Cross-sectional Dependence in Idiosyncratic Volatility*

Ilze Kalnina†
North Carolina State University

Kokouvi Tewou‡
Concordia University

First version: February 25, 2015
This version: May 30, 2019

Abstract

This paper introduces an econometric framework for analyzing 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. Naïve estimators of these measures are biased due to the use of the error-laden estimates of idiosyncratic volatilities. We provide bias-corrected estimators and the relevant asymptotic theory. Next, we introduce an idiosyncratic volatility factor model, in which we decompose the variation in idiosyncratic volatilities into two parts: the variation related to the systematic factors such as the market volatility, and the residual variation. Again, naïve estimators of the decomposition are biased, and we provide bias-corrected estimators. We also provide the asymptotic theory that allows us to test whether the residual (non-systematic) components of the idiosyncratic volatilities exhibit cross-sectional dependence. 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 (non-systematic) idiosyncratic volatilities across the stocks.

Keywords: factor model, systematic risk, networks of risk, residual idiosyncratic volatility, 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, Silvia Gonçalves, Peter Hansen, Jean Jacod, Dennis Kristensen, Joon Park, Benoit Perron, and Dacheng Xiu. We thank seminar participants at University of Amsterdam, Bank of Canada, Concordia, HEC Montreal, Indiana, LSE, McGill, NC State, Pennsylvania, Surrey, Toulouse, UCL, Warwick, Western Ontario, as well as participants of various conferences.

†Department of Economics, North Carolina State University. E-mail address: ilze_kalnina@ncsu.edu. Kalnina’s research was supported by IFSID. She is grateful to UCL and CeMMaP for their hospitality and support. She is also grateful to the Economics Department, the Gregory C. Chow Econometrics Research Program, and the Bendheim Center for Finance at Princeton University for their hospitality.

‡Department of Economics, Concordia University. E-mail address: kokouvi.tewou@concordia.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, such as the Fama-French factors. In this return factor model (R-FM), the total volatility of an asset return can be decomposed into two parts: a component due to the exposure to the common return factors (the systematic volatility), and a residual component termed the idiosyncratic volatility (IdioVol). 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)).\textsuperscript{1} Also, idiosyncratic volatility measures the exposure to the idiosyncratic risk in imperfectly diversified portfolios. A recent observation is that the IdioVols seem to be strongly correlated in the cross-section of stocks.\textsuperscript{2} Herskovic, Kelly, Lustig, and Nieuwerburgh (2016) argue this is due to a common IdioVol factor, which they relate to household risk. Moreover, cross-sectional dependence in IdioVols is important for option pricing, see Gourier (2016).

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

To study idiosyncratic volatilities, we introduce the idiosyncratic volatility factor model (IdioVol-FM). Just like a return factor model, R-FM, such as the Fama-French model, decomposes returns into common and idiosyncratic returns, the IdioVol-FM decomposes the IdioVols into systematic and residual (non-systematic) components. The IdioVol factors may or may not be related to the return factors. The IdioVol factors can include the volatility of the return factors, or,

\textsuperscript{1}A stock is said to be mispriced with respect to a given model if the expected value of the return on the stock is not consistent with the model.

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). We propose bias-corrected estimators of the components of the IdioVol-FM model.

We provide the asymptotic theory for this model. For example, it allows us to test whether the residual (non-systematic) components of the IdioVols exhibit cross-sectional dependence. This allows us to identify the network of dependencies in the residual IdioVols across stocks.

Our bias-corrected estimators and inference results are an application of a new asymptotic theory that we develop for general estimators of quadratic covariation of vector-valued transformations of spot covariance matrices. This theoretical contribution is of its own interest. An example of alternative applications is the study of cross-sectional dependence of asset betas. Two features make the development of this asymptotic theory difficult. First, preliminary estimation of volatility results in first-order biases even in the special case of quadratic variation of the volatility one stock without any transformations, as in Vetter (2015). Second, we consider general nonlinear functionals in multivariate settings, which substantially complicates the analysis.

Throughout the paper, we use factors that are specified by the researcher. An example of our Return Factor Model is the so-called Fama-French factor model, which has three observable factors, or the CAPM, which has one observable factor (the market portfolio return). An example of our IdioVol factors is the market volatility, which can be estimated from the market index. Thus, our setup is different from settings such as PCA where factors are identified from the cross-section of the assets studied. The treatment of the latter case adds an additional layer of complexity to the model and is beyond the scope of the current paper.

We apply our methodology to high-frequency data on the 30 Dow Jones Industrial Average components. We study the IdioVols 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 IdioVols 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 IdioVol-FM. We use two sets of IdioVol factors: the market volatility alone and the market volatility together with

---

3The high frequency Fama-French factors are provided by Aït-Sahalia, Kalnina, and Xiu (2019).
volatilities of nine industry ETFs. With the market volatility as the only IdioVol factor, the average pairwise correlation between residual (non-systematic) IdioVols is substantially lower (0.25) than between the total IdioVols. With the additional nine industry ETF volatilities as IdioVol factors, average correlation between the residual IdioVols decreases further (to 0.18). However, neither of the two sets of the IdioVol factors can fully explain the cross-sectional dependence in the IdioVols. We map out the network of dependencies in residual IdioVols across all stocks.

This paper analyzes 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 (2006). A continuous-time factor model for asset returns with observable return factors was first studied in Mykland and Zhang (2006). Various return factor models with observable factors have been studied by, among others, Bollerslev and Todorov (2010), Fan, Furger, and Xiu (2016), Li, Todorov, and Tauchen (2017a,b), and Aït-Sahalia, Kalnina, and Xiu (2019). Emerging literature also studies the cross-sectional dependence in returns using high-frequency data and latent return factors, see Aït-Sahalia and Xiu (2019, 2017) and Pelger (2019a,b). Importantly, the models in the above papers are silent on the cross-sectional dependence structure in the IdioVols.

The Realized Beta GARCH model of Hansen, Lunde, and Voev (2014) imposes a structure on the cross-sectional dependence in IdioVols. This structure is tightly linked with the return factor model parameters, whereas our stochastic volatility framework allows separate specification of the return factors and the IdioVol factors.4

Our inference theory is related to several results in the existing literature. First, as mentioned above, we generalize the result of Vetter (2015). Jacod and Rosenbaum (2013, 2015), Li, Todorov, and Tauchen (2016) and Li, Liu, and Xiu (2019) estimate integrated functionals of volatilities, which includes idiosyncratic volatilities. The latter problem is simpler than the problem of the current paper 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

---

4In the Beta GARCH model, the IdioVol of a stock is a product of its own (total) volatility, and one minus the square of the correlation between the stock return and the market return.
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 (2017), Kalnina and Xiu (2017) 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 (2018), and Kanaya and Kristensen (2016).

In the empirical section, we define a network of dependencies using (functions of) quadratic covariations of IdioVols. This approach can be compared with the network connectedness measures of Diebold and Yilmaz (2014). The latter measures are based on forecast error variance decompositions from vector autoregressions. They capture co-movements in forecast errors. In contrast, we assume a general semimartingale setting, and our framework captures realized co-movements in idiosyncratic volatilities, while accounting for the measurement errors in these volatilities.

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 uses high-frequency stock return data to study the cross-sectional dependence in IdioVols using our framework. Section 6 contains Monte Carlo simulations. The Appendix contains all proofs and additional figures.

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 (IdioVol-FM). In this framework, we proceed to define the cross-sectional measures of dependence between the total IdioVols, as well as the residual IdioVols, which take into account the dependence induced by the IdioVol 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 semimartin-
gale, i.e., $Y_t$ follows

$$Y_t = Y_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t,$$  \hspace{1cm} (1)

where $W$ is a $dW$-dimensional Brownian motion ($dW \geq d$), $\sigma_s$ is a $d \times dW$ 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,\(^5\)

$$C_t = C_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dW_s.$$  \hspace{1cm} (2)

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 $Y$, and $C_{U,t}$ to refer to $C_{UU,t}$.

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

**Definition (Factor Model for Returns, R-FM).** For all $0 \leq t \leq T$ and $j = 1, \ldots, d_S$,\(^6\)

$$dS_{j,t} = \beta_{j,t}^c dF_{c,t} + \tilde{\beta}_{j,t}^d dF_{d,t} + dZ_{j,t} \text{ with } [Z_{j,t}, F]_t = 0.$$

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

\(^5\)Note that assuming that $Y$ and $C$ are driven by the same $dW$-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).

\(^6\)The quadratic covariation of two vector-valued Itô semimartingales $X$ and $Y$, over the time span $[0, T]$, is defined as

$$[X, Y]_T = p\lim_{M \to \infty} \sum_{s=0}^{M-1} (X_{t_{s+1}} - X_{t_s})(Y_{t_{s+1}} - Y_{t_s})^\top,$$

for any sequence $t_0 < t_1 < \ldots < t_M = T$ with $\sup_{s} \{t_{s+1} - t_s\} \to 0$ as $M \to \infty$, where $p\lim$ stands for the probability limit.
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 Ait-Sahalia, Kalnina, and Xiu (2019).

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

We define the idiosyncratic Volatility (IdioVol) as the spot volatility of the $Z_{j,t}$ process and
denote it by $C_{Zj}$. Notice that the R-FM in (3) implies that the factor loadings $\beta_t$ as well as IdioVol
are functions of the total spot covariance matrix $C_t$. In particular, the vector of factor loadings satisfies

$$\beta_{jt} = (C_{F,t})^{-1}C_{FSj,t},$$

for $j = 1, \ldots, d_S$, where $C_{F,t}$ 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 IdioVol of stock $j$ is also
a function of the total spot covariance matrix $C_t$,

$$C_{Zj,t} = C_{Yj,t} - (C_{FSj,t})^\top (C_{F,t})^{-1}C_{FSj,t}.$$  

By the Itô lemma, (4) and (5) imply that factor loadings and IdioVols are also Itô semimartingales with their characteristics related to those of $C_t$.

We now introduce the Idiosyncratic Volatility Factor model (IdioVol-FM). In the IdioVol-FM,
the cross-sectional dependence in the IdioVol shocks can be potentially explained by certain IdioVol
factors. We assume the IdioVol factors are known functions of the matrix $C_t$. In the empirical
application, we use the market volatility as the IdioVol factor, which has been used in Herskovic,
Kelly, Lustig, and Nieuwerburgh (2016) and Gourier (2016); we discuss other possibilities below.
We allow the IdioVol 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,

$$\mathcal{G}(p) = \{ H : H \text{ is three-times continuously differentiable and for some } K > 0, \| \partial^j H(x) \| \leq K(1 + \| x \|)^{p-j}, j = 0, 1, 2, 3 \}, \text{ for some } p \geq 3. \quad (6)$$

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

$$dC_{Zj,t} = \gamma_{Zj}^\top d\Pi_t + dC_{Zj,t}^{\text{resid}} \text{ with}$$

$$[C_{Zj}^{\text{resid}}, \Pi]_t = 0, \quad (7)$$

where $\Pi_t = (\Pi_{1t}, \ldots, \Pi_{d\Pi t})$ is a $\mathbb{R}^{d\Pi}$-valued vector of IdioVol 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 $C_{Zj,t}^{\text{resid}}$ the residual IdioVol of asset $j$. Our assumptions imply that the components of the IdioVol-FM, $C_{Zj,t}, \Pi_t$ and $C_{Zj,t}^{\text{resid}}$, are continuous Itô semimartingales. We remark that both the dependent variable and the regressors in our IdioVol-FM are not directly observable and have to be estimated, and our asymptotic theory takes that into account. 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 IdioVol factors permitted by our theory is rather wide as it includes general non-linear transforms of the spot covolatility process $C_t$. For example, IdioVol 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 IdioVol factors are linear combinations of the IdioVols, such as the equally-weighted average of the IdioVols, which Herskovic, Kelly, Lustig, and Nieuwerburgh (2016) denote by the “CIV”. The IdioVol 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) IdioVols. Their estimation is discussed in Section 3.
Before considering the effect of IdioVol factors by using the IdioVol-FM decomposition, one may be interested in quantifying the dependence between the IdioVols of two stocks \( j \) and \( s \). A natural measure of dependence is the quadratic-covariation based correlation between the two IdioVol processes,

\[
\text{Corr} (C_{Zj}, C_{Zs}) = \frac{[C_{Zj}, C_{Zs}]_T}{\sqrt{[C_{Zj}, C_{Zj}]_T} \sqrt{[C_{Zs}, C_{Zs}]_T}}.
\] (8)

Alternatively, one may consider the quadratic covariation \([C_{Zj}, C_{Zs}]_T\) 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 IdioVols.

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

\[
\text{Corr} \left( C^{\text{resid}}_{Zj}, C^{\text{resid}}_{Zs} \right) = \frac{[C^{\text{resid}}_{Zj}, C^{\text{resid}}_{Zs}]_T}{\sqrt{[C^{\text{resid}}_{Zj}, C^{\text{resid}}_{Zj}]_T} \sqrt{[C^{\text{resid}}_{Zs}, C^{\text{resid}}_{Zs}]_T}}.
\] (9)

In Section 4.4, we use the quadratic covariation between the two residual IdioVol processes \([C^{\text{resid}}_{Zj}, C^{\text{resid}}_{Zs}]_T\) without normalization for testing purposes.

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

\[
R^2_{\text{IdioVol-FM}} = \frac{\gamma_{Zj}^T |\Pi| \gamma_{Zj}}{[C_{Zj}, C_{Zj}]_T}.
\] (10)

It is interesting to compare the correlation measure between IdioVols in equation (8) with the correlation between the residual parts of IdioVols in (9). We consider their difference,

\[
\text{Corr} (C_{Zj}, C_{Zs}) - \text{Corr} \left( C^{\text{resid}}_{Zj}, C^{\text{resid}}_{Zs} \right)
\] (11)

to see how much of the dependence between IdioVols can be attributed to the IdioVol 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 IdioVols is the following quantity,

\[ Q_{Z_j, Z_s}^{IdioVol-FM} = \frac{\gamma_{Z_j}^\top [\Pi, \Pi]^T \gamma_{Z_s}}{[C_{Z_j}, C_{Z_s}]_T}. \]  

(12)

This measure is bounded by 1 if the covariations between residual IdioVols are nonnegative and smaller than the covariations between IdioVols, 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 (2016), \( H_0 : C_{Z_j, t} = a_{Z_j} + \gamma_{Z_j}^\top \Pi_t, 0 \leq t \leq T \). This \( H_0 \) implies that the IdioVol is a deterministic function of the factors, which does not allow for an error term. In particular, this null hypothesis implies \( R_{Z_j}^{2, IdioVol-FM} = 1 \).

3 Estimation

We now discuss the problem of the identification and estimation of the quantities of interest introduced in Section 2. We do so by showing that this problem is a special case of a more general problem, which is of its own interest, and solving the latter problem.

For the identification, the strategy is to show that each of the quantities of interest introduced in Section 2,

\[ [C_{Z_j}, C_{Z_s}]_T, \text{ Corr} (C_{Z_j}, C_{Z_s}), \gamma_{Z_j}, [C_{Z_j}^{resid}, C_{Z_j}^{resid}]_T, \text{ Corr} (C_{Z_j}^{resid}, C_{Z_s}^{resid}), Q_{Z_j, Z_s}^{IdioVol-FM}, \text{ and } R_{Z_j}^{2, IdioVol-FM}, \]

(13)

for \( j, s = 1, \ldots, d_S \), can be written as

\[ \varphi ([H_1(C), G_1(C)]_T, \ldots, [H_\kappa(C), G_\kappa(C)]_T), \]

(14)

where \( \varphi \) as well as \( H_r \) and \( G_r \), for \( r = 1, \ldots, \kappa \), are known real-valued functions. Each element in (14) is of the form \( [H(C), G(C)]_T \), i.e., it is a quadratic covariation between functions of \( C_t \). \( [H(C), G(C)]_T \) is observable from continuous-record observations on \( Y \) in (1), which means it can be estimated from (discrete) high-frequency observations on \( Y \).
While the identification is relatively simple, the estimation problem has to address the biases due to preliminary estimation of (idiosyncratic) volatility. To this end, we introduce two estimators of $[H(C), G(C)]_T$. Section 4 derives the joint asymptotic distribution of several objects of this type, $[H_r(C), G_r(C)]_T$ for $r = 1, \ldots, \kappa$. The asymptotic distribution of the general estimand in (14), and hence of every quantity of interest in equation (13), follows by the Delta method.

We start by discussing the identification of the first estimand in (13), which is the quadratic covariation between $j$th and $s$th IdioVol, $[C_{Zj}, C_{Zs}]_T$. It can be written as $[H(C), G(C)]_T$ if we choose $H(C_t) = C_{Zj,t}$ and $G(C_t) = C_{Zs,t}$. By (5), both $C_{Zj,t}$ and $C_{Zs,t}$ are functions of $C_t$. Next, consider $\text{Corr} (C_{Zj}, C_{Zs})$ defined in (8). 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, $\text{Corr} (C_{Zj}, C_{Zs})$ is itself a known function of three objects of the form $[H(C), G(C)]_T$.

To show that the remaining quantities in (13) can also be expressed in terms of objects of the form $[H(C), G(C)]_T$, note that the IdioVol-FM implies

$$\gamma_{Zj} = ([\Pi, \Pi]_T)^{-1} [\Pi, C_{Zj}]_T \text{ and } [C^\text{resid}_{Zj}, C^\text{resid}_{Zs}]_T = [C_{Zj}, C_{Zs}]_T - \gamma_{Zj} [\Pi, \Pi]_T \gamma_{Zs},$$

for $j, s = 1, \ldots, d_S$. Since $C_{Zj,t}$, $C_{Zs,t}$ and every element in $\Pi_t$ are real-valued functions of $C_t$, the above equalities imply that all quantities of interest in (13) can be written as real-valued, known functions of a finite number of quantities of the form $[H(C), G(C)]_T$.

We now turn to the estimation of $[H(C), G(C)]_T$. Suppose we have discrete observations on $Y_t$ over an interval $[0, T]$. Denote by $\Delta_n$ 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,

$$\hat{C}_{i\Delta_n} = \frac{1}{k_n\Delta_n} \sum_{m=0}^{k_n-1} (\Delta_{i+m}^n Y)(\Delta_{i+m}^n Y)^\top 1_{\{\|\Delta_{i+m}^n Y\| \leq \chi \Delta_n^p\}}, \quad (15)$$

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.

Throughout the paper we set $\hat{C}_{i\Delta_n} = (\hat{C}_{ab,i\Delta_n})_{1 \leq a, b \leq d}$.\footnote{It is also possible to define more flexible kernel-based estimators as in Kristensen (2010).}

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

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

(16)

where the factor $3/2$ and the last term correct for the biases arising due to the preliminary estimation of volatility $C_t$.

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

\[
[\hat{H}(C), \hat{G}(C)]_{T} = \sum_{g,h,a,b=1}^{d} \int_{0}^{T} (\partial_{gh} H \partial_{ab} G)(C_t) \overline{C}_{t}^{gh,ab} dt,
\]

(17)

where $\overline{C}_{t}^{gh,ab}$ 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 second estimator is a bias-corrected version of the sample counterpart of the “linearized” expression in (17),

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

(18)

Consistency for a similar estimator has been established by Jacod and Rosenbaum (2015).\(^8\) We go beyond their result by deriving the asymptotic distribution and proposing a consistent estimator of its asymptotic variance.

If we had observations on $C_{i\Delta_n}$, the estimators of $[\hat{H}(C), \hat{G}(C)]_{T}$ would not need any bias-correction terms. However, due to the replacement of $C_{i\Delta_n}$ by its estimate $\hat{C}_{i\Delta_n}$, two types of

\[^8\text{Jacod and Rosenbaum (2015) derive the probability limit of the following estimator:}\]

\[
\frac{3}{2k_n} \sum_{g,h,a,b=1}^{d} \sum_{i=1}^{[T/\Delta_n] - 2k_n + 1} (\partial_{gh,ab} H)(\hat{C}_{i\Delta_n}) \left( (\hat{C}_{(i+k_n)\Delta_n} - \hat{C}_{i\Delta_n})(\hat{C}_{(i+k_n)\Delta_n} - \hat{C}_{i\Delta_n}) - \frac{2}{k_n} (\hat{C}_{ga,i\Delta_n} \hat{C}_{gb,i\Delta_n} + \hat{C}_{gb,i\Delta_n} \hat{C}_{ha,i\Delta_n}) \right).
\]
bias-correction terms arise: a multiplicative correction $3/2$, as well as an additive bias-correction term

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

(19)

We remark that this additive bias correction 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 (2015), which was extended to allow for jumps in Jacod and Rosenbaum (2015). 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 IdioVols, 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 $F_t$.

**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 $\mu$ 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 $\sigma$-
finite measure $\lambda$ on $E$. The process $b_t$ is $\mathbb{R}^d$-valued optional, $\sigma_t$ is $\mathbb{R}^d \times \mathbb{R}^{dW}$-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 $(\tau_m)$ is a localizing sequence of stopping times and, for some $r \in [0, 1]$, the function $\Gamma_m$ 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,$^9$ 

$$C_t = C_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dW_s. \tag{20}$$

where $\tilde{b}$ 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 co-
variation of the continuous martingale parts of $X$ and $c$ are defined as follows,

$$C_{gh,ab}^t = \sum_{m=1}^{dW} \tilde{\sigma}_{g,m}^h \tilde{\sigma}_{a,m}^b, \quad C_{g,ab}^t = \sum_{m=1}^{dW} \tilde{\sigma}_{g,m}^a \tilde{\sigma}_{b,m}^b. \tag{21}$$

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

**Assumption 2.** $\tilde{\sigma}_t$ 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 (2015), Kalnina and Xiu (2017) and Wang and Mykland (2014).

