Q}$  to the dynamics under the objective measure  $\mathbb{P}$  are necessary. In this section, we assume that we have access to what we call a volatility instrument.

**Assumption 2.** *There exists an observable variable  $Z_t$ , which is a monotone increasing and twice differentiable function of  $c_t$  and a smooth function of  $t$ , that is,  $Z_t = f(t, c_t)$ . In addition,  $[Z, Z]_s^c$  is almost surely not vanishing and  $f'(s, x) > 0$ , for all  $0 \leq s \leq t$  and  $x > 0$ .<sup>10</sup> We call such a variable a volatility instrument.*

Assumption 2 implies that the spot leverage effect between  $X_t$  and  $c_t$  equals the spot leverage

<sup>9</sup>Unlike price jumps, volatility jumps are much more complicated to deal with. Proving the asymptotic properties of this estimator may be possible by combining our results with a method used by [Jacod and Rosenbaum \(2012\)](#), which would complicate the proof even further.

<sup>10</sup>We use  $f'$  here to denote the derivative of  $f$  with respect to the second argument.

effect between  $X_t$  and  $Z_t$ ,

$$\rho_s = \frac{[X, c]_s^{c'}}{\sqrt{[X, X]_s^{c'}} \sqrt{[c, c]_s^{c'}}} = \frac{[X, Z]_s^{c'}}{\sqrt{[X, X]_s^{c'}} \sqrt{[Z, Z]_s^{c'}}}.$$

The last equality is true by Itô's lemma for general Itô semimartingales, see, e.g., Proposition 8.19 of [Cont and Tankov \(2004\)](#). This implies that the  $IL$  measure is invariant with respect to the functional transformation  $f$ . If the researcher has access to high-frequency data on a volatility instrument, this property can be used to eliminate the effect of risk premia embedded in the function  $f$ , and hence to estimate  $IL$ .

**Remark.** *A weaker assumption is sufficient for all of the results in this paper. If we decompose the stochastic process  $Z_t$  into its continuous and jump components, we only require  $Z_t^c = \tilde{f}(t, c_t)$ , where  $Z_t^c$  is the continuous component of  $Z_t$ . This means we need no assumptions on the jump part of the volatility instrument other than some mild conditions on its activity as in Assumption 1. For ease of presentation and interpretation, we adopt the current Assumption 2.*

Assumption 2 imposes restrictions on the volatility dynamics under the risk-neutral measure  $\mathbb{Q}$  and is therefore a high-level assumption. It holds for a number of models in the derivative pricing literature. We summarize two classes of such risk-neutral models for the spot variance, of which the implied variance,  $Z_t$ , defined as the squared VIX up to some scalar, is a monotone increasing and differentiable function.<sup>11</sup> The first class of models, henceforth *Type-I models*, has the following risk-neutral dynamics:

$$\text{Type-I Model: } \sigma_t^2 = \sigma_0^2 + \int_0^t \kappa(\bar{\xi}^2 - \sigma_s^2) ds + M_t^{\mathbb{Q}},$$

where  $M^{\mathbb{Q}}$  is a martingale under the  $\mathbb{Q}$ -measure and  $\kappa, \bar{\xi}$  are model parameters. Type-I models include those studied by [Bakshi, Cao, and Chen \(1997\)](#), [Bates \(2000\)](#), [Pan \(2002\)](#), [Eraker \(2004\)](#), [Eraker, Johannes, and Polson \(2003\)](#), [Broadie, Chernov, and Johannes \(2007\)](#), and [Bates \(2012\)](#), among others, where jumps may be driven by the compound Poisson process with time-varying intensity or the CGMY process ([Carr, Geman, Madan, and Yor \(2003\)](#)). Type-I models also include non-Gaussian OU processes considered by [Barndorff-Nielsen and Shephard \(2001\)](#); see also [Shephard](#)

<sup>11</sup>For both classes of models, the dynamics of the logarithm of the index  $X_t$ , under the risk-neutral measure  $\mathbb{Q}$ , is assumed to follow

$$X_t = X_0 + \int_0^t b_s^{\mathbb{Q}} ds + \int_0^t \sigma_s dW_s^{\mathbb{Q}} + \int_0^t \int_{\mathbb{R}} z (N(ds, dz) - \nu(V_s, dz) ds),$$

where the drift  $b_s^{\mathbb{Q}}$  is determined by the no-arbitrage condition and is irrelevant for the pricing of the VIX,  $W^{\mathbb{Q}}$  is a  $\mathbb{Q}$ -Brownian motion, and  $N(ds, dz)$  is the jump measure of  $X_t$  with compensator  $\nu(V_s, dz) ds$  that may depend on the spot variance. We assume there exist constants  $\eta_0$  and  $\eta_1$ , such that for all  $v > 0$ ,  $\int_{\mathbb{R}} z^2 \nu(v, dz) = \eta_0 + \eta_1 v$ .

(2005) for a collection of similar models. For Type-I models, we can show that  $Z_t = a + b\sigma_t^2$ , where  $a$  and  $b$  depend on parameters that appear in the risk-neutral dynamics.

Second, *Type-II models* impose an exponential-affine structure for the  $\mathbb{Q}$ -dynamics of  $\sigma_t^2$ :

$$\text{Type-II Model: } \log \sigma_t^2 = \alpha + \beta F_t, \quad F_t = F_0 - \int_0^t \kappa F_s ds + L_t^{\mathbb{Q}},$$

where  $L_t^{\mathbb{Q}}$  is a finite variational Lévy martingale with diffusive coefficient  $\sigma$  and Lévy measure  $\tilde{\nu}$ , and  $\alpha, \beta, \kappa$  are model parameters. This model dates back to Nelson (1990), who introduces it as a continuous-time limit of the discrete EGARCH model. Andersen, Bollerslev, and Meddahi (2005) and Chernov, Gallant, Ghysels, and Tauchen (2003) have employed this model in their empirical work. One can show that Type-II models imply the following pricing formula for the implied variance,

$$Z_t = \eta_0 + \frac{1}{\tau} \int_0^\tau (\eta_1 + 1) \exp(\alpha + e^{-\kappa u}(\log \sigma_t^2 - \alpha) + C(u)) du,$$

where  $\tau = 21$  trading days is the horizon of the VIX,  $C(u) = \int_0^u \varphi(\beta e^{-\kappa v}) dv$ ,  $\varphi(u) = \sigma^2 u^2 / 2 + \int_{\mathbb{R}} (e^{uz} - 1 - uz) \tilde{\nu}(dz)$ , and  $\eta_0$  and  $\eta_1$  are constants defined in footnote (11). It can be easily verified that  $Z_t$  is an increasing differentiable function of  $\sigma_t^2$ . In fact,  $Z_t \approx a + b\sigma_t^{2d}$ , where  $a, b$ , and  $d$  are some constants that depend on the parameters in the risk-neutral dynamics.

Assumption 2 effectively says that to use the VIX as an instrument, the VIX should only be driven by the unobserved volatility. Notice that this does not necessarily imply that the objective volatility process can only be driven by one factor. However, it does rule out a few models in the recent empirical finance literature, see, e.g., Mencia and Sentana (2013), Christoffersen, Heston, and Jacobs (2009), and Andersen, Fusari, and Todorov (2015). For example, a common two-factor volatility process under the risk-neutral measure  $\mathbb{Q}$  is

$$d\sigma_t^2 = \left( \eta + \kappa \xi_t^2 - \kappa^{\mathbb{Q}} \sigma_t^2 \right) dt + \gamma \sigma_t d\tilde{W}_t^{\mathbb{Q}}, \quad d\xi_t^2 = \left( \kappa_\xi \bar{\xi}^2 - \kappa_\xi^{\mathbb{Q}} \xi_t^2 \right) dt + \gamma_\xi \xi_t d\tilde{B}_t^{\mathbb{Q}}, \quad (9)$$

where  $E^{\mathbb{Q}} \left( d\tilde{B}_t^{\mathbb{Q}} d\tilde{W}_t^{\mathbb{Q}} \right) = \tilde{\rho}_t dt$ , and  $E^{\mathbb{Q}} \left( d\tilde{B}_t^{\mathbb{Q}} dW_t^{\mathbb{Q}} \right) = \bar{\rho}_t dt$ . It is straightforward to derive that  $VIX_t^2 = a + b\sigma_t^2 + c\xi_t^2$ , where  $a, b$ , and  $c$  depend on parameters under the risk-neutral measure  $\mathbb{Q}$ .

The VIX-based *IRL* estimator converges to

$$\frac{1}{t} \int_0^t \frac{b\gamma\rho_s\sigma_s + c\gamma_\xi\bar{\rho}_s\xi_s}{\sqrt{b^2\gamma^2\sigma_s^2 + c^2\gamma_\xi^2\xi_s^2}} ds \neq \frac{1}{t} \int_0^t \rho_s ds = IL_t.$$

The typical justification for one of the two factors, say  $\xi_t^2$ , is that it captures the dynamics of the long-term variance, which has a much smaller  $\gamma_\xi$  compared to  $\gamma$ , as implied from the empirical data, see, e.g., Song and Xiu (2014). In addition,  $c$  is smaller relative to  $b$ , since the maturity of the VIX is only 1 month. As a result, the *IRL* estimator may approximate *IL* very well, despite being

inconsistent (this is also what we find in our simulations).

Besides the VIX, alternative volatility instruments can also be used. For example, for the S&P 500 index, we can use the intraday VIX futures or VIX options, which require a stronger assumption that the state-price densities of the implied variance only depend on  $\sigma_t^2$ . For individual equities, the VIX can be calculated.<sup>12</sup> Alternatively, the Black-Scholes implied volatility with a fixed maturity and moneyness can be used as a volatility instrument. Assumption 2 is satisfied if the marginal state price density of  $X_t$  only depends on  $\sigma_t^2$ , and the option price is homogenous of degree 1, see, e.g., Joshi (2002) and Song and Xiu (2014). More specifically, suppose that

$$C(X, \sigma_t^2, k, T - t) = \exp(X) \cdot C(0, \sigma_t^2, m, T - t),$$

where  $C$  is the price of a European call option with a log strike price  $k$  and a fixed maturity date  $T$ , and  $m = k - X$  is a fixed log-moneyness.<sup>13</sup> It then follows that the Black-Scholes implied volatility is a deterministic function of  $\sigma_t^2$ .

## 4.2 IRL Estimator and its Asymptotic Distribution

We now develop the asymptotic theory for the Instrument-based Realized Leverage (*IRL*), which exploits additional information embedded in the instrument. Our asymptotic results of the *IRL* estimator draw extensively from the general theoretical framework of Jacod and Rosenbaum (2013).

Suppose we have a sample of equidistant observations on  $X$  and  $Z$  over the interval  $[0, t]$ . Denote the distance between adjacent observations by  $\Delta_n$ . Partition all observations into non-overlapping blocks, indexed by  $i = 0, 1, 2, \dots, [t/(k_n \Delta_n)]$ , so that each block contains  $k_n$  observations. We can then estimate the spot leverage effect at time  $ik_n \Delta_n$  using the local truncated realized correlation between  $X$  and  $Z$ ,

$$\hat{\rho}_{ik_n+1} = \frac{\sum_{j=0}^{k_n-1} (\Delta_{ik_n+1+j}^n Z) (\Delta_{ik_n+1+j}^n X) \cdot 1_{\{|\Delta_{ik_n+1+j}^n Z| \leq u'_n\}} 1_{\{|\Delta_{ik_n+1+j}^n X| \leq u_n\}}}{\sqrt{\sum_{j=0}^{k_n-1} (\Delta_{ik_n+1+j}^n Z)^2 1_{\{|\Delta_{ik_n+1+j}^n Z| \leq u'_n\}}} \sqrt{\sum_{j=0}^{k_n-1} (\Delta_{ik_n+1+j}^n X)^2 1_{\{|\Delta_{ik_n+1+j}^n X| \leq u_n\}}}}, \quad (10)$$

where we truncate jumps in  $X$  and  $Z$  using thresholds  $u_n$  and  $u'_n$ . The choice of  $u_n$  and  $u'_n$  is standard in the literature, see, e.g., Aït-Sahalia and Jacod (2014). It is crucial to truncate jumps here as the true spot correlation is only defined for the continuous components. As the length of the

<sup>12</sup>The CBOE has already calculated the VIX for a limited number of stocks since January 7, 2011.

<sup>13</sup>Since  $X$  is time-varying, fixed log-moneyness implies that  $k$  needs to be changed over time.

local window  $k_n \Delta_n$  shrinks to zero,  $\hat{\rho}_{ik_n+1}$  approximates the spot leverage  $\rho_{ik_n \Delta_n}$ ,

$$\begin{aligned} \hat{\rho}_{ik_n+1} &\approx \frac{[Z, X]_s^{c'}}{\sqrt{[Z, Z]_s^{c'}} \sqrt{[X, X]_s^{c'}}} = \frac{f'(s, \sigma_s^2) [X, \sigma^2]_s^{c'}}{\sqrt{[f'(s, \sigma_s^2)]^2 [\sigma^2, \sigma^2]_s^{c'}} \sqrt{[X, X]_s^{c'}}} \\ &= \frac{[X, \sigma^2]_s^{c'}}{\sqrt{[\sigma^2, \sigma^2]_s^{c'}} \sqrt{[X, X]_s^{c'}}} = \rho_s, \text{ where } s = ik_n \Delta_n. \end{aligned} \quad (11)$$

Notice how this relationship holds for any smooth function  $f$ . In particular, the spot leverage  $\rho_s$  is invariant to functional transformations of  $\sigma^2$  and  $X$ , a property that is analogous to the linear invariant property of the standard correlation of two random variables. This property is important in that it eliminates the impact of risk premia embedded in the pricing function  $f$ . A consistent estimator of the  $IL$  is thereby a Riemann sum of the estimators of the spot leverage, but it has an asymptotic bias.

Define the Instrument-based Realized Leverage effect,  $IRL$ , as follows,

$$IRL_t = \frac{k_n \Delta_n}{t} \sum_{i=0}^{\lfloor t/k_n \Delta_n \rfloor - 1} \left( \hat{\rho}_{ik_n+1} - \frac{1}{2k_n} \left( (\hat{\rho}_{ik_n+1})^3 - \hat{\rho}_{ik_n+1} \right) \right). \quad (12)$$

The following theorem presents the asymptotic distribution of the  $IRL$  estimator.

**Theorem 2.** *Suppose Assumptions 1 and 2 hold. Let  $u_n \asymp \Delta_n^\varpi$ ,  $u'_n \asymp \Delta_n^\varpi$ , and suppose  $5/(12-2r) \leq \varpi < 1/2$ . Then, for any fixed  $t$ , and as  $\Delta_n \rightarrow 0$ ,  $k_n^2 \Delta_n \rightarrow 0$  and  $k_n^3 \Delta_n \rightarrow \infty$ , we have*

$$\Delta_n^{-1/2} (IRL_t - IL_t) \xrightarrow{L-\mathcal{S}} \mathcal{Z} \sqrt{V_t^{IRL}}, \quad \text{where } V_t^{IRL} = \frac{1}{t^2} \int_0^t (1 - 2\rho_s^2 + \rho_s^4) ds,$$

and  $\mathcal{Z}$  is a standard normal random variable defined on the extension of the original probability space.<sup>14</sup>

The estimator of the asymptotic variance can be constructed similarly,

$$\hat{V}_t^{IRL} = \frac{k_n \Delta_n}{t^2} \sum_{i=0}^{\lfloor t/k_n \Delta_n \rfloor - 1} \left( 1 - 2(\hat{\rho}_{ik_n+1})^2 + (\hat{\rho}_{ik_n+1})^4 \right).$$

An alternative way to estimate the asymptotic variance is by applying the subsampling method of [Kalnina \(2011\)](#) and [Kalnina \(2015\)](#). While the sampling period above is fixed at  $[0, t]$  such as 1 month, we typically apply the  $IRL$  estimator to a sequence of time periods in practice. Therefore, we provide uniform confidence bands of  $IL$  across different periods. Suppose the  $\tau$ th sampling month is  $[(\tau-1)t, \tau t]$ , and there are  $N$  months in total. Let  $IL(\tau)$  denote the integrated leverage at month  $\tau$ , and  $IRL(\tau)$  and  $V^{IRL}(\tau)$  be its estimator and the asymptotic variance. Then a nominal level  $1 - \alpha$

<sup>14</sup>The convergence here is stable in law, see, e.g., Chapter 2.2.1 of [Jacod and Protter \(2012\)](#) for a detailed review.

uniform confidence band for  $IL(\tau)$  is given by

$$CI_n(\tau; 1 - \alpha) = \left[ IRL(\tau) - z_{N,\alpha} \Delta^{1/2} \sqrt{\hat{V}^{IRL}(\tau)}, IRL(\tau) + z_{N,\alpha} \Delta^{1/2} \sqrt{\hat{V}^{IRL}(\tau)} \right],$$

where  $z_{N,\alpha}$  is the  $1 - \alpha$  quantile of the variable  $\max_{1 \leq \tau \leq N} |\mathcal{N}_\tau|$  and  $\{\mathcal{N}_\tau : 1 \leq \tau \leq N\}$  are independent standard normal variables. The asymptotic property of the above confidence band follows from the asymptotic independence of the estimation errors across non-overlapping periods, as is typical in the high frequency setting.

For the purpose of estimating the spot correlation, only the rate  $\Delta_n^{-1/2}$  of the tuning parameter  $k_n$  is compatible with the optimal rate of convergence, see [Alvarez, Panloup, Pontier, and Savy \(2012\)](#). Interestingly, [Theorem 2](#) and its proof show that for the purpose of estimating the integrated correlation, a slower rate for  $k_n$  ensures the feasibility of the inference by diminishing asymptotic biases, while maintaining the same asymptotic variance as in the case when  $k_n$  is of order  $\Delta_n^{-1/2}$ . The final estimator is easy to construct and has a simple expression for the asymptotic variance. Our Monte Carlo simulations suggest it also has good finite sample performance.

An alternative estimator can be constructed by using overlapping blocks of size  $k_n$ . Such an estimator shares the same asymptotic distribution with the non-overlapping estimator. This feature has been shown by [Jacod and Rosenbaum \(2013\)](#) for their estimator. We use the non-overlapping implementation because it is faster. We find that both implementations have similar finite sample performance.

## 5 Specification Test

In the presence of a valid volatility instrument, the *IRL* is more efficient than the *PRL*, in the sense that it has a faster rate of convergence. This increase in precision does not come without a cost. The *IRL* estimator requires the availability of a volatility instrument, which might not be available for all stocks. It also assumes that [Assumption 2](#) adequately describes the data used. Besides, the *IRL* has two practical disadvantages. First, the use of two high-frequency time series creates biases due to asynchronicity. Second, the derivatives are typically less liquid than the underlying stocks, which results in higher levels of the market microstructure noise. Both issues can be mitigated by using lower sampling frequencies, which leads to data loss.

It is of interest to test whether the two estimators, the *PRL* and the *IRL*, converge to the same limit. We use a Durbin-Wu-Hausman statistic, i.e., a standardized difference between the *PRL* and the *IRL* estimators. It can be used as a specification test of [Assumption 2](#).

**Theorem 3.** *Suppose Assumptions 1 and 2 hold. In addition,  $\sigma_t$  is a continuous Itô semimartingale. Let  $u_n \asymp \Delta_n^\varpi$ ,  $u'_n \asymp \Delta_n^\varpi$ , and suppose  $0 \leq r < 1/3$  and  $5/(12 - 6r) \leq \varpi < 1/2$ . Let  $PRL$  be defined by equation (7) with  $k'_n$  observations in each of the long blocks and  $l_n$  observations in each of the short blocks. Let  $IRL$  be defined by equation (12) with  $k_n$  observations in each block. Suppose  $\beta = l_n \sqrt{\Delta_n} \in (0, \infty)$ . Then, for any fixed  $t$ , and as  $\Delta_n \rightarrow 0$ ,  $k'_n \Delta_n \rightarrow 0$ ,  $k_n \Delta_n \rightarrow \infty$ ,  $(k'_n)^3 \Delta_n^2 \rightarrow \infty$ , and  $(k'_n)^4 \Delta_n^3 \rightarrow 0$ , we have*

$$\Delta_n^{-1/4} \frac{PRL_t - IRL_t}{\sqrt{\hat{V}_t^{PRL}}} \xrightarrow{L\text{-}s} \mathcal{Z}, \quad (13)$$

where  $\hat{V}_t^{PRL}$  is given in equation (8) and  $\mathcal{Z}$  is a standard normal random variable defined on the extension of the original probability space.

This test is consistent against those alternatives, under which our asymptotic theory for the  $PRL$  estimator continues to hold, while the  $IRL$  estimator becomes inconsistent.

## 6 Simulations

This section considers the finite sample properties of the  $PRL$ , and of the  $IRL$  using the VIX as a volatility instrument.

We first consider two different models of price and volatility dynamics, the generalized Heston and the log volatility (LogV) models. In the Heston model, the log-price  $X$  under the objective measure satisfies

$$\begin{aligned} dX_t &= (\mu_0 + \mu_1 \sigma_t^2) dt + \sigma_t dW_t + J_X dN_t - \lambda_t \mu_X dt, & d\rho_t^2 &= \kappa_\rho (\bar{\rho} - \rho_t) dt + \gamma_\rho \sqrt{1 - \rho_t^2} dB_t, \\ d\sigma_t^2 &= \kappa (\xi - \sigma_t^2) dt + \gamma \sigma_t d\tilde{W}_t + J_{\sigma^2} dN_t - \beta_{\sigma^2} \lambda_t dt, \end{aligned}$$

where  $W_t$  and  $\tilde{W}_t$  are two Brownian motions with correlation  $\rho_t$ ,  $B_t$  is another independent Brownian Motion,  $N_t$  is a Poisson process with state-dependent intensity  $\lambda_t = \lambda_0 + \lambda_1 \sigma_t^2$ ,  $J_X$  is a random jump size of  $X$  satisfying  $J_X \sim N(\mu_X, \sigma_X^2)$ , and  $J_{\sigma^2}$  is a random jump size of  $\sigma_t^2$  satisfying  $J_{\sigma^2} \sim \exp(\beta_{\sigma^2})$ . We set the parameters to the values typically used in the literature:  $\kappa = 5$ ,  $\gamma = 0.35$ ,  $\kappa_\rho = 4$ ,  $\gamma_\rho = 0.2$ ,  $\bar{\rho} = -0.8$ ,  $\xi = 0.06$ ,  $\mu_X = 0$ ,  $\sigma_X = 0.05$ ,  $\beta_{\sigma^2} = 0.01$ ,  $\lambda_0 = 30$ ,  $\lambda_1 = 60$ ,  $\mu_0 = 0.05$ , and  $\mu_1 = 0.5$ . We assume the risk-neutral dynamics follows the same model, but with different parameters. In this case,  $VIX_t^2 = a + b\sigma_t^2$  where the constants  $a$  and  $b$  depend on the parameters of the risk-neutral dynamics of  $X$ . Since our analysis depends on the risk-neutral parameters through the use of the VIX, we only need to choose the values of  $a$  and  $b$ . According to our calibration, we



set  $\text{VIX}_t^2 = 100^2 \cdot (0.06 + 0.63\sigma_t^2)$ .

In the LogV model, the log-price  $X$  under the objective measure satisfies

$$\begin{aligned} dX_t &= (\mu_0 + \mu_1 \sigma_t^2)dt + \sigma_t dW_t + J_X dN_t - \lambda \mu_X dt, & d\rho_t^2 &= \kappa_\rho(\bar{\rho} - \rho_t)dt + \gamma_\rho \sqrt{1 - \rho_t^2} dB_t, \\ dF_t &= \kappa F_t dt + \sigma d\widetilde{W}_t + J_F dN_t - \mu_F \lambda dt, & \sigma_t^2 &= \exp(\alpha + \beta F_t), \end{aligned}$$

where again  $W_t$  and  $\widetilde{W}_t$  are two Brownian motions with correlation  $\rho_t$ ,  $B_t$  is another independent Brownian Motion,  $N_t$  is a Poisson process with intensity  $\lambda$ ,  $J_X \sim N(\mu_X, \sigma_X^2)$ , and  $J_F \sim N(\mu_F, \sigma_F^2)$ . We set the values of the parameters to  $\alpha = -2.8$ ,  $\beta = 3$ ,  $\bar{\rho} = -0.8$ ,  $\mu_X = 0$ ,  $\sigma_X = 0.05$ ,  $\mu_F = 0.02$ ,  $\sigma_F = 0.02$ ,  $\sigma = 0.8$ ,  $\kappa = -4$ ,  $\lambda = 20$ ,  $\mu_0 = 0.05$ , and  $\mu_1 = 0.5$ . If the risk-neutral dynamics of  $X$  follows the same model, then  $\text{VIX}_t^2 = a + b\sigma_t^{2c}$  for some constants  $a, b$ , and  $c$ , where  $a, b$ , and  $c$  depends on the risk-neutral parameters. We set  $\text{VIX}_t^2 = 100^2 \cdot (0.1 + 0.75\sigma_t^{2 \times 0.8})$ .

