Cross-sectional Dependence in Idiosyncratic Volatility*

Ilze Kalnina† | Kokouvi Tewou‡
University College London | Université de Montréal

November 15, 2016

Abstract

Idiosyncratic volatility is the volatility of asset returns once the impact of common factors has been removed. The empirical evidence suggests the idiosyncratic volatilities are cross-sectionally correlated. This paper introduces an econometric framework for analysis of cross-sectional dependence in the idiosyncratic volatilities of assets using high frequency data. We first consider the estimation of standard measures of dependence in the idiosyncratic volatilities such as covariances and correlations. Next, we study an idiosyncratic volatility factor model, in which we decompose the variation in idiosyncratic volatilities into two parts: the variation related to the common factors such as the market volatility, and the residual variation. When using high frequency data, naive estimators of all of the above measures are biased due to the use of error-laden estimates of idiosyncratic volatilities. We provide bias-corrected estimators and establish their asymptotic properties. We apply our methodology to the 30 Dow Jones Industrial Average components, and document strong cross-sectional dependence in their idiosyncratic volatilities. We consider two different sets of idiosyncratic volatility factors, and find that neither can fully account for the cross-sectional dependence in idiosyncratic volatilities. We map out the network of dependencies in residual idiosyncratic volatilities across the stocks.

Keywords: network of risk; systematic risk; idiosyncratic risk; risk management; high frequency data.

JEL Codes: C58, C22, C14, G11.

*We benefited from discussions with Marine Carrasco, Yoosoon Chang, Valentina Corradi, Russell Davidson, Jean-Marie Dufour, Prosper Dovonon, Kirill Evdokimov, Sílvia Gonçalves, Jean Jacod, Dennis Kristensen, Joon Park, Benoit Perron, and Dacheng Xiu. We thank seminar participants at Concordia University, HEC Montréal, Indiana University, McGill University, Surrey University, and University College London, as well as participants of the CESG 2016 conference, SoFiE 2016 conference, Financial Econometrics Conference at the Toulouse School of Economics, the 49th Annual Conference of the Canadian Economics Association, and the 11th World Congress of the Econometric Society.

†Department of Economics, University College London. E-mail address: i.kalnina@ucl.ac.uk. Kalnina's research was supported by *Institut de la finance structuré et des instruments dérivés de Montréal.* She is grateful to the Economics Department, the Gregory C. Chow Econometrics Research Program, and the Bendheim Center for Finance at Princeton University for their hospitality and support in the Spring of 2015.

‡Département de sciences économiques, Université de Montréal. E-mail address: kokouvi.tewou@umontreal.ca.

1 Introduction

In a panel of assets, returns are generally cross-sectionally dependent. This dependence is usually modelled using the exposure of assets to some common return factors. For example, the Capital Asset Pricing Model of Lintner (1965) and Sharpe (1964) has one return factor (the market portfolio), while the model of Fama and French (1993) has three return factors. The total volatility of an asset return can be decomposed into two parts: a component due to the exposure to the return factors, and a residual component termed the Idiosyncratic Volatility (IV). These two components of the volatility of returns are the most popular measures of the systematic risk and idiosyncratic risk of an asset.

Idiosyncratic volatility is important in economics and finance for several reasons. For example, when arbitrageurs exploit the mispricing of an individual asset, they are exposed to the idiosyncratic risk of the asset and not the systematic risk (see, e.g., Campbell, Lettau, Malkiel, and Xu (2001)). Also, idiosyncratic volatility measures the exposure to the idiosyncratic risk in imperfectly diversified portfolios. A recent observation is that the IVs seem to be strongly correlated in the cross-section of stocks.[1] Herskovic, Kelly, Lustig, and Nieuwerburgh (2016) argue this is due to a common IV factor, which they relate to household risk. Moreover, cross-sectional dependence in IVs is important for option pricing, see Gourier (2016).

This paper provides an econometric framework for studying the cross-sectional dependence in IVs using high frequency data. We show that the naive estimators, such as covariances and correlations of estimated IVs used by several empirical studies, are substantially biased. The bias arises due to the use of error-laden estimates of IVs. We provide the bias-corrected estimators.

We then study an Idiosyncratic Volatility Factor Model (IV-FM). (Throughout the paper, in accordance with the finance literature, we use the term "factor model" to denote a regression model as in Fama and French (1993).) Just like the Return Factor Model (R-FM), such as the Fama-French model, decomposes returns into systematic and idiosyncratic returns, the IV-FM decomposes the IVs into systematic (common) and residual components. The IV factors can include the volatility of the return factors, or, more generally, (possibly non-linear) transformations of the spot covariance matrices of any observable variables, such as the average variance and average correlation factors of Chen and Petkova (2012). The naive estimators of this decomposition also need to be bias-corrected, and we provide valid estimators. We also provide the asymptotic theory that allows us to test whether the residual (non-systematic) components of the IVs exhibit cross-sectional dependence. This allows us to identify the network of unexplained dependencies in the IVs across all stocks.

To provide the bias-corrected estimators and inference results, we develop a new asymptotic theory for general estimators of quadratic covariation of vector-valued and possibly nonlinear trans-

[1] See Duarte, Kamara, Siegel, and Sun (2014), Christoffersen, Fournier, and Jacobs (2015), and Herskovic, Kelly, Lustig, and Nieuwerburgh (2016))

formations of spot covariance matrices. This theoretical contribution is of its own interest. Two factors make the development of this asymptotic theory difficult. First, the preliminary estimation of volatility results in the first-order biases even for the univariate linear functional, as in Vetter (2012). Considering general nonlinear functionals in the multivariate setting substantially complicates the analysis.

We apply our methodology to high-frequency data on the 30 Dow Jones Industrial Average components. We study the IVs with respect to two models for asset returns: the CAPM and the three-factor Fama-French model. In both cases, the average pairwise correlation between the IVs is high (0.55). We verify that this dependence cannot be explained by the missing return factors. This confirms the recent findings of Herskovic, Kelly, Lustig, and Nieuwerburgh (2016) who use low frequency (daily and monthly) return data. We then consider the IV-FM. We use two sets of IV factors: the market volatility alone and the market volatility together with volatilities of nine industry ETFs. With the market volatility as the only IV factor, the average pairwise correlation between residual (non-systematic) IVs is substantially lower (0.25) than between the total IVs. With the additional nine industry ETF volatilities as IV factors, average correlation between the residual IVs decreases further (to 0.18). However, neither of the two sets of the IV factors can fully explain the cross-sectional dependence in the IVs. We map out the network of dependencies in residual IVs across all stocks.

The goal of this paper is to study cross-sectional dependence in idiosyncratic volatilities. This should not be confused with the analysis of cross-sectional dependence in total and idiosyncratic returns. A growing number of papers study the latter question using high frequency data. These date back to the analysis of realized covariances and their transformations, see, e.g., Barndorff-Nielsen and Shephard (2004) and Andersen, Bollerslev, Diebold, and Wu (2004). A continuous-time factor model for asset returns with observable return factors was first studied in Mykland and Zhang (2006). It was extended to multiple factors and jumps in Aït-Sahalia, Kalnina, and Xiu (2014). Related specifications with observable return factors are considered in Li, Todorov, and Tauchen (2014) and Bollerslev and Todorov (2010); see also Ait-Sahalia and Xiu (2015, 2016), and Pelger (2015). Importantly, the above papers do not consider the cross-sectional dependence structure in the IVs. The Beta GARCH model of Hansen, Lunde, and Voev (2014) implies that the IVs exhibit nonlinear cross-sectional dependencies driven by the market volatility and certain realized measures. Their model allows for some return factors to be omitted and hence tested for, but the IV factors are fixed. Our framework allows a general specification of both the return factors and the IV factors.

Our empirical analysis requires the availability of return factors at high frequency. The Fama-French factors are available on the website of Kenneth French only at the daily frequency. The high frequency Fama-French factors are provided by Aït-Sahalia, Kalnina, and Xiu (2014).

Our inference theory is related to several results in the existing literature. First, as mentioned above, it generalizes the result of Vetter (2012). Jacod and Rosenbaum (2012, 2013) estimate integrated functionals of volatilities, so they also use transformations of covariance matrices. However, the latter setting is simpler in the sense that $\sqrt{n}$-consistent estimation is possible, and no first-order bias terms due to preliminary estimation of volatilities arise. The need for a first-order bias correction due to preliminary estimation of volatility has also been observed in the literature on the estimation of the leverage effect, see Aït-Sahalia, Fan, and Li (2013), Aït-Sahalia, Fan, Laeven, Wang, and Yang (2013), Kalnina and Xiu (2015) and Wang and Mykland (2014). The biases due to preliminary estimation of volatility can be made theoretically negligible when an additional, long-span, asymptotic approximation is used. This requires the assumption that the frequency of observations is high enough compared to the time span, see, e.g., Corradi and Distaso (2006), Bandi and Renò (2012), Li and Patton (2015), and Kanaya and Kristensen (2015).

Cross-sectional dependence structure in the IVs is important for option pricing. For example, Gourier (2016) studies risk premia embedded in options using a parametric model with a factor structure in the IVs. She uses one IV factor, the market volatility. She finds that a factor structure in the IVs is a crucial feature of the model. By relying on high frequency data, our methods offer a nonparametric and computationally straightforward way of testing whether a given set of IV factors is sufficient to explain all the cross-sectional dependence in the IVs for a given data set. Empirically, we reject the hypothesis that the market volatility as the sole IV factor is sufficient for the data set of 30 DJIA stocks. Another related paper is Christoffersen, Fournier, and Jacobs (2015) who apply principal component analysis to stock option data. While their model is agnostic about the cross-sectional dependence in IVs, they report empirically high cross-sectional correlations in IVs that motivate our study.

The remainder of the paper is organized as follows. Section 2 introduces the model and the quantities of interest. Section 3 describes the identification and estimation. Section 4 presents the asymptotic properties of our estimators. Section 5 contains a Monte Carlo study. Section 6 uses high-frequency stock return data to study the cross-sectional dependence in IVs using our framework. All proofs are in the Appendix A.

2 Model and Quantities of Interest

We first describe a general factor model for the returns (R-FM), which allows us to define the idiosyncratic volatility. We then introduce the idiosyncratic volatility factor model (IV-FM). In this framework, we proceed to define the cross-sectional measures of dependence between the total IVs, as well as the residual IVs, which take into account the dependence induced by the IV factors.

We start by introducing some notation. Suppose we have (log) prices on d_S assets such as stocks and on d_F observable factors. We stack them into the d-dimensional process $Y_t = (S_{1,t}, \ldots, S_{d_S,t}, F_{1,t}, \ldots, F_{d_F,t})^\top$ where $d = d_S + d_F$. The observable factors $F_1, \ldots F_{d_F}$ are used in the R-FM model below. We assume that all observable variables jointly follow an Itô semimartingale, i.e., Y_t follows

$$Y_t = Y_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t,$$

where W is a d^W-dimensional Brownian motion ($d^W \geq d$), σ_s is a $d \times d^W$ stochastic volatility process, and J_t denotes a finite variation jump process. The reader can find the full list of assumptions in Section 4.1. We also assume that the spot covariance matrix process $C_t = \sigma_t \sigma_t^\top$ of Y_t is a continuous Itô semimartingale,[2]

$$C_t = C_0 + \int_0^t \widetilde{b}_s ds + \int_0^t \widetilde{\sigma}_s dW_s. \tag{1}$$

We denote $C_t = (C_{ab,t})_{1 \leq a,b \leq d}$. For convenience, we also use the alternative notation $C_{UV,t}$ to refer to the spot covariance between two elements U and V of Y.

We assume a standard continuous-time factor model for the asset returns.

Definition (Factor Model for Returns, R-FM). *For for all* $0 \leq t \leq T$ *and* $j = 1, \ldots, d_S$,[3]

$$\begin{aligned} dS_{j,t} &= \beta_{j,t}^\top dF_t^c + \tilde{\beta}_{j,t}^\top dF_t^d + dZ_{j,t} \quad \textit{with} \\ [Z_j, F]_t &= 0. \end{aligned} \tag{2}$$

In the above, $dZ_{j,t}$ is the idiosyncratic return of stock j. The superscripts c and d indicate the continuous and jump part of the processes, so that $\beta_{j,t}$ and $\tilde{\beta}_{j,t}$ are the continuous and jump factor loadings. For example, the k-th component of $\beta_{j,t}$ corresponds to the time-varying loading of the continuous part of the return on stock j to the continuous part of the return on the k-th factor. We set $\beta_t = (\beta_{1,t}, \ldots, \beta_{d_S,t})^\top$ and $Z_t = (Z_{1,t}, \ldots, Z_{d_S,t})^\top$.

We do not need the return factors F_t to be the same across assets to identify the model, but without loss of generality, we keep this structure as it is standard in empirical finance. These return factors are assumed to be observable, which is also standard. For example, in the empirical application, we use two sets of return factors: the market portfolio and the three Fama-French factors, which are constructed in Aït-Sahalia, Kalnina, and Xiu (2014).

A continuous-time factor model for returns with observable factors was originally studied in Mykland and Zhang (2006) in the case of one factor and in the absence of jumps. A burgeoning literature uses related models to study the cross-sectional dependence of total and/or idiosyncratic

[2]Note that assuming that Y and C are driven by the same d^W-dimensional Brownian motion W is without loss of generality provided that d' is large enough, see, e.g., equation (8.12) of Aït-Sahalia and Jacod (2014).

[3]The quadratic covariation of two vector-valued Itô semimartingales X and Y, over the time span $[0, T]$, is defined as

$$[X, Y]_T = \underset{M \to \infty}{p\text{-lim}} \sum_{j=0}^{M-1} (X_{t_{j+1}} - X_{t_j})(Y_{t_{j+1}} - Y_{t_j})^\top,$$

for any sequence $t_0 < t_1 < \ldots < t_M = T$ with $\sup_j \{t_{j+1} - t_j\} \to 0$ as $M \to \infty$, where p-lim stands for the probability limit.

returns, see Section 1 for details. This literature does not consider the cross-sectional dependence in the IVs. Below, we use the R-FM to define the IV, and proceed to study the cross-sectional dependence of IVs using the IV Factor Model.

We define the idiosyncratic Volatility (IV) as the spot volatility of the $Z_{j,t}$ process and denote it by C_{ZjZj}. Notice that the R-FM in (2) implies that the factor loadings β_t as well as IV are functions of the total spot covariance matrix C_t. In particular, the vector of factor loadings satisfies

$$\beta_{jt} = (C_{FF,t})^{-1} C_{FS_j,t}, \tag{3}$$

for $j = 1, \ldots, d_S$, where $C_{FF,t}$ denotes the spot covariance matrix of the factors F, which is the lower $d_F \times d_F$ sub-matrix of C_t; and $C_{FSj,t}$ denotes the covariance of the factors and the j^{th} stock, which is a vector consisting of the last d_F elements of the j^{th} column of C_t. The IV of stock j is also a function of the total spot covariance matrix C_t,

$$\underbrace{C_{ZjZj,t}}_{\text{IV of stock j}} = \underbrace{C_{YjYj,t}}_{\text{total volatility of stock j}} - (C_{FSj,t})^\top (C_{FF,t})^{-1} C_{FSj,t}. \tag{4}$$

By the Itô lemma, (3) and (4) imply that factor loadings and IVs are also Itô semimartingales with their characteristics related to those of C_t.

We now introduce the Idiosyncratic Volatility Factor model (IV-FM). In the IV-FM, the cross-sectional dependence in the IV shocks can be potentially explained by certain IV factors. We assume the IV factors are known functions of the matrix C_t. In the empirical application, we use the market volatility as the IV factor, which has been used in Herskovic, Kelly, Lustig, and Nieuwerburgh (2016) and Gourier (2016); we discuss other possibilities below. We allow the IV factors to be any known functions of C_t as long as they satisfy a certain polynomial growth condition in the sense of being in the class $\mathcal{G}(p)$ below,

$$\begin{aligned} \mathcal{G}(p) = \{H : H \text{ is three-times continuously differentiable and for some } K > 0, \\ \|\partial^j H(x)\| \leq K(1 + \|x\|)^{p-j}, j = 0, 1, 2, 3\}, \quad \text{for some } \ p \geq 3. \end{aligned} \tag{5}$$

Definition (Idiosyncratic Volatility Factor Model, IV-FM). *For all $0 \leq t \leq T$ and $j = 1, \ldots, d_S$, the idiosyncratic volatility C_{ZjZj} follows,*

$$\begin{aligned} dC_{ZjZj,t} &= \gamma_{Zj}^\top d\Pi_t + dC_{ZjZj,t}^{resid} \quad \textit{with} \\ [C_{ZjZj}^{resid}, \Pi]_t &= 0, \end{aligned} \tag{6}$$

where $\Pi_t = (\Pi_{1t}, \ldots, \Pi_{d_\Pi t})$ is a $\mathbb{R}^{d_\Pi}$-valued vector of IV factors, which satisfy $\Pi_{kt} = \Pi_k(C_t)$ with the function $\Pi_k(\cdot)$ belonging to $\mathcal{G}(p)$ for $k = 1, \ldots, d_\Pi$.

We call the residual term $C_{ZjZj,t}^{resid}$ the residual IV of asset j. Our assumptions imply that the components of the IV-FM, $C_{ZjZj,t}, \Pi_t$ and $C_{ZjZj,t}^{resid}$, are continuous Itô semimartingales. We remark that both the dependent variable and the regressors in our IV-FM are not directly observable and have to be estimated. As will see in Section 3, this preliminary estimation implies that the naive estimators of all the dependence measures defined below are biased. One of the contributions of this paper is to quantify this bias and provide the bias-corrected estimators for all the quantities of interest.

The class of IV factors permitted by our theory is rather wide as it includes general non-linear transforms of the spot covolatility process C_t. For example, IV factors can be linear combinations of the total volatilities of stocks, see, e.g., the average variance factor of Chen and Petkova (2012). Other examples of IV factors are linear combinations of the IVs, such as the equally-weighted average of the IVs, which Herskovic, Kelly, Lustig, and Nieuwerburgh (2016) denote by the "CIV". The IV factors can also be the volatilities of any other observable processes.

Having specified our econometric framework, we now provide the definitions of some natural measures of dependence for (residual) IVs. Their estimation is discussed in Section 3.

Before considering the effect of IV factors by using the IV-FM decomposition, one may be interested in quantifying the dependence between the IVs of two stocks i and j. A natural measure of dependence is the quadratic-covariation based correlation between the two IV processes,

$$\rho_{Zi,Zj} = \frac{[C_{ZiZi}, C_{ZjZj}]_T}{\sqrt{[C_{ZiZi}, C_{ZiZi}]_T}\sqrt{[C_{ZjZj}, C_{ZjZj}]_T}}. \tag{7}$$

Alternatively, one may consider the quadratic covariation $[C_{ZiZi}, C_{ZjZj}]_T$ without any normalization. In Section 4.4, we use the estimator of the latter quantity to test for the presence of cross-sectional dependence in IVs.

To measure the residual cross-sectional dependence between the IVs of two stocks after accounting for the effect of the IV factors, we use again the quadratic-covariation based correlation,

$$\rho^{resid}_{Zi,Zj} = \frac{[C^{resid}_{ZiZi}, C^{resid}_{ZjZj}]_T}{\sqrt{[C^{resid}_{ZiZi}, C^{resid}_{ZiZi}]_T}\sqrt{[C^{resid}_{ZjZj}, C^{resid}_{ZjZj}]_T}}. \tag{8}$$

In Section 4.4, we use the quadratic covariation between the two residual IV processes $[C^{resid}_{ZiZi}, C^{resid}_{ZjZj}]_T$ without normalization for testing purposes.

We want to capture how well the IV factors explain the time variation of IV of the j^{th} asset. For this purpose, we use the quadratic-covariation based analog of the coefficient of determination. For $j = 1, \ldots, d_S$,

$$R^{2,IV\text{-}FM}_{Zj} = \frac{\gamma^\top_{Zj}[\Pi, \Pi]_T \gamma_{Zj}}{[C_{ZjZj}, C_{ZjZj}]_T}. \tag{9}$$

It is interesting to compare the correlation measure between IVs in equation (7) with the correlation between the residual parts of IVs in (8). We consider their difference,

$$\rho_{Zi,Zj} - \rho^{resid}_{Zi,Zj}, \tag{10}$$

to see how much of the dependence between IVs can be attributed to the IV factors. In practice, if we compare assets that are known to have positive covolatilities (typically, stocks have that property), another useful measure of the common part in the overall covariation between IVs is the following quantity,

$$Q^{IV\text{-}FM}_{Zi,Zj} = \frac{\gamma^\top_{Zi}[\Pi,\Pi]_T\gamma_{Zj}}{[C_{ZiZi},C_{ZjZj}]_T}. \tag{11}$$

This measure is bounded by 1 if the covariations between residual IVs are nonnegative and smaller than the covariations between IVs, which is what we find for every pair in our empirical application with high-frequency observations on stock returns.

We remark that our framework can be compared with the following null hypothesis studied in Li, Todorov, and Tauchen (2013), $H_0:\ C_{ZjZj,t} = a_{Zj} + \gamma^\top_{Zj}\Pi_t,\ 0 \le t \le T$. This H_0 implies that the IV is a deterministic function of the factors, which does not allow for an error term. In particular, this null hypothesis implies $R^{2,IV\text{-}FM}_{Zj} = 1$.

3 Estimation

We now discuss the identification and estimation of the quantities of interest introduced in Section 2. The identification arguments are relatively simple. However, the estimation has to address the biases due to preliminary estimation of (idiosyncratic) volatility. The current section proposes two classes of bias-corrected estimators. Section 4 establishes their asymptotic properties.

We would like to estimate the following quantities defined in Section 2,

$$[C_{ZiZi},C_{ZjZj}]_T,\ \rho_{Zi,Zj},\ [C^{resid}_{ZjZj},\ C^{resid}_{ZjZj}]_T,\ \rho^{resid}_{Zi,Zj},\ Q^{IV\text{-}FM}_{Zi,Zj},\ \text{and}\ R^{2,IV\text{-}FM}_{Zi}, \tag{12}$$

for $i,j = 1,\dots,d_S$. The first two quantities in the above are defined even if only the R-FM holds; the last four need both the R-FM and IV-FM to hold to be well defined.

We first show that each of the quantities in (12) can be written as

$$\varphi\left([H_1(C),G_1(C)]_T,\dots,[H_\kappa(C),G_\kappa(C)]_T\right),$$

where φ as well as H_r and G_r, for $r = 1,\dots,\kappa$, are known real-valued functions. Each element in this expression is of the form $[H(C),G(C)]_T$, i.e., it is a quadratic covariation between functions of C_t. Afterwards, we present two methods to estimate $[H(C),G(C)]_T$.

First, consider the quadratic covariation between i^{th} and j^{th} IV, $[C_{ZiZi},C_{ZjZj}]_T$. It can be written as $[H(C),G(C)]_T$ if we choose $H(C_t) = C_{ZiZi,t}$ and $G(C_t) = C_{ZjZj,t}$. By (4), both $C_{ZiZi,t}$ and $C_{ZjZj,t}$ are functions of C_t. Next, consider the correlation $\rho_{Zi,Zj}$ defined in (7). By the same argument, its numerator and each of the two components in the denominator can be written as $[H(C),G(C)]_T$ for different functions H and G. Therefore, $\rho_{Zi,Zj}$ is itself a known function of three objects of the form $[H(C),G(C)]_T$.

To show that the remaining quantities in (12) can also be expressed in terms of objects of the form $[H(C), G(C)]_T$, note that the IV-FM implies

$$\gamma_{Zj} = ([\Pi, \Pi]_T)^{-1} [\Pi, C_{ZjZj}]_T \quad \text{and} \quad [C^{resid}_{ZiZi}, C^{resid}_{ZjZj}]_T = [C_{ZiZi}, C_{ZjZj}]_T - \gamma_{Zi}^\top [\Pi, \Pi]_T \gamma_{Zj},$$

for $i, j = 1, \ldots, d_S$. Since $C_{ZiZi,t}$, $C_{ZjZj,t}$ and every element in Π_t are real-valued functions of C_t, the above equalities imply that all quantities of interest in (12) can be written as real-valued, known functions of a finite number of quantities of the form $[H(C), G(C)]_T$.

We now discuss the estimation of $[H(C), G(C)]_T$. Suppose we have discrete observations on Y_t over an interval $[0, T]$. Denote by Δ_n the distance between observations. It is well known that we can estimate the spot covariance matrix C_t at time $(i-1)\Delta_n$ with a local truncated realized volatility estimator (Mancini (2001)),

$$\widehat{C}_{i\Delta_n} = \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n - 1} (\Delta^n_{i+j} Y)(\Delta^n_{i+j} Y)^\top 1_{\{\|\Delta^n_{i+j} Y\| \leq \chi \Delta_n^{\varpi}\}}, \tag{13}$$

where $\Delta^n_i Y = Y_{i\Delta_n} - Y_{(i-1)\Delta_n}$ and where k_n is the number of observations in a local window.[4] Throughout the paper we set $\widehat{C}_{i\Delta_n} = (\widehat{C}_{ab,i\Delta_n})_{1 \leq a,b \leq d}$.

We propose two estimators for the general quantity $[H(C), G(C)]_T$.[5] The first is based on the analog of the definition of quadratic covariation between two Itô processes,

$$\begin{aligned} \widehat{[H(C), G(C)]}^{AN}_T = \frac{3}{2k_n} \sum_{i=1}^{[T/\Delta_n] - 2k_n + 1} \Bigg(& \Big(H(\widehat{C}_{(i+k_n)\Delta_n}) - H(\widehat{C}_{i\Delta_n})\Big)\Big(G(\widehat{C}_{(i+k_n)\Delta_n}) - G(\widehat{C}_{i\Delta_n})\Big) \\ & - \frac{2}{k_n} \sum_{g,h,a,b=1}^{d} (\partial_{gh} H \partial_{ab} G)(\widehat{C}_{i\Delta_n}) \Big(\widehat{C}_{ga,i\Delta_n} \widehat{C}_{gb,i\Delta_n} + \widehat{C}_{gb,i\Delta_n} \widehat{C}_{ha,i\Delta_n}\Big)\Bigg), \end{aligned} \tag{14}$$

where the factor $3/2$ and last term correct for the biases arising due to the estimation of volatility C_t. The increments used in the above expression are computed over overlapping blocks, which results in a smaller asymptotic variance compared to the version using non-overlapping blocks.

Our second estimator is based on the following equality, which follows by the Itô lemma,

$$[H(C), G(C)]_T = \sum_{g,h,a,b=1}^{d} \int_0^T \big(\partial_{gh} H \partial_{ab} G\big)(C_t) \overline{C}^{gh,ab}_t dt, \tag{15}$$

where $\overline{C}^{gh,ab}_t$ denotes the covariation between the volatility processes $C_{gh,t}$ and $C_{ab,t}$. The quantity is thus a non-linear functional of the spot covariance and spot volatility of volatility matrices. Our

[4] It is also possible to define more flexible kernel-based estimators as in Kristensen (2010).

[5] As evident from their formulas, the computation time required for the calculation of the two estimators is increasing with the number of stocks and factors d. To ease the implementation of the procedure, we compute all the quantities of interest for pairs of stocks which means practically one needs only to set $d_S = 2$ so that $d = d_F + 2$.

second estimator is based on this "linearized" expression,

$$
\begin{aligned}
[H(\widehat{C)}, \widehat{G}(C)]_T^{LIN} &= \frac{3}{2k_n} \sum_{g,h,a,b=1}^{d} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} (\partial_{gh} H \partial_{ab} G)(\widehat{C}_{i\Delta_n}) \times \\
\Big((\widehat{C}_{gh,(i+k_n)\Delta_n} - \widehat{C}_{gh,i\Delta_n})&(\widehat{C}_{ab,(i+k_n)\Delta_n} - \widehat{C}_{ab,i\Delta_n}) - \frac{2}{k_n}(\widehat{C}_{ga,i\Delta_n}\widehat{C}_{gb,i\Delta_n} + \widehat{C}_{gb,i\Delta_n}\widehat{C}_{ha,i\Delta_n}) \Big).
\end{aligned} \tag{16}
$$

Consistency for a similar estimator has been established by Jacod and Rosenbaum (2012).[6] We go beyond their result by deriving the asymptotic distribution and proposing a consistent estimator of its asymptotic variance.

Note that the same additive bias-correcting term,

$$
-\frac{3}{k_n^2} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} \left(\sum_{g,h,a,b=1}^{d} (\partial_{gh} H \partial_{ab} G)(\widehat{C}_{i\Delta_n}) \Big(\widehat{C}_{ga,i\Delta_n}\widehat{C}_{gb,i\Delta_n} + \widehat{C}_{gb,i\Delta_n}\widehat{C}_{ha,i\Delta_n} \Big) \right), \tag{17}
$$

is used for the two estimators. This term is (up to a scale factor) an estimator of the asymptotic covariance between the sampling errors embedded in estimators of $\int_0^T H(C_t)dt$ and $\int_0^T G(C_t)dt$ defined in Jacod and Rosenbaum (2013).

The two estimators are identical when H and G are linear, for example, when estimating the covariation between two volatility processes. In the univariate case $d = 1$, when $H(C) = G(C) = C$, our estimator coincides with the volatility of volatility estimator of Vetter (2012), which was extended to allow for jumps in Jacod and Rosenbaum (2012). Our contribution is the extension of this theory to the multivariate $d > 1$ case with nonlinear functionals.

4 Asymptotic Properties

In this section, we first present the full list of assumptions for our asymptotic results. We then state the asymptotic distribution for the general functionals introduced in the previous section, and develop estimators for the asymptotic variance. Finally, to illustrate the application of the general theory, we describe three statistical tests about the IVs, which we later implement in the empirical and Monte Carlo analysis.

4.1 Assumptions

Recall that the d-dimensional process Y_t represents the (log) prices of stocks, S_t, and factors F_t.

[6] Jacod and Rosenbaum (2012) derive the probability limit of the following estimator:

$$
\frac{3}{2k_n} \sum_{g,h,a,b=1}^{d} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} (\partial^2_{gh,ab} H)(\widehat{C}_{i\Delta_n}) \Big((\widehat{C}_{(i+k_n)\Delta_n} - \widehat{C}_{i\Delta_n})(\widehat{C}_{(i+k_n)\Delta_n} - \widehat{C}_{i\Delta_n}) - \frac{2}{k_n}(\widehat{C}_{ga,i\Delta_n}\widehat{C}_{gb,i\Delta_n} + \widehat{C}_{gb,i\Delta_n}\widehat{C}_{ha,i\Delta_n}) \Big).
$$

Assumption 1. *Suppose Y is an Itô semimartingale on a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$,*

$$Y_t = Y_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_E \delta(s, z)\mu(ds, dz),$$

where W is a d^W-dimensional Brownian motion ($d^W \geq d$) and μ is a Poisson random measure on $\mathbb{R}_+ \times E$, with E an auxiliary Polish space with intensity measure $\nu(dt, dz) = dt \otimes \lambda(dz)$ for some σ-finite measure λ on E. The process b_t is $\mathbb{R}^d$-valued optional, σ_t is $\mathbb{R}^d \times \mathbb{R}^{d^W}$-valued, and $\delta = \delta(w, t, z)$ is a predictable $\mathbb{R}^d$ -valued function on $\Omega \times \mathbb{R}_+ \times E$. Moreover, $\|\delta(w, t\wedge\tau_m(w), z)\| \wedge 1 \leq \Gamma_m(z)$, for all (w,t,z), where (τ_m) is a localizing sequence of stopping times and, for some $r \in [0, 1]$, the function Γ_m on E satisfies $\int_E \Gamma_m(z)^r \lambda(dz) < \infty$. The spot volatility matrix of Y is then defined as $C_t = \sigma_t\sigma_t^\top$. We assume that C_t is a continuous Itô semimartingale,[7]

$$C_t = C_0 + \int_0^t \widetilde{b}_s ds + \int_0^t \widetilde{\sigma}_s dW_s. \tag{18}$$

where $\widetilde{b}$ is $\mathbb{R}^d \times \mathbb{R}^d$-valued optional.

