# Cross-sectional Dependence in Idiosyncratic Volatility

Ilze Kalnina* Kokouvi Tewou†

May 7, 2025

**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. Naive 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, naive 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 S&P 100 index constituents, 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. For each model, we map out the network of dependencies in residual (non-systematic) idiosyncratic volatilities across all stocks.

**Keywords**: factor model, systematic risk, networks of risk, residual idiosyncratic volatility, (co-)volatility of volatility, high frequency data.

JEL Codes: C58, C22, C14, G11.

*Corresponding author. Department of Economics, NC State University. E-mail address: `ikalnin@ncsu.edu`.

†National Bank of Canada. E-mail address: `kokouvi.tewou@bnc.ca`.

## 1 Introduction

In a panel of assets, returns are generally cross-sectionally dependent. This dependence is usually modeled 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)).[1] Also, Idiosyncratic Volatility measures the exposure to the idiosyncratic risk in imperfectly diversified portfolios. The cross-sectional dependence in IdioVols is also important for option pricing, see Gourier (2016). The attention to IdioVols in empirical finance literature is exemplified by two IdioVol puzzles, see Campbell, Lettau, Malkiel, and Xu (2001) and Ang, Hodrick, Xing, and Zhang (2006). A recent observation is that the IdioVols seem to be strongly correlated in the cross-section of stocks.[2] We propose methods to formally study this empirical phenomenon with high-frequency data, while fully accounting for the measurement errors in IdioVols.

This paper provides an econometric framework for studying the cross-sectional dependence in the Idiosyncratic Volatilities using high frequency data. The analysis is based on a new general asymptotic theory that we develop for estimators of quadratic covariations between nonlinear functions of spot volatility matrices. We show that naive estimators, such as covariances and correlations, are biased. The bias arises due to the use of error-laden estimates of the spot volatility matrices. We provide the bias-corrected estimators. We derive the asymptotic distribution of these estimators, and propose consistent estimators of the asymptotic variances. We apply this new asymptotic theory to construct tests of dependence between IdioVols and map out the network of

[1] An asset is said to be mispriced with respect to a given model if the expected value of the return on the asset is not consistent with the model.

[2] See, e.g., Connor, Korajczyk, and Linton (2006), Duarte, Kamara, Siegel, and Sun (2014), Herskovic, Kelly, Lustig, and Nieuwerburgh (2016), and Christoffersen, Fournier, and Jacobs (2018).

dependencies in IdioVols in a panel of assets.

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, 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.

Reduced-form analysis of total and idiosyncratic volatilities can be useful to inform the formulation of structural asset pricing models. For example, Herskovic, Kelly, Lustig, and Nieuwerburgh (2016) document strong dependence in firm IdioVols, and propose an incomplete markets asset pricing model, where IdioVol behavior is explained by the idiosyncratic risk faced by households. When documenting the cross-sectional dependence in IdioVol, Herskovic, Kelly, Lustig, and Nieuwerburgh (2016) estimate several volatility factor models, for example, they regress IdioVols on average firm volatilities, where the IdioVols are defined with respect to the market return factor or the Fama-French factors. Our framework can be used to estimate high-frequency regressions with these variables, on a fixed time interval, while fully capturing the effect of the measurement error from the preliminary estimation of both the dependent variable and the factor.

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 S&P 100 index constituents. We study the IdioVols with respect to two models for asset returns: the CAPM and the three-factor Fama-French model.[3] In both cases, the average pairwise correlation between the IdioVols is high (0.35). 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 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 (0.21) than between the total IdioVols. With the additional industry ETF volatilities as IdioVol factors, average correlation between the residual IdioVols decreases further (to 0.17). However, neither of the two sets of the IdioVol factors can fully explain the cross-sectional dependence in the IdioVols. For each model, 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 (2020). 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 (2019, 2020). Importantly, the models in the above papers are silent on the cross-sectional dependence structure in the IdioVols.

While this paper focuses on the study of cross-sectional dependence of IdioVols, our new asymptotic theory can be used in various other applications. For example, we can estimate dependence measures, in the form of co-volatilities or the corresponding correlations, between the time-varying

[3] The high frequency Fama-French factors are provided by Aït-Sahalia, Kalnina, and Xiu (2020).

asset betas.[4] While it is well-known that asset betas vary over time in practice, there is no consensus as to what common factors drive this variation, so accurate dependence measures of asset beta co-movement can be helpful. Another example is the estimation of dependence measures between total volatilities or systematic volatilities of asset returns. In addition, we can estimate high-frequency regressions of one element of a spot volatility matrix on other elements, such as regression of the asset volatility on market volatility. Finally, we can estimate high-frequency regressions of total asset volatility on average asset volatility, which mirrors one more of the specifications considered in Herskovic, Kelly, Lustig, and Nieuwerburgh (2016), in addition to the specifications described earlier.