We fix the truncation levels (values of  $u_n$  and  $u'_n$ ) to be  $3\sqrt{[Y, \hat{Y}]_t^c / t \Delta_n^{0.47}}$ , where  $Y$  is either the return of the stock or the change in the VIX, and  $[Y, Y]_t^c$  is estimated by bipower variation, see [Barndorff-Nielsen and Shephard \(2004\)](#) and [Aït-Sahalia and Jacod \(2012\)](#). For the *IRL* estimator, the rule-of-thumb choice of the block size  $k_n$  is approximately between  $0.1 \times \Delta_n^{-0.5}$  and  $0.5 \times \Delta_n^{-0.5}$ , corresponding to a window that is typically a couple of hours, as used in [Aït-Sahalia and Xiu \(2014\)](#) and [Li and Xiu \(2015\)](#). Since we use non-overlapping windows, we try to choose one among the divisors of the total number of observations within the desired ranges of  $k_n$ s. This helps improve the finite sample performance. For the *PRL*, the size of the smaller block  $l_n$  is around  $5 \times \Delta_n^{-0.5}$ , whereas the rule-of-thumb range of the larger block size  $k_n$  is approximately between  $2 \times \Delta_n^{-0.7}$  and  $10 \times \Delta_n^{-0.7}$ .

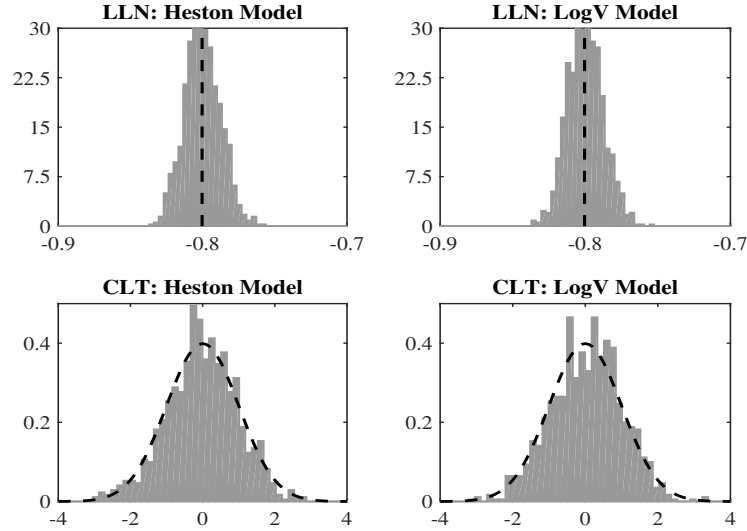
In [Figure 1](#), we provide histograms based on 1,000 repetitions of the *IRL* estimator using 1-minute returns spanning a 5-day window in the Heston and LogV models. In both cases considered, the *IRL* estimator can recover the integrated leverage precisely, and its finite sample distribution is close to the asymptotic distribution.

[Table 1](#) provides a robustness check for the performance of the *IRL* estimator for different frequencies, time spans, and bandwidths ( $k_n$ s). In all cases, the *IRL* estimator performs well, with the best performance in the case of 1 week, 1 minute data, which has the largest number of observations among the scenarios in [Table 1](#).

Finally, we investigate the performance of both leverage estimators using a multi-factor volatility model. Under the objective measure, we have

$$dX_t = (\mu_0 + \mu_1 \sigma_t^2)dt + \sigma_t dW_t + J_X dN_t - \lambda_t \mu_X dt, \quad d\rho_t^2 = \kappa_\rho(\bar{\rho} - \rho_t)dt + \gamma_\rho \sqrt{1 - \rho_t^2} dB_t,$$

Figure 1: Simulation Results: The *IRL* estimator



Note: The upper panel provides the histograms of the *IRL*, whereas the bottom panel plots the standardized histograms. The solid lines denote  $\bar{\rho}$  or the standard normal density. The parameters are  $\Delta_n = 1$  minute,  $T = 5$  days, and  $k_n = 78$ .

$$d\sigma_t^2 = (\eta + \kappa(\xi_t^2 - \sigma_t^2)) dt + \gamma\sigma_t d\widetilde{W}_t, \quad d\xi_t^2 = \kappa_\xi(\bar{\xi}^2 - \xi_t^2)dt + \gamma_\xi\xi_t d\widetilde{B}_t,$$

where  $W_t$  and  $\widetilde{W}_t$  are two Brownian motions with correlation  $\rho_t$ ,  $B_t$  and  $\widetilde{B}_t$  are two additional independent Brownian motions,  $N_t$  is a Poisson process with state-dependent intensity  $\lambda_t = \lambda_0 + \lambda_1\sigma_t^2$ ,  $J_X$  is a random jump size of  $X$  satisfying  $J_X \sim N(\mu_X, \sigma_X^2)$ . We choose parameters as follows:  $\kappa = 5$ ,  $\eta = 0.5$ ,  $\kappa_\xi = 2$ ,  $\kappa_\rho = 4$ ,  $\gamma = 0.35$ ,  $\gamma_\xi = 0.3$ ,  $\gamma_\rho = 0.2$ ,  $\bar{\rho} = -0.8$ ,  $\bar{\xi}^2 = 0.06$ ,  $\mu_X = 0$ ,  $\sigma_X = 0.05$ ,  $\lambda_0 = 15$ ,  $\lambda_1 = 60$ ,  $\mu_0 = 0.05$ , and  $\mu_1 = 0.5$ . If the risk-neutral dynamics follows the same model, then  $\text{VIX}_t^2 = a + b\sigma_t^2 + c\xi_t^2$ , where the constants  $a$ ,  $b$ , and  $c$  depend on the parameters of the risk-neutral dynamics of  $X$ . We set them to  $\text{VIX}_t^2 = 100^2 \cdot (0.3 + 0.75\sigma_t^2 + 0.15\xi_t^2)$ , in order to match the observed empirical data.

In this multi-factor volatility model, Assumption 2 is violated. Therefore, our results for the *IRL* estimator do not apply, while the results for the *PRL* estimator continue to hold. Table 2 and Figure 2 demonstrate the behavior of the two estimators in this model. Figure 2 shows that *IRL* is biased (top left plot); the finite sample distribution is therefore shifted away from the asymptotic distribution (bottom left plot). On the other hand, the finite sample distribution of the *PRL* estimator is very close to the asymptotic distribution. The bias of the *PRL* estimator is small relative to the standard deviation (top right plot), but it has clearly much larger variability than the *IRL* estimator. Table 2

Table 1: Simulation Results: The *IRL* estimator

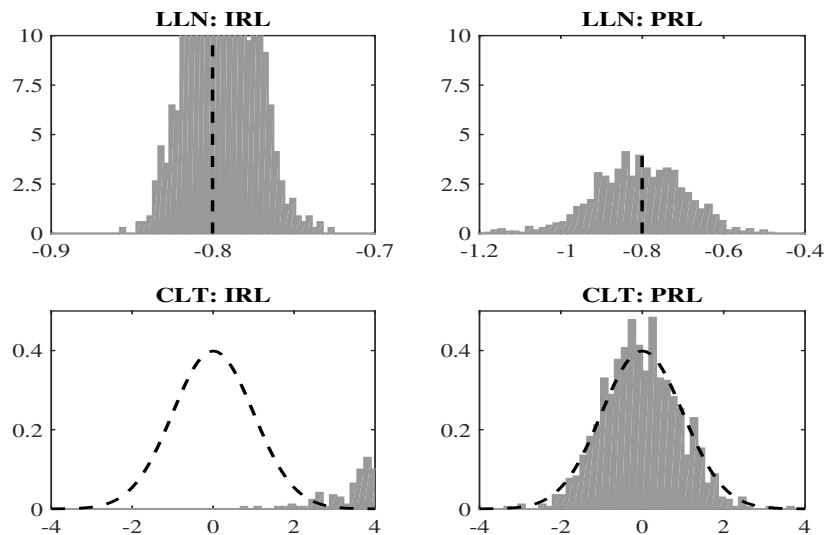
	<i>Heston: 1 week, 1 minute</i>			<i>LogV: 1 week, 1 minute</i>		
bias	0.000	0.000	0.000	0.001	0.001	0.001
RMSE	0.018	0.015	0.017	0.026	0.026	0.026
	<i>Heston: 1 week, 5 minutes</i>			<i>LogV: 1 week, 5 minutes</i>		
bias	0.002	0.002	0.002	0.004	0.004	0.004
RMSE	0.045	0.045	0.045	0.067	0.068	0.067
	<i>Heston: 1 month, 5 minutes</i>			<i>LogV: 1 month, 5 minutes</i>		
bias	0.002	0.002	0.002	0.003	0.003	0.003
RMSE	0.048	0.047	0.048	0.052	0.052	0.052
	<i>Heston: 1 month, 30 minutes</i>			<i>LogV: 1 month, 30 minutes</i>		
bias	0.025	0.026	0.030	0.025	0.026	0.030
RMSE	0.161	0.164	0.178	0.161	0.164	0.178

Note: Rows “bias” and “IQR” contain the average and interquartile range across simulations of the estimation error  $IRL_t - IL_t$ . For a given model, the three columns correspond to different bandwidths  $k_n$ , which are as follows: 39, 79, and 117 for 1 minute sampling; 26, 39, and 78 for 5 minute sampling; 13, 26, and 39 for 30 minute sampling.

considers multiple settings and choices of the bandwidth. The absolute bias of the *IRL* estimator is sometimes smaller and sometimes larger than that of the *PRL* estimator. However, due to the much smaller variability of the *IRL* estimator, it has a much smaller RMSE in all settings considered.

The choice between the *IRL* and the *PRL* estimator therefore depends on the exact loss function of the researcher. The *PRL* estimator does not rely on Assumption 2, hence it is valid in a much larger class of models. The cost is larger variability of the resulting estimates compared to the *IRL*.

Figure 2: Simulation Results: the Multi-factor Volatility Model



Note: The top panel provides the histograms of the *IRL* and *PRL* estimators, whereas the right panel plots the standardized histograms. The gray areas show the histograms of the *PRL* estimators. The solid lines denote  $\bar{\rho}$  or the standard normal density. The parameters are  $\Delta_n = 1$  minute and  $T = 2$  years. For the *PRL*, the small block (with  $l_n$  observations) is 3 days long, and the large block (with  $k_n$  observations) is 2 months long. For the *IRL*,  $k_n = 78$ .

Table 2: Simulation Results: the Multi-factor Volatility Model

	<i>PRL: 1 month, 1 second</i>			<i>IRL: 1 month, 1 second</i>		
bias	0.002	-0.009	-0.001	0.009	0.009	0.009
RMSE	1.472	0.174	0.194	0.097	0.097	0.096
% outside $[-1, 1]$	6.300	11.700	11.000	0.000	0.000	0.000
	<i>PRL: 2 years, 1 minute</i>			<i>IRL: 2 years, 1 minute</i>		
bias	-0.013	0.012	0.022	0.005	0.005	0.005
RMSE	0.069	0.157	0.187	0.074	0.073	0.073
% outside $[-1, 1]$	2.000	0.500	0.100	0.000	0.000	0.000
	<i>PRL: 5 years, 5 minutes</i>			<i>IRL: 5 years, 5 minutes</i>		
bias	-0.036	0.034	0.049	0.007	0.007	0.007
RMSE	1.304	0.320	0.276	0.082	0.081	0.081
% outside $[-1, 1]$	2.500	0.900	0.300	0.000	0.000	0.000

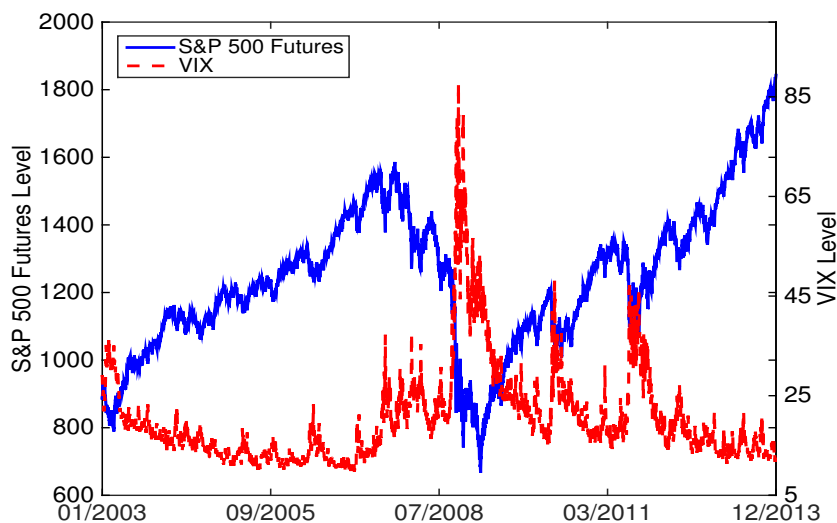
Note: For a given time span and frequency, the three columns correspond to different bandwidths. For the *IRL* they are:  $k_n = 150, 3120$ , and  $6240$  for 1 second sampling,  $k_n = 39, 78$ , and  $117$  for 1 minute sampling, and  $k_n = 26, 39$ , and  $78$  for 5 minute sampling. Choice of parameters for the *PRL*: for 1 second data, the small block length is 1/2 day, the long subsamples are 3, 4, and 5 days; for 1 minute data, the small block is 3 days, and the large blocks are 2, 3, and 4 months; for 5 minute data, the small block is 7 days, and the large blocks are 4, 5, and 6 months.

## 7 Empirical Results

### 7.1 Data and Preliminary Analysis

We consider two different empirical applications: a time-series analysis that uses the *IRL* estimator, and a cross-sectional analysis that uses the *PRL* estimator. For this purpose, we collect two sets of data. For the time series analysis, we use intraday time series of E-mini S&P 500 futures and the VIX from the Tick Data Inc. The VIX sample period starts from September 22, 2003, when the CBOE began disseminating prices for the VIX with the new methodology, and ends on December 31, 2013. We extend the intraday VIX series to January 1, 2003, by calculating the VIX using intraday options, based on the method developed by the CBOE, so that the full sample period covers 11 years in total.<sup>15</sup> We obtain a time series of the E-mini S&P 500 future prices by rolling over the front contracts. After removing non-trading days or half-trading days, our sample contains 2,769 days. Overnight returns are excluded from our data. Figure 3 plots the time series of the intraday E-mini S&P 500 futures and the VIX from January 1, 2003 to December 31, 2013. After investigation of the signature plots of the *IRL*, we choose the sampling frequency of 30-min for further time series analysis.

Figure 3: Time Series of the E-mini S&P 500 Future Prices and the VIX



Note: This figure plots the time series of the intraday E-mini S&P 500 futures and the VIX from January 1, 2003 to December 31, 2013.

For the cross-sectional analysis, we consider all stocks that have been constituents of the the S&P

<sup>15</sup>See the CBOE White paper on VIX at <http://www.cboe.com/micro/vix/vixwhite.pdf>.

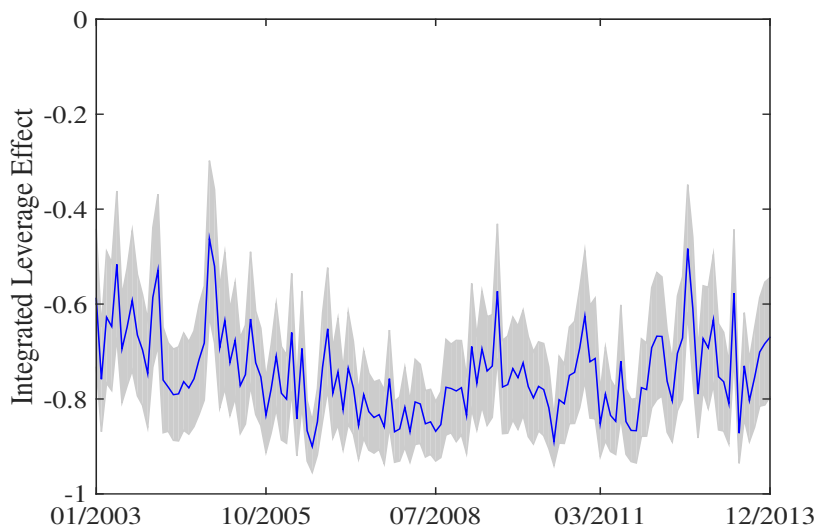
500 index over 2003-2012. We obtain 30-min return observations from the NYSE TAQ database. We next link this database with the CRSP-CompuStat merged database in order to obtain the long-term debt, current liabilities, preferred stock, and stockholders' equity for these stocks. Following [Christie \(1982\)](#), we define the total debt as the sum of the first three variables, and the debt to equity ratio as the ratio of total debt and stockholders' equity. We only consider stocks, for which none of the accounting variables are missing, and which have at least 2 years of observations during 2003-2012. In total, our cross-section contains 4,318 stocks.

## 7.2 Time Series Analysis

The current section implements a time series analysis of the leverage effect of the S&P500 index. We use the *IRL* estimator. The *PRL* estimator is noisy over short time intervals such as one month and hence not very informative about the variation of the true leverage effect across time. We use the VIX index as a volatility instrument.

We first plot the monthly time series based on the *IRL* estimator in [Figure 4](#), calculated using 30 minute frequency. The average for the VIX-based *IRL* is -0.745. Both the time series pattern and the average leverage effect are not sensitive to the choice of  $\theta$  and sampling frequencies (15-minute and lower).

Figure 4: Time Series of Leverage Effect Estimates



Note: This figure plots the monthly integrated leverage effect estimates of the E-mini S&P 500 futures for the years 2003 - 2013 using 30-minute data, together with the uniform confidence intervals in grey.

Table 3: Time Series Regression Results

Variable	(1)	(2)	(3)	(4)	(5)	(6)
DEF		-0.066* (0.034)				0.019 (0.047)
ILLIQ			-0.139** (0.061)			-0.145** (0.056)
DER				-0.230* (0.126)		-0.329** (0.131)
NBER					-0.030** (0.014)	-0.031** (0.015)
AR	0.460*** (0.084)	0.439*** (0.085)	0.448*** (0.080)	0.464*** (0.083)	0.427*** (0.084)	0.426*** (0.078)
adj. $R^2$	0.209	0.215	0.248	0.209	0.217	0.251

Note: Column (6) contains the results of the time series regression (14) with the VIX-based  $IRL$  as the dependent variable. Columns (1)-(5) contain results of the corresponding simple regressions. DER denotes the log total debt-to-total-equity ratio of the S&P 500 index, DEF denotes the Default spread, ILLIQ is a monthly illiquidity measure, and NBER denotes the crisis dummy. The reported standard errors are based on the Newey-West procedure. All regressions include an intercept, which is not reported.

The time series plot suggests two observations. First, our estimates indicate that the Brownian co-movement between the S&P 500 and its volatility is very pronounced and cannot be ignored. This challenges the pure jump specification of volatility process as suggested by [Todorov and Tauchen \(2011\)](#). Second, the correlation between the driving Brownian Motions of the price and the volatility is clearly negative and time-varying.

It is interesting to explore if the documented time variation in the leverage effect estimates is related to variation in financial variables. For this purpose, we conduct the following regression analysis. First, we select the following economic variables: a measure of the credit risk, the default spread (DEF), calculated as the monthly difference between Moody’s Seasoned BAA and AAA corporate bond yields from the FRED; an illiquidity measure (ILLIQ), which is constructed by a monthly value-weighted firm-wise Amihud measure using CRSP data ([Amihud \(2002\)](#)); a crisis dummy variable (NBER, 1 = crisis), constructed according to the NBER’s Business Cycle Dating Committee; and the logarithm of the monthly total debt-to-total-equity ratio (DER) of the S&P 500 index, downloaded from Bloomberg.<sup>16</sup> Our sample consists of monthly leverage effects over 11 years,

<sup>16</sup>The debt-to-equity ratio for S&P 500 index is construed by Bloomberg, as the sum of short term and long term borrowings divided by the total shareholder’s equity, where the latter is equal to the sum of preferred equity, minority interest, and total common equity. We have also constructed the TED spread, the difference between the three-month LIBOR and the three-month T-Bill interest rate obtained from the FRED, a liquidity measure from [Pastor and Stambaugh \(2003\)](#), as well as an equally-weighted firm-wise Amihud illiquidity measure. The economic interpretation of the regression results is robust to the choice of different variables.



which is a total of 132 observations.

We estimate the following AR(1) regression:

$$LEV_t = \beta_0 + \beta_1 \varepsilon_{DEF,t} + \beta_2 \varepsilon_{ILLIQ,t} + \beta_3 \varepsilon_{DER,t} + \beta_4 NBER_t + \beta_5 LEV_{t-1} + \varepsilon_t, \quad (14)$$

where  $\varepsilon_{.,t}$  denotes the corresponding AR(1) innovation of each covariate. From the regression results in Table 3, we find that the credit and liquidity factors are relevant to the leverage effect and their coefficients have signs consistent with the economic intuition (recall the dependent variable is negative). They imply that the leverage effect is magnified in bad times, i.e., a 1 percentage drop of stock price when credit risk is high and liquidity is low, may lead to a larger percentage increase of risk. The same conclusion holds with NBER dummy, i.e., crisis periods display a larger leverage effect. Moreover, the debt-to-equity ratio is significant, which supports the financial leverage hypothesis of Black (1976) that the debt-to-equity ratio is one of the determinants of the leverage effect. The latter finding is robust with respect to various alternative specifications, which are omitted for brevity.

### 7.3 Specification Test

We now implement the Durbin-Wu-Hausman test of Section 5. The test statistic is given in equation (13). We use 30-minute observations of E-mini S&P 500 future prices and the VIX index observations at the same frequency. The *PRL* estimator is -0.634 (with the standard error 0.073), while the *IRL* estimator is -0.745. The p-value of the two-sided test is 0.13, so we do not reject the null hypothesis of *PRL* and *IRL* yielding the same estimates in large samples.

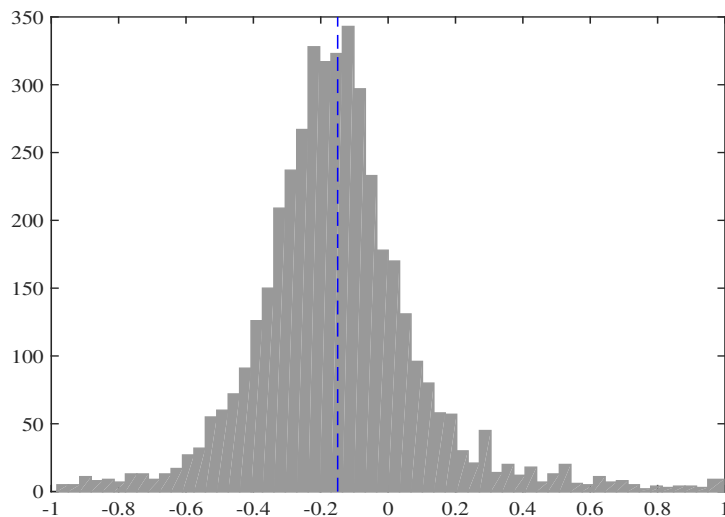
### 7.4 Cross-Sectional Analysis

In this section, we consider the cross-sectional variation of the leverage effect of individual companies. For this purpose, the *PRL* estimator seems particularly well suited because it is hard to obtain reliable intraday data on volatility instruments for a large number of companies. Also, by using a time span of ten years, we can ensure that the *PRL* delivers reasonably precise estimates.

We first plot the histogram of the *PRL* estimates for each individual stock in Figure 5. The estimates are robust to tuning parameters, so we report the results based on  $l_n = 7$  days and  $k_n = 6$  months. There are two immediate observations. First, a large portion of stocks exhibit a negative correlation between their returns and changes in volatilities. Second, the cross-sectional average of the leverage effect is negative, around -0.2, which is significantly smaller than the *PRL* estimate using the E-mini S&P 500 futures. This may suggest that the leverage effect of the S&P 500 index

is not only affected by each constituent on its own, but is also amplified by their comovement.