### 4.2 Asymptotic Distribution

We have seen in Section 3 that all quantities of interest in (13) 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. The current section presents this asymptotic distribution.

$^9$Note that $\tilde{\sigma}_s = (\tilde{\sigma}_s^{gh,m})$ is $(d \times d \times dW)$-dimensional and $\tilde{\sigma}_s dW_s$ is $(d \times d)$-dimensional with $(\tilde{\sigma}_s dW_s)^{gh} = \sum_{m=1}^{dW} \tilde{\sigma}_s^{gh,m} dW_s^m$. 

14
Let $H_1, G_1, \ldots, H_k, G_n$ be some arbitrary elements of $\mathcal{G}(p)$ defined in equation (6). We are interested in the asymptotic behavior of vectors

$$\left( [H_1(C), G_1(C)]^T, \ldots, [H_n(C), G_n(C)]^T \right)^{AN} \quad \text{and} \quad \left( [H_1(C), G_1(C)]^T, \ldots, [H_n(C), G_n(C)]^T \right)^{LIN}.$$

The following theorem summarizes the joint asymptotic behavior of the estimators.

**Theorem 1.** Let $[H_r(C), G_r(C)]_T$ be either $[H_r(C), G_r(C)]^T$ or $[H_r(C), G_r(C)]^T$ defined in (16) and (18), respectively. Suppose Assumption 1 and Assumption 2 hold. Fix $k = \theta \Delta_n^{-1/2}$ for some $\theta \in (0, \infty)$ and set $(8p-1)/4(4p-r) \leq \varpi < \frac{1}{2}$. Then, as $\Delta_n \to 0$,

$$\Delta_n^{-1/4} \begin{bmatrix}
[H_1(C), G_1(C)]_T - [H_1(C), G_1(C)]_T \\
\vdots \\
[H_n(C), G_n(C)]_T - [H_n(C), G_n(C)]_T
\end{bmatrix} \xrightarrow{L^\infty} MN(0, \Sigma_T), \quad (22)$$

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

$$\Sigma^{r,s}_T = \Sigma^{r,s,(1)}_T + \Sigma^{r,s,(2)}_T + \Sigma^{r,s,(3)}_T,$$

$$\Sigma^{r,s,(1)}_T = \frac{6}{\theta^3} \sum_{g, h, a, b = 1}^d \sum_{j, k, l, m = 1}^d \int_0^T \left( \partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s (C_t) \right) \left[ C_t(ab) C_t(gh, lm) \right] dt,$$

$$\Sigma^{r,s,(2)}_T = \frac{151 \theta}{140} \sum_{g, h, a, b = 1}^d \sum_{j, k, l, m = 1}^d \int_0^T \left( \partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s (C_t) \right) \left[ \overline{C}_t^{gh, jk} \overline{C}_t^{ab, lm} \right] \left[ \overline{C}_t^{ab, jk} \overline{C}_t^{gh, lm} \right] dt,$$

$$\Sigma^{r,s,(3)}_T = \frac{3}{20} \sum_{g, h, a, b = 1}^d \sum_{j, k, l, m = 1}^d \int_0^T \left( \partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s (C_t) \right) \left[ C_t(ab) \overline{C}_t^{gh, jk} + C_t(gh, lm) \overline{C}_t^{ab, jk} + C_t(ab, jk) \overline{C}_t^{gh, lm} \right] dt,$$
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 \( \theta \) 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 \( \theta \) in a Monte Carlo experiment (see Section 6).

### 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:

\[
\tilde{\Omega}_{T}^{r,s,(1)} = \Delta_n \sum_{g,h,a,b=1}^{d} \sum_{j,k,i,m=1}^{d} (\partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s (\tilde{C}_i \Delta_n)) \times [\tilde{C}_i \Delta_n (gh, jk) \tilde{C}_i \Delta_n (ab, lm) + \tilde{C}_i \Delta_n (ab, jk) \tilde{C}_i \Delta_n (gh, lm)],
\]

\[
\tilde{\Omega}_{T}^{r,s,(2)} = \Delta_n \sum_{g,h,a,b=1}^{d} \sum_{j,k,i,m=1}^{d} (\partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s (\tilde{C}_i \Delta_n)) \times [\tilde{C}_i \Delta_n (gh, jk) \tilde{C}_i \Delta_n (ab, lm) + \tilde{C}_i \Delta_n (ab, jk) \tilde{C}_i \Delta_n (gh, lm)],
\]

\[
\tilde{\Omega}_{T}^{r,s,(3)} = \Delta_n \sum_{g,h,a,b=1}^{d} \sum_{j,k,i,m=1}^{d} (\partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s (\tilde{C}_i \Delta_n)) \times [\tilde{C}_i \Delta_n (gh, jk) \tilde{C}_i \Delta_n (ab, lm) + \tilde{C}_i \Delta_n (ab, jk) \tilde{C}_i \Delta_n (gh, lm)],
\]

with \( \tilde{n}_{jk} = \tilde{C}_{ij+k} - \tilde{C}_{ij} \) and \( \tilde{C}_i \Delta_n (gh, jk) = (\tilde{C}_{gj,i} \Delta_n \tilde{C}_{hk,i} \Delta_n + \tilde{C}_{gk,i} \Delta_n \tilde{C}_{hj,i} \Delta_n) \). The following result holds,
Theorem 2. Suppose the assumptions of Theorem 1 hold, then, as $\Delta_n \to 0$
\begin{align*}
\frac{6}{\theta^3} \hat{\Omega}^{r,s,(1)}_{T} & \xrightarrow{P} \hat{\Sigma}^{r,s,(1)}_{T} \\
\frac{3}{2\theta} \hat{\Omega}^{r,s,(3)}_{T} - \frac{6}{\theta^2} \hat{\Omega}^{r,s,(1)}_{T} & \xrightarrow{P} \hat{\Sigma}^{r,s,(3)}_{T} \\
\frac{151\theta}{140} \left[ \hat{\Omega}^{r,s,(2)}_{T} + \frac{4}{3\theta^2} \hat{\Omega}^{r,s,(1)}_{T} - \frac{4}{3} \hat{\Omega}^{r,s,(3)}_{T} \right] & \xrightarrow{P} \hat{\Sigma}^{r,s,(2)}_{T}.
\end{align*}

The estimated matrix $\hat{\Sigma}_T$ is symmetric but is not guaranteed to be positive semi-definite. By Theorem 1, $\hat{\Sigma}_T$ 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 (2015) and a straightforward extension of Theorem 1 can be used to show that the rate of convergence in equation (23) is $\Delta_n^{-1/2}$, and the rate of convergence in (25) is $\Delta_n^{-1/4}$. The rate of convergence in (24) can be shown to be $\Delta_n^{-1/4}$.

Remark 2: In the one-dimensional case ($d = 1$), much simpler estimators of $\Sigma_{r,s}^{(2)}$ can be constructed using the quantities $\hat{\lambda}_i^{n,jk} \hat{\lambda}_i^{n,lm} \hat{\lambda}_i^{n,gh} \hat{\lambda}_i^{n,xy}$ or $\hat{\lambda}_i^{n,jk} \hat{\lambda}_i^{n,lm} \hat{\lambda}_i^{n,gh} \hat{\lambda}_i^{n,xy}$ as in Vetter (2015). However, in the multidimensional case, the latter quantities do not identify separately the quantity $C_{t}^{jk,lm}C_{t}^{gh,xy}$ since the combination $C_{t}^{jk,lm}C_{t}^{gh,xy} + C_{t}^{jk,gh}C_{t}^{lm,xy} + C_{t}^{jk,xy}C_{t}^{gh,lm}$ shows up in a non-trivial way in the limit of the estimator.

Corollary 3. For $1 \leq r \leq \kappa$, let $[H_r(C), G_r(C)]_T$ be either $[H_r(C), G_r(C)]_{AN}^T$ or $[H_r(C), G_r(C)]_{LIN}^T$ defined in (18) and (16), respectively. Suppose the assumptions of theorem 1 hold. Then,
\begin{equation}
\Delta_n^{-1/4} \hat{\Sigma}_T^{-1/2} \begin{pmatrix}
[H_1(C), G_1(C)]_T - [H_1(C), G_1(C)]_T \\
\vdots \\
[H_\kappa(C), G_\kappa(C)]_T - [H_\kappa(C), G_\kappa(C)]_T
\end{pmatrix} \xrightarrow{L} N(0, I_\kappa).
\end{equation}

In the above, we use the notation $L$ to denote the convergence in distribution and $I_\kappa$ the identity matrix of order $\kappa$. 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
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 IdioVols of the returns on assets \(j\) and \(s\),

\[
H^1_0 : [C_{Zj}, C_{Zs}]_T = 0 \; \text{vs} \; H^1_1 : [C_{Zj}, C_{Zs}]_T \neq 0.
\]

The null hypothesis \(H^1_0\) is rejected whenever the t-test exceeds the \(\alpha/2\)-quantile of the standard normal distribution, \(Z_{\alpha/2}\),

\[
\Delta_n^{-1/4} \frac{[C_{Zj}, C_{Zs}]_T}{\sqrt{\text{AVAR}(C_{Zj}, C_{Zs})}} > Z_{\alpha/2}.
\]

Second, we can test for all the IdioVol factors \(\Pi\) being irrelevant to explain the dynamics of IdioVol shocks of stock \(j\),

\[
H^2_0 : [C_{Zj}, \Pi]_T = 0 \; \text{vs} \; H^2_1 : [C_{Zj}, \Pi]_T \neq 0.
\]

Under this null hypothesis, the vector of IdioVol factor loadings equals zero, \(\gamma_{Zj} = 0\). The null hypothesis \(H^2_0\) is rejected when

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

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

\[ H_0^3 : [\hat{C}_{Zj}^{\text{resid}}, \hat{C}_{Zs}^{\text{resid}}]_T = 0 \text{ vs } H_1^3 : [\hat{C}_{Zj}^{\text{resid}}, \hat{C}_{Zs}^{\text{resid}}]_T \neq 0. \]

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

\[
\Delta_n^{-1/4} \frac{\left| [\hat{C}_{Zj}^{\text{resid}}, \hat{C}_{Zs}^{\text{resid}}]_T \right|}{\sqrt{\hat{\text{AVAR}}(\hat{C}_{Zj}^{\text{resid}}, \hat{C}_{Zs}^{\text{resid}})}} > Z_{\alpha/2}. \tag{28}
\]

Each of the above estimators

\[ [\hat{C}_{Zj}, \hat{C}_{Zs}]_T, [\hat{C}_{Zj}, \hat{\Pi}]_T, \text{ and } [\hat{C}_{Zj}^{\text{resid}}, \hat{C}_{Zs}^{\text{resid}}]_T \]

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

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

For the first two tests, the expression for the true asymptotic variance, \( \text{AVAR} \), 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, \( \hat{\text{AVAR}} \), 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 \( \alpha \), and their power approaches 1. The same properties apply for the tests of the null hypotheses \( H_0^3 \) with our R-FM and IdioVol-FM representations.

Theoretically, it is possible to test for absence of dependence in the IdioVols at each point in time. In this case the null hypothesis is \( H_0^{1'} : [\hat{C}_{Zj}, \hat{C}_{Zs}]_t = 0 \) for all \( 0 \leq t \leq T \), which is, in theory, stronger than our \( H_0^1 \). In particular, Theorem 1 can be used to set up Kolmogorov-Smirnov type of tests for \( H_0^{1'} \) in the same spirit as Vetter (2015). 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 IdioVols, which means that in practice, it is not too restrictive to assume \([\hat{C}_{Zj}, \hat{C}_{Zs}]_t \geq 0 \forall t\), under which \( H_0^1 \) and \( H_0^{1'} \) are
5 Empirical Analysis

We apply our methods to study the cross-sectional dependence in IdioVol 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 1. 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 (2019). 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 (2015).

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 (\( 3\hat{\sigma}_t \Delta_n^{0.49} \) where \( \hat{\sigma}_t^2 \) is the bipower variation).

Figures D.1 and D.2 contain plots of the time series of the estimated \( R_{Y,j}^2 \) of the return factor model (R-FM) for each stock.\(^\text{10}\) Each plot contains monthly \( R_{Y,j}^2 \) from two return factor models,\(^\text{10}\)

\(^{10}\)For the \( j^{th} \) stock, our analog of the coefficient of determination in the R-FM is \( R_{Y,j}^2 = 1 - \frac{\int \hat{f}_{Y,j}(t) - \hat{f}_{Y,j}(\hat{Y}_{Y,j}^0)\, \text{d}t}{\int \hat{f}_{Y,j}(\hat{Y}_{Y,j}^0)\, \text{d}t} \). We estimate \( R_{Y,j}^2 \) using the general method of Jacod and Rosenbaum (2013). The resulting estimator of \( R_{Y,j}^2 \) requires a choice of a
CAPM and the Fama-French regression with market, size, and value factors. Figures D.1 and D.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 $R^2_{I_{YJ}}$ indicates that the systematic risk is relatively more important, which is typical during crises. $R^2_{I_{YJ}}$ 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 IdioVols, $\text{Corr} (C_{Zi}, C_{Zj})$. All pairwise correlations are positive in our sample, and their average is 0.55. Figure 1 maps the network of dependency in the IdioVol. We simultaneously test 435 hypotheses of no correlation, and Figure 1 connects only the assets, for which the null is rejected. We account for multiple testing by controlling the false discovery rate at 5%. Overall, Figure 1 shows that the cross-sectional dependence between the IdioVols is very strong.

Figure 1: The network of dependencies in total IdioVols. The color and thickness of each line is proportional to the estimated value of $\text{Corr} (C_{Zi}, C_{Zj})$, the quadratic-covariation based correlation between the IdioVols, defined in equation (8) (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.

Could missing factors in the R-FM provide an explanation? Omitted return factors in the 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 $l_n^2 \Delta_n \to \infty$, so it is of smaller order than the number of observations $k_n$ in our estimators of Section 3).
R-FM are captured by the idiosyncratic returns, and can therefore induce correlation between the estimated IdioVols, provided these missing return factors have non-negligible volatility of volatility. To investigate this possibility, we consider the correlations between idiosyncratic returns, $\text{Corr}(Z_i, Z_j)$.\footnote{Our measure of correlation between the idiosyncratic returns $dZ_i$ and $dZ_j$ is}

$$
\text{Corr}(Z_i, Z_j) = \frac{\int_0^T C_{Z_iZ_j,t} \, dt}{\sqrt{\int_0^T C_{Z_i,t} \, dt \sqrt{\int_0^T C_{Z_j,t} \, dt}}}, \quad i, j = 1, \ldots, dS,
$$

(29)

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

Table 2 presents a summary of how estimates $\text{Corr}(Z_i, Z_j)$ are related to the estimates of correlation in IdioVols, $\text{Corr}(C_{Z_i}, C_{Z_j})$. In particular, different rows in Table 2 display average values of $\text{Corr}(C_{Z_i}, C_{Z_j})$ among those pairs, for which $\text{Corr}(Z_i, Z_j)$ is below some threshold. For example, the last-but-one row in Table 2 indicates that there are 56 pairs of stocks with $\text{Corr}(Z_i, Z_j) < 0.01$, and among those stocks, the average correlation between IdioVols, $\text{Corr}(C_{Z_i}, C_{Z_j})$, is estimated to be 0.579. We observe that $\text{Corr}(C_{Z_i}, C_{Z_j})$ is virtually the same compared to pairs of stocks with high $\text{Corr}(Z_i, Z_j)$. These results suggest that missing return factors cannot explain dependence in IdioVols 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 IdioVols, we consider the Idiosyncratic Volatility Factor Model (IdioVol-FM) of Section 2. We first use the market volatility as the only IdioVol factor.\footnote{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.} Table 3 reports the estimates of the IdioVol loading ($\gamma_{Z_i}$) and the $R^2$ of the IdioVol-FM ($R^2_{Z_i,\text{IdioVol-FM}}$, see equation (10)). Table 3 uses two different definitions of IdioVol, one defined with respect to CAPM, and a second IdioVol defined with respect to Fama-French three factor model. For every stock, the estimated IdioVol factor loading is positive, suggesting that the idiosyncratic volatility co-moves with the market volatility. Next, Figure 2 shows the implications for the cross-section of the one-factor IdioVol-FM when the IdioVol is defined with respect to CAPM. The average pairwise correlations between the residual IdioVols, $\hat{\text{Corr}}(C_{Z_i}, C_{Z_j})$, decrease to 0.25. However, the market volatility cannot explain all cross-sectional dependence in residual IdioVols, as evidenced by the remaining links in Figure 2.
Finally, we consider an IdioVol-FM with ten IdioVol factors, market volatility and the volatilities of nine industry ETFs. Figure 3 shows the implications for the cross-section of this ten-factor IdioVol-FM when the IdioVol is defined with respect to CAPM. The average pairwise correlations between the residual IdioVols, \( \hat{\text{Corr}}(C_{Zi}, C_{Zj}) \), decrease further to 0.18. However, significant dependence between the residual IdioVols remains, as evidenced by the remaining links in Figure 2. Our results suggest that there is room for considering the construction of additional IdioVol factors based on economic theory, for example, along the lines of the heterogeneous agents model of Herskovic, Kelly, Lustig, and Nieuwerburgh (2016).
Figure 2: The network of dependencies in residual IdioVols with a single IdioVol factor: the market variance.

Figure 3: The network of dependencies in residual IdioVols with ten IdioVol factors: the market variance and the variances of nine industry ETFs.

In both figures, the color and thickness of each line is proportional to the estimated value of $\text{Corr}(C_{Zi}^{\text{resid}}, C_{Zj}^{\text{resid}})$, the quadratic-covariation based correlation between the IdioVols, defined in (9), 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.
<table>
<thead>
<tr>
<th>Sector</th>
<th>Stock</th>
<th>Ticker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial</td>
<td>American International Group, Inc.</td>
<td>AIG</td>
</tr>
<tr>
<td></td>
<td>American Express Company</td>
<td>AXP</td>
</tr>
<tr>
<td></td>
<td>Citigroup Inc.</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>JPMorgan Chase &amp; Co.</td>
<td>JPM</td>
</tr>
<tr>
<td>Energy</td>
<td>Chevron Corp.</td>
<td>CVX</td>
</tr>
<tr>
<td></td>
<td>Exxon Mobil Corp.</td>
<td>XOM</td>
</tr>
<tr>
<td>Consumer Staples</td>
<td>Coca Cola Company</td>
<td>KO</td>
</tr>
<tr>
<td></td>
<td>Altria</td>
<td>MO</td>
</tr>
<tr>
<td></td>
<td>The Procter &amp; Gamble Company</td>
<td>PG</td>
</tr>
<tr>
<td></td>
<td>Wal-Mart Stores</td>
<td>WMT</td>
</tr>
<tr>
<td>Industrials</td>
<td>Boeing Company</td>
<td>BA</td>
</tr>
<tr>
<td></td>
<td>Caterpillar Inc.</td>
<td>CAT</td>
</tr>
<tr>
<td></td>
<td>General Electric Company</td>
<td>GE</td>
</tr>
<tr>
<td></td>
<td>Honeywell International Inc</td>
<td>HON</td>
</tr>
<tr>
<td></td>
<td>3M Company</td>
<td>MMM</td>
</tr>
<tr>
<td></td>
<td>United Technologies</td>
<td>UTX</td>
</tr>
<tr>
<td>Technology</td>
<td>Hewlett-Packard Company</td>
<td>HPQ</td>
</tr>
<tr>
<td></td>
<td>International Bus. Machines</td>
<td>IBM</td>
</tr>
<tr>
<td></td>
<td>Intel Corp.</td>
<td>INTC</td>
</tr>
<tr>
<td></td>
<td>Microsoft Corporation</td>
<td>MSFT</td>
</tr>
<tr>
<td>Health Care</td>
<td>Johnson &amp; Johnson</td>
<td>JNJ</td>
</tr>
<tr>
<td></td>
<td>Merck &amp; Co.</td>
<td>MRK</td>
</tr>
<tr>
<td></td>
<td>Pfizer Inc.</td>
<td>PFE</td>
</tr>
<tr>
<td>Consumer Discretionary</td>
<td>The Walt Disney Company</td>
<td>DIS</td>
</tr>
<tr>
<td></td>
<td>Home Depot Inc</td>
<td>HD</td>
</tr>
<tr>
<td></td>
<td>McDonald’s Corporation</td>
<td>MCD</td>
</tr>
<tr>
<td>Materials</td>
<td>Alcoa Inc.</td>
<td>AA</td>
</tr>
<tr>
<td></td>
<td>E.I. du Pont de Nemours &amp; Company</td>
<td>DD</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>AT&amp;T Inc.</td>
<td>T</td>
</tr>
<tr>
<td>Services</td>
<td>Verizon Communications Inc.</td>
<td>VZ</td>
</tr>
</tbody>
</table>