With the above notation, the elements of the spot volatility of volatility matrix and spot covariation of the continuous martingale parts of X and c are defined as follows,

$$\overline{C}_t^{gh,ab} = \sum_{m=1}^{d^W} \widetilde{\sigma}_t^{gh,m} \widetilde{\sigma}_t^{ab,m}, \quad \overline{C}_t^{\prime g,ab} = \sum_{m=1}^{d^W} \sigma_t^{gm} \widetilde{\sigma}_t^{ab,m}. \tag{19}$$

We assume the following for the process $\widetilde{\sigma}_t$:

Assumption 2. *$\widetilde{\sigma}_t$ is a continuous Itô semimartingale with its characteristics satisfying the same requirements as that of C_t.*

Assumption 1 is very general and nests most of the multivariate continuous-time models used in economics and finance. It allows for potential stochastic volatility and jumps in returns. Assumption 2 is required to obtain the asymptotic distribution of estimators of the quadratic covariation between functionals of the spot covariance matrix C_t. It is not needed to prove consistency. This assumption also appears in Vetter (2012), Kalnina and Xiu (2015) and Wang and Mykland (2014).

4.2 Asymptotic Distribution

We have seen in Section 3 that all quantities of interest in (12) are functions of multiple objects of the form $[H(C), G(C)]_T$. Therefore, if we can obtain a multivariate asymptotic distribution for a vector with elements of the form $[H(C), G(C)]_T$, the asymptotic distributions for all our estimators follow by the delta method. Presenting this asymptotic distribution is the purpose of the current section.

[7]Note that $\widetilde{\sigma}_s = (\widetilde{\sigma}_s^{gh,m})$ is $(d \times d \times d^W)$-dimensional and $\widetilde{\sigma}_s dW_s$ is $(d \times d)$-dimensional with $(\widetilde{\sigma}_s dW_s)^{gh} = \sum_{m=1}^{d^W} \widetilde{\sigma}_s^{gh,m} dW_s^m$.

Let $H_1, G_1, \ldots, H_\kappa, G_\kappa$ be some arbitrary elements of $\mathcal{G}(p)$ defined in equation (5). We are interested in the asymptotic behavior of vectors

$$\left([H_1(\widehat{C), G_1}(C)]_T^{AN}, \ldots, [H_\kappa(\widehat{C), G_\kappa}(C)]_T^{AN}\right)^\top \text{ and } \left([H_1(\widehat{C), G_1}(C)]_T^{LIN}, \ldots, [H_\kappa(\widehat{C), G_\kappa}(C)]_T^{LIN}\right)^\top.$$

The smoothness requirement on the different functions H_j and G_j is useful for obtaining the asymptotic distribution of the bias correcting terms (see for example Jacod and Rosenbaum (2012) and Jacod and Rosenbaum (2013)). The following theorem summarizes the joint asymptotic behavior of the estimators.

Theorem 1. *Let $[H_r(\widehat{C), G_r}(C)]_T$ be either $[H_r(\widehat{C), G_r}(C)]_T^{AN}$ or $[H_r(\widehat{C), G_r}(C)]_T^{LIN}$ defined in (14) and (16), respectively. Suppose Assumption 1 and Assumption 2 hold. Fix $k_n = \theta\Delta_n^{-1/2}$ for some $\theta \in (0, \infty)$ and set $(8p-1)/4(4p-r) \leq \varpi < \frac{1}{2}$. Then, as $\Delta_n \to 0$,*

$$\Delta_n^{-1/4} \begin{pmatrix} [H_1(\widehat{C), G_1}(C)]_T - [H_1(C), G_1(C)]_T \\ \vdots \\ [H_\kappa(\widehat{C), G_\kappa}(C)]_T - [H_\kappa(C), G_\kappa(C)]_T \end{pmatrix} \xrightarrow{L\text{-}s} MN(0, \Sigma_T), \tag{20}$$

where $\Sigma_T = \left(\Sigma_T^{r,s}\right)_{1 \leq r,s \leq \kappa}$ denotes the asymptotic covariance between the estimators $[H_r(\widehat{C), G_r}(C)]_T$ and $[H_s(\widehat{C), G_s}(C)]_T$. The elements of the matrix Σ_T are

$$\begin{aligned}
\Sigma_T^{r,s} &= \Sigma_T^{r,s,(1)} + \Sigma_T^{r,s,(2)} + \Sigma_T^{r,s,(3)}, \\
\Sigma_T^{r,s,(1)} &= \frac{6}{\theta^3} \sum_{g,h,a,b=1}^{d} \sum_{j,k,l,m=1}^{d} \int_0^T \left(\partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(C_s)\right) \Big[C_t(gh, jk) C_t(ab, lm) \\
&\qquad + C_t(ab, jk) C_t(gh, lm)\Big] dt, \\
\Sigma_T^{r,s,(2)} &= \frac{151\theta}{140} \sum_{g,h,a,b=1}^{d} \sum_{j,k,l,m=1}^{d} \int_0^T \left(\partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(C_t)\right) \Big[\overline{C}_t^{gh,jk} \overline{C}_t^{ab,lm} + \overline{C}_t^{ab,jk} \overline{C}_t^{gh,lm}\Big] dt, \\
\Sigma_T^{r,s,(3)} &= \frac{3}{2\theta} \sum_{g,h,a,b=1}^{d} \sum_{j,k,l,m=1}^{d} \int_0^T \left(\partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(C_t)\right) \Big[C_t(gh, jk) \overline{C}_t^{ab,lm} + C_t(ab, lm) \overline{C}_t^{gh,jk} \\
&\qquad + C_t(gh, lm) \overline{C}_t^{ab,jk} + C_t(ab, jk) \overline{C}_t^{gh,lm}\Big] dt,
\end{aligned}$$

with

$$C_t(gh, jk) = C_{gj,t} C_{hk,t} + C_{gk,t} C_{hj,t}.$$

The convergence in Theorem 1 is stable in law (denoted L-s, see for example Aldous and Eagleson (1978) and Jacod and Protter (2012)). The limit is mixed gaussian and the precision of the estimators depends on the paths of the spot covariance and the volatility of volatility process. The rate of convergence $\Delta_n^{-1/4}$ has been shown to be the optimal for volatility of volatility estimation (under the assumption of no volatility jumps).

The asymptotic variance of the estimators depends on the tuning parameter θ whose choice may be crucial for the reliability of the inference. We document the sensitivity of the inference theory to the choice of the parameter θ in a Monte Carlo experiment (see Section 5).

4.3 Estimation of the Asymptotic Covariance Matrix

To provide a consistent estimator for the element $\Sigma_T^{r,s}$ of the asymptotic covariance matrix in Theorem 1, we introduce the following quantities:

$$\widehat{\Omega}_T^{r,s,(1)} = \Delta_n \sum_{g,h,a,b=1}^{d} \sum_{j,k,l,m=1}^{d} \sum_{i=1}^{[T/\Delta_n]-4k_n+1} \left(\partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(\widehat{C}_{i\Delta_n})\right) \Big[\widetilde{C}_{i\Delta_n}(gh,jk)\widetilde{C}_{i\Delta_n}(ab,lm)$$
$$+ \widetilde{C}_{i\Delta_n}(ab,jk)\widetilde{C}_{i\Delta_n}(gh,lm)\Big],$$
$$\widehat{\Omega}_T^{r,s,(2)} = \sum_{g,h,a,b=1}^{d} \sum_{j,k,l,m=1}^{d} \sum_{i=1}^{[T/\Delta_n]-4k_n+1} \left(\partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(\widehat{C}_{i\Delta_n})\right) \Big[\frac{1}{2}\widehat{\lambda}_i^{n,gh}\widehat{\lambda}_i^{n,jk}\widehat{\lambda}_{i+2k_n}^{n,ab}\widehat{\lambda}_{i+2k_n}^{n,lm} +$$
$$\frac{1}{2}\widehat{\lambda}_i^{n,ab}\widehat{\lambda}_i^{n,lm}\widehat{\lambda}_{i+2k_n}^{n,gh}\widehat{\lambda}_{i+2k_n}^{n,jk} + \frac{1}{2}\widehat{\lambda}_i^{n,ab}\widehat{\lambda}_i^{n,jk}\widehat{\lambda}_{i+2k_n}^{n,gh}\widehat{\lambda}_{i+2k_n}^{n,lm} + \frac{1}{2}\widehat{\lambda}_i^{n,gh}\widehat{\lambda}_i^{n,lm}\widehat{\lambda}_{i+2k_n}^{n,ab}\widehat{\lambda}_{i+2k_n}^{n,jk}\Big],$$
$$\widehat{\Omega}_T^{r,s,(3)} = \frac{3}{2k_n} \sum_{g,h,a,b=1}^{d} \sum_{j,k,l,m=1}^{d} \sum_{i=1}^{[T/\Delta_n]-4k_n+1} \left(\partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(\widehat{C}_{i\Delta_n})\right)\times$$
$$\Big[\widetilde{C}_{i\Delta_n}(gh,jk)\widehat{\lambda}_i^{n,ab}\widehat{\lambda}_i^{n,lm} + \widetilde{C}_{i\Delta_n}(ab,lm)\widehat{\lambda}_i^{n,gh}\widehat{\lambda}_i^{n,jk} + \widetilde{C}_{i\Delta_n}(gh,lm)\widehat{\lambda}_i^{n,ab}\widehat{\lambda}_i^{n,jk} + (\widetilde{C}_{i\Delta_n}(ab,jk)\widehat{\lambda}_i^{n,gh}\widehat{\lambda}_i^{n,lm}\Big],$$

with $\widehat{\lambda}_i^{n,jk} = \widehat{C}_{i+k_n}^{n,jk} - \widehat{C}_i^{n,jk}$ and $\widetilde{C}_{i\Delta_n}(gh,jk) = (\widehat{C}_{gj,i\Delta_n}\widehat{C}_{hk,i\Delta_n} + \widehat{C}_{gk,i\Delta_n}\widehat{C}_{hj,i\Delta_n})$.
The following result holds,

Theorem 2. *Suppose the assumptions of Theorem 1 hold, then, as* $\Delta_n \longrightarrow 0$

$$\frac{6}{\theta^3}\widehat{\Omega}_T^{r,s,(1)} \xrightarrow{\mathbb{P}} \Sigma_T^{r,s,(1)} \tag{21}$$
$$\frac{3}{2\theta}[\widehat{\Omega}_T^{r,s,(3)} - \frac{6}{\theta}\widehat{\Omega}_T^{r,s,(1)}] \xrightarrow{\mathbb{P}} \Sigma_T^{r,s,(3)} \tag{22}$$
$$\frac{151\theta}{140}\frac{9}{4\theta^2}[\widehat{\Omega}_T^{r,s,(2)} + \frac{4}{\theta^2}\widehat{\Omega}_T^{r,s,(1)} - \frac{4}{3}\widehat{\Omega}_T^{r,s,(3)}] \xrightarrow{\mathbb{P}} \Sigma_T^{r,s,(2)}. \tag{23}$$

The estimated matrix $\widehat{\Sigma}_T$ is symmetric but is not guaranteed to be positive semi-definite. By Theorem 1, $\widehat{\Sigma}_T$ is positive semi-definite in large samples. An interesting question is the estimation of the asymptotic variance using subsampling or bootstrap methods, and we leave it for future research.

Remark 1: Results of Jacod and Rosenbaum (2012) and a straightforward extension of Theorem 1 can be used to show that the rate of convergence in equation (21) is $\Delta_n^{-1/2}$, and the rate of convergence in (23) is $\Delta_n^{-1/4}$. The rate of convergence in (22) can be shown to be $\Delta_n^{-1/4}$.

Remark 2: In the one-dimensional case ($d = 1$), much simpler estimators of $\Sigma_T^{r,s,(2)}$ can be constructed using the quantities $\widehat{\lambda}_i^{n,jk}\widehat{\lambda}_i^{n,lm}\widehat{\lambda}_{i+k_n}^{n,gh}\widehat{\lambda}_{i+k_n}^{n,xy}$ or $\widehat{\lambda}_i^{n,jk}\widehat{\lambda}_i^{n,lm}\widehat{\lambda}_i^{n,gh}\widehat{\lambda}_i^{n,xy}$ as in Vetter (2012).

However, in the multidimensional case, the latter quantities do not identify separately the quantity $\overline{C_t}^{jk,lm}\overline{C_t}^{gh,xy}$ since the combination $\overline{C_t}^{jk,lm}\overline{C_t}^{gh,xy} + \overline{C_t}^{jk,gh}\overline{C_t}^{lm,xy} + \overline{C_t}^{jk,xy}\overline{C_t}^{gh,lm}$ shows up in a non-trivial way in the limit of the estimator.

Corollary 3. *For $1 \leq r \leq \kappa$, let $[H_r(\widehat{C), G_r}(C)]_T$ be either $[H_r(\widehat{C), G_r}(C)]_T^{AN}$ or $[H_r(\widehat{C), G_r}(C)]_T^{LIN}$ defined in (16) and (14), respectively. Suppose the assumptions of theorem 1 hold. Then,*

$$\Delta_n^{-1/4}\widehat{\Sigma}_T^{-1/2}\begin{pmatrix} [H_1(\widehat{C), G_1}(C)]_T - [H_1(C), G_1(C)]_T \\ \vdots \\ [H_\kappa(\widehat{C), G_\kappa}(C)]_T - [H_\kappa(C), G_\kappa(C)]_T \end{pmatrix} \xrightarrow{L} N(0, I_\kappa). \tag{24}$$

In the above, we use the notation L to denote the convergence in distribution and I_κ the identity matrix of order κ. Corollary 3 states the standardized asymptotic distribution, which follows directly from the properties of stable-in-law convergence. Similarly, by the delta method, standardized asymptotic distribution can also be derived for the estimators of the quantities in (12). These standardized distributions allow the construction of confidence intervals for all the latent quantities of the form $[H_r(C), G_r(C)]_T$ and, more generally, functions of these quantities.

4.4 Tests

As an illustration of application of the general theory, we provide three tests about the dependence of idiosyncratic volatility. Our framework allows to test general hypotheses about the joint dynamics of any subset of the available stocks. The three examples below are stated for one pair of stocks, and correspond to the tests we implement in the empirical and Monte Carlo studies.

First, one can test for the absence of dependence between the IVs of the returns on assets i and j,

$$H_0^1 : [C_{ZiZi}, C_{ZjZj}]_T = 0 \text{ vs } H_1^1 : [C_{ZiZi}, C_{ZjZj}]_T \neq 0.$$

The null hypothesis H_0^1 is rejected whenever the t-test exceeds the $\alpha/2$-quantile of the standard normal distribution, Z_α,

$$\Delta_n^{-1/4}\frac{\left|[C_{ZiZi}, \widehat{C_{ZjZj}}]_T\right|}{\sqrt{\widehat{\text{AVAR}}\left(C_{ZiZi}, C_{ZjZj}\right)}} > Z_{\alpha/2}.$$

Second, we can test for all the IV factors Π being irrelevant to explain the dynamics of IV shocks of stock j,

$$H_0^2 : [C_{ZjZj}, \Pi]_T = 0 \text{ vs } H_1^2 : [C_{ZjZj}, \Pi]_T \neq 0.$$

Under this null hypothesis, the vector of IV factor loadings equals zero, $\gamma_{Z_j} = 0$. The null hypothesis H_0^2 is rejected when

$$\Delta_n^{-1/4}\left(\widehat{[C_{ZjZj},\Pi]}_T\right)^\top\left(\widehat{\text{AVAR}}\left(C_{ZjZj},\Pi\right)\right)^{-1}\widehat{[C_{ZjZj},\Pi]}_T>\mathcal{X}^2_{d_\Pi,1-\alpha},\tag{25}$$

where d_Π denotes the number of IV factors, and where $\mathcal{X}^2_{d_q,1-\alpha}$ is the $(1-\alpha)$ quantile of the $\mathcal{X}^2_{d_q}$ distribution. One can of course also construct a t-test for irrelevance of any one particular IV factor. The final example is a test for absence of dependence between the residual IVs,

$$H_0^3:[C^{resid}_{ZiZi},C^{resid}_{ZjZj}]_T=0\ \text{ vs }\ H_1^3:[C^{resid}_{ZiZi},C^{resid}_{ZjZj}]_T\neq 0.$$

The null can be rejected when the following t-test exceeds the critical value,

$$\Delta_n^{-1/4}\frac{\left|\widehat{[C^{resid}_{ZiZi},C^{resid}_{ZjZj}]}_T\right|}{\sqrt{\widehat{\text{AVAR}}\left(C^{resid}_{ZiZi},C^{resid}_{ZjZj}\right)}}>Z_{\alpha/2}.\tag{26}$$

Each of the above estimators

$$\widehat{[C_{ZiZi},C_{ZjZj}]}_T,\ \widehat{[C_{ZjZj},\Pi]}_T,\ \text{and}\ \widehat{[C^{resid}_{ZiZi},C^{resid}_{ZjZj}]}_T$$

can be obtained by choosing appropriate pair(s) of transformations H and G in the general estimator $\widehat{[H(C),G(C)]}_T$, see Section 3 for details. Any of the two types of the latter estimator can be used,

$$\widehat{[H(C),G(C)]}_T^{AN}\ \text{ or }\ \widehat{[H(C),G(C)]}_T^{LIN}.$$

For the first two tests, the expression for the true asymptotic variance, AVAR, is obtained using Theorem 1 and its estimation follows from Theorem 2. The asymptotic variance in the third test is obtained by applying the delta method to the joint convergence result in Theorem 1. The expression for the estimator of the asymptotic variance, $\widehat{\text{AVAR}}$, follows from Theorem 2. Under R-FM and the assumptions of Theorem 1, Corollary 3 implies that the asymptotic size of the two types of tests for the null hypotheses H_0^1 and H_0^2 is α, and their power approaches 1. The same properties apply for the tests of the null hypotheses H_0^3 with our R-FM and IV-FM representations.

Theoretically, it is possible to test for absence of dependence in the IVs at each point in time. In this case the null hypothesis is $H_0^{1\prime}:[C_{ZiZi},C_{ZjZj}]_t=0$ for all $0\leq t\leq T$, which is, in theory, stronger than our $H_0^{1\prime}$. In particular, Theorem 1 can be used to set up Kolmogorov-Smirnov type of tests for $H_0^{\prime 1}$ in the same spirit as Vetter (2012). However, we do not pursue this direction in the current paper for two reasons. First, the testing procedure would be more involved. Second, empirical evidence suggests nonnegative dependence between IVs, which means that in practice, it is not too restrictive to assume $[C_{ZiZi},C_{ZjZj}]_t\geq 0\ \forall t$, under which H_0^1 and $H_0^{1\prime}$ are equivalent.

5 Monte Carlo

This section investigates the finite sample properties of our estimators and tests. The data generating process (DGP) is similar to that of Li, Todorov, and Tauchen (2013) and is constructed as follows. Denote by Y_1 and Y_2 log-prices of two individual stocks, and by X the log-price of the market portfolio. Recall that the superscript c indicates the continuous part of a process. We assume

$$dX_t = dX_t^c + dJ_{3,t}, \quad dX_t^c = \sqrt{c_{X,t}} dW_t,$$

and, for $j = 1, 2$,

$$dY_{j,t} = \beta_t dX_t^c + d\widetilde{Y}_{j,t}^c + dJ_{j,t}, \quad d\widetilde{Y}_{j,t}^c = \sqrt{c_{Zj,t}} d\widetilde{W}_{j,t}.$$

In the above, c_X is the spot volatility of the market portfolio, $\widetilde{W}_1$, and $\widetilde{W}_2$ are Brownian motions with $\mathrm{Corr}(d\widetilde{W}_{1,t}, d\widetilde{W}_{2,t}) = 0.4$, and W is an independent Brownian motion; J_1, J_2, and J_3 are independent compound Poisson processes with intensity equal to 2 jumps per year and jump size distribution N(0,0.02^2). The beta process is time-varying and is specified as $\beta_t = 0.5 + 0.1\sin(100t)$.

We next specify the volatility processes. As our building blocks, we first generate four processes $f_1, \ldots, f_4$ as mutually independent Cox-Ingersoll-Ross processes,

$$\begin{aligned} df_{1,t} &= 5(0.09 - f_{1,t})dt + 0.35\sqrt{f_{1,t}}\Big(-0.8dW_t + \sqrt{1-0.8^2}dB_{1,t}\Big), \\ df_{j,t} &= 5(0.09 - f_{j,t})dt + 0.35\sqrt{f_{1,t}}dB_{j,t} \quad , \text{ for } j = 2, 3, 4, \end{aligned}$$

where $B_1, \ldots, B_4$ and independent standard Brownian Motions, which are also independent from the Brownian Motions of the return Factor Model.[8] We use the first process f_1 as the market volatility, i.e., $c_{X,t} = f_{1,t}$. We use the other three processes f_2, f_3, and f_4 to construct three different specifications for the IV processes $c_{Z1,t}$ and $c_{Z2,t}$, see Table 1 for details. The common Brownian Motion W_t in the market portfolio price process X_t and its volatility process $c_{X,t} = f_{1,t}$ generates a leverage effect for the market portfolio. The value of the leverage effect is -0.8, which is standard in the literature, see Kalnina and Xiu (2015), Aït-Sahalia, Fan, and Li (2013) and Aït-Sahalia, Fan, Laeven, Wang, and Yang (2013).

	$c_{Z1Z1,t}$	$c_{Z2Z2,t}$
Model 1	$0.1 + 1.5f_{2,t}$	$0.1 + 1.5f_{3,t}$
Model 2	$0.1 + 0.6c_{XX,t} + 0.4f_{2,t}$	$0.1 + 0.5c_{XX,t} + 0.5f_{3,t}$
Model 3	$0.1 + 0.45c_{XX,t} + f_{2,t} + 0.4f_{4,t}$	$0.1 + 0.35c_{XX,t} + 0.3f_{3,t} + 0.6f_{4,t}$

Table 1: Different specifications for the Idiosyncratic Volatility processes $c_{Z1,t}$and $c_{Z2,t}$.

We set the time span T equal 1260 or 2520 days, which correspond approximately to 5 and 10

[8]The Feller property is satisfied implying the positiveness of the processes $(f_{j,t})_{1\le j\le 4}$.

business years. These values are close to those typically used in the nonparametric leverage effect estimation literature (see Aït-Sahalia, Fan, and Li (2013) and Kalnina and Xiu (2015)), which is related to the problem of volatility of volatility estimation. Each day consists of 6.5 trading hours. We consider two different values for the sampling frequency, $\Delta_n = 1$ minute and $\Delta_n = 5$ minutes. We follow Li, Todorov, and Tauchen (2013) and set the truncation threshold u_n in day t at $3\widehat{\sigma}_t\Delta_n^{0.49}$, where $\widehat{\sigma}_t$ is the squared root of the annualized bipower variation of Barndorff-Nielsen and Shephard (2004). We use 10 000 Monte Carlo replications in all the experiments.

We first investigate the finite sample properties of the estimators under Model 3. The considered estimators include:

- the IV factor loading of the first stock (γ_{Z1}),
- the contribution of the market volatility to the variation of the IV of the first stock $(R_{Z1}^{2,IV\text{-}FM})$,
- the correlation between the idiosyncratic volatilities of stocks 1 and 2 $(\rho_{Z1,Z2})$,
- the correlation between residual idiosyncratic volatilities $(\rho_{Z1,Z2}^{resid})$,

The interpretation of simulation results is much simpler when the quantities of interest do not change across simulations. To achieve that, we generate once and keep fixed the paths of the processes $C_{XX,t}$ and $(f_{j,t})_{0\leq j\leq 27}$ and replicate several times the other parts of the DGP. In Table 2, we report the bias and the interquartile range (IQR) of the two type of estimators for each quantity using 5 minutes data sampled over 10 years. We choose four different values for the width of the subsamples, which corresponds to $\theta = 1.5, 2, 2.5$ and 3 (recall that the number of observations in a window is $k_n = \theta/\sqrt{\Delta_n}$). It seems that larger values of the parameters produce better results. Next, we investigate how these results change when we increase the sampling frequency. In Table 3, we report the results with $\Delta_n = 1$ minute in the same setting. We note a reduction of the bias and IQR at all levels of significance. However, the magnitude of the decrease of the IQR is very small. Finally, we conduct the same experiment using data sampled at one minute over 5 years. Despite using more than twice as many observations than in the first experiment, the precision is not as good. In other words, increasing the time span is more effective for precision gain than increasing the sampling frequency. This result is typical for $\Delta_n^{1/4}$-convergent estimators, see, e.g., Kalnina and Xiu (2015).

Next, we study the size and power of the three statistical tests as outlined in Section 4.4. We use Model 1 to study the size properties of the first two tests: the test of the absence of dependence between the IVs ($H_0^1 : [C_{Z1Z1}, C_{Z2Z2}]_T = 0$), and the absence of dependence between the IV of the first stock and the market volatility ($H_0^2 : [C_{Z1Z1}, C_{XX}]_T = 0$). We use Model 2 to study the size properties of the third test ($H_0^3 : [C_{Z1Z1}^{resid}, C_{Z2Z2}^{resid}]_T = 0$). Finally, we use Model 3 to study power properties of all three tests.

The upper panel Tables 5, 6, and 7 reports the size results while the lower panels shows the results for the power. We present the results for the two sampling frequencies ($\Delta_n = 1$ minute and

θ	AN 1.5	2	2.5	3	LIN 1.5	2	2.5	3
				Median Bias				
$\widehat{\gamma}_{Z1}$	-0.047	-0.025	-0.011	-0.003	-0.006	0.001	0.009	0.015
$\widehat{R}^{2,IV\text{-}FM}_{Z1}$	0.176	0.130	0.103	0.085	0.181	0.140	0.112	0.092
$\widehat{\rho}_{Z1,Z2}$	-0.288	-0.212	-0.163	-0.133	-0.249	-0.190	-0.146	-0.120
$\widehat{\rho}^{resid}_{Z1,Z2}$	-0.189	-0.113	-0.064	-0.034	-0.150	-0.091	-0.047	-0.021
				IQR				
$\widehat{\gamma}_{Z1}$	0.222	0.166	0.138	0.121	0.226	0.168	0.139	0.122
$\widehat{R}^{2,IV\text{-}FM}_{Z1}$	0.210	0.188	0.172	0.152	0.181	0.166	0.152	0.140
$\widehat{\rho}_{Z1,Z2}$	0.404	0.325	0.263	0.223	0.338	0.283	0.237	0.205
$\widehat{\rho}^{resid}_{Z1,Z2}$	0.456	0.384	0.315	0.272	0.388	0.337	0.285	0.250

Table 2: Finite sample properties of our estimators using 10 years of data sampled at 5 minutes. The true values are $\gamma_{Z1} = 0.450$, $R^{IV\text{-}FM}_{Z1} = 0.342$, $\rho_{Z1,Z2} = 0.523$, $\rho^{resid}_{Z1,Z2} = 0.424$.

θ	AN 1.5	2	2.5	3	LIN 1.5	2	2.5	3
				Median Bias				
$\widehat{\gamma}_{Z1}$	-0.022	-0.012	-0.003	0.004	-0.003	-0.000	0.006	0.012
$\widehat{R}^{IV\text{-}FM}_{Z1}$	0.107	0.091	0.073	0.056	0.113	0.095	0.075	0.058
$\widehat{\rho}_{Z1,Z2}$	-0.147	-0.104	-0.073	-0.048	-0.133	-0.097	-0.067	-0.042
$\widehat{\rho}^{resid}_{Z1,Z2}$	-0.135	-0.086	-0.058	-0.039	-0.119	-0.078	-0.052	-0.032
				IQR				
$\widehat{\gamma}_{Z1}$	0.156	0.112	0.088	0.075	0.157	0.112	0.088	0.075
$\widehat{R}^{IV\text{-}FM}_{Z1}$	0.201	0.146	0.118	0.100	0.184	0.138	0.113	0.096
$\widehat{\rho}_{Z1,Z2}$	0.340	0.238	0.184	0.150	0.309	0.226	0.177	0.145
$\widehat{\rho}^{resid}_{Z1,Z2}$	0.417	0.291	0.228	0.184	0.378	0.274	0.217	0.177

Table 3: Finite sample properties of our estimators using 10 years of data sampled at 1 minute. The true values are $\gamma_{Z1} = 0.450$, $R^{2,IV\text{-}FM}_{Z1} = 0.336$, $\rho_{Z1,Z2} = 0.514$, $\rho^{resid}_{Z1,Z2} = 0.408$.

θ	AN 1.5	AN 2	AN 2.5	AN 3	LIN 1.5	LIN 2	LIN 2.5	LIN 3
				Median Bias				
$\widehat{\gamma}_{Z1}$	-0.019	-0.011	-0.007	0.000	-0.001	-0.001	0.002	0.008
$\widehat{R}^{2,IV\text{-}FM}_{Z1}$	0.115	0.096	0.081	0.069	0.119	0.100	0.084	0.071
$\widehat{\rho}_{Z1,Z2}$	-0.168	-0.101	-0.064	-0.038	-0.149	-0.092	-0.057	-0.033
$\widehat{\rho}^{resid}_{Z1,Z2}$	-0.141	-0.079	-0.035	-0.007	-0.127	-0.067	-0.029	-0.001
				IQR				
$\widehat{\gamma}_{Z1}$	0.215	0.159	0.128	0.110	0.216	0.158	0.129	0.110
$\widehat{R}^{2,IV\text{-}FM}_{Z1}$	0.282	0.204	0.168	0.144	0.260	0.194	0.161	0.139
$\widehat{\rho}_{Z1,Z2}$	0.472	0.337	0.263	0.213	0.436	0.319	0.252	0.206
$\widehat{\rho}^{resid}_{Z1,Z2}$	0.541	0.412	0.324	0.266	0.510	0.391	0.311	0.256

Table 4: Finite sample properties of our estimators using 5 years of data sampled at 1 minute. The true values are $\gamma_{Z1} = 0.450$, $R^{2,IV\text{-}FM}_{Z1} = 0.35$, $\rho_{Z1,Z2} = 0.517$, $\rho^{resid}_{Z1,Z2} = 0.417$.

$\Delta_n = 5$ minutes) and the two type of tests (AN and LIN). We observe that the size of three tests are reasonably close to their nominal levels. The rejection probabilities under the alternatives are rather high, except when the data is sampled at 5 minutes frequency and the nominal level at 1%. We note that the tests based on LIN estimators have better testing power compared to those that build on AN estimators. Increasing the window length induces some size distortions but is very effective for power gain. Consistent with the asymptotic theory, the size of the three tests are closer to the nominal levels and the power is higher at the one minute sampling frequency. Clearly, the test of absence of dependence between IV and the market volatility has the best power, followed by the test of absence of dependence between the two IVs. This ranking is compatible with the notion that the finite sample properties of the tests deteriorate with the degree of latency embedded in each null hypothesis.