Figure 5: The Cross-Sectional Distribution of Leverage Effect Estimates



Note: This figure plots the histogram of the average  $PRL$  of 4,318 stocks for the years 2003 - 2012 using 30-minute data. The dotted line marks the cross-sectional average of the estimates.

We next run the following cross-sectional regression:

$$LEV_i = \beta_0 + \beta_1 \log(DER_i) + \varepsilon_i, \quad (15)$$

where  $LEV_i$  is the  $PRL$  estimate of the  $i$ th stock using data sampled within 2003 - 2012, and  $DER_i$  is the logarithm of the average debt-to-equity ratio over the same sampling period. We use the logarithm of debt-to-equity ratios because of the wide range of debt-to-equity ratios in the cross-section. The regression estimate  $\hat{\beta}_1$  is  $-0.0245$ , and the standard error is  $0.003$ . The  $R^2$  of the regression is  $1.43\%$ . Hence, the debt-to-equity ratio seems to be a relevant factor related to the cross-sectional variation of the leverage effect.

## 8 Conclusion

We propose two nonparametric estimators of the leverage effect. The first estimator,  $PRL$ , only uses the price data, and corrects for biases that arise due to the preliminary estimation of volatility. Our proof of the asymptotic distribution of the  $PRL$  extends several key methods in the literature. Our second estimator,  $IRL$ , uses the data from two sources, the stock price as well as a volatility instrument such as Black-Scholes implied volatility or the VIX. We provide the asymptotic theory

for the *IRL* estimator as well. The two estimators we develop are complementary to each other and have their own advantages and disadvantages. The *PRL* estimator is valid in a very general class of models, while the *IRL* estimator has a faster rate of convergence. Empirically, we find the *PRL* estimator has much larger standard errors than the *IRL* estimator, but is nevertheless useful for estimating the integrated leverage effect over the span of several years. We conduct a cross-sectional study with the *PRL*, and a time-series study with the *IRL*. Both studies find significant relationship between the leverage effect and the debt-to-equity ratio, which supports the leverage hypothesis of [Black \(1976\)](#).

## References

- AÏT-SAHALIA, Y., J. FAN, R. LAEVEN, D. C. WANG, AND X. YANG (2015): “The Estimation of Continuous and Discontinuous Leverage Effects,” Economics papers, Princeton University.
- AÏT-SAHALIA, Y., J. FAN, AND Y. LI (2013): “The Leverage Effect Puzzle: Disentangling Sources of Bias at High Frequency,” *Journal of Financial Economics*, *forthcoming*.
- AÏT-SAHALIA, Y., AND J. JACOD (2012): “Analyzing the Spectrum of Asset Returns: Jump and Volatility Components in High Frequency Data,” *Journal of Economic Literature*, 50, 1005–1048.
- (2014): *High-Frequency Financial Econometrics*. Princeton University Press.
- AÏT-SAHALIA, Y., I. KALNINA, AND D. XIU (2014): “The Idiosyncratic Volatility Puzzle: A Re-assessment at High Frequency,” Discussion paper, The University of Chicago.
- AÏT-SAHALIA, Y., AND R. KIMMEL (2007): “Maximum Likelihood Estimation of Stochastic Volatility Models,” *Journal of Financial Economics*, 83, 413–452.
- AÏT-SAHALIA, Y., AND D. XIU (2014): “Principal Component Analysis of High Frequency Data,” Discussion paper, Princeton University and the University of Chicago.
- ALVAREZ, A., F. PANLOUP, M. PONTIER, AND N. SAVY (2012): “Estimation of the Instantaneous Volatility,” *Statistical Inference and Stochastic Processes*, 15, 27–59.
- AMIHUD, Y. (2002): “Illiquidity and Stock Returns: Cross-Section and Time Series Effects,” *Journal of Financial Markets*, 5, 31–56.

- ANDERSEN, T., T. BOLLERSLEV, AND N. MEDDAHI (2005): “Correcting the Errors: Volatility Forecast Evaluation Using High Frequency Data and Realized Volatilities,” *Econometrica*, 73, 279–296.
- ANDERSEN, T., O. BONDARENKO, AND M. T. GONZALEZ-PEREZ (2015): “Uncovering Novel Features of Equity-Index Return Dynamics via Corridor Implied Volatility,” *Review of Financial Studies*, forthcoming.
- ANDERSEN, T., N. FUSARI, AND V. TODOROV (2015): “The Risk Premia Embedded in Index Options,” *Journal of Financial Economics*, forthcoming.
- ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2003): “Modeling and Forecasting Realized Volatility,” *Econometrica*, 71, 579–625.
- BAKSHI, G., C. CAO, AND Z. CHEN (1997): “Empirical Performance of Alternative Option Pricing Models,” *The Journal of Finance*, 52, 2003–2049.
- BANDI, F., AND R. RENÒ (2012): “Time-Varying Leverage Effects,” *Journal of Econometrics*, 12, 94–113.
- BANDI, F. M., AND R. RENÒ (2011): “Nonparametric Stochastic Volatility,” Discussion paper, Johns Hopkins University.
- (2014): “Price and Volatility Co-Jumps,” *Journal of Financial Economics*, forthcoming.
- BARNDORFF-NIELSEN, O. E., AND N. SHEPHARD (2001): “Non-Gaussian Ornstein-Uhlenbeck-Based Models And Some Of Their Uses In Financial Economics,” *Journal of the Royal Statistical Society, B*, 63, 167–241.
- (2004): “Power and bipower variation with stochastic volatility and jumps,” *Journal of Financial Econometrics*, 2, 1–37.
- BATES, D. (2012): “U.S. stock market crash risk, 1926-2010.,” *Journal of Financial Economics*, 105, 229–259.
- BATES, D. S. (2000): “Post-’87 Crash Fears in the S&P 500 Futures Option Market,” *Journal of Econometrics*, 94, 181–238.
- BLACK (1976): “Studies of Stock Price Volatility Changes,” *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics*, 6, 177–181.

- BOLLERSLEV, T., U. KRESCHMER, C. PIGORSCH, AND G. E. TAUCHEN (2009): “A Discrete-Time Model for Daily S&P 500 Returns and Realized Variations: Jumps and Leverage Effects,” *Journal of Econometrics*, 150, 151–166.
- BOLLERSLEV, T., S. Z. LI, AND V. TODOROV (2015): “Roughing Up Beta: Continuous vs Discontinuous Betas, and the Cross-Section of Expected Stock Returns,” *Journal of Financial Economics*, forthcoming.
- BOLLERSLEV, T., J. LITVINOVA, AND G. TAUCHEN (2006): “Leverage and Volatility Feedback Effects in High-Frequency Data,” *Journal of Financial Econometrics*, 4(3), 353–384.
- BOLLERSLEV, T., N. SIZOVA, AND G. TAUCHEN (2012): “Volatility in Equilibrium: Asymmetries and Dynamic Dependencies,” *Review of Finance*, 16, 31–80.
- BROADIE, M., M. CHERNOV, AND M. S. JOHANNES (2007): “Model specification and risk premia: Evidence from futures options,” *The Journal of Finance*, 62, 1453–1490.
- CARR, P., H. GEMAN, D. B. MADAN, AND M. YOR (2003): “Stochastic Volatility for Lévy Processes,” *Mathematical Finance*, 13, 345–342.
- CARR, P., AND L. WU (2009): “Variance Risk Premiums,” *Review of Financial Studies*, 22, 1311–1341.
- CHERNOV, M., A. R. GALLANT, E. GHYSELS, AND G. T. TAUCHEN (2003): “Alternative Models for Stock Price Dynamics,” *Journal of Econometrics*, 116, 225–257.
- CHRISTIE (1982): “The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects,” *Journal of Financial Economics*, 3, 145–166.
- CHRISTOFFERSEN, P., S. HESTON, AND K. JACOBS (2009): “The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well,” *Management Science*, 55, 1914–1932.
- COMTE, F., AND E. RENAULT (1998): “Long memory in continuous-time stochastic volatility models,” *Mathematical Finance*, 8, 291–323.
- CONT, R., AND P. TANKOV (2004): *Financial Modelling with Jump Processes*. Chapman & Hall/CRC.

- DURBIN, J. (1954): “Errors in Variables,” *Review of the International Statistical Institute*, 22, 23–32.
- ERAKER, B. (2004): “Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices,” *The Journal of Finance*, 59, 1367–1404.
- ERAKER, B., M. S. JOHANNES, AND N. POLSON (2003): “The Impact of Jumps in Equity Index Volatility and Returns,” *The Journal of Finance*, 58, 1269–1300.
- HARVEY, A. C., AND N. SHEPHARD (1994): “Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns,” *Journal of Business and Economic Statistics*, 14, 429–434.
- HAUSMAN, J. (1978): “Specification Tests in Econometrics,” *Econometrica*, 46, 1251–1271.
- JACOD, J., AND P. PROTTER (2012): *Discretization of Processes*. Springer.
- JACOD, J., AND M. ROSENBAUM (2012): “Estimation of Volatility Functionals: the Case of a  $\sqrt{n}$  Window,” Discussion paper, Université de Paris-6.
- JACOD, J., AND M. ROSENBAUM (2013): “Quarticity and Other Functionals of Volatility: Efficient Estimation,” *Annals of Statistics*, 41, 1462–1484.
- JACOD, J., AND A. N. SHIRYAEV (2003): *Limit Theorems for Stochastic Processes*. Springer-Verlag, second edn.
- JACQUIER, E., N. G. POLSON, AND P. E. ROSSI (2004): “Bayesian analysis of stochastic volatility models with fat-tails and correlated errors,” *Journal of Econometrics*, 122, 185–212.
- JOSHI, M. (2002): “Log-type models, Homogeneity of Option Prices and Convexity,” Discussion paper, Melbourne University.
- KALNINA, I. (2011): “Subsampling High Frequency Data,” *Journal of Econometrics*, 161, 262–283.
- (2015): “Inference for Nonparametric High-Frequency Estimators with an Application to Time Variation in Betas,” Discussion paper, University of Montreal.
- KRISTENSEN, D. (2010): “Nonparametric Filtering of the Realised Spot Volatility: A Kernel-Based Approach,” *Econometric Theory*, 26, 60–93.
- LI, J., AND A. PATTON (2015): “Asymptotic Inference about Predictive Accuracy using High Frequency Data,” Discussion paper, Duke University.

- LI, J., V. TODOROV, AND G. TAUCHEN (2014): “Adaptive Estimation of Continuous-Time Regression Models using High-Frequency Data,” Discussion paper, Duke University.
- LI, J., AND D. XIU (2015): “Generalized Methods of Integrated Moments for High-Frequency Data,” Discussion paper, Duke University and University of Chicago.
- MENCIA, J., AND E. SENTANA (2013): “Valuation of VIX Derivatives,” *Journal of Financial Economics*, 108(2), 367–391.
- NELSON, D. B. (1990): “ARCH Models as Diffusion Approximations,” *Journal of Econometrics*, 45, 7–38.
- (1991): “Conditional Heteroskedasticity in Asset Returns: A New Approach,” *Econometrica*, 59, 347–370.
- PAN, J. (2002): “The jump-risk premia implicit in options: Evidence from an integrated time-series study,” *Journal of Financial Economics*, 63, 3–50.
- PASTOR, L., AND R. F. STAMBAUGH (2003): “Liquidity Risk and Expected Stock Returns,” *Journal of Political Economy*, 111, 642–685.
- PATTON, A., AND M. VERARDO (2012): “Does Beta Move with News? Firm-Specific Information Flows and Learning about Profitability,” *The Review of Financial Studies*, 25, 2789–2839.
- PITT, M. K., S. MALIK, AND A. DOUCET (2014): “Simulated likelihood inference for stochastic volatility models using continuous particle filtering,” *Annals of the Institute of Statistical Mathematics*, 66(3), 527–552.
- REISS, M., V. TODOROV, AND G. E. TAUCHEN (2015): “Nonparametric Test for a Constant Beta between Itô Semi-martingales based on High-Frequency Data,” *Stochastic Processes and their Applications*, forthcoming.
- SEGAL, G., I. SHALIASTOVICH, AND A. YARON (2015): “Good and Bad Uncertainty: Macroeconomic and Financial Market Implications,” *Journal of Financial Economics*, 117, 369–397.
- SHEPHARD, N. (2005): *Stochastic volatility*. Oxford University Press.
- SONG, Z., AND D. XIU (2014): “A Tale of Two Option Markets: State-Price Densities and Volatility Risk,” Discussion paper, Federal Reserve Board and University of Chicago.

- TODOROV, V., AND G. E. TAUCHEN (2011): “Volatility Jumps,” *Journal of Business & Economic Statistics*, 29(3), 356–371.
- VETTER, M. (2012a): “Estimation of Correlation for Continuous Semimartingales,” *Scandinavian Journal of Statistics*, 39, 757–771.
- VETTER, M. (2012b): “Estimation of Integrated Volatility of Volatility with Applications to Goodness-of-fit Testing,” Discussion paper, Ruhr-Universität Bochum.
- WANG, D. C., AND P. A. MYKLAND (2014): “The Estimation of Leverage Effect with High Frequency Data,” *Journal of the American Statistical Association*, 109, 197–215.
- WU, D.-M. (1973): “Alternative Tests of Independence between Stochastic Regressor and Disturbances,” *Econometrica*, 41, 733–750.
- YU, J. (2005): “On Leverage in a Stochastic Volatility Model,” *Journal of Econometrics*, 127, 165–178.



# Appendix

## A Proof of Theorem 1

We start by introducing some notation. Our notation is similar to that of the proofs for Chapter 8 in [Ait-Sahalia and Jacod \(2014\)](#) whenever possible.

### A.1 Notation

We denote the continuous component of  $X$  as  $X'$ , that is,  $X'_t = \int_0^t b_s ds + \int_0^t \sigma_s dW_s$ , and use prime to denote the estimators based on  $X'$ :

$$\begin{aligned}\hat{c}'_i &= \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n-1} (\Delta_{i+j}^n X')^2, \\ \hat{c}_i^{(X,c)'} &= \frac{1}{l_n k_n \Delta_n} \sum_{j=0}^{k_n-1} (\hat{\sigma}_{i+j+l_n}^{2l} - \hat{\sigma}_{i+j}^{2l}) (X'_{(i+j-1+2l_n)\Delta_n} - X'_{(i+j-1)\Delta_n}), \\ \hat{c}_i^{(c)'} &= \frac{3}{2l_n k_n \Delta_n} \sum_{j=0}^{k_n-1} \left[ (\hat{\sigma}_{i+j+l_n}^{2l} - \hat{\sigma}_{i+j}^{2l})^2 - \frac{4}{l_n} (\hat{\sigma}_{i+j}^{2l})^2 \right], \\ \hat{\sigma}_i^{2l} &= \frac{1}{l_n \Delta_n} \sum_{j=0}^{l_n-1} (\Delta_{i+j}^n X')^2.\end{aligned}$$

With this notation,  $\hat{C}'_i$  is a symmetric matrix with (1,1), (1,2) and (2,2) elements being  $\hat{c}'_i$ ,  $\hat{c}_i^{(X,c)'}$ , and  $\hat{c}_i^{(c)'}$ , respectively. In addition, we define

$$\bar{C}_{(i-1)\Delta_n} = \frac{1}{k_n} \sum_{s=0}^{k_n-1} C_{(i+s-1)\Delta_n}, \quad \alpha_i = \hat{C}'_i - \bar{C}_{(i-1)\Delta_n}, \quad \bar{\beta}_i = \bar{C}_{(i-1)\Delta_n} - C_{(i-1)\Delta_n}, \quad \beta_i = \alpha_i + \bar{\beta}_i.$$

Next, we introduce some notations that will be used to rewrite  $\alpha_i$ . To do this, we define

$$\begin{aligned}\varepsilon(1)_s^n &= \begin{cases} -1, & \text{if } 0 \leq s < l_n \\ 1, & \text{if } l_n \leq s < 2l_n \end{cases}, \quad \varepsilon(2)_s^n = (s+1) \wedge (2l_n - s - 1), \quad \varepsilon(3)_s^n = 1, \\ y_{u,v}^n &= \begin{cases} \frac{3}{2l_n^3}, & \text{if } u, v \in \{1, 2\} \\ \frac{1}{l_n^2}, & \text{otherwise.} \end{cases}, \quad \bar{\varepsilon}(1)_s^n = 1, \quad \bar{\varepsilon}(2)_s^n = l_n - s - 1, \\ \zeta(1)_i^n &= \frac{1}{\Delta_n} (\Delta_i^n X')^2 - c_{(i-1)\Delta_n}, \quad \zeta(2)_i^n = \Delta_i^n c, \quad \zeta(3)_i^n = \Delta_i^n X'.\end{aligned}$$

Moreover, we introduce

$$A(1; u, v)_i^n = \sum_{s=0}^{k_n-1} \Gamma(u, v)_0^n \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n,$$

$$\begin{aligned}
A(2; u, v)_i^n &= \sum_{s=0}^{k_n+2l_n-2} \left( \gamma(u, v; 0)_{s-k_n+1, s+1}^n - 1_{\{s \leq k_n-1\}} \Gamma(u, v)_0^n \right) \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n, \\
A(3; u, v)_i^n &= \sum_{s=0}^{k_n+2l_n-2} \left( \sum_{m=1}^{2l_n-1} \gamma(u, v; m)_{s-k_n+1, s+1}^n \zeta(u)_{i+s-m}^n - 1_{\{s \geq 2l_n-1\}} \rho(u, v)_{i+s}^n \right) \zeta(v)_{i+s}^n, \\
A(4; u, v)_i^n &= \sum_{s=2l_n-1}^{k_n+2l_n-2} \rho(u, v)_{i+s} \zeta'(v)_{i+s}^n,
\end{aligned}$$

where we use

$$\begin{aligned}
\gamma(u, v; m)_{j, L}^n &= y_{u, v}^n \sum_{q=0 \vee (j-m)}^{(L-m-1) \wedge (2l_n-m-1)} \varepsilon(u)_q^n \varepsilon(v)_{m+q}^n, \\
\Gamma(u, v)_m^n &= \gamma(u, v; m)_{0, 2l_n}^n = y_{u, v}^n \sum_{q=0}^{2l_n-m-1} \varepsilon(u)_q^n \varepsilon(v)_{m+q}^n, \\
Z(u, v)_i^n &= \sum_{j=2l_n-1}^{k_n+2l_n-2} \rho(u, v)_{i+j} \zeta''(v)_{i+j}^n, \quad \rho(u, v)_i^n = \sum_{m=1}^{2l_n-1} \Gamma(u, v)_m^n \zeta(u)_{i-m}^n.
\end{aligned}$$

For a similar purpose, we define:

$$\begin{aligned}
\bar{A}(0)_i &= \frac{6}{l_n^2} \sum_{j=0}^{k_n-1} (c_{(i+j-1)\Delta_n})^2, \\
\bar{A}(1; u)_i^n &= \frac{12}{l_n^3} \sum_{j=0}^{k_n-1} \left( c_{(i+j-1)\Delta_n} \sum_{m=0}^{l_n-1} \bar{\varepsilon}(u)_m^n \zeta(u)_{i+j+m}^n \right), \\
\bar{A}(2; u, v)_i^n &= \sum_{j=1}^{k_n+l_n-2} \bar{\gamma}(u, v; 0)_{j-k_n+1, j}^n \zeta(u)_{i+j}^n \zeta(v)_{i+j}^n, \\
\bar{A}(3; u, v)_i^n &= \sum_{j=1}^{k_n+l_n-2} \sum_{m=1}^{(l_n-1) \wedge j} \bar{\gamma}(u, v; m)_{j-k_n+1, j}^n \zeta(u)_{i+j-m}^n \zeta(v)_{i+j}^n, \\
\bar{\gamma}(u, v; m)_{j, s}^n &= \frac{6}{l_n^4} \sum_{q=0 \vee (j-m)}^{(s-m) \wedge (l_n-m-1)} \bar{\varepsilon}(u)_q^n \bar{\varepsilon}(v)_{q+m}^n.
\end{aligned}$$

Now we can write  $\alpha_i$  as the sum of  $\eta_i$  and  $\kappa_i$ , where  $\kappa_i^{lm}$  and  $\eta_i^{lm}$  are defined as follows,

$$\begin{aligned}
\kappa_i^{12} &= \frac{1}{k_n \Delta_n} \sum_{u=1}^2 (A(1; u, 3)_i^n + A(2; u, 3)_i^n + A(3; u, 3)_i^n + A(4; u, 3)_i^n + A(3; 3, u)_i^n + A(4; 3, u)_i^n) \\
&\quad - \frac{1}{k_n} \sum_{s=0}^{k_n-1} C_{(i+s-1)\Delta_n}^{12}, \\
\kappa_i^{22} &= \frac{1}{k_n \Delta_n} \sum_{u, v=1}^2 (A(1; u, v)_i + A(2; u, v)_i + 2A(3; u, v)_i + 2A(4; u, v)_i - \bar{A}(2; u, v)_i^n - 2\bar{A}(3; u, v)_i^n)
\end{aligned}$$

$$\begin{aligned}
& -\frac{1}{k_n \Delta_n} \bar{A}(0)_i^n - \frac{1}{k_n \Delta_n} \sum_{u=1}^2 \bar{A}(1; u)_i^n - \frac{1}{k_n} \sum_{s=0}^{k_n-1} C_{(i+s-1)\Delta_n}^{22}, \\
\kappa_i^{11} &= \alpha_i^{11}, \quad \kappa_i^{12} = \kappa_i^{21}, \\
\eta_i^{12} &= \frac{1}{k_n \Delta_n} \sum_{u=1}^2 (Z(u, 3)_i^n + Z(3, u)_i^n), \quad \eta_i^{22} = 2 \frac{1}{k_n \Delta_n} \sum_{u,v=1}^2 Z(u, v)_i, \quad \eta_i^{11} = 0, \quad \eta_i^{12} = \eta_i^{21}.
\end{aligned}$$

We develop the asymptotic distribution of these estimators below. First, we obtain that  $\Gamma(u, v)_m^n$  equals (first column for the case when  $m \leq l_n - 1$ , second column for the case when  $m \geq l_n$ ):

$$\begin{aligned}
\Gamma(1, 1)_m^n &= \frac{6l_n - 9m}{2l_n^3}, & -\frac{6l_n - 3m}{2l_n^3}, \\
\Gamma(1, 3)_m^n &= -\frac{m}{l_n^2}, & -\frac{2l_n - m}{l_n^2}, \\
\Gamma(3, 1)_m^n &= \frac{m}{l_n^2}, & \frac{2l_n - m}{l_n^2}, \\
\Gamma(1, 2)_m^n &= -\frac{12l_n m - 9m^2 + 6l_n - 9m}{4l_n^3}, & -\frac{3(2l_n - m)(2l_n - m - 1)}{4l_n^3}, \\
\Gamma(2, 1)_m^n &= \frac{12l_n m - 9m^2 - 6l_n + 9m}{4l_n^3}, & \frac{3(2l_n - m)(2l_n - m + 1)}{4l_n^3}, \\
\Gamma(2, 2)_m^n &= \frac{4l_n^3 - 6l_n m^2 + 3m^3 + 2l_n - 3m}{4l_n^3}, & \frac{(2l_n - m)^3 - 2l_n + m}{4l_n^3}, \\
\Gamma(2, 3)_m^n &= \frac{2l_n^2 - m^2 + m}{2l_n^2}, & \frac{(2l_n - m)(2l_n - m + 1)}{2l_n^2}, \\
\Gamma(3, 2)_m^n &= \frac{2l_n^2 - m^2 - m}{2l_n^2}, & \frac{(2l_n - m)(2l_n - m - 1)}{2l_n^2}.
\end{aligned}$$