Our inference theory is related to several estimators in the existing literature. The closest are the volatility of volatility estimator of Vetter (2015) and one of the asymptotic bias estimators of Jacod and Rosenbaum (2015). Vetter (2015) proposes an estimator of volatility of volatility of the returns of one asset, and derives the relevant theory for inference.[5] We extend the analysis to the multivariate case with nonlinear transformations, return jumps, and volatility jumps. While Jacod and Rosenbaum (2015) focus on a different problem, one of the asymptotic bias terms in their paper coincides with our quantity of interest in a special case, see Section 3.1 for details. The setting in Jacod and Rosenbaum (2015) is multivariate and robust to return and volatility jumps, but they only establish consistency of the relevant estimator, and do not provide any asymptotic distribution theory. In contrast, we derive the asymptotic distribution, as well as the consistency of the estimator of the asymptotic variance. See also Li, Liu, and Zhang (2022) who extend the results in Vetter (2015) to allow for price jumps and market microstructure noise. They do not consider the multivariate case, nonlinear transformations, or volatility jumps. Finally, Chong and Todorov (2024) propose nonparametric estimators of the volatility of volatility and leverage effect using high-frequency data on short-dated options.

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

[4]Here, asset betas are the loadings of asset returns on return factors; these are distinct from the asset volatility betas that we describe in the next section.

[5]This estimator is also studied in Aït-Sahalia and Jacod (2014) (Section 8.3) under similar assumptions to Vetter (2015). Aït-Sahalia and Jacod (2014) cite 2011 working paper version of Vetter (2015).

estimation is possible, and the estimators are consistent without a bias correction (see Section 3.1 for details). In the literature on the estimation of the leverage effect, preliminary estimation of volatility also creates a bias, which also needs to be corrected to achieve consistency, see Aït-Sahalia, Fan, and Li (2013), Aït-Sahalia, Fan, Laeven, Wang, and Yang (2017), Kalnina and Xiu (2017) and Wang and Mykland (2014).

One of the reasons why we can account for the measurement error from preliminary estimation of volatilities is the fact that our framework only uses one (in-fill) asymptotic approximation. It is interesting to contrast this approach with the analysis of two-step estimators using joint in-fill and long-span asymptotics, see, e.g., Corradi and Distaso (2006), Todorov (2009), Bandi and Renò (2012), Kanaya and Kristensen (2016), and Li and Patton (2018). In these double asymptotic settings, the inference methods for the second step typically do not depend on the first-step measurement error. This provides a good approximation as long as the number of high-frequency observations in every low-frequency period is large enough. A notable early exception is Bollerslev and Zhou (2002) who use a simple parametric model for the first-step measurement error.

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.[6]

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

[6] In 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.

the cross-sectional dependence in IdioVols using our framework. Section 6 contains Monte Carlo simulations. The Online Supplementary 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.

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

$$Y_t = Y_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t^Y, \tag{1}$$

where $W$ is a $d^W$-dimensional Brownian motion ($d^W \geq d$), $C_t = \sigma_t \sigma_t^\top$ is the spot covariance process, and $J_t^Y$ denotes a finite variation jump process. The spot covariance matrix process $C_t$ of $Y_t$ is a continuous Itô semimartingale,[7]

$$C_t = C_0 + \int_0^t \widetilde{b}_s ds + \int_0^t \widetilde{\sigma}_s dW_s + J_t^\sigma. \tag{2}$$

We refer to the $(C_t)_{a,b}$ element of the matrix $C_t$ as $C_{ab,t}$. 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.