	$\Delta_n = 5$ minutes						$\Delta_n = 1$ minute					
	$\theta = 1.5$		$\theta = 2.0$		$\theta = 2.5$		$\theta = 1.5$		$\theta = 2.0$		$\theta = 2.5$	
Type of the test	AN	LIN	AN	LIN	AN	LIN	AN	LIN	AN	LIN	AN	LIN
						Panel A : Size Analysis-Model 1						
$\alpha = 10\%$	9.7	10.6	10.6	12.6	9.7	10.3	10.2	9.7	10.0	10.2	9.8	10.2
$\alpha = 5\%$	4.7	5.1	4.5	5.3	4.8	5.6	5.3	5.3	5.2	5.3	4.9	5.1
$\alpha = 1\%$	0.9	1.1	0.9	1.2	0.9	1.1	1.1	1.1	1.2	1.1	1.0	1.0
						Panel B : Power Analysis-Model 3						
$\alpha = 10\%$	20.5	31.5	35.7	48.3	53.3	65.8	33.9	41.0	65.6	71.6	88.0	91.2
$\alpha = 5\%$	11.9	21.0	23.9	35.76	40.6	53.4	22.3	29.5	52.8	59.8	79.6	84.4
$\alpha = 1\%$	3.3	6.9	8.7	15.6	18.4	28.6	8.9	12.4	28.6	34.5	57.4	64.1

Table 5: Size and Power of the test of absence of dependence between idiosyncratic volatilities for $T =$ 10 years.

	$\Delta_n = 5$ minutes						$\Delta_n = 1$ minute					
	$\theta = 1.5$		$\theta = 2.0$		$\theta = 2.5$		$\theta = 1.5$		$\theta = 2.0$		$\theta = 2.5$	
Type of test	AN	LIN	AN	LIN	AN	LIN	AN	LIN	AN	LIN	AN	LIN
						Panel A : Size Analysis-Model 1						
$\alpha = 10\%$	12.1	10.2	10.0	10.6	9.8	11.0	11.0	10.4	10.3	10.4	10.4	10.4
$\alpha = 5\%$	6.2	5.0	4.5	5.2	4.6	5.4	5.5	5.4	5.2	5.1	5.2	5.3
$\alpha = 1\%$	1.5	1.0	0.8	1.0	0.9	1.2	1.1	1.1	1.0	0.9	0.8	1.0
						Panel B : Power Analysis-Model 3						
$\alpha = 10\%$	60.0	69.0	84.0	88.3	94.6	96.1	91.1	93.3	99.2	99.4	100	100
$\alpha = 5\%$	47.7	57.2	75.0	81.0	89.6	92.6	84.9	88.2	98.2	98.6	100	100
$\alpha = 1\%$	24.1	32.3	52.2	60.1	73.7	78.9	67.7	72.0	93.0	94.5	99.2	99.4

Table 6: Size and Power of the test of absence of dependence between the idiosyncratic volatility and the market volatility for $T = 10$ years.

	$\Delta_n = 5$ minutes						$\Delta_n = 1$ minute					
	$\theta = 1.5$		$\theta = 2.0$		$\theta = 2.5$		$\theta = 1.5$		$\theta = 2.0$		$\theta = 2.5$	
Type of test	AN	LIN	AN	LIN	AN	LIN	AN	LIN	AN	LIN	AN	LIN
Panel A : Size Analysis-Model 2												
$\alpha = 10\%$	10.0	10.1	12.1	10.8	9.9	12.6	10.1	10.3	10.6	11.3	10.1	11.4
$\alpha = 5\%$	5.0	6.3	5.1	6.3	5.1	6.7	5.5	5.5	5.3	5.9	5.2	6.0
$\alpha = 1\%$	1.1	1.5	0.8	1.6	1.1	1.4	1.1	1.2	1.3	1.3	1.3	1.5
Panel B : Power Analysis-Model 3												
$\alpha = 10\%$	13.7	19.2	16.8	23.0	28.1	36.9	19.0	22.2	35.0	39.4	53.4	58.3
$\alpha = 5\%$	7.4	11.3	9.3	14.2	18.3	25.2	11.0	13.7	23.9	28.0	40.0	44.9
$\alpha = 1\%$	1.6	3.1	2.3	3.9	6.0	9.5	2.9	4.0	9.3	11.6	18.8	22.2

Table 7: Size and Power of the test of absence of dependence between residual IVs for $T = 10$ years.

6 Empirical Analysis

We apply our methods to study the cross-sectional dependence in IV using high frequency data. One of our main findings is that stocks' idiosyncratic volatilities co-move strongly with the market volatility. This is a quite surprising finding. It is of course well known that the total volatility of stocks moves with the market volatility. However, we stress that we find that the strong effect is still present when considering the idiosyncratic volatilities.

We use full record transaction prices from NYSE TAQ database for 30 constituents of the DJIA index over the time period 2003-2012, see Table 8. After removing the non-trading days, our sample contains 2517 days. The selected stocks were the constituents of the DJIA index in 2007. We also use the high-frequency data on nine industry Exchange-Traded Funds, ETFs (Consumer Discretionary, Consumer Staples, Energy, Financial, Health Care, Industrial, Materials, Technology, and Utilities), and the high-frequency size and value Fama-French factors, see Aït-Sahalia, Kalnina, and Xiu (2014). For each day, we consider data from the regular exchange opening hours from time stamped between 9:30 a.m. until 4 p.m. We clean the data following the procedure suggested by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008), remove the overnight returns and then sample at 5 minutes. This sparse sampling has been widely used in the literature because the effect of the microstructure noise and potential asynchronicity of the data is less important at this frequency, see also Liu, Patton, and Sheppard (2014).

The parameter choices for the estimators are as follows. Guided by our Monte Carlo results, we set the length of window to be approximately one week for the estimators in Section 3 (this corresponds to $\theta = 2.5$ where $k_n = \theta\Delta_n^{-1/2}$ is the number of observations in a window). The truncation threshold for all estimators is set as in the Monte Carlo study ($3\widehat{\sigma}_t\Delta_n^{0.49}$ where $\widehat{\sigma}_t^2$ is the bipower variation).

Figures 1 and 2 contain plots of the time series of the estimated R^2_{Yj} of the return factor model (R-FM) for each stock.[9] Each plot contains monthly R^2_{Yj} from two return factor models, CAPM and the Fama-French regression with market, size, and value factors. Figures 1 and 2 show that these time series of all stocks follow approximately the same trend with a considerable increase in the contribution around the crisis year 2008. Higher R^2_{Yj} indicates that the systematic risk is relatively more important, which is typical during crises. R^2_{Yj} is consistently higher in the Fama-French regression model compared to the CAPM regression model, albeit not by much. We proceed to investigate the dynamic properties of the panel of idiosyncratic volatilities.

We first investigate the dependence in the (total) idiosyncratic volatilities. Our panel has 435 pairs of stocks. For each pair of stocks, we compute the correlation between the IVs, $\rho_{Zi,Zj}$. All

[9] For the j^{th} stock, our analog of the coefficient of determination in the R-FM is $R^2_{Yj} = 1 - \frac{\int_0^T c_{ZjZj,t}dt}{\int_0^T c_{YjYj,t}dt}$. We estimate R^2_{Yj} using the general method of Jacod and Rosenbaum (2013). The resulting estimator of R^2_{Yj} requires a choice of a block size for the spot volatility estimation; we choose two hours in practice (the number of observations in a block, say l_n, has to satisfy $l_n^2\Delta_n \to 0$ and $l_n^3\Delta_n \to \infty$, so it is of smaller order than the number of observations k_n in our estimators of Section 3).

pairwise correlations are positive in our sample, and their average is 0.55. Figure 3 maps the network of dependency in the IV. We simultaneously test 435 hypotheses of no correlation, and Figure 3 connects only the assets, for which the null is rejected. Overall, Figure 3 shows that the cross-sectional dependence between the IVs is very strong.

Could missing factors in the R-FM provide an explanation? Omitted return factors in the R-FM are captured by the idiosyncratic returns, and can therefore induce correlation between the estimated IVs, provided these missing return factors have non-negligible volatility of volatility. To investigate this possibility, we consider the correlations between idiosyncratic returns, $\mathrm{Corr}(Z_i, Z_j)$.[10] Table 9 presents a summary of how estimates $\mathrm{Corr}(Z_i, Z_j)$ are related to the estimates of correlation in IVs, $\rho_{Zi,Zj}$. In particular, different rows in Table 9 display average values of $\widehat{\rho}_{Zi,Zj}$ among those pairs, for which $\widehat{Corr}(Z_i, Z_j)$ is below some threshold. For example, the last-but-one row in Table 9 indicates that there are 56 pairs of stocks with $\widehat{Corr}(Z_i, Z_j) < 0.01$, and among those stocks, the average correlation between IVs, $\rho_{Zi,Zj}$, is estimated to be 0.579. We observe that $\widehat{\rho}_{Zi,Zj}$ is virtually the same compared to pairs of stocks with high $\mathrm{Corr}(Z_i, Z_j)$. These results suggest that missing return factors cannot explain dependence in IVs for all considered stocks. This finding is in line with the empirical analysis of Herskovic, Kelly, Lustig, and Nieuwerburgh (2016) with daily and monthly returns.

To understand the source of the strong cross-sectional dependence in the IVs, we consider the Idiosyncratic Volatility Factor Model (IV-FM) of Section 2. We first use the market volatility as the only IV factor.[11] Table 10 reports the estimates of the IV loading ($\widehat{\gamma}_{Zi}$) and the R^2 of the IV-FM ($R^{2,IV\text{-}FM}_{Zi}$, see equation (9)). Table 10 uses two different definitions of IV, one defined with respect to CAPM, and a second IV defined with respect to Fama-French three factor model. For every stock, the estimated IV factor loading is positive, suggesting that the idiosyncratic volatility co-moves with the market volatility. Next, Figure 4 shows the implications for the cross-section of the one-factor IV-FM when the IV is defined with respect to CAPM. The average pairwise correlations between the residual IVs, $\widehat{\rho}_{Zi,Zj}$, decrease to 0.25. However, the market volatility cannot explain all cross-sectional dependence in residual IVs, as evidenced by the remaining links in Figure 4.

Finally, we consider an IV-FM with ten IV factors, market volatility and the volatilities of nine industry ETFs. Figure 5 shows the implications for the cross-section of this ten-factor IV-FM when the IV is defined with respect to CAPM. The average pairwise correlations between the residual

[10] Our measure of correlation between the idiosyncratic returns dZ_i and dZ_j is

$$\mathrm{Corr}(Z_i, Z_j) = \frac{\int_0^T C_{ZiZj,t}dt}{\sqrt{\int_0^T C_{ZiZi,t}dt}\sqrt{\int_0^T C_{ZjZj,t}dt}}, \quad i,j = 1,\ldots,d_S, \tag{27}$$

where $C_{ZiZj,t}$ is the spot covariation between Z_i and Z_j. Similarly to R^2_{Yj}, we estimate $\mathrm{Corr}(Z_i, Z_j)$ using the method of Jacod and Rosenbaum (2013).

[11] We also considered the volatility of size and value Fama-French factors. However, both these factors turned out to have very low volatility of volatility and therefore did not significantly change the results.

IVs, $\widehat{\rho}_{Zi,Zj}$, decrease further to 0.18. However, significant dependence between the residual IVs remains, as evidenced by the remaining links in Figure 4. Our results suggest that there is room for considering the construction of additional IV factors based on economic theory, for example, along the lines of the heterogeneous agents model of Herskovic, Kelly, Lustig, and Nieuwerburgh (2016).

Sector	Stock	Ticker
Financial	American International Group, Inc.	AIG
	American Express Company	AXP
	Citigroup Inc.	C
	JPMorgan Chase & Co.	JPM
Energy	Chevron Corp.	CVX
	Exxon Mobil Corp.	XOM
Consumer Staples	Coca Cola Company	KO
	Altria	MO
	The Procter & Gamble Company	PG
	Wal-Mart Stores	WMT
Industrials	Boeing Company	BA
	Caterpillar Inc.	CAT
	General Electric Company	GE
	Honeywell International Inc	HON
	3M Company	MMM
	United Technologies	UTX
Technology	Hewlett-Packard Company	HPQ
	International Bus. Machines	IBM
	Intel Corp.	INTC
	Microsoft Corporation	MSFT
Health Care	Johnson & Johnson	JNJ
	Merck & Co.	MRK
	Pfizer Inc.	PFE
Consumer Discretionary	The Walt Disney Company	DIS
	Home Depot Inc	HD
	McDonald's Corporation	MCD
Materials	Alcoa Inc.	AA
	E.I. du Pont de Nemours & Company	DD
Telecommunications Services	AT&T Inc.	T
	Verizon Communications Inc.	VZ

Table 8: The table lists the stocks used in the empirical application. They are the 30 constituents of DJIA in 2007. The first column provides the Global Industry Classification Standard (GICS) sectors, the second the names of the companies and the third their tickers.

$\lvert\widehat{\text{Corr}}(Z_i, Z_j)\rvert$	CAPM			FF3 Model		
	Pairs	Avg $\lvert\widehat{\text{Corr}}(Z_i, Z_j)\rvert$	Avg $\widehat{\rho}_{Zi,Zj}$	Pairs	Avg $\lvert\widehat{\text{Corr}}(Z_i, Z_j)\rvert$	Avg $\widehat{\rho}_{Zi,Zj}$
< 0.6	435	0.038	0.510	435	0.038	0.512
< 0.5	434	0.036	0.509	434	0.037	0.512
< 0.4	434	0.036	0.509	434	0.037	0.512
< 0.3	434	0.036	0.509	434	0.037	0.512
< 0.2	431	0.035	0.508	430	0.035	0.511
< 0.1	403	0.028	0.503	404	0.029	0.506
< 0.075	383	0.025	0.500	382	0.026	0.502
< 0.050	315	0.018	0.487	316	0.019	0.489
< 0.025	177	0.006	0.447	178	0.007	0.452
< 0.010	80	0.001	0.415	81	0.002	0.414
< 0.005	43	0.000	0.385	41	0.001	0.409

Table 9: Each row in this table describes the subset of pairs of stocks with $|\widehat{Corr(Z_i, Z_j)}|$ below a threshold in column one. The table considers two R-FMs: the left panel defines the IV with respect to CAPM, and the right panel defines the IV with respect to the three-factor Fama-French model. In both cases, the market volatility is the only IV factor. Each panel reports three quantities for the given subset of pairs: the number of pairs, average absolute pairwise correlation in idiosyncratic returns, and average pairwise correlation between IVs.

	CAPM			FF3 Model		
Stock	$\widehat{\gamma}_z$	$\widehat{R}_Z^{2,IV\text{-}FM}$	p-val	$\widehat{\gamma}_z$	$\widehat{R}_Z^{2,IV\text{-}FM}$	p-val
AIG	1.49	0.02	0.093	1.53	0.02	0.085
AXP	3.02	0.27	0.146	2.98	0.27	0.149
C	3.46	0.108	0.007	3.48	0.11	0.007
JPM	2.44	0.20	0.007	2.46	0.21	0.006
CVX	1.08	0.51	0.030	1.07	0.51	0.030
XOM	0.60	0.48	0.044	0.61	0.49	0.043
KO	0.33	0.58	0.012	0.33	0.58	0.011
MO	0.44	0.35	0.001	0.44	0.35	0.001
PG	0.43	0.63	0.001	0.43	0.63	0.002
WMT	0.45	0.58	0.006	0.45	0.56	0.008
BA	0.47	0.42	0.003	0.48	0.44	0.003
CAT	0.69	0.49	0.009	0.69	0.48	0.009
GE	1.14	0.26	0.003	1.15	0.26	0.002
HON	0.53	0.44	0.014	0.53	0.43	0.014
MMM	0.39	0.55	0.000	0.38	0.54	0.000
UTX	0.50	0.52	0.003	0.50	0.53	0.004
HPQ	0.65	0.33	0.004	0.66	0.34	0.004
IBM	0.35	0.48	0.011	0.35	0.47	0.012
INTC	0.46	0.46	0.003	0.46	0.46	0.003
MSFT	0.68	0.52	0.008	0.67	0.51	0.010
JNJ	0.41	0.68	0.007	0.40	0.67	0.007
MRK	0.54	0.32	0.001	0.54	0.32	0.001
PFE	0.43	0.34	0.002	0.43	0.34	0.001
DIS	0.57	0.48	0.001	0.58	0.49	0.001
HD	0.66	0.45	0.010	0.66	0.45	0.010
MCD	0.29	0.29	0.003	0.29	0.29	0.003
AA	3.03	0.41	0.019	3.04	0.42	0.018
DD	0.61	0.59	0.001	0.61	0.59	0.001
T	0.76	0.45	0.003	0.76	0.44	0.003
VZ	0.54	0.55	0.000	0.54	0.54	0.001

Table 10: Estimates of the IV factor loading ($\widehat{\gamma}_Z$, see equation (6)), and the contribution of the market volatility to the variation in the IVs ($\widehat{R}_Z^{2,IV\text{-}FM}$, see equation (9)). The table considers two R-FMs: the left panel defines the IV with respect to CAPM, and the right panel defines the IV with respect to the three-factor Fama-French model. In both cases, the market volatility is the only IV factor. P-val is the p-value of the test of the absence of dependence between the IV and the market volatility for a given individual stock, see equation (25).

References

Aït-Sahalia, Y., J. Fan, R. Laeven, D. C. Wang, and X. Yang (2013): "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 (2014): *High-Frequency Financial Econometrics*. Princeton University Press.

Aït-Sahalia, Y., I. Kalnina, and D. Xiu (2014): "The Idiosyncratic Volatility Puzzle: A Reassessment at High Frequency," Discussion paper, The University of Chicago.

Ait-Sahalia, Y., and D. Xiu (2015): "Principal Component Analysis of High Frequency Data," Discussion paper, Princeton University and University of Chicago.

——— (2016): "Using Principal Component Analysis to Estimate a High Dimensional Factor Model with High-Frequency Data," Discussion paper, Princeton University and University of Chicago.

Aldous, D. G., and G. K. Eagleson (1978): "On Mixing and Stability of Limit Theorems," *Annals of Probability*, 6, 325–331.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and J. Wu (2004): "Realized beta: Persistence and predictability," Discussion paper.

Bandi, F., and R. Renò (2012): "Time-Varying Leverage Effects," *Journal of Econometrics*, 12, 94–113.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008): "Designing Realized Kernels to Measure Ex-post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76, 1481–1536.

Barndorff-Nielsen, O. E., and N. Shephard (2004): "Econometric analysis of realised covariation: high frequency based covariance, regression and correlation in financial economics," *Econometrica*, 72, 885–925.

Bollerslev, T., and V. Todorov (2010): "Jumps and Betas: A New Framework for Disentangling and Estimating Systematic Risks," *Journal of Econometrics*, 157, 157, 220235.

Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu (2001): "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance*, 56(1), 143.

Chen, Z., and R. Petkova (2012): "Does idiosyncratic volatility proxy for risk exposure?," *Reviews of Financial Studies*, 25(9), 2745–2787.

Christoffersen, P., M. Fournier, and K. Jacobs (2015): "The Factor Structure in Equity Options," Discussion paper, University of Toronto, HEC Montreal, and University of Houston.

Corradi, V., and W. Distaso (2006): "Semi-Parametric Comparison of Stochastic Volatility Models using Realized Measures," *Review of Economic Studies*, 73, 635–667.

Duarte, J., A. Kamara, S. Siegel, and C. Sun (2014): "The Systematic Risk of Idiosyncratic Volatility," Discussion paper.

Fama, E. F., and K. R. French (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, 3–56.

Gourier, E. (2016): "Pricing of Idiosyncratic Equity and Variance Risks," Discussion paper.

Hansen, P. R., A. Lunde, and V. Voev (2014): "Realized Beta GARCH: a Multivariate GARCH Model with Realized Measures of Volatility," *Journal of Applied Econometrics*, 29, 774–799.

Herskovic, B., B. Kelly, H. Lustig, and S. V. Nieuwerburgh (2016): "he Common Factor in Idiosyncratic Volatility: Quantitative Asset Pricing Implications," *Journal of Financial Economics*, 119, 249283.

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.

Kalnina, I., and D. Xiu (2015): "Nonparametric Estimation of the Leverage Effect using Information from Derivatives Markets," Discussion paper, University of Chicago and University of Montreal.

Kanaya, S., and D. Kristensen (2015): "Estimation of Stochastic Volatility Models by Nonparametric Filtering," *Econometric Theory, forthcoming.*

Kristensen, D. (2010): "Nonparametric Filtering of the Realized Spot Volatility: A Kernel-Based Approach," *Econometric Theory*, 26.

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 (2013): "Inference Theory on Volatility Functional Dependencies," Discussion paper, Duke University.

——— (2014): "Jump Regressions," Discussion paper, Duke University.

Lintner, J. (1965): “Security Prices, Risk, and Maximal Gains from Diversification,” *Journal of Finance*, 20, 587615.

Liu, L., A. Patton, and K. Sheppard (2014): “Does Anything Beat 5-Minute RV? A Comparison of Realized Measures Across Multiple Asset Classes,” *Journal of Econometrics, forthcoming.*

Mancini, C. (2001): “Disentangling the jumps of the di usion in a geometric jumping Brownian motion,” *Giornale dellIstituto Italiano degli Attuari LXIV*, pp. 19–47.

Mykland, P. A., and L. Zhang (2006): “ANOVA for diffusions and Itô processes,” *Annals of Statistics*, 34(4), 19311963.

Pelger, M. (2015): “Large-dimensional factor modeling based on high-frequency observations,” Discussion paper, Stanford University.

Sharpe, W. F. (1964): “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk,” *Journal of Finance*, 19, 42542.

Vetter, M. (2012): “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.

Appendix

Appendix A contains the proofs, and Appendix B contains Tables and Figures.

A Proofs

Throughout, we denote by K a generic constant, which may change from line to line. When it depends on a parameter p we use the notation K_p instead. We let by convention $\sum_{i=a}^{a'} = 0$ when $a > a'$.

A.1 Proof of Theorem 1

We prove this theorem in three steps. For simplicity, in the first two steps we focus on the estimation of $[H(C), G(C)]_T$ with $H, G \in \mathcal{G}(p)$. The joint estimation is discussed in Step 3.
By a localization argument (See Lemma 4.4.9 of Jacod and Protter (2012)), there exists a π-integrable function J on E and a constant such that the stochastic processes in (18) and (19) satisfy

$$\|b\|, \|\widetilde{b}\|, \|c\|, \|\widetilde{c}\|, J \leq A, \|\delta(w,t,z)\|^r \leq J(z). \tag{28}$$

Setting $b'_t = b_t - \int \delta(t,z) 1_{\{\|\delta(t,z)\| \leq 1\}} \pi(dz)$ and $Y'_t = \int_0^t b'_s ds + \int_0^t \sigma_s dW_s$, we have

$$Y_t = Y_0 + Y'_t + \sum_{s \leq t} \Delta Y_s.$$

The local estimator of the spot variance of the unobservable process Y' is given by,

$$\widehat{C}_i'^n = \frac{1}{k_n \Delta_n} \sum_{u=0}^{k_n - 1} (\Delta_{i+u}^n Y')(\Delta_{i+u}^n Y)'^\top = (\widehat{C}_i'^{n,gh})_{1 \leq g,h \leq d}. \tag{29}$$

Note that no jump truncation in needed in the definition of $\widehat{C}_i'^n$ since the process Y' is continuous. Therefore, it is more convenient to work with $\widehat{C}_i'^n$ rather than $\widehat{C}_i^n$ (defined in (13)). Let $[\widehat{H(C), G(C)}]_T^{LIN'}$ and $[\widehat{H(C), G(C)}]_T^{AN'}$ be the infeasible estimators obtained by replacing $\widehat{C}_i^n$ by $\widehat{C}_i'^n$ in the definition of $[\widehat{H(C), G(C)}]_T^{LIN}$ and $[\widehat{H(C), G(C)}]_T^{AN}$.

Step1: Dealing with price jumps

We prove that, as long as $(8p-1)/4(4p-r) \leq \varpi < \frac{1}{2}$, we have

$$\Delta_n^{-1/4}\Big([\widehat{H(C), G(C)}]_T^{LIN} - [\widehat{H(C), G(C)}]_T^{LIN'}\Big) \xrightarrow{\mathbb{P}} 0 \quad \text{and} \quad \Delta_n^{-1/4}\Big([\widehat{H(C), G(C)}]_T^{AN} - [\widehat{H(C), G(C)}]_T^{AN'}\Big) \xrightarrow{\mathbb{P}} 0. \tag{30}$$

To show this result, let us define the functions

$$R(x,y) = \sum_{g,h,a,b=1}^{d} (\partial_{gh} H \partial_{ab} G)(x)(y^{gh} - x^{gh})(y^{ab} - x^{ab}), \ S(x,y) = (H(y) - H(x))(G(y) - G(x))$$
$$U(x) = \sum_{g,h,a,b=1}^{d} (\partial_{gh} H \partial_{ab} G)(x)(x^{ga} x^{hb} + x^{gb} x^{ha}),$$

for any $\mathbb{R}^d \times \mathbb{R}^d$ matrices x and y. The following decompositions hold,

$$[H(\widehat{C),G}(C)]_T^{AN} - [H(\widehat{C),G}(C)]_T^{AN'} = \frac{3}{2k_n} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} \Big[\big(S(\widehat{C}_i^n, \widehat{C}_{i+k_n}^n) - S(\widehat{C}_i'^n, \widehat{C}_{i+k_n}'^n)\big) - \frac{2}{k_n}\big(U(\widehat{C}_i^n) - U(\widehat{C}_i'^n)\big)\Big],$$

$$[H(\widehat{C),G}(C)]_T^{LIN} - [H(\widehat{C),G}(C)]_T^{LIN'} = \frac{3}{2k_n} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} \Big[\big(R(\widehat{C}_i^n, \widehat{C}_{i+k_n}^n) - R(\widehat{C}_i'^n, \widehat{C}_{i+k_n}'^n)\big) - \frac{2}{k_n}\big(U(\widehat{C}_i^n) - U(\widehat{C}_i'^n)\big)\Big].$$

By (3.11) in Jacod and Rosenbaum (2012), there exists a sequence of real numbers a_n converging to zero such that

$$\mathbb{E}(\|\widehat{C}_i^n - \widehat{C}_i'^n\|^q) \leq K_q a_n \Delta_n^{(2q-r)\varpi+1-q}, \text{ for any } q > 0. \tag{31}$$

Since H and $G \in \mathcal{G}(p)$, it is easy to see that the functions R and S are continuously differentiable and satisfy

$$\|\partial J(x,y)\| \leq K(1 + \|x\| + \|y\|)^{2p-1} \text{ for } 1 \leq g,h,a,b \leq d \text{ and } J \in \{S,R\}, \tag{32}$$

$$\|\partial U(x)\| \leq K(1 + \|x\|)^{2p-1}, \tag{33}$$

where ∂J (respectively, ∂U) is a vector that collects the first order partial derivatives of the function J (respectively, U) with respect to all the elements of (x,y) (resp x). By Taylor expansion, Jensen inequality, (32) and (33), it can be shown that, for $J \in \{S,R\}$,

$$\begin{aligned}
|J(\widehat{C}_i^n, \widehat{C}_{i+k_n}^n) - J(\widehat{C}_i'^n, \widehat{C}_{i+k_n}'^n)| &\leq K(1 + \|\widehat{C}_i'^n\|^{2p-1} + \|\widehat{C}_{i+k_n}'^n\|^{2p-1})(\|\widehat{C}_i^n - \widehat{C}_i'^n\| + \|\widehat{C}_{i+k_n}^n - \widehat{C}_{i+k_n}'^n\|) \\
&\quad + K\|\widehat{C}_i^n - \widehat{C}_i'^n\|^{2p} + K\|\widehat{C}_{i+k_n}^n - \widehat{C}_{i+k_n}'^n\|^{2p} \text{ and} \\
|U(\widehat{C}_i^n) - U(\widehat{C}_i'^n)| &\leq K(1 + \|\widehat{C}_i'^n\|^{2p-1})(\|\widehat{C}_i^n - \widehat{C}_i'^n\|) + K\|\widehat{C}_i^n - \widehat{C}_i'^n\|^{2p}.
\end{aligned}$$

By (3.20) in Jacod and Rosenbaum (2012), we have $\mathbb{E}(\|\widehat{C}_i'^n\|^v) \leq K_v$, for any $v \geq 0$. Hence by Hölder inequality, for $\epsilon > 0$ fixed,

$$\begin{aligned}
\mathbb{E}(\|\widehat{C}_i'^n\|^{2p-2}\|\widehat{C}_i^n - \widehat{C}_i'^n\|) &\leq \Big(\mathbb{E}(\|\widehat{C}_i^n - \widehat{C}_i'^n\|^{(1+\epsilon)})\Big)^{1/1+\epsilon}\Big(\mathbb{E}(\|\widehat{C}_i'^n\|^{(2p-2)(1+\epsilon)/\epsilon})\Big)^{\epsilon/1+\epsilon} \\
&\leq K_p\Big(\mathbb{E}(\|\widehat{C}_i^n - \widehat{C}_i'^n\|^{(1+\epsilon)})\Big)^{1/1+\epsilon} \\
&\leq K_p a_n \Delta_n^{(2-\frac{1}{1+\epsilon})\varpi + \frac{1}{1+\epsilon} - 1}
\end{aligned}$$

Using the above result and (31), it easy to see that for (30) to hold, the following conditions are sufficient:

$$(2 - \frac{r}{1+\epsilon})\varpi + \frac{1}{1+\epsilon} - 1 - \frac{3}{4} \geq 0, \quad (4p - r)\varpi + 1 - 2p - \frac{3}{4} \geq 0, \quad \text{and} \quad (2-r)\varpi + -\frac{3}{4} \geq 0.$$

Using the fact that $0 < \varpi < \frac{1}{2}$, and taking ϵ sufficiently close to zero, we can see that (30) holds if $(8p-1)/4(4p-r) \leq \varpi < \frac{1}{2}$, which completes the proof.