Moreover, we define

$$H(u, v; u'v')_n = z_{u,u'}^n z_{v,v'}^n \sum_{m=0}^{2l_n-1} \Gamma(u, v)_m^n \Gamma(u', v')_m^n, \quad \text{where } z_{u,v}^n = \begin{cases} 1/\Delta_n & \text{if } u = v = 1 \\ 1 & \text{otherwise} \end{cases},$$

and derive its limit as

$$\sqrt{\Delta_n} H(u, v; u'v')_n \rightarrow \begin{cases} 3/\beta^3 & \text{if } (u, v, u', v') = (1, 1, 1, 1) \\ 3/4\beta & \text{if } (u, v, u', v') = (1, 2, 1, 2), (2, 1, 2, 1) \\ 5/8\beta & \text{if } (u, v, u', v') = (1, 2, 1, 3), (2, 1, 3, 1) \\ 2/3\beta & \text{if } (u, v, u', v') = (1, 3, 1, 3), (3, 1, 3, 1) \\ 151\beta/280 & \text{if } (u, v, u', v') = (2, 2, 2, 2) \\ 151\beta/240 & \text{if } (u, v, u', v') = (2, 2, 2, 3), (2, 2, 3, 2) \\ 23\beta/30 & \text{if } (u, v, u', v') = (2, 3, 2, 3), (2, 3, 3, 2), (3, 2, 3, 2) \\ 0 & \text{otherwise} \end{cases}.$$

Next, for convenience we define

$$V(u, v)_t = \begin{cases} 2(C_t^{11})^2 & \text{if } (u, v) = (1, 1) \\ C_t^{22} & \text{if } (u, v) = (2, 2) \\ C_t^{12} & \text{if } (u, v) = (2, 3), (3, 2) \\ C_t^{11} & \text{if } (u, v) = (3, 3) \\ 0 & \text{otherwise} \end{cases}.$$

In addition, we define

$$l(u, v'; u', v'; C_s) = \begin{cases} \frac{12}{\beta^3} (C_s^{11})^4 & \text{if } (u, v; u', v') = (1, 1; 1, 1) \\ \frac{3}{2\beta} (C_s^{11})^2 C_s^{22} & \text{if } (u, v; u', v') = (1, 2; 1, 2), (2, 1; 2, 1) \\ \frac{151\beta}{280} (C_s^{22})^2 & \text{if } (u, v; u', v') = (2, 2; 2, 2) \\ \frac{5}{4\beta} (C_s^{11})^2 C_s^{12} & \text{if } \begin{cases} (u, v; u', v') = (1, 2; 1, 2), \\ (2, 1; 2, 1), (1, 3; 1, 1), (3, 1; 2, 1) \end{cases} \\ \frac{4}{3\beta} (C_s^{11})^3 & \text{if } (u, v; u', v') = (1, 3; 1, 3), (3, 1; 3, 1) \\ \frac{151\beta}{240} C_s^{22} C_s^{12} & \text{if } \begin{cases} (u, v; u', v') = (2, 2; 2, 3), \\ (2, 2; 3, 2), (2, 3; 2, 2), (3, 2; 2, 2) \end{cases} \\ \frac{23\beta}{30} C_s^{22} C_s^{11} & \text{if } (u, v; u', v') = (2, 3; 2, 3), (3, 2; 3, 2) \\ \frac{23\beta}{30} (C_s^{12})^2 & \text{if } (u, v; u', v') = (2, 3; 2, 3) \\ 0 & \text{otherwise.} \end{cases},$$

and

$$h(j, k; l, m; C_t) \tag{A.1}$$

$$= \begin{cases} 0, & \text{if } (j, k) \text{ or } (l, m) = (1, 1); \\ \sum_{(u,v) \in A} \sum_{(u',v') \in A} l(u, v, u', v'; C_t), & \text{if } (j, k, l, m) = (1, 2, 1, 2), (2, 1, 2, 1), \\ & (2, 1, 1, 2), (1, 2, 2, 1); \\ \sum_{(u,v) \in A} \sum_{(u',v') \in B} l(u, v, u', v'; C_t), & \text{if } (j, k, l, m) = (1, 2, 2, 2), (2, 1, 2, 2), \\ & (2, 2, 2, 1), (2, 2, 1, 2); \\ \sum_{(u,v) \in B} \sum_{(u',v') \in B} l(u, v, u', v'; C_t), & \text{if } (j, k, l, m) = (2, 2, 2, 2). \end{cases}$$

$$= \begin{cases} 0, & \text{if } (j, k) \text{ or } (l, m) = (1, 1); \\ \frac{8}{3\beta} c_s^3 + \frac{23\beta}{15} c_s^{(c)} c_s + \frac{23\beta}{30} (c_s^{(X,c)})^2, & \text{if } (j, k, l, m) = (1, 2, 1, 2), (2, 1, 2, 1), \\ & (2, 1, 1, 2), (1, 2, 2, 1); \\ \frac{5}{\beta} (c_s)^2 c_s^{(X,c)} + \frac{151\beta}{120} c_s^{(c)} c_s^{(X,c)}, & \text{if } (j, k, l, m) = (1, 2, 2, 2), (2, 1, 2, 2), \\ & (2, 2, 2, 1), (2, 2, 1, 2); \\ \frac{48}{\beta^3} c_s^4 + \frac{12}{\beta} (c_s)^2 c_s^{(c)} + \frac{151\beta}{70} (c_s^{(c)})^2, & \text{if } (j, k, l, m) = (2, 2, 2, 2). \end{cases},$$

where  $A = \{(1, 3), (2, 3), (3, 1), (3, 2)\}$  and  $B = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$ .

## A.2 Auxiliary Results

We provide a useful and general theorem here, which will be used to prove Theorem 1. The proof of this theorem is given in Appendix A.3. To cope with a general  $C^3$  function without imposing growth conditions of Jacod and Rosenbaum (2013), we need a spatial localization assumption similar to

those introduced in [Li, Todorov, and Tauchen \(2014\)](#) and [Li and Xiu \(2015\)](#), which can be easily verified for the special case in [Theorem 1](#).

**Theorem A1.** *There exists a localizing sequence of stopping times  $(\tau_m)_{m \geq 1}$  and a sequence of convex compact subsets  $\mathcal{K}_m \subset \mathcal{M}_2$ , a  $2 \times 2$ -dimensional matrix space, such that  $C_t \in \mathcal{K}_m$  for  $t \leq \tau_m$  and  $g$  is  $C^3$  on  $\mathcal{K}_m^\varepsilon \equiv \{M \in \mathcal{M}_2 : \inf_{A \in \mathcal{K}_m} \|M - A\| \leq \varepsilon\}$ , for some  $\varepsilon > 0$ . In addition, suppose  $0 \leq r < 1/3$  and  $5/(12 - 6r) \leq \varpi < 1/2$ . Then as  $k_n^3 \Delta_n^2 \rightarrow \infty$ ,  $k_n^4 \Delta_n^3 \rightarrow 0$ ,*

$$\begin{aligned} & \Delta_n^{-1/4} \left( k_n \Delta_n \sum_{i=0}^{N_t^n} \left( g(\hat{C}_{ik_n+1}) - \frac{1}{2k_n \sqrt{\Delta_n}} \sum_{j,k,l,m}^2 \partial_{jk,lm} g(\hat{C}_{ik_n+1}) h(j, k; l, m; \hat{C}_{ik_n+1}) \right) \right. \\ & \left. - \int_0^t g(C_s) ds \right) \xrightarrow{L^{-s}} \mathcal{Z} \sqrt{V_t^g}, \end{aligned} \quad (\text{A.2})$$

where  $\mathcal{Z}$  is a standard normal random variable defined on the extension of the original probability space, the function  $h$  is defined in [\(A.1\)](#), and the variance  $V_t^g$  is given by

$$V_t^g = \int_0^t \sum_{j,k,l,m}^2 \partial_{jk} g(C_s) \partial_{lm} g(C_s) h(j, k, l, m; C_s) ds.$$

In addition, we collect a few technical lemmas below, which will be used for the proof of [Theorem A1](#). We postpone the proofs of [Lemmas A2](#) and [A3](#) to the end of the appendix. [Lemma A4](#) collects results from [Aït-Sahalia and Jacod \(2014\)](#), hence its proof is omitted.

**Lemma A2.** *Under the same assumptions as in [Theorem A1](#), we have*

$$\|\mathbf{E}_{i-1}(\beta_i)\| \leq K k_n \Delta_n + K \Delta_n^{1/4} \eta_{i,2k_n}^n, \quad (\text{A.3})$$

$$\mathbf{E}_{i-1} \|\beta_i\|^q \leq K \left( k_n^{-1} \Delta_n^{-1/2} \right)^{q/2} + K k_n \Delta_n, \quad q \geq 2, \quad (\text{A.4})$$

$$\mathbf{E}_{i-1} \|\alpha_i\|^q \leq K \left( k_n^{-1} \Delta_n^{-1/2} \right)^{q/2}, \quad q \geq 1. \quad (\text{A.5})$$

**Lemma A3.** *Under the same assumptions as in [Theorem A1](#), we have, for all  $(j, k, l, m)$ ,*

$$\left| \mathbf{E}_{i-1} \left( \beta_i^{jl} \beta_i^{km} - k_n^{-1} \Delta_n^{-1/2} h(j, k; l, m; C_{(i-1)\Delta_n}) \right) \right| \leq K k_n \Delta_n.$$

**Lemma A4.** *Under the same assumptions as in [Theorem A1](#), we have*

$$|\zeta'(1)_i^n| \leq K \sqrt{\Delta_n} \left( \sqrt{\Delta_n} + \eta_i^n \right) \leq K \sqrt{\Delta_n}, \quad (\text{A.6})$$

$$\mathbf{E}_{i-1} |\zeta(1)_i^n|^q \leq K_q, \quad (\text{A.7})$$

$$\left| \zeta'(2)_i^n - b_{(i-1)\Delta_n}^{(c)} \Delta_n \right| + \left| \zeta'(3)_i^n - b_{(i-1)\Delta_n} \Delta_n \right| \leq K \sqrt{\Delta_n} \left( \sqrt{\Delta_n} + \eta_i^n \right) \leq K \sqrt{\Delta_n}, \quad (\text{A.8})$$

$$\mathbf{E}_{i-1} (|\zeta(2)_i^n|^q) + \mathbf{E}_{i-1} (|\zeta(3)_i^n|^q) \leq K_q \Delta_n^{q/2}, \quad (\text{A.9})$$

$$\left| \mathbf{E}_{i-1} \left( (\bar{\zeta}(1)_i^n)^2 - 2c_{(i-1)\Delta_n}^2 \right) \right| \leq K \sqrt{\Delta_n}, \quad (\text{A.10})$$

$$\left| \mathbf{E}_{i-1} \left( (\bar{\zeta}(2)_i^n)^2 - c_{(i-1)\Delta_n}^{(c)} \Delta_n \right) \right| \leq K \left( \Delta_n^2 + \Delta_n^{3/2} \eta_i^n \right), \quad (\text{A.11})$$

$$\left| \mathbf{E}_{i-1} \left( (\bar{\zeta}(3)_i^n)^2 - c_{(i-1)\Delta_n} \Delta_n \right) \right| \leq K \left( \Delta_n^2 + \Delta_n^{3/2} \eta_i^n \right), \quad (\text{A.12})$$

$$\left| \mathbf{E}_{i-1} \left( \bar{\zeta}(2)_i^n \bar{\zeta}(3)_i^n - c_{(i-1)\Delta_n}^{(X,c)} \Delta_n \right) \right| \leq K \left( \Delta_n^2 + \Delta_n^{3/2} \eta_i^n \right), \quad (\text{A.13})$$

$$\left| \mathbf{E}_{i-1} \left( \bar{\zeta}(1)_i^n \bar{\zeta}(2)_i^n \right) \right| \leq K \Delta_n, \quad (\text{A.14})$$

$$\left| \mathbf{E}_{i-1} \left( \bar{\zeta}(1)_i^n \bar{\zeta}(3)_i^n \right) \right| \leq K \Delta_n, \quad (\text{A.15})$$

$$\left\| \mathbf{E}_{i-1} \bar{\beta}_i \right\| \leq K k_n \Delta_n, \quad (\text{A.16})$$

$$\mathbf{E}_{i-1} \left( \left\| \bar{\beta}_i \right\|^q \right) \leq K (k_n \Delta_n)^{q/2} + K k_n \Delta_n, \quad (\text{A.17})$$

$$\left| \mathbf{E}_{i-1} \frac{1}{k_n} \sum_{j=0}^{k_n-1} \left( \frac{1}{\Delta_n} (\zeta(2)_i^n) (\zeta(3)_i^n) - C_{(i-1)\Delta_n}^{12} \right) \right| \leq K \sqrt{\Delta_n}, \quad (\text{A.18})$$

$$\left| \mathbf{E}_{i-1} \frac{1}{k_n} \sum_{j=0}^{k_n-1} \left( \frac{1}{\Delta_n} (\zeta(2)_i^n)^2 - C_{(i-1)\Delta_n}^{22} \right) \right| \leq K \sqrt{\Delta_n}, \quad (\text{A.19})$$

$$\mathbf{E}_{i-1} \left| \frac{1}{k_n} \sum_{j=0}^{k_n-1} \left( \frac{1}{\Delta_n} (\zeta(2)_i^n) (\zeta(3)_i^n) - C_{(i-1)\Delta_n}^{12} \right) \right|^q \leq K \left( \Delta_n^{q/2} + k_n^{-q/2} \right), \quad (\text{A.20})$$

$$\mathbf{E}_{i-1} \left| \frac{1}{k_n} \sum_{j=0}^{k_n-1} \left( \frac{1}{\Delta_n} (\zeta(2)_i^n)^2 - C_{(i-1)\Delta_n}^{22} \right) \right|^q \leq K \left( \Delta_n^{q/2} + k_n^{-q/2} \right), \quad (\text{A.21})$$

where, writing  $Y = (b, b^{(c)}, \sigma^{(c)}, c^{(c)}, c^{(X,c)})$ ,

$$\eta_{i,j}^n = \sqrt{\mathbf{E}_{i-1} \left( \sup_{s \in (0, j\Delta_n]} \|Y_{(i-1)\Delta_n+s} - Y_{(i-1)\Delta_n}\|^2 \right)}, \quad \eta_i^n = \eta_{i,1}^n. \quad (\text{A.22})$$

For all  $i \leq i' < i + j \leq i' + 2k_n$  we have

$$\mathbf{E}_{i-1} (\eta_{i',j}^n) \leq 2\eta_{i,2k_n}^n, \quad (\text{A.23})$$

and for all  $t$  we have

$$k_n \Delta_n \mathbf{E} \left( \sum_{i=0}^{N_t^n} \eta_{ik_n, 2k_n}^n \right) \rightarrow 0.$$

Additionally, if  $a_i^n$ 's are reals, all bounded by some constant  $L$ , then for all  $q \geq 2$  we have

$$\mathbf{E} \left( \left| \sum_{j=0}^{2l_n-1} a_j^n \zeta(u)_{i+j}^n \right|^q \right) \leq \begin{cases} K_q L^q l_n^{q/2} & \text{if } u = 1 \\ K_q L^q l_n^{-q/2} & \text{if } u = 2, 3 \end{cases}. \quad (\text{A.24})$$

Finally, for all  $q \geq 2$ ,

$$\mathbf{E}_{i-1} |\rho(u, 3)_{i+2l_n-1}|^q \leq \begin{cases} K l_n^{-3q/2} & \text{if } v = 1 \\ K l_n^{-q/2} & \text{if } v = 2, 3 \end{cases}. \quad (\text{A.25})$$

### A.3 Proof of Theorem A1

(a) We first establish the uniform convergence of  $\left\| \hat{C}_{ik_n+1} - \bar{C}_{ik_n\Delta_n} \right\|$  to 0 over  $i \in \{0, 1, 2, \dots, N_t^n\}$ .

By (4.8) of [Jacod and Rosenbaum \(2013\)](#), there exists a sequence of positive numbers  $\phi_n$  going to 0, such that

$$\mathbb{E} \left( \left| \hat{c}_{ik_n+1} - \hat{c}'_{ik_n+1} \right| \right) \leq K \phi_n \Delta_n^{(2-r)\varpi}. \quad (\text{A.26})$$

Since we have

$$\begin{aligned} \hat{c}_{ik_n+1}^{(X,c)} - \hat{c}'_{ik_n+1}^{(X,c)} &= \frac{1}{l_n^2 k_n \Delta_n^2} \sum_{j=0}^{k_n-1} \sum_{l=0}^{2l_n-1} \sum_{k=0}^{l_n-1} \left\{ (\Delta_{ik_n+j+l+1}^n X) (\Delta_{ik_n+j+l_n+k+1}^n X)^2 \times \right. \\ &\quad \left. \mathbb{1}_{\{|\Delta_{ik_n+j+l+1}^n X| \leq u_n\}} \mathbb{1}_{\{|\Delta_{ik_n+j+l_n+k+1}^n X| \leq u_n\}} - (\Delta_{ik_n+j+l+1}^n X') (\Delta_{ik_n+j+l_n+k+1}^n X')^2 \right\} \\ &\quad - \frac{1}{l_n^2 k_n \Delta_n^2} \sum_{j=0}^{k_n-1} \sum_{l=0}^{2l_n-1} \sum_{k=0}^{l_n-1} \left\{ (\Delta_{ik_n+j+l+1}^n X) (\Delta_{ik_n+j+k+1}^n X)^2 \times \right. \\ &\quad \left. \mathbb{1}_{\{|\Delta_{ik_n+j+l+1}^n X| \leq u_n\}} \mathbb{1}_{\{|\Delta_{ik_n+j+k+1}^n X| \leq u_n\}} - (\Delta_{ik_n+j+l+1}^n X') (\Delta_{ik_n+j+k+1}^n X')^2 \right\}, \end{aligned}$$

it follows that

$$\begin{aligned} \left| \hat{c}_{ik_n+1}^{(X,c)} - \hat{c}'_{ik_n+1}^{(X,c)} \right| &\leq \frac{1}{l_n^2 k_n \Delta_n^2} \sum_{j=0}^{k_n-1} \sum_{l=0}^{2l_n-1} \sum_{k=0}^{l_n-1} \left| (\Delta_{ik_n+j+l+1}^n X) (\Delta_{ik_n+j+k+1}^n X)^2 \times \right. \\ &\quad \left. \mathbb{1}_{\{|\Delta_{ik_n+j+l+1}^n X| \leq u_n\}} \mathbb{1}_{\{|\Delta_{ik_n+j+k+1}^n X| \leq u_n\}} - (\Delta_{ik_n+j+l+1}^n X') (\Delta_{ik_n+j+k+1}^n X')^2 \right|. \end{aligned}$$

Then we apply Lemma 13.2.6 of [Jacod and Protter \(2012\)](#). Using their notation, we define

$$\bar{X}_{i,j}^n = \left( \frac{\Delta_i^n X}{\sqrt{\Delta_n}}, \dots, \frac{\Delta_{i+j-1}^n X}{\sqrt{\Delta_n}}, \frac{\Delta_{i+j}^n X'}{\sqrt{\Delta_n}}, \dots, \frac{\Delta_{i+k-1}^n X'}{\sqrt{\Delta_n}} \right),$$

and consider the following functions:

$$F(x_1, x_2, \dots, x_k) = x_1 x_{j+1}^2, \quad F_u(x_1, x_2, \dots, x_k) = x_1 x_{j+1}^2 \mathbb{1}_{\{x_1 \leq u\}} \mathbb{1}_{\{x_{j+1} \leq u\}},$$

$$\eta_{i,j}^n = F_{u_n/\sqrt{\Delta_n}}(\bar{X}_{i,j+1}^n) - F_{u_n/\sqrt{\Delta_n}}(\bar{X}_{i,j}^n).$$

Then, by Lemma 13.2.6 of [Jacod and Protter \(2012\)](#) with  $m = 1, p' = 2, s = 1, s' = 2$ , applied to  $\eta_{i,0}^n$  and  $\eta_{i,j+1}^n$ , there exists a sequence of positive numbers  $\phi_n$ , such that

$$\begin{aligned} \mathbb{E}_{i-1} \left| \left( \frac{\Delta_i^n X}{\sqrt{\Delta_n}} \right) \left( \frac{\Delta_{i+j}^n X}{\sqrt{\Delta_n}} \right)^2 \mathbb{1}_{\{|\Delta_i^n X| \leq u_n\}} \mathbb{1}_{\{|\Delta_{i+j}^n X| \leq u_n\}} \right. \\ \left. - \left( \frac{\Delta_i^n X'}{\sqrt{\Delta_n}} \right) \left( \frac{\Delta_{i+j}^n X'}{\sqrt{\Delta_n}} \right)^2 \mathbb{1}_{\{|\Delta_i^n X'| \leq u_n\}} \mathbb{1}_{\{|\Delta_{i+j}^n X'| \leq u_n\}} \right| \leq \phi_n \left( \Delta_n^{(2-r)/2} + \Delta_n^{(2-r)\varpi} \right). \end{aligned}$$

With  $i$  and  $i + j$  above replaced by  $ik_n + 1 + l$  and  $ik_n + 1 + k$ , and further note that the set  $A_i^n = \{|\Delta_i^n X'| \leq u_n\}$  satisfies  $\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbf{P}((A_i^n)^c) < \infty$ , hence on the set  $\Omega_t^n = \cap_{1 \leq i \leq \lfloor t/\Delta_n \rfloor} A_i^n$ , whose probability goes to 1 as  $n \rightarrow \infty$ , we have:

$$\mathbf{E} \left| \hat{c}_{ik_n+1}^{(X,c)} - \hat{c}_{ik_n+1}^{(X,c)'} \right| \leq K \phi_n \left( \Delta_n^{\frac{1-r}{2}} + \Delta_n^{(2-r)\varpi - \frac{1}{2}} \right). \quad (\text{A.27})$$

We apply the same strategy to the function  $F(x_1, x_2, \dots, x_k) = x_1^2 x_j^2$  and use Lemma 13.2.6 of [Jacod and Protter \(2012\)](#) again with  $m = 1, p' = 3, s = 1, s' = 2$ , then a similar result to (A.27) holds for  $\hat{c}_{ik_n+1}^{(c)}$ , i.e.,

$$\mathbf{E} \left| \hat{c}_{ik_n+1}^{(c)} - \hat{c}_{ik_n+1}^{(c)'} \right| \leq K \phi_n \left( \Delta_n^{\frac{1-r}{2}} + \Delta_n^{(2-r)\varpi - \frac{1}{2}} \right). \quad (\text{A.28})$$

Therefore, by a maximal inequality, we have

$$\mathbf{E} \left( \sup_{i=\{0,1,2,\dots,N_t^n\}} \left\| \hat{C}_{ik_n+1} - \hat{C}_{ik_n+1}' \right\| \right) \leq K \phi_n \left( \Delta_n^{(2-r)\varpi} + \Delta_n^{\frac{1-r}{2}} + \Delta_n^{(2-r)\varpi - \frac{1}{2}} \right) (k_n \Delta_n)^{-1} \rightarrow 0,$$

as long as  $r < 1/3$  and  $5/(12 - 6r) \leq \varpi < 1/2$ .

Moreover, by (A.5) of Lemma A2, applying the maximal inequality with  $q \geq 4$ , we have

$$\mathbf{E} \left( \sup_{i=\{0,1,2,\dots,N_t^n\}} \left\| \hat{C}_{ik_n+1}' - \bar{C}_{ik_n+1} \right\|^q \right) \leq K \left( k_n^{-1} \Delta_n^{-1/2} \right)^{q/2} (k_n \Delta_n)^{-1} \rightarrow 0.$$

This establishes the desired uniform convergence.

By the localization assumption, we can assume that  $C_s \in \mathcal{K}$ , for any  $0 \leq s \leq t$  and some convex set  $\mathcal{K} \subset \mathcal{M}_2$ . By convexity,  $\bar{C}_s \in \mathcal{K}$ , for  $0 \leq s \leq t$ . Moreover,  $g$  is  $C^3$  on  $\mathcal{K}^{2\varepsilon}$ , for some  $\varepsilon > 0$ . By the uniform convergence established above,  $\hat{C}_{ik_n \Delta_n}' \in \mathcal{K}^\varepsilon$  with probability approaching 1, uniformly over  $i = \{0, 1, 2, \dots, N_t^n\}$  for  $n$  large enough. Therefore, we can restrict the domain of  $g$  on the compact set  $\bar{\mathcal{K}}^\varepsilon \subset \mathcal{K}^{2\varepsilon}$ , in which we have  $\|g(C)\|$ ,  $\|\partial g(C)\|$ , and  $\|\partial^2 g(C)\|$  are all bounded. Therefore, we can assume these properties of  $g$  in the following proof.