*Table 1:* The table lists the stocks used in the empirical application (for the time period 2003-2012). 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.
| $|\widehat{\text{Corr}}(Z_i, Z_j)|$ | Pairs | Avg $|\widehat{\text{Corr}}(Z_i, Z_j)|$ | Avg $\widehat{\text{Corr}}(C_{Z_i}, C_{Z_j})$ | Pairs | Avg $|\widehat{\text{Corr}}(Z_i, Z_j)|$ | Avg $\widehat{\text{Corr}}(C_{Z_i}, C_{Z_j})$ |
|---|---|---|---|---|---|---|
| $< 0.6$ | 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 2: Each row in this table describes the subset of pairs of stocks with $|\text{Corr}(Z_i, Z_j)|$ below a threshold in column one. The table considers two R-FMs: the left panel defines the IdioVol with respect to CAPM, and the right panel defines the IdioVol with respect to the three-factor Fama-French model. In both cases, the market volatility is the only IdioVol 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 IdioVols.
Table 3: Estimates of the IdioVol factor loading ($\hat{\gamma}_z$, see equation (7)), and the contribution of the market volatility to the variation in the IdioVols ($\hat{R}^2_{Z,IdioVol-FM}$, see equation (10)). The table considers two R-FMs: the left panel defines the IdioVol with respect to CAPM, and the right panel defines the IdioVol with respect to the three-factor Fama-French model. In both cases, the market volatility is the only IdioVol factor. P-val is the p-value of the test of the absence of dependence between the IdioVol and the market volatility for a given individual stock, see equation (27).
6 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 $j = 1, 2$,

$$dY_{j,t} = \beta_t dX_t^c + d\tilde{Y}_{j,t}^c + dJ_{j,t}, \quad d\tilde{Y}_{j,t}^c = \sqrt{C_{Z_{j,t}}} d\tilde{W}_{j,t}.$$

In the above, $C_X$ is the spot volatility of the market portfolio, $\tilde{W}_1$, and $\tilde{W}_2$ are Brownian motions with $\text{Corr}(d\tilde{W}_{1,t}, d\tilde{W}_{2,t}) = 0.4$, 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,

$$df_{1,t} = 5(0.09 - f_{1,t}) dt + 0.35\sqrt{f_{1,t}} \left( -0.8 dW_t + \sqrt{1-0.8^2} dB_{1,t} \right),$$

$$df_{j,t} = 5(0.09 - f_{j,t}) dt + 0.35\sqrt{f_{1,t}} dB_{j,t} \quad \text{for } j = 2, 3, 4,$$

where $B_1, \ldots, B_4$ and independent standard Brownian Motions, which are also independent from the Brownian Motions of the return Factor Model. 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 IdioVol processes $C_{Z_{1,t}}$ and $C_{Z_{2,t}}$, see Table 4 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

\[ \text{The Feller property is satisfied implying the positiveness of the processes } (f_j,t)_{1 \leq j \leq 4}. \]
is $-0.8$, which is standard in the literature, see Kalnina and Xiu (2017), Aït-Sahalia, Fan, and Li (2013) and Aït-Sahalia, Fan, Laeven, Wang, and Yang (2017).

\[
\begin{array}{ccc}
C_{Z1,t} & C_{Z2,t} \\
\hline
\text{Model 1} & 0.1 + 1.5f_{z1,t} & 0.1 + 1.5f_{z2,t} \\
\text{Model 2} & 0.1 + 0.6c_{X,1,t} + 0.4f_{z1,t} & 0.1 + 0.5c_{X,1,t} + 0.5f_{z2,t} \\
\text{Model 3} & 0.1 + 0.45c_{X,1,t} + f_{z1,t} + 0.4f_{z4,t} & 0.1 + 0.35c_{X,1,t} + 0.3f_{z4,t} + 0.6f_{z4,t} \\
\end{array}
\]

Table 4: Different specifications for the Idiosyncratic Volatility processes $C_{Z1,t}$ and $C_{Z2,t}$.

We set the time span $T$ equal 1,260 or 2,520 days, which correspond approximately to 5 and 10 business years. These values are standard in the nonparametric leverage effect estimation literature (see Aït-Sahalia, Fan, and Li (2013) and Kalnina and Xiu (2017)), where the rate of convergence is also $\Delta^{-1/4}$. 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 (2016) and set the truncation threshold $u_n$ in day $t$ at $3\hat{\sigma}_t\Delta_0^{0.49}$, where $\hat{\sigma}_t$ 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 estimands include:

- the IdioVol factor loading of the first stock, $\gamma_{Z1}$,
- the contribution of the market volatility to the variation of the IdioVol of the first stock $R^{2\text{IdioVol-FM}}_{Z1}$,
- the correlation between the idiosyncratic volatilities of stocks 1 and 2, $\text{Corr} \left( C_{Zj}, C_{Zs} \right)$,
- the correlation between residual idiosyncratic volatilities, $\text{Corr} \left( C_{\text{resid}Zj}, C_{\text{resid}Zs} \right)$.

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 $C_{X,t}$ and $(f_{jt})_{0 \leq j \leq 27}$ and replicate several times the other parts of the DGP. In Table 5, 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 6, 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, see Table 7. 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 (2017).

<table>
<thead>
<tr>
<th>( \hat{\theta} )</th>
<th>AN</th>
<th>LIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\gamma}_{Z1} )</td>
<td>1.5</td>
<td>2</td>
</tr>
<tr>
<td>Median Bias</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.047</td>
<td>-0.025</td>
<td>-0.011</td>
</tr>
<tr>
<td>( \hat{\gamma}_{Z1} )</td>
<td>0.176</td>
<td>0.130</td>
</tr>
<tr>
<td>( \hat{R}_{Z1}^{2,IdioVol-FM} )</td>
<td>-0.288</td>
<td>-0.212</td>
</tr>
<tr>
<td>( \hat{Corr} (C_{Z1}, C_{Z2}) )</td>
<td>-0.189</td>
<td>-0.113</td>
</tr>
<tr>
<td>( \hat{Corr} (C_{Z1}^\text{resid}, C_{Z2}^\text{resid}) )</td>
<td>0.222</td>
<td>0.166</td>
</tr>
<tr>
<td>IQR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.176</td>
<td>0.130</td>
<td>0.103</td>
</tr>
<tr>
<td>( \hat{R}_{Z1}^{2,IdioVol-FM} )</td>
<td>0.210</td>
<td>0.188</td>
</tr>
<tr>
<td>( \hat{Corr} (C_{Z1}, C_{Z2}) )</td>
<td>0.404</td>
<td>0.325</td>
</tr>
<tr>
<td>( \hat{Corr} (C_{Z1}^\text{resid}, C_{Z2}^\text{resid}) )</td>
<td>0.456</td>
<td>0.384</td>
</tr>
</tbody>
</table>

Table 5: Finite sample properties of our estimators using 10 years of data sampled at 5 minutes. The true values are \( \gamma_{Z1} = 0.450, R_{Z1}^{2,IdioVol-FM} = 0.342, Corr (C_{Z1}, C_{Z2}) = 0.523, Corr (C_{Z1}^\text{resid}, C_{Z2}^\text{resid}) = 0.424 \). Model 3.

Next, we study the empirical rejection probabilities of the three statistical tests as outlined in Section 4.4. The first null hypothesis is the absence of dependence between the IdioVols (for which we use Model 1), \( H_0^1 : [C_{Z1}, C_{Z2}]_T = 0 \). The second null hypothesis we test is the absence of dependence between the IdioVol of the first stock and the market volatility (for which we use Model 1), \( H_0^2 : [C_{Z1}, C_X]_T = 0 \). The third null hypothesis is the absence of dependence in the two
<table>
<thead>
<tr>
<th>$\hat{\theta}$</th>
<th>AN</th>
<th>LIN</th>
<th>Median Bias</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\gamma}_{Z1}$</td>
<td>1.5</td>
<td>2</td>
<td>2.5</td>
<td>3</td>
</tr>
<tr>
<td>$\hat{\gamma}_{Z1}$</td>
<td>-0.022</td>
<td>-0.012</td>
<td>-0.003</td>
<td>0.004</td>
</tr>
<tr>
<td>$\hat{R}^{IdioVol-FM}_{Z1}$</td>
<td>0.107</td>
<td>0.091</td>
<td>0.073</td>
<td>0.056</td>
</tr>
<tr>
<td>$\hat{Corr}(C_{Z1},C_{Z2})$</td>
<td>-0.147</td>
<td>-0.104</td>
<td>-0.073</td>
<td>-0.048</td>
</tr>
<tr>
<td>$\hat{Corr}(C_{Z1}^{resid},C_{Z2}^{resid})$</td>
<td>-0.135</td>
<td>-0.086</td>
<td>-0.058</td>
<td>-0.039</td>
</tr>
</tbody>
</table>

**Table 6:** Finite sample properties of our estimators using 10 years of data sampled at 1 minute. The true values are $\gamma_{Z1} = 0.450$, $R^{IdioVol-FM}_{Z1} = 0.336$, $Corr(C_{Z1},C_{Z2}) = 0.514$, $Corr(C_{Z1}^{resid},C_{Z2}^{resid}) = 0.408$. Model 3.

<table>
<thead>
<tr>
<th>$\hat{\theta}$</th>
<th>AN</th>
<th>LIN</th>
<th>Median Bias</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\gamma}_{Z1}$</td>
<td>1.5</td>
<td>2</td>
<td>2.5</td>
<td>3</td>
</tr>
<tr>
<td>$\hat{\gamma}_{Z1}$</td>
<td>0.156</td>
<td>0.112</td>
<td>0.088</td>
<td>0.075</td>
</tr>
<tr>
<td>$\hat{R}^{IdioVol-FM}_{Z1}$</td>
<td>0.201</td>
<td>0.146</td>
<td>0.118</td>
<td>0.100</td>
</tr>
<tr>
<td>$\hat{Corr}(C_{Z1},C_{Z2})$</td>
<td>0.340</td>
<td>0.238</td>
<td>0.184</td>
<td>0.150</td>
</tr>
<tr>
<td>$\hat{Corr}(C_{Z1}^{resid},C_{Z2}^{resid})$</td>
<td>0.417</td>
<td>0.291</td>
<td>0.228</td>
<td>0.184</td>
</tr>
</tbody>
</table>

**Table 7:** Finite sample properties of our estimators using 5 years of data sampled at 1 minute. The true values are $\gamma_{Z1} = 0.450$, $R^{IdioVol-FM}_{Z1} = 0.35$, $Corr(C_{Z1},C_{Z2}) = 0.517$, $Corr(C_{Z1}^{resid},C_{Z2}^{resid}) = 0.417$. Model 3.
residual IdioVols (for which we use Model 2), $H_0^3 : [C_{Z1Z1}^{resid}, C_{Z2Z2}^{resid}]_T = 0$.

The three panels of Table 8 contain the empirical rejection probabilities for the three null hypotheses. We present the results for two sampling frequencies ($\Delta_n = 1$ minute and $\Delta_n = 5$ minutes) and the two type of estimators (AN and LIN). We see that the empirical rejection probabilities are reasonably close to the nominal size of the test. Neither type of estimator (AN or LIN) seems to dominate the other. Consistent with the asymptotic theory, the empirical rejection probabilities of the three tests become closer to the nominal size of the test when frequency is higher.

<table>
<thead>
<tr>
<th>$\Delta_n = 5$ minutes</th>
<th>$\Delta_n = 1$ minute</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta = 1.5$</td>
<td>$\theta = 1.5$</td>
</tr>
<tr>
<td>AN</td>
<td>LIN</td>
</tr>
</tbody>
</table>

**Panel A : $H_0^1 : [C_{Z1}, C_{Z2}]_T = 0$, Model 1**

- $\alpha = 10\%$: 9.7 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 : $H_0^2 : [C_{Z1}, C_X]_T = 0$, 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 C : $H_0^3 : [C_{Z1Z1}^{resid}, C_{Z2Z2}^{resid}]_T = 0$, 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

Table 8: Panel A contains the empirical rejection probabilities of the test of absence of dependence between IdioVols. Panel B contains the empirical rejection probabilities of the test of absence of dependence between the IdioVol and the market volatility. Panel C contains the empirical rejection probabilities of the test absence of dependence between residual IdioVols. $T = 10$ years. $\alpha$ denotes the nominal size of the test.
7 Conclusion

We introduce an econometric framework for analysis of cross-sectional dependence in the IdioVols of assets using high frequency data. First, we provide bias-corrected estimators of standard measures of dependence between IdioVols, as well as the associated asymptotic theory. Second, we study an IdioVol factor model, in which we decompose the variation in IdioVols into two parts: the variation related to the systematic factors such as the market volatility, and the residual variation. We provide the asymptotic theory that allows us to test, for example, whether the residual (non-systematic) components of the IdioVols exhibit cross-sectional dependence.

To provide the bias-corrected estimators and inference results, we develop a new asymptotic theory for general estimators of quadratic covariation of vector-valued (possibly) nonlinear transformations of the spot covariance matrices. This theoretical contribution is of its own interest, and can be applied in other contexts. For example, our results can be used to conduct inference for the cross-sectional dependence in asset betas.

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 (non-systematic) idiosyncratic volatilities across the stocks.
References


Appendix

Sections A, B, and C contain all the proofs, and Section D contains additional figures for the empirical application.

A Proof of Theorem 1

We start by introducing some notation. Our notation is similar to that of the proofs of Jacod and Rosenbaum (2015) whenever possible.

A.1 Notation

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 $a > a'$.

By the usual localization argument, there exists a $\pi$-integrable function $J$ on $E$ and a constant such that the stochastic processes in (20) and (21) satisfy

$$\|b\|, \|\tilde{b}\|, \|c\|, \|\tilde{c}\|, J \leq A, \|\delta(w, t, z)||^r \leq J(z).$$  \hspace{1cm} (A.1)

For any càdlàg bounded process $Z$, we set

$$\eta_{t,s}(Z) = \sqrt{E \left( \sup_{0 \leq u \leq s} \|Z_{t+u} - Z_t\|^2 |F_t \right)}, \quad \text{and}$$

$$\eta_{i,j}^n(Z) = \sqrt{E \left( \sup_{0 \leq u \leq j \Delta_n} \|Z_{(i-1)\Delta_n + u} - Z_{(i-1)\Delta_n}\|^2 |F_{i\Delta_n} \right)}.$$

For convenience, we decompose $Y_t$ as

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

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

Let $\hat{C}^m_i$ be the local estimator of the spot variance of the unobservable process $Y'$, that is

$$\hat{C}^m_i = \frac{1}{k_n \Delta_n} \sum_{u=0}^{k_n-1} (\Delta^n_{i+u} Y') (\Delta^n_{i+u} Y')^T = (\hat{C}^{m,gh}_i)_{1 \leq g, h \leq d}. \hspace{1cm} (A.2)$$

There is no jump truncation applied in the definition of $\hat{C}^m_i$ since the process $Y'$ is continuous. Hence, it is more convenient to work with $\hat{C}^{m}_{i}$ rather than $\hat{C}_i$ (defined in (15)). Let’s also define

$$\alpha^n_i = (\Delta^n_i Y') (\Delta^n_i Y')^T - C^n_i \Delta_n, \quad \nu^n_i = \hat{C}^n_i - C^n_i, \quad \text{and} \quad \lambda^n_i = \hat{C}^n_{i+k_n} - \hat{C}^n_i,$$

which satisfy

$$\nu^n_i = \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n-1} (\alpha^n_{i+j} + (C^n_{i+j} - C^n_i) \Delta_n) \quad \text{and} \quad \lambda^n_i = \nu^n_{i+k_n} - \nu^n_i + \Delta_n(C^n_{i+k_n} - C^n_i). \hspace{1cm} (A.4)$$

The following multidimensional quantities will be used in the sequel

$$\zeta(1)_i^n = \frac{1}{\Delta_n} \Delta^n_i Y'(\Delta^n_i Y')^T - C^n_{i-1}, \quad \zeta(2)_i^n = \Delta^n_i C.$$
We provide some useful theorems and lemmas here, which are used to prove Theorem 1. These theorems and lemmas are proved in Appendix C below.

**Theorem A1.** Let \([H\hat{(}C),\hat{G}(C)]_T^{LIN} \) and \([H\hat{(}C),\hat{G}(C)]_T^{AN} \) be the infeasible estimators obtained by replacing \( \hat{C}_i^n \) by \( \hat{C}_i^n \) in the definition of \([H(C),G(C)]_T^{LIN} \) and \([H(C),G(C)]_T^{AN} \) in (18) and (16). As long as \((8p - 1)/(4p - r) \leq \varpi < \frac{1}{2}\), we have

\[
\Delta_n^{-1/4} \left( [H(C),\hat{G}(C)]_T^{LIN} - [H(C),\hat{G}(C)]_T^{LIN'} \right) \overset{p}{\rightarrow} 0
\]

and

\[
\Delta_n^{-1/4} \left( [H(C),\hat{G}(C)]_T^{AN} - [H(C),\hat{G}(C)]_T^{AN'} \right) \overset{p}{\rightarrow} 0.
\]

Theorem A1 allows, in particular, to focus on the derivation of the asymptotic distributions of \([H(C),\hat{G}(C)]_T^{LIN'} \) and \([H(C),\hat{G}(C)]_T^{AN'} \). The next theorem connects the two estimators that we have intro-
duced. To state the theorem, define

\[
[H(C), G(C)]_T^A = \frac{3}{2^k_n} \sum_{g,h,a,b=1}^d \left( \frac{T/\Delta_n}{2^{k_n+1}} \right) \left( \partial_{gh} H \partial_{ab} G \right) (C_1^n)(C_i^n, gh) (\tilde{C}_i^n, gh) (\tilde{C}_i^n, ab) - 2 \frac{\partial_{gh} H \partial_{ab} G (\tilde{C}_i^n, gh) (\tilde{C}_i^n, ab)}{k_n}
\]

with \( C_i^n = C_{i-1} \Delta_n \), and the superscript \( A \) stands for “approximated”. For simplicity, we do not index the above quantity by a prime although it depends on \( \tilde{C}_i^n \), instead of \( \tilde{C}_i^n \).

**Theorem A2.** Under the assumptions of Theorem 1, we have

\[
\Delta_n^{-1/4} \left( [H(C), G(C)]_T^{LIN'} - [H(C), G(C)]_T^A \right) \overset{p}{\rightarrow} 0 \quad \text{and} \quad \Delta_n^{-1/4} \left( [H(C), G(C)]_T^{AN'} - [H(C), G(C)]_T^A \right) \overset{p}{\rightarrow} 0.
\]

(A.8)

Theorem A2 shows that the two estimators \([H(C), G(C)]_T^{LIN'} \) and \([H(C), G(C)]_T^{AN'} \) can be approximated by a certain quantity with an error of approximation of order smaller than \( \Delta_n^{-1/4} \).