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

**Definition (Factor Model for Returns, R-FM).** *For all* $0 \le t \le T$ *and* $j = 1, \ldots, d_S$,[8]

$$\begin{aligned} dS_{j,t} &= \beta_{j,t}^{\top} dF_t^c + \tilde{\beta}_{j,t}^{\top} dF_t^d + dZ_{j,t} \quad \textit{with} \\ [Z_j, F]_t &= 0. \end{aligned} \tag{3}$$

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

We do not need the return factors $F_t$ to be the same across assets to identify the model, but without loss of generality, we keep this structure as it is standard in empirical finance. These return factors are assumed to be observable, which is also standard. For example, in the empirical application, we use two sets of return factors: the market portfolio and the three Fama-French factors, which are constructed in Aït-Sahalia, Kalnina, and Xiu (2020).

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. However, this literature does not consider the cross-sectional dependence in the IdioVols.

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

$$\beta_{jt} = (C_{F,t})^{-1} C_{FSj,t}, \tag{4}$$

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

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

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

for any $t_0 < t_1 < \ldots < t_M = T$ with $\sup_s |t_{s+1} - t_s| \to 0$ as $M \to \infty$.

Intuitively, quadratic covariation can be thought of as the integrated covariance between the increments $dX_t$ and $dY_t$.

is a vector consisting of the last $d_F$ elements of the $j^{th}$ column of $C_t$. The IdioVol of stock $j$ is then also a function of the total spot covariance matrix $C_t$,

$$\underbrace{C_{Zj,t}}_{\text{IdioVol of stock j}} = \underbrace{C_{Yj,t}}_{\text{total volatility of stock j}} - (C_{FSj,t})^\top (C_{F,t})^{-1} C_{FSj,t}. \tag{5}$$

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

We now introduce the Idiosyncratic Volatility Factor model (IdioVol-FM). In IdioVol-FM, the cross-sectional dependence in the IdioVol shocks can be potentially explained by certain IdioVol factors we denote as $\Pi_t$. A simple example of IdioVol factor is the market volatility. Our model allows IdioVol factors to be any given smooth functions of the matrix $C_t$; we discuss examples below.

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

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

*where $\Pi_t = (\Pi_{1t}, \ldots, \Pi_{d_\Pi t})$ is a $\mathbb{R}^{d_\Pi}$-valued vector of IdioVol factors. IdioVol factors satisfy*

$$\Pi_{kt} = \Pi_k(C_t) \tag{7}$$

*with the function $\Pi_k(\cdot)$ being three times continuously differentiable for $k = 1, \ldots, d_\Pi$.*

$\Pi(\cdot)$ is a smooth function of $C_t$. For example, often $\Pi(C_t)$ is $C_{F,t}$, i.e., $\Pi(\cdot)$ selects the components of $C_t$ that correspond to the volatilities of the observable factors $F_t$. More generally, $\Pi_t$ may also include the volatilities and covolatilities of other assets beyond $F_t$. Even more generally, our theory permits a rather wide class of IdioVol factors, since it includes general non-linear transforms of the spot covariance matrix process $C_t$. For example, IdioVol factors can be linear combinations of the total volatilities of assets, see, e.g., the average variance factor of Chen and Petkova (2012). Another example is the common IdioVol factor, or "CIV", which is studied in Herskovic, Kelly, Lustig, and Nieuwerburgh (2016). CIV is defined as the cross-sectional average of the firm IdioVols from CAPM. The IdioVol factors can also be the volatilities of any other observable processes.

We call the residual term $C^{resid}_{Zj,t}$ in the IdioVol-FM the residual IdioVol of asset $j$. Our assumptions imply that the components of the IdioVol-FM, $C_{Zj,t}, \Pi_t$ and $C^{resid}_{Zj,t}$, are 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.

Having specified our econometric framework, we now provide the definitions of some natural measures of dependence of (the continuous parts of) the (total) IdioVols and the residual IdioVols. We consider the estimation of these measures in Section 3.

Before studying the decomposition of the IdioVol-FM model, one may be interested in quantifying the dependence between the (total) IdioVols of two stocks $j$ and $s$. Quadratic covariation $[C_{Zj}, C_{Zs}]^c_T$ is one natural measure of dependence between the (continuous parts of) the IdioVols $C_{Zj}$ and $C_{Zs}$. Another natural and scale invariant measure is the quadratic-covariation-based correlation between the two IdioVol processes over a given time period $[0, T]$,

$$Corr\left(C_{Zj}, C_{Zs}\right) = \frac{[C_{Zj}, C_{Zs}]^c_T}{\sqrt{[C_{Zj}, C_{Zj}]^c_T}\sqrt{[C_{Zs}, C_{Zs}]^c_T}}. \tag{8}$$

Correlation-based measure is more convenient for reporting the strength of dependence, while the quadratic covariation $[C_{Zj}, C_{Zs}]^c_T$ without normalization is more convenient for testing for the presence of cross-sectional dependence in IdioVols. We consider such tests in Section 4.4.