(b) We now decompose the left-hand side of (A.2) as:

$$\begin{aligned} & \Delta_n^{-1/4} \left( k_n \Delta_n \sum_{i=0}^{N_t^n} \left( g(\hat{C}_{ik_n+1}) - \frac{1}{2k_n \sqrt{\Delta_n}} \sum_{j,k,l,m} \partial_{j,k,l,m}^2 g(\hat{C}_{ik_n+1}) h(j, k; l, m; \hat{C}_{ik_n+1}) \right) \right) - \int_0^t g(C_s) ds \\ & = U(1)_t^n + U(2)_t^n + U(3)_t^n + U(4)_t^n + U(5)_t^n + U(6)_t^n, \end{aligned}$$

where

$$U(1)_t^n = \Delta_n^{-1/4} \left( \sum_{i=1}^{N_t^n} \int_{(i-1)k_n \Delta_n}^{ik_n \Delta_n} (g(C_{(i-1)k_n \Delta_n}) - g(C_s)) ds \right) - \Delta_n^{-1/4} \int_{k_n \Delta_n}^{N_t^n} g(C_s) ds,$$



$$U(2)_t^n = k_n \Delta_n^{3/4} \sum_{i=0}^{N_t^n} \sum_{j,k} \partial_{jk} g(C_{ik_n \Delta_n}) \bar{\beta}_{ik_n+1}^{jk},$$

$$U(3)_t^n = k_n \Delta_n^{3/4} \sum_{i=0}^{N_t^n} \left( g(\hat{C}'_{ik_n+1}) - g(C_{ik_n \Delta_n}) - \sum_{j,k} \partial_{jk} g(C_{ik_n \Delta_n}) (\alpha_{ik_n+1}^{jk} + \bar{\beta}_{ik_n+1}^{jk}) \right. \\ \left. - \frac{1}{2k_n \sqrt{\Delta_n}} \sum_{j,k,l,m}^2 \partial_{jk,lm}^2 g(\hat{C}'_{ik_n+1}) h(j,k;l,m; \hat{C}'_{ik_n+1}) \right),$$

$$U(4)_t^n = k_n \Delta_n^{3/4} \sum_{i=0}^{N_t^n} \sum_{j,k} \partial_{jk} g(C_{ik_n \Delta_n}) \kappa_{ik_n+1}^{jk},$$

$$U(5)_t^n = k_n \Delta_n^{3/4} \sum_{i=0}^{N_t^n} \sum_{j,k} \partial_{jk} g(C_{ik_n \Delta_n}) \eta_{ik_n+1}^{jk},$$

$$U(6)_t^n = k_n \Delta_n^{3/4} \sum_{i=0}^{N_t^n} \left( g(\hat{C}_{ik_n+1}) - g(\hat{C}'_{ik_n+1}) \right) - \Delta_n^{1/4} \sum_{i=0}^{N_t^n} \sum_{j,k,l,m}^2 \partial_{jk,lm}^2 g(\hat{C}_{ik_n+1}) h(j,k;l,m; \hat{C}_{ik_n+1}) \\ + \Delta_n^{1/4} \sum_{i=0}^{N_t^n} \sum_{j,k,l,m}^2 \partial_{jk,lm}^2 g(\hat{C}'_{ik_n+1}) h(j,k;l,m; \hat{C}'_{ik_n+1}).$$

We thereby need to analyze these terms respectively.

We start with  $U(1)_t^n$ . Let  $\xi_{ik_n} = \int_{(i-1)k_n \Delta_n}^{ik_n \Delta_n} (g(C_{(i-1)k_n \Delta_n}) - g(C_s)) ds$ , which is  $\mathcal{F}_{ik_n \Delta_n}$ -measurable. Also, we have  $|\mathbb{E}_{(i-1)k_n}(\xi_{ik_n})| \leq K(k_n \Delta_n)^2$  and  $\mathbb{E}_{(i-1)k_n} |\xi_{ik_n}|^2 \leq K(k_n \Delta_n)^3$ . By applying Doob's inequality to the first term, we obtain

$$\mathbb{E} \left( \sup_{s \leq t} |U(1)_s^n| \right) \leq K \Delta_n^{-1/4} \left( \sum_{i=1}^{N_t^n} |\mathbb{E}_{(i-1)k_n}(\xi_{ik_n})| + \left( \sum_{i=1}^{N_t^n} \mathbb{E}_{(i-1)k_n} |\xi_{ik_n}|^2 \right)^{1/2} \right) + K k_n \Delta_n^{3/4} \rightarrow 0.$$

Similar argument applies to  $U(2)_t^n$ . Denote

$$\xi_{ik_n} = \sum_{j,m} \partial_{jm} g(C_{ik_n \Delta_n}) \sum_{s=0}^{k_n-1} \left( C_{(ik_n+s)\Delta_n}^{jm} - C_{ik_n \Delta_n} \right),$$

which again is  $\mathcal{F}_{(i+1)k_n}$ -measurable. We also have  $|\mathbb{E}_{ik_n}(\xi_{ik_n})| \leq K k_n^2 \Delta_n$  and  $\mathbb{E}_{ik_n} |\xi_{ik_n}|^2 \leq K k_n^3 \Delta_n$ .

Therefore, by Doob's inequality, we have

$$\mathbb{E} \left( \sup_{s \leq t} |U(2)_s^n| \right) \leq K \Delta_n^{3/4} \left( \sum_{i=1}^{N_t^n} |\mathbb{E}_{ik_n}(\xi_{ik_n})| + \left( \sum_{i=1}^{N_t^n} \mathbb{E}_{ik_n} |\xi_{ik_n}|^2 \right)^{1/2} \right) \leq K k_n \Delta_n^{3/4} \rightarrow 0.$$

Next, to analyze  $U(3)_t^n$ , we denote  $\omega_{ik_n} = \xi_{ik_n} - v_{ik_n}$ , where

$$\xi_{ik_n} = g(\hat{C}_{ik_n+1}) - g(C_{ik_n \Delta_n}) - \sum_{j,m} \partial_{jm} g(C_{ik_n \Delta_n}) \beta_{ik_n+1}^{jm}$$

$$\begin{aligned}
& - \frac{1}{2k_n\sqrt{\Delta_n}} \sum_{j,k,l,m} \partial_{jk,lm}^2 \left( \hat{C}_{ik_n+1} \right) h \left( j, k; l, m; \hat{C}_{ik_n+1} \right), \\
v_{ik_n} &= \frac{1}{2} \sum_{j,k,l,m} \partial_{jk,lm}^2 (C_{ik_n\Delta_n}) \left( \beta_{ik_n+1}^{jl} \beta_{ik_n+1}^{km} - k_n^{-1} \Delta_n^{-1/2} \sum_{j,k,l,m} h(j, k; l, m; C_{ik_n\Delta_n}) \right),
\end{aligned}$$

By the mean-value theorem, Jensen's inequality, and Lemma A2, we have that

$$\mathbb{E}_{ik_n} |\omega_{ik_n}| \leq K \mathbb{E}_{ik_n} \|\beta_{ik_n}\|^3 + K k_n^{-1} \Delta_n^{-1/2} \mathbb{E}_{ik_n} \|\beta_{ik_n}\| \leq K k_n \Delta_n + K k_n^{-3/2} \Delta_n^{-3/4} + K k_n^{-1/2}.$$

Also, by Lemmas A2 and A3, we have  $|\mathbb{E}_{ik_n}(v_{ik_n})| \leq K k_n \Delta_n$  and  $\mathbb{E}_{ik_n} |v_{ik_n}|^2 \leq K k_n^{-2} \Delta_n^{-1}$ . Therefore,

$$\mathbb{E} \left( \sup_{s \leq t} |U(3)_s^n| \right) \leq K k_n \Delta_n^{3/4} \left( \sum_{i=1}^{N_t^n} |\mathbb{E}_{ik_n}(v_{ik_n})| + \sum_{i=1}^{N_t^n} \mathbb{E}_{ik_n} |\omega_{ik_n}| + \left( \sum_{i=1}^{N_t^n} \mathbb{E}_{ik_n} |v_{ik_n}|^2 \right)^{1/2} \right) \rightarrow 0.$$

The next term is  $U(4)_t^n$ , which comprises many terms in each  $\kappa_{ik_n+1}^{lm}$ , for which we have already obtained the desired bounds in the proof of Lemma A2. In summary, we have shown

$$|\mathbb{E}_{ik_n}(\kappa_{ik_n+1}^{lm})| \leq k_n \Delta_n + K \Delta_n^{1/4} \eta_{ik_n, 2k_n}^n, \quad \mathbb{E}_{ik_n} |\kappa_{ik_n+1}^{lm}|^2 \leq k_n \Delta_n.$$

This leads to  $U(4)_t^n = o_p(1)$  because by Doob's inequality,

$$\begin{aligned}
\mathbb{E} \left( \sup_{s \leq t} |U(4)_s^n| \right) &\leq k_n \Delta_n^{3/4} \sum_{i=0}^{N_t^n} \sum_{j,m} \mathbb{E} |\partial_{jm} g(C_{ik_n\Delta_n})| \left| \mathbb{E}_{ik_n} \kappa_{ik_n+1}^{lm} \right| \\
&\quad + k_n \Delta_n^{3/4} \left( \sum_{i=0}^{N_t^n} \sum_{j,m} \mathbb{E} |\partial_{jm} g(C_{ik_n\Delta_n})| \mathbb{E}_{ik_n} \left| \kappa_{ik_n+1}^{lm} \right|^2 \right)^{1/2} \\
&\leq k_n \Delta_n \sum_{i=0}^{N_t^n} \eta_{ik_n, 2k_n}^n + K k_n \Delta_n^{3/4} \rightarrow 0.
\end{aligned}$$

Next, we investigate  $U(5)_t^n$ , which delivers the asymptotic distribution. Denote

$$\xi_{j,k}^n(u, v) = k_n \Delta_n^{3/4} \sum_{i=1}^{N_t^n} \partial_{jk} g(C_{ik_n\Delta_n}) \frac{1}{k_n \Delta_n} Z(u, v)_{ik_n+1}^n.$$

We first show that for any  $(j, k), (u, v)$ , we have

$$(\xi_{j,k}^n(u, v)) \xrightarrow{\mathcal{L}-s} \xi = (\xi_{j,k}(u, v)), \tag{A.29}$$

where  $\xi$ , conditionally on  $\mathcal{F}$ , is a continuous centered Gaussian martingale with

$$\mathbb{E}(\xi_{j,k}(u, v) \xi_{l,m}(u', v') | \mathcal{F}) = \int_0^t \partial_{jk} g(C_s) \partial_{lm} g(C_s) l(u, v, u', v'; C_s) ds.$$

In fact, by rewriting the summations, we obtain

$$\begin{aligned}\xi_{j,k}^n(u,v) &= \Delta_n^{-1/4} \sum_{i=0}^{N_t^n} \partial_{jk} g(C_{ik_n \Delta_n}) \sum_{j=2l_n-1}^{k_n+2l_n-2} \rho(u,v)_{ik_n+1+j} \zeta''(v)_{ik_n+1+j}^n \\ &= \Delta_n^{-1/4} \sum_{i=1}^{k_n(N_t^n+1)} \partial_{jk} g(C_{[(i-1)/k_n]k_n \Delta_n}) \rho(u,v)_{i+2l_n-1} \zeta''(v)_{i+2l_n-1}^n.\end{aligned}$$

Similar to the proof of (B.104) - (B.106) of [Ait-Sahalia and Jacod \(2014\)](#), and using the fact that  $g(\cdot)$  and its derivatives are bounded, we can show that

$$\begin{aligned}& \Delta_n^{-1/2} \sum_{i=1}^{k_n(N_t^n+1)} \partial_{jk} g(C_{[(i-1)/k_n]k_n \Delta_n}) \partial_{lm} g(C_{[(i-1)/k_n]k_n \Delta_n}) \rho(u,v)_{i+2l_n-1} \\ & \times \rho(u',v')_{i+2l_n-1} \mathbf{E}_{i-1} (\zeta''(v)_{i+2l_n-1}^n \zeta''(v')_{i+2l_n-1}^n) \xrightarrow{p} \int_0^t \partial_{jk} g(C_s) \partial_{lm} g(C_s) l(u,v,u',v'; C_s) ds, \\ & \Delta_n^{-1/2} \sum_{i=1}^{k_n(N_t^n+1)} \|\partial g(C_{[(i-1)/k_n]k_n \Delta_n})\|^4 \|\rho(u,v)_{i+2l_n-1}\|^4 \mathbf{E}_{i-1} (\|\zeta''(v)_{i+2l_n-1}^n\|^4) \xrightarrow{p} 0, \\ & \Delta_n^{-1/4} \sum_{i=1}^{k_n(N_t^n+1)} \partial_{jk} g(C_{[(i-1)/k_n]k_n \Delta_n}) \rho(u,v)_{i+2l_n-1} \mathbf{E}_{i-1} (\zeta''(v)_{i+2l_n-1}^n \Delta_{i+2l_n-1}^n N) \xrightarrow{p} 0,\end{aligned}$$

where  $N$  is any bounded martingale orthogonal to  $W$  or  $W^l$  for some  $l$ . Therefore, [\(A.29\)](#) follows immediately from Theorem IX.7.28 of [Jacod and Shiryaev \(2003\)](#).

To derive the asymptotic variance for  $U(5)_t^n$ , we recall that

$$\eta_i^{12} = \frac{1}{k_n \Delta_n} \sum_{u=1}^2 (Z(u,3)_i^n + Z(3,u)_i^n), \quad \eta_i^{22} = 2 \frac{1}{k_n \Delta_n} \sum_{u,v=1}^2 Z(u,v)_i, \quad \text{and} \quad \eta_i^{11} = 0.$$

Therefore, the following expression leads to the desired formula:

$$h(j,k,l,m; C_s) = \begin{cases} 0, & \text{if } (j,k) \text{ or } (l,m) = (1,1); \\ 2 \sum_{u,u',v'=1}^2 (l(u,3,u',v') + l(3,u,u',v')), & \text{if } (j,k,l,m) = (1,2,1,2), (2,1,2,1), \\ & (2,1,1,2), (1,2,2,1); \\ 2 \sum_{u,u',v'=1}^2 (l(u,3,u',v') + l(3,u,u',v')), & \text{if } (j,k,l,m) = (1,2,2,2), (2,1,2,2), \\ & (2,2,2,1), (2,2,1,2); \\ 4 \sum_{u,v,u',v'=1}^2 l(u,v,u',v'), & \text{if } (j,k,l,m) = (2,2,2,2). \end{cases}$$

$$= \begin{cases} 0, & \text{if } (j, k) \text{ or } (l, m) = (1, 1); \\ \frac{8}{3\beta} c_s^3 + \frac{23\beta}{15} c_s^{(c)} c_s + \frac{23\beta}{30} (c_s^{(X,c)})^2, & \text{if } (j, k, l, m) = (1, 2, 1, 2), (2, 1, 2, 1), \\ & (2, 1, 1, 2), (1, 2, 2, 1); \\ \frac{5}{\beta} (c_s)^2 c_s^{(X,c)} + \frac{151\beta}{120} c_s^{(c)} c_s^{(X,c)}, & \text{if } (j, k, l, m) = (1, 2, 2, 2), (2, 1, 2, 2), \\ & (2, 2, 2, 1), (2, 2, 1, 2); \\ \frac{48}{\beta^3} c_s^4 + \frac{12}{\beta} (c_s)^2 c_s^{(c)} + \frac{151\beta}{70} (c_s^{(c)})^2, & \text{if } (j, k, l, m) = (2, 2, 2, 2). \end{cases}$$

Finally, we turn to  $U(6)_t^n$ . By (A.26), (A.27), (A.28), and under the condition that  $r < 1/2$  and  $3/(8 - 4r) \leq \varpi < 1/2$ , we have

$$\begin{aligned} k_n \Delta_n^{3/4} \mathbb{E} \sum_{i=0}^{N_t^n} |\hat{c}_{ik_n+1} - \hat{c}'_{ik_n+1}| &\leq K \phi_n \Delta_n^{(2-r)\varpi - \frac{1}{4}} \rightarrow 0, \\ k_n \Delta_n^{3/4} \mathbb{E} \sum_{i=0}^{N_t^n} |\hat{c}_{ik_n+1}^{(X,c)} - \hat{c}'_{ik_n+1}{}^{(X,c)}| &\leq K \phi_n \left( \Delta_n^{(1-2r)/4} + \Delta_n^{(2-r)\varpi - 3/4} \right) \rightarrow 0, \\ k_n \Delta_n^{3/4} \mathbb{E} \sum_{i=0}^{N_t^n} |\hat{c}_{ik_n+1}^{(c)} - \hat{c}'_{ik_n+1}{}^{(c)}| &\leq K \phi_n \left( \Delta_n^{(1-2r)/4} + \Delta_n^{(2-r)\varpi - 3/4} \right) \rightarrow 0. \end{aligned}$$

Then by the mean-value theorem and the boundedness of  $g$  and its derivatives, we have

$$k_n \Delta_n^{3/4} \sum_{i=0}^{N_t^n} \left( g \left( \hat{C}_{ik_n+1} \right) - g \left( \hat{C}'_{ik_n+1} \right) \right) = o_p(1).$$

Applying the mean-value theorem again yields:

$$\Delta_n^{1/4} \sum_{i=0}^{N_t^n} \left( \sum_{j,k,l,m} \partial_{jklm} g \left( \hat{C}'_{ik_n+1} \right) h \left( j, k; l, m; \hat{C}'_{ik_n+1} \right) - \partial_{jklm} g \left( \hat{C}_{ik_n+1} \right) h \left( j, k; l, m; \hat{C}_{ik_n+1} \right) \right),$$

which is  $o_p(k_n^{-1} \Delta_n^{-1/2})$ . Therefore,  $U(6)_t^n = o_p(1)$ , which concludes the proof.

#### A.4 Return to the Proof of Theorem 1

Theorem 1 is a special case of Theorem A1 when  $g(C_s) = C_s^{12} (C_s^{11})^{-1/2} (C_s^{22})^{-1/2}$ . Apparently,  $g$  is a  $C^3$  function on any convex subset of  $\mathcal{M}_2$  that satisfies  $C_s^{11} > 0$  and  $C_s^{22} > 0$ , therefore the spatial localization assumption is satisfied. Moreover, we have

$$\begin{aligned} \partial_{12} g(C_s) &= \frac{1}{\sqrt{C_s^{22} C_s^{11}}}, & \partial_{22} g(C_s) &= -\frac{C_s^{12}}{2\sqrt{(C_s^{22})^3 C_s^{11}}}, & \partial_{11} g(C_s) &= -\frac{C_s^{12}}{2\sqrt{(C_s^{11})^3 C_s^{22}}}. \\ \partial_{12,12}^2 g(C_i) &= 0, & \partial_{12,22}^2 g(C_i) &= -\frac{1}{2} \frac{1}{\sqrt{(C_s^{22})^3 C_s^{11}}}, & \partial_{22,22}^2 g(C_i) &= \frac{3}{4} \frac{C_s^{12}}{\sqrt{(C_s^{22})^5 C_s^{11}}}. \end{aligned}$$

Recall that  $C_s^{11}$ ,  $C_s^{12}$  and  $C_s^{22}$  can be written as  $c_s$ ,  $c_s^{(X,c)}$ , and  $c_s^{(c)}$ , respectively. Therefore, we can recover the expression for the bias-correction part of Theorem 1 from Theorem A1,

$$\begin{aligned} & \sum_{j,k,l,m}^2 \partial_{jk,lm} g(\hat{C}_{ik_n+1}) h(j, k; l, m; \hat{C}_{ik_n+1}) \\ &= -\frac{5}{\beta} \frac{c_s^{3/2} c_s^{(X,c)}}{(c_s^{(c)})^{3/2}} - \frac{151\beta}{120} \frac{c_s^{(X,c)}}{(c_s)^{1/2} (c_s^{(c)})^{1/2}} + \frac{36}{\beta} \frac{c_s^{(X,c)} c_s^{7/2}}{(c_s^{(c)})^{3/2}} + \frac{9}{\beta} \frac{c_s^{(X,c)} (c_s)^{3/2}}{(c_s^{(c)})^{1/2}} + \frac{453}{280} \beta \frac{c_s^{(X,c)} (c_s^{(c)})^{1/2}}{c_s^{1/2}}. \end{aligned}$$

Similarly, we obtain the expression for  $V_t^{PRL}$  in Theorem 1 from Theorem A1,

$$\begin{aligned} V_t^{PRL} &= \int_0^t \sum_{j,k,l,m}^2 \partial_{jk} g(C_s) \partial_{lm} g(C_s) h(j, k, l, m; C_s) ds \\ &= \int_0^t \left( \frac{23}{15} \beta + \frac{8}{3\beta} \frac{c_s^2}{c_s^{(c)}} - \frac{2}{\beta} \frac{c_s (c_s^{(X,c)})^2}{(c_s^{(c)})^2} + \frac{12 c_s^3 (c_s^{(X,c)})^2}{\beta^3 (c_s^{(c)})^3} + \frac{1}{21} \beta \frac{(c_s^{(X,c)})^2}{c_s c_s^{(c)}} \right) ds. \end{aligned}$$

## B Proof of Theorem 2

Below we introduce  $K$  as a generic constant which may change from line to line. Let  $v_t$  be the spot covariation between the continuous parts of  $X$  and  $Z$ . By the localization procedure (see Lemma 4.4.9 in Jacod and Protter (2012)), we can assume that  $\|v_t\| \leq A_0$  and  $\underline{\lambda}(v_t) > A_0^{-1}$ , where  $\underline{\lambda}(\cdot) = \min(v_{11}, v_{22})$ , for any  $t$ .

Next, introduce the ‘‘correlation’’ function  $g(\cdot)$  and its regularized version  $g_A(\cdot)$ , for any  $A > 0$ , as follows,

$$g(v) = \frac{v_{12}}{\sqrt{v_{11}}\sqrt{v_{22}}} 1_{\{v_{11}>0, v_{22}>0\}} \quad \text{and} \quad g_A(v) = \begin{cases} g(v) & \text{if } \underline{\lambda}(v) > (2A)^{-1} \\ 0 & \text{if } \underline{\lambda}(v) \leq (4A)^{-1}. \end{cases}$$

Notice that  $g_A(\cdot)$  is a  $C^\infty$  function which satisfies the polynomial growth condition in Jacod and Rosenbaum (2013), namely,  $|\partial^j g_A(v)| \leq K(1 + \|v\|^{3-j})$ , for  $j = 0, 1, 2, 3$ , where  $\partial^j$  denotes the  $j$ th derivatives.