Step 2 : First approximation for the estimators

Taking advantage of Step 1, it is enough to derive the asymptotic distributions of $[H(\widehat{C),G}(C)]_T^{LIN\prime}$ and $[H(\widehat{C),G}(C)]_T^{AN\prime}$. We show that the two estimators $[H(\widehat{C),G}(C)]_T^{LIN'}$ and $[H(\widehat{C),G}(C)]_T^{AN'}$ can be approximated by a certain quantity with an error of approximation of order smaller than $\Delta_n^{-1/4}$. To see this, we set

$$[H(\widehat{C),G}(C)]_T^A = \frac{3}{2k_n}\sum_{g,h,a,b=1}^{d}\sum_{i=1}^{[T/\Delta_n]-2k_n+1}\Bigg(\Big(\partial_{gh}H\partial_{ab}G\Big)(C_i^n)\Big[(\widehat{C}'^{n,gh}_{i+k_n}-\widehat{C}'^{n,gh}_i)(\widehat{C}'^{n,ab}_{i+k_n}-\widehat{C}'^{n,ab}_i)$$
$$-\frac{2}{k_n}(\widehat{C}'^{n,ga}_i\widehat{C}'^{n,hb}_i+\widehat{C}'^{n,gb}_i\widehat{C}'^{n,ha}_i)\Big]\Bigg),$$

with $C_i^n = C_{(i-1)\Delta_n}$ and the superscript A being a short for the word "approximate". For notational simplicity, we do not index the above quantity by a prime although it depends on $\widehat{C}'^n_i$ instead $\widehat{C}^n_i$. We aim to prove that

$$\Delta_n^{-1/4}\Big([H(\widehat{C),G}(C)]_T^{LIN'}-[H(\widehat{C),G}(C)]_T^{A}\Big)\xrightarrow{\mathbb{P}}0 \quad\text{and}\quad \Delta_n^{-1/4}\Big([H(\widehat{C),G}(C)]_T^{AN'}-[H(\widehat{C),G}(C)]_T^{A}\Big)\xrightarrow{\mathbb{P}}0. \tag{34}$$

To prove (34), we introduce some new notation. Following Jacod and Rosenbaum (2012), we define

$$\alpha_i^n = (\Delta_i^n Y')(\Delta_i^n Y')^\top - C_i^n\Delta_n,\ \ \nu_i^n = \widehat{C}'^n_i - C_i^n,\ \ \text{and}\ \ \lambda_i^n = \widehat{C}'^n_{i+k_n}-\widehat{C}'^n_i, \tag{35}$$

which satisfy

$$\nu_i^n = \frac{1}{k_n\Delta_n}\sum_{j=0}^{k_n-1}(\alpha_{i+j}^n + (C_{i+j}^n - C_i^n)\Delta_n)\ \ \text{and}\ \ \lambda_i^n = \nu_{i+k_n}-\nu_i^n+\Delta_n(C_{i+k_n}^n - C_i^n). \tag{36}$$

We have

$$[H(\widehat{C),G}(C)]_T^{LIN'} - [H(\widehat{C),G}(C)]_T^{A} = \frac{3}{2k_n}\sum_{g,h,a,b=1}^{d}\sum_{i=1}^{[T/\Delta_n]-2k_n+1}\psi_i^n(g,h,a,b),$$
$$[H(\widehat{C),G}(C)]_T^{AN'} - [H(\widehat{C),G}(C)]_T^{A} = \frac{3}{2k_n}\sum_{i=1}^{[T/\Delta_n]-2k_n+1}\Big(\chi_i^n - \sum_{g,h,a,b=1}^{d}(\partial_{gh}H\partial_{ab}G)(C_i^n)\lambda_i^{n,gh}\lambda_i^{n,ab}\Big),$$

with

$$\psi_i^n(g,h,a,b) = \Big(\big(\partial_{gh}H\partial_{ab}G\big)(\widehat{C}'^n_i) - \big(\partial_{gh}H\partial_{ab}G\big)(C_i^n)\Big)\lambda_i^{n,gh}\lambda_i^{n,ab},$$
$$\chi_i^n = \Big(H(\widehat{C}'^n_{i+k_n}) - H(\widehat{C}'^n_i)\Big)\Big(G(\widehat{C}'^n_{i+k_n}) - G(\widehat{C}'^n_i)\Big).$$

By Taylor expansion, we have

$$\big(\partial_{gh}S\partial_{ab}G\big)(\widehat{C}'^n_i) - \big(\partial_{gh}S\partial_{ab}G\big)(C_i^n) = \sum_{x,y=1}^{d}\Big(\partial^2_{xy,gh}S\partial_{ab}G + \partial^2_{xy,ab}G\partial_{gh}S\Big)(C_i^n)\nu_i^{n,xy}$$
$$+\frac{1}{2}\sum_{j,k,x,y=1}^{d}\Big(\partial^3_{jk,xy,gh}S\partial_{ab}G + \partial^2_{xy,gh}S\partial^2_{jk,ab}G + \partial^3_{jk,xy,ab}G\partial_{gh}S + \partial^2_{xy,ab}G\partial^2_{jk,gh}S\Big)(\widetilde{c}_i^n)\nu_i^{n,xy}\nu_i^{n,jk}$$

and

$$S(\widehat{C}'^n_{i+k_n}) - S(\widehat{C}'^n_i) = \sum_{gh}\partial_{gh}S(C_i^n)\lambda_i^{n,gh} + \sum_{j,k,g,h}\partial^2_{jk,gh}S(C_i^n)\lambda_i^{n,gh}\nu_i^{n,jk} + \frac{1}{2}\sum_{x,y,g,h}\partial^2_{xy,gh}S(C_i^n)\lambda_i^{n,gh}\lambda_i^{n,xy}$$
$$+\frac{1}{2}\sum_{x,y,j,k,g,h}\partial^3_{xy,jk,gh}S(CC_i^{n,S})\lambda_i^{n,gh}\nu_i^{n,xy}\nu_i^{n,jk} + \frac{1}{6}\sum_{j,k,x,y,g,h}\partial^3_{jk,xy,gh}S(C_i^{n,S})\lambda_i^{n,jk}\lambda_i^{n,gh}\lambda_i^{n,xy},$$

for $S \in \{H, G\}$, $\widetilde{c}_i^n = \pi C_i^n + (1-\pi)\widehat{C}_i'^n$, $C_i^{n,S} = \pi_S \widehat{C}_i'^n + (1-\pi_S)\widehat{C}_{i+k_n}'^n$, $CC_i^{n,S} = \mu_S C_i^n + (1-\mu_S)\widehat{C}_i'^n$ for $\pi, \pi_H, \mu_H, \pi_G, \mu_G \in [0,1]$. Although $\widetilde{c}_i^n$ and π depend on g, h, a, and b, we do not emphasize this in our notation to simplify the exposition.

We remind the reader some well-known results. For any continuous Itô process Z_t, we have

$$\mathbb{E}\Big(\sup_{w\in[0,s]} \Big\|Z_{t+w} - Z_t\Big\|^q \Big|\mathcal{F}_t\Big) \le K_q s^{q/2}, \text{ and } \Big\|\mathbb{E}\Big(Z_{t+s} - Z_t\Big)\Big|\mathcal{F}_t\Big\| \le Ks. \tag{37}$$

Set $\mathcal{F}_i^n = \mathcal{F}_{(i-1)\Delta_n}$. By (4.10) in Jacod and Rosenbaum (2013) we have,

$$\mathbb{E}\Big(\Big\|\alpha_i^n\Big\|^q\Big|\mathcal{F}_i^n\Big) \le K_q \Delta_n^q \text{ for all } q \ge 0 \text{ and } \mathbb{E}\Big(\Big|\sum_{j=0}^{k_n-1} \alpha_{i+j}^n\Big|^q\Big|\mathcal{F}_i^n\Big) \le K_q \Delta_n^q k_n^{q/2} \text{ whenever } q \ge 2. \tag{38}$$

Combining (46), (44), (45) with $Z = c$ and the Hölder inequality yields for $q \ge 2$,

$$\mathbb{E}\Big(\Big\|\nu_i^n\Big\|^q\Big|\mathcal{F}_i^n\Big) \le K_q \Delta^{q/4}, \text{ and } \mathbb{E}\Big(\Big\|\lambda_i^n\Big\|^q\Big|\mathcal{F}_i^n\Big) \le K_q \Delta^{q/4}. \tag{39}$$

The bound in the first equation of (47) is tighter than that in (4.11) of Jacod and Rosenbaum (2012) due to the absence of volatility jumps. This tighter bound will be useful later for deriving the asymptotic distribution for the approximate estimator (**Step** 3). By the boundedness of C_t and the polynomial growth assumption, we have

$$\Big|(\partial^3_{jk,xy,ab} G \partial_{gh} H + \partial^2_{xy,gh} H \partial^2_{jk,ab} G)(\widetilde{c}_i^n)\nu_i^{n,xy}\nu_i^{n,jk}\lambda_i^{n,gh}\lambda_i^{n,ab}\Big| \le K(1 + \|\widetilde{c}_i^n\|)^{2(p-2)}\|\nu_i^n\|^2\|\lambda_i^n\|^2.$$

Recalling $\widetilde{c}_i^n = \pi C_i^n + (1-\pi)\widehat{C}_i'^n$ and using the convexity of the function $x^{2(p-2)}$, we can refine the last inequality as follows:

$$\Big|(\partial^3_{jk,xy,ab} G \partial_{gh} H + \partial^2_{xy,gh} H \partial^2_{jk,ab} G)(\widetilde{c}_i^n)\nu_i^{n,xy}\nu_i^{n,jk}\lambda_i^{n,gh}\lambda_i^{n,ab}\Big| \le K\big(1 + \|\nu_i^n\|^{2(p-2)}\big)\|\nu_i^n\|^2\|\lambda_i^n\|^2. \tag{40}$$

By Taylor expansion, the polynomial growth assumption and using similar idea as for (40), we have

$$\chi_i^n - \sum_{g,h,a,b} (\partial_{gh} H \partial_{ab} G)(C_i^n)\lambda_i^{n,gh}\lambda_i^{n,ab} = \sum_{g,h,a,b,j,k} (\partial_{gh} H \partial^2_{jk,xy} G + \partial_{gh} G \partial^2_{jk,xy} H)(C_i^n)(\lambda_i^{n,gh} + \frac{1}{2}\nu_i^{n,gh})\lambda_i^{n,ab}\lambda_i^{n,jk} + \varphi_i^n$$

$$\sum_{g,h,a,b} \big(\partial_{gh} H \partial_{ab} G\big)(\widehat{C}_i'^n) - \big(\partial_{gh} H \partial_{ab} G\big)(C_i^n) = \sum_{g,h,a,b,x,y} (\partial_{gh} H \partial^2_{ab,xy} G + \partial_{ab} G \partial^2_{gh,xy} G)(C_i^n)(\nu_i^{n,xy})\lambda_i^{n,gh}\lambda_i^{n,ab} + \delta_i^n$$

with $\mathbb{E}(|\varphi_i^n||\mathcal{F}_i^n) \le K\Delta_n$ and $\mathbb{E}(|\delta_i^n||\mathcal{F}_i^n) \le K\Delta_n$ which follow by the Cauchy-Schwartz inequality together with (47). Given that $k_n = \theta(\Delta_n)^{-1/2}$, a direct implication of the previous inequalities is

$$\frac{3\Delta_n^{-1/4}}{2k_n} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} \varphi_i^n \xrightarrow{\mathbb{P}} 0 \quad \text{and} \quad \frac{3\Delta_n^{-1/4}}{2k_n} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} \delta_i^n \xrightarrow{\mathbb{P}} 0.$$

Therefore, in order to prove the two claims in (34), it suffices to show

$$\frac{3\Delta_n^{-1/4}}{2k_n} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} \sum_{g,h,a,b,j,k} (\partial_{gh} H \partial^2_{jk,ab} G + \partial_{gh} H \partial^2_{jk,ab} G)(C_i^n)\lambda_i^{n,gh}\lambda_i^{n,ab}\lambda_i^{n,jk} \xrightarrow{\mathbb{P}} 0, \tag{41}$$

$$\frac{3\Delta_n^{-1/4}}{2k_n}\sum_{i=1}^{[T/\Delta_n]-2k_n+1}\sum_{g,h,a,b,j,k}(\partial_{gh}H\partial^2_{jk,ab}G+\partial_{gh}H\partial^2_{jk,ab}G)(C_i^n)\nu_i^{n,gh}\lambda_i^{n,ab}\lambda_i^{n,jk}\xrightarrow{\mathbb{P}}0. \quad (42)$$

For any càdlàg bounded process Z, we set

$$\eta_{t,s}(Z)=\sqrt{\mathbb{E}\Big(\sup_{0<u\le s}\|Z_{t+u}-Z_t\|^2|\mathcal{F}_i^n\Big)},$$
$$\eta_{i,j}^n(Z)=\sqrt{\mathbb{E}\Big(\sup_{0\le u\le j\Delta_n}\|Z_{(i-1)\Delta_n+u}-Z_{(i-1)\Delta_n}\|^2|\mathcal{F}_i^n\Big)}.$$

In order to prove (41) and (42), we introduce the following lemmas.

Lemma 1. *For any càdlàg bounded process Z, for all $t,s>0$, $j,k\ge 0$, set $\eta_{t,s}=\eta_{t,s}(Z)$. Then,*

$$\Delta_n\mathbb{E}\Big(\sum_{i=1}^{[t/\Delta_n]}\eta_{i,k_n}\Big)\longrightarrow 0,\qquad \Delta_n\mathbb{E}\Big(\sum_{i=1}^{[t/\Delta_n]}\eta_{i,2k_n}\Big)\longrightarrow 0,$$
$$\mathbb{E}\Big(\eta_{i+j,k}|\mathcal{F}_i^n\Big)\le\eta_{i,j+k}\ \ \text{and}\qquad \Delta_n\mathbb{E}\Big(\sum_{i=1}^{[t/\Delta_n]}\eta_{i,4k_n}\Big)\longrightarrow 0.$$

The first three claims of Lemma 6 are proved in Jacod and Rosenbaum (2012). The last result can be proved similarly to the first two.

Lemma 2. *Let Z be a continuous Itô process with drift b_t^Z and spot variance process C_t^Z, and set $\eta_{t,s}=\eta_{t,s}(b^Z,c^Z)$. Then, the following bounds hold:*

$$|\mathbb{E}(Z_t|\mathcal{F}_0)-tb_0^Z|\le Kt\eta_{0,t}$$
$$|\mathbb{E}(Z_t^jZ_t^k-tC_0^{Z,jk}|\mathcal{F}_0)|\le Kt^{3/2}(\sqrt{\Delta_n}+\eta_{0,t})$$
$$|\mathbb{E}\big((Z_t^jZ_t^k-tC_0^{Z,jk})(C_t^{Z,lm}-C_0^{Z,lm})|\mathcal{F}_0)|\le Kt^2$$
$$|\mathbb{E}(Z_t^jZ_t^kZ_t^lZ_t^m|\mathcal{F}_0)-\Delta_n^2(C_0^{Z,jk}C_0^{Z,lm}+C_0^{Z,jl}C_0^{Z,km}+C_0^{Z,jm}C_0^{Z,kl})|\le Kt^{5/2}$$
$$|\mathbb{E}(Z_t^jZ_t^kZ_t^l|\mathcal{F}_0)|\le Kt^2$$
$$|\mathbb{E}(\prod_{l=1}^{6}Z_t^{j_l}|\mathcal{F}_0)-\frac{\Delta_n^3}{6}\sum_{l<l'}\sum_{k<k'}\sum_{m<m'}C_0^{Z,j_lj_{l'}}C_0^{Z,j_kj_{k'}}C_0^{Z,j_mj_{m'}}|\le Kt^{7/2}$$

The first four claims of Lemma 7 are parts of Lemma 4.1 in Jacod and Rosenbaum (2012). The two remaining statements can be shown similarly.

Lemma 3. *Let ζ_i^n be a r-dimensional $\mathcal{F}_i^n$ measurable process satisfying $\|\mathbb{E}(\zeta_i^n|\mathcal{F}_{i-1}^n)\|\le L'$ and $\mathbb{E}\Big(\|\zeta_i^n\|^q\Big|\mathcal{F}_{i-1}^n\Big)\le L_q$. Also, let φ_i^n be a real-valued $\mathcal{F}_i^n$-measurable process with $\mathbb{E}\Big(\|\varphi_{i+j-1}^n\|^q\Big|\mathcal{F}_{i-1}^n\Big)\le L^q$ for $q\ge 2$ and $1\le j\le 2k_n-1$. Then, we have*

$$\mathbb{E}\Bigg(\Big\|\sum_{j=1}^{2k_n-1}\varphi_{i+j-1}^n\zeta_{i+j}^n\Big\|^q\Bigg|\mathcal{F}_{i-1}^n\Bigg)\le K_qL^q(L_qk_n^{q/2}+L'^qk_n^q).$$

Proof of Lemma 5

Set

$$\xi_i^n=\varphi_{i-1}^n\zeta_i^n,\qquad \xi_i'^n=\mathbb{E}(\xi_i|\mathcal{F}_{i-1}^n)=\mathbb{E}(\varphi_{i-1}^n\zeta_i^n|\mathcal{F}_{i-1}^n)=\varphi_{i-1}^n\mathbb{E}(\zeta_i^n|\mathcal{F}_{i-1}^n),\ \text{ and }\ \xi_i''^n=\xi_i^n-\xi_i'^n.$$

Given that $\|\mathbb{E}(\zeta_i^n|\mathcal{F}_{i-1}^n)\| \leq L'$, we have $\|\xi_i^{'n}\| \leq L'|\varphi_{i-1}^n|$. By the convexity of the function x^q, which holds for $q \geq 2$, we have

$$\|\sum_{j=1}^{2k_n-1} \xi_{i+j}^n\|^q \leq K\Big(\|\sum_{j=1}^{2k_n-1} \xi_{i+j}^{'n}\|^q + \|\sum_{j=1}^{2k_n-1} \xi_{i+j}^{''n}\|^q\Big).$$

Therefore, on the one hand we have

$$\|\sum_{j=1}^{2k_n-1} \xi_{i+j}^{'n}\|^q \leq Kk_n^{q-1} \sum_{j=1}^{2k_n-1} \|\xi_{i+j}^{'n}\|^q \leq Kk_n^{q-1}L'^q \sum_{j=1}^{2k_n-1} |\varphi_{i+j-1}^n|^q,$$

which by $\mathbb{E}\Big(\|\varphi_{i+j-1}^n\|^q\Big|\mathcal{F}_{i-1}^n\Big) \leq L^q$, satisfies

$$\mathbb{E}(\|\sum_{j=1}^{2k_n-1} \xi_{i+j}^{'n}\|^q|\mathcal{F}_{i-1}^n) \leq KL'^q k_n^{q-1} \sum_{j=1}^{2k_n-1} \mathbb{E}(|\varphi_{i+j-1}^n|^q|\mathcal{F}_{i-1}^n) \leq KL'^q k_n^q L^q.$$

On the other hand, we have $\mathbb{E}(\|\xi_{i+j}^{''n}\|^q|\mathcal{F}_{i-1}^n) \leq \mathbb{E}(\|\xi_{i+j}^n\|^q|\mathcal{F}_{i-1}^n) \leq L_q L^q$ and $\mathbb{E}(\xi_{i+j}^{''n}|\mathcal{F}_{i-1}^n) = 0$, where the first inequality is a consequence of $\mathbb{E}(\|\xi_{i+j}^{'n}\|^q|\mathcal{F}_{i-1}^n) \leq \mathbb{E}(\|\xi_{i+j}^n\|^q|\mathcal{F}_{i-1}^n) \leq L_q L^q$, which follows by the Jensen inequality and the law of iterated expectation. Hence, by Lemma B.2 of Aït-Sahalia and Jacod (2014) we have

$$\mathbb{E}(\|\sum_{j=1}^{2k_n-1} \xi_{i+j}^{''n}\|^q|\mathcal{F}_{i-1}^n) \leq K_q L^q L_q k_n^{q/2}.$$

To see the latter, we first prove that the required condition $\mathbb{E}(\|\xi_i^n\|^q|\mathcal{F}_{i-1}^n) \leq L_q L^q)$ in the Lemma B.2 of Aït-Sahalia and Jacod (2014) can be replaced by $\mathbb{E}(\|\xi_{i+j}^n\|^q|\mathcal{F}_{i-1}^n) \leq L_q L^q)$ for $1 \leq j \leq 2k_n - 1$ without altering the result.

Lemma 4. *We have:*

$$\begin{aligned}
&\Big|\mathbb{E}(\lambda_i^{n,jk}\lambda_i^{n,lm}\lambda_{i+2k_n}^{n,gh}\lambda_{i+2k_n}^{n,ab}|\mathcal{F}_i^n) - \frac{4}{k_n^2}(C_i^{n,ga}C_i^{n,hb} + C_i^{n,gb}C_i^{n,ha})(C_i^{n,jl}C_i^{n,km} + C_i^{n,jm}C_i^{n,kl}) \\
&- \frac{4\Delta_n}{3}(C_i^{n,jl}C_i^{n,km} + C_i^{n,jm}C_i^{n,kl})\overline{C}_i^{n,gh,ab} - \frac{4\Delta_n}{3}(C_i^{n,ga}C_i^{n,hb} - C_i^{n,gb}C_i^{n,ha})\overline{C}_i^{n,jk,lm} \\
&- \frac{4(k_n\Delta_n)^2}{9}\overline{C}_i^{n,gh,ab}\overline{C}_i^{n,jk,lm}\Big| \leq K\Delta_n(\Delta_n^{1/8} + \eta_{i,4k_n}^n).
\end{aligned}$$

Throughout, we use the expression "successive conditioning" to refer to the following equalities,

$$\begin{aligned}
&x_1y_1 - x_0y_0 = x_0(y_1 - y_0) + y_0(x_1 - x_0) + (x_1 - x_0)(y_1 - y_0), \\
&x_1y_1z_1 - x_0y_0z_0 = x_0y_0(z_1 - z_0) + x_0z_0(y_1 - y_0) + y_0z_0(x_1 - x_0) + x_0(y_0 - y_1)(z_0 - z_1) \\
&\qquad + y_0(x_0 - x_1)(z_0 - z_1) + z_0(x_0 - x_1)(y_0 - y_1) + (x_1 - x_0)(y_1 - y_0)(z_1 - z_0),
\end{aligned}$$

which hold for any real numbers x_0, y_0, z_0, x_1, y_1, and z_1.

Proof of Lemma 4

To prove Lemma 4, we first note that $\lambda_i^{n,jk}\lambda_i^{n,lm}$ is $\mathcal{F}_{i+2k_n}^n$-measurable. Then, by the law of iterated expectations, we have

$$\mathbb{E}\Big(\lambda_i^{n,jk}\lambda_i^{n,lm}\lambda_{i+2k_n}^{n,gh}\lambda_{i+2k_n}^{n,ab}|\mathcal{F}_i^n\Big) = \mathbb{E}\Big(\lambda_i^{n,jk}\lambda_i^{n,lm}\mathbb{E}(\lambda_{i+2k_n}^{n,gh}\lambda_{i+2k_n}^{n,ab}|\mathcal{F}_{i+2k_n}^n)|\mathcal{F}_i^n\Big).$$

By equation (3.27) in Jacod and Rosenbaum (2012), we have

$$|\mathbb{E}(\lambda_{i+2k_n}^{n,gh}\lambda_{i+2k_n}^{n,ab}|\mathcal{F}_{i+2k_n}^n) - \frac{2}{k_n}(C_{i+2k_n}^{n,ga}C_{i+2k_n}^{n,hb} + C_{i+2k_n}^{n,gb}C_{i+2k_n}^{n,ha}) - \frac{2k_n\Delta_n}{3}\overline{C}_{i+2k_n}^{n,gh,ab}| \le K\sqrt{\Delta_n}(\Delta_n^{1/8} + \eta_{i+2k_n,2k_n}^n),$$
$$|\mathbb{E}(\lambda_i^{n,jk}\lambda_i^{n,lm}|\mathcal{F}_i^n) - \frac{2}{k_n}(C_i^{n,jl}C_i^{n,km} + C_i^{n,jm}C_i^{n,kl}) - \frac{2k_n\Delta_n}{3}\overline{C}_i^{n,jk,lm}| \le K\sqrt{\Delta_n}(\Delta_n^{1/8} + \eta_{i,2k_n}^n).$$

Also,

$$\begin{aligned}
&|\mathbb{E}\Big(\lambda_i^{n,jk}\lambda_i^{n,lm}\Big[\mathbb{E}(\lambda_{i+2k_n}^{n,gh}\lambda_{i+2k_n}^{n,ab}\Big|\mathcal{F}_{i+2k_n}^n) - \frac{2}{k_n}(C_{i+2k_n}^{n,ga}C_{i+2k_n}^{n,hb} + C_{i+2k_n}^{n,gb}C_{i+2k_n}^{n,ha}) - \frac{2k_n\Delta_n}{3}\overline{C}_{i+2k_n}^{n,gh,ab}\Big]\Big|\mathcal{F}_i^n\Big)| \\
&\le \sqrt{\Delta_n}\mathbb{E}(|\lambda_i^{n,jk}||\lambda_i^{n,lm}|(\Delta_n^{1/8} + \eta_{i+2k_n,2k_n}^n)|\Big|\mathcal{F}_i^n) \le K\sqrt{\Delta_n}\Delta_n^{1/8}\mathbb{E}(|\lambda_i^{n,jk}||\lambda_i^{n,lm}|\Big|\mathcal{F}_i^n) \\
&+ K\sqrt{\Delta_n}\mathbb{E}(|\lambda_i^{n,jk}||\lambda_i^{n,lm}|\eta_{i+2k_n,2k_n}^n|\Big|\mathcal{F}_i^n) \le K\Delta_n(\Delta_n^{1/8} + \eta_{i,4k_n}^n),
\end{aligned}$$

where the last inequality follows from Lemma 6. Using (45) successively with $Z = c$ and $Z = \overline{C}$ (recall that the latter holds under Assumption 2), together with the successive conditioning, we have

$$\begin{aligned}
&|\mathbb{E}\Big(\lambda_i^{n,jk}\lambda_i^{n,lm}\Big[\frac{2}{k_n}(C_{i+2k_n}^{n,ga}C_{i+2k_n}^{n,hb} + C_{i+2k_n}^{n,gb}C_{i+2k_n}^{n,ha}) + \frac{2k_n\Delta_n}{3}\overline{C}_{i+2k_n}^{n,gh,ab} - \frac{2}{k_n}(C_i^{n,ga}C_i^{n,hb} + C_i^{n,gb}C_i^{n,ha}) \\
&- \frac{2k_n\Delta_n}{3}\overline{C}_i^{n,gh,ab}\Big]\Big|\mathcal{F}_i^n\Big)| \le K\Delta_n\Delta_n^{1/4}, \\
&|\mathbb{E}\Big(\lambda_i^{n,jk}\lambda_i^{n,lm}\Big[\frac{2}{k_n}(C_i^{n,ga}C_i^{n,hb} + C_i^{n,gb}C_i^{n,ha}) + \frac{2k_n\Delta_n}{3}\overline{C}_i^{n,gh,ab}\Big] - \Big[\frac{2}{k_n}(C_i^{n,jl}C_i^{n,km} + C_i^{n,jm}C_i^{n,kl}) + \frac{2k_n\Delta_n}{3}\overline{C}_i^{n,jk,lm}\Big] \\
&\times \Big[\frac{2}{k_n}(C_i^{n,ga}C_i^{n,hb} + C_i^{n,gb}C_i^{n,ha}) + \frac{2k_n\Delta_n}{3}\overline{C}_i^{n,gh,ab}\Big]\Big|\mathcal{F}_i^n\Big)| \le K\Delta_n(\Delta_n^{1/8} + \eta_{i,2k_n}^n).
\end{aligned}$$

The last inequality yields the result.

Lemma 5. *Let ζ_i^n be a r-dimensional $\mathcal{F}_i^n$-measurable process satisfying $\|\mathbb{E}(\zeta_i^n|\mathcal{F}_{i-1}^n)\| \le L'$ and $\mathbb{E}\Big(\|\zeta_i^n\|^q\Big|\mathcal{F}_{i-1}^n\Big) \le L_q$. Also, let φ_i^n be a real-valued $\mathcal{F}_i^n$-measurable process with $\mathbb{E}\Big(\|\varphi_{i+j-1}^n\|^q\Big|\mathcal{F}_{i-1}^n\Big) \le L^q$ for $q \ge 2$ and $1 \le j \le 2k_n - 1$. Then,*

$$\mathbb{E}\Bigg(\Big\|\sum_{j=1}^{2k_n-1}\varphi_{i+j-1}^n\zeta_{i+j}^n\Big\|^q\Bigg|\mathcal{F}_{i-1}^n\Bigg) \le K_qL^q(L_qk_n^{q/2} + L'^qk_n^q).$$

We introduce some new notation. Following Jacod and Rosenbaum (2012), we define

$$\alpha_i^n = (\Delta_i^nY')(\Delta_i^nY')^\top - C_i^n\Delta_n, \;\; \nu_i^n = \widehat{C}_i'^n - C_i^n, \;\; \text{and} \;\; \lambda_i^n = \widehat{C}_{i+k_n}'^n - \widehat{C}_i'^n, \tag{43}$$

which satisfy

$$\nu_i^n = \frac{1}{k_n\Delta_n}\sum_{j=0}^{k_n-1}(\alpha_{i+j}^n + (C_{i+j}^n - C_i^n)\Delta_n) \;\; \text{and} \;\; \lambda_i^n = \nu_{i+k_n} - \nu_i^n + \Delta_n(C_{i+k_n}^n - C_i^n). \tag{44}$$

We remind some well-known results. For any continuous Itô process Z_t, we have

$$\mathbb{E}\Big(\sup_{w\in[0,s]}\Big\|Z_{t+w}-Z_t\Big\|^q\Big|\mathcal{F}_t\Big)\leq K_q s^{q/2},\text{ and }\Big\|\mathbb{E}\Big(Z_{t+s}-Z_t\Big)\Big|\mathcal{F}_t\Big\|\leq Ks. \tag{45}$$

Set $\mathcal{F}_i^n=\mathcal{F}_{(i-1)\Delta_n}$. By (4.10) in Jacod and Rosenbaum (2013), we have

$$\mathbb{E}\Big(\Big\|\alpha_i^n\Big\|^q\Big|\mathcal{F}_i^n\Big)\leq K_q\Delta_n^q\text{ for all }q\geq 0\text{ and }\mathbb{E}\Big(\Big|\sum_{j=0}^{k_n-1}\alpha_{i+j}^n\Big\|^q\Big|\mathcal{F}_i^n\Big)\leq K_q\Delta_n^q k_n^{q/2}\text{ whenever }q\geq 2. \tag{46}$$

Combining (46), (44), (45) with $Z=c$ and the Hölder inequality yields, for $q\geq 2$,

$$\mathbb{E}\Big(\Big\|\nu_i^n\Big\|^q\Big|\mathcal{F}_i^n\Big)\leq K_q\Delta^{q/4},\quad\text{and}\quad\mathbb{E}\Big(\Big\|\lambda_i^n\Big\|^q\Big|\mathcal{F}_i^n\Big)\leq K_q\Delta^{q/4}. \tag{47}$$

For any càdlàg bounded process Z, we set

$$\eta_{t,s}(Z)=\sqrt{\mathbb{E}\Big(\sup_{0<u\leq s}\|Z_{t+u}-Z_t\|^2|\mathcal{F}_i^n\Big)},$$
$$\eta_{i,j}^n(Z)=\sqrt{\mathbb{E}\Big(\sup_{0\leq u\leq j\Delta_n}\|Z_{(i-1)\Delta_n+u}-Z_{(i-1)\Delta_n}\|^2|\mathcal{F}_i^n\Big)}.$$

Lemma 6. *For any càdlàg bounded process Z, for all $t,s>0$, $j,k\geq 0$, and set $\eta_{t,s}=\eta_{t,s}(Z)$. Then,*

$$\Delta_n\mathbb{E}\Big(\sum_{i=1}^{[t/\Delta_n]}\eta_{i,k_n}\Big)\longrightarrow 0,\qquad\Delta_n\mathbb{E}\Big(\sum_{i=1}^{[t/\Delta_n]}\eta_{i,2k_n}\Big)\longrightarrow 0,$$
$$\mathbb{E}\Big(\eta_{i+j,k}|\mathcal{F}_i^n\Big)\leq\eta_{i,j+k}\quad\textit{and}\quad\Delta_n\mathbb{E}\Big(\sum_{i=1}^{[t/\Delta_n]}\eta_{i,4k_n}\Big)\longrightarrow 0.$$

The first three claims of Lemma 6 are proved in Jacod and Rosenbaum (2012). The last result can be proved similarly to the first two.