Now, we decompose the approximated estimator as follows

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

A.9

with

\[
[H(C), G(C)]_T^{(A1)} = \frac{3}{2^k_n} \sum_{g,h,a,b=1}^d \left( \frac{T/\Delta_n}{2^{k_n+1}} \right) \left( \partial_{gh} H \partial_{ab} G \right) (C_1^n)(C_i^n, gh) (\tilde{C}_i^n, gh) (\tilde{C}_i^n, ab) - 2 \frac{\partial_{gh} H \partial_{ab} G (\tilde{C}_i^n, gh) (\tilde{C}_i^n, ab)}{k_n}
\]

and

\[
[H(C), G(C)]_T^{(A2)} = \frac{3}{2^k_n} \sum_{g,h,a,b=1}^d \left( \frac{T/\Delta_n}{2^{k_n+1}} \right) \left( \partial_{gh} H \partial_{ab} G \right) (\tilde{C}_i^n) (\tilde{C}_i^n, gh) (\tilde{C}_i^n, ab) + 2 \frac{\partial_{gh} H \partial_{ab} G (\tilde{C}_i^n, gh) (\tilde{C}_i^n, ab)}{k_n}
\]

The following theorem holds:

**Theorem A3.** Under the assumptions of Theorem 1, we have

\[
\frac{1}{\Delta_n^{1/4}} \left( [H(C), G(C)]_T^{(A1)} - \sum_{g,h,a,b=1}^d \sum_{u,v=1}^2 \overline{A_1 \Pi}(H, gh, u; G, ab, v) \right) + \overline{A_1 \Pi}(G, ab, v; H, gh, u) \overset{p}{\rightarrow} 0.
\]

Lemmma A1. For any càdlàg bounded process \( Z \), for all \( t, s > 0, j, k \geq 0 \), set \( \eta_{t,s} = \eta_{t,s}(Z) \). Then,

\[
\Delta_n E \left( \sum_{i=1}^{[t/\Delta_n]} \eta_{i,k} \right) \rightarrow 0, \quad \Delta_n E \left( \sum_{i=1}^{[t/\Delta_n]} \eta_{i,2k} \right) \rightarrow 0,
\]

\[
E \left( \eta_{i+j,k} | \mathcal{F}_{i}^n \right) \leq \eta_{i,j+k} \quad \text{and} \quad \Delta_n E \left( \sum_{i=1}^{[t/\Delta_n]} \eta_{i,4k} \right) \rightarrow 0.
\]

39
Lemma A2. Let $Z$ be a continuous Itô process with drift $b^Z_t$ and spot variance process $C^Z_t$, and set $\eta_{t,s} = \eta_{t,s}(b^Z_t, c^Z_t)$. Then, the following bounds hold:

\[
\begin{align*}
&\mathbb{E}(Z^Z_t | F_0) - tb^Z_0 \leq K t \eta_{0,t} \\
&\mathbb{E}(Z^Z_t Z^Z_t - tC^Z_0 Z^Z_t) \leq K t^{3/2}(\sqrt{\Delta} + \eta_{0,t}) \\
&\mathbb{E}((Z^Z_t Z^Z_t - tC^Z_0 Z^Z_t)(C^Z_0 - C^Z_0)) \leq K t^2 \\
&\mathbb{E}(Z^Z_t Z^Z_t | F_0) - \Delta_n Z^Z_0 C^Z_0 + C^Z_0 C^Z_0 + C^Z_0 C^Z_0 \leq K t^{5/2} \\
&\mathbb{E}(|Z^Z_t Z^Z_t| | F_0) \leq K t^2 \\
&\mathbb{E}\left(\sum_{l=1}^{6} Z^Z_t | F_0 - \frac{\Delta_n^3}{6} \sum_{l=1}^{3} \sum_{k=1}^{k} \sum_{l=1}^{k'} \sum_{m=1}^{m'} C^Z_{i,l} C^Z_{i,k} C^Z_{i,l} C^Z_{i,k} \right) \leq K t^{7/2} \\
&\mathbb{E}\left(\sup_{w \in [0,s]} \left|Z_{t+w} - Z_t\right|^q | F_t\right) \leq K q s^{q/2}, \text{ and } \mathbb{E}\left(\left|Z_{t+s} - Z_t\right| | F_t\right) \leq K s. \tag{A.10}
\end{align*}
\]

Lemma A3. Let $\zeta^\eta_i$ be a $r$-dimensional $F^n_t$-measurable process satisfying $\mathbb{E}(\zeta^\eta_i | F^n_{t-1}) \leq L'$ and $\mathbb{E}(\zeta^\eta_i | F^n_{t-1}) \leq L_q$. Also, let $\varphi^\eta_i$ be a real-valued $F^n_t$-measurable process with $\mathbb{E}(\varphi^\eta_i | F^n_{t-1}) \leq L^q$ for $q \geq 2$ and $1 \leq j \leq 2k_n - 1$. Then,

\[
\mathbb{E}\left(\left\|\sum_{j=1}^{2k_n-1} \varphi^\eta_{i+j} \zeta^\eta_i \right\|^q | F^n_{t-1}\right) \leq K q \left(L_q k^{q/2} + L^q k^q\right). \tag{A.11}
\]

Lemma A4. Under the assumptions of Theorem 1, we have:

\[
\begin{align*}
&\mathbb{E}\left(\lambda_i^{n,j} \lambda_i^{n,k} \lambda_i^{n,g} \lambda_i^{n,h} \lambda_i^{n,ab} | F^n_t \right) - \frac{4}{k^n} \left(C_i^{n,ga} C_i^{n,gb} + C_i^{n,gb} C_i^{n,ha} + C_i^{n,jl} C_i^{n,km} + C_i^{n,jm} C_i^{n,kl}\right) \\
&- \frac{4\Delta_n^3}{3} \left(C_i^{n,jl} C_i^{n,km} + C_i^{n,jm} C_i^{n,kl}\right) \mathbb{E}\left(\varphi^\eta_i^{n,gh,ab} | F^n_t \right) \leq K \Delta_n \left(\Delta_n^{1/8} + \eta_{t,4k_n}\right).
\end{align*}
\]

Lemma A5. Under the assumptions of Theorem 1, we have:

\[
\begin{align*}
&\mathbb{E}\left(\varphi^\eta_i^{n,jk} \varphi^\eta_i^{n,lm} \varphi^\eta_i^{n,gh} | F^n_t \right) \leq K \Delta_n^{3/4} \left(\Delta_n^{1/4} + \eta_{t,k_n}\right), \tag{A.12}
\end{align*}
\]

\[
\begin{align*}
&\mathbb{E}\left(\varphi^\eta_i^{n,jk} \varphi^\eta_i^{n,lm} (c_i^{n,gh} - c_i^{n,gh}) | F^n_t \right) \leq K \Delta_n^{3/4} \left(\Delta_n^{1/4} + \eta_{t,k_n}\right), \tag{A.13}
\end{align*}
\]

\[
\begin{align*}
&\mathbb{E}\left(\varphi^\eta_i^{n,jk} (c_i^{n,lm} - c_i^{n,lm}) (c_i^{n,gh} - c_i^{n,gh}) | F^n_t \right) \leq K \Delta_n^{3/4} \left(\Delta_n^{1/4} + \eta_{t,k_n}\right), \tag{A.14}
\end{align*}
\]

\[
\begin{align*}
&\mathbb{E}\left(\varphi^\eta_i^{n,jk} \lambda_i^{n,lm} \lambda_i^{n,gh} | F^n_t \right) \leq K \Delta_n^{3/4} \left(\Delta_n^{1/4} + \eta_{t,2k_n}\right), \tag{A.15}
\end{align*}
\]

\[
\begin{align*}
&\mathbb{E}\left(\lambda_i^{n,jk} \lambda_i^{n,lm} \lambda_i^{n,gh} | F^n_t \right) \leq K \Delta_n^{3/4} \left(\Delta_n^{1/4} + \eta_{t,2k_n}\right). \tag{A.16}
\end{align*}
\]
Lemma A6. Under the assumptions of Theorem 1, we have:

\[ \frac{1}{\Delta_n^{1/4}} \sum_{i=2k_n}^{[T/\Delta_n]} (\partial_{gh} H \partial_{ab} G)(C_{i-2k_n}^n) \rho_{gh}(u,v)_{i}^{n} \zeta_n(v)_{i} \xrightarrow{p} 0, \quad \forall \ (u,v) \]  \hspace{1cm} (A.17)

\[ \frac{1}{\Delta_n^{1/4}} \left( \mathcal{A} \Pi(H, gh; u; G, ab, v) - \int_0^T (\partial_{gh} H \partial_{ab} G)(C_t) C_t^{gh, ab} dt \right) \xrightarrow{p} 0 \text{ when } (u,v) = (2,2) \]  \hspace{1cm} (A.18)

\[ \frac{1}{\Delta_n^{1/4}} \left( \mathcal{A} \Pi(H, gh; u; G, ab, v) - \frac{3}{6} \int_0^T (\partial_{gh} H \partial_{ab} G)(C_t)(C_t^{ga}C_t^{hb} + C_t^{gb}C_t^{ha}) dt \right) \xrightarrow{p} 0 \]  \hspace{1cm} (A.19)

when \ (u,v) = (1,1),

\[ \frac{1}{\Delta_n^{1/4}} \mathcal{A} \Pi(H, gh; u; G, ab, v) \xrightarrow{p} 0 \text{ when } (u,v) = (1,2), (2,1) \]  \hspace{1cm} (A.20)

### A.3 Return to the Proof of Theorem 1

We now continue the proof of Theorem 1. By Theorem A3, we have

\[ \frac{1}{\Delta_n^{1/4}} \left( [H(C), G(C)]^T - \sum_{g,h,a,b=1}^{d} \sum_{u,v=1}^{2} \mathcal{A} \Pi(H, gh; u; G, ab, v)_{T}^{n} + \mathcal{A} \Pi(G, ab; v; H, gh, u)_{T}^{n} \right) \xrightarrow{p} 0. \]

Recalling the definition of \( \mathcal{A} \Pi(G, ab; v; H, gh, u)_{T}^{n} \) from A.6, Lemma A6 implies that

\[ \frac{1}{\Delta_n^{1/4}} \left( [H(C), G(C)]^T - \sum_{g,h,a,b=1}^{d} \sum_{u,v=1}^{2} \mathcal{A} \Pi(H, gh; u; G, ab, v)_{T}^{n} - \frac{3}{2k_n^3} \sum_{g,h,a,b=1}^{d} \sum_{u,v=1}^{2} \sum_{i=2k_n}^{[T/\Delta_n]} \right) \left( (\partial_{gh} H \partial_{ab} G)(C_{i-2k_n}^n) \rho_{gh}(u,v)_{i}^{n} \zeta_n(v)_{i}^{n} + (\partial_{ab} H \partial_{gh} G)(C_{i-2k_n}^n) \rho_{ab}(v,u)_{i}^{n} \zeta_n^\prime(v)_{i}^{n} \right) \xrightarrow{p} 0. \]  \hspace{1cm} (A.21)

Next, define

\[ \xi(H, gh; u; G, ab, v)_{i}^{n} = \frac{1}{\Delta_n^{1/4}} (\partial_{gh} H \partial_{ab} G)(C_{i-2k_n}^n) \rho_{gh}(u,v)_{i}^{n} \zeta_n(v)_{i}^{n}, \]

\[ Z(H, gh; u; G, ab, v)_{i}^{n} = \frac{1}{\Delta_n^{1/4}} \sum_{i=2k_n}^{[T/\Delta_n]} \xi(H, gh; u; G, ab, v)_{i}^{n}. \]

Notice that (A.21) implies

\[ \frac{1}{\Delta_n^{1/4}} \left( [H(C), G(C)]^T - [H(C), G(C)]_{T} \right) \xrightarrow{\mathcal{F}_i^{n}} \sum_{g,h,a,b=1}^{d} \sum_{u,v=1}^{2} \frac{1}{\Delta_n^{1/4}} (Z(H, gh; u; G, ab, v)_{T}^{n}). \]  \hspace{1cm} (A.22)

Next, observe that to derive the asymptotic distribution of \( \left( [H_1(C), G_1(C)]^T, \ldots, [H_k(C), G_k(C)]^T \right) \), it suffices to study the joint asymptotic behavior of the family of processes \( \frac{1}{\Delta_n^{1/4}} Z(H, gh; u; G, ab, v)_{T}^{n} \). Notice that \( \xi(H, gh; u; G, ab, v)_{i}^{n} \) are martingale increments relative to the discrete filtration \( \mathcal{F}_i^{n} \). Therefore, to
obtain the joint asymptotic distribution of \( \frac{1}{\Delta_n} Z(H, gh, u; G, ab, v)_{n} \), it is enough to prove the following three properties:

\[
A\left((H, gh, u; G, ab, v), (H', g'h', u'; G', a'b', v')\right)_{t} \rightarrow \sum_{i=2k_n}^{\left[\frac{t}{\Delta_n}\right]} \mathbb{E}\left(\xi(H, gh, u; G, ab, v)\xi(H', g'h', u'; G', a'b', v')|\mathcal{F}_{i-1}^{n}\right)
\]

\[
\Rightarrow A\left((H, gh, u; G, ab, v), (H', g'h', u'; G', a'b', v')\right)_{t},
\]

(A.23)

\[
\sum_{i=2k_n}^{\left[\frac{t}{\Delta_n}\right]} \mathbb{E}\left(\xi(H, gh, u; G, ab, v)\right|_{i}^{\left[\frac{t}{\Delta_n}\right]} \mathcal{F}_{i-1}^{n}\right) \Rightarrow 0, \text{ and}
\]

(A.24)

\[
B(N; H, gh, u; G, ab, v)_{n} := \sum_{i=2k_n}^{\left[\frac{t}{\Delta_n}\right]} \mathbb{E}\left(\xi(H, gh, u; G, ab, v)|\mathcal{F}_{i-1}^{n}\right) \Rightarrow 0,
\]

(A.25)

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 \( W^{j} \).

Using the polynomial growth assumption imposed on \( H_{r} \) and \( G_{r} \), (A.24) and (A.25) can be proved by a simple extension of (B.105) and (B.106) in Aït-Sahalia and Jacod (2014) to handle multivariate processes.

Next, define

\[
V_{ab}^{v'}(v, v')_{t} = \begin{cases} 
(C_{g'a'}^{C_{bb'}} + C_{g'b}^{C_{ba'}}) & \text{if } (v, v') = (1, 1) \\
C_{t}^{ab,a'b'} & \text{if } (v, v') = (2, 2) \\
0 & \text{otherwise,}
\end{cases}
\]

and

\[
\nabla_{gh}^{g'h'}(u, u')_{t} = \begin{cases} 
(C_{g'h'}^{C_{hh'}} + C_{t}^{g'h'} C_{t}^{b'g'}) & \text{if } (u, u') = (1, 1) \\
0 & \text{otherwise.}
\end{cases}
\]

Using again the polynomial growth assumption on \( H_{r} \) and \( G_{r} \), we can show that,

\[
A\left((H, gh, u; G, ab, v), (H', g'h', u'; G', a'b', v')\right)_{t} = 

M(u, v; u', v') \int_{0}^{t} \left(\partial_{gh} H \partial_{ab} G \partial_{g'h'} H \partial_{a'b'} G\right)(C_{s}) V_{ab}^{v'}(v, v')_{s} \nabla_{gh}^{g'h'}(u, u')_{s} ds,
\]

with

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

Therefore, we have

\[
A\left((H, gh, u; G, ab, v), (H', g'h', u'; G', a'b', v')\right)_{T} = 

\begin{cases} 
\frac{3}{\theta^{3}} \int_{0}^{T} \left(\partial_{gh} H \partial_{ab} G \partial_{g'h'} H \partial_{a'b'} G\right)(C_{t}) (C_{t}^{g'a'} C_{t}^{b'h'} + C_{t}^{g'h'} C_{t}^{b'g'}) (C_{t}^{g'a'} C_{t}^{b'h'} + C_{t}^{g'h'} C_{t}^{b'g'}) dt, & \text{if } (u, v; u', v') = (1, 1; 1, 1) \\
\frac{3}{4\theta} \int_{0}^{T} \left(\partial_{gh} H \partial_{ab} G \partial_{g'h'} H \partial_{a'b'} G\right)(C_{t}) (C_{t}^{g'a'} C_{t}^{b'h'} + C_{t}^{g'h'} C_{t}^{b'g'}) C_{t}^{a'b'} dt, & \text{if } (u, v; u', v') = (1, 2; 1, 2), (2, 1; 2, 1) \\
\frac{151}{280} \theta \int_{0}^{T} \left(\partial_{gh} H \partial_{ab} G \partial_{g'h'} H \partial_{a'b'} G\right)(C_{t}) C_{t}^{a'b'} dt, & \text{if } (u, v; u', v') = (2, 2; 2, 2) \\
0 & \text{otherwise.}
\end{cases}
\]
Using (A.22), we deduce that the asymptotic covariance between \([H_r(C), G_r(C)]_T^{(A)}\) and \([H_s(C), G_s(C)]_T^{(A)}\) is given by

\[
\sum_{g,h,a,b=1}^{d} \sum_{i,j,k,l,m=1}^{d} \sum_{j,k,l,m=1}^{2} \left( A \left( (H_r, gh; u, v; G_r, ab; v), (H_s, g'h', u'; G_s, ab'; v') \right) \right)_T \\
+ A \left( (H_r, gh; u, v; G_r, ab; v), (H_s, a'b', v'; G_s, g'h', u') \right)_T \\
+ A \left( (H_r, ab; v; H_r, gh; u), (H_s, a'b', v'; G_s, g'h', u') \right)_T \\
+ A \left( (H_r, ab; v; H_r, gh; u), (H_s, a'b', v'; G_s, g'h', u') \right)_T.
\]

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

\[
\sum_{g,h,a,b=1}^{d} \sum_{i,j,k,l,m=1}^{d} \left( \frac{6}{T^2} \int_0^T \left( \partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(C_i) \right) \left[ (C_i^{gh,jk} C_i^{ab,lm}) + (C_i^{gh,jk} C_i^{ab,lm}) \right] dT \\
+ \frac{151\theta}{140} \int_0^T \left( \partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(C_i) \right) \left[ \left( C_i^{gh,jk} C_i^{ab,lm} \right) + \left( C_i^{gh,jk} C_i^{ab,lm} \right) \right] dT \\
+ \frac{3}{2\theta} \int_0^T \left( \partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(C_i) \right) \left[ \left( C_i^{gh,jk} C_i^{ab,lm} \right) + \left( C_i^{gh,jk} C_i^{ab,lm} \right) \right] dT \\
+ \left( C_i^{gh,jk} C_i^{ab,lm} \right) dT \right).
\]

which completes the proof.

**B Proof of Theorem 2**

Using the polynomial growth assumption on \(H_r, G_r, H_s\) and \(G_s\) and Theorem 2.2 in Jacob and Rosenbaum (2015), one can show that

\[
\frac{6}{T^2} \tilde{\Omega}_T^{r,s,1} - \frac{3}{2\theta} \hat{\Omega}_T^{r,s,1} \rightarrow \Sigma_T^{r,s,1}.
\]

Next, by equation (3.27) in Jacob and Rosenbaum (2015), we have

\[
\frac{3}{2\theta} \hat{\Omega}_T^{r,s,1} - \frac{3}{2\theta} \hat{\Omega}_T^{r,s,1} \rightarrow \Sigma_T^{r,s,1}.
\]

Finally, to show that

\[
\frac{151\theta}{140} \frac{9}{4\theta^2} \hat{\Omega}_T^{r,s,1} + \frac{4}{\theta^2} \hat{\Omega}_T^{r,s,1} - \frac{3}{2\theta} \hat{\Omega}_T^{r,s,1} \rightarrow \Sigma_T^{r,s,1},
\]

we first observe that the approximation error induced by replacing \(\hat{C}_i^n\) by \(\tilde{C}_i^n\) in Theorem 2 is negligible. For \(1 \leq g, h, a, b, j, k, l, m \leq d\) and \(1 \leq r, s \leq d\), we define

\[
\hat{W}_i^n = \sum_{|j| \leq \Delta_n} \left( \partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(C_i) \right) \left( \hat{C}_i^n \right) \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm} |F_i^n|^n,
\]

\[
\hat{w}(1)_i^n = \left( \partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(C_i) \right) \left( \hat{C}_i^n \right) \left( \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm} \right) |F_i^n|^n,
\]

43
\[\tilde{w}(2)_i^n = (\partial_{gh}H_r \partial_{ab}G_r \partial_{jk}H_s \partial_{lm}G_s)(C^n_i) (\lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm} - \Delta^{(h) \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm}} | F^n_i)),\]
\[\tilde{w}(3)_i^n = \left( (\partial_{gh}H_r \partial_{ab}G_r \partial_{jk}H_s \partial_{lm}G_s)(\hat{C}^n_i) - (\partial_{gh}H_r \partial_{ab}G_r \partial_{jk}H_s \partial_{lm}G_s)(C^n_i) \right) \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm},\]
\[
\hat{W}(u)_i^n = \sum_{i=1}^{[T/\Delta_n] - 4k_n + 1} \tilde{w}_i(u), \quad u = 1, 2, 3.
\]

Now, note that we also have \(\hat{W}_i^n = \hat{W}(1)_i^n + \hat{W}(2)_i^n + \hat{W}(3)_i^n\). By Taylor expansion and using repeatedly the boundedness of \(C_i\), we obtain
\[
|\tilde{w}(3)_i^n| \leq (1 + \|\nu_i^n\|^{(p-1)}) \|\nu_i^n\| \|\lambda_i^n\| \|\lambda_i^{n+2k_n}\|^2,
\]
which implies \(E(|\tilde{w}(3)_i^n|) \leq K \Delta_n^{5/4}\) and hence \(\hat{W}(3)_i^n \xrightarrow{p} 0\). Using Cauchy-Schwartz inequality and the bound \(E(|\tilde{w}(2)_i^n|) \leq K \Delta_n^{3/4}\), we have \(E(|\tilde{w}(2)_i^n|) \leq K \Delta_n^{3/4}\). Observing furthermore that \(\tilde{w}(2)_i^n\) is \(F_{i+4k_n}\)-measurable, Lemma B.8 in Ait-Sahalia and Jacod (2014) implies \(\hat{W}(2)_i^n \xrightarrow{p} 0\).

Next, define
\[
w_i^n = (\partial_{gh}H_r \partial_{ab}G_r \partial_{jk}H_s \partial_{lm}G_s)(C^n_i) \left[ \frac{4}{k_n^2 \Delta_n^5} \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm} \right] + \frac{4}{3} \left[ (\partial_{gh}H_r \partial_{ab}G_r \partial_{jk}H_s \partial_{lm}G_s)(C^n_i) \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm} \right] + \frac{4}{9} \left[ (\partial_{gh}H_r \partial_{ab}G_r \partial_{jk}H_s \partial_{lm}G_s)(C^n_i) \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm} \right] + \frac{4 \Delta_n^{1/8}}{9} \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm}.
\]
Using the cadlag property of \(c\) and \(C\), \(k_n \sqrt{\Delta_n} \rightarrow \theta\), and the Riemann integral convergence, we conclude that \(W_T^n \xrightarrow{p} W_T\) where
\[
W_T = \int_0^T (\partial_{gh}H_r \partial_{ab}G_r \partial_{jk}H_s \partial_{lm}G_s)(C^n_i) \left[ \frac{4}{k_n^2 \Delta_n^5} \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm} \right] + \frac{4}{3} \left[ (\partial_{gh}H_r \partial_{ab}G_r \partial_{jk}H_s \partial_{lm}G_s)(C^n_i) \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm} \right] + \frac{4 \Delta_n^{1/8}}{9} \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm} dt.
\]
In addition, by Lemma A4, it holds that
\[
E(|\hat{W}(1)_i^n - W_T^n|) \leq \Delta_n E \left( \sum_{i=1}^{[T/\Delta_n] - 4k_n + 1} (\Delta_n^{1/8} + \eta_i^{4k_n}) \right).
\]
Hence, by the third result of Lemma A1 we have \(\hat{W}_i^n \xrightarrow{p} W_i\), from which it follows that
\[
\frac{9}{4 \Delta_n^2} \hat{W}(1)_i^n + \frac{4}{k_n^2 \Delta_n^5} \sum_{i=1}^{[T/\Delta_n] - 4k_n + 1} (\partial_{gh}H_r \partial_{ab}G_r \partial_{jk}H_s \partial_{lm}G_s)(\hat{C}^n_i) [C^n_i (j,k,l,m) (\tilde{C}^n_i (gh,ab)]
\]
\[
- \frac{2}{k_n} \sum_{i=1}^{[T/\Delta_n] - 4k_n + 1} (\partial_{gh}H_r \partial_{ab}G_r \partial_{jk}H_s \partial_{lm}G_s)(\hat{C}^n_i) \left( \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm} \right)
\]
\[
- \frac{2}{k_n} \sum_{i=1}^{[T/\Delta_n] - 4k_n + 1} (\partial_{gh}H_r \partial_{ab}G_r \partial_{jk}H_s \partial_{lm}G_s)(\hat{C}^n_i) \left( \lambda_i^{n,gh} \lambda_i^{n,jk} \lambda_i^{n,ab} \lambda_i^{n,lm} \right)
\]
\[
\text{for } i = 1, \ldots, p.
\]
The result follows from the above convergence, the already invoked symmetry argument, and straightforward calculations.