Our spot leverage estimator is  $\hat{\rho}_i = g(\hat{v}_i)$ , where

$$\hat{v}_i = \begin{pmatrix} \hat{v}_i^{11} & \hat{v}_i^{12} \\ \hat{v}_i^{21} & \hat{v}_i^{22} \end{pmatrix}, \quad \text{and}$$

$$\begin{aligned} \hat{v}_i^{11} &= \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n-1} (\Delta_{i+j}^n Z)^2 \cdot 1_{\{|\Delta_{i+j}^n Z| \leq u'_n\}}, \\ \hat{v}_i^{12} &= \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n-1} (\Delta_{i+j}^n Z) (\Delta_{i+j}^n X) \cdot 1_{\{|\Delta_{i+j}^n Z| \leq u'_n\}} 1_{\{|\Delta_{i+j}^n X| \leq u'_n\}}, \end{aligned}$$

$$\hat{v}_i^{22} = \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n-1} (\Delta_{i+j}^n X)^2 \cdot 1_{\{|\Delta_{i+j}^n X| \leq u'_n\}}.$$

We then apply Theorem 3.2 in [Jacod and Rosenbaum \(2013\)](#), which concludes that for any  $C^3$  function  $g$  that satisfies the above polynomial growth condition, we have

$$\Delta_n^{-1/2} \left\{ k_n \Delta_n \sum_{i=0}^{[t/k_n \Delta_n]-1} \left( g(\hat{v}_{ik_n+1}) - \int_0^t g(v_s) ds - \frac{1}{2k_n} \sum_{j,k,l,m=1}^2 \partial_{jk,lm}^2 g(\hat{v}_{ik_n+1}) \right. \right. \\ \left. \left. (\hat{v}_{ik_n+1}^{jl} \hat{v}_{ik_n+1}^{km} + \hat{v}_{ik_n+1}^{jm} \hat{v}_{ik_n+1}^{kl}) \right) \right\} \xrightarrow{L\text{-}s} \sqrt{V_t^{IRL}} \mathcal{Z}_t,$$

where  $\mathcal{Z}$  is a standard normal random variable defined on the extension of the original probability space, and  $V_t^{IRL}$  is the conditional variance given by

$$V_t^{IRL}(g) = \sum_{j,k,l,m=1}^2 \int_0^t \partial_{jk} g(v_s) \partial_{lm} g(v_s) (v_s^{jl} v_s^{km} + v_s^{jm} v_s^{kl}) ds.$$

Denote the bias-corrected estimators by  $\hat{G}_A$  and  $\hat{G}$ , which are based on  $g_A(\cdot)$  and  $g(\cdot)$ , respectively. The above theorem does not directly apply to  $\hat{G}$ , but it does establish the asymptotic distribution of  $\hat{G}_A$ . Note that on  $\Omega_t^{A_0} = \{\underline{\lambda}(v_s) > A_0^{-1}, \text{ for any } s \in [0, t]\}$ , we have

$$\int_0^t g_{A_0}(v_s) ds = \int_0^t g(v_s) ds \quad \text{and} \quad V_t^{IRL}(g_{A_0}) = V_t^{IRL}(g).$$

Therefore, in order to establish the asymptotic distribution for  $\hat{G}$ , it remains to show that

$$\Delta_n^{-1/2} \left\{ k_n \Delta_n \sum_{i=0}^{[t/k_n \Delta_n]-1} \left( g(\hat{v}_{ik_n+1}) - g_{A_0}(\hat{v}_{ik_n+1}) \right) \right\} \xrightarrow{p} 0 \quad \text{and} \quad (\text{B.30})$$

$$\Delta_n^{-1/2} \left\{ \Delta_n \sum_{i=0}^{[t/k_n \Delta_n]-1} \sum_{j,k,l,m=1}^2 \partial_{jk,lm}^2 g(\hat{v}_{ik_n+1}) (\hat{v}_{ik_n+1}^{jl} \hat{v}_{ik_n+1}^{km} + \hat{v}_{ik_n+1}^{jm} \hat{v}_{ik_n+1}^{kl}) \right\} \xrightarrow{p} 0. \quad (\text{B.31})$$

In fact, by the Lipschitz continuity of  $\underline{\lambda}(\cdot)$ , there exists some  $\epsilon > 0$  such that for any  $i$ ,

$$\|\hat{v}_{i\Delta_n} - v_{i\Delta_n}\| \leq \epsilon \text{ implies } |\underline{\lambda}(\hat{v}_{i\Delta_n}) - \underline{\lambda}(v_{i\Delta_n})| \leq (2A_0)^{-1},$$

which implies  $\underline{\lambda}(\hat{v}_{i\Delta_n}) > (2A_0)^{-1}$  on  $\Omega_t^{A_0}$ . Hence, we have

$$\Delta_n^{1/2} k_n \sum_{i=0}^{[t/k_n \Delta_n]-1} |g(\hat{v}_{ik_n+1}) - g_{A_0}(\hat{v}_{ik_n+1})| \\ \leq \Delta_n^{1/2} k_n \sum_{i=0}^{[t/k_n \Delta_n]-1} (|g(\hat{v}_{ik_n+1})| + |g_{A_0}(\hat{v}_{ik_n+1})|) 1_{\{|\underline{\lambda}(\hat{v}_{ik_n+1})| \leq (2A_0)^{-1}\}}$$

$$\leq 2\Delta_n^{1/2}k_n \sum_{i=0}^{[t/k_n\Delta_n]-1} 1_{\{\|\hat{v}_{ik_n+1}-v_{ik_n+1}\|>\epsilon'\}}.$$

On the other hand, for any  $\epsilon' > 0$ ,

$$\Delta_n^{1/2}k_n \sum_{i=0}^{[t/k_n\Delta_n]-1} 1_{\{\|\hat{v}_{ik_n+1}-v_{ik_n+1}\|>\epsilon'\}} \xrightarrow{p} 0,$$

since

$$\begin{aligned} & \Delta_n^{1/2}k_n \sum_{i=0}^{[t/k_n\Delta_n]-1} \mathbb{E} \left( 1_{\{\|\hat{v}_{ik_n+1}-v_{ik_n+1}\|>\epsilon'\}} \right) = \Delta_n^{1/2}k_n \sum_{i=0}^{[t/k_n\Delta_n]-1} \mathbb{P}(\|\hat{v}_{ik_n+1} - v_{ik_n+1}\| > \epsilon') \\ & \leq \Delta_n^{1/2}k_n \frac{K}{\epsilon'^3} \sum_{i=0}^{[t/k_n\Delta_n]-1} \left\{ \mathbb{E} \|\hat{v}_{ik_n+1} - \hat{v}'_{ik_n+1}\|^3 + \mathbb{E} \|\hat{v}'_{ik_n+1} - v_{ik_n+1}\|^3 \right\} \\ & \leq \frac{K}{\epsilon'^3} \left( a_n \Delta_n^{(6-r)\varpi-3+1/2} + k_n^{-3/2} \Delta_n^{-1/2} + k_n \Delta_n^{1/2} \right) \rightarrow 0, \end{aligned}$$

where the last inequality follows by (4.8) and (4.11) of [Jacod and Rosenbaum \(2013\)](#), the fact that  $\varpi \geq 5/(12 - 2r)$ , and that  $a_n \rightarrow 0$  and  $k_n^2 \Delta_n \rightarrow 0$  as  $\Delta_n \rightarrow 0$ . Hence, by (2.2.35) of [Jacod and Protter \(2012\)](#), we have

$$\Delta_n^{1/2}k_n \sum_{i=0}^{[t/k_n\Delta_n]-1} |g(\hat{v}_{ik_n+1}) - g_{A_0}(\hat{v}_{ik_n+1})| \xrightarrow{p} 0,$$

and (B.30) holds. To show (B.31), note that

$$\sum_{j,k,l,m=1}^2 \partial_{jk,lm}^2 g(\hat{v}_{ik_n+1}) (\hat{v}_{ik_n+1}^{jl} \hat{v}_{ik_n+1}^{km} + \hat{v}_{ik_n+1}^{jm} \hat{v}_{ik_n+1}^{kl}) = (\hat{\rho}_i^n)^3 - \hat{\rho}_i^n$$

and that

$$\begin{aligned} & \Delta_n^{1/2} \left| \sum_{i=0}^{[t/k_n\Delta_n]-1} \sum_{j,k,l,m=1}^2 \partial_{jk,lm}^2 \left( g(\hat{v}_{ik_n+1}) - g_{A_0}(\hat{v}_{ik_n+1}) \right) 1_{\{\lambda(\hat{v}_{i\Delta_n}) \leq (2A_0)^{-1}\}} (\hat{v}_{ik_n+1}^{jl} \hat{v}_{ik_n+1}^{km} + \hat{v}_{ik_n+1}^{jm} \hat{v}_{ik_n+1}^{kl}) \right| \\ & \leq \Delta_n^{1/2} \sum_{i=0}^{[t/k_n\Delta_n]-1} \left| (\hat{\rho}_{ik_n+1}^n)^3 - \hat{\rho}_{ik_n+1}^n - \sum_{j,k,l,m=1}^2 \partial_{jk,lm}^2 g_{A_0}(\hat{v}_{ik_n+1}) (\hat{v}_{ik_n+1}^{jl} \hat{v}_{ik_n+1}^{km} + \hat{v}_{ik_n+1}^{jm} \hat{v}_{ik_n+1}^{kl}) \right| \\ & \quad \times 1_{\{\lambda(\hat{v}_{ik_n+1}) \leq (2A_0)^{-1}\}} \\ & \leq K \Delta_n^{1/2} \sum_{i=0}^{[t/k_n\Delta_n]-1} \left( 1 + (1 + \|\hat{v}_{ik_n+1}\|) \|\hat{v}_{ik_n+1}\|^2 \right) 1_{\{\|\hat{v}_{ik_n+1}-v_{ik_n+1}\|>\epsilon'\}}, \end{aligned}$$

and that for any  $\epsilon' > 0$ , by triangle inequality  $\|\hat{v}_{ik_n+1}\| \leq \|\hat{v}_{ik_n+1} - v_{ik_n+1}\| + \|v_{ik_n+1}\|$ , and

$$1_{\{\|\hat{v}_{ik_n+1}-v_{ik_n+1}\|>\epsilon'\}} \leq \frac{\|\hat{v}_{ik_n+1} - v_{ik_n+1}\|^j}{\epsilon'^j}, \text{ for } j = 0, 1, 2, 3,$$

we have

$$\begin{aligned} & \Delta_n^{1/2} \sum_{i=0}^{[t/k_n \Delta_n]-1} \mathbb{E} \left( \left( 1 + (1 + \|\hat{v}_{ik_n+1}\|) \|\hat{v}_{ik_n+1}\|^2 \right) 1_{\{\|\hat{v}_{ik_n+1} - v_{ik_n+1}\| > \epsilon'\}} \right) \\ & \leq K \Delta_n^{1/2} \sum_{i=0}^{[t/k_n \Delta_n]-1} \left( \mathbb{E} \|\hat{v}_{ik_n+1} - \hat{v}'_{ik_n+1}\|^3 + \mathbb{E} \|\hat{v}'_{ik_n+1} - v_{ik_n+1}\|^3 \right). \end{aligned}$$

Hence, (B.31) follows from the same argument as above.

As a result, the CLT holds for  $\hat{G}$ . By direct calculations, we have

$$\sum_{j,k,l,m=1}^2 \int_0^t \partial_{jk} g(v_s) \partial_{lm} g(v_s) (v_s^{jl} v_s^{km} + v_s^{jm} v_s^{kl}) ds = \int_0^t (1 - 2\rho_s^2 + \rho_s^4) ds.$$

This establishes the asymptotic distribution of the *IRL* estimator.

To estimate the asymptotic variance, we define

$$\hat{V}_t^{IRL} = \frac{\Delta_n k_n}{t^2} \sum_{i=0}^{[t/k_n \Delta_n]-1} (1 - 2\hat{\rho}_{ik_n+1}^2 + \hat{\rho}_{ik_n+1}^4). \quad (\text{B.32})$$

Then, by Theorem 9.4.1 of [Jacod and Protter \(2012\)](#) and the same localization argument, we have

$$\hat{V}_t^{IRL} \xrightarrow{p} \frac{1}{t^2} \int_0^t (1 - 2\rho_s^2 + \rho_s^4) ds.$$

## C Proof of Theorem 3

The proof follows immediately from Theorems 1 and 2.

## D Proof of Lemma A2

For the (A.5), we can assume  $q \geq 2$  without loss of generality. The case with  $q \geq 1$  then follows by Jensen's inequality.

Each element of the term  $\beta_i = \hat{C}_i - C_{(i-1)\Delta_n} = \kappa_i + \eta_i + \bar{\beta}_i$  is a sum of the 10 types of terms below, and each element of the term  $\alpha_i = \kappa_i + \eta_i$  is a sum of the first 9 types of terms below,



- (a)  $\frac{1}{k_n \Delta_n} A(1; 2, 3)_i^n - \bar{C}_{(i-1)\Delta_n}^{12}, \frac{1}{k_n \Delta_n} A(1; 2, 2)_i^n - \bar{C}_{(i-1)\Delta_n}^{22}, \frac{1}{k_n \Delta_n} A(1; 1, 1)_i^n - \frac{1}{k_n \Delta_n} \bar{A}(0)_i^n,$
- (b)  $\frac{1}{k_n \Delta_n} \bar{A}(1; u)_i^n, u = 1, 2,$
- (c)  $\frac{1}{k_n \Delta_n} \bar{A}(2; u, v)_i^n, u = 1, 2, v = 1, 2,$  also,  $\bar{A}(3; u, v)_i^n, u = 1, 2, v = 1, 2,$
- (d)  $\frac{1}{k_n \Delta_n} A(1; u, v)_i^n, (u, v) = (1, 2), (2, 1), (1, 3),$
- (e)  $\frac{1}{k_n \Delta_n} A(2; u, v)_i^n$  for all  $(u, v)$  except  $(3, 3),$
- (f)  $\frac{1}{k_n \Delta_n} A(3; u, v)_i^n$  for all  $(u, v)$  except  $(3, 3),$
- (g)  $\frac{1}{k_n \Delta_n} A(4; u, v)_i^n$  for all  $(u, v)$  except  $(3, 3),$
- (h)  $\frac{1}{k_n \Delta_n} Z(u, 3)_i^n, \frac{1}{k_n \Delta_n} Z(3, u)_i^n$  for  $u = 1, 2,$
- (i)  $\frac{1}{k_n \Delta_n} Z(u, v)_i^n$  for  $(u, v) = (1, 2), (2, 1), (1, 1), (2, 2),$
- (j)  $\beta_i^{lm}, l = 1, 2, m = 1, 2.$

Therefore, to prove (A.4) it is sufficient to show that for each of the terms above in (a)-(j), the bound in (A.4) holds. Similarly, to prove (A.5) it is sufficient to show that for each of the terms above in (a)-(i), the bound in (A.5) holds. For most of the terms (all except cases (h) and (i)) we will in fact establish a tighter bound than in (A.4). These tighter bounds will be useful for the proof of Lemma A3.

(a) By (A.16), (A.17), (A.18), and (A.20), we have

$$\begin{aligned} \left| \mathbb{E}_{i-1} \frac{1}{k_n \Delta_n} A(1; 2, 3)_i^n - \bar{C}_{(i-1)\Delta_n}^{12} \right| &\leq K \Delta_n^{1/2}, \\ \mathbb{E}_{i-1} \left| \frac{1}{k_n \Delta_n} A(1; 2, 3)_i^n - \bar{C}_{(i-1)\Delta_n}^{12} \right|^q &\leq K \left( \Delta_n^{q/2} + k_n^{-q/2} \right). \end{aligned}$$

The proof for the second term in (a) is similar. As to the third term, note that by (A.10),

$$\begin{aligned} &\left| \mathbb{E}_{i-1} \left( \frac{1}{k_n \Delta_n} A(1; 1, 1)_i^n - \frac{1}{k_n \Delta_n} \bar{A}(0)_i^n \right) \right| \\ &= \frac{1}{l_n^2} \frac{1}{k_n \Delta_n} \sum_{s=0}^{k_n-1} \left| \mathbb{E}_{i-1} \left( (\zeta(1)_{i+s}^n)^2 - 2(c_{(i+s-1)\Delta_n})^2 \right) \right| \leq K \frac{1}{l_n^2} \frac{1}{k_n \Delta_n} k_n \sqrt{\Delta_n} \leq K \Delta_n^{1/2}. \end{aligned}$$

In addition, we have

$$\begin{aligned} &\mathbb{E}_{i-1} \left| \frac{1}{k_n \Delta_n} A(1; 1, 1)_i^n - \frac{1}{k_n \Delta_n} \bar{A}(0)_i^n \right|^q \\ &\leq K \left( \frac{1}{k_n \Delta_n} \right)^q l_n^{-2q} \mathbb{E}_{i-1} \left| \sum_{s=0}^{k_n-1} \left[ \mathbb{E}_{i+s-1} \left( (\zeta(1)_{i+s}^n)^2 - 2(c_{(i+s-1)\Delta_n})^2 \right) \right] \right|^q \\ &+ K \left( \frac{1}{k_n \Delta_n} \right)^q l_n^{-2q} \mathbb{E}_{i-1} \left| \sum_{s=0}^{k_n-1} \left[ (\zeta(1)_{i+s}^n)^2 - \mathbb{E}_{i+s-1} \left( (\zeta(1)_{i+s}^n)^2 \right) \right] \right|^q \\ &\leq K \left( \frac{1}{k_n \Delta_n} \right)^q l_n^{-2q} k_n^q \Delta_n^{q/2} + K \left( \frac{1}{k_n \Delta_n} \right)^q l_n^{-2q} k_n^{q/2-1} \sum_{s=0}^{k_n-1} \mathbb{E}_{i-1} |\zeta(1)_{i+s}^n|^{2q} \end{aligned}$$

$$\begin{aligned}
& + K \left( \frac{1}{k_n \Delta_n} \right)^q l_n^{-2q} k_n^{q/2-1} \sum_{s=0}^{k_n-1} \mathbb{E}_{i-1} \left| \mathbb{E}_{i+s-1} \left( \zeta(1)_{i+s}^n \right)^2 \right|^q \\
& \leq K l_n^{-2q} \Delta_n^{-q/2} + K \Delta_n^{-q} l_n^{-2q} k_n^{-q/2} \\
& \leq K k_n^{-q/2} \leq K (k_n \Delta_n)^{q/2}.
\end{aligned}$$

In above, the second inequality follows by Burkholder-Gundy inequality and (A.10), and the third inequality follows by (A.7).

(b) Rewrite  $\bar{A}(1; u)_i^n$  as follows,

$$\begin{aligned}
\bar{A}(1; u)_i^n & = \frac{12}{l_n^3} \sum_{m=0}^{l_n-1} \sum_{j=0}^{k_n-1} c_{(i+j-1)\Delta_n} \bar{\varepsilon}(u)_m^n \zeta(u)_{i+j+m}^n \\
& = \frac{12}{l_n^3} \sum_{s=0}^{k_n+l_n-2} \sum_{j=0 \vee (s-l_n+1)}^{(k_n-1) \wedge s} c_{(i+j-1)\Delta_n} \bar{\varepsilon}(u)_{s-j}^n \zeta(u)_{i+s}^n \\
& = \frac{12}{l_n^3} \sum_{s=0}^{k_n+l_n-2} a_{i,s} \zeta(u)_{i+s}^n,
\end{aligned}$$

where  $a_{i,s} = \sum_{j=0 \vee (s-l_n+1)}^{(k_n-1) \wedge s} c_{(i+j-1)\Delta_n} \bar{\varepsilon}(u)_{s-j}^n$ . By (A.6) and (A.8), we have

$$\left| \mathbb{E}_{i-1} \left( \bar{A}(1; u)_i^n \right) \right| \leq K l_n^{-3} l_n \sum_{s=0}^{k_n+l_n-2} \sup |\bar{\varepsilon}(u)^n| \left| \mathbb{E}_{i-1} \zeta(u)_{i+s}^n \right| \leq K l_n^{-2} k_n \Delta_n^{1/2},$$

hence it follows that

$$\left| \mathbb{E}_{i-1} \left( \frac{1}{k_n \Delta_n} \bar{A}(1; u)_i^n \right) \right| \leq K l_n^{-1}.$$

Moreover, by (A.24),

$$\mathbb{E}_{i-1} \left| \bar{A}(1; u)_i^n \right|^q = \left( \frac{12}{l_n^3} \right)^q \mathbb{E}_{i-1} \left| \sum_{s=0}^{k_n-1} a_{i,s} \zeta(u)_{i+s}^n \right|^q \leq \begin{cases} K l_n^{-2q} k_n^{q/2} & \text{if } u = 1 \\ K l_n^{-q} k_n^{-q/2} & \text{if } u = 2 \end{cases}.$$

Hence,

$$\mathbb{E}_{i-1} \left| \frac{1}{k_n \Delta_n} \bar{A}(1; u)_i^n \right|^q \leq \begin{cases} K k_n^{-q/2} & \text{if } u = 1 \\ K k_n^{-3q/2} \Delta_n^{-q/2} & \text{if } u = 2 \end{cases}.$$

(c) The first term, by Hölder and Cauchy-Schwartz inequalities and (A.7) and (A.9), can be bounded by

$$\mathbb{E}_{i-1} \left| \bar{A}(2; u, v)_i^n \right|^q \leq k_n^{q-1} \sup_{s,j} \left| \bar{\gamma}(u, v; 0)_{s,j}^n \right|^q \sum_{j=1}^{k_n+l_n-2} \sqrt{\mathbb{E}_{i-1} \left| \zeta(u)_{i+j}^n \right|^{2q}} \sqrt{\mathbb{E}_{i-1} \left| \zeta(v)_{i+j}^n \right|^{2q}},$$

which is bounded by  $Kk_n^q \Delta_n^{3q/2}$  for every  $(u, v) = (1, 2), (2, 1), (2, 2), (1, 1)$ . Hence,

$$\left| \mathbb{E}_{i-1} \left( \frac{1}{k_n \Delta_n} \bar{A}(2; u, v)_i^n \right) \right| \leq K \Delta_n^{1/2}.$$

As to  $\bar{A}(3; u, v)_i^n$ , we have

$$\begin{aligned} & \left| \mathbb{E}_{i-1} \left( \sum_{m=1}^{(l_n-1) \wedge j} \bar{\gamma}(u, v; m)_{j-k_n+1, j}^n \zeta(u)_{i+j-m}^n \zeta(v)_{i+j}^n \right) \right| \\ & \leq \sum_{m=1}^{(l_n-1) \wedge j} \bar{\gamma}(u, v; m)_{j-k_n+1, j}^n \sqrt{\mathbb{E}_{i-1} \left( \zeta(u)_{i+j-m}^n \right)^2} \sqrt{\mathbb{E}_{i-1} \left( \zeta(v)_{i+j}^n \right)^2}, \end{aligned}$$

which is bounded by  $K \Delta_n^{3/2}$ . Therefore, we have

$$\left| \mathbb{E}_{i-1} \left( \frac{1}{k_n \Delta_n} \bar{A}(2; u, v)_i^n \right) \right| \leq K \Delta_n^{1/2}.$$

Moreover, note that by (A.24),

$$\mathbb{E}_{i-1} \left| \sum_{m=1}^{(l_n-1) \wedge j} \bar{\gamma}(u, v; m)_{j-k_n+1, j}^n \zeta(u)_{i+j-m}^n \right|^q \leq \begin{cases} Kl_n^{-3q/2} & \text{if } v = 2 \\ Kl_n^{-5q/2} & \text{if } v = 1 \end{cases}.$$

Therefore, we can obtain

$$\mathbb{E}_{i-1} \left| \bar{A}(3; u, v)_i^n \right|^q \leq K k_n^{q-1} \sum_{j=1}^{k_n+l_n-2} \left[ \mathbb{E}_{i-1} \left| \sum_{m=1}^{(l_n-1) \wedge j} \bar{\gamma}(u, v; m)_{j-k_n+1, j}^n \zeta(u)_{i+j-m}^n \right|^q \mathbb{E}_{i+j-1} \left| \zeta(v)_{i+j}^n \right|^q \right]$$

which is bounded by  $Kk_n^q \Delta_n^{5q/4}$  for every  $(u, v) = (1, 2), (2, 1), (2, 2), (1, 1)$ . The first inequality follows by the Hölder inequality; the final bound by (A.7) and (A.9) and the result just above.