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

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

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

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

$$R_{Zj}^{2,\textit{IdioVol-FM}} = \frac{\gamma_{Zj}^\top [\Pi, \Pi]_T^c \gamma_{Zj}}{[C_{Zj}, C_{Zj}]_T^c}. \tag{10}$$

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,

$$Corr\left(C_{Zj}, C_{Zs}\right) - Corr\left(C_{Zj}^{resid}, C_{Zs}^{resid}\right) \tag{11}$$

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,

$$Q_{Zj,Zs}^{\textit{IdioVol-FM}} = \frac{\gamma_{Zj}^\top [\Pi, \Pi]_T^c \gamma_{Zs}}{[C_{Zj}, C_{Zs}]_T^c}. \tag{12}$$

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), $H_0 : \; C_{Zj,t} = a_{Zj} + \gamma_{Zj}^\top \Pi_t, \; 0 \leq t \leq T$. 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 $R_{Zj}^{2,\textit{IdioVol-FM}} = 1$. Our framework allows for testing stochastic relationships, i.e., null hypotheses $H_0 : \gamma_{Zj}^\top = 0$ in the presence of an error term.

# 3 Estimation

As we show below, the quantities of interest in Section 2 can be expressed in terms of the continuous quadratic covariation between two functions of the spot covariance matrix $C_t$,

$$[H(C), G(C)]^c_T \, . \tag{13}$$

Section 3.1 proposes estimators of this general functional, and Section 3.2 explains how to use these formulas to obtain estimators of the quantities of interest in Section 2.

## 3.1 Estimation of a General Functional

This section proposes estimators of the continuous quadratic covariation between two functions of the spot covariance matrix $[H(C), G(C)]^c_T$, where $H$ and $G$ are given real-valued smooth functions. Recall that $C_t$ is the spot covariance matrix of the observable variables, see equations (1)-(2).

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,

$$\widehat{C}_{i\Delta_n} = \frac{1}{k_n \Delta_n} \sum_{m=0}^{k_n - 1} \left(\Delta^n_{i+m} Y\right) \left(\Delta^n_{i+m} Y\right)^\top 1_{\{\|\Delta^n_{i+m} Y\| \leq u_n\}}, \tag{14}$$

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.[9] We refer to the $\left(\widehat{C}_{i\Delta_n}\right)_{a,b}$ element of the matrix $\widehat{C}_{i\Delta_n}$ as $\widehat{C}_{ab,i\Delta_n}$.

If $C_{i\Delta_n}$ was observed and in the absence of volatility jumps, we could estimate $[H(C), G(C)]_T$ by the realized covariance between $G(C_{i\Delta_n})$ and $H(C_{i\Delta_n})$, which is the sample analog of the definition of $[H(C), G(C)]_T$. However, we do not observe $C_{i\Delta_n}$. If we replace it with $\widehat{C}_{i\Delta_n}$ in (14), we obtain the plug-in estimator

$$\widehat{[H(C), G(C)]}^{Naive}_T = \frac{1}{k_n} \sum_{i=1}^{[T/\Delta_n] - 2k_n + 1} \left( H(\widehat{C}_{(i+k_n)\Delta_n}) - H(\widehat{C}_{i\Delta_n}) \right) \left( G(\widehat{C}_{(i+k_n)\Delta_n}) - G(\widehat{C}_{i\Delta_n}) \right). \tag{15}$$

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

However, it turns out that due to the measurement errors in $\widehat{C}_{i\Delta_n}$, this estimator is inconsistent.