C Proofs of Auxiliary Results

This section is devoted to the proofs of the auxiliary theorems and lemmas (listed in Section A.2) that were used to prove Theorem 1 and Theorem 2.

C.1 Proof of Theorem A1

To show this result, let us define the functions

\[ R(x, y) = \sum_{g,h,a,b=1}^{d} \left( \partial_{g h} H \partial_{a b} G \right)(x) \left( y^{g h} - x^{g h} \right) \left( y^{a b} - x^{a b} \right) \]

\[ S(x, y) = \left( H(y) - H(x) \right) \left( G(y) - G(x) \right) \]

\[ U(x) = \sum_{g,h,a,b=1}^{d} \left( \partial_{g h} H \partial_{a b} G \right)(x) \left( x^{g a} x^{h b} + x^{g b} x^{h a} \right) \]

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

\[ [H(C), G(C)]_{T}^{AN} - [H(C), G(C)]_{T}^{AN'} = \frac{3}{2k_n} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} \left[ (S(C_i, \hat{C}_i^{n+k_n}) - S(\hat{C}_i^{n}, \hat{C}_i^{n+k_n})) - \frac{2}{k_n} (U(C_i^{n}) - U(\hat{C}_i^{n})) \right] \]

\[ [H(C), G(C)]_{T}^{LIN} - [H(C), G(C)]_{T}^{LIN'} = \frac{3}{2k_n} \sum_{i=1}^{[T/\Delta_n]-2k_n+1} \left[ (R(C_i, \hat{C}_i^{n+k_n}) - R(\hat{C}_i^{n}, \hat{C}_i^{n+k_n})) - \frac{2}{k_n} (U(C_i^{n}) - U(\hat{C}_i^{n})) \right] \]

By (3.11) in Jacob and Rosenbaum (2015), there exists a sequence of real numbers \( a_n \) converging to zero such that

\[ \mathbb{E}(\|C_i^{n} - \hat{C}_i^{n}\|^q) \leq K q a_n \Delta_n^{2q-2\alpha+1} q, \text{ for any } q > 0. \]  

(C.26)

Since \( H \) and \( G \in G(p) \), the functions \( R \) and \( S \) are continuously differentiable and satisfy

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

(C.27)

\[ \|\partial U(x)\| \leq K(1 + \|x\|)^{2p-1}, \]  

(C.28)

where \( \partial J (\text{respectively, } \partial U) \) 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 \( x \)). Using the Taylor expansion, the Jensen's inequality, (C.27) and (C.28), it holds that, for \( J \in \{S, R\}, \)

\[ |J(C_i^{n}, \hat{C}_i^{n+k_n}) - J(\hat{C}_i^{n}, C_i^{n+k_n})| \leq K(1 + \|\hat{C}_i^{n}\|^{2p-1} + \|\hat{C}_i^{n+k_n}\|^{2p-1}) \]

\[ \times (\|\hat{C}_i^{n} - C_i^{n}\| + \|\hat{C}_i^{n+k_n} - C_i^{n+k_n}\| + K\|\hat{C}_i^{n} - C_i^{n}\|^2 + K\|\hat{C}_i^{n+k_n} - C_i^{n+k_n}\|^2) \]  

and

\[ |U(C_i^{n}) - U(\hat{C}_i^{n})| \leq K(1 + \|\hat{C}_i^{n}\|^{2p-1}(\|\hat{C}_i^{n} - C_i^{n}\| + K\|\hat{C}_i^{n} - C_i^{n}\|^2). \]  

45
By (3.20) in Jacob and Rosenbaum (2015), we have \( \mathbb{E}(\|\tilde{C}_i^{tn}\|^r) \leq K_r \), for any \( r \geq 0 \). Hence by Hölder inequality, for \( \epsilon > 0 \) fixed,

\[
\mathbb{E}(\|\tilde{C}_i^{tn}\|^{2p-2}\|\tilde{C}_i^{tn} - \tilde{C}_i^{tn}\|) \leq \left( \mathbb{E}(\|\tilde{C}_i^{tn}\|^{(1+\epsilon)}) \right)^{1/1+\epsilon} \left( \mathbb{E}(\|\tilde{C}_i^{tn}\|^{2p-2(1+\epsilon)/\epsilon}) \right)^{1/1+\epsilon} \\
\leq K_p \left( \mathbb{E}(\|\tilde{C}_i^{tn} - \tilde{C}_i^{tn}\|^{(1+\epsilon)}) \right)^{1/1+\epsilon} \\
\leq K_p a_n \Delta_n (2^{-\frac{1}{1+\epsilon}})^{w+\frac{1}{1+\epsilon}-1}.
\]

Using the above result and (C.26), the following conditions are sufficient for Theorem A1 to hold:

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

Using the fact that \( 0 < \omega < \frac{1}{2} \), and taking \( \epsilon \) sufficiently close to zero, we can see that Theorem A1 holds when \( (8p - 1)/4(4p-r) \leq \omega < \frac{1}{2} \), which completes the proof.

### C.2 Proof of Theorem A2

Note that we have

\[
[H(\tilde{C}), G(C)]_{T}^{L1N'} - [H(\tilde{C}), G(C)]_{T}^{A} = \frac{3}{2k_n} \sum_{g,h,a,b=1}^{d} [T/\Delta_n]^{-2k_n+1} \psi_i^n(g, h, a, b),
\]

\[
[H(\tilde{C}), G(C)]_{T}^{AN'} - [H(\tilde{C}), G(C)]_{T}^{A} = \frac{3}{2k_n} \sum_{i=1}^{d} [T/\Delta_n]^{-2k_n+1} \chi_i^n - \sum_{g,h,a,b=1}^{d} (\partial_{gh}H\partial_{ab}G)(C^n_i) \lambda_i^{n,gh} \lambda_i^{n,ab},
\]

with

\[
\psi_i^n(g, h, a, b) = \left( (\partial_{gh}H\partial_{ab}G)(\tilde{C}_i^n) - (\partial_{gh}H\partial_{ab}G)(C^n_i) \right) \lambda_i^{n,gh} \lambda_i^{n,ab},
\]

\[
\chi_i^n = \left( H(\tilde{C}_{i+k,n}) - H(\tilde{C}_i^n) \right) \left( G(\tilde{C}_{i+k,n}) - G(\tilde{C}_i^n) \right).
\]

By Taylor expansion, we have

\[
(\partial_{gh}S\partial_{ab}G)(\tilde{C}_i^n) - (\partial_{gh}S\partial_{ab}G)(C^n_i) = \sum_{x,y=1}^{d} \left( \partial_{xy,gh}S\partial_{ab}G + \partial_{xy,gh}^2G\partial_{gh}S \right)(C^n_i) \nu_i^{n,xy}
\]

\[
+ \frac{1}{2} \sum_{j,k,x,y=1}^{d} \left( \partial_{j,k,xy,gh}^3S\partial_{ab}G + \partial_{j,k,xy,gh}^2G\partial_{j,k,gh}S \right)(C^n_i) \nu_i^{n,xy} \nu_i^{n,jk}
\]

and

\[
S(\tilde{C}_{i+k,n}) - S(\tilde{C}_i^n) = \sum_{gh} \partial_{gh}S(C^n_i) \lambda_i^{n,gh} + \sum_{j,k,g,h} \partial_{j,k,gh}^2S(C^n_i) \lambda_i^{n,gh} \nu_i^{n,jk}
\]

\[
+ \frac{1}{2} \sum_{x,y,g,h} \partial_{x,y,gh}^2S(C^n_i) \lambda_i^{n,gh} \nu_i^{n,xy} + \frac{1}{2} \sum_{x,y,j,k,g,h} \partial_{x,y,j,k,gh}^3S(C^n_i) \lambda_i^{n,gh} \nu_i^{n,xy} \nu_i^{n,jk}
\]

\[
+ \frac{1}{6} \sum_{j,k,x,y,g,h} \partial_{j,k,x,y,gh}^3S(C^n_i) \lambda_i^{n,jk} \lambda_i^{n,gh} \lambda_i^{n,xy},
\]

for \( S \in \{ H, G \} \), \( \tilde{c}_i^n = \pi C_i^n + (1 - \pi) \tilde{C}_i^n \), \( C_i^{n,S} = \pi S \tilde{C}_i^n + (1 - \pi S) \tilde{C}_i^{n,k} \), \( CC_i^{n,S} = \mu S C_i^n + (1 - \mu S) \tilde{C}_i^{n,k} \) for \( \pi, \mu_H, \mu_G, \mu_G \in [0,1] \). Although \( \tilde{c}_i^n \) and \( \pi \) depend on \( g, h, a, b \), we do not emphasize this in our
notation to simplify the exposition.

Set \( F_n^i = F_{(i-1)} \Delta_n \). By (4.10) in Jacob and Rosenbaum (2013) we have

\[
\mathbb{E}\left(\left\| \alpha_i^n \right\|_q F_n^i \right) \leq K \Delta_n^q \quad \text{for all } q \geq 0 \quad \text{and} \quad \mathbb{E}\left( \left| \sum_{j=0}^{k-1} \alpha_{i+j}^n \right| F_n^i \right) \leq K \Delta_n^q k_n^{q/2} \quad \text{whenever } q \geq 2. \tag{C.29}
\]

Combining (C.29), (A.4), (A.10) with \( Z = c \) and the Hölder inequality yields for \( q \geq 2 \),

\[
\mathbb{E}\left( \left\| \nu_i^n \right\|_q F_n^i \right) \leq K \Delta_n^{q/4} \quad \text{and} \quad \mathbb{E}\left( \left\| \lambda_i^n \right\|_q F_n^i \right) \leq K \Delta_n^{q/4} \tag{C.30}
\]

The bound in the first equation of (C.30) is tighter than that in (4.11) of Jacob and Rosenbaum (2015) due to the absence of volatility jumps. This tighter bound will be useful later in deriving the asymptotic distribution for the approximated estimator. By the boundedness of \( C_t \) and the polynomial growth assumption, we have

\[
\left| (\partial_{jk,xy,ab}^3 G \partial_{gh} H + \partial_{xy,gh}^2 H \partial_{jk,ab}^2 G)(\tilde{C}_i^n) \nu_i^n,xy \nu_i^n,jk \lambda_i^n,gh \lambda_i^n,ab \right| \leq K (1 + \| \nu_i^n \|)^{2(p-2)} \| \nu_i^n \|^2 \| \lambda_i^n \|^2.
\]

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

\[
\left| (\partial_{jk,xy,ab}^3 G \partial_{gh} H + \partial_{xy,gh}^2 H \partial_{jk,ab}^2 G)(\tilde{C}_i^n) \nu_i^n,xy \nu_i^n,jk \lambda_i^n,gh \lambda_i^n,ab \right| \leq K (1 + \| \nu_i^n \|^{2(p-2)}) \| \nu_i^n \|^2 \| \lambda_i^n \|^2. \tag{C.31}
\]

Using the Taylor expansion, the polynomial growth assumption and using similar idea as for (C.31), we have

\[
\chi_i^n - \sum_{g,h,a,b} (\partial_{gh} H \partial_{ab} G)(C_i^n) \lambda_i^{n,gh} \lambda_i^{n,ab} =
\sum_{g,h,a,b,j,k} (\partial_{gh} H \partial_{jk,xy}^2 G + \partial_{gh} G \partial_{jk,xy}^2 H)(C_i^n) (\lambda_i^{n,gh} + \frac{1}{2} \nu_i^{n,gh}) \lambda_i^{n,ab} \lambda_i^{n,jk} + \phi_i^n, \text{ and}
\sum_{g,h,a,b} (\partial_{gh} H \partial_{ab} G)(\tilde{C}_i^n) - (\partial_{gh} H \partial_{ab} G)(C_i^n) =
\sum_{g,h,a,b,j,k} (\partial_{gh} H \partial_{jk,xy}^2 G + \partial_{gh} G \partial_{jk,xy}^2 H)(C_i^n)(\nu_i^{n,xy}) \lambda_i^{n,gh} \lambda_i^{n,ab} + \delta_i^n
\]

with \( \mathbb{E}(\| \nu_i^n \| F_n^i) \leq K \Delta_n \) and \( \mathbb{E}(\| \delta_i^n \| F_n^i) \leq K \Delta_n \) which follow by the Cauchy-Schwarz inequality together with (C.30). Given that \( k_n = \theta(\Delta_n)^{-1/2} \), the previous inequalities imply

\[
\frac{3 \Delta_n^{-1/4} |T/\Delta_n - 2k_n + 1|}{2k_n} \sum_{i=1}^{3 \Delta_n^{-1/4} |T/\Delta_n - 2k_n + 1|} \phi_i^n \overset{p}{\longrightarrow} 0 \quad \text{and} \quad \frac{3 \Delta_n^{-1/4} |T/\Delta_n - 2k_n + 1|}{2k_n} \sum_{i=1}^{3 \Delta_n^{-1/4} |T/\Delta_n - 2k_n + 1|} \delta_i^n \overset{p}{\longrightarrow} 0.
\]

Therefore, it suffices to show that

\[
\frac{3 \Delta_n^{-1/4} |T/\Delta_n - 2k_n + 1|}{2k_n} \sum_{i=1}^{3 \Delta_n^{-1/4} |T/\Delta_n - 2k_n + 1|} (\partial_{gh} H \partial_{jk,ab}^2 G + \partial_{gh} H \partial_{jk,ab}^2 G)(C_i^n) \lambda_i^{n,gh} \lambda_i^{n,ab} \lambda_i^{n,jk} \overset{p}{\longrightarrow} 0, \tag{C.32}
\]

\[
\frac{3 \Delta_n^{-1/4} |T/\Delta_n - 2k_n + 1|}{2k_n} \sum_{i=1}^{3 \Delta_n^{-1/4} |T/\Delta_n - 2k_n + 1|} (\partial_{gh} H \partial_{jk,ab}^2 G + \partial_{gh} H \partial_{jk,ab}^2 G)(C_i^n) \nu_i^{n,gh} \lambda_i^{n,ab} \lambda_i^{n,jk} \overset{p}{\longrightarrow} 0. \tag{C.33}
\]

These results hold by the bounds in A.5.
C.3 Proof of Theorem A3

First, we decompose the approximated estimator as

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

with

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

and

$$[H(C), G(C)]^{(A2)}_T = \frac{3}{2k_n} \sum_{g, h, a, b=1}^{d} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor - 2k_n + 1} (\partial_{gh} H \partial_{ab} G)(C_i^n) (\hat{C}_{i}^{n, ab} - \hat{C}_{i}^{n, ha}).$$

In this section, we use the notation $C_{i-1}^n = C_{(i-1)\Delta_n}$ and $F_i = F_{(i-1)\Delta_n}$ 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 (2015) we have

$$\frac{1}{\Delta_n^{1/4}} \left( [H(C), G(C)]^{(A2)}_T - \frac{3}{2g} \sum_{g, h, a, b=1}^{d} \int_0^T (\partial_{gh} H \partial_{ab} G)(C_i^n) (\hat{C}_{i}^{n, ab} + \hat{C}_{i}^{n, ha}) dt \right) = O_p(1),$$

which yields

$$\frac{1}{\Delta_n^{1/4}} \left( [H(C), G(C)]^{(A2)}_T - \frac{3}{2g} \sum_{g, h, a, b=1}^{d} \int_0^T (\partial_{gh} H \partial_{ab} G)(C_i^n) (\hat{C}_{i}^{n, ab} + \hat{C}_{i}^{n, ha}) dt \right) \xrightarrow{p} 0.$$

Using the multivariate quantities defined in Section A.1, we can show that the following decompositions hold:

$$\hat{C}_{i}^{n} = C_{i-1}^n + \frac{1}{k_n} \sum_{j=0}^{k_n-1} \sum_{u=1}^{2} \varepsilon(u)_j^n \zeta(u)_{i+j}, \quad \hat{C}_{i+k_n}^{n} - \hat{C}_{i}^{n} = \frac{1}{k_n} \sum_{j=0}^{2k_n-1} \sum_{u=1}^{2} \varepsilon(u)_j^n \zeta(u)_{i+j},$$

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

$$+ \sum_{j=0}^{2k_n-2} \sum_{q=0}^{j+1} \varepsilon(u)_j^n \varepsilon(v)_q^n \zeta(u)_{i+j} \zeta(v)_{i+q} + \sum_{j=0}^{2k_n-2} \sum_{q=0}^{j+1} \varepsilon(u)_j^n \varepsilon(v)_q^n \zeta(u)_{i+j} \zeta(v)_{i+q}.$$
Now, set

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

and

$$\mathcal{A}_{11}(H, gh, u; G, ab, v)_{i}^{n} = \frac{3}{2k_{n}^{3}} \sum_{i=1}^{[T/D_{n}] - 2k_{n} - 2k_{n} - 1} \sum_{j=0}^{k_{n}} \sum_{q=0}^{(i-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1}^{n}) \varepsilon(u)_{j}^{n} \varepsilon(v)_{q}^{n} \zeta(u)_{i+j}^{n, gh} \zeta(v)_{i+j}^{n, ab},$$

$$\mathcal{A}_{12}(H, gh, u; G, ab, v)_{i}^{n} = \frac{3}{2k_{n}^{3}} \sum_{i=1}^{[T/D_{n}] - 2k_{n} - 2k_{n} - 1} \sum_{j=0}^{k_{n}} \sum_{q=0}^{(i-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1}^{n}) \varepsilon(u)_{j}^{n} \varepsilon(v)_{q}^{n} \zeta(u)_{i+j}^{n, gh} \zeta(v)_{i+j}^{n, ab},$$

$$\mathcal{A}_{13}(H, gh, u; G, ab, v)_{i}^{n} = \frac{3}{2k_{n}^{3}} \sum_{i=1}^{[T/D_{n}] - 2k_{n} - 2k_{n} - 1} \sum_{j=0}^{k_{n}} \sum_{q=0}^{(i-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1}^{n}) \varepsilon(u)_{j}^{n} \varepsilon(v)_{q}^{n} \zeta(u)_{i+j}^{n, gh} \zeta(v)_{i+j}^{n, ab},$$

where we clearly have $\mathcal{A}_{13}(H, gh, u; G, ab, v)_{i}^{n} = \mathcal{A}_{12}(G, ab, v; H, gh, u)_{i}^{n}$. By a change of the order of the summation,

$$\mathcal{A}_{11}(H, gh, u; G, ab, v)_{i}^{n} = \frac{3}{2k_{n}^{3}} \sum_{i=1}^{[T/D_{n}] - 2k_{n} - 1 \wedge (i-1)} \sum_{j=0}^{(k_{n}-1) \wedge (i-1)} \sum_{q=0}^{(i-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1}^{n}) \varepsilon(u)_{j}^{n} \varepsilon(v)_{q}^{n} \zeta(u)_{i+j}^{n, gh} \zeta(v)_{i+j}^{n, ab},$$

$$\mathcal{A}_{12}(H, gh, u; G, ab, v)_{i}^{n} = \frac{3}{2k_{n}^{3}} \sum_{i=1}^{[T/D_{n}] - 2k_{n} - 1 \wedge (i-1)} \sum_{j=0}^{(k_{n}-1) \wedge (i-1)} \sum_{q=0}^{(i-1) \wedge (k_{n}-1)} \sum_{m=1}^{k_{n}} \sum_{j=0}^{(i-1) \wedge (k_{n}-1) \wedge (i-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1}^{n}) \varepsilon(u)_{j}^{n} \varepsilon(v)_{j+m}^{n} \zeta_{gh}(u)_{i-m}^{n} \zeta_{ab}(v)_{i}^{n}.$$
with
\[ \rho_{gh}(u, v)^n_i = \sum_{m=1}^{2k_n-1} \lambda(u, v)_m^n \zeta_{gh}(u)^{n-m}_i. \]

We show below that the following results hold:
\[ \frac{1}{\Delta_n^{1/4}} \left( \tilde{A}\bar{1} w(H, gh, u; G, ab, v)^n_i - \tilde{A}\bar{1} w(H, gh, u; G, ab, v)^n_T \right) \xrightarrow{P} 0 \quad (C.35) \]
\[ \frac{1}{\Delta_n^{1/4}} \left( \tilde{A}\bar{1} w(H, gh, u; G, ab, v)^n_i - \tilde{A}\bar{1} w(H, gh, u; G, ab, v)^n_T \right) \xrightarrow{P} 0 \quad (C.36) \]
for all \((H, gh, u, G, ab, v)\) and \(w = 1, 2\).