Hence, for  $j = 1, 2$  and  $(u, v) = (1, 2), (2, 1), (2, 2), (1, 1)$ ,

$$\mathbb{E}_{i-1} \left| \frac{1}{k_n \Delta_n} \bar{A}(j; u, v)_i^n \right|^q \leq K \Delta_n^{q/4} \leq K (k_n \Delta_n)^{q/2}.$$

(d) Here we look at the term  $A(1; u, v)_i^n$ , say for  $(u, v) = (1, 2)$ . We have

$$\mathbb{E}_{i-1} |A(1; 1, 2)_i^n|^q \leq K l_n^{-2q} k_n^{q-1} \sum_{s=0}^{k_n-1} \sqrt{\mathbb{E}_{i-1} |\zeta(1)_{i+s}^n|^{2q}} \sqrt{\mathbb{E}_{i-1} |\zeta(2)_{i+s}^n|^{2q}} \leq K l_n^{-2q} k_n^q \Delta_n^{q/2},$$

by the Hölder and Cauchy-Schwartz inequalities, (A.7), and (A.9). Hence,

$$\mathbb{E}_{i-1} \left| \frac{1}{k_n \Delta_n} A(1; 1, 2)_i^n \right|^q \leq K \Delta_n^{q/2}.$$

A similar argument shows the result for  $(u, v) = (2, 1)$  and  $(1, 3)$ . This also establishes that

$$\left| \mathbb{E}_{i-1} \left( \frac{1}{k_n \Delta_n} A(1; 1, 2)_i^n \right) \right| \leq K \Delta_n^{1/2}.$$

(e) To prove (A.3), notice that for  $s = 0, \dots, 2l_n - 1$ ,

$$\gamma(u, v; 0)_{s-k_n+1, s+1}^n + \gamma(u, v; 0)_{s+1, s+k_n+1}^n = y_{u,v}^n \sum_{q=0}^s \varepsilon(u)_q^n \varepsilon(v)_q^n + y_{u,v}^n \sum_{q=s+1}^{2l_n-1} \varepsilon(u)_q^n \varepsilon(v)_q^n = \Gamma(u, v)_i^n.$$

Therefore, we can rewrite  $A(2; u, v)_i^n$  as

$$A(2; u, v)_i^n = \sum_{s=0}^{2l_n-1} \left( \gamma(u, v; 0)_{s+1, s+k_n+1}^n - \Gamma(u, v)_0^n \right) \left( \zeta(u)_{i+s+k_n}^n \zeta(v)_{i+s+k_n}^n - \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n \right).$$

By (A.10), (A.11), and (A.12), we have

$$\begin{aligned} & \left| \mathbf{E}_{i-1} \left( \zeta(u)_{i+s+k_n}^n \zeta(v)_{i+s+k_n}^n - \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n \right) \right| \\ & \leq \left| \mathbf{E}_{i-1} \left( \zeta(u)_{i+s+k_n}^n \zeta(v)_{i+s+k_n}^n - z_{u,v}^n \Delta_n V(u, v)_{i+s+k_n}^n \right) \right| + z_{u,v}^n \Delta_n \left| \mathbf{E}_{i-1} \left( V(u, v)_{i+s+k_n}^n - V(u, v)_{i+s}^n \right) \right| \\ & + \left| \mathbf{E}_{i-1} \left( \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n - z_{u,v}^n \Delta_n V(u, v)_{i+s}^n \right) \right| \\ & \leq \begin{cases} K k_n \Delta_n^2, & \text{if } (u, v) = (2, 2), (2, 3), (3, 2) \\ K \Delta_n, & \text{if } (u, v) = (1, 2), (2, 1), (1, 3), (3, 1) \\ K k_n \Delta_n, & \text{if } (u, v) = (1, 1) \end{cases}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} & \left| \mathbf{E}_{i-1} \left( \frac{1}{k_n \Delta_n} A(2; u, v)_i^n \right) \right| \\ & \leq \frac{1}{k_n \Delta_n} \sum_{s=0}^{2l_n-1} \left| \gamma(u, v; 0)_{s+1, s+k_n+1}^n - \Gamma(u, v)_0^n \right| \left| \mathbf{E}_{i-1} \left( \zeta(u)_{i+s+k_n}^n \zeta(v)_{i+s+k_n}^n - \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n \right) \right| \\ & \leq K \Delta_n^{1/2}. \end{aligned}$$

Now we prove (A.4) and (A.5). Decompose  $A(2; u, v)_i^n$  into five parts arising from writing

$$\begin{aligned} & \zeta(u)_{i+s+k_n}^n \zeta(v)_{i+s+k_n}^n - \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n \\ & = \left[ \zeta(u)_{i+s+k_n}^n \zeta(v)_{i+s+k_n}^n - \mathbf{E}_{i+s+k_n} \zeta(u)_{i+s+k_n}^n \zeta(v)_{i+s+k_n}^n \right] \\ & - \left[ \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n - \mathbf{E}_{i+s-1} \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n \right] \\ & + \left[ \mathbf{E}_{i+s+k_n} \zeta(u)_{i+s+k_n}^n \zeta(v)_{i+s+k_n}^n - z_{u,v}^n \Delta_n V(u, v)_{i+s+k_n}^n \right] \\ & - \left[ \mathbf{E}_{i+s-1} \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n - z_{u,v}^n \Delta_n V(u, v)_{i+s}^n \right] \\ & + z_{u,v}^n \Delta_n V(u, v)_{i+s+k_n}^n - z_{u,v}^n \Delta_n V(u, v)_{i+s}^n. \end{aligned}$$

The first component of  $A(2; u, v)_i^n$  satisfies, by the Burkholder-Gundy, Hölder, Cauchy-Schwartz inequalities and (A.25), (A.7) and (A.9)

$$\mathbf{E}_{i-1} \left| \sum_{s=0}^{2l_n-1} \left( \gamma(u, v; 0)_{s+1, s+k_n+1}^n - \Gamma(u, v)_0^n \right) \left( \zeta(u)_{i+s+k_n}^n \zeta(v)_{i+s+k_n}^n - \mathbf{E}_{i+s+k_n} \zeta(u)_{i+s+k_n}^n \zeta(v)_{i+s+k_n}^n \right) \right|^q$$

$$\leq l_n^{q/2-1} \sum_{s=0}^{2l_n-1} \left| \gamma(u, v; 0)_{s+1, s+k_n+1}^n - \Gamma(u, v)_0^n \right|^q \sqrt{\mathbb{E}_{i-1} |\zeta(u)_{i+s+k_n}^n|^{2q}} \sqrt{\mathbb{E}_{i-1} |\zeta(v)_{i+s+k_n}^n|^{2q}},$$

which is bounded by  $K\Delta_n^{3q/4}$  for every  $(u, v)$ . Second component of  $A(2; u, v)_i^n$  is very similar. The third satisfies, by Hölder inequality, (A.10), (A.11), and (A.12),

$$\begin{aligned} & \mathbb{E}_{i-1} \left| \sum_{s=0}^{2l_n-1} \left( \gamma(u, v; 0)_{s+1, s+k_n+1}^n - \Gamma(u, v)_0^n \right) \left( \mathbb{E}_{i+s-1} \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n - z_{u,v}^n V(u, v)_{i+s}^n \right) \right|^q \\ & \leq Kl_n^{q-1} \sum_{s=0}^{2l_n-1} \left| \gamma(u, v; 0)_{s+1, s+k_n+1}^n - \Gamma(u, v)_0^n \right|^q \mathbb{E}_{i-1} \left( \left| \mathbb{E}_{i+s-1} \zeta(u)_{i+s}^n \zeta(v)_{i+s}^n - z_{u,v}^n V(u, v)_{i+s}^n \right|^q \right), \end{aligned}$$

which is bounded by  $K\Delta_n^q$ . The fourth term is similar. The fifth and final component of  $A(2; u, v)_i^n$  satisfies, by the Hölder inequality and (A.25),

$$\begin{aligned} & \mathbb{E}_{i-1} \left| \sum_{s=0}^{2l_n-1} \left( \gamma(u, v; 0)_{s+1, s+k_n+1}^n - \Gamma(u, v)_0^n \right) \left( z_{u,v}^n \Delta_n V(u, v)_{i+s+k_n}^n - z_{u,v}^n \Delta_n V(u, v)_{i+s}^n \right) \right|^q \\ & \leq l_n^{q-1} \sum_{s=0}^{2l_n-1} \left| \gamma(u, v; 0)_{s+1, s+k_n+1}^n - \Gamma(u, v)_0^n \right|^q |z_{u,v}^n \Delta_n|^q \mathbb{E}_{i-1} |V(u, v)_{i+s+k_n}^n - V(u, v)_{i+s}^n|^q, \end{aligned}$$

which is bounded by  $K\Delta_n^{q/2}$  for every  $(u, v)$ . Therefore,

$$\mathbb{E}_{i-1} \left| \frac{1}{k_n \Delta_n} A(2; u, v)_i^n \right|^q \leq K (k_n^{-2} \Delta_n^{-1})^{q/2}.$$

(f) We rewrite  $A(3; u, v)_i^n$  as follows,

$$A(3; u, v)_i^n = \sum_{s=0}^{k_n+2l_n-2} \eta(i, s) \zeta(v)_{i+s}^n = \sum_{s=0}^{2l_n-2} \eta(i, s) \zeta(v)_{i+s}^n + \sum_{s=k_n}^{k_n+2l_n-2} \eta(i, s) \zeta(v)_{i+s}^n, \quad (\text{D.33})$$

where  $\eta(i, s) = \sum_{m=1}^{(2l_n-1) \wedge s} \gamma(u, v; m)_{s-k_n+1, s+1}^n \zeta(u)_{i+s-m}^n - 1_{\{s \geq 2l_n-1\}} \rho(u, v)_{i+s}^n$ . Note that  $\eta(i, s)$  is  $\mathcal{F}_{i+s-1}$ -measurable. We will use the fact that by (A.7), (A.9), and (A.24), we have for all  $q \geq 2$  and all  $s$ ,

$$\begin{aligned} & \mathbb{E}_{i-1} \left| \left( \sum_{m=1}^{(2l_n-1) \wedge s} \gamma(u, v; m)_{s-k_n+1, s+1}^n \zeta(u)_{i+s-m}^n \right) \right|^q \\ & \leq \begin{cases} Kl_n^{-q/2} & \text{if } (u, v) = (2, 2), (3, 2), (2, 3) \\ K (l_n^{-1})^q l_n^{q/2} & \text{if } (u, v) = (1, 2), (1, 3) \\ K (l_n^{-1})^q l_n^{-q/2} & \text{if } (u, v) = (2, 1), (3, 1) \\ K (l_n^{-2})^q l_n^{q/2} & \text{if } (u, v) = (1, 1) \end{cases} = \begin{cases} Kl_n^{-q/2} & \text{if } v = 2, 3 \\ Kl_n^{-3q/2} & \text{if } v = 1 \end{cases}, \end{aligned}$$

which implies

$$\mathbb{E}_{i-1} |\eta(i, s)|^q \leq \begin{cases} Kl_n^{-q/2} & \text{if } v = 2, 3 \\ Kl_n^{-3q/2} & \text{if } v = 1 \end{cases}. \quad (\text{D.34})$$

We first prove (A.3). By (D.34), (A.6) and (A.8), the first part of  $A(3; u, v)_i^n$  satisfies

$$\left| \mathbb{E}_{i-1} \left( \sum_{s=0}^{2l_n-2} \eta(i, s) \zeta(v)_{i+s}^n \right) \right| \leq K \sum_{s=0}^{2l_n-2} \sqrt{\mathbb{E}_{i-1} (\eta(i, s))^2} \sqrt{\mathbb{E}_{i-1} (\zeta'(v)_{i+s}^n)^2} \leq K \Delta_n^{1/2}.$$

The same steps can be used for the second part of  $A(3; u, v)_i^n$ . Therefore,

$$\left| \mathbb{E}_{i-1} \left( \frac{1}{k_n \Delta_n} A(3; u, v)_i^n \right) \right| \leq K k_n^{-1} \Delta_n^{-1/4} \leq K k_n \Delta_n.$$

We now prove (A.4) and (A.5). The first part of  $A(3; u, v)_i^n$  satisfies, by the Burkholder-Gundy inequality, (A.6)-(A.9) and (D.34),

$$\mathbb{E}_{i-1} \left| \sum_{s=0}^{2l_n-2} \eta(i, s) \zeta(v)_{i+s}^n \right|^q \leq K \sum_{s=0}^{2l_n-2} \mathbb{E}_{i-1} \left( |\eta(i, s)|^q \left( l_n^{q-1} |\zeta'(v)_{i+s}^n|^q + l_n^{q/2-1} |\zeta''(v)_{i+s}^n|^q \right) \right),$$

which is bounded by  $\leq K \Delta_n^{q/2}$ . The same steps apply to the second part of  $A(3; u, v)_i^n$ . Hence,

$$\mathbb{E}_{i-1} \left| \frac{1}{k_n \Delta_n} A(3; u, v)_i^n \right|^q \leq K (k_n^{-2} \Delta_n^{-1})^{q/2}.$$

(g) First, consider  $v = 1$ . By (A.6) and (A.25), we have

$$\begin{aligned} \left| \frac{1}{k_n \Delta_n} \mathbb{E}_{i-1} A(4; u, 1)_i^n \right| &\leq \frac{1}{k_n \Delta_n} \sum_{s=2l_n-1}^{k_n+2l_n-2} \mathbb{E}_{i-1} [|\rho(u, 1)_{i+s}| \mathbb{E}_{i-1} |\zeta'(1)_{i+s}^n|] \\ &\leq \frac{K}{k_n \Delta_n} k_n l_n^{-3/2} \Delta_n^{1/2} \left( \Delta_n^{1/2} + \mathbb{E}_{i-1} \eta_{i+s}^n \right) \leq K \Delta_n^{3/4} + K \Delta_n^{1/4} \eta_{i, 2k_n}^n. \end{aligned}$$

Moreover, by Hölder inequality, (A.6), and (A.25) again, we have

$$\mathbb{E}_{i-1} |A(4; u, 1)_i^n|^q \leq k_n^{q-1} \sum_{s=2l_n-1}^{k_n+2l_n-2} \mathbb{E}_{i-1} [|\rho(u, 1)_{i+s}|^q \mathbb{E}_{i-1} |\zeta'(1)_{i+s}^n|^q] \leq K k_n^q \Delta_n^{5q/4}.$$

Next, consider  $v = 2, 3$ . Introduce the notation  $V(2) = b^{(c)}$  and  $V(3) = b$ . First notice that

$$\begin{aligned} |\mathbb{E}_{i-1} A(4; u, v)_i^n| &\leq \left| \mathbb{E}_{i-1} \sum_{s=2l_n-1}^{k_n+2l_n-2} \rho(u, v)_{i+s} \left( \zeta'(v)_{i+s}^n - V(v)_{(i+s-2l_n)\Delta_n} \Delta_n \right) \right| \\ &\quad + \left| \mathbb{E}_{i-1} \sum_{s=2l_n-1}^{k_n+2l_n-2} \rho(u, v)_{i+s} V(v)_{(i+s-2l_n)\Delta_n} \Delta_n \right|. \end{aligned}$$

On the one hand, we have

$$\left| \mathbb{E}_{i-1} \sum_{s=2l_n-1}^{k_n+2l_n-2} \rho(u, v)_{i+s} \left( \zeta'(v)_{i+s}^n - V(v)_{(i+s-2l_n)\Delta_n} \Delta_n \right) \right|$$

$$\leq \sum_{s=2l_n-1}^{k_n+2l_n-2} \sqrt{\mathbf{E}_{i-1} |\rho(u, v)_{i+s}|^2} \sqrt{\mathbf{E}_{i-1} \left| \zeta'(v)_{i+s}^n - V(v)_{(i+s-2l_n)\Delta_n} \Delta_n \right|^2} \leq K k_n \Delta_n^{5/4} \eta_{i,2k_n}^n,$$

since by (A.8), (A.22), and (A.23),

$$\begin{aligned} & \mathbf{E}_{i-1} \left( \left| \zeta'(v)_{i+s}^n - V(v)_{(i+s-2l_n)\Delta_n} \Delta_n \right|^2 \right) \\ & \leq K \mathbf{E}_{i-1} \left( \left| \zeta'(v)_{i+s}^n - V(v)_{(i+s-1)\Delta_n} \Delta_n \right|^2 \right) + K \Delta_n^2 \mathbf{E}_{i-1} \left( \left| V(v)_{(i+s-1)\Delta_n} - V(v)_{(i+s-2l_n)\Delta_n} \right|^2 \right) \\ & \leq \mathbf{E}_{i-1} \left[ \Delta_n \left( \sqrt{\Delta_n} + \eta_{i+s}^n \right) \right]^2 + K \Delta_n^2 (\eta_{i,2k_n}^n)^2 \leq K \Delta_n^2 (\eta_{i,2k_n}^n)^2. \end{aligned}$$

On the other hand, by Cauchy-Schwartz inequality, (A.7), and (A.9),

$$\begin{aligned} & \left| \mathbf{E}_{i-1} \sum_{s=2l_n-1}^{k_n+2l_n-2} \rho(u, v)_{i+s} V(v)_{(i+s-2l_n)\Delta_n} \Delta_n \right| \\ & \leq \Delta_n \sum_{s=2l_n-1}^{k_n+2l_n-2} \sum_{m=1}^{2l_n-1} \sup_s |V(v)_s| \sqrt{\mathbf{E}_{i-1} |\Gamma(u, v)_m^n|^2} \sqrt{\mathbf{E}_{i-1} |\zeta'(u)_{i+s-m}^n|^2} \leq K k_n \Delta_n^{3/2}. \end{aligned}$$

Therefore, we obtain

$$\frac{1}{k_n \Delta_n} |\mathbf{E}_{i-1} A(4; u, v)_i^n| \leq \frac{K}{k_n \Delta_n} \left( k_n \Delta_n^{5/4} \eta_{i,2k_n}^n + k_n \Delta_n^{3/2} \right) = K \Delta_n^{1/4} \eta_{i,2k_n}^n + K \Delta_n^{1/2}.$$

To prove (A.4)-(A.5), we have, by Hölder inequality, (A.6), and (A.25),

$$\mathbf{E}_{i-1} |A(4; u, v)_i^n|^q \leq k_n^{q-1} \sum_{s=2l_n-1}^{k_n+2l_n-2} \mathbf{E}_{i-1} \left[ |\rho(u, v)_{i+s}|^q \mathbf{E}_{i+s-1} |\zeta'(v)_{i+s}^n|^q \right] \leq K k_n^q \Delta_n^{5q/4}.$$

We conclude that for  $v = 1, 2$ , and  $3$ ,  $\mathbf{E}_{i-1} \left| \frac{1}{k_n \Delta_n} A(4; u, v)_i^n \right|^q \leq K \Delta_n^{q/4}$ .

(h) The first statement, (A.3), is trivial for any pair of  $u$  and  $v$ :

$$\mathbf{E}_{i-1} Z(u, v)_i^n = \mathbf{E}_{i-1} \sum_{j=2l_n-1}^{k_n+2l_n-2} \sum_{m=1}^{2l_n-1} \Gamma(u, v)_m^n \zeta(u)_{i+j-m}^n \zeta''(v)_{i+j}^n = 0.$$

To prove (A.4)-(A.5), we start with the term  $Z(u, 3)_i^n$  for  $u = 1, 2$ . We have

$$\begin{aligned} \mathbf{E}_{i-1} |Z(u, 3)_i^n|^q & \leq K \mathbf{E}_{i-1} \left| \sum_{j=2l_n-1}^{k_n+2l_n-2} \left( \rho(u, 3)_{i+j} \zeta''(3)_{i+j}^n \right) \right|^{q/2} \\ & \leq K k_n^{\frac{q}{2}-1} \sum_{j=2l_n-1}^{k_n+2l_n-2} \mathbf{E}_{i-1} \left[ \left( \rho(u, 3)_{i+j} \right)^q \mathbf{E}_{i+j-1} \left( \left( \zeta(3)_{i+j}^n \right)^q + \left( \zeta'(3)_{i+j}^n \right)^q \right) \right] \\ & \leq K k_n^{\frac{q}{2}} l_n^{-q/2} \Delta_n^{q/2} + K k_n^{\frac{q}{2}} l_n^{-q/2} \Delta_n^q. \end{aligned}$$

The first inequality follows by the Burkholder-Gundy inequality, second by Hölder inequality, and last by (A.8), (A.9), and (A.25). Therefore,  $\mathbf{E}_{i-1} \left| \frac{1}{k_n \Delta_n} Z(u, 3)_i^n \right|^q \leq K \left( k_n^{-1} \Delta_n^{-1/2} \right)^{q/2}$ . Next term

is  $Z(3, u)_i^n$  for  $u = 1, 2$ . Similar to the above,  $\mathbb{E}_{i-1} |Z(3, u)_i^n|^q$  is bounded by

$$K k_n^{\frac{q}{2}-1} \sum_{j=2l_n-1}^{k_n+2l_n-2} \mathbb{E}_{i-1} \left[ \left( \rho(3, u)_{i+j} \right)^q \mathbb{E}_{i+j-1} \left( \left( \zeta(u)_{i+j}^n \right)^q + \left( \zeta'(u)_{i+j}^n \right)^q \right) \right],$$

which is bounded by  $K \left( k_n \Delta_n^{3/2} \right)^{q/2}$ . Therefore,  $\mathbb{E}_{i-1} \left| \frac{1}{k_n \Delta_n} Z(3, u)_i^n \right|^q \leq K \left( k_n^{-1} \Delta_n^{-1/2} \right)^{q/2}$ .

(i) We only need to prove (A.4)-(A.5) regarding  $Z(u, v)_i^n$  when  $(u, v) = (1, 2)$ . We have

$$\begin{aligned} \mathbb{E}_{i-1} |Z(1, 2)_i^n|^q &\leq K k_n^{\frac{q}{2}-1} \sum_{j=2l_n-1}^{k_n+2l_n-2} \mathbb{E}_{i-1} \left[ \left( \rho(1, 2)_{i+j} \right)^q \mathbb{E}_{i+j-1} \left( \zeta(2)_{i+j}^n \right)^q \right] \\ &\quad + K k_n^{\frac{q}{2}-1} \sum_{j=2l_n-1}^{k_n+2l_n-2} \mathbb{E}_{i-1} \left[ \left( \rho(1, 2)_{i+j} \right)^q \mathbb{E}_{i+j-1} \left( \zeta'(2)_{i+j}^n \right)^q \right] \\ &\leq K k_n^{\frac{q}{2}} l_n^{-q/2} \Delta_n^{q/2} + K k_n^{\frac{q}{2}} l_n^{-q/2} \Delta_n^q \leq K \left( k_n \Delta_n^{3/2} \right)^{q/2}. \end{aligned}$$

In above the first inequality follows by Hölder inequality, the second by (A.8), (A.9), and (A.25). Hence, we obtain  $\mathbb{E}_{i-1} \left| \frac{1}{k_n \Delta_n} Z(1, 2)_i^n \right|^q \leq K \left( k_n^{-1} \Delta_n^{-1/2} \right)^{q/2}$ . Similar arguments give the result for  $Z(u, v)_i^n$  when  $(u, v) = (2, 1), (1, 1)$ , and  $(2, 2)$ .