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

$$\widehat{[H(C), G(C)]_T^c}^{AN} = \frac{3}{2k_n} \sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1} \Bigg( \left(H(\widehat{C}_{(i+k_n)\Delta_n}) - H(\widehat{C}_{i\Delta_n})\right) \left(G(\widehat{C}_{(i+k_n)\Delta_n}) - G(\widehat{C}_{i\Delta_n})\right) - \frac{2}{k_n} \sum_{g,h,a,b=1}^{d} (\partial_{gh} H \partial_{ab} G)(\widehat{C}_{i\Delta_n}) \left(\widehat{C}_{ga,i\Delta_n} \widehat{C}_{hb,i\Delta_n} + \widehat{C}_{gb,i\Delta_n} \widehat{C}_{ha,i\Delta_n}\right) \Bigg) 1_{\{A_i \cap A_{i+k_n}\}}, \tag{16}$$

where the indicator function should only be applied if we are concerned about volatility jumps, and thus we want to truncate them. In the above, we denote by $A_i$ the event of not detecting a volatility jump in the interval $(i\Delta_n, (i+k_n)\,\Delta_n]$, defined as $A_i \equiv \{||\widehat{C}_{(i+k_n)\Delta_n} - \widehat{C}_{(i-k_n)\Delta_n}|| < u_n'\}$, where $u_n'$ is some threshold.

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

$$[H(C), G(C)]_T^c = \sum_{g,h,a,b=1}^{d} \int_0^T \left(\partial_{gh} H \partial_{ab} G\right)(C_t) \overline{C}_t^{gh,ab} dt, \tag{17}$$

where $\overline{C}_t^{gh,ab}$ denotes the continuous 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),[10]

$$\begin{aligned} &\widehat{[H(C), G(C)]_T^c}^{LIN} \\ = \ &\frac{3}{2k_n} \sum_{g,h,a,b=1}^{d} \sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1} (\partial_{gh} H \partial_{ab} G)(\widehat{C}_{i\Delta_n}) \Big( (\widehat{C}_{gh,(i+k_n)\Delta_n} - \widehat{C}_{gh,i\Delta_n})(\widehat{C}_{ab,(i+k_n)\Delta_n} - \widehat{C}_{ab,i\Delta_n}) \\ &- \frac{2}{k_n} (\widehat{C}_{ga,i\Delta_n} \widehat{C}_{hb,i\Delta_n} + \widehat{C}_{gb,i\Delta_n} \widehat{C}_{ha,i\Delta_n}) \Big) 1_{\{A_i \cap A_{i+k_n}\}}. \end{aligned} \tag{18}$$

We now provide the intuition for the bias terms. Suppose volatility is continuous. If we had observations on $C_{i\Delta_n}$, the estimators of $[H(C), G(C)]_T$ would not need any bias-correction

[10] The computation time for any of our two estimators is increasing with the number of stocks and factors $d$. In practice, we compute all the quantities of interest for pairs of stocks, so $d_S = 2$ and thus $d = d_F + 2$.

terms. It is useful to think of $\widehat{C}_{i\Delta_n}$ as an estimator of integrated volatility matrix, $\widehat{C}_{i\Delta_n} = \frac{1}{k_n\Delta_n}\int_{i\Delta_n}^{(i+k_n)\Delta_n} C_s ds + U_{i\Delta_n}$, where $U_{i\Delta_n}$ is the estimation error. The first part of the bias-correction in (16) and (18) is an additive term

$$-\frac{3}{k_n^2}\sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1}\left(\sum_{g,h,a,b=1}^{d}(\partial_{gh}H\partial_{ab}G)(\widehat{C}_{i\Delta_n})\left(\widehat{C}_{ga,i\Delta_n}\widehat{C}_{hb,i\Delta_n}+\widehat{C}_{gb,i\Delta_n}\widehat{C}_{ha,i\Delta_n}\right)\right). \tag{19}$$

This term arises because of the estimation error $U_{i\Delta_n}$. Intuitively, estimation of, e.g., variance of functionals of $C_{i\Delta_n}$ by variance of functionals of $\widehat{C}_{i\Delta_n}$ overestimates it due to the additional variability of $U_{i\Delta_n}$. In particular, one can show that the additive bias-correction term in (19) is, up to a scale factor, an estimator of the asymptotic covariance between the estimators of $\int_0^T H(C_t)dt$ and $\int_0^T G(C_t)dt$.