### C.3.1 Proof of (C.35) for \(w = 1\)

To prove this result, first, notice that the \(\zeta(u)^{n, gh \zeta(v)}^n_i\) are scaled by random variables rather than constant real numbers. Next, observe that we can write

\[ \tilde{A}\bar{1} - \bar{A}\bar{1} = \tilde{A}\bar{1}(1) + \tilde{A}\bar{1}(2) + \tilde{A}\bar{1}(3) \text{ with} \]

\[
\tilde{A}\bar{1}(1) = \sum_{i=1}^{(2k_n-1)\wedge[T/\Delta_n]} \frac{3}{2k_n^3} \sum_{j=0}^{(2k_n-1)\wedge(i-1)} (\partial_{gh} H \partial_{ab} G)(C^{n}_{i-j-1}) \varepsilon(u)^n_j \varepsilon(v)^n_j \zeta(u)^{n, gh \zeta(v)}^{n, ab}_i.
\]

\[ \tilde{A}\bar{1}(2) = \sum_{i=[T/\Delta_n]-2k_n+2}^{T/\Delta_n} \frac{3}{2k_n^3} \left( \sum_{j=0}^{(2k_n-1)\wedge(i-1)} (\partial_{gh} H \partial_{ab} G)(C^{n}_{i-j-1}) \varepsilon(u)^n_j \varepsilon(v)^n_j - \sum_{j=0}^{(2k_n-1)} (\partial_{gh} H \partial_{ab} G)(C^{n}_{i-j-1}) \varepsilon(u)^n_j \varepsilon(v)^n_j \zeta(u)^{n, gh \zeta(v)}^{n, ab}_i \right), \]

\[ \tilde{A}\bar{1}(3) = \sum_{i=2k_n+1}^{[T/\Delta_n]-2k_n+1} \frac{3}{2k_n^3} \left( \sum_{j=0}^{(2k_n-1)\wedge(i-1)} (\partial_{gh} H \partial_{ab} G)(C^{n}_{i-j-1}) \varepsilon(u)^n_j \varepsilon(v)^n_j - \sum_{j=0}^{(2k_n-1)-1} (\partial_{gh} H \partial_{ab} G)(C^{n}_{i-j-1}) \varepsilon(u)^n_j \varepsilon(v)^n_j \zeta(u)^{n, gh \zeta(v)}^{n, ab}_i \right). \]

It is easy to see that \(\tilde{A}\bar{1}(3) = 0\). Using (A.10) with \(Z = c\) and (C.29), we obtain

\[ E(||(1)^n_i||q F_{i-1}^n) \leq K_q, \quad E(||(2)^n_i||q F_{i-1}^n) \leq K_q \Delta_n^{q/2}. \quad (C.37) \]

The polynomial growth assumption on \(H\) and \(G\) and the boundedness of \(C_t\) imply that \(|(\partial_{gh} H \partial_{ab} G)(C^{n}_{i-j-1})| \leq K\).

Hence, the random quantities \(\left\{ \frac{3}{2k_n^3} \sum_{j=0}^{(2k_n-1)\wedge(i-1)} (\partial_{gh} H \partial_{ab} G)(C^{n}_{i-j-1}) \varepsilon(u)^n_j \varepsilon(v)^n_j \right\} \) and \(\frac{3}{2k_n^3} \sum_{j=0}^{(2k_n-1)-1} (\partial_{gh} H \partial_{ab} G)(C^{n}_{i-j-1}) \varepsilon(u)^n_j \varepsilon(v)^n_j \) are \(F_{i-1}^n\) measurable and are bounded by \(\lambda_{u, v}^n\) defined as

\[
\lambda_{u, v}^n = \begin{cases} 
K, & \text{if } (u, v) = (2, 2) \\
K/k_n, & \text{if } (u, v) = (1, 2), (2, 1) \\
K/k_n^2, & \text{if } (u, v) = (1, 1). 
\end{cases}
\]
Similarly, the quantity
\[
\frac{3}{2k_n^3} \left( \sum_{j=0}^{(2k_n-1)^2(i-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1-j-m}^n) \varepsilon(u)_j^n \varepsilon(v)_j^n - \sum_{j=0}^{(2k_n-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1-j-m}^n) \varepsilon(u)^n_j \varepsilon(v)_j^n \right)
\]
is \( F_{i-1}^m \) measurable and bounded by \( \tilde{\lambda}_{u,v}^n \). Note also that, by (C.37) and the Cauchy Schwartz inequality, we have
\[
E(|\varsigma(u)^n_i \varsigma(v)^n_i \hat{A}_{i-1}^n| \leq E(\|\varsigma(u)^n_i \|_2^2 | F_{i-1}^m)^{1/2} E(\|\varsigma(v)^n_i \|_2^2 | F_{i-1}^m)^{1/2}
\]

\[
\begin{cases}
K \Delta_n & \text{if } (u,v) = (2,2) \\
K \Delta_n^{1/2} & \text{if } (u,v) = (1,2), (2,1) \\
K & \text{if } (u,v) = (1,1).
\end{cases}
\]

The above bounds, together with the fact that \( k_n = \theta \Delta_n^{-1/2} \), imply \( E(A_{11}(1)) \leq K \Delta_n^{1/2} \) and \( E(A_{11}(2)) \leq K \Delta_n^{1/2} \) for all \((u,v)\). These two results together imply \( A_{11}(1) = o(\Delta_n^{-1/4}) \) and \( A_{11}(2) = o(\Delta_n^{-1/4}) \), which yields the result.

C.3.2 Proof of (C.35) for \( w = 2 \)

First, observe that \( \tilde{A}_{12} - \tilde{A}_{12} = \tilde{A}_{12}(1) + \tilde{A}_{12}(2) \), with
\[
\tilde{A}_{12}(1) = \sum_{i=2}^{(2k_n-1)^2(i-1)} \left( \sum_{j=0}^{(2k_n-1)^2(i-1)} \left( \frac{3}{2k_n^3} \sum_{j=0}^{(2k_n-1)^2(i-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1-j-m}^n) \varepsilon(u)_j^n \varepsilon(v)_j^n \right) \varsigma_{gh}(u)_i^n \right)
\]
\[
\times \varsigma_{gh}(u)_i^n,
\]
\[
\tilde{A}_{12}(2) = \sum_{i=[T/\Delta_n]-2k_n+2}^{[T/\Delta_n]} \left( \sum_{j=0}^{(2k_n-1)^2(i-1)} \left( \frac{3}{2k_n^3} \sum_{j=0}^{(2k_n-1)^2(i-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1-j-m}^n) \varepsilon(u)_j^n \varepsilon(v)_j^n \right) \varsigma_{gh}(u)_i^n \right)
\]
\times \varepsilon(v)_j^n,m - \sum_{j=0}^{(2k_n-1)^2(i-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1-j-m}^n) \varepsilon(u)_j^n \varepsilon(v)_j^n \varsigma_{gh}(u)_i^n \right) \varsigma_{gh}(u)_i^n.
\]

Notice that the quantity
\[
\kappa_{i,m}^n = \frac{3}{2k_n^3} \left( \sum_{j=0}^{(2k_n-1)^2(i-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1-j-m}^n) \varepsilon(u)_j^n \varepsilon(v)_j^n \right)
\]
is \( F_{i-m}^m \) measurable and bounded by \( \tilde{\lambda}_{u,v}^n \). Let
\[
\kappa_{i}^n = \sum_{m=1}^{(2k_n-1)^2(i-1)} \left( \frac{3}{2k_n^3} \sum_{j=0}^{(2k_n-1)^2(i-1)} (\partial_{gh} H \partial_{ab} G)(C_{i-1-j-m}^n) \varepsilon(u)_j^n \varepsilon(v)_j^n \right) \varsigma_{gh}(u)^n_{i-m}.
\]

It follows that \( \kappa_{i}^n \) is \( F_{i-1}^m \)-measurable and we have
\[
E(|\kappa_{i,m}^n|^2 | F_0) \leq \tilde{\lambda}_{u,v}^n^2,
\]
\[
|E(\varsigma(u)^n_i \varsigma(v)^n_i | F_{i-1}^m)| \leq \begin{cases}
K \sqrt{\Delta_n} & \text{if } u = 1 \\
K \Delta_n & \text{if } u = 2.
\end{cases}
\]

51
\[
E(||\zeta(u)_{i-m}^n||^2|\mathcal{F}_{i-m-1}) \leq \begin{cases} 
K_z & \text{if } u = 1 \\
K_z\Delta_n^{3/2} & \text{if } u = 2 
\end{cases}
\]

Using Lemma A3, we deduce that for \( z \geq 2 \),
\[
E(|\kappa_i^{|z^2}|^2) \leq \begin{cases} 
K_zk_n^{-3/2} & \text{if } u = 1 \\
K_zk_n^{-3/2} & \text{if } v = 1 \\
K_zk_n^{-3/2} & \text{if } v = 2 
\end{cases}
\]

Using the above result, we obtain \( \frac{1}{\Delta_n^{0.4}} \# A12(1) \overset{p}{\Rightarrow} 0 \). A similar argument yields \( \frac{1}{\Delta_n^{0.4}} \# A12(2) \overset{p}{\Rightarrow} 0 \), which completes the proof of (C.35) for \( w = 2 \).

C.3.3 Proof of (C.36) for \( w = 1 \)

Define
\[
\Theta(u, v)_0^{(C), i, n} = \frac{3}{2\kappa^2_n} \sum_{j=0}^{2k_n-1} \left( (\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n) - (\partial_{gh}H\partial_{ab}G)(C_{i-2k_n}^n) \right) \varepsilon(u)^n_iz(v)^n_j.
\]

By Taylor expansion, the polynomial growth assumption on \( H \) and \( G \) and using (A.10) with \( Z = c \), we have
\[
\mathbb{E}\left( \left| (\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n) - (\partial_{gh}H\partial_{ab}G)(C_{i-2k_n}^n) \right| \mathcal{F}_{i-2k_n}^n \right) \leq K(k_n\Delta_n) \leq K\sqrt{\Delta_n}.
\]

Using (A.10), we have \( \mathbb{E}\left( \left| \Theta(u, v)_0^{(C), i, n} \right| \mathcal{F}_{i-2k_n}^n \right) \leq K\Delta_n^{1/4}\tilde{\lambda}_{u,v}^n \). Hence, \( \mathbb{E}\left( \left| \Theta(u, v)_0^{(C), i, n} \right| \mathcal{F}_{i-2k_n}^n \right) \leq K\Delta_n^{1/4}\tilde{\lambda}_{u,v}^n \) for \( q \geq 2 \) and \( j = 0, \ldots, 2k_n - 1 \). Next, observe that \( \Theta(u, v)_0^{(C), i, n} \) is \( \mathcal{F}_{i-1}^n \)-measurable and satisfies \( |\Theta(u, v)_0^{(C), i, n}| \mathcal{F}_{i-2k_n}^n = \mathbb{E}\left( \Theta(u, v)_0^{(C), i, n} | \mathcal{F}_{i-2k_n}^n \right) \leq K\Delta_n^{1/4}\tilde{\lambda}_{u,v}^n \). Therefore, \( \mathbb{E}(\Theta(u, v)_0^{(C), i, n} | \mathcal{F}_{i-2k_n}^n) \leq K\Delta_n^{1/4}\tilde{\lambda}_{u,v}^n \) for all \( H \), \( G \), \( g \), \( a \), \( b \) with \( u, v = 1, 2 \).

To show this result, we first introduce the following quantities:
\[
\hat{E}(1) = \frac{1}{\Delta_n^{1/4}} \left[ \sum_{i=2k_n}^{[T/\Delta_n]} \Theta(u, v)_0^{(C), i, n} \mathbb{E}(\zeta(u)^n_i g \zeta(v)^{n, ab}_i | \mathcal{F}_{i-1}^n) \right],
\]

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

with \( \hat{E} = \hat{E}(1) + \hat{E}(2) \). By Cauchy-Schwartz inequality, we have
\[
\mathbb{E}(\zeta(u)^n_i g \zeta(v)^{n, ab}_i | \mathcal{F}_{i-1}^n) \leq (\tilde{\lambda}_{u,v}^n)^{q/2},\text{ where } \tilde{\lambda}_{u,v}^n = \begin{cases} 
K & \text{if } (u, v) = (1, 1) \\
K\Delta_n & \text{if } (u, v) = (1, 2), (2, 1) \\
K\Delta_n^2 & \text{if } (u, v) = (2, 2) 
\end{cases}
\]

Since \( \zeta(u)^n_i g \zeta(v)^{n, ab}_i \) is \( \mathcal{F}_{i-1}^n \)-measurable, the martingale property of \( \zeta(u)^n_i g \zeta(v)^{n, ab}_i | \mathcal{F}_{i-1}^n \) implies, for all \( u, v \),
\[
\mathbb{E}(\hat{E}(2)^2) \leq K\Delta_n^{-3/2}(\Delta_n^{1/4}\tilde{\lambda}_{u,v}^n)^2\tilde{\lambda}_{u,v}^n \leq K\Delta_n.
\]

52
The latter inequality implies \( \hat{E}(2) \xrightarrow{p} 0 \) for all \((u, v)\). It remains to show that \( \hat{E}(1) \xrightarrow{p} 0 \).

Here, we recall some bounds under Assumption 2,

\[
\begin{align*}
|E(\zeta(1)n,gh(2)\mid F^n_i)| &\leq K\Delta_n, \\
|E(\zeta(1)n,gh(1)i,ab \mid F^n_i) - (C_i\Delta_n + C_i^1)\Delta_n| &\leq K\Delta_n^{1/2}, \\
|E(\zeta(2)i,gh(2)\mid F^n_i - C_i\Delta_n^n)| &\leq K\Delta_n^{3/2}(\sqrt{\Delta_n} + n_i).
\end{align*}
\]  

(C.38) (C.39) (C.40)

Case \((u, v) \in \{1, 2\}, (2, 1)\). By (C.38) we have

\[
E(\hat{E}(1)) \leq K\frac{T}{\Delta_n} \frac{1}{\Delta_n^{1/4}} (\Delta_n^{1/4} + \Delta_n) \leq K\Delta_n^{1/2} \quad \text{so} \quad \hat{E}(1) \xrightarrow{p} 0.
\]

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

\[
\begin{align*}
\hat{E}'(1) &= \frac{1}{\Delta_n^{1/4}} \left[ \sum_{i \in [2k_n]} \Theta(u, v)^{0,C,i,n} V^n_i \right], \\
\hat{E}''(1) &= \frac{1}{\Delta_n^{1/4}} \left[ \sum_{i \in [2k_n]} \Theta(u, v)^{0,C,i,n} (V^n_i - V^n_{i-2k_n}) \right], \\
\hat{E}'''(1) &= \frac{1}{\Delta_n^{1/4}} \left[ \sum_{i \in [2k_n]} \Theta(u, v)^{0,C,i,n} \left( E(\zeta(u)n,gh(2)v)\mid F^n_i - V^n_i \right) \right],
\end{align*}
\]

where

\[
V^n_{i-1} = \begin{cases} 
C_i^1 \Delta_n & \text{if } (u, v) = (2, 2) \\
C_i^1 \Delta_n & \text{if } (u, v) = (1, 1) \\
0 & \text{otherwise}
\end{cases}
\]

Note that we have \( \hat{E}(1) = \hat{E}'(1) + \hat{E}''(1) + \hat{E}'''(1) \). Using (C.39) and (C.40), it can be shown that

\[
E(\hat{E}'''(1)) \leq \begin{cases} 
K\frac{1}{\Delta_n^{1/4}} (\Delta_n^{1/4} + \Delta_n)^{1/2} & \text{if } (u, v) = (1, 1) \\
K\frac{1}{\Delta_n^{1/4}} (\Delta_n^{1/4} + \Delta_n)^{3/2} & \text{if } (u, v) = (2, 2)
\end{cases} \leq K\Delta_n^{1/2} \quad \text{in all cases}.
\]

Next, we prove \( \hat{E}'(1) \xrightarrow{p} 0 \). To this end, write

\[
\hat{E}'(1) = \frac{1}{\Delta_n^{1/4}} \left[ \sum_{i=1}^{[T/\Delta_n]} \Theta(u, v)^{0,C,i-1,2k_n,n} V^n_{(i-1)\Delta_n} \right].
\]

Using the \( F^n_{i+2k_n-2} \)-measurability of the last sum, we are able to show

\[
\begin{align*}
\frac{1}{\Delta_n^{1/4}} \left[ \sum_{i=1}^{[T/\Delta_n]} |E(\Theta(u, v)^{0,C,i-1,2k_n,n} V^n_{(i-1)\Delta_n} \mid F^n_i)| \right] \xrightarrow{p} 0 \\
\frac{2k_n - 2}{\Delta_n^{1/2}} \left[ \sum_{i=1}^{[T/\Delta_n]} E\left( |\Theta(u, v)^{0,C,i-1,2k_n,n} V^n_{(i-1)\Delta_n}|^2 \right) \right] \xrightarrow{p} 0.
\end{align*}
\]

The first result readily follows from the inequality

\[
|E(\Theta(u, v)^{0,C,i-1,2k_n,n} V^n_{(i-1)\Delta_n} \mid F^n_i)| \leq \begin{cases} 
K\Delta_n^{1/2} \Delta_n^{1/4} & \text{if } (u, v) = (1, 1) \\
K\Delta_n^{1/2} \Delta_n^{1/4} & \text{if } (u, v) = (2, 2)
\end{cases} \leq K\Delta_n^{3/2} \quad \text{in all cases},
\]

53
while the second is a direct consequence of
\[
\mathbb{E}(|\Theta(u,v)(C,i-1+2\kappa_n)^n V(i-1) \Delta_n|^2) \leq \begin{cases} 
K\Delta_{n}^{1/2} (\lambda_{u,v}^{n})^2 & \text{if } (u,v) = (1,1) \\
K\Delta_{n}^{1/2} (\lambda_{u,v}^{n})^2 \Delta_n^2 & \text{if } (u,v) = (2,2) 
\end{cases}
\]
in all cases.

Finally, to prove that \( \hat{E}''(1) \overset{P}{\longrightarrow} 0 \), we use the fact that
\[
\mathbb{E}(|\Theta(u,v)_{0}^{C,i,n}(V(i-1)\Delta_n - V(i-2k_n)\Delta_n)|) \leq \mathbb{E}(|\Theta(u,v)(C,i,n|^{2})^{1/2} \mathbb{E}(|V(i-1)\Delta_n - V(i-2k_n)\Delta_n|^{2})^{1/2} 
\leq \begin{cases} 
K\Delta_{n}^{1/2} \lambda_{u,v}^{n} & \text{if } (u,v) = (1,1) \\
K\Delta_{n}^{1/4} \lambda_{u,v}^{n} \Delta_n^{1/4} & \text{if } (u,v) = (2,2) 
\end{cases},
\]

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

**C.3.4 Proof of (C.36) for \( w = 2 \)**

Our aim here is to show that
\[
\hat{E}'(2) = \frac{1}{\Delta_n^{1/4}} \sum_{i=2k_n}^{[T/\Delta_n]} \sum_{j=0}^{2k_n-1} \left( \sum_{m=1}^{2k_n-m-1} \frac{3}{2k_n} \sum_{j=0}^{m} \left[ (\partial_{gh} H \partial_{ab} G)(c_{i,j,m-1}) - (\partial_{gh} H \partial_{ab} G)(c_{i,j-2k_n}) \right] \varepsilon(u)_{j} \varepsilon(v)^{n}_{j+m} \right) \times 
\zeta(u)^{n,gh}_{i-m} \zeta(v)^{n,ab}_{i} \overset{P}{\longrightarrow} 0.
\]

For this purpose, we introduce some new notation. For any \( 0 \leq m \leq 2k_n - 1 \), set
\[
\Theta(u,v)(C,i,n) = \frac{3}{2k_n} \sum_{j=0}^{2k_n-m-1} \left[ (\partial_{gh} H \partial_{ab} G)(c_{i,j,m-1}) - (\partial_{gh} H \partial_{ab} G)(c_{i,j-2k_n}) \right] \varepsilon(u)_{j} \varepsilon(v)^{n}_{j+m} 
\zeta(u)^{n,gh}_{i-m}.
\]

It is easy to see that \( \Theta(u,v)(C,i,n) \) is \( F_{i-m-1}^{n} \) measurable and satisfies, by Hölder inequality,
\[
|\Theta(u,v)(C,i,n)| \leq \lambda_{u,v}^{n} \quad \text{and} \quad \mathbb{E}(|\Theta(u,v)(C,i,n)|^{q} | F_{i-2k_n}^{n}) \leq K_q \Delta_{n}^{q/4} (\lambda_{u,v}^{n})^{q}.
\]

Lemma A3 implies that for \( q \geq 2 \),
\[
\mathbb{E}(|\rho(u,v)(C,i,n,gh)|^{q}) \leq \begin{cases} 
K_q (\Delta_{n}^{1/4} \lambda_{u,v}^{n})^{q} & \text{if } u = 1 \\
K_q (\Delta_{n}^{1/4} \lambda_{u,v}^{n})^{q} / \kappa_n^{q/2} & \text{if } u = 2 
\end{cases} \leq \begin{cases} 
K_q / \kappa_n^{q/2} & \text{if } v = 1 \\
K_q \kappa_n^{q/2} & \text{if } v = 2 \end{cases}. \quad (C.41)
\]

Set
\[
\hat{E}'(2) = \frac{1}{\Delta_n^{1/4}} \sum_{i=2k_n}^{[T/\Delta_n]} \rho(u,v)(C,i,n,gh) \mathbb{E}(\zeta(v)^{n,ab}_{i} | F_{i-1}^{n}), 
\]
\[
\hat{E}''(2) = \frac{1}{\Delta_n^{1/4}} \sum_{i=2k_n}^{[T/\Delta_n]} \rho(u,v)(C,i,n,gh) (\zeta(v)^{n,ab}_{i} - \mathbb{E}(\zeta(v)^{n,ab}_{i} | F_{i-1}^{n})).
\]
The martingale increments property implies \( \mathbb{E}((\tilde{E}'(2))^2) \leq K \Delta_n^{1/2} \) in all the cases, which in turn implies \( \tilde{E}'(2) \overset{p}{\to} 0 \). Next, using the bounds on \( \rho(u,v)(C_{i,n,\rho}) \), we obtain that \( \tilde{E}'(2) \overset{p}{\to} 0 \).