Lemma 7. *Let Z be a continuous Itô process with drift term b_t^Z and spot variance process C_t^Z, and set $\eta_{t,s}=\eta_{t,s}(b^Z,c^Z)$. Then, the following bounds hold:*

$$|\mathbb{E}(Z_t|\mathcal{F}_0)-tb_0^Z|\leq Kt\eta_{0,t}$$
$$|\mathbb{E}(Z_t^jZ_t^k-tC_0^{Z,jk}|\mathcal{F}_0)|\leq Kt^{3/2}(\sqrt{\Delta_n}+\eta_{0,t})$$
$$|\mathbb{E}\big((Z_t^jZ_t^k-tC_0^{Z,jk})(C_t^{Z,lm}-C_0^{Z,lm})|\mathcal{F}_0\big)|\leq Kt^2$$
$$|\mathbb{E}(Z_t^jZ_t^kZ_t^lZ_t^m|\mathcal{F}_0)-\Delta_n^2(C_0^{Z,jk}C_0^{Z,lm}+C_0^{Z,jl}C_0^{Z,km}+C_0^{Z,jm}C_0^{Z,kl})|\leq Kt^{5/2}$$
$$|\mathbb{E}(Z_t^jZ_t^kZ_t^l|\mathcal{F}_0)|\leq Kt^2$$
$$|\mathbb{E}(\prod_{l=1}^{6}Z_t^{j_l}|\mathcal{F}_0)-\frac{\Delta_n^3}{6}\sum_{l<l'}\sum_{k<k'}\sum_{m<m'}C_0^{Z,j_lj_{l'}}C_0^{Z,j_kj_{k'}}C_0^{Z,j_mj_{m'}}|\leq Kt^{7/2}$$

The first four claims of Lemma 7 are parts of Lemma 4.1 in Jacod and Rosenbaum (2012). The two remaining statements can be shown similarly.

Lemma 8. *The following results hold:*

$$|\mathbb{E}(\nu_i^{n,jk}\nu_i^{n,lm}\nu_i^{n,gh}|\mathcal{F}_i^n)|\leq K\Delta_n^{3/4}(\Delta_n^{1/4}+\eta_{i,k_n}^n), \tag{48}$$

$$|\mathbb{E}(\nu_i^{n,jk}\nu_i^{n,lm}(c_{i+k_n}^{n,gh}-c_i^{n,gh})|\mathcal{F}_i^n)| \leq K\Delta_n^{3/4}(\Delta_n^{1/4}+\eta_{i,k_n}^n), \tag{49}$$
$$|\mathbb{E}(\nu_i^{n,jk}(c_{i+k_n}^{n,lm}-c_i^{n,lm})(c_{i+k_n}^{n,gh}-c_i^{n,gh})|\mathcal{F}_i^n)| \leq K\Delta_n^{3/4}(\Delta_n^{1/4}+\eta_{i,k_n}^n), \tag{50}$$
$$|\mathbb{E}(\nu_i^{n,jk}\lambda_i^{n,lm}\lambda_i^{n,gh}|\mathcal{F}_i^n)| \leq K\Delta_n^{3/4}(\Delta_n^{1/4}+\eta_{i,2k_n}^n), \tag{51}$$
$$|\mathbb{E}(\lambda_i^{n,jk}\lambda_i^{n,lm}\lambda_i^{n,gh}|\mathcal{F}_i^n)| \leq K\Delta_n^{3/4}(\Delta_n^{1/4}+\eta_{i,2k_n}^n). \tag{52}$$

Proof of (48) in Lemma 8

We start by obtaining some useful bounds for some quantities of interest. First, using the second statement in Lemma 7 applied to $Z = Y'$, we have

$$|\mathbb{E}(\alpha_i^{n,jk}|\mathcal{F}_i^n)| \leq K\Delta_n^{3/2}(\sqrt{\Delta_n}+\eta_{i,1}^n). \tag{53}$$

Second, by repeated application of the Cauchy-Schwartz inequality and making use of the third and last statements in Lemma 7 as well as (45) with $Z = c$, it can be shown that

$$\left|\mathbb{E}(\alpha_i^{n,jk}\alpha_i^{n,lm}|\mathcal{F}_i^n) - \Delta_n^2\Big(C_i^{n,jl}C_i^{n,km}+C_i^{n,jm}C_i^{n,kl}\Big)\right| \leq K\Delta_n^{5/2}. \tag{54}$$

Next, by successive conditioning and using the bound in (45) for $Z = c$ as well as (53) and (54) , we have for $0 \leq u \leq k_n - 1$,

$$\left|\mathbb{E}(\alpha_{i+u}^{n,jk}|\mathcal{F}_i^n)\right| \leq K\Delta_n^{3/2}(\sqrt{\Delta_n}+\eta_{i,u}^n), \tag{55}$$

$$\left|\mathbb{E}(\alpha_{i+u}^{n,jk}\alpha_{i+u}^{n,lm}|\mathcal{F}_i^n) - \Delta_n^2\Big(C_i^{n,jl}C_i^{n,km}+C_i^{n,jm}C_i^{n,kl}\Big)\right| \leq K\Delta_n^{5/2}, \tag{56}$$

To show (48), we first observe that $\nu_i^{n,jk}\nu_i^{n,lm}\nu_i^{n,gh}$ can be decomposed as

$$\begin{aligned}
\nu_i^{n,jk}\nu_i^{n,lm}\nu_i^{n,gh} &= \frac{1}{k_n^3\Delta_n^3}\sum_{u=0}^{k_n-1}\zeta_{i,u}^{n,jk}\zeta_{i,u}^{n,lm}\zeta_{i,u}^{n,gh} + \frac{1}{k_n^3\Delta_n^3}\sum_{u=0}^{k_n-2}\sum_{v=u+1}^{k_n-1}\Big[\zeta_{i,u}^{n,jk}\zeta_{i,v}^{n,lm}\zeta_{i,v}^{n,gh}+\zeta_{i,u}^{n,gh}\zeta_{i,v}^{n,jk}\zeta_{i,v}^{n,lm}\\
&+\zeta_{i,u}^{n,lm}\zeta_{i,v}^{n,gh}\zeta_{i,v}^{n,jk}\Big] + \frac{1}{k_n^3\Delta_n^3}\sum_{u=0}^{k_n-2}\sum_{v=u+1}^{k_n-1}\big[\zeta_{i,u}^{n,jk}\zeta_{i,u}^{n,lm}\zeta_{i,v}^{n,gh}+\zeta_{i,u}^{n,gh}\zeta_{i,u}^{n,jk}\zeta_{i,v}^{n,lm}+\zeta_{i,u}^{n,lm}\zeta_{i,u}^{n,gh}\zeta_{i,v}^{n,jk}\big]\\
&+\frac{1}{k_n^3\Delta_n^3}\sum_{u=0}^{k_n-3}\sum_{v=u+1}^{k_n-2}\sum_{w=v+1}^{k_n-1}\Big[\zeta_{i,u}^{n,jk}\zeta_{i,v}^{n,lm}\zeta_{i,w}^{n,gh}+\zeta_{i,u}^{n,jk}\zeta_{i,v}^{n,gh}\zeta_{i,w}^{n,lm}+\zeta_{i,u}^{n,lm}\zeta_{i,v}^{n,jk}\zeta_{i,w}^{n,gh}+\zeta_{i,u}^{n,lm}\zeta_{i,v}^{n,gh}\zeta_{i,w}^{n,jk}\\
&+\zeta_{i,u}^{n,gh}\zeta_{i,v}^{n,lm}\zeta_{i,w}^{n,jk}+\zeta_{i,u}^{n,gh}\zeta_{i,v}^{n,jk}\zeta_{i,w}^{n,lm}\Big],
\end{aligned}$$

with $\zeta_{i,u}^n = \alpha_{i+u}^n + (C_{i+u}^n - C_i^n)\Delta_n$, which satisfies $\mathbb{E}(\|\zeta_{i,u}^n\|^q|\mathcal{F}_i^n) \leq K\Delta_n^q$ for $q \geq 2$.
Set

$$\xi_i^n(1) = \frac{1}{k_n^3\Delta_n^3}\sum_{u=0}^{k_n-1}\zeta_{i,u}^{n,jk}\zeta_{i,u}^{n,lm}\zeta_{i,u}^{n,gh}, \quad \xi_i^n(2) = \frac{1}{k_n^3\Delta_n^3}\sum_{u=0}^{k_n-2}\sum_{v=u+1}^{k_n-1}\zeta_{i,u}^{n,jk}\zeta_{i,v}^{n,lm}\zeta_{i,v}^{n,gh}$$
$$\xi_i^n(3) = \frac{1}{k_n^3\Delta_n^3}\sum_{u=0}^{k_n-2}\sum_{v=u+1}^{k_n-1}\zeta_{i,u}^{n,jk}\zeta_{i,u}^{n,lm}\zeta_{i,v}^{n,gh} \quad \text{and} \quad \xi_i^n(4) = \frac{1}{k_n^3\Delta_n^3}\sum_{u=0}^{k_n-3}\sum_{v=u+1}^{k_n-2}\sum_{w=v+1}^{k_n-1}\zeta_{i,u}^{n,jk}\zeta_{i,v}^{n,lm}\zeta_{i,w}^{n,gh}.$$

The following bounds can be established,

$$|\mathbb{E}(\xi_i^n(1)|\mathcal{F}_i^n)| \leq K\Delta_n, \quad |\mathbb{E}(\xi_i^n(2)|\mathcal{F}_i^n)| \leq K\Delta_n, \quad |\mathbb{E}(\xi_i^n(3)|\mathcal{F}_i^n)| \leq K\Delta_n \quad \text{and}$$

$$|\mathbb{E}(\xi_i^n(4)|\mathcal{F}_i^n)| \le K\Delta_n^{3/4}(\Delta_n^{1/4} + \eta_{i,k_n}).$$

Proof of $|\mathbb{E}(\xi_i^n(1)|\mathcal{F}_i^n)| \le K\Delta_n$

The result readily follows from an application of the Cauchy Schwartz inequality together with the bound $\mathbb{E}(\|\zeta_{i+u}^n\|^q|\mathcal{F}_i^n) \le K_q\Delta_n^q$ for $q \ge 2$.

Proof of $|\mathbb{E}(\xi_i^n(2)|\mathcal{F}_i^n)| \le K\Delta_n$

Using the law of iterated expectation, we have, for $u < v$,

$$\mathbb{E}(\zeta_{i+u}^{n,jk}\zeta_{i+v}^{n,lm}\zeta_{i+v}^{n,gh}|\mathcal{F}_i^n) = \mathbb{E}(\zeta_{i+u}^{n,jk}\mathbb{E}(\zeta_{i+v}^{n,lm}\zeta_{i+v}^{n,gh}|\mathcal{F}_{i+u+1}^n)|\mathcal{F}_i^n). \tag{57}$$

By successive conditioning, (54), and the Cauchy-Schwartz inequality, we also have

$$|\mathbb{E}(\zeta_{i,v}^{n,lm}\zeta_{i,v}^{n,gh}|\mathcal{F}_{i+u+1}^n) - \Delta_n^2(C_{i+u+1}^{n,lg}C_{i+u+1}^{n,mh} + C_{i+u+1}^{n,lh}C_{i+u+1}^{n,mg}) - \Delta_n^2(C_{i+u+1}^{n,gh} - C_i^{n,gh})(C_{i+u+1}^{n,lm} - C_i^{n,lm})| \le K\Delta_n^{5/2}.$$

Given that $\mathbb{E}(|\zeta_{i+u}^{n,jk}|^q|\mathcal{F}_i^n) \le \Delta_n^q$, the approximation error involved in replacing $\mathbb{E}(\zeta_{i+v}^{n,lm}\zeta_{i+v}^{n,gh}|\mathcal{F}_{i+u+1}^n)$ by $\Delta_n^2(C_{i+u+1}^{n,lg}C_{i+u+1}^{n,mh} + C_{i+u+1}^{n,lh}C_{i+u+1}^{n,mg}) + \Delta_n^2(C_{i+u+1}^{n,gh} - C_i^{n,gh})(C_{i+u+1}^{n,lm} - C_i^{n,lm})$ in (57) is smaller than $\Delta_n^{7/2}$. From (3.9) in Jacod and Rosenbaum (2012) we have

$$|\mathbb{E}(\alpha_{i+u}^{n,jk}(C_{i+u+1}^{n,lm} - C_{i+u}^{n,lm})|\mathcal{F}_i^n)| \le K\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_{i,k_n}^n). \tag{58}$$

Since $(C_{i+u}^n - C_i^n)$ is $\mathcal{F}_{i+u}^n$-measurable, we use the successive conditioning, the Cauchy-Schwartz inequality, (53), (54), and the fifth statement in Lemma 7 applied to $Z = c$ to obtain

$$\begin{aligned}
&|\mathbb{E}(\alpha_{i+u}^{n,gh}(C_{i+u}^{n,lm} - C_i^{n,lm})(C_{i+u}^{n,jk} - C_i^{n,jk})|\mathcal{F}_i^n)| \le K\Delta_n^{5/2}\\
&|\mathbb{E}(\alpha_{i+u}^{n,jk}\alpha_{i+u}^{n,lm}(C_{i+u}^{n,gh} - C_i^{n,gh})|\mathcal{F}_i^n)| \le K\Delta_n^{5/2}\\
&|\mathbb{E}\big((C_{i+u}^{n,lm} - C_i^{n,lm})(C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+u}^{n,gh} - C_i^{n,gh})\big)|\mathcal{F}_i^n)| \le K\Delta_n,
\end{aligned} \tag{59}$$

which can be proved using . The following inequalities can be established easily using (53), the successive conditioning together with (45) for $Z = c$,

$$\Big|\mathbb{E}(\alpha_{i+u}^{n,jk}(C_{i+u+1}^{n,lg}C_{i+u+1}^{n,mh} + C_{i+u+1}^{n,lh}C_{i+u+1}^{n,mg})|\mathcal{F}_i^n)\Big| \le K\Delta_n^{3/2}$$

$$\Big|\mathbb{E}\Big((C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+u+1}^{n,lg}C_{i+u+1}^{n,mh} + C_{i+u+1}^{n,lh}C_{i+u+1}^{n,mg})|\mathcal{F}_i^n\Big)\Big| \le K\Delta_n^{1/2}$$

$$\Big|\mathbb{E}(\alpha_{i+u}^{n,jk}(C_{i+u+1}^{n,gh} - C_i^{n,gh})(C_{i+u+1}^{n,lm} - C_i^{n,lm})|\mathcal{F}_i^n)\Big| \le K\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_{i,k_n}^n).$$

The last three inequalities together yield $|\mathbb{E}(\xi_i^n(2)|\mathcal{F}_i^n)| \le K\Delta_n$.

Proof of $|\mathbb{E}(\xi_i^n(3)|\mathcal{F}_i^n)| \le K\Delta_n$

First, note that, for $u < v$, we have

$$\mathbb{E}(\zeta_{i+u}^{n,jk}\zeta_{i+u}^{n,lm}\zeta_{i+v}^{n,gh}|\mathcal{F}_i^n) = \mathbb{E}(\zeta_{i+u}^{n,jk}\zeta_{i+u}^{n,lm}\mathbb{E}(\zeta_{i+v}^{n,gh}|\mathcal{F}_{i+u+1}^n)|\mathcal{F}_i^n). \tag{60}$$

By successive conditioning and (53) , we have

$$|\mathbb{E}(\alpha_{i+w}^{n,gh}|\mathcal{F}_{i+v+1}^n)| \le K\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_{i+v+1,w-v}). \tag{61}$$

Using the first statement of Lemma applied to $Z = c$, it can be shown that

$$|\mathbb{E}\big((C^{n,gh}_{i+w} - C^{n,gh}_{i+v+1}))|\mathcal{F}^n_i\big) - \Delta_n(w-v-1)\widetilde{b}^{n,gh}_{i+v+1}| \leq K(w-v-1)\Delta_n\eta_{i+v+1,w-v} \leq K\Delta_n^{1/2}\eta_{i+v+1,w-v}.$$

The last two inequalities together imply

$$\Big|\mathbb{E}\Big(\zeta^{n,gh}_{i+w}|\mathcal{F}^n_{i+v+1}\Big) - (C^{n,gh}_{i+v+1} - C^{n,gh}_i)\Delta_n - \Delta_n^2(w-v-1)\widetilde{b}^{n,gh}_{i+v+1}\Big| \leq K\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_{i+v+1,w-v}). \quad (62)$$

Since $\mathbb{E}(|\zeta^{n,jk}_{i,u}|^q|\mathcal{F}^n_i) \leq \Delta_n^q$, the error induced by replacing $\mathbb{E}(\zeta^{n,gh}_{i+v}|\mathcal{F}^n_{i+u+1})$ by $(C^{n,gh}_{i+v+1} - C^{n,gh}_i)\Delta_n + \Delta_n^2(w - v-1)\widetilde{b}^{n,gh}_{i+v+1}$ in (60) is smaller that $\Delta_n^{7/2}$.
Using Cauchy Schwartz inequality, successive conditioning, (59), (45) for $Z = c$ and the boundedness of $\widetilde{b}_t$ and C_t we obtain

$$\begin{aligned}
&\Big|\mathbb{E}\Big(\alpha^{n,jk}_{i+u}\alpha^{n,lm}_{i+u}(C^{n,jk}_{i+u+1} - C^{n,gh}_i)|\mathcal{F}^n_{i+u}\Big)\Big| \leq K\Delta_n^{5/2}\\
&\Big|\mathbb{E}\Big(\alpha^{n,jk}_{i+u}\alpha^{n,lm}_{i+u}\widetilde{b}^{n,gh}_{i+u+1}|\mathcal{F}^n_{i+u}\Big)\Big| \leq K\Delta_n^2\\
&\Big|\mathbb{E}\Big(\alpha^{n,jk}_{i+u}(C^{n,lm}_{i+u} - C^{n,lm}_i)(C^{n,gh}_{i+u+1} - C^{n,gh}_i)|\mathcal{F}^n_i\Big)\Big| \leq K\Delta_n^{1/4}\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta^n_{i,k_n})\\
&\Big|\mathbb{E}\Big(\alpha^{n,jk}_{i+u}(C^{n,lm}_{i+u} - C^{n,lm}_i)\widetilde{b}^{n,gh}_{i+u+1}|\mathcal{F}^n_i\Big)\Big| \leq \Delta_n^{5/4}\\
&\Big|\mathbb{E}\Big((C^{n,jk}_{i+u} - C^{n,gh}_i)(C^{n,lm}_{i+u} - C^{n,lm}_i)\widetilde{b}^{n,gh}_{i+u+1}|\mathcal{F}^n_i\Big)\Big| \leq K\Delta_n^{1/2}\\
&\Big|\mathbb{E}\Big((C^{n,jk}_{i+u} - C^{n,jk}_i)(C^{n,lm}_{i+u} - C^{n,lm}_i)(C^{n,gh}_{i+u+1} - C^{n,gh}_i)|\mathcal{F}^n_i\Big)\Big| \leq K\Delta_n.
\end{aligned}$$

The above inequalities together yield $|\mathbb{E}(\xi^n_i(3)|\mathcal{F}^n_i)| \leq K\Delta_n$.

Proof of $|\mathbb{E}(\xi^n_i(4)\big|\mathcal{F}^n_i)| \leq K\Delta_n^{3/4}(\Delta_n^{1/4} + \eta^n_{i,k_n})$

We first observe that $\xi^n_i(4)$ can be rewritten as

$$\xi^n_i(4) = \frac{1}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\sum_{u=0}^{v-1}\zeta^{n,jk}_{i+u}\zeta^{n,lm}_{i+v}\zeta^{n,gh}_{i+w},$$

where

$$\begin{aligned}
\zeta^{n,jk}_{i+u}\zeta^{n,lm}_{i+v}\zeta^{n,gh}_{i+w} = \Bigg[&\alpha^{n,jk}_{i+u}\alpha^{n,lm}_{i+v}\alpha^{n,gh}_{i+w} + \alpha^{n,jk}_{i+u}\Delta_n\alpha^{n,lm}_{i+v}(C^{n,gh}_{i+w} - C^{n,gh}_i) + \alpha^{n,jk}_{i+u}\Delta_n(C^{n,lm}_{i+v} - C^{n,lm}_i)\alpha^{n,gh}_{i+w}\\
&+ \Delta_n^2\alpha^{n,jk}_{i+u}(C^{n,lm}_{i+v} - C^{n,lm}_i)(C^{n,gh}_{i+w} - C^{n,gh}_i) + \Delta_n(C^{n,jk}_{i+u} - C^{n,jk}_i)\alpha^{n,lm}_{i+v}\alpha^{n,gh}_{i+w} + \Delta_n^2(C^{n,jk}_{i+u} - C^{n,jk}_i)\alpha^{n,lm}_{i+v}(C^{n,gh}_{i+w} - C^{n,gh}_i)\\
&+ \Delta_n^2(C^{n,jk}_{i+u} - C^{n,jk}_i)(C^{n,lm}_{i+v} - C^{n,lm}_i)\alpha^{n,gh}_{i+w} + \Delta_n^3(C^{n,jk}_{i+u} - C^{n,jk}_i)(C^{n,lm}_{i+v} - C^{n,lm}_i)(C^{n,gh}_{i+w} - C^{n,gh}_i)\Bigg].
\end{aligned}$$

Based on the above decomposition, we set

$$\xi^n_i(4) = \sum_{j=1}^{8}\chi(j),$$

with $\chi(j)$ defined below. We aim to show that $|\mathbb{E}(\chi(j)\big|\mathcal{F}_i^n)| \leq K\Delta_n^{3/4}(\Delta_n^{1/4} + \eta_{i,k_n}^n)$, $j = 1, \ldots, 8$. First, set

$$\chi(1) = \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} \alpha_{i+v}^{n,lm} \alpha_{i+w}^{n,gh}.$$

Upon changing the order of the summation, we have

$$\chi(1) = \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk}\Big) \alpha_{i+v}^{n,lm} \alpha_{i+w}^{n,gh}.$$

Define also

$$\chi'(1) = \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk}\Big) \alpha_{i+v}^{n,lm} \mathbb{E}(\alpha_{i+w}^{n,gh} | \mathcal{F}_{i+v+1}^n).$$

Note that $\mathbb{E}(\chi(1)|\mathcal{F}_i^n) = \mathbb{E}(\chi'(1)|\mathcal{F}_i^n)$.
It is easy to see that by Lemma 5, we have for $q \geq 2$,

$$\mathbb{E}\Big(\Big\| \sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} \Big\|^q \Big| \mathcal{F}_i^n\Big) \leq K_q \Delta_n^{3q/4}.$$

The Cauchy-Schwartz inequality yields,

$$\mathbb{E}\bigg(\Big| \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk}\Big) \alpha_{i+v}^{n,lm} \mathbb{E}(\alpha_{i+w}^{n,gh} | \mathcal{F}_{i+v+1}^n)\Big| \bigg| \mathcal{F}_i^n\bigg) \leq K k_n^2 \Big[\mathbb{E}\Big(\Big| \sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} \Big|^4 \Big| \mathcal{F}_i^n\Big)\Big]^{1/4} \Big[\mathbb{E}\Big(\Big| \alpha_{i+v}^{n,lm} \Big|^4 \Big| \mathcal{F}_i^n\Big)\Big]^{1/4}$$
$$\times \Big[\mathbb{E}\Big(\Big| \mathbb{E}(\alpha_{i+w}^{n,gh} | \mathcal{F}_{i+v+1}^n)\Big|^2 \Big| \mathcal{F}_i^n\Big)\Big]^{1/2} \leq K \Delta_n k_n^2 \Delta_n^{3/4} \Delta_n^{3/2} (\sqrt{\Delta_n} + \eta_{i,k_n}^n),$$

where the last iteration is obtained using (61) as well as the inequality $(a+b)^{1/2} \leq a^{1/2} + b^{1/2}$, which holds for positive real numbers a and b, and the third statement in Lemma 6.
It follows from this result that

$$|\mathbb{E}\Big(\chi(1)\big|\mathcal{F}_i^n\Big)| \leq K\Delta_n^{3/4}(\sqrt{\Delta_n} + \eta_{i,k_n}^n).$$

Next, we introduce

$$\chi(2) = \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \Delta_n(C_{i+u}^{n,jk} - C_i^{n,jk})\Big) \alpha_{i+v}^{n,lm} \alpha_{i+w}^{n,gh},$$
$$\chi(3) = \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \alpha_{i+v}^{n,jk}\Big) \Delta_n(C_{i+u}^{n,lm} - C_i^{n,lm}) \alpha_{i+w}^{n,gh},$$
$$\chi(4) = \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \Delta_n(C_{i+u}^{n,jk} - C_i^{n,jk})\Big) \Delta_n(C_{i+u}^{n,lm} - C_i^{n,lm}) \alpha_{i+w}^{n,gh}.$$

Given that for $q \geq 2$, we have

$$\mathbb{E}\Big(\Big\| \sum_{u=0}^{v-1} \Delta_n(C_{i+u}^{n,jk} - C_i^{n,jk})\Big\|^q \Big| \mathcal{F}_i^n\Big) \leq K_q \Delta_n^{3q/4} \quad \text{and} \quad \mathbb{E}(\|C_{i+u}^{n,jk} - C_i^{n,jk}\|^q \big| \mathcal{F}_i^n) \leq K_q \Delta_n^{q/4},$$

one can follow essentially the same steps as for $\chi(1)$ to show that

$$|\mathbb{E}(\chi(2)|\mathcal{F}_i^n)| \leq K\Delta_n^{3/4}(\sqrt{\Delta_n} + \eta_{i,k_n}^n) \quad \text{and} \quad |\mathbb{E}(\chi(j)|\mathcal{F}_i^n)| \leq K\Delta_n(\sqrt{\Delta_n} + \eta_{i,k_n}^n) \quad \text{for} \quad j = 3, 4.$$

Define

$$\begin{aligned}
\chi(5) &= \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} \Big) \alpha_{i+v}^{n,lm} \Delta_n (C_{i+w}^{n,gh} - C_i^{n,gh}) \\
\chi'(5) &= \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} \Big) \alpha_{i+v}^{n,lm} \Delta_n \mathbb{E}\big((C_{i+w}^{n,gh} - C_i^{n,gh})|\mathcal{F}_{i+v+1}^n\big) \\
\chi(6) &= \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \Delta_n (C_{i+u}^{n,jk} - C_i^{n,jk}) \Big) \alpha_{i+v}^{n,lm} \Delta_n (C_{i+w}^{n,gh} - C_i^{n,gh}) \\
\chi(7) &= \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} \Big) \Delta_n (C_{i+v}^{n,lm} - C_i^{n,lm}) \Delta_n (C_{i+w}^{n,gh} - C_i^{n,gh}),
\end{aligned}$$

where we have $\mathbb{E}(\chi(5)|\mathcal{F}_i^n) = \mathbb{E}(\chi'(5)|\mathcal{F}_i^n)$. Recalling (62), we further decompose $\chi'(5)$ as,

$$\chi'(5) = \sum_{j=1}^{5} \chi(5)[j],$$

with

$$\begin{aligned}
\chi\prime(5)[1] &= \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} \Big) \alpha_{i+v}^{n,lm} \Big(\mathbb{E}\Big(C_{i+w}^{n,gh} - C_i^{n,gh} | \mathcal{F}_{i+v+1}^n \Big) - (C_{i+v+1}^{n,gh} - C_i^{n,gh})\Delta_n - \widetilde{b}_{i+v+1}^{n,gh} \Delta_n^2 (w - v - 1) \Big) \\
\chi\prime(5)[2] &= \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Delta_n (C_{i+v}^{n,gh} - C_i^{n,gh}) \Big(\sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} \Big) \alpha_{i+v}^{n,lm} \\
\chi\prime(5)[3] &= \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} \Big) \Delta_n (C_{i+v+1}^{n,gh} - C_{i+v}^{n,gh}) \alpha_{i+v}^{n,lm} \\
\chi\prime(5)[4] &= \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Big(\sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} \Big) \Delta_n^2 (w - v - 1)(\widetilde{b}_{i+v+1}^{n,gh} - \widetilde{b}_{i+v}^{n,gh}) \alpha_{i+v}^{n,lm} \\
\chi\prime(5)[5] &= \frac{1}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \Delta_n^2 (w - v - 1) \widetilde{b}_{i+v}^{n,gh} \Big(\sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} \Big) \alpha_{i+v}^{n,lm}.
\end{aligned}$$

Using (62), (61), (58) and following the same strategy proof as for $\chi(1)$, it can be shown that

$$|\mathbb{E}\Big(\chi\prime(5)[j]|\mathcal{F}_i^n\Big)| \leq K\Delta_n^{3/4}(\sqrt{\Delta_n} + \eta_{i,k_n}^n), \quad \text{for} \quad j = 1, \ldots, 5,$$

which in turn implies

$$|\mathbb{E}\Big(\chi(5)|\mathcal{F}_i^n\Big)| \leq K\Delta_n^{3/4}(\sqrt{\Delta_n} + \eta_{i,k_n}^n), \quad \text{for} \quad j = 1, \ldots, 5.$$

The term $\chi(6)$ can be handled similarly to $\chi(5)$, hence we conclude that

$$|\mathbb{E}\Big(\chi(6)|\mathcal{F}_i^n\Big)| \leq K\Delta_n^{3/4}(\sqrt{\Delta_n} + \eta_{i,k_n}^n).$$

Next, we set

$$\chi(7) = \frac{1}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\Bigg(\sum_{v=0}^{w-1}\Big(\sum_{u=0}^{v-1}\alpha_{i+u}^{n,jk}\Big)\Delta_n(C_{i+v}^{n,lm}-C_i^{n,lm})\Delta_n(C_{i+w}^{n,gh}-C_i^{n,gh})\Bigg).$$

Define

$$\chi(7)[1] = \frac{1}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\Bigg(\sum_{v=0}^{w-1}\Big(\sum_{u=0}^{v-1}\alpha_{i+u}^{n,jk}\Big)\Delta_n(C_{i+v}^{n,lm}-C_i^{n,lm})\Delta_n(C_{i+v+1}^{n,gh}-C_{i+v}^{n,gh})\Bigg)$$

$$\chi(7)[2] = \frac{1}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\Bigg(\sum_{v=0}^{w-1}\Big(\sum_{u=0}^{v-1}\alpha_{i+u}^{n,jk}\Big)\Delta_n(C_{i+v}^{n,lm}-C_i^{n,lm})\Delta_n(C_{i+v}^{n,gh}-C_i^{n,gh})\Bigg)$$

$$\chi(7)[3] = \frac{1}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\Bigg(\sum_{v=0}^{w-1}\Big(\sum_{u=0}^{v-1}\alpha_{i+u}^{n,jk}\Big)\Delta_n(C_{i+v}^{n,lm}-C_i^{n,lm})\Delta_n^2(w-v-1)(\widetilde{b}_{i+v+1}^{n,gh}-\widetilde{b}_{i+v}^{n,gh})\Bigg)$$

$$\chi(7)[4] = \frac{1}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\Bigg(\sum_{v=0}^{w-1}\Delta_n^2(w-v-1)\widetilde{b}_{i+v}^{n,gh}\Big(\sum_{u=0}^{v-1}\alpha_{i+u}^{n,jk}\Big)\Delta_n(C_{i+v}^{n,lm}-C_i^{n,lm})\Bigg),$$

so that

$$\chi(7) = \sum_{j=1}^{4}\chi(7)[j].$$

Similar to calculations used for $\chi(1)$, it can be shown that

$$|\mathbb{E}(\chi(7)[j]\big|\mathcal{F}_i^n)| \leq K\Delta_n^{1/4}(\Delta_n^{1/4}+\eta_{i,k_n}), \quad \text{for} \quad j=1,\dots,3.$$