The second part of the bias-correction in (16) and (18) is the multiplicative correction factor 3/2. This correction factor is needed because of a smoothing bias that arises due to the replacement of $C_{i\Delta_n}$ by $\frac{1}{\Delta_n}\int_{i\Delta_n}^{(i+k_n)\Delta_n} C_s ds$. To gain some intuition, consider the special case of $d=1$ and $H(\cdot)=G(\cdot)=\cdot$ . Suppose we had observations on $\frac{1}{\Delta_n}\int_{i\Delta_n}^{(i+k_n)\Delta_n} C_s ds$. The $i^{th}$ summand in the naive estimator of $[C,C]_T$ would be

$$\left(\int_{(i+k_n)\Delta_n}^{(i+2k_n)\Delta_n} C_s ds - \int_{i\Delta_n}^{(i+k_n)\Delta_n} C_s ds\right)^2 = \left(\int_{i\Delta_n}^{(i+k_n)\Delta_n}\left(C_{s+\Delta_n k_n}-C_s\right)ds\right)^2, \tag{20}$$

divided by $\Delta_n^2 k_n^3$. Consider the weights that the integral $\int_{i\Delta_n}^{(i+k_n)\Delta_n}\left(C_{s+\Delta_n k_n}-C_s\right)ds$ puts on $\Delta_n$-increments of the volatility $C_t$: these weights are triangular, i.e., $\left(\Delta_n k_n - |\Delta_n k_n + i\Delta_n - s|\right) I\left\{s\in[i\Delta_n,(i+2k_n)\Delta_n]\right\}$. One can show that the squared integral in (20) is proportional to the integral of the squared triangular weights, $\frac{1}{(\Delta_n k_n)^3}\int_{i\Delta_n}^{(2k_n+i)\Delta_n}\left(\Delta_n k_n - |\Delta_n k_n + i\Delta_n - s|\right)^2 ds$. The latter integral equals $\frac{2}{3}$, hence the estimator needs a multiplicative correction factor $\frac{3}{2}$.

When $H(\cdot)=G(\cdot)$, the estimand is nonnegative, $[H(C),G(C)]_T^c \geq 0$, so our estimators are nonnegative in large samples. However, due to the presence of an additive bias-correction, our estimators are not guaranteed to be nonnegative in finite samples. We remark that Vetter (2015) constructs a univariate volatility of volatility estimator that is guaranteed to be nonnegative, at the cost of a slower rate of convergence.

Our two estimators, AN in equation (16) and LIN in (18), 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 $H(\cdot) = G(\cdot) = \cdot$ , and when one assumes no price or volatility jumps and omits the price and volatility jump truncation, both of our estimators coincide with the volatility of volatility estimator of Vetter (2015).

While Jacod and Rosenbaum (2015) focus on a different problem, one of the asymptotic bias terms in their paper is of the form $[H(C), H(C)]_T^c$. In the special case $H(\cdot) = G(\cdot)$, aside from a scale factor, the end-effects, and the form of the volatility jump truncation, our LIN estimator in equation (18) coincides with their estimator. Our approach to volatility jumps differs as we truncate these jump from below, while Jacod and Rosenbaum (2015) truncate from above, and we use a simpler form of truncation that in finite samples is robust to consecutive volatility jumps. Jacod and Rosenbaum (2015) only establish consistency of the relevant estimator, and do not provide any asymptotic distribution theory. In contrast, we derive the asymptotic distribution of the estimators of $[H(C), G(C)]_T^c$, and provide a consistent estimator of the asymptotic variance.

### 3.2 Estimation in R-FM and IdioVol-FM models

In this section, we explain how to use the formulas in equations (16) and (18) to obtain estimators for the objects of interest in Section 2, see equations (6)–(12). In particular, each of these objects of interest,

$$\begin{gathered} [C_{Zj}, C_{Zs}]_T^c,\ Corr\left(C_{Zj}, C_{Zs}\right),\ \gamma_{Z_j},\ [C_{Zj}^{resid}, C_{Zs}^{resid}]_T^c, \\ Corr\left(C_{Zj}^{resid}, C_{Zs}^{resid}\right),\ Q_{Zj,Zs}^{IdioVol\text{-}FM},\ \text{and}\ R_{Zj}^{2,IdioVol\text{-}FM}, \end{gathered} \tag{21}$$

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

$$\varphi\left([H_1(C), G_1(C)]_T^c, \ldots, [H_\kappa(C), G_\kappa(C)]_T^c\right), \tag{22}$$

for some smooth, real-valued functions $\varphi$, $H_r$, $G_r$, $r = 1, \ldots, \kappa$. Each element in (22) is of the form $[H_r(C), G_r(C)]_T^c$, i.e., it is the continuous part of a quadratic covariation between functions of $C_t$, and hence can be estimated using the estimators proposed in Section 3.1.