We refer to Jacod and Rosenbaum (2015) for the proofs of Lemma A1 and Lemma A2.

C.4 Proof of Lemma A3

Set

\[ \xi^n_i = \varphi^n_{i-1} \xi^n_i, \quad \xi^n_i = \mathbb{E}(\xi_i | F^n_{i-1}) = \mathbb{E}(\varphi^n_{i-1} \xi^n_i | F^n_{i-1}) = \varphi^n_{i-1} \mathbb{E}(\xi^n_i | F^n_{i-1}), \quad \text{and} \quad \xi^n_i = \xi^n_i - \xi^n_i. \]

Given that \( \|\mathbb{E}(\xi^n_i | F^n_{i-1})\| \leq L' \), we have \( \|\xi^n_i\| \leq L'|\varphi^n_{i-1}| \). By the convexity of the function \( x^q \), which holds for \( q \geq 2 \), we have

\[
\| \sum_{j=1}^{2k_n-1} \xi^n_{i+j} \|^q \leq K \left( \| \sum_{j=1}^{2k_n-1} \xi^n_{i+j} \|^q + \| \sum_{j=1}^{2k_n-1} \xi^n_{i+j} \|^q \right).
\]

Therefore, on the one hand we have

\[
\| \sum_{j=1}^{2k_n-1} \xi^n_{i+j} \|^q \leq K k_n^{q-1} \sum_{j=1}^{2k_n-1} \xi^n_{i+j} \|^q \leq K k_n^{q-1} L'^q \sum_{j=1}^{2k_n-1} |\varphi^n_{i+j-1}|^q,
\]

which by \( \mathbb{E}(\|\varphi^n_{i+j-1}\|^q | F^n_{i-1}) \leq L'^q \), satisfies

\[
\mathbb{E}(\| \sum_{j=1}^{2k_n-1} \xi^n_{i+j} \|^q | F^n_{i-1}) \leq KL'^q k_n^{q-1} \sum_{j=1}^{2k_n-1} \mathbb{E}(\|\varphi^n_{i+j-1}\|^q | F^n_{i-1}) \leq KL'^q k_n^{q} L'^q.
\]

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

\[
\mathbb{E}(\| \sum_{j=1}^{2k_n-1} \xi^n_{i+j} \|^q | F^n_{i-1}) \leq K_q L'^q L'^q k_n^{q/2}.
\]

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

C.5 Proof of Lemma A4

We use the terminology “successive conditioning” to refer to either of the following two equalities,

\[
\begin{align*}
x_1 y_1 - x_0 y_0 &= x_0 (y_1 - y_0) + y_0 (x_1 - x_0) + (x_1 - x_0) (y_1 - y_0), \\
x_1 y_1 z_1 - x_0 y_0 z_0 &= x_0 y_0 (z_1 - z_0) + x_0 z_0 (y_1 - y_0) + y_0 z_0 (x_1 - x_0) + x_0 (y_0 - y_1) (z_0 - z_1) \\
&\quad + y_0 (x_0 - x_1) (z_0 - z_1) + z_0 (x_0 - x_1) (y_0 - y_1) + (x_1 - x_0) (y_1 - y_0) (z_1 - z_0),
\end{align*}
\]

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

To prove Lemma A4, we first note that \( \lambda^n_{i,j,k} \lambda^n_{i,j,m} \) is \( \mathcal{F}^n_{i+2k_n} \)-measurable. Therefore, by the law of iterated
expectations, we have
\[ E\left(\lambda^{n,jk}_{i+2k} \lambda^{n,gh}_{i+2k} \lambda^{n,ab}_{i+2k} | \mathcal{F}^{n}_{i+2k} \right) = E\left(\lambda^{n,jk}_{i} \lambda^{n,gh}_{i} \lambda^{n,ab}_{i} | \mathcal{F}^{n}_{i+2k} \right) \].

By equation (3.27) in Jacod and Rosenbaum (2015), we have
\[ |E(\lambda^{n,jk}_{i} \lambda^{n,lm}_{i} | \mathcal{F}^{n}_{i}) - \frac{2}{k_{n}} (C^{n,ga}_{i+2k} C^{n,hb}_{i+2k} + C^{n,gb}_{i+2k} C^{n,ha}_{i+2k})| \leq K \sqrt{\Delta_{n}^{1/8} + \eta_{i+2k}^{2}}. \]

Now, using (A.10) successively with \( Z = c \) and \( Z = \mathcal{C} \) (recall that the latter holds under Assumption 2), together with the successive conditioning, we also have
\[ |E(\lambda^{n,jk}_{i} \lambda^{n,lm}_{i} | \mathcal{F}^{n}_{i}) - \frac{2}{k_{n}} (C^{n,ga}_{i+2k} C^{n,hb}_{i+2k} + C^{n,gb}_{i+2k} C^{n,ha}_{i+2k})| \leq K \Delta_{n}^{1/4} |\lambda^{n,jk}_{i} | |\lambda^{n,lm}_{i} | |\mathcal{F}^{n}_{i} |. \]

The result derives from the last inequality.

C.5.1 Proof of (A.12) in Lemma A5

We start by obtaining some useful bounds for some important quantities. First, using the second statement in Lemma A2 applied to \( Z = Y' \), we have
\[ |E(\alpha^{n,jk}_{i} | \mathcal{F}^{n}_{i} |) \leq K \Delta_{n}^{1/2} (\sqrt{\Delta_{n}^{1/8} + \eta_{i+1}^{n}}). \] (C.42)

Second, by repeated application of the Cauchy-Schwarz inequality and making use of the third and last statements in Lemma A2 as well as (A.10) with \( Z = c \), it can be shown that
\[ |E(\alpha^{n,jk}_{i} \alpha^{n,lm}_{i} | \mathcal{F}^{n}_{i} |) - \Delta_{n}^{2} (C^{n,jl}_{i} C^{n,km}_{i} + C^{n,jm}_{i} C^{n,kl}_{i})| \leq K \Delta_{n}^{5/2}. \] (C.43)
Next, by successive conditioning and using the bound in (A.10) for $Z = c$ as well as (C.42) and (C.43), we have for $0 \leq u \leq k_n - 1$,
\[
|\mathbb{E}(\alpha_{i+u}^{n,jk}|F_i^n)| \leq K\Delta_n^{3/2} (\sqrt{\Delta_n} + \eta_{i,u}),
\]
(C.44)
\[
|\mathbb{E}(\alpha_{i+u}^{n,jk},\alpha_{i+u}^{n,lm}|F_i^n) - \Delta_n^2 \left(C_{i,j}^{n,jl} C_{i,k}^{n,kl} + C_{i,l}^{n,jm} C_{i,m}^{n,kl}\right)| \leq K\Delta_n^{5/2}.
\]
(C.45)

To show (A.12), we first observe that $\nu_i^{n,jk,\nu_i^{n,lm,\nu_i^{n,gh}}}$ can be decomposed as
\[
\nu_i^{n,jk,\nu_i^{n,lm,\nu_i^{n,gh}}} = \frac{1}{k_n^3 \Delta_n^3} \sum_{u=0}^{k_n-1} \sum_{v=0}^{k_n-1} \sum_{w=0}^{k_n-1} \left[ \xi_{i,u}^{n,jk} \xi_{i,v}^{n,lm} \xi_{i,w}^{n,gh} + \xi_{i,u}^{n,jk} \xi_{i,v}^{n,lm} \xi_{i,w}^{n,gh} + \xi_{i,u}^{n,jk} \xi_{i,v}^{n,lm} \xi_{i,w}^{n,gh} + \xi_{i,u}^{n,jk} \xi_{i,v}^{n,lm} \xi_{i,w}^{n,gh} \right]
\]
with $\xi_{i,u} = \alpha_{i+u}^{n} + (C_{i,u}^{n} - C_{i,u}^{n}) \Delta_n$, which satisfies $\mathbb{E}(||\xi_{i,u}^{n}||^q|F_i^n) \leq K\Delta_n^{q}$ for $q \geq 2$.

Set
\[
\xi_{i}^{n}(1) = \frac{1}{k_n^3 \Delta_n^3} \sum_{u=0}^{k_n-1} \sum_{v=0}^{k_n-1} \sum_{w=0}^{k_n-1} \xi_{i,u}^{n,jk} \xi_{i,v}^{n,lm} \xi_{i,w}^{n,gh}, \quad \xi_{i}^{n}(2) = \frac{1}{k_n^3 \Delta_n^3} \sum_{u=0}^{k_n-1} \sum_{v=0}^{k_n-1} \sum_{w=0}^{k_n-1} \xi_{i,u}^{n,jk} \xi_{i,v}^{n,lm} \xi_{i,w}^{n,gh}
\]
\[
\xi_{i}^{n}(3) = \frac{1}{k_n^3 \Delta_n^3} \sum_{u=0}^{k_n-1} \sum_{v=0}^{k_n-1} \sum_{w=0}^{k_n-1} \xi_{i,u}^{n,jk} \xi_{i,v}^{n,lm} \xi_{i,w}^{n,gh} \quad \text{and} \quad \xi_{i}^{n}(4) = \frac{1}{k_n^3 \Delta_n^3} \sum_{u=0}^{k_n-1} \sum_{v=0}^{k_n-1} \sum_{w=0}^{k_n-1} \xi_{i,u}^{n,jk} \xi_{i,v}^{n,lm} \xi_{i,w}^{n,gh}
\]

The following bounds can be established,
\[
|\mathbb{E}(\xi_{i}^{n}(1)|F_i^n)| \leq K\Delta_n
\]
(C.46)
\[
|\mathbb{E}(\xi_{i}^{n}(2)|F_i^n)| \leq K\Delta_n
\]
(C.47)
\[
|\mathbb{E}(\xi_{i}^{n}(3)|F_i^n)| \leq K\Delta_n
\]
(C.48)
\[
|\mathbb{E}(\xi_{i}^{n}(4)|F_i^n)| \leq K\Delta_n^{9/4}(\Delta_n^{1/4} + \eta_{i,k_n}).
\]
(C.49)

**C.5.2 Proof of (C.46)**
The result readily follows from an application of the Cauchy Schwartz inequality coupled with the bound $\mathbb{E}(||\xi_{i,u}^{n}||^q|F_i^n) \leq K\Delta_n^{q}$ for $q \geq 2$.

**C.5.3 Proof of (C.47)**
Using the law of iterated expectation, we have, for $u < v$,
\[
\mathbb{E}(\nu_i^{n,jk,\nu_i^{n,lm,\nu_i^{n,gh}}}|F_i^n) = \mathbb{E}(\nu_i^{n,jk}|\mathbb{E}(\nu_i^{n,lm,\nu_i^{n,gh}}|F_{i+u}^n)|F_i^n).
\]
(C.50)

By successive conditioning, (C.43), and the Cauchy-Schwarz inequality, we also have
\[
|\mathbb{E}(\nu_i^{n,lm,\nu_i^{n,gh}}|F_{i+u}^n) - \Delta_n^2 (C_{i+u+1}^{n,lg} C_{i+u+1}^{n,ml} + C_{i+u+1}^{n,lm} C_{i+u+1}^{n,mg}) - \Delta_n^2 (C_{i+u+1}^{n,gh} - C_{i+u+1}^{n,lm})| \leq K\Delta_n^{5/2}.
\]
Given that $\mathbb{E}(|\zeta_{n,j}^{i,k}|^{q} F_{n}^{q}) \leq \Delta_{n}^{q}$, the approximation error involved in replacing $\mathbb{E}(\zeta_{n,l}^{i,m} \zeta_{n,g}^{i,h} F_{n}^{i+1})$ by $\Delta_{n}^{2}(C_{i+u}^{n,l}C_{i+u+1}^{n,mh} + C_{i+u+1}^{n,lm} + \Delta_{n}^{2}(C_{i+u}^{n,g} - C_{i}^{n,g})(C_{i+u+1}^{n,lm} - C_{i}^{n,lm})$ in (C.50) is smaller than $\Delta_{n}^{7/2}$.

We can also easily show that

$$|\mathbb{E}(\alpha_{i+u}^{n,g} (C_{n+1}^{i+u} - C_{i}^{n,lm}) | F_{n}^{i})| \leq K \Delta_{n}^{3/2}/(\sqrt{\Delta_{n} + \eta_{n,k}^{n}}). \tag{C.51}$$

Since $(C_{n+1}^{i+u} - C_{i}^{n})$ is $F_{i+u}^{n}$-measurable, we use the successive conditioning, the Cauchy-Schwartz inequality, (C.42), (C.43), and the fifth statement in Lemma A2 applied to $Z = c$ to obtain

$$|\mathbb{E}(\alpha_{i+u}^{n,g} (C_{n+1}^{i+u} - C_{i}^{n,lm}) | F_{n}^{i})| \leq K \Delta_{n}^{3/2}/(\sqrt{\Delta_{n} + \eta_{n,k}^{n}}). \tag{C.52}$$

The following inequalities can be established using (C.42), the successive conditioning together with (A.10) for $Z = c$,

$$|\mathbb{E}(\alpha_{i+u}^{n,g} (C_{n+1}^{i+u} - C_{i}^{n,lm}) | F_{n}^{i})| \leq K \Delta_{n}^{3/2}/(\sqrt{\Delta_{n} + \eta_{n,k}^{n}}). \tag{C.53}$$

The last three inequalities together yield $|\mathbb{E}(\zeta_{n}^{i}(2) | F_{n}^{i})| \leq K \Delta_{n}$. 

**C.5.4 Proof of (C.48)**

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

$$\mathbb{E}(\zeta_{n,j}^{i,k} \zeta_{n,m}^{i,v} F_{n}^{i}) = \mathbb{E}(\zeta_{n,j}^{i,k} \zeta_{n,m}^{i,v} \mathbb{E}(\zeta_{n,g}^{i,h} | F_{n}^{i+1}) F_{n}^{i}). \tag{C.54}$$

By successive conditioning and (C.42), we have

$$|\mathbb{E}(\alpha_{i+u}^{n,g} | F_{n}^{i+1})| \leq K \Delta_{n}^{3/2}/(\sqrt{\Delta_{n} + \eta_{n,v+1,v-w}}). \tag{C.55}$$

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

$$|\mathbb{E}(\zeta_{n,g}^{i} | F_{n}^{i}) - \Delta_{n}(w - v - 1) \bar{b}_{n,v+1}^{n,gh}| \leq K(w - v - 1) \Delta_{n} \eta_{n,v+1,v-w} \leq K \Delta_{n}^{1/2} \eta_{n,v+1,v-w}. \tag{C.56}$$

The last two inequalities together imply

$$|\mathbb{E}(\zeta_{n,g}^{i} | F_{n}^{i}) - \Delta_{n}(w - v - 1) \bar{b}_{n,v+1}^{n,gh}| \leq K \Delta_{n}^{3/2}/(\sqrt{\Delta_{n} + \eta_{n,v+1,v-w}}). \tag{C.57}$$

Since $|\mathbb{E}(\zeta_{n,j}^{i,k} | F_{n}^{i})| \leq \Delta_{n}^{q}$, the error induced by replacing $\mathbb{E}(\zeta_{n,g}^{i} | F_{n}^{i})$ by $(\zeta_{n,j}^{i,k} | F_{n}^{i})$ in (C.53) is smaller than $\Delta_{n}^{7/2}$.

Using Cauchy Schwartz inequality, successive conditioning, (C.52), (A.10) for $Z = c$ and the boundedness of $\bar{b}_{u}$ and $C_{i}$ we obtain

$$|\mathbb{E}(\alpha_{i+u}^{n,g} | F_{n}^{i+1})| \leq K \Delta_{n}^{5/2}, \tag{C.58}$$

$$|\mathbb{E}(\alpha_{i+u}^{n,g} | F_{n}^{i+1})| \leq K \Delta_{n}^{5/2}, \tag{C.59}$$

$$|\mathbb{E}(\alpha_{i+u}^{n,g} | F_{n}^{i+1})| \leq K \Delta_{n}^{1/4} \Delta_{n}^{3/2}/(\sqrt{\Delta_{n} + \eta_{n,k}^{n}}). \tag{C.60}$$
By Lemma A3, we have for $E$

We first observe that $C.5.5$ Proof of (C.49)

The above inequalities together yield $|E(\xi_n^3(F^n_i))| \leq K \Delta_n$.

C.5.5 Proof of (C.49)

We first observe that $\xi_n^1(4)$ can be rewritten as

$$\xi_n^1(4) = \frac{1}{(kn \Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \sum_{u=0}^{v-1} \alpha_{n,jk} \alpha_{n,lm} \alpha_{n,gh} \alpha_{n,lm} \alpha_{n,gh} \alpha_{n,lm} \alpha_{n,gh}$$

where

$$\alpha_{n,jk} \alpha_{n,lm} \alpha_{n,gh} = \alpha_{n,jk} \alpha_{n,lm} \alpha_{n,gh} + \alpha_{n,jk} \Delta_n \alpha_{n,lm} \alpha_{n,gh} - \alpha_{n,jk} \alpha_{n,lm} \alpha_{n,gh} - \alpha_{n,jk} \alpha_{n,lm} \alpha_{n,gh}$$

Based on the above decomposition, we set

$$\xi_n^1(4) = \sum_{j=1}^{8} \chi(j)$$

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

First, set

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

Upon changing the order of the summation, we have

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

Define also

$$\chi'(1) = \frac{1}{(kn \Delta_n)^3} \sum_{w=2}^{k_n-1} \sum_{v=0}^{w-1} \left( \sum_{u=0}^{v-1} \alpha_{n,jk} \alpha_{n,lm} \alpha_{n,gh} \right)$$

Note that $E(\chi(1)|F^n_i) = E(\chi'(1)|F^n_i)$.