To handle the remaining term $\chi(7)[4]$, we set

$$\chi(7)[4][1] = \frac{\Delta_n^2}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\sum_{u=0}^{v-1}\alpha_{i+u}^{n,jk}(C_{i+u+1}^{n,lm}-C_{i+u}^{n,lm})(C_{i+u+1}^{n,gh}-C_{i+u}^{n,gh})$$

$$\chi(7)[4][2] = \frac{\Delta_n^2}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\sum_{u=0}^{v-1}(C_{i+u}^{n,gh}-C_i^{n,gh})\alpha_{i+u}^{n,jk}(C_{i+u+1}^{n,lm}-C_{i+u}^{n,lm})$$

$$\chi'(7)[4][2] = \frac{\Delta_n^2}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\sum_{u=0}^{v-1}(C_{i+u}^{n,gh}-C_i^{n,gh})\mathbb{E}(\alpha_{i+u}^{n,jk}(C_{i+u+1}^{n,lm}-C_{i+u}^{n,lm})|\mathcal{F}_{i+u}^n)$$

$$\chi(7)[4][3] = \frac{\Delta_n^2}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\sum_{u=0}^{v-1}(C_{i+u}^{n,lm}-C_i^{n,lm})\alpha_{i+u}^{n,jk}(C_{i+u+1}^{n,gh}-C_{i+u}^{n,gh})$$

$$\chi(7)[4][4] = \frac{\Delta_n^2}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\sum_{u=0}^{v-1}(C_{i+u}^{n,lm}-C_i^{n,lm})(C_{i+u}^{n,gh}-C_i^{n,gh})\alpha_{i+u}^{n,jk}$$

$$\chi(7)[4][5] = \frac{\Delta_n^2}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\sum_{u=0}^{v-1}(C_{i+u}^{n,lm}-C_i^{n,lm})\alpha_{i+u}^{n,jk}(C_{i+v}^{n,gh}-C_{i+u+1}^{n,gh})$$

$$\chi'(7)[2][5] = \frac{\Delta_n^2}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\sum_{u=0}^{v-1}(C_{i+u}^{n,lm}-C_i^{n,lm})\alpha_{i+u}^{n,jk}\mathbb{E}((C_{i+v}^{n,gh}-C_{i+u+1}^{n,gh}|\mathcal{F}_{i+u}^n)$$

$$\chi(7)[4][6] = \frac{\Delta_n^2}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\sum_{u=0}^{v-1}\alpha_{i+u}^{n,jk}(C_{i+u+1}^{n,lm}-C_{i+u}^{n,lm})(C_{i+v}^{n,gh}-C_{i+u+1}^{n,gh})$$

$$\chi(7)[4][7] = \frac{\Delta_n^2}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \sum_{u=0}^{v-1} (C_{i+u}^{n,gh} - C_i^{n,gh}) \alpha_{i+u}^{n,jk} (C_{i+v}^{n,lm} - C_{i+u+1}^{n,lm})$$
$$\chi(7)[4][8] = \frac{\Delta_n^2}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} (C_{i+u+1}^{n,gh} - C_{i+u}^{n,gh})(C_{i+v}^{n,lm} - C_{i+u+1}^{n,lm})$$
$$\chi(7)[4][9] = \frac{\Delta_n^2}{(k_n\Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \sum_{u=0}^{v-1} \alpha_{i+u}^{n,jk} (C_{i+v}^{n,lm} - C_{i+u+1}^{n,lm})(C_{i+v}^{n,gh} - C_{i+u+1}^{n,gh}),$$

which satisfy,

$$\chi(7)[4] = \sum_{j=1}^{9} \chi(7)[4][j].$$

By using arguments similar to those used for $\chi(1)$, it can be shown that

$$|\mathbb{E}(\chi(7)[4][j] \big| \mathcal{F}_i^n)| \le K\Delta_n^{1/4}(\Delta_n^{1/4} + \eta_{i,k_n}), \quad \text{for} \quad j = 1, \dots, 8,$$

which yields

$$|\mathbb{E}(\chi(7) \big| \mathcal{F}_i^n)| \le K\Delta_n^{1/4}(\Delta_n^{1/4} + \eta_{i,k_n}).$$

Next, define

$$\chi(8) = \frac{1}{k_n^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \sum_{u=0}^{v-1} (C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+v}^{n,lm} - C_i^{n,lm})(C_{i+w}^{n,gh} - C_i^{n,gh}).$$

This term can be further decomposed into 6 components. Successive conditioning and existing bounds give

$$|\mathbb{E}\Big((C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+v}^{n,lm} - C_{i+u}^{n,lm})(C_{i+w}^{n,gh} - C_{i+v}^{n,gh}) \big| \mathcal{F}_i^n\Big)| \le K\Delta_n$$
$$|\mathbb{E}\Big((C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+v}^{n,lm} - C_{i+u}^{n,lm})(C_{i+v}^{n,gh} - C_{i+u}^{n,gh}) \big| \mathcal{F}_i^n\Big)| \le K\Delta_n^{3/4}(\Delta_n^{1/4} + \eta_{i,k_n})$$
$$|\mathbb{E}\Big((C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+v}^{n,lm} - C_{i+u}^{n,lm})(C_{i+u}^{n,gh} - C_i^{n,gh}) \big| \mathcal{F}_i^n\Big)| \le K\Delta_n$$
$$|\mathbb{E}\Big((C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+u}^{n,lm} - C_i^{n,lm})(C_{i+w}^{n,gh} - C_{i+v}^{n,gh}) \big| \mathcal{F}_i^n\Big)| \le K\Delta_n$$
$$|\mathbb{E}\Big((C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+u}^{n,lm} - C_i^{n,lm})(C_{i+v}^{n,gh} - C_{i+u}^{n,gh}) \big| \mathcal{F}_i^n\Big)| \le K\Delta_n$$
$$|\mathbb{E}\Big((C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+u}^{n,lm} - C_i^{n,lm})(C_{i+u}^{n,gh} - C_i^{n,gh}) \big| \mathcal{F}_i^n\Big)| \le K\Delta_n$$

These bounds can be used to deduce

$$|\mathbb{E}(\chi(8) \big| \mathcal{F}_i^n)| \le K\Delta_n.$$

This completes the proof.

Proof of (49) and (50) in Lemma 8

Observe that

$$\nu_i^{n,jk}(C_{i+k_n}^{n,lm}-C_i^{n,lm})(C_{i+k_n}^{n,gh}-C_i^{n,gh})=\frac{1}{k_n\Delta_n}\sum_{u=0}^{k_n-1}\varsigma_{i,u}^{n,jk}(C_{i+k_n}^{n,lm}-C_i^{n,lm})(C_{i+k_n}^{n,gh}-C_i^{n,gh}),$$

$$\begin{aligned}\nu_i^{n,jk}\nu_i^{n,lm}(C_{i+k_n}^{n,gh}-C_i^{n,gh})&=\frac{1}{k_n^2\Delta_n^2}\sum_{u=0}^{k_n-1}\varsigma_{i,u}^{n,jk}\varsigma_{i,u}^{n,lm}(C_{i+k_n}^{n,gh}-C_i^{n,gh})+\frac{1}{k_n^2\Delta_n^2}\sum_{u=0}^{k_n-2}\sum_{v=0}^{k_n-1}\varsigma_{i,u}^{n,jk}\varsigma_{i,v}^{n,lm}(C_{i+k_n}^{n,gh}-C_i^{n,gh})\\&+\frac{1}{k_n^2\Delta_n^2}\sum_{u=0}^{k_n-2}\sum_{v=0}^{k_n-1}\varsigma_{i,u}^{n,lm}\varsigma_{i,v}^{n,jk}(C_{i+k_n}^{n,gh}-C_i^{n,gh}).\end{aligned}$$

Hence, (49) and (50) can be proved using the same strategy as for (48).

Proof of (51) and (52) in Lemma 8

Note that we have

$$\begin{aligned}&\lambda_i^{n,jk}\lambda_i^{n,lm}\nu_i^{n,gh}=\nu_i^{n,gh}\nu_{i+k_n}^{n,jk}\nu_{i+k_n}^{n,lm}+\nu_i^{n,gh}\nu_i^{n,jk}\nu_i^{n,lm}-\nu_i^{n,gh}\nu_i^{n,lm}\nu_{i+k_n}^{n,jk}-\nu_i^{n,gh}\nu_i^{n,lm}\nu_{i+k_n}^{n,jk}\\&+\nu_i^{n,gh}\nu_{i+k_n}^{n,jk}(C_{i+k_n}^{n,lm}-C_i^{n,lm})-\nu_i^{n,gh}\nu_i^{n,jk}(C_{i+k_n}^{n,lm}-C_i^{n,lm})+\nu_i^{n,gh}\nu_{i+k_n}^{n,lm}(C_{i+k_n}^{n,jk}-C_i^{n,jk})-\nu_i^{n,gh}\nu_i^{n,lm}(C_{i+k_n}^{n,jk}-C_i^{n,jk})\\&+\nu_i^{n,gh}(C_{i+k_n}^{n,jk}-C_i^{n,jk})(C_{i+k_n}^{n,lm}-C_i^{n,lm}),\end{aligned}$$

and

$$\begin{aligned}&\lambda_i^{n,gh}\lambda_i^{n,jk}\lambda_i^{n,lm}=\nu_{i+k_n}^{n,gh}\nu_{i+k_n}^{n,jk}\nu_{i+k_n}^{n,lm}+\nu_{i+k_n}^{n,gh}\nu_i^{n,jk}\nu_i^{n,lm}-\nu_{i+k_n}^{n,gh}\nu_i^{n,lm}\nu_{i+k_n}^{n,jk}-\nu_{i+k_n}^{n,gh}\nu_i^{n,lm}\nu_{i+k_n}^{n,jk}\\&+\nu_{i+k_n}^{n,gh}\nu_{i+k_n}^{n,jk}(C_{i+k_n}^{n,lm}-C_i^{n,lm})-\nu_{i+k_n}^{n,gh}\nu_i^{n,jk}(C_{i+k_n}^{n,lm}-C_i^{n,lm})+\nu_{i+k_n}^{n,gh}\nu_{i+k_n}^{n,lm}(C_{i+k_n}^{n,jk}-C_i^{n,jk})-\nu_{i+k_n}^{n,gh}\nu_i^{n,lm}(C_{i+k_n}^{n,jk}-C_i^{n,jk})\\&+\nu_{i+k_n}^{n,gh}(C_{i+k_n}^{n,jk}-C_i^{n,jk})(C_{i+k_n}^{n,lm}-C_i^{n,lm})-\nu_i^{n,gh}\nu_{i+k_n}^{n,jk}\nu_{i+k_n}^{n,lm}-\nu_i^{n,gh}\nu_i^{n,jk}\nu_i^{n,lm}+\nu_i^{n,gh}\nu_i^{n,lm}\nu_{i+k_n}^{n,jk}+\nu_i^{n,gh}\nu_i^{n,lm}\nu_{i+k_n}^{n,jk}\\&-\nu_i^{n,gh}\nu_{i+k_n}^{n,jk}(C_{i+k_n}^{n,lm}-C_i^{n,lm})+\nu_i^{n,gh}\nu_i^{n,jk}(C_{i+k_n}^{n,lm}-C_i^{n,lm})-\nu_i^{n,gh}\nu_{i+k_n}^{n,lm}(C_{i+k_n}^{n,jk}-C_i^{n,jk})+\nu_i^{n,gh}\nu_i^{n,lm}(C_{i+k_n}^{n,jk}-C_i^{n,jk})\\&-\nu_i^{n,gh}(C_{i+k_n}^{n,jk}-C_i^{n,jk})(C_{i+k_n}^{n,lm}-C_i^{n,lm})+\nu_{i+k_n}^{n,jk}\nu_{i+k_n}^{n,lm}(C_{i+k_n}^{n,gh}-C_i^{n,gh})+\nu_i^{n,jk}\nu_i^{n,lm}(C_{i+k_n}^{n,gh}-C_i^{n,gh})\\&-\nu_i^{n,lm}\nu_{i+k_n}^{n,jk}(C_{i+k_n}^{n,gh}-C_i^{n,gh})-\nu_i^{n,lm}\nu_{i+k_n}^{n,jk}(C_{i+k_n}^{n,gh}-C_i^{n,gh})+\nu_{i+k_n}^{n,jk}(C_{i+k_n}^{n,lm}-C_i^{n,lm})(C_{i+k_n}^{n,gh}-C_i^{n,gh})\\&-\nu_i^{n,jk}(C_{i+k_n}^{n,lm}-C_i^{n,lm})(C_{i+k_n}^{n,gh}-C_i^{n,gh})+\nu_{i+k_n}^{n,lm}(C_{i+k_n}^{n,jk}-C_i^{n,jk})(C_{i+k_n}^{n,gh}-C_i^{n,gh})\\&-\nu_i^{n,lm}(C_{i+k_n}^{n,jk}-C_i^{n,jk})(C_{i+k_n}^{n,gh}-C_i^{n,gh})+(C_{i+k_n}^{n,jk}-C_i^{n,jk})(C_{i+k_n}^{n,lm}-C_i^{n,lm})(C_{i+k_n}^{n,gh}-C_i^{n,gh}).\end{aligned}$$

From (44), notice that ν_i^n is $\mathcal{F}_{i+k_n}^n$-measurable and satisfies $\|\mathbb{E}(\nu_i^n|\mathcal{F}_i^n)\|\le K\Delta_n^{1/2}$.
Using the law of iterated expectations and existing bounds, it can be shown that

$$\begin{aligned}&|\mathbb{E}(\nu_i^{n,lm}\nu_{i+k_n}^{n,jk}|\mathcal{F}_i^n)|\le K\Delta_n^{3/4}.\\&|\mathbb{E}(\nu_i^{n,lm}\nu_i^{n,gh}\nu_{i+k_n}^{n,jk}|\mathcal{F}_i^n)|\le K\Delta_n\\&|\mathbb{E}(\nu_i^{n,lm}(C_{i+k_n}^{n,gh}-C_i^{n,gh})\nu_{i+k_n}^{n,jk}|\mathcal{F}_i^n)|\le K\Delta_n\\&|\mathbb{E}(\nu_{i+k_n}^{n,lm}(C_{i+k_n}^{n,jk}-C_i^{n,jk})|\mathcal{F}_i^n)|\le K\Delta_n^{3/4}\\&|\mathbb{E}((C_{i+k_n}^{n,jk}-C_i^{n,jk})(C_{i+k_n}^{n,lm}-C_i^{n,lm})(C_{i+k_n}^{n,gh}-C_i^{n,gh})|\mathcal{F}_i^n)|\le K\Delta_n.\end{aligned}\tag{63}$$

By Lemma 3.3 in Jacod and Rosenbaum (2012), we have

$$|\mathbb{E}(\nu_{i+k_n}^{n,gh}\nu_{i+k_n}^{n,ab}|\mathcal{F}_{i+k_n}^n)-\frac{1}{k_n}(C_{i+k_n}^{n,ga}C_{i+k_n}^{n,hb}+C_{i+k_n}^{n,gb}C_{i+k_n}^{n,ha})-\frac{k_n\Delta_n}{3}\overline{C}_{i+k_n}^{n,gh,ab}|\le K\sqrt{\Delta_n}(\Delta_n^{1/8}+\eta_{i+k_n,k_n}^n).$$

Hence, for $\varphi_i^{n,gh} \in \{\nu_i^{n,gh}, C_{i+k_n}^{n,gh} - C_i^{n,gh}\}$, which satisfies $\mathbb{E}(|\varphi_i^{n,gh}|^q \big| \mathcal{F}_i^n) \leq K\Delta_n^{q/4}$ and $\mathbb{E}(\varphi_i^{n,gh}|\mathcal{F}_i^n) \leq K\Delta_n^{1/2}$, it can be proved that

$$|\mathbb{E}(\varphi_i^{n,gh}\nu_{i+k_n}^{n,jk}\nu_{i+k_n}^{n,lm}|\mathcal{F}_i^n) - \mathbb{E}\Big(\varphi_i^{n,gh}\Big[\frac{1}{k_n}(C_{i+k_n}^{n,jl}C_{i+k_n}^{n,km} + C_{i+k_n}^{n,jm}C_{i+k_n}^{n,kl}) - \frac{k_n\Delta_n}{3}\overline{C}_{i+k_n}^{n,jk,lm}\Big]|\mathcal{F}_i^n\Big)| \leq K\Delta_n^{3/4}(\Delta_n^{1/4} + \eta_{i,2k_n}^n).$$

Next, successive conditioning and existing bounds give

$$\begin{aligned}
&|\mathbb{E}(\varphi_i^{n,gh}\overline{C}_{i+k_n}^{n,jk,lm})| \leq K\Delta_n^{1/4}(\Delta_n^{1/4} + \eta_{i,k_n}^n) \\
&|\mathbb{E}(\varphi_i^{n,gh}C_{i+k_n}^{n,jl}C_{i+k_n}^{n,km})| \leq K\Delta_n^{1/2},
\end{aligned}$$

which implies

$$|\mathbb{E}(\varphi_i^{n,gh}\nu_{i+k_n}^{n,jk}\nu_{i+k_n}^{n,lm}|\mathcal{F}_i^n)| \leq K\Delta_n^{3/4}(\Delta_n^{1/4} + \eta_{i,2k_n}^n). \tag{64}$$

It is easy to see that (48), (63) and (64) and the inequality $\eta_{i,k_n}^n \leq \eta_{i,2k_n}^n$ together yield (51) and (52).

Step 3: Asymptotic Distribution of the approximate estimator

First, we decompose the approximate estimator as

$$[H(\widehat{C), G}(C)]_T^{(A)} = [H(\widehat{C), G}(C)]_T^{(A1)} - [H(\widehat{C), G}(C)]_T^{(A2)},$$

with

$$[H(\widehat{C), G}(C)]_T^{(A1)} = \frac{3}{2k_n}\sum_{g,h,a,b=1}^{d}\sum_{i=1}^{[T/\Delta_n]-2k_n+1}(\partial_{gh}H\partial_{ab}G)(C_{i-1}^n)(\widehat{C}_{i+k_n}^{\prime n,gh} - \widehat{C}_i^{\prime n,gh})(\widehat{C}_{i+k_n}^{\prime n,ab} - \widehat{C}_i^{\prime n,ab}),$$

and

$$[H(\widehat{C), G}(C)]_T^{(A2)} = \frac{3}{k_n^2}\sum_{g,h,a,b=1}^{d}\sum_{i=1}^{[T/\Delta_n]-2k_n+1}(\partial_{gh}H\partial_{ab}G)(\widehat{C}_i^{\prime n})(\widehat{C}_i^{\prime n,ga}\widehat{C}_i^{\prime n,hb} + \widehat{C}_i^{\prime n,gb}\widehat{C}_i^{\prime n,ha}).$$

In this section, we use the notation $C_{i-1}^n = C_{(i-1)\Delta_n}$ and $\mathcal{F}_i = \mathcal{F}_{(i-1)\Delta_n}$ to simplify the exposition. Given the polynomial growth assumption satisfied by H and G and the fact that $k_n = \theta(\Delta_n)^{-1/2}$, by Theorem 2.2 in Jacod and Rosenbaum (2012) we have

$$\frac{1}{\sqrt{\Delta_n}}\left([H(\widehat{C), G}(C)]_T^{(A2)} - \frac{3}{\theta^2}\sum_{g,h,a,b=1}^{d}\int_0^T(\partial_{gh}H\partial_{ab}G)(C_t)(c_t^{ga}c_t^{hb} + c_i^{gb}c_t^{ha})dt\right) = O_p(1),$$

which yields

$$\frac{1}{\Delta_n^{1/4}}\left([H(\widehat{C), G}(C)]_T^{(A2)} - \frac{3}{\theta^2}\sum_{g,h,a,b=1}^{d}\int_0^T(\partial_{gh}H\partial_{ab}G)(C_t)(c_t^{ga}c_t^{hb} + c_i^{gb}c_t^{ha})dt\right) \xrightarrow{\mathbb{P}} 0.$$

To study the asymptotic behavior of $[H(\widehat{C), G}(C)]_T^{(A1)}$, we follow Aït-Sahalia and Jacod (2014) and define the following multidimensional quantities

$$\zeta(1)_i^n = \frac{1}{\Delta_n}\Delta_i^nY'(\Delta_i^nY')^\top - C_{i-1}^n, \qquad \zeta(2)_i^n = \Delta_i^nc,$$

$$\zeta'(u)_i^n = \mathbb{E}(\zeta(u)_i^n|\mathcal{F}_{i-1}^n), \qquad \zeta''(u)_i^n = \zeta(u)_i^n - \zeta'(u)_i^n,$$

with

$$\zeta^r(u)_i^n = \left(\zeta^r(u)_i^{n,gh}\right)_{1\leq g,h\leq d}.$$

We also define, for $m \in \{0, \dots, 2k_n - 1\}$ and $j, l \in \mathbb{Z}$,

$$\varepsilon(1)_m^n = \begin{cases} -1 & if\ 0 \leq m < k_n \\ +1 & if\ k_n \leq m < 2k_n, \end{cases}$$

$$\varepsilon(2)_m^n = \sum_{q=m+1}^{2k_n-1} \varepsilon(1)_q^n = (m+1) \wedge (2k_n - m - 1),$$

$$z_{u,v}^n = \begin{cases} 1/\Delta_n & if\ u = v = 1 \\ 1 & \text{otherwise}, \end{cases}$$

$$\lambda(u,v;m)_{j,l}^n = \frac{3}{2k_n^3} \sum_{q=0\vee(j-m)}^{(l-m-1)\vee(2k_n-m-1)} \varepsilon(u)_q^n \varepsilon(u)_{q+m}^n, \qquad \lambda(u,v)_m^n = \lambda(u,v;m)_{0,2k_n}^n,$$
$$M(u,v;u',v')_n = z_{u,v}^n z_{u',v'}^n \sum_{m=1}^{2k_n-1} \lambda(u,v)_m^n \lambda(u',v')_m^n.$$

The following decompositions hold,

$$\widehat{C}_i'^n = C_{i-1}^n + \frac{1}{k_n}\sum_{j=0}^{k_n-1}\sum_{u=1}^{2} \overline{\varepsilon}(u)_j^n \zeta(u)_{i+j}^n, \qquad \widehat{C}_{i+k_n}'^n - \widehat{C}_i'^n = \frac{1}{k_n}\sum_{j=0}^{2k_n-1}\sum_{u=1}^{2} \varepsilon(u)_j^n \zeta(u)_{i+j}^n,$$
$$\lambda_i^{n,gh}\lambda_i^{n,ab} = \frac{1}{k_n^2}\sum_{u=1}^{2}\sum_{v=1}^{2}\left(\sum_{j=0}^{2k_n-1} \varepsilon(u)_j^n\varepsilon(v)_j^n\zeta(u)_{i+j}^{n,gh}\zeta(v)_{i+j}^{n,ab} + \sum_{j=0}^{2k_n-2}\sum_{q=j+1}^{2k_n-1} \varepsilon(u)_j^n\varepsilon(v)_q^n\zeta(u)_{i+j}^{n,gh}\zeta(v)_{i+q}^{n,ab} \right.$$
$$\left. + \sum_{j=1}^{2k_n-1}\sum_{q=0}^{j-1} \varepsilon(u)_j^n\varepsilon(v)_q^n\zeta(u)_{i+j}^{n,gh}\zeta(v)_{i+q}^{n,ab}\right).$$

A change of the order of the summation in the last term gives

$$\lambda_i^{n,gh}\lambda_i^{n,ab} = \frac{1}{k_n^2}\sum_{u=1}^{2}\sum_{v=1}^{2}\left(\sum_{j=0}^{2k_n-1} \varepsilon(u)_j^n\varepsilon(v)_j^n\zeta(u)_{i+j}^{n,gh}\zeta(v)_{i+j}^{n,ab} + \sum_{j=0}^{2k_n-2}\sum_{q=j+1}^{2k_n-1} \varepsilon(u)_j^n\varepsilon(v)_q^n\zeta(u)_{i+j}^{n,gh}\zeta(v)_{i+q}^{n,ab} \right.$$
$$\left. + \sum_{j=0}^{2k_n-2}\sum_{q=j+1}^{2k_n-1} \varepsilon(v)_j^n\varepsilon(u)_q^n\zeta(v)_{i+j}^{n,ab}\zeta(u)_{i+q}^{n,gh}\right).$$

Therefore, we can further rewrite $[\widehat{H(C),G(C)}]_T^{(A1)}$ as

$$[\widehat{H(C),G(C)}]_T^{(A1)} = [\widehat{H(C),G(C)}]_T^{(A11)} + [\widehat{H(C),G(C)}]_T^{(A12)} + [\widehat{H(C),G(C)}]_T^{(A13)}, \text{with}$$

$$[\widehat{H(C),G(C)}]_T^{(A1w)} = \sum_{g,h,a,b=1}^{d} \sum_{u,v=1}^{2} \widehat{A1w}(H,gh,u;G,ab,v)_T^n, \quad w=1,2,3,$$

and,

$$\widehat{A11}(H,gh,u;G,ab,v)_T^n = \frac{3}{2k_n^3} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} \sum_{j=0}^{2k_n-1} (\partial_{gh}H\partial_{ab}G)(C_{i-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n\zeta(u)_{i+j}^{n,gh}\zeta(v)_{i+j}^{n,ab},$$

$$\widehat{A12}(H,gh,u;G,ab,v)_T^n = \frac{3}{2k_n^3} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} \sum_{j=0}^{2k_n-2} \sum_{q=j+1}^{2k_n-1} (\partial_{gh}H\partial_{ab}G)(C_{i-1}^n)\varepsilon(u)_j^n\varepsilon(v)_q^n\zeta(u)_{i+j}^{n,gh}\zeta(v)_{i+q}^{n,ab},$$

$$\widehat{A13}(H,gh,u;G,ab,v)_T^n = \frac{3}{2k_n^3} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} \sum_{j=0}^{2k_n-2} \sum_{q=j+1}^{2k_n-1} (\partial_{gh}H\partial_{ab}G)(C_{i-1}^n)\varepsilon(v)_j^n\varepsilon(u)_q^n\zeta(v)_{i+j}^{n,ab}\zeta(u)_{i+q}^{n,gh},$$

where we clearly have $\widehat{A13}(H,gh,u;G,ab,v)_T^n = \widehat{A12}(G,ab,v;H,gh,u)_T^n$. By a change of the order of the summation,

$$\widehat{A11}(H,gh,u;G,ab,v)_T^n = \frac{3}{2k_n^3} \sum_{i=1}^{[T/\Delta_n]} \sum_{j=0\vee(i+2k_n-1-[T/\Delta_n])}^{(2k_n-1)\wedge(i-1)} (\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab},$$

$$\widehat{A12}(H,gh,u;G,ab,v)_T^n = \frac{3}{2k_n^3} \sum_{i=2}^{[T/\Delta_n]} \sum_{m=1}^{(i-1)\wedge(2k_n-1)} \sum_{j=0\vee(i+2k_n-1-m-[T/\Delta_n])}^{(2k_n-m-1)\wedge(i-m-1)} (\partial_{gh}H\partial_{ab}G)(C_{i-1-j-m}^n)\times$$
$$\varepsilon(u)_j^n\varepsilon(v)_{j+m}^n\zeta_{gh}(u)_{i-m}^n\zeta_{ab}(v)_i^n.$$

Set

$$\widetilde{A11}(H,gh,u;G,ab,v)_T^n = \frac{3}{2k_n^3} \sum_{i=2k_n}^{[T/\Delta_n]} \sum_{j=0}^{2k_n-1} (\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab},$$

$$\widetilde{A12}(H,gh,u;G,ab,v)_T^n = \frac{3}{2k_n^3} \sum_{i=2k_n}^{[T/\Delta_n]} \sum_{m=1}^{(i-1)\wedge(2k_n-1)} \sum_{j=0}^{(2k_n-m-1)} (\partial_{gh}H\partial_{ab}G)(C_{i-j-1-m}^n)\varepsilon(u)_j^n\varepsilon(v)_{j+m}^n\zeta_{gh}(u)_{i-m}^n\zeta_{ab}(v)_i^n,$$

and

$$\begin{aligned}\overline{A11}(H,gh,u;G,ab,v)_T^n &= \frac{3}{2k_n^3} \sum_{i=2k_n}^{[T/\Delta_n]} \Big(\sum_{j=0}^{2k_n-1} \varepsilon(u)_j^n\varepsilon(v)_j^n\Big)(\partial_{gh}H\partial_{ab}G)(C_{i-2k_n}^n)\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab} \\ &= \lambda(u,v)_0^n \sum_{i=2k_n}^{[T/\Delta_n]} (\partial_{gh}H\partial_{ab}G)(C_{i-2k_n}^n)\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}, \\ \overline{A12}(H,gh,u;G,ab,v)_T^n &= \frac{3}{2k_n^3} \sum_{i=2k_n}^{[T/\Delta_n]} (\partial_{gh}H\partial_{ab}G)(C_{i-2k_n}^n) \sum_{m=1}^{(i-1)\wedge(2k_n-1)} \sum_{j=0}^{(2k_n-m-1)} \varepsilon(u)_j^n\varepsilon(v)_{j+m}^n\zeta_{gh}(u)_{i-m}^n\zeta_{ab}(v)_i^n \\ &= \sum_{i=2k_n}^{[T/\Delta_n]} (\partial_{gh}H\partial_{ab}G)(C_{i-2k_n}^n)\rho_{gh}(u,v)_i^n\zeta_{ab}(v)_i^n,\end{aligned}$$

with

$$\rho_{gh}(u,v)_i^n = \sum_{m=1}^{2k_n-1} \lambda(u,v)_m^n \zeta_{gh}(u)_{i-m}^n.$$

The following results hold:

$$\frac{1}{\Delta_n^{1/4}}\Big(\widehat{A1w}(H,gh,u;G,ab,v)_T^n - \widetilde{A1w}(H,gh,u;G,ab,v)_T^n\Big) \xrightarrow{\mathbb{P}} 0 \quad \text{for all} \quad (H,gh,u,G,ab,v) \text{ and } w=1,2. \tag{65}$$

$$\frac{1}{\Delta_n^{1/4}}\Big(\widetilde{A1w}(H,gh,u;G,ab,v)_T^n - \overline{A1w}(H,gh,u;G,ab,v)_T^n\Big) \xrightarrow{\mathbb{P}} 0 \quad \text{for all} \quad (H,gh,u,G,ab,v) \text{ and } w=1,2. \tag{66}$$

Proof of (65) for $w = 1$

The proof is similar to Step 5 on page 548 of Aït-Sahalia and Jacod (2014). Our proof deviates from the latter reference by the fact that, in all the sums, the terms $\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}$ are scaled by random variables rather that constant real numbers. First, observe that we can write

$$\widehat{A11} - \widetilde{A11} = \widetilde{\widehat{A11}}(1) + \widetilde{\widehat{A11}}(2) + \widetilde{\widehat{A11}}(3) \quad \text{with}$$

$$\widetilde{\widehat{A11}}(1) = \sum_{i=1}^{(2k_n-1)\wedge[T/\Delta_n]} \left(\frac{3}{2k_n^3}\sum_{j=0\vee(i+2k_n-1-[T/\Delta_n])}^{(2k_n-1)\wedge(i-1)} (\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n\right)\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab},$$

$$\widetilde{\widehat{A11}}(2) = \sum_{i=[T/\Delta_n]-2k_n+2}^{[T/\Delta_n]} \frac{3}{2k_n^3}\left(\sum_{j=0\vee(i+2k_n-1-[T/\Delta_n])}^{(2k_n-1)\wedge(i-1)} (\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n - \sum_{j=0}^{(2k_n-1)} (\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n\right)\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab},$$

$$\widetilde{\widehat{A11}}(3) = \sum_{i=2k_n}^{[T/\Delta_n]-2k_n+1} \frac{3}{2k_n^3}\left(\sum_{j=0\vee(i+2k_n-1-[T/\Delta_n])}^{(2k_n-1)\wedge(i-1)} (\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n - \sum_{j=0}^{(2k_n-1)} (\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n\right)\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}.$$

It is easy to see that $\widetilde{\widehat{A12}}(3) = 0$. Using (45) with $Z = c$ and (46), it can be shown that

$$\mathbb{E}(\|\zeta(1)_i^n\|^q|\mathcal{F}_{i-1}^n) \le K_q, \quad \mathbb{E}(\|\zeta(2)_i^n\|^q|\mathcal{F}_{i-1}^n) \le K_q\Delta_n^{q/2}. \tag{67}$$

The polynomial growth assumption on H and G and the boundedness of C_t imply that $|(\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)| \le K$. Hence, the random quantities $\left(\frac{3}{2k_n^3}\sum_{j=0\vee(i+2k_n-1-[T/\Delta_n])}^{(2k_n-1)\wedge(i-1)}(\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n\right)$ and $\frac{3}{2k_n^3}\sum_{j=0}^{(2k_n-1)}(\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n$ are $\mathcal{F}_{i-1}^n-$ measurable and are bounded by $\widetilde{\lambda}_{u,v}^n$ defined as

$$\widetilde{\lambda}_{u,v}^n = \begin{cases} K & \text{if } (u,v) = (2,2) \\ K/k_n & \text{if } (u,v) = (1,2),(2,1) \\ K/k_n^2 & \text{if } (u,v) = (1,1). \end{cases}$$

Similarly, the quantity,

$$\frac{3}{2k_n^3}\left(\sum_{j=0\vee(i+2k_n-1-[T/\Delta_n])}^{(2k_n-1)\wedge(i-1)}(\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n-\sum_{j=0}^{(2k_n-1)}(\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n\right),$$

is $\mathcal{F}_{i-1}^n-$ measurable and bounded by $2\widetilde{\lambda}_{u,v}^n$. Note also that, by (67) and the Cauchy Schwartz inequality, we have,

$$\mathbb{E}(|\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}||\mathcal{F}_{i-1}^n)\leq\mathbb{E}(\|\zeta(u)_i^n\|^2|\mathcal{F}_{i-1}^n)^{1/2}\mathbb{E}(\|\zeta(v)_i^n\|^2|\mathcal{F}_{i-1}^n)^{1/2}\leq\begin{cases}K\Delta_n & \text{if } (u,v)=(2,2)\\ K\Delta_n^{1/2} & \text{if } (u,v)=(1,2),(2,1)\\ K & \text{if } (u,v)=(1,1).\end{cases}$$

The above bounds, together with the fact that $k_n=\theta\Delta_n^{-1/2}$, give $\mathbb{E}(|\widetilde{\widehat{A11}}(1)|)\leq K\Delta_n^{1/2}$ and $\mathbb{E}(|\widetilde{\widehat{A11}}(2)|)\leq K\Delta_n^{1/2}$ for all (u,v). These two results together imply $\widetilde{\widehat{A11}}(1)=o(\Delta_n^{-1/4})$ and $\widetilde{\widehat{A11}}(2)=o(\Delta_n^{-1/4})$, which yields the result.