(j) By (A.16) and (A.17), we have

$$\begin{aligned} \mathbb{E}_{i-1} \left| \bar{\beta}_i^{lm} \right|^q &= \mathbb{E}_{i-1} \left| \frac{1}{k_n} \sum_{s=0}^{k_n-1} C_{(i+s-1)\Delta_n}^{lm} - C_{(i-1)\Delta_n}^{lm} \right|^q \leq K (k_n \Delta_n)^{q/2} + K k_n \Delta_n, \\ \left| \mathbb{E}_{i-1} \left( \bar{\beta}_i^{lm} \right) \right| &\leq K k_n \Delta_n. \end{aligned}$$

## E Proof of Lemma A3

As in Lemma A2, we can decompose each of  $\beta_i^{jl}$  and  $\beta_i^{km}$  for all  $(j, l)$  and  $(k, m)$ , into 10 parts, respectively, denoted as (a)-(j) in Lemma A2. An inspection of the proof of Lemma A2, for  $q = 2$ , shows that the cross-products of almost all terms are bounded by  $K k_n \Delta_n$ . There are three exceptions. First, the cross-products between terms in (f) and (h) – (i) (i.e., the terms in (E.36) below). Second, the cross-products between terms in (j) and (h) – (i) (i.e., the terms in (E.37) below). Third, product of terms in (h) – (i), consisting of  $Z(u, v)_i^n Z(u', v')_i^n$  for all  $u, v, u', v' = 1, 2, 3$ . These latter terms exactly correspond to the contributions of  $\eta_i^{jl}$  towards  $\beta_i^{jl}$  for each  $(j, l)$ . These are the terms that will absorb the centering above  $k_n^{-1} \Delta_n^{-1/2} h(j, k; l, m; C_{(i-1)\Delta_n})$ . Therefore, by the proof of Lemma A2 and the Cauchy-Schwartz inequality, it is sufficient to prove the following three results, for all



$(u, v, u', v')$  and  $(j, l)$ ,

$$\left| \mathbf{E}_{i-1} \left( \eta_i^{jl} \eta_i^{km} - k_n^{-1} \Delta_n^{-1/2} h(j, k; l, m; C_{(i-1)\Delta_n}) \right) \right| \leq K k_n \Delta_n, \quad (\text{E.35})$$

$$\left| \mathbf{E}_{i-1} \left( \frac{1}{k_n \Delta_n} A(3; u, v)_i^n \frac{1}{k_n \Delta_n} Z(u', v')_i^n \right) \right| \leq K k_n \Delta_n, \quad (\text{E.36})$$

$$\left| \mathbf{E}_{i-1} \left( \bar{\beta}_i^{jl} \frac{1}{k_n \Delta_n} Z(u', v')_i^n \right) \right| \leq K k_n \Delta_n. \quad (\text{E.37})$$

We start by showing (E.35). By the definition of  $\eta_i^{jl}$ , we know that

$$\eta_i^{jk} \eta_i^{lm} = \begin{cases} \sum_{(u,v) \in A} \sum_{(u',v') \in A} \frac{1}{k_n^2 \Delta_n^2} Z(u, v)_i^n Z(u', v')_i^n & \text{if } (j, k, l, m) = (1, 2, 1, 2), (1, 2, 2, 1), \\ & (2, 1, 1, 2), (2, 1, 2, 1) \\ \sum_{(u,v) \in A} \sum_{(u',v') \in B} \frac{1}{k_n^2 \Delta_n^2} Z(u, v)_i^n Z(u', v')_i^n & \text{if } (j, k, l, m) = (1, 2, 2, 2), (2, 1, 2, 2), \\ & (2, 2, 1, 2), (2, 2, 2, 1) \\ \sum_{(u,v) \in B} \sum_{(u',v') \in B} \frac{1}{k_n^2 \Delta_n^2} Z(u, v)_i^n Z(u', v')_i^n & \text{if } (j, k, l, m) = (2, 2, 2, 2). \\ 0 & \text{otherwise} \end{cases},$$

where  $A = \{(1, 3), (2, 3), (3, 1), (3, 2)\}$ ,  $B = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$ . Therefore, by comparing it with the definition of  $h$  in equation (A.1), to prove that

$$\left| \mathbf{E}_{i-1} \left( \eta_i^{jl} \eta_i^{km} - k_n^{-1} \Delta_n^{-1/2} h(j, k; l, m; C_{(i-1)\Delta_n}) \right) \right| \leq K k_n \Delta_n$$

for every  $(j, l, k, m)$ , it is sufficient to prove that

$$\left| \mathbf{E}_{i-1} \left[ G_i^n - k_n^{-1} \Delta_n^{-1/2} l(u, v; u', v'; C_{(i-1)\Delta_n}) \right] \right| \leq K k_n \Delta_n$$

for every  $(u, v, u', v')$ , where

$$G_i^n = \left( \frac{1}{k_n \Delta_n} Z(u, v)_i^n \right) \left( \frac{1}{k_n \Delta_n} Z(u', v')_i^n \right).$$

We prove this below. We have

$$\mathbf{E}_{i-1} \left[ G_i^n - k_n^{-1} \Delta_n^{-1/2} l(u, v; u', v'; C_{(i-1)\Delta_n}) \right] = \mathbf{E}_{i-1} \left[ G_n^{(1)} - k_n^{-1} \Delta_n^{-1/2} l(u, v; u', v'; C_{(i-1)\Delta_n}) \right],$$

where

$$G_n^{(1)} = \left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1} \left( \rho(u, v)_{i+j} \zeta''(v)_{i+j}^n \right) \left( \rho(u', v')_{i+j} \zeta''(v')_{i+j}^n \right),$$

which follows because  $\mathbf{E}_{i+j-1} \zeta''(v')_{i+j}^n = 0$ . Next, define

$$\bar{G}_n = \left[ \left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1} \rho(u, v)_{i+j} \rho(u', v')_{i+j} \Delta_n z_{v, v'}^n V(v, v')_{i+j} - k_n^{-1} \Delta_n^{-1/2} l(u, v; u', v'; C_{(i-1)\Delta_n}) \right].$$

It suffices to show the bound with  $\overline{G}_n$  instead of  $G_n^{(1)}$  since  $\left| \mathbf{E}_{i-1} \left( G_n^{(1)} - \overline{G}_n \right) \right|$  is bounded by

$$\left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1}^{2k_n + 2l_n - 2} \left| \mathbf{E}_{i-1} \rho(u, v)_{i+j} \rho(u', v')_{i+j} \mathbf{E}_{i+j-1} \left[ \zeta''(v)_{i+j}^n \zeta''(v')_{i+j}^n - \Delta_n z_{v, v'}^n V(v, v')_{i+j} \right] \right|,$$

which is bounded by  $Kk_n^{-1}$  by (A.10)–(A.15), and (A.25) Next, define

$$\overline{G}_{t,i}^n = \left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1}^{2k_n + 2l_n - 2} \rho(u, v)_{i+j} \rho(u', v')_{i+j} \Delta_n z_{v, v'}^n V(v, v')_{i+j-2l_n} - k_n^{-1} \Delta_n^{-1/2} l(u, v; u', v'; C_{(i-1)\Delta_n}).$$

It suffices to show the bound with  $\overline{G}_{t,i}^n$  instead of  $\overline{G}_n$  since  $\left| \mathbf{E}_{i-1} \left( \overline{G}_n - \overline{G}_{t,i}^n \right) \right|$  is bounded by

$$\Delta_n z_{v, v'}^n \left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1}^{2k_n + 2l_n - 2} \sqrt{\mathbf{E}_{i-1} \left( \rho(u, v)_{i+j} \rho(u', v')_{i+j} \right)^2} \sqrt{\mathbf{E}_{i-1} \left( V(v, v')_{i+j} - V(v, v')_{i+j-2l_n} \right)^2},$$

by the Cauchy-Schwartz inequality, which is in turn bounded by  $Kk_n^{-1}$  using (A.25) and the standard estimates of Itô semimartingales applied to each value of the pair  $(v, v')$ . Next, decompose  $\overline{G}_{t,i}^n = \hat{G}(1)_i + \hat{G}(2)_i + \hat{G}(3)_i$  where

$$\begin{aligned} \hat{G}(1)_i &= \left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1}^{2k_n + 2l_n - 2} \sum_{m=1}^{2l_n-1} \Gamma(u, v)_m^n \Gamma(u', v')_m^n \zeta(u)_{i+j-m}^n \zeta(u')_{i+j-m}^n \Delta_n z_{v, v'}^n V_{i+j-2l_n}, \\ \hat{G}(2)_i &= \left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1}^{2k_n + 2l_n - 2} \sum_{m=1}^{2l_n-1} \Gamma(u, v)_m^n \zeta(u)_{i+j-m}^n \sum_{m' > m} \Gamma(u', v')_{m'}^n \zeta(u')_{i+j-m'}^n \Delta_n z_{v, v'}^n V_{i+j-2l_n}, \\ \hat{G}(3)_i &= \left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1}^{2k_n + 2l_n - 2} \sum_{m=1}^{2l_n-1} \Gamma(u, v)_m^n \zeta(u)_{i+j-m}^n \sum_{m' < m} \Gamma(u', v')_{m'}^n \zeta(u')_{i+j-m'}^n \Delta_n z_{v, v'}^n V(v, v')_{i+j-2l_n}. \end{aligned}$$

First, define  $\xi_{i+j} = \sum_{m=1}^{2l_n-1} \Gamma(u, v)_m^n V(v, v')_{i+j-2l_n} \sum_{m' > m} \Gamma(u', v')_{m'}^n \zeta(u')_{i+j-m'}^n$ , so that

$$\mathbf{E}_{i-1} \hat{G}(2)_i = \Delta_n z_{v, v'}^n \left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1}^{2k_n + 2l_n - 2} \mathbf{E}_{i-1} \xi_{i+j} \mathbf{E}_{i+j-m-1} \zeta(u)_{i+j-m}^n.$$

We have

$$\begin{aligned} \mathbf{E}_{i-1} \left( |\xi_{i+j}| \right) &\leq \sum_{m=1}^{2l_n-1} \mathbf{E}_{i-1} \left| \Gamma(u, v)_m^n V(v, v')_{i+j-2l_n} \right| \left| \sum_{m' > m} \Gamma(u', v')_{m'}^n \zeta(u')_{i+j-m'}^n \right| \\ &\leq \begin{cases} Kl_n^{-5/2} & \text{if } (u, v) = (1, 1) \text{ and } v' = 1 \\ Kl_n^{-3/2} & \text{if } (u, v) = (1, 1) \text{ and } v' = 2, 3 \\ Kl_n^{-3/2} & \text{if } (u, v) = (1, 2), (2, 1), (1, 3), (3, 1) \text{ and } v' = 1 \\ Kl_n^{-1/2} & \text{if } (u, v) = (1, 2), (2, 1), (1, 3), (3, 1) \text{ and } v' = 2, 3 \\ Kl_n^{-1/2} & \text{if } (u, v) = (2, 3), (3, 2), (2, 2) \text{ and } v' = 1 \\ Kl_n^{1/2} & \text{if } (u, v) = (2, 3), (3, 2), (2, 2) \text{ and } v' = 2, 3 \end{cases}, \end{aligned}$$

by applying (A.7) and (A.9) to each possibility of  $(u, v)$  and  $v'$ . Together with (A.6), (A.8), and (A.24), we have  $|\mathbf{E}_{i-1} \hat{G}(2)_i| \leq K \Delta_n^{-1/4} k_n^{-1}$ . The bound for  $|\mathbf{E}_{i-1} \hat{G}(3)_i|$  is the same by similarity of  $\hat{G}(2)_i$  to  $\hat{G}(3)_i$ . Therefore, to show the bound for  $\overline{G}_{t,i}^n$ , it is sufficient to show the bound for  $\hat{G}(1)_i$ .

We decompose it as  $\hat{G}(1)_i = \hat{G}'' + \hat{G}'$ , where

$$\begin{aligned} \hat{G}'' &= \left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1}^{2k_n+2l_n-2} \sum_{m=1}^{2l_n-1} \Gamma(u, v)_m^n \Gamma(u', v')_m^n \\ &\quad \times [\zeta(u)_{i+j-m}^n \zeta(u')_{i+j-m}^n - V(u, u')_{i+j-m} \Delta_n z_{u,u'}^n] \Delta_n z_{v,v'}^n V(v, v')_{i+j-2l_n}, \\ \hat{G}' &= \left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1}^{2k_n+2l_n-2} \sum_{m=1}^{2l_n-1} \Gamma(u, v)_m^n \Gamma(u', v')_m^n \times [V(u, u')_{i+j-m} \Delta_n z_{u,u'}^n] \Delta_n z_{v,v'}^n V(v, v')_{i+j-2l_n}. \end{aligned}$$

We have

$$\begin{aligned} |\mathbf{E}_{i-1} \hat{G}''| &\leq 4 \left( \frac{1}{k_n \Delta_n} \right)^{2k_n + 2l_n - 2} \sum_{j=2l_n-1}^{2k_n+2l_n-2} \sum_{m=1}^{2l_n-1} \Gamma(u, v)_m^n \Gamma(u', v')_m^n \Delta_n z_{v,v'}^n V(v, v')_{i+j-2l_n} \\ &\quad \times |\mathbf{E}_{i-1} [\zeta(u)_{i+j-m}^n \zeta(u')_{i+j-m}^n - V(u, u')_{i+j-m} \Delta_n z_{u,u'}^n]| \leq K k_n^{-1}, \end{aligned}$$

by applying (A.10)–(A.15) to each value of  $(u, v, u', v')$ . It remains to show the bound for  $|\mathbf{E}_{i-1} \hat{G}'_i|$ .

Define  $\hat{G}_i = \left( \frac{1}{k_n \Delta_n} \right)^2 \sum_{j=2l_n-1}^{k_n+2l_n-2} H(u, v; u', v')_n \Delta_n^2 V(u, u')_{i-1} V(v, v')_{i-1}$ . It is sufficient to show  $|\mathbf{E}_{i-1} \hat{G}_i - k_n^{-1} \Delta_n^{-1/2} l(u, v; u', v'; C_{(i-1)\Delta_n})| \leq K k_n \Delta_n$  because

$$\begin{aligned} |\mathbf{E}_{i-1} (\hat{G}_i - \hat{G}'_i)| &= \left| \mathbf{E}_{i-1} \left[ \left( \frac{1}{k_n \Delta_n} \right)^2 \sum_{j=2l_n-1}^{k_n+2l_n-2} H(u, v; u', v')_n \Delta_n^2 [V(u, u')_{i+j-m} V(v, v')_{i+j-2l_n} \right. \right. \\ &\quad \left. \left. - V(u, u')_{i-1} V_{i-1}(v, v') \right] \right| \leq K k_n^{-1} \Delta_n^{-1/4}, \end{aligned}$$

using the standard estimates because  $V(v, v')_s V(u, u')_s$  is an Itô semimartingale. Before continuing, note that  $l(u, v; u', v'; C_{(i-1)\Delta_n})$  satisfies

$$l(u, v; u', v'; C_s) = V_{i-1}(v, v') V(u, u')_{i-1} \lim_{n \rightarrow \infty} \left[ H(u, v; u', v'; C_s)_n \Delta_n^{1/2} \right],$$

and we can easily verify

$$\left| \mathbf{E}_{i-1} \left[ H(u, v; u', v')_n \Delta_n^{1/2} - \lim_{n \rightarrow \infty} (H(u, v; u', v')_n \Delta_n^{1/2}) \right] \right| \leq K \Delta_n^{1/2}.$$

Therefore,

$$\begin{aligned} &\left| \mathbf{E}_{i-1} \hat{G}_i - k_n^{-1} \Delta_n^{-1/2} l(u, v; u', v'; C_{(i-1)\Delta_n}) \right| \\ &= \left| k_n^{-1} \Delta_n^{-1/2} H(u, v; u', v')_n \Delta_n^{1/2} V(u, u')_{i-1} V(v, v')_{i-1} - k_n^{-1} \Delta_n^{-1/2} l(u, v; u', v'; C_{(i-1)\Delta_n}) \right| \\ &\quad + K k_n^{-2} \Delta_n^{-1/2} l_n \leq K \Delta_n^{1/2} + K k_n^{-2} \Delta_n^{-1/2} l_n \leq K k_n \Delta_n. \end{aligned}$$

This completes the proof of (E.35).

We now prove (E.36). Note that the expectation of the cross-term between  $Z(u, v)_i^n$  and the first part of  $A(3; u, v)_i^n$  in (D.33) is zero, by the law of iterated expectations. Therefore, to prove (E.36), we need to obtain a bound for the cross term between  $Z(u, v)_i^n$  and the second part of  $A(3; u, v)_i^n$  in (D.33), which is

$$\begin{aligned}
& \mathbb{E}_{i-1} \left( Z(u, v)_i^n \sum_{s=k_n}^{k_n+2l_n-2} \eta(i, s) \zeta(v)_{i+s}^n \right) \\
&= \mathbb{E}_{i-1} \left( \sum_{j=2l_n-1}^{k_n+2l_n-2} \rho(u, v)_{i+j} \zeta''(v)_{i+j}^n \sum_{s=k_n}^{k_n+2l_n-2} (-\eta(i+k_n, s-k_n)) \zeta(v)_{i+s}^n \right) \\
&= \mathbb{E}_{i-1} \left( \sum_{s=k_n}^{k_n+2l_n-2} \sum_{j=2l_n-1}^{s-1} \xi(s, j)_i^n \right) + \mathbb{E}_{i-1} \left( \sum_{s=k_n}^{k_n+2l_n-2} \xi(s, s)_i^n \right), \tag{E.38}
\end{aligned}$$

where first equality follows by the fact that for any  $s = k_n, \dots, k_n + 2l_n - 2$ ,  $\eta(i+k_n, s-k_n) + \eta(i, s) = 0$ , and where  $\xi(s, j)_i^n = -\rho(u, v)_{i+j} \eta(i+k_n, s-k_n) \zeta''(v)_{i+j}^n \zeta(v)_{i+s}^n$ , which clearly satisfies  $\mathbb{E}_{i-1} \xi(s, j)_i^n = 0$  whenever  $j > s$ . The first part in (E.38) satisfies

$$\begin{aligned}
& \left| \mathbb{E}_{i-1} \left( \sum_{s=k_n}^{k_n+2l_n-2} \sum_{j=2l_n-1}^{s-1} \xi(s, j)_i^n \right) \right| \\
&\leq K \sum_{s=k_n}^{k_n+2l_n-2} \sqrt{\sum_{j=2l_n-1}^{s-1} \mathbb{E}_{i-1} \left( \left| \rho(u, v)_{i+j} \right|^2 \left| \zeta''(v)_{i+j}^n \right|^2 \right)} \sqrt{\mathbb{E}_{i-1} \left( \left| \eta(i+k_n, s-k_n) \right|^2 \left| \zeta'(v)_{i+s}^n \right|^2 \right)} \\
&\leq \begin{cases} Kl_n \sqrt{k_n l_n}^{-3/2} \times 1 \times l_n^{-3/2} \Delta_n^{1/2} & \text{if } v = 1 \\ Kl_n \sqrt{k_n l_n}^{-1/2} \times \Delta_n^{1/2} \times l_n^{-1/2} \Delta_n & \text{if } v = 2, 3 \end{cases} \leq K \sqrt{k_n} \Delta_n^{3/2}. \tag{E.39}
\end{aligned}$$

The first inequality follows by the law of iterated expectations and Burkholder-Gundy inequality, second by (A.6)-(A.9), (D.34), and (A.25). We next show that the bound for the second part in (E.38) satisfies

$$\left| \mathbb{E}_{i-1} \left( \sum_{s=k_n}^{k_n+2l_n-2} \xi(s, s)_i^n \right) \right| \leq K \Delta_n. \tag{E.40}$$

We have

$$\begin{aligned}
& \left| \mathbb{E}_{i-1} \left( \sum_{s=k_n}^{k_n+2l_n-2} \xi(s, s)_i^n \right) \right| \tag{E.41} \\
&\leq \left| \mathbb{E}_{i-1} \sum_{s=k_n}^{k_n+2l_n-2} \rho(u, v)_{i+s} \eta(i+k_n, s-k_n) (\zeta''(v)_{i+s}^n \zeta''(v)_{i+s}^n - z_{v,v}^n \Delta_n V(v, v)_{i+s}^n) \right|
\end{aligned}$$

$$\begin{aligned}
& + \left| \mathbf{E}_{i-1} \sum_{s=k_n}^{k_n+2l_n-2} \rho(u, v)_{i+s} \eta(i+k_n, s-k_n) z_{v,v}^n \Delta_n (V(v, v)_{i+s}^n - V(v, v)_{i+s-2l_n}^n) \right| \\
& + \left| \mathbf{E}_{i-1} \sum_{s=k_n}^{k_n+2l_n-2} \rho(u, v)_{i+s} \eta(i+k_n, s-k_n) z_{v,v}^n \Delta_n V(v, v)_{i+s-2l_n}^n \right|.
\end{aligned}$$

The first term in (E.41) is smaller or equal than

$$\sum_{s=k_n}^{k_n+2l_n-2} \mathbf{E}_{i-1} (|\rho(u, v)_{i+s}| |\eta(i+k_n, s-k_n)| |\mathbf{E}_{i+s-1} (\zeta''(v)_{i+s}^n \zeta''(v)_{i+s}^n - z_{v,v}^n \Delta_n V(v, v)_{i+s}^n)|),$$

which is bounded by  $K \Delta_n^{3/2}$ . The second term in (E.41) is smaller or equal than

$$\begin{aligned}
& z_{v,v}^n \Delta_n \sum_{s=k_n}^{k_n+2l_n-2} \left( \mathbf{E}_{i-1} |\rho(u, v)_{i+s}|^4 \right)^{\frac{1}{4}} \left( \mathbf{E}_{i-1} |\eta(i+k_n, s-k_n)|^4 \right)^{\frac{1}{4}} \left( \mathbf{E}_{i-1} |V(v, v)_{i+s}^n - V(v, v)_{i+s-2l_n}^n|^2 \right)^{\frac{1}{2}} \\
& \leq \begin{cases} K \Delta_n k_n \left( l_n^{-1/2} \right)^2 \sqrt{l_n \Delta_n} & \text{if } v = 2, 3 \\ K k_n \left( l_n^{-3/2} \right)^2 \sqrt{l_n \Delta_n} & \text{if } v = 1 \end{cases} \leq K \Delta_n^{7/4} k_n.
\end{aligned}$$

For the third term in (E.41), we first obtain the bound, for  $s = k_n, \dots, k_n + 2l_n - 2$

$$\begin{aligned}
& |\mathbf{E}_{i+s-2l_n} (\rho(u, v)_{i+s} \eta(i+k_n, s-k_n))| \\
& = \left| \mathbf{E}_{i+s-2l_n} \left( \sum_{q=1}^{2l_n-1} \Gamma(u, v)_q^n \zeta(u)_{i+s-q}^n \left[ \sum_{m=1}^{(2l_n-1) \wedge (s-k_n)} \gamma(u, v; m)_{s-2k_n+1, s-k_n+1}^n \zeta(u)_{i+s-m}^n \right] \right) \right| \\
& \leq \begin{cases} K \sqrt{\Delta_n} & \text{if } v = 2, 3 \\ K \Delta_n^{3/2} & \text{if } v = 1 \end{cases}.
\end{aligned}$$

The last equality follows by considering separately the terms with  $m = q$ ,  $m > q$ , and  $m < q$ . For the  $m = q$  terms, one uses (A.10)–(A.15). For the  $m > q$  terms, one first uses the law of iterated expectations to replace  $\zeta(u)_{i+s-m}^n$  by  $\zeta'(u)_{i+s-m}^n$ ; the final bound follows by (A.6)–(A.9). Similar steps apply to the  $m < q$  terms. Therefore, the third term in (E.41) satisfies

$$\begin{aligned}
& \left| \mathbf{E}_{i-1} \sum_{s=k_n}^{k_n+2l_n-2} \rho(u, v)_{i+s} \eta(i+k_n, s-k_n) z_{v,v}^n \Delta_n V(v, v)_{i+s-2l_n}^n \right| \\
& \leq z_{v,v}^n \Delta_n \sum_{s=k_n}^{k_n+2l_n-2} \sup_s |V_s| \mathbf{E}_{i-1} |\mathbf{E}_{i+s-2l_n} \rho(u, v)_{i+s} \eta(i+k_n, s-k_n)|,
\end{aligned}$$

which is bounded by  $\leq K \Delta_n$ . This concludes the proof of (E.40), and together with (E.39), we find

$$\left| \mathbf{E}_{i-1} \left( \frac{1}{k_n \Delta_n} A(3; u, v)_i^n \frac{1}{k_n \Delta_n} Z(u', v')_i^n \right) \right| \leq K k_n^{-2} \Delta_n^{-1} \leq K k_n \Delta_n.$$

This concludes the proof of (E.36). We now prove (E.37). For any  $(j, l)$ ,

$$\begin{aligned}
& \left| \mathbf{E}_{i-1} \left( \bar{\beta}_i^{jl} Z(u, v)_i^n \right) \right| \\
&= \left| \mathbf{E}_{i-1} \left( \frac{1}{k_n} \sum_{s=0}^{k_n-1} \sum_{q=s+2l_n-1}^{k_n-1} (k_n - q - 1) \rho(u, v)_{i+s+2l_n-1} \zeta''(v)_{i+s+2l_n-1}^n \mathbf{E}_{i+q-1} \Delta_{i+q}^n C^{jl} \right) \right| \\
&\leq \frac{1}{k_n} \sum_{q=2l_n-1}^{k_n-1} \mathbf{E}_{i-1} \left| \sum_{q=s+2l_n-1}^{k_n-1} \rho(u, v)_{i+s+2l_n-1} \zeta''(v)_{i+s+2l_n-1}^n \right| \left| (k_n - q - 1) \mathbf{E}_{i+q-1} \Delta_{i+q}^n C^{jl} \right| \\
&\leq \Delta_n \sum_{q=2l_n-1}^{k_n-1} \left| \mathbf{E}_{i-1} \sum_{q=s+2l_n-1}^{k_n-1} |\rho(u, v)_{i+s+2l_n-1}|^2 |\zeta''(v)_{i+s+2l_n-1}^n|^2 \right|^{1/2} \\
&\leq \begin{cases} K \Delta_n k_n (k_n \times l_n^{-3} \times 1)^{1/2} & \text{if } v = 1 \\ K \Delta_n k_n (k_n \times l_n^{-1} \times \Delta_n)^{1/2} & \text{if } v = 2, 3 \end{cases} \leq K \Delta_n^{7/4} k_n^{3/2}.
\end{aligned}$$

Second transition follows by the fact that summands with  $i + s + 2l_n - 1 > i + s + 2l_n - 1$  are zero by the law of iterated expectations. Third transition follows by the Burkholder-Gundy inequality and standard estimates for Itô Semimartingales. Fourth transition follows by (A.25), (A.7), and (A.9). Therefore, for any  $(u, v)$  and any  $(j, l)$ ,

$$\left| \mathbf{E}_{i-1} \left( \bar{\beta}_i^{jl} \frac{1}{k_n \Delta_n} Z(u, v)_i^n \right) \right| \leq K \Delta_n^{3/4} k_n^{1/2} \leq K k_n \Delta_n.$$

This concludes the proof of (E.36) and hence the lemma.