By Lemma A3, we have for $q \geq 2$,

$$E\left( \left\| \sum_{u=0}^{v-1} \alpha_{n,jk} \right\|^q \left| F^n_i \right. \right) \leq K_q \Delta_n^{3q/4}$$

59
The Cauchy-Schwartz inequality yields
\[
\mathbb{E}\left( \left| \sum_{u=2}^{k_n-1} \sum_{v=0}^{w-1} \left( \sum_{u=0}^{v-1} \alpha_{i+v}^{n,jk} \right) \alpha_{i+v}^{n,lm} \mathbb{E}(\alpha_{i+w}^{n,gh} | \mathcal{F}_i^n) \right| \mathcal{F}_i^n \right) \leq K k_n^2 \mathbb{E}\left( \left( \sum_{u=0}^{v-1} \alpha_{i+w}^{n,jk} \right)^4 | \mathcal{F}_i^n \right)^{1/4} \\
\times \left[ \mathbb{E}\left( \left( \sum_{i+w}^{n,lm} \mathcal{F}_i^n \right)^{1/4} \right) \times \left[ \mathbb{E}\left( \mathbb{E}(\alpha_{i+w}^{n,gh} | \mathcal{F}_i^n) \right)^2 | \mathcal{F}_i^n \right) \right]^{1/2} \leq K \Delta_n k_n^2 \Delta_n^{3/4} \Delta_n^{3/2} (\sqrt{\Delta_n} + \eta_{n,k_n}^n),
\]
where the last iteration is obtained using (C.54) 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 A1. It follows that
\[
|\mathbb{E}(\chi(1)| \mathcal{F}_i^n)| \leq K \Delta_n^{3/4} (\sqrt{\Delta_n} + \eta_{n,k_n}^n).
\]

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

Given that for \(q \geq 2\), we have
\[
\mathbb{E}\left( \left\| \sum_{u=0}^{v-1} \Delta_n(C_{i+u}^{n,jk} - C_{i}^{n,jk}) \right\|_q^{q} | \mathcal{F}_i^n \right) \leq K \Delta_n^{3/4} \quad \text{and} \quad \mathbb{E}(||C_{i+u}^{n,jk} - C_{i}^{n,jk}||_q | \mathcal{F}_i^n) \leq K \Delta_n^{q/4}.
\]

Similar steps to \(\chi(1)\) lead to
\[
|\mathbb{E}(\chi(2)| \mathcal{F}_i^n)| \leq K \Delta_n^{3/4} (\sqrt{\Delta_n} + \eta_{n,k_n}^n) \quad \text{and} \quad |\mathbb{E}(\chi(j)| \mathcal{F}_i^n)| \leq K \Delta_n (\sqrt{\Delta_n} + \eta_{n,k_n}^n) \quad \text{for} \quad j = 3, 4.
\]

Define
\[
\chi(5) = \frac{1}{(k_n \Delta_n)^3} \sum_{u=0}^{k_n-1} \sum_{v=0}^{w-1} \left( \sum_{u=0}^{v-1} \alpha_{i+v}^{n,jk} \right) \alpha_{i+v}^{n,lm} \Delta_n(C_{i+w}^{n,gh} - C_{i}^{n,gh}) \\
\chi'(5) = \frac{1}{(k_n \Delta_n)^3} \sum_{u=0}^{k_n-1} \sum_{v=0}^{w-1} \left( \sum_{u=0}^{v-1} \alpha_{i+v}^{n,jk} \right) \alpha_{i+v}^{n,lm} \Delta_n((C_{i+w}^{n,gh} - C_{i}^{n,gh}) | \mathcal{F}_i^n) \\
\chi(6) = \frac{1}{(k_n \Delta_n)^3} \sum_{u=0}^{k_n-1} \sum_{v=0}^{w-1} \left( \sum_{u=0}^{v-1} \Delta_n(C_{i+u}^{n,jk} - C_{i}^{n,jk}) \right) \alpha_{i+v}^{n,lm} \Delta_n(C_{i+w}^{n,gh} - C_{i}^{n,gh}) \\
\chi(7) = \frac{1}{(k_n \Delta_n)^3} \sum_{u=0}^{k_n-1} \sum_{v=0}^{w-1} \left( \sum_{u=0}^{v-1} \alpha_{i+v}^{n,jk} \right) \Delta_n(C_{i+w}^{n,lm} - C_{i}^{n,lm}) \Delta_n(C_{i+w}^{n,gh} - C_{i}^{n,gh}),
\]
where we have \(\mathbb{E}(\chi(5)| \mathcal{F}_i^n) = \mathbb{E}(\chi'(5)| \mathcal{F}_i^n)\). Recalling (C.55), we further decompose \(\chi'(5)\) as,
\[
\chi'(5) = \sum_{j=1}^{5} \chi(5)[j],
\]
60
with

\[ \chi(5)[1] = \frac{1}{(k_\Delta \Delta)^3} \sum_{w=2}^{k_\Delta-1} \sum_{v=0}^{w-1} \left( \sum_{u=0}^{v-1} \alpha_{i+u}^{n} \right) \Delta_n (C_i^{n,lm} - C_i^{n,gh}) \left( \frac{E(C_{i+w}^{n,gh} - C_i^{n,gh})}{F_{i+w}^n} \right) \]

\[ \chi(5)[2] = \frac{1}{(k_\Delta \Delta)^3} \sum_{w=2}^{k_\Delta-1} \sum_{v=0}^{w-1} \Delta_n (C_i^{n,gh} - C_i^{n,gh}) \left( \sum_{u=0}^{v-1} \alpha_{i+u}^{n} \right) \Delta_n (C_i^{n,gh} - C_i^{n,gh}) \alpha_{i+v}^{n,lm} \]

\[ \chi(5)[3] = \frac{1}{(k_\Delta \Delta)^3} \sum_{w=2}^{k_\Delta-1} \sum_{v=0}^{w-1} \left( \sum_{u=0}^{v-1} \alpha_{i+u}^{n} \right) \Delta_n (C_i^{n,gh} - C_i^{n,gh}) \alpha_{i+v}^{n,lm} \]

\[ \chi(5)[4] = \frac{1}{(k_\Delta \Delta)^3} \sum_{w=2}^{k_\Delta-1} \sum_{v=0}^{w-1} \left( \sum_{u=0}^{v-1} \alpha_{i+u}^{n} \right) \Delta_n (C_i^{n,gh} - C_i^{n,gh}) \alpha_{i+v}^{n,lm} \]

\[ \chi(5)[5] = \frac{1}{(k_\Delta \Delta)^3} \sum_{w=2}^{k_\Delta-1} \sum_{v=0}^{w-1} \Delta_n^2 (w-v-1) \Delta_n (C_i^{n,gh} - C_i^{n,gh}) \alpha_{i+v}^{n,lm} \]

Using (C.55), (C.54), (C.51) and following the same strategy proof as for \( \chi(1) \), it can be shown that

\[ |E(\chi(5)[j]|F_i^n |) \leq K \Delta_n^{3/4} (\sqrt{\Delta_n} + \eta_{i,k_n}) \text{, for } j = 1, \ldots, 5, \]

which in turn implies

\[ |E(\chi(5)|F_i^n |) \leq K \Delta_n^{3/4} (\sqrt{\Delta_n} + \eta_{i,k_n}) \text{, for } j = 1, \ldots, 5. \]

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

\[ |E(\chi(6)|F_i^n |) \leq K \Delta_n^{3/4} (\sqrt{\Delta_n} + \eta_{i,k_n}). \]

Next, we set

\[ \chi(7) = \frac{1}{(k_\Delta \Delta)^3} \sum_{w=2}^{k_\Delta-1} \left( \sum_{v=0}^{w-1} \Delta_n (C_{i+w}^{n,lm} - C_{i+u}^{n,lm}) \Delta_n (C_i^{n,gh} - C_i^{n,gh}) \right). \]

Define

\[ \chi(7)[1] = \frac{1}{(k_\Delta \Delta)^3} \sum_{w=2}^{k_\Delta-1} \left( \sum_{v=0}^{w-1} \Delta_n (C_{i+w}^{n,lm} - C_{i+u}^{n,lm}) \Delta_n (C_i^{n,gh} - C_i^{n,gh}) \right) \]

\[ \chi(7)[2] = \frac{1}{(k_\Delta \Delta)^3} \sum_{w=2}^{k_\Delta-1} \left( \sum_{v=0}^{w-1} \Delta_n (C_{i+w}^{n,lm} - C_{i+u}^{n,lm}) \Delta_n (C_i^{n,gh} - C_i^{n,gh}) \right) \]

\[ \chi(7)[3] = \frac{1}{(k_\Delta \Delta)^3} \sum_{w=2}^{k_\Delta-1} \left( \sum_{v=0}^{w-1} \Delta_n (C_{i+w}^{n,lm} - C_{i+u}^{n,lm}) \Delta_n (C_i^{n,gh} - C_i^{n,gh}) \right) \]

\[ \chi(7)[4] = \frac{1}{(k_\Delta \Delta)^3} \sum_{w=2}^{k_\Delta-1} \left( \sum_{v=0}^{w-1} \Delta_n (C_{i+w}^{n,lm} - C_{i+u}^{n,lm}) \Delta_n (C_i^{n,gh} - C_i^{n,gh}) \right). \]
It is easy to see that
\[
\chi(7) = \sum_{j=1}^{4} \chi(7)[j].
\]

Similarly to calculations used for \(\chi(1)\), it can be shown that
\[
|E(\chi(7)[j]|\mathcal{F}_i^n)\) \leq K \Delta_i^{1/4} (\Delta_i^{1/4} + \eta_{i,k_n}), \text{ for } j = 1, \ldots, 3.
\]

To handle the remaining term \(\chi(7)[4]\), we decompose it \(\chi(7)[4] = \sum_{j=1}^{9} \chi(7)[4][j]\), where
\[
\begin{align*}
\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_{n,jk} n_{i,u} (C_{i,u+1}^{n,m} - C_{i,u+1}^{n,n}) (C_{i,u+1}^{n,g} - C_{i,u+1}^{n,g}),
\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+1}^{n,g} - C_{i,u+1}^{n,g}) (C_{i,u+1}^{n,m} - C_{i,u+1}^{n,m}),
\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+1}^{n,g} - C_{i,u+1}^{n,g}) (C_{i,u+1}^{n,m} - C_{i,u+1}^{n,m}),
\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+1}^{n,m} - C_{i,u+1}^{n,m}) (C_{i,u+1}^{n,g} - C_{i,u+1}^{n,g}),
\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+1}^{n,m} - C_{i,u+1}^{n,m}) (C_{i,u+1}^{n,g} - C_{i,u+1}^{n,g}) \alpha_{i,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+1}^{n,m} - C_{i,u+1}^{n,m}) (C_{i,u+1}^{n,g} - C_{i,u+1}^{n,g}) \alpha_{i,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+1}^{n,m} - C_{i,u+1}^{n,m}) (C_{i,u+1}^{n,g} - C_{i,u+1}^{n,g}) \alpha_{i,jk},
\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} (C_{i,u+1}^{n,m} - C_{i,u+1}^{n,m}) (C_{i,u+1}^{n,g} - C_{i,u+1}^{n,g}) \alpha_{i,jk},
\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+1}^{n,m} - C_{i,u+1}^{n,m}) (C_{i,u+1}^{n,g} - C_{i,u+1}^{n,g}) \alpha_{i,jk},
\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} (C_{i,u+1}^{n,m} - C_{i,u+1}^{n,m}) (C_{i,u+1}^{n,g} - C_{i,u+1}^{n,g}) \alpha_{i,jk},
\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} (C_{i,u+1}^{n,m} - C_{i,u+1}^{n,m}) (C_{i,u+1}^{n,g} - C_{i,u+1}^{n,g}).
\end{align*}
\]

Using arguments similar to those involved for the treatment of \(\chi(1)\), it can be shown that
\[
|E(\chi(7)[4][j]|\mathcal{F}_i^n)\) \leq K \Delta_i^{1/4} (\Delta_i^{1/4} + \eta_{i,k_n}), \text{ for } j = 1, \ldots, 8,
\]
which yields
\[
|E(\chi(7)|\mathcal{F}_i^n)| \leq K \Delta_i^{1/4} (\Delta_i^{1/4} + \eta_{i,k_n}).
\]
Next, define
\[
\chi(8) = \frac{1}{k^n} \sum_{u=2}^{k-1} \sum_{v=0}^{n-1} \sum_{u=0}^{k-1} (C_{i+u}^{n,jk} - C_{i}^{n,jk})(C_{i+v}^{n,lm} - C_{i}^{n,lm})(C_{i+u}^{n,gh} - C_{i}^{n,gh}).
\]

This term can be further decomposed into six components. Successive conditioning and existing bounds give
\[
\begin{align*}
&|E\left((C_{i+u}^{n,jk} - C_{i}^{n,jk})(C_{i+v}^{n,lm} - C_{i}^{n,lm})(C_{i+u}^{n,gh} - C_{i}^{n,gh})|F_i^{n}\right)| \leq K\Delta_n \\
&|E\left((C_{i+u}^{n,jk} - C_{i}^{n,jk})(C_{i+v}^{n,lm} - C_{i}^{n,lm})(C_{i+u}^{n,gh} - C_{i}^{n,gh})|F_i^{n}\right)| \leq K\Delta_n^{3/4}(\Delta_n^{1/4} + \eta, k_n) \\
&|E\left((C_{i+u}^{n,jk} - C_{i}^{n,jk})(C_{i+v}^{n,lm} - C_{i}^{n,lm})(C_{i+u}^{n,gh} - C_{i}^{n,gh})|F_i^{n}\right)| \leq K\Delta_n \\
&|E\left((C_{i+u}^{n,jk} - C_{i}^{n,jk})(C_{i+v}^{n,lm} - C_{i}^{n,lm})(C_{i+u}^{n,gh} - C_{i}^{n,gh})|F_i^{n}\right)| \leq K\Delta_n \\
&|E\left((C_{i+u}^{n,jk} - C_{i}^{n,jk})(C_{i+v}^{n,lm} - C_{i}^{n,lm})(C_{i+u}^{n,gh} - C_{i}^{n,gh})|F_i^{n}\right)| \leq K\Delta_n
\end{align*}
\]

These bounds can be used to deduce
\[
|E(\chi(8)|F_i^{n})| \leq K\Delta_n.
\]

This completes the proof.

C.5.6 Proof of (A.13) and (A.14) in Lemma A5

Observe that
\[
\nu_i^{n,jk}(C_{i+k_n}^{n,lm} - C_{i}^{n,lm})(C_{i+k_n}^{n,gh} - C_{i}^{n,gh}) = \frac{1}{k^n\Delta_n^2} \sum_{u=0}^{k_n-1} \zeta_{i,u}^{n,jk}(C_{i+u}^{n,lm} - C_{i}^{n,lm})(C_{i+u}^{n,gh} - C_{i}^{n,gh}),
\]
\[
\nu_i^{n,jk}(C_{i+k_n}^{n,lm} - C_{i}^{n,lm})(C_{i+k_n}^{n,gh} - C_{i}^{n,gh}) = \frac{1}{k^n\Delta_n^2} \sum_{u=0}^{k_n-1} \zeta_{i,u}^{n,jk}(C_{i+u}^{n,gh} - C_{i}^{n,gh}),
\]
\[
+ \frac{1}{k^n\Delta_n^2} \sum_{u=0}^{k_n-2} \zeta_{i,u}^{n,jk}(C_{i-k_n}^{n,lm} - C_{i}^{n,lm}) + \frac{1}{k^n\Delta_n^2} \sum_{u=0}^{k_n-2} \zeta_{i,u}^{n,jk}(C_{i-k_n}^{n,gh} - C_{i}^{n,gh}).
\]

Hence, (A.13) and (A.14) can be proved using the same strategy as for (A.12).

C.5.7 Proof of (A.15) and (A.16) in Lemma A5

Note that we have
\[
\lambda_i^{n,jk}(C_{i+k_n}^{n,lm} - C_{i}^{n,lm})(C_{i+k_n}^{n,gh} - C_{i}^{n,gh}) = \nu_i^{n,gh}(C_{i+k_n}^{n,jk} + C_{i}^{n,jk}) + \nu_i^{n,gh}(C_{i+k_n}^{n,jk} - C_{i}^{n,jk}) + \nu_i^{n,gh}(C_{i+k_n}^{n,jk} - C_{i}^{n,jk}),
\]
and
\[
\lambda_i^{n,gh}(C_{i+k_n}^{n,lm} - C_{i}^{n,lm})(C_{i+k_n}^{n,jk} - C_{i}^{n,jk}) + \nu_i^{n,gh}(C_{i+k_n}^{n,jk} - C_{i}^{n,jk})(C_{i+k_n}^{n,lm} - C_{i}^{n,lm})
\]

63
Hence, for the next step, by combining the successive conditioning together with existing bounds, we have
\[
\phi_n \leq \frac{1}{k_n} (C_{i+n,k} C_{i,k} + C_{i+k,n} C_{i,n}) - \frac{k_n \Delta_n}{3} C_{i+k,n}.
\]
From (A.4), notice that \( \nu_i \) is \( F_{i+k,n} \)-measurable and satisfies \( \|E(\nu_i | F^n)\| \leq K \Delta_n^{1/2} \). The law of iterated expectations and existing bounds imply
\[
|E(\nu_i, i \rightarrow i+k, i \rightarrow i+k) | \leq K \Delta_n^{3/4},
\]
\[
|E(\nu_i, i \rightarrow i+k) | \leq K \Delta_n^{1/2},
\]
\[
|E((C_{i+k,n} - C_{i,n})/C_{i+k,n}) | \leq K \Delta_n^{1/2},
\]
\[
|E((C_{i+k,n} - C_{i,n})/C_{i+k,n}) | \leq K \Delta_n^{1/2},
\]

It can also be readily verified that
\[
|E(\nu_i, i \rightarrow i+k, i \rightarrow i+k) | \leq K \Delta_n^{3/4} + \eta_{i+k,n,k}.n.
\]
Hence, for \( \phi_n \in [\nu_i, i \rightarrow i+k, i \rightarrow i+k] \), which satisfies \( E(\phi_n | F^n) \leq K \Delta_n^{1/4} \) and \( E(\phi_n | F^n) \leq K \Delta_n^{1/2} \). One can show that
\[
|E(\nu_i, i \rightarrow i+k, i \rightarrow i+k) | \leq K \Delta_n^{3/4} (\Delta_n^{1/4} + \eta_{i+k,n,k}).n.
\]
Next, by combining the successive conditioning together with existing bounds, we have
\[
|E(\nu_i, i \rightarrow i+k, i \rightarrow i+k) | \leq K \Delta_n^{1/4} (\Delta_n^{1/4} + \eta_{i+k,n,k}),
\]
which together imply
\[
|E(\phi_n, i \rightarrow i+k, i \rightarrow i+k) | \leq K \Delta_n^{3/4} (\Delta_n^{1/4} + \eta_{i+k,n,k}).n.
\]
It is easy to see that (A.12), (C.56) and (C.57) and the inequality \( \eta_{i,k,n} \leq \eta_{i+k,n,k} \) together yield (A.15) and (A.16).
C.6 Proof of Lemma A6

(A.17) can be proved easily using the bounds of ρ(u, v)\(^{n,gh}\) in (C.41). To show (A.18), (A.19) and (A.20), we set

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

Then it holds that

\[ \frac{1}{\Delta_n^{1/4}} \left( \overline{\Pi}(H, gh, u; G, ab, v) - \overline{\Pi}(H, gh, u; G, ab, v) \right) \Rightarrow 0. \]

The above result is proved following similar steps as for (C.35) in case \( w = 1 \) by replacing \( \Theta(u, v)_0^{(C),i,n} \) by \( \lambda(u, v)_0^{(C),i,n} (\partial_{gh} H \partial_{ab} G)(C_{i-1}) - (\partial_{gh} H \partial_{ab} G)(C_{i-2k_n}) \), which has the same bounds as the former. Next, decompose \( \overline{\Pi} \) as follows,

\[ \overline{\Pi}(H, gh, u; G, ab, v) = \lambda(u, v)_0^{[T/\Delta_n]} \sum_{i=2k_n} (\partial_{gh} H \partial_{ab} G)(C_{i-1}) V_i^{n-1} + \sum_{i=2k_n} (\partial_{gh} H \partial_{ab} G)(C_{i-1}) \left( \mathbb{E}(\zeta(u)_i^{n,gh} \zeta(v)_i^{n,ab}|F_{i-1}) - V_i^{n-1} \right) + \sum_{i=2k_n} (\partial_{gh} H \partial_{ab} G)(C_{i-1}) \left( \zeta(u)_i^{n,gh} \zeta(v)_i^{n,ab} - \mathbb{E}(\zeta(u)_i^{n,gh} \zeta(v)_i^{n,ab}|F_{i-1}) \right). \]

We follow the proof of (C.36) for \( w = 1 \), and we replace \( \Theta(u, v)_0^{(C),i,n} \) by \( \lambda(u, v)_0^{(C),i,n} (\partial_{gh} H \partial_{ab} G)(C_{i-1}) \), which satisfies only the condition \( |\lambda(u, v)_0^{(C),i,n} (\partial_{gh} H \partial_{ab} G)(C_{i-1})| \leq \lambda_n^{n,v} \). This calculation shows that the last two terms in the above decomposition vanish at a rate faster than \( \Delta_n^{1/4} \). Therefore,

\[ \frac{1}{\Delta_n^{1/4}} \left( \overline{\Pi}(H, gh, u; G, ab, v) - \lambda(u, v)_0^{[T/\Delta_n]} \sum_{i=2k_n} (\partial_{gh} H \partial_{ab} G)(C_{i-1}) V_i^{n-1} \right) \Rightarrow 0. \]

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

\[ \frac{1}{\Delta_n^{1/4}} \overline{\Pi}(H, gh, u; G, ab, v) \Rightarrow 0. \]

The results follow from the following observation,

\[ \frac{1}{\Delta_n^{1/4}} \left( \sum_{g,h,a,b=1}^d \sum_{i=2k_n} (\partial_{gh} H \partial_{ab} G)(C_{i-1}) V_i^{n-1}(u, v) \right) - \frac{3}{\theta^2} \int_0^T (\partial_{gh} H \partial_{ab} G)(C_t)(C_t^{agh} C_t^{a,b} + C_t^{agh} C_t^{ah,b}) dt \Rightarrow 0, \quad \text{for } (u, v) = (2, 2), \]

\[ \frac{1}{\Delta_n^{1/4}} \left( \sum_{g,h,a,b=1}^d \sum_{i=2k_n} (\partial_{gh} H \partial_{ab} G)(C_{i-1}) V_i^{n-1}(u, v) - [H(C), G(C)]_T \right) \Rightarrow 0, \quad \text{for } (u, v) = (1, 1). \]
Figure D.1: Monthly $R^2$ of two return factor models ($\hat{R}^2_{Yj}$): the CAPM (the blue dotted line) and the Fama-French three factor model (the red solid line). Stocks are represented by tickers (see Table 1 for full stock names).
Figure D.2: Monthly $R^2$ of two return factor models ($\hat{R}^2_{Y,j}$): the CAPM (the blue dotted line) and the Fama-French three factor model (the red solid line). Stocks are represented by tickers (see Table 1 for full stock names).