Proof of (65) for $w=2$

We proceed similarly to Step 6 on page 548 of Aït-Sahalia and Jacod (2014). First, observe that we have

$$\widehat{A12}-\widetilde{A12}=\widetilde{\widehat{A12}}(1)+\widetilde{\widehat{A12}}(2)\quad\text{with}$$

$$\widetilde{\widehat{A12}}(1)=\sum_{i=2}^{(2k_n-1)\wedge[T/\Delta_n]}\left(\sum_{m=1}^{(i-1)}\frac{3}{2k_n^3}\Big(\sum_{j=0\vee(i+2k_n-1-m-[T/\Delta_n])}^{(2k_n-m-1)\wedge(i-m-1)}(\partial_{gh}H\partial_{ab}G)(C_{i-1-j-m}^n)\varepsilon(u)_j^n\varepsilon(v)_{j+m}^n\Big)\zeta_{gh}(u)_{i-m}^n\right)\zeta_{ab}(v)_i^n,$$

$$\widetilde{\widehat{A12}}(2)=\sum_{i=[T/\Delta_n]-2k_n+2}^{[T/\Delta_n]}\left(\sum_{m=1}^{(i-1)\wedge(2k_n-1)}\Big(\frac{3}{2k_n^3}\sum_{j=0\vee(i+2k_n-1-m-[T/\Delta_n])}^{(2k_n-m-1)\wedge(i-m-1)}(\partial_{gh}H\partial_{ab}G)(C_{i-1-j-m}^n)\varepsilon(u)_j^n\varepsilon(v)_{j+m}^n\Big)\right.$$
$$\left.-\sum_{j=0}^{(2k_n-m-1)}(\partial_{gh}H\partial_{ab}G)(C_{i-1-j-m}^n)\varepsilon(u)_j^n\varepsilon(v)_{j+m}^n\Big)\zeta_{gh}(u)_{i-m}^n\right)\zeta_{ab}(v)_i^n.$$

It is easy to see that the quantity

$$\kappa_i^{m,n}=\frac{3}{2k_n^3}\Big(\sum_{j=0\vee(i+2k_n-1-m-[T/\Delta_n])}^{(2k_n-m-1)\wedge(i-m-1)}(\partial_{gh}H\partial_{ab}G)(C_{i-1-j-m}^n)\varepsilon(u)_j^n\varepsilon(v)_{j+m}^n\Big)$$

is $\mathcal{F}_{i-m-1}^n$ measurable and bounded by $\widetilde{\lambda}_{u,v}^n$. Let

$$\kappa_i^n=\sum_{m=1}^{(i-1)}\frac{3}{2k_n^3}\Big(\sum_{j=0\vee(i+2k_n-1-m-[T/\Delta_n])}^{(2k_n-m-1)\wedge(i-m-1)}(\partial_{gh}H\partial_{ab}G)(C_{i-1-j-m}^n)\varepsilon(u)_j^n\varepsilon(v)_{j+m}^n\Big)\zeta_{gh}(u)_{i-m}^n.$$

It follows that κ_i^n is $\mathcal{F}_{i-1}^n$-measurable. We have

$$\mathbb{E}(|\kappa_i^{m,n}|^z|\mathcal{F}_0)\leq(\widetilde{\lambda}_{u,v}^n)^z$$

$$|\mathbb{E}(\zeta(u)^n_{i-m}|\mathcal{F}_{i-m-1})| \leq \begin{cases} K\sqrt{\Delta_n} & \text{if } u=1 \\ K\Delta_n & \text{if } u=2 \end{cases}, \qquad \mathbb{E}(\|\zeta(u)^n_{i-m}\|^z|\mathcal{F}_{i-m-1}) \leq \begin{cases} K_z & \text{if } u=1 \\ K_z\Delta_n^{z/2} & \text{if } u=2 \end{cases}$$

Using Lemma 5, we deduce that for $z \geq 2$,

$$\mathbb{E}(|\kappa^n_i|^z) \leq \begin{cases} K_z(\widetilde{\lambda}^n_{u,v})^z k_n^{z/2} & \text{if } u=1 \\ K_z(\widetilde{\lambda}^n_{u,v})^z/k_n^{z/2} & \text{if } u=2 \end{cases} \leq \begin{cases} K_z/k_n^{-3z/2} & \text{if } v=1 \\ K_z k_n^{-z/2} & \text{if } v=2 \end{cases}$$

Using the above result, and similarly to step 6 on page 548 of Aït-Sahalia and Jacod (2014), we obtain that $\frac{1}{\Delta_n^{1/4}}\widetilde{\widehat{A12}}(1) \overset{\mathbb{P}}{\Rightarrow} 0$. A similar argument yields $\frac{1}{\Delta_n^{1/4}}\widetilde{\widehat{A12}}(2) \overset{\mathbb{P}}{\Rightarrow} 0$, which completes the proof of (65) for $w=2$.

Proof of (66) for $w = 1$

Define

$$\Theta(u,v)^{(C),i,n}_0 = \frac{3}{2k_n^3}\sum_{j=0}^{2k_n-1}\Big((\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1}) - (\partial_{gh}H\partial_{ab}G)(C^n_{i-2k_n})\Big)\varepsilon(u)^n_j\varepsilon(v)^n_j.$$

By Taylor expansion, the polynomial growth assumption on H and G and using (45) with $Z = c$, we have

$$\Big|\mathbb{E}\Big((\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1}) - (\partial_{gh}H\partial_{ab}G)(C^n_{i-2k_n})|\mathcal{F}^n_{i-2k_n}\Big)\Big| \leq K(k_n\Delta_n) \leq K\sqrt{\Delta_n} \quad \text{for} \quad j=0,\ldots,2k_n-1$$
$$\mathbb{E}(|(\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1}) - (\partial_{gh}H\partial_{ab}G)(C^n_{i-2k_n})|^q|\mathcal{F}^n_{i-2k_n})| \leq K(k_n\Delta_n)^{q/2} \leq K\Delta_n^{q/4} \quad \text{for} \quad q \geq 2$$

Next, observe that $\Theta(u,v)^{(C),i,n}_0$ is $\mathcal{F}^n_{i-1}$ -measurable and satisfies $|\Theta(u,v)^{(C),i,n}_0| \leq \widetilde{\lambda}^n_{u,v}$, $|\mathbb{E}\Big(\Theta(u,v)^{(C),i,n}_0|\mathcal{F}^n_{i-2k_n}\Big)| \leq K\Delta_n^{1/2}\widetilde{\lambda}^n_{u,v}$ and $\mathbb{E}\Big(|\Theta(u,v)^{(C),i,n}_0|^q|\mathcal{F}^n_{i-2k_n}\Big) \leq K_q\Delta_n^{q/4}(\widetilde{\lambda}^n_{u,v})^q$ where the latter follows from the Hölder inequality. We aim to prove that

$$\widehat{E} = \frac{1}{\Delta_n^{1/4}}\left[\sum_{i=2k_n}^{[T/\Delta_n]}\Theta(u,v)^{(C),i,n}_0\zeta(u)^{n,gh}_i\zeta(v)^{n,ab}_i\right]$$

converges to zero in probability for any H, G, g, h, a, and b with $u,v = 1,2$.
To show this result, we first introduce the following quantities:

$$\widehat{E}(1) = \frac{1}{\Delta_n^{1/4}}\left[\sum_{i=2k_n}^{[T/\Delta_n]}\Theta(u,v)^{(C),i,n}_0\mathbb{E}(\zeta(u)^{n,gh}_i\zeta(v)^{n,ab}_i|\mathcal{F}^n_{i-1})\right]$$
$$\widehat{E}(2) = \frac{1}{\Delta_n^{1/4}}\left[\sum_{i=2k_n}^{[T/\Delta_n]}\Theta(u,v)^{(C),i,n}_0\big(\zeta(u)^{n,gh}_i\zeta(v)^{n,ab}_i - \mathbb{E}(\zeta(u)^{n,gh}_i\zeta(v)^{n,ab}_i|\mathcal{F}^n_{i-1})\big)\right],$$

with $\widehat{E} = \widehat{E}(1) + \widehat{E}(2)$. By Cauchy-Schwartz inequality, we have

$$\mathbb{E}(|\zeta(u)^{n,gh}_i\zeta(v)^{n,ab}_i|^q) \leq (\widehat{\lambda}^n_{u,v})^{q/2}, \text{where } \widehat{\lambda}^n_{u,v} = \begin{cases} K & \text{if } (u,v) = (1,1) \\ K\Delta_n & \text{if } (u,v) = (1,2),(2,1) \\ K\Delta_n^2 & \text{if } (u,v) = (2,2) \end{cases}$$

Since $\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}$ is $\mathcal{F}_i^n$-measurable, the martingale property of $\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab} - \mathbb{E}(\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}|\mathcal{F}_{i-1}^n)$ implies, for all (u,v),

$$\mathbb{E}(|\widehat{E}(2)|^2) \leq K\Delta_n^{-3/2}(\Delta_n^{1/4}\widetilde{\lambda}_{u,v}^n)^2\widehat{\lambda}_{u,v}^n \leq K\Delta_n.$$

The latter inequality implies $\widehat{E}(2) \overset{\mathbb{P}}{\Rightarrow} 0$ for all (u,v). It remains to show that $\widehat{E}(1) \overset{\mathbb{P}}{\Rightarrow} 0$.
We remind some bounds under Assumption 2, see (B.83) in Aït-Sahalia and Jacod (2014),

$$|\mathbb{E}(\zeta(1)_i^{n,gh}\zeta(2)_i^{n,ab}|\mathcal{F}_{i-1}^n)| \leq K\Delta_n, \tag{68}$$
$$|\mathbb{E}(\zeta(1)_i^{n,gh}\zeta(1)_i^{n,ab}|\mathcal{F}_{i-1}^n) - (C_{i-1}^{n,ga}C_{i-1}^{n,hb} + C_{i-1}^{n,gb}C_{i-1}^{n,ha})| \leq K\Delta_n^{1/2}, \tag{69}$$
$$|\mathbb{E}(\zeta(2)_i^{n,gh}\zeta(2)_i^{n,ab}|\mathcal{F}_{i-1}^n - \overline{C}_{i-1}^{n,gh,ab}\Delta_n)| \leq K\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_i^n). \tag{70}$$

Case $(u,v) \in \{(1,2),(2,1)\}$. By (68) we have

$$\mathbb{E}(|\widehat{E}(1)|) \leq K\frac{T}{\Delta_n}\frac{1}{\Delta_n^{1/4}}(\Delta_n^{1/4}\widetilde{\lambda}_{u,v}^n\Delta_n) \leq K\Delta_n^{1/2} \quad \text{so} \quad \widehat{E}(1) \overset{\mathbb{P}}{\Rightarrow} 0.$$

Case $(u,v) \in \{(1,1),(2,2)\}$. Set

$$\widehat{E}'(1) = \frac{1}{\Delta_n^{1/4}}\left[\sum_{i=2k_n}^{[T/\Delta_n]}\Theta(u,v)_0^{(C),i,n}V_{i-2k_n}^n\right]$$
$$\widehat{E}''(1) = \frac{1}{\Delta_n^{1/4}}\left[\sum_{i=2k_n}^{[T/\Delta_n]}\Theta(u,v)_0^{(C),i,n}(V_{i-1}^n - V_{i-2k_n}^n)\right]$$
$$\widehat{E}'''(1) = \frac{1}{\Delta_n^{1/4}}\left[\sum_{i=2k_n}^{[T/\Delta_n]}\Theta(u,v)_0^{(C),i,n}\Big(\mathbb{E}(\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}|\mathcal{F}_{i-1}^n) - V_{i-1}^n\Big)\right]$$

where

$$V_{i-1}^n = \begin{cases} C_{i-1}^{n,ga}C_{i-1}^{n,hb} + C_{i-1}^{n,gb}C_{i-1}^{n,ha} & \text{if } (u,v) = (2,2) \\ \overline{C}_{i-1}^{n,gh,ab}\Delta_n & \text{if } (u,v) = (1,1) \\ 0 & \text{otherwise} \end{cases}$$

Note that we have $\widehat{E}(1) = \widehat{E}'(1) + \widehat{E}''(1) + \widehat{E}'''(1)$. Using (69) and (70), it can be shown that

$$\mathbb{E}(|\widehat{E}'''(1)|) \leq \begin{cases} K\frac{1}{\Delta_n^{5/4}}(\Delta_n^{1/4}\widetilde{\lambda}_{u,v}^n)\Delta_n^{1/2} & \text{if } (u,v) = (1,1) \\ K\frac{1}{\Delta_n^{5/4}}(\Delta_n^{1/4}\widetilde{\lambda}_{u,v}^n)\Delta_n^{3/2} & \text{if } (u,v) = (2,2) \end{cases} \leq K\Delta_n^{1/2} \quad \text{in all cases.}$$

Next, we prove $\widehat{E}'(1) \overset{\mathbb{P}}{\Rightarrow} 0$. To this end, write

$$\widehat{E}'(1) = \frac{1}{\Delta_n^{1/4}}\left[\sum_{i=1}^{[T/\Delta_n]-2k_n+1}\Theta(u,v)_0^{(C),i-1+2k_n,n}V_{(i-1)\Delta_n}\right].$$

The fact that the summand in the last sum is $\mathcal{F}_{i+2k_n-2}^n$-measurable and lemma B.8 in Aït-Sahalia and Jacod (2014) imply that it is sufficient to show

$$\frac{1}{\Delta_n^{1/4}}\left[\sum_{i=1}^{[T/\Delta_n]-2k_n+1}|\mathbb{E}(\Theta(u,v)_0^{(C),i-1+2k_n,n}V_{(i-1)\Delta_n}|\mathcal{F}_{i-1}^n)|\right] \overset{\mathbb{P}}{\Rightarrow} 0 \quad \text{and}$$

$$\frac{2k_n - 2}{\Delta_n^{1/2}} \left[\sum_{i=1}^{[T/\Delta_n]-2k_n+1} \mathbb{E}\Big(|\Theta(u,v)_0^{(C),i-1+2k_n,n} V_{(i-1)\Delta_n})|^2 \Big) \right] \Rightarrow 0.$$

The first result readily follows from the inequality

$$|\mathbb{E}(\Theta(u,v)_0^{(C),i-1+2k_n,n} V_{(i-1)\Delta_n} | \mathcal{F}_{i-1}^n)| \leq \begin{cases} K\Delta_n^{1/2} \widetilde{\lambda}_{u,v}^n & \text{if } (u,v) = (1,1) \\ K\Delta_n^{1/2} \widetilde{\lambda}_{u,v}^n \Delta_n & \text{if } (u,v) = (2,2) \end{cases} \leq K\Delta_n^{3/2} \quad \text{in all cases}$$

while the second is a direct consequence of

$$\mathbb{E}(|\Theta(u,v)_0^{(C),i-1+2k_n,n} V_{(i-1)\Delta_n}|^2) \leq \begin{cases} K\Delta_n^{1/2} (\widetilde{\lambda}_{u,v}^n)^2 & \text{if } (u,v) = (1,1) \\ K\Delta_n^{1/2} (\widetilde{\lambda}_{u,v}^n)^2 \Delta_n^2 & \text{if } (u,v) = (2,2) \end{cases} \leq K\Delta_n^{5/2} \quad \text{in all cases.}$$

Finally, to prove that $\widehat{E}''(1) \stackrel{\mathbb{P}}{\Longrightarrow} 0$, we use the fact that

$$\begin{aligned} \mathbb{E}(|\Theta(u,v)_0^{(C),i,n} (V_{(i-1)\Delta_n} - V_{(i-2k_n)\Delta_n})|) &\leq \mathbb{E}(|\Theta(u,v)_0^{(C),i,n}|^2)^{1/2} \mathbb{E}(|V_{(i-1)\Delta_n} - V_{(i-2k_n)\Delta_n}|^2)^{1/2} \\ &\leq \begin{cases} K\Delta_n^{1/2} \widetilde{\lambda}_{u,v}^n & \text{if } (u,v) = (1,1) \\ K\Delta_n^{1/4} \widetilde{\lambda}_{u,v}^n \Delta_n \Delta_n^{1/4} & \text{if } (u,v) = (2,2) \end{cases}, \end{aligned}$$

which follows by the Cauchy-Schwartz inequality and earlier bounds. In particular, successive conditioning together with Assumption 2 imply that for $(u,v) = (1,1)$ and $(2,2)$, $\mathbb{E}(|V_{(i-1)\Delta_n} - V_{(i-2k_n)\Delta_n}|^2) \leq \Delta_n^{1/2}$.

Proof of (66) for $w = 2$

Our aim here is to show that

$$\widehat{E}(2) = \frac{1}{\Delta_n^{1/4}} \sum_{i=2k_n}^{[T/\Delta_n]} \left(\sum_{m=1}^{2k_n-1} \Big(\frac{3}{2k_n^3} \sum_{j=0}^{2k_n-m-1} \big[(\partial_{gh} H \partial_{ab} G)(c_{i-j-m-1}^n) - (\partial_{gh} H \partial_{ab} G)(c_{i-2k_n}^n) \big] \varepsilon(u)_j^n \varepsilon(v)_{j+m}^n \Big) \times \zeta(u)_{i-m}^{n,gh} \right) \zeta(v)_i^{n,ab} \stackrel{\mathbb{P}}{\Longrightarrow} 0.$$

For this purpose, we introduce some new notation. For any $0 \leq m \leq 2k_n - 1$, set

$$\Theta(u,v)_m^{(C),i,n} = \frac{3}{2k_n^3} \sum_{j=0}^{2k_n-m-1} \big[(\partial_{gh} H \partial_{ab} G)(c_{i-j-m-1}^n) - (\partial_{gh} H \partial_{ab} G)(c_{i-2k_n}^n) \big] \varepsilon(u)_j^n \varepsilon(v)_{j+m}^n$$

$$\rho(u,v)^{(C),i,n,gh} = \sum_{m=1}^{2k_n-1} \Theta(u,v)_m^{(C),i,n} \zeta(u)_{i-m}^{n,gh}.$$

It is easy to see that $\Theta(u,v)_m^{(C),i,n}$ is $\mathcal{F}_{i-m-1}^n$ measurable and satisfies, by Hölder inequality,

$$|\Theta(u,v)_m^{(C),i,n}| \leq \widetilde{\lambda}_{u,v}^n \quad \text{and} \quad \mathbb{E}\Big(|\Theta(u,v)_m^{(C),i,n}|^q \big| \mathcal{F}_{i-2k_n}^n \Big) \leq K_q \Delta_n^{q/4} (\widetilde{\lambda}_{u,v}^n)^q.$$

Lemma 5 implies that for $q \geq 2$,

$$\mathbb{E}(|\rho(u,v)^{(C),i,n,gh}|^q) \leq \begin{cases} K_q(\Delta_n^{1/4}\widetilde{\lambda}_{u,v}^n)^q k_n^{q/2} & \text{if } u=1 \\ K_q(\Delta_n^{1/4}\widetilde{\lambda}_{u,v}^n)^q / k_n^{q/2} & \text{if } u=2 \end{cases} \leq \begin{cases} K_q/k_n^{2q} & \text{if } v=1 \\ K_q k_n^q & \text{if } v=2 \end{cases}. \tag{71}$$

Set

$$\widehat{E}'(2) = \frac{1}{\Delta_n^{1/4}} \sum_{i=2k_n}^{[T/\Delta_n]} \rho(u,v)^{(C),i,n,gh}\mathbb{E}(\zeta(v)_i^{n,ab}|\mathcal{F}_{i-1}^n),$$

$$\widehat{E}''(2) = \frac{1}{\Delta_n^{1/4}} \sum_{i=2k_n}^{[T/\Delta_n]} \rho(u,v)^{(C),i,n,gh}(\zeta(v)_i^{n,ab} - \mathbb{E}(\zeta(v)_i^{n,ab}|\mathcal{F}_{i-1}^n)).$$

The martingale increments property implies $\mathbb{E}(|\widehat{E}''(2)|^2) \leq K\Delta_n^{1/2}$ in all the cases, which in turn implies $\widehat{E}''(2) \stackrel{\mathbb{P}}{\Longrightarrow} 0$. Next, using the bounds on $\rho(u,v)^{(C),i,n,gh}$ and similarly to step 7 on page 549 of Aït-Sahalia and Jacod (2014), we obtain that $\widehat{E}'(2) \stackrel{\mathbb{P}}{\Longrightarrow} 0$.

Return to the proof of Theorem 1

So far, we have proved that

$$\frac{1}{\Delta_n^{1/4}}\Big([H(\widehat{C),G}(C)]_T^{(A1)} - \sum_{g,h,a,b=1}^{d}\sum_{u,v=1}^{2} \overline{A11}(H,gh,u;G,ab,v)_T^n + \overline{A12}(H,gh,u;G,ab,v)_T^n$$
$$+ \overline{A12}(G,ab,v;H,gh,u)_T^n\Big) \stackrel{\mathbb{P}}{\longrightarrow} 0.$$

We next show that,

$$\frac{1}{\Delta_n^{1/4}}\sum_{i=2k_n}^{[T/\Delta_n]}(\partial_{gh}H\partial_{ab}G)(C_{i-2k_n}^n)\rho_{gh}(u,v)_i^n\zeta'_{ab}(v)_i^n \stackrel{\mathbb{P}}{\Longrightarrow} 0, \quad \forall \quad (u,v) \tag{72}$$

$$\frac{1}{\Delta_n^{1/4}}\Big(\overline{A11}(H,gh,u;G,ab,v) - \int_0^T(\partial_{gh}H\partial_{ab}G)(C_t)\overline{C}_t^{gh,ab}dt\Big) \stackrel{\mathbb{P}}{\Longrightarrow} 0 \text{ when } (u,v)=(2,2) \tag{73}$$

$$\frac{1}{\Delta_n^{1/4}}\Big(\overline{A11}(H,gh,u;G,ab,v) - \frac{3}{\theta^2}\int_0^T(\partial_{gh}H\partial_{ab}G)(C_t)(C_t^{ga}C_t^{hb}+C_t^{gb}C_t^{ha})dt\Big) \stackrel{\mathbb{P}}{\Longrightarrow} 0 \text{ when } (u,v)=(1,1) \tag{74}$$

$$\frac{1}{\Delta_n^{1/4}}\overline{A11}(H,gh,u;G,ab,v) \stackrel{\mathbb{P}}{\Longrightarrow} 0 \text{ when } (u,v)=(1,2),(2,1) \tag{75}$$

which will in turn imply

$$\frac{1}{\Delta_n^{1/4}}\Big([H(\widehat{C),G}(C)]_T^{(A)} - [H(C),G(C)]_T - \frac{3}{2k_n^3}\sum_{g,h,a,b}^{d}\sum_{u,v=1}^{2}\sum_{i=2k_n}^{[T/\Delta_n]}\Big[(\partial_{gh}H\partial_{ab}G)(C_{i-2k_n}^n)\rho_{gh}(u,v)_i^n\zeta''_{ab}(v)_i^n \tag{76}$$

$$+ (\partial_{ab}H\partial_{gh}G)(C_{i-2k_n}^n)\rho_{ab}(v,u)_i^n\zeta''_{gh}(v)_i^n\Big]\Big) \stackrel{\mathbb{P}}{\Longrightarrow} 0. \tag{77}$$

(72) can be proved easily following steps similar to step 7 on page 549 of Aït-Sahalia and Jacod (2014) and using the bounds of $\rho(u,v)_i^{n,gh}$ in (71) . To show (73),(74) and (75), we set

$$\overline{\overline{A11}}(H,gh,u;G,ab,v)=\lambda(u,v)_0^n\sum_{i=2k_n}^{[T/\Delta_n]}(\partial_{gh}H\partial_{ab}G)(C_{i-1})\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}.$$

Then it holds that,

$$\frac{1}{\Delta_n^{1/4}}\Big(\overline{\overline{A11}}(H,gh,u;G,ab,v)-\overline{A11}(H,gh,u;G,ab,v)\Big)\overset{\mathbb{P}}{\Rightarrow}0.$$

This result can be proved following similar steps as for (65) in case $w=1$ by replacing $\Theta(u,v)_0^{(C),i,n}$ by $\lambda(u,v)_0^n((\partial_{gh}H\partial_{ab}G)(C_{i-1})-(\partial_{gh}H\partial_{ab}G)(C_{i-2k_n}))$, which has the same bounds as the former. Next, decompose $\overline{\overline{A11}}$ as follows,

$$\begin{aligned}\overline{\overline{A11}}(H,gh,u;G,ab,v)=\lambda(u,v)_0^n\Bigg[&\sum_{i=2k_n}^{[T/\Delta_n]}(\partial_{gh}H\partial_{ab}G)(C_{i-1})V_{i-1}^n\\&+\sum_{i=2k_n}^{[T/\Delta_n]}(\partial_{gh}H\partial_{ab}G)(C_{i-1})\Big(\mathbb{E}(\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}|\mathcal{F}_{i-1}^n)-V_{i-1}^n\Big)\\&+\sum_{i=2k_n}^{[T/\Delta_n]}(\partial_{gh}H\partial_{ab}G)(C_{i-1})\Big(\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}-\mathbb{E}(\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}|\mathcal{F}_{i-1}^n)\Big)\Bigg].\end{aligned}$$

We follow the proof of (66) for $w=1$, and we replace $\Theta(u,v)_0^{(C),i,n}$ by $\lambda(u,v)_0^n(\partial_{gh}H\partial_{ab}G)(C_{i-1})$, which satisfies only the condition $|\lambda(u,v)_0^n(\partial_{gh}H\partial_{ab}G)(C_{i-1})|\leq\widetilde{\lambda}_{u,v}^n$. This calculation shows that the last two terms in the above decomposition of vanish at a rate slower that $\Delta_n^{1/4}$. Therefore,

$$\frac{1}{\Delta_n^{1/4}}\left(\overline{\overline{A11}}(H,gh,u;G,ab,v)-\lambda(u,v)_0^n\Big(\sum_{i=2k_n}^{[T/\Delta_n]}(\partial_{gh}H\partial_{ab}G)(C_{i-1})V_{i-1}^n\Big)\right)\Rightarrow0.$$

As a consequence, for $(u,v)=(1,2)$ and $(2,1)$,

$$\frac{1}{\Delta_n^{1/4}}\overline{\overline{A11}}(H,gh,u;G,ab,v)\Rightarrow0.$$

The results follow from the following observation,

$$\frac{1}{\Delta_n^{1/4}}\left(\lambda(u,v)_0^n\Big(\sum_{g,h,a,b=1}^{d}\sum_{i=2k_n}^{[T/\Delta_n]}(\partial_{gh}H\partial_{ab}G)(C_{i-1})V_{i-1}^n(u,v)\Big)-\frac{3}{\theta^2}\int_0^T(\partial_{gh}H\partial_{ab}G)(C_t)(C_t^{ga}C_t^{hb}+C_t^{gb}C_t^{ha})dt\right)\Rightarrow0,$$

$$\text{for }(u,v)=(2,2)$$

$$\frac{1}{\Delta_n^{1/4}}\left(\sum_{g,h,a,b=1}^{d}\lambda(u,v)_0^n\Big(\sum_{i=2k_n}^{[T/\Delta_n]}(\partial_{gh}H\partial_{ab}G)(C_{i-1})V_{i-1}^n(u,v)\Big)-[H(C),G(C)]_T\right)\Rightarrow0,\quad\text{for }(u,v)=(1,1).$$

Set

$$\xi(H,gh,u;G,ab,v)_i^n=\frac{1}{\Delta_n^{1/4}}(\partial_{gh}H\partial_{ab}G)(C_{i-2k_n}^n)\rho_{gh}(u,v)_i^n\zeta_{ab}''(v)_i^n,$$

$$Z(H, gh, u; G, ab, v)^n_t = \Delta_n^{1/4} \sum_{i=2k_n}^{[t/\Delta_n]} \xi(H, gh, u; G, ab, v)^n_i.$$

Notice that (76) implies

$$\frac{1}{\Delta_n^{1/4}}\Big([H(\widehat{C), G}(C)]^{(A)}_T - [H(C), G(C)]_T\Big) \overset{\mathcal{L}}{=} \sum_{g,h,a,b=1}^{d} \sum_{u,v=1}^{2} \frac{1}{\Delta_n^{1/4}}\Big(Z(H, gh, u; G, ab, v)^n_T + Z(H, ab, v; G, gh, u)^n_T\Big). \tag{78}$$

Next, observe that to derive the asymptotic distribution of $\Big([H_1(\widehat{C), G_1}(C)]^{(A)}_T, \ldots, [H_\kappa(\widehat{C), G_\kappa}(C)]^{(A)}_T\Big)$, it suffices to study the joint asymptotic behavior of the family of processes $\frac{1}{\Delta_n^{1/4}} Z(H, gh, u; G, ab, v)^n_T$.
It is easy to see that $\xi(H, gh, u; G, ab, v)^n_i$ are martingale increments, relative to the discrete filtration $(\mathcal{F}^n_i)$. Therefore, by Theorem 2.2.15 of Jacod and Protter (2012), to obtain the joint asymptotic distribution of $\frac{1}{\Delta_n^{1/4}} Z(H, gh, u; G, ab, v)^n_T$, it is enough to prove the following three properties, for all $t > 0$, all $(H, gh, u; G, ab, v), (H', g'h', u'; G', a'b', v')$ and all martingales N which are either bounded and orthogonal to W, or equal to one component W^j,

$$A\Big((H, gh, u; G, ab, v), (H', g'h', u'; G', a'b', v')\Big)^n_t := \sum_{i=2k_n}^{[t/\Delta_n]} \mathbb{E}(\xi(H, gh, u; G, ab, v)^n_i \xi(H', g'h', u'; G', a'b', v')^n_i | \mathcal{F}^n_{i-1})$$
$$\overset{\mathbb{P}}{\Longrightarrow} A\Big((H, gh, u; G, ab, v), (H', g'h', u'; G', a'b', v')\Big)_t$$
$$\sum_{i=2k_n}^{[t/\Delta_n]} \mathbb{E}(|\xi(H, gh, u; G, ab, v)^n_i|^4 | \mathcal{F}^n_{i-1}) \overset{\mathbb{P}}{\Longrightarrow} 0$$
$$B(N; H, gh, u; G, ab, v)^n_t := \sum_{i=2k_n}^{[t/\Delta_n]} \mathbb{E}(\xi(H, gh, u; G, ab, v)^n_i \Delta^n_i N | \mathcal{F}^n_{i-1}) \overset{\mathbb{P}}{\Longrightarrow} 0.$$

Using the polynomial growth assumption on H_r and G_r, the second and the third results can be proved by a natural extension to the multivariate case of (B.105) and (B.106) in Aït-Sahalia and Jacod (2014).
Define

$$V^{a'b'}_{ab}(v, v')_t = \begin{cases} (C^{aa'}_t C^{bb'}_t + C^{ab'}_t C^{ba'}_t) & \text{if} \quad (v, v') = (1, 1) \\ \overline{C}^{ab,a'b'}_t & \text{if} \quad (v, v') = (2, 2) \\ 0 & \text{otherwise}, \end{cases}$$

and

$$\overline{V}^{g'h'}_{gh}(u, u')_t = \begin{cases} (C^{gg'}_t C^{hh'}_t + C^{gh'}_t C^{hg'}_t) & \text{if} \quad (u, u') = (1, 1) \\ \overline{C}^{gh,g'h'}_t & \text{if} \quad (u, u') = (2, 2) \\ 0 & \text{otherwise}. \end{cases}$$

Once again using the polynomial growth assumption on H_r and G_r and following steps similar to the proof of (B.104) in Aït-Sahalia and Jacod (2014), one can show that

$$A\Big((H, gh, u; G, ab, v), (H', g'h', u'; G', a'b', v')\Big)_t =$$
$$M(u, v; u', v') \int_0^t (\partial_{gh} H \partial_{ab} G \partial_{g'h'} H \partial_{a'b'} G)(C_s) V^{a'b'}_{ab}(v, v')_s \overline{V}^{g'h'}_{gh}(u, u')_s ds,$$

with

$$M(u,v;u',v') = \begin{cases} 3/\theta^3 & \text{if} \quad (u,v;u',v') = (1,1;1,1) \\ 3/4\theta & \text{if} \quad (u,v;u',v') = (1,2;1,2),(2,1;2,1) \\ 151\theta/280 & \text{if} \quad (u,v;u',v') = (2,2;2,2) \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, we have
$A\Big((H,gh,u;G,ab,v),(H',g'h',u';G',a'b',v')\Big)_T =$

$$\begin{cases} \frac{3}{\nu^3}\int_0^T(\partial_{gh}H\partial_{ab}G\partial_{g'h'}H'\partial_{a'b'}G')(C_t)(C_t^{gg'}C_t^{hh'}+C_t^{gh'}C_t^{hg'})(C_t^{aa'}C_t^{bb'}+C_t^{ab'}C_t^{ba'})dt & \text{if} \quad (u,v;u',v') = (1,1;1,1) \\ \frac{3}{4\nu}\int_0^T(\partial_{gh}H\partial_{ab}G\partial_{g'h'}H'\partial_{a'b'}G')(C_t)(C_t^{gg'}C_t^{hh'}+C_t^{gh'}C_t^{hg'})\overline{C}_t^{ab,a'b'}dt & \text{if} \quad (u,v;u',v') = (1,2;1,2) \\ \frac{3}{4\nu}\int_0^T(\partial_{gh}H\partial_{ab}G\partial_{g'h'}H'\partial_{a'b'}G')(C_t)(C_t^{aa'}C_t^{bb'}+C_t^{ab'}C_s^{ba'})\overline{t}_s^{gh,g'h'}dt & \text{if} \quad (u,v;u',v') = (2,1;2,1) \\ \frac{151\nu}{280}\int_0^T(\partial_{gh}H\partial_{ab}G\partial_{g'h'}H'\partial_{a'b'}G')(C_t)\overline{C}_s^{ab,a'b'}\overline{C}_t^{gh,g'h'}dt & \text{if} \quad (u,v;u',v') = (2,2;2,2) \\ 0 & \text{otherwise.} \end{cases}$$

Using (78), we deduce that the asymptotic covariance between $[H_r(\widehat{C),G_r}(C)]_T^{(A)}$ and $[H_s(\widehat{C),G_s}(C)]_T^{(A)}$ is given by

$$\sum_{g,h,a,b=1}^{d}\sum_{g',h',a',b'=1}^{d}\sum_{u,v,u',v'=1}^{2}\Bigg(A\Big((H_r,gh,u;G_r,ab,v),(H_s,g'h',u';G_s,a'b',v')\Big)_T$$
$$+A\Big((H_r,gh,u;G_r,ab,v),(H_s,a'b',v';G_s,g'h',u')\Big)_T+A\Big((H_r,ab,v;G_r,gh,u),(H_s,g'h',u';G_s,a'b',v')\Big)_T$$
$$+A\Big((H_r,ab,v;H_r,gh,u),(H_s,a'b',v';G_s,g'h',u')\Big)_T\Bigg).$$

After some simple calculations, the above expression can be rewritten as

$$\sum_{g,h,a,b=1}^{d}\sum_{j,k,l,m=1}^{d}\Bigg(\frac{6}{\theta^3}\int_0^T\left(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s(C_t)\right)\Big[(C_t^{gj}C_t^{hk}+C_t^{gk}C_t^{hj})(C_t^{al}C_t^{bm}+C_t^{am}C_t^{bl})$$
$$+(C_t^{aj}C_t^{bk}+C_t^{ak}C_t^{bj})(C_t^{gl}C_t^{hm}+C_t^{gm}C_t^{hl})\Big]dt$$
$$+\frac{151\theta}{140}\int_0^t\left(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s(C_t)\right)\Big[\overline{C}^{gh,jk}\overline{C}^{ab,lm}+\overline{C}^{ab,jk}\overline{C}^{gh,lm}\Big]dt$$
$$+\frac{3}{2\theta}\int_0^t\left(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s(C_t)\right)\Big[(C_t^{gj}C_t^{hk}+C_t^{gk}C_t^{hj})\overline{C}_t^{ab,lm}+(C_t^{al}C_t^{bm}+C_t^{am}C_t^{bl})\overline{C}_t^{gh,jk}$$
$$+(C_t^{gl}C_s^{hm}+C_t^{gm}C_s^{hl})\overline{C}_t^{ab,jk}+(C_t^{aj}C_t^{bk}+C_t^{ak}C_t^{bj})\overline{C}_t^{gh,lm}\Big]dt\Bigg),$$

which completes the proof.

A.2 Proof of Theorem 2

Using the polynomial growth assumption on H_r, G_r, H_s and G_s and Theorem 2.2 in Jacod and Rosenbaum (2012), one can show that

$$\frac{6}{\theta^3}\widehat{\Omega}_T^{r,s,(1)} \xrightarrow{\mathbb{P}} \Sigma_T^{r,s,(1)}.$$

Next, by equation (3.27) in Jacod and Rosenbaum (2012), we have

$$\frac{3}{2\theta}[\widehat{\Omega}_T^{r,s,(3)} - \frac{6}{\theta}\widehat{\Omega}_T^{r,s,(1)}] \xrightarrow{\mathbb{P}} \Sigma_T^{r,s,(3)}.$$

Finally, to show that

$$\frac{151\theta}{140}\frac{9}{4\theta^2}[\widehat{\Omega}_T^{r,s,(2)} + \frac{4}{\theta^2}\widehat{\Omega}_T^{r,s,(1)} - \frac{4}{3}\widehat{\Omega}_T^{r,s,(3)}] \xrightarrow{\mathbb{P}} \Sigma_T^{r,s,(2)},$$

we first observe that as in Step 1, the approximation error induced by replacing $\widehat{C}_i^n$ by $\widehat{C}_i'^n$ is negligible. For $1 \le g, h, a, b, j, k, l, m \le d$ and $1 \le r, s \le d$, we define

$$\begin{aligned}
\widehat{W}_T^n &= \sum_{i=1}^{[T/\Delta_n]-4k_n+1} (\partial_{gh}H_r\partial_{ab}G_r\partial_{gh}H_s\partial_{lm}G_s)(\widehat{C}_i^n)\lambda_i^{n,gh}\lambda_i^{n,jk}\lambda_{i+2k_n}^{n,ab}\lambda_{i+2k_n}^{n,lm} \\
\widehat{w}(1)_i^n &= (\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(C_i^n)\mathbb{E}(\lambda_i^{n,gh}\lambda_i^{n,jk}\lambda_{i+2k_n}^{n,ab}\lambda_{i+2k_n}^{n,lm}|\mathcal{F}_i^n) \\
\widehat{w}(2)_i^n &= (\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(C_i^n)(\lambda_i^{n,gh}\lambda_i^{n,jk}\lambda_{i+2k_n}^{n,ab}\lambda_{i+2k_n}^{n,lm} - \mathbb{E}(\lambda_i^{n,gh}\lambda_i^{n,jk}\lambda_{i+2k_n}^{n,ab}\lambda_{i+2k_n}^{n,lm}|\mathcal{F}_i^n)) \\
\widehat{w}(3)_i^n &= \Big((\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(\widehat{C}_i^n) - (\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(C_i^n)\Big)\lambda_i^{n,gh}\lambda_i^{n,jk}\lambda_{i+2k_n}^{n,ab}\lambda_{i+2k_n}^{n,lm} \\
\widehat{W}(u)_t^n &= \sum_{i=1}^{[T/\Delta_n]-4k_n+1} \widehat{w}_i(u), \ u = 1, 2, 3.
\end{aligned}$$

Note that we have $\widehat{W}_t^n = \widehat{W}(1)_t^n + \widehat{W}(2)_t^n + \widehat{W}(3)_t^n$. By Taylor expansion and using repeatedly the boundedness of C_t, we have

$$|\widehat{w}(3)_i^n| \le (1 + \|\nu_i^n\|^{4(p-1)})\|\nu_i^n\|\|\lambda_i^n\|^2\|\lambda_{i+2k_n}^n\|^2,$$

which implies $\mathbb{E}(|\widehat{w}(3)_i^n|) \le K\Delta_n^{5/4}$ and $\widehat{W}(3)_t^n \xrightarrow{\mathbb{P}} 0$. Using Cauchy-Schwartz inequality and the bound $\mathbb{E}(\|\lambda_i^n\|^q|\mathcal{F}_i^n) \le K\Delta_n^{q/4}$, we have $\mathbb{E}(|\widehat{w}(2)_i^n|^2) \le K\Delta_n^2$. Observing furthermore that $\widehat{w}(2)_i^n$ is $\mathcal{F}_{i+4k_n}-$measurable, we use Lemma B.8 in Aït-Sahalia and Jacod (2014) to show that $\widehat{W}(2)_t^n \xrightarrow{\mathbb{P}} 0$. Also, define

$$\begin{aligned}
w_i^n &= (\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(C_i^n)\Big[\frac{4}{k_n^2\Delta_n}(C_i^{n,ga}C_i^{n,hb} + C_i^{n,gb}C_i^{n,ha})(C_i^{n,jl}C_i^{n,km} + C_i^{n,jm}C_i^{n,kl}) \\
&+ \frac{4}{3}(C_i^{n,jl}C_i^{n,km} + C_i^{n,jm}C_i^{n,kl})\overline{C}_i^{n,gh,ab} + \frac{4}{3}(C_i^{n,ga}C_i^{n,hb} + C_i^{n,gb}C_i^{n,ha})\overline{C}_i^{n,jk,lm} + \frac{4(k_n^2\Delta_n)}{9}\overline{C}_i^{n,gh,ab}\overline{C}_i^{n,jk,lm}\Big], \\
W_T^n &= \Delta_n \sum_{i=1}^{[T/\Delta_n]-4k_n+1} w_i^n.
\end{aligned}$$

The cadlag property of c and $\overline{C}$, $k_n\sqrt{\Delta_n} \longrightarrow \theta$, and the Riemann integral argument imply $W_T^n \xrightarrow{\mathbb{P}} W_T$ where

$$\begin{aligned}
W_T &= \int_0^T (\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(C_t)\Big[\frac{4}{\theta^2}(C_t^{ga}C_t^{hb} + C_t^{gb}C_t^{ha})(C_t^{jl}C_t^{km} + C_t^{jm}C_t^{kl}) + \frac{4}{3}(C_t^{jl}C_t^{km} + C_t^{jm}C_t^{kl})\overline{C}_t^{gh,ab} \\
&+ \frac{4}{3}(C_t^{ga}C_i^{hb} + C_t^{gb}C_t^{ha})\overline{C}_t^{jk,lm} + \frac{4\theta^2}{9}\overline{C}_t^{gh,ab}\overline{C}_t^{jk,lm}\Big]dt.
\end{aligned}$$

In addition, by Lemma 4, we have

$$\mathbb{E}(|\widehat{W}(1)_T^n - W_T^n|) \le \Delta_n\mathbb{E}\bigg(\sum_{i=1}^{[T/\Delta_n]-4k_n+1}(\Delta_n^{1/8} + \eta_{i,4k_n})\bigg).$$

Hence, by the third result of Lemma 6 we have $\widehat{W}_T^n \xrightarrow{\mathbb{P}} W_t$, from which it can be deduced that

$$\frac{9}{4\theta^2}\Big[\widehat{W}(1)_T^n + \frac{4}{k_n^2}\sum_{i=1}^{[T/\Delta_n]-4k_n+1}(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(\widehat{C}_i^n)[C_i^n(jk,lm)C_i^n(gh,ab)]$$
$$-\frac{2}{k_n}\sum_{i=1}^{[T/\Delta_n]-4k_n+1}(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(\widehat{C}_i^n)C_i^n(gh,ab)\lambda_i^{n,jk}\lambda_i^{n,lm}$$
$$-\frac{2}{k_n}\sum_{i=1}^{[T/\Delta_n]-4k_n+1}(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(\widehat{C}_i^n)C_i^n(jk,lm)\lambda_i^{n,gh}\lambda_i^{n,ab}\Big]$$
$$\xrightarrow{\mathbb{P}}\int_0^T(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(C_t)\overline{C}_t^{gh,ab}\overline{C}_t^{jk,lm}dt.$$

The result follows from the above convergence, a symmetry argument, and straightforward calculations.

B Tables and Figures

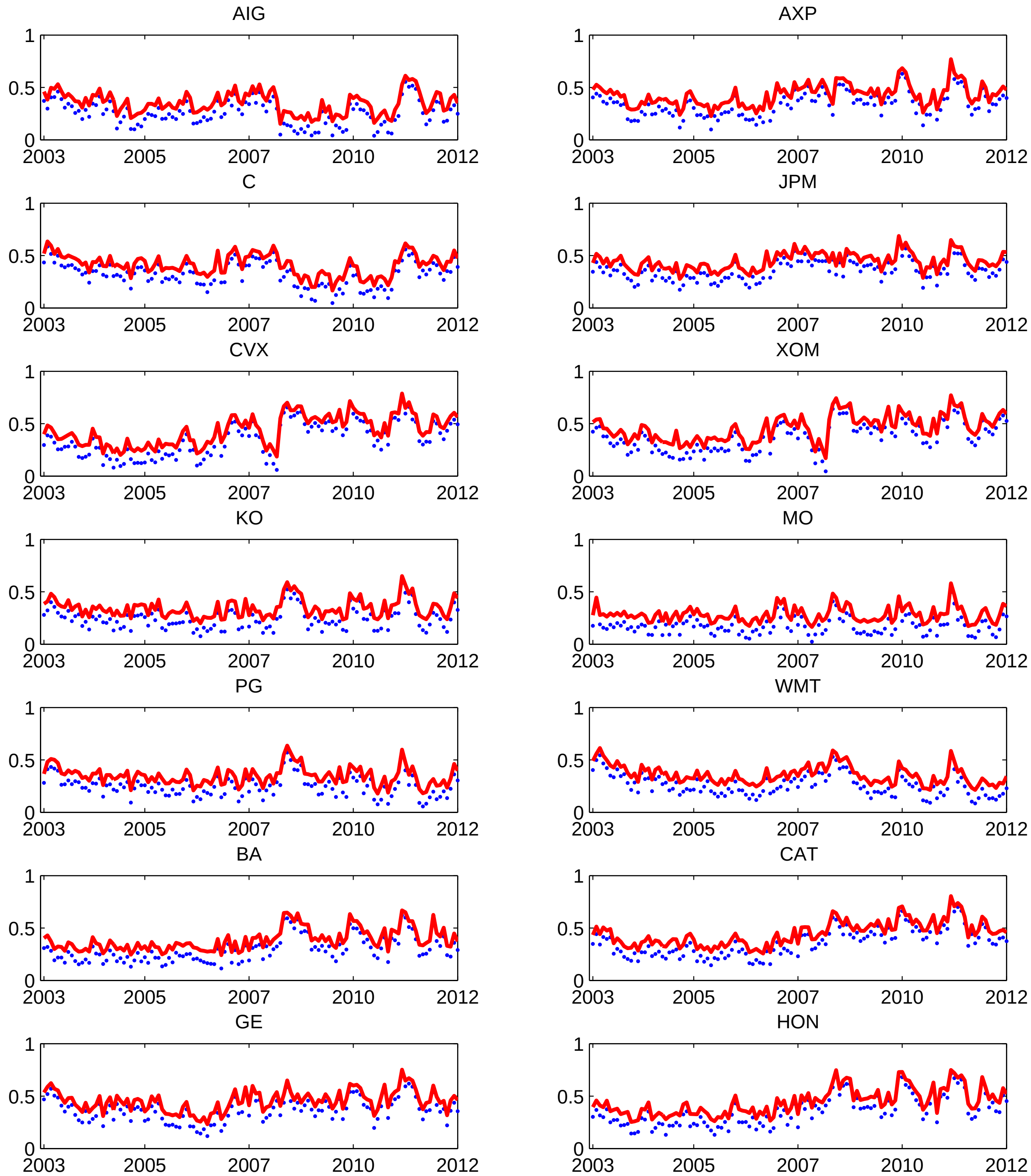

Figure 1: Monthly R^2 of two return factor models ($\widehat{R}^2_{Y,j}$): the CAPM (the blue dotted line) and the Fama-French three factor model (the red solid line). Stocks are represented by tickers (see Table 8 for full stock names).

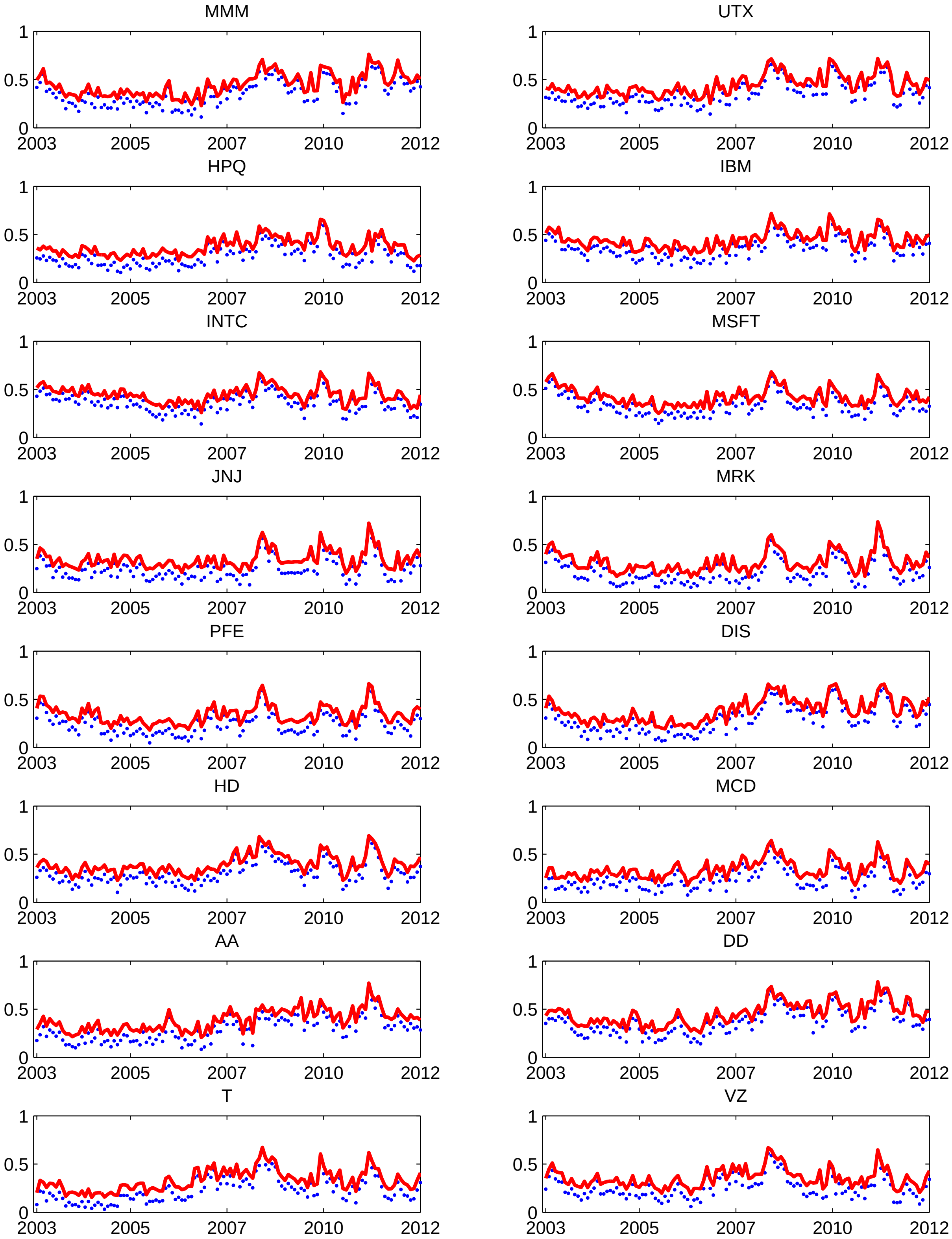

Figure 2: Monthly R^2 of two return factor models ($\widehat{R}^2_{Y_j}$): the CAPM (the blue dotted line) and the Fama-French three factor model (the red solid line). Stocks are represented by tickers (see Table 8 for full stock names).

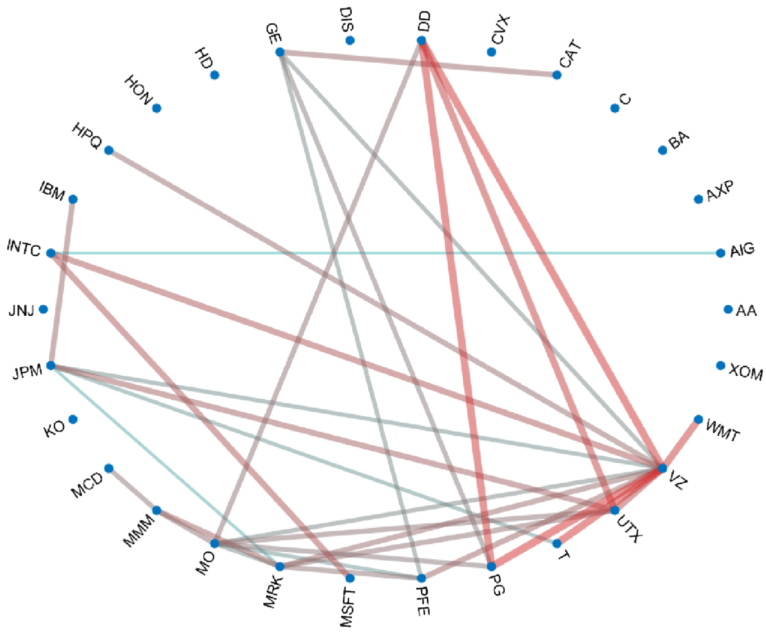

Figure 3: The network of dependencies in total IVs. The color and thickness of each line is proportional to the estimated value of $\rho_{Zi,Zj}$, the quadratic-covariation based correlation between the IVs, defined in equation (7) (red and thick lines indicate high correlation). We simultaneously test 435 null hypotheses of no correlation, and the lines are only plotted when the null is rejected.

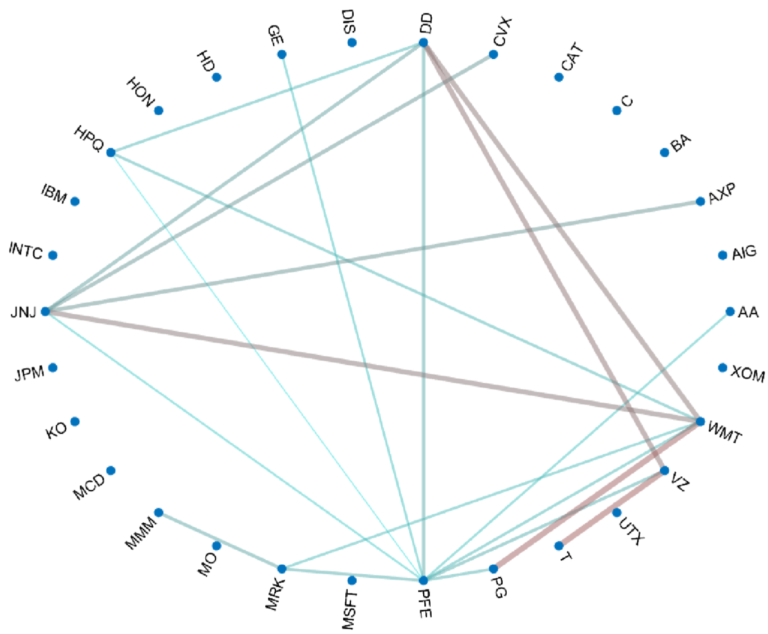

Figure 4: The network of dependencies in residual IVs when the market variance is the only IV factor. The color and thickness of each line is proportional to the estimated value of $\rho^{resid}_{Zi,Zj}$ the quadratic-covariation based correlation between the IVs, defined in equation (8), of each pair of stocks (red and thick lines indicate high correlation). We simultaneously test 435 null hypotheses of no correlation, and the lines are only plotted when the null is rejected.

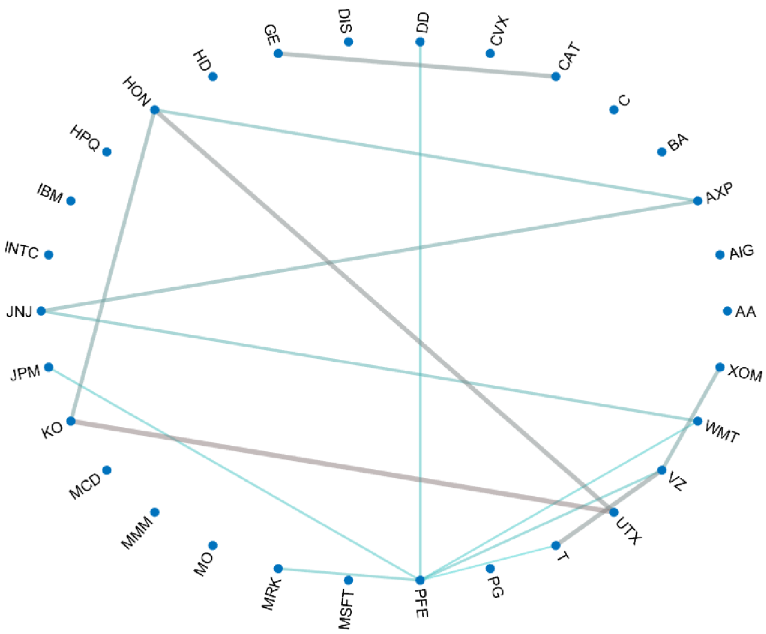

Figure 5: The network of dependencies in residual IVs with ten IV factors: the market variance and the variances of nine industry ETFs. The color and thickness of each line is proportional to the estimated value of $\rho^{resid}_{Zi,Zj}$ *the quadratic-covariation based correlation between the IVs, defined in equation (8), of each pair of stocks (red and thick lines indicate high correlation). We simultaneously test 435 null hypotheses of no correlation, and the lines are only plotted when the null is rejected.*