Consider the first quantity in equation (21), which is the continuous part of the quadratic covariation between $j^{th}$ and $s^{th}$ IdioVol, $[C_{Zj}, C_{Zs}]^c_T$. By equation (5), $C_{Z\ell} = C_{Y\ell,t} - (C_{FS\ell,t})^\top (C_{F,t})^{-1} C_{FS\ell,t}$, and the quantity is of the form $[C_{Zj}, C_{Zs}]^c_T = [H(C_t), G(C_t)]^c_T$, where

$$
\begin{aligned}
H(C_t) &= C_{Yj,t} - (C_{FSj,t})^\top (C_{F,t})^{-1} C_{FSj,t} \\
G(C_t) &= C_{Ys,t} - (C_{FSs,t})^\top (C_{F,t})^{-1} C_{FSs,t}.
\end{aligned}
$$

As per equation (8), $Corr(C_{Zj}, C_{Zs})$ is also of the form of equation (22).

Next, note that IdioVol-FM implies

$$
\gamma_{Zj} = \left([\Pi, \Pi]^c_T\right)^{-1} [\Pi, C_{Zj}]^c_T, \quad \text{and} \tag{23}
$$

$$
[C^{resid}_{Zj}, C^{resid}_{Zs}]^c_T = [C_{Zj}, C_{Zs}]^c_T - \gamma^\top_{Zj} [\Pi, \Pi]^c_T \gamma_{Zs} \tag{24}
$$

for $j, s = 1, \ldots, d_S$. Recall that $C_{Zj,t}$, $C_{Zs,t}$, and every element of $\Pi_t$ are given real-valued functions of $C_t$. For example, if volatility factors are the volatilities of return factors $F_t$, we have $\Pi(C_t) = C_{F,t}$, so $\Pi(\cdot)$ selects the last $d_F$ diagonal elements from $C_t$ (recall that $F_t$ are the last $d_F$ elements of vector $Y_t$). Thus, the right-hand-sides of (23) and (24) have the form of equation (22) for a finite number of quantities of the form $[H_r(C), G_r(C)]^c_T$.

Finally, the remaining quantities in equation (21), $Corr\left(C^{resid}_{Zj}, C^{resid}_{Zs}\right)$, $Q^{IdioVol\text{-}FM}_{Zj,Zs}$ and $R^{2,IdioVol\text{-}FM}_{Zj}$, are smooth functions of $[C^{resid}_{Zj}, C^{resid}_{Zj}]^c_T$, $[C_{Zj}, C_{Zs}]^c_T$, $\gamma_{Zj}$, and $[\Pi, \Pi]^c_T$, each of which is of the form of equation (22), and hence are themselves of the form of equation (22).

## 4 Asymptotic Properties

In this section, we first present the full list of assumptions for our asymptotic results. We then obtain the joint asymptotic distribution between the general functionals $[H_r(C), G_r(C)]^c_T$ for $r = 1, \ldots, \kappa$ introduced in Section 3.1. We also develop estimators for the asymptotic variance-covariance matrix. The asymptotic distributions of the estimators of $Corr(C_{Zi}, C_{Zj})$ and other quantities of interest in Section 2 follow by the Delta method (see Section 3.2 for details). 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 $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$,*

$$Y_t = Y_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_E \delta(s,z)\mu(ds,dz), \tag{25}$$

*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 $\mathbb{R}^d \times \mathbb{R}^{d^W}$-valued, and $\delta = \delta(w,t,z)$ is a predictable $\mathbb{R}^d$ -valued function on $\Omega \times \mathbb{R}_+ \times E$. Moreover, $\|\delta(w, t \wedge \tau_m(w), z)\| \wedge 1 \leq \Gamma_m(z)$, for all (w,t,z), where $(\tau_m)$ is a localizing sequence of stopping times and, for some $r \in [0, 1/2)$, 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 an Itô semimartingale,*[11]

$$C_t = C_0 + \int_0^t \widetilde{b}_s ds + \int_0^t \widetilde{\sigma}_s dW_s + J_t^\sigma, \tag{26}$$

*where $\widetilde{b}$ is $\mathbb{R}^d \times \mathbb{R}^d$-valued optional, and $J_t^\sigma$ is a finite activity jump process. $C_t$ takes values in the space $\mathcal{M}_d$ consisting of $d \times d$ positive definite matrices. For a sequence of convex compact subsets $(\mathcal{K}_m)_{m\geq 1}$ of $\mathcal{M}_d$, $C_t \in \mathcal{K}_m$ for all $t \leq \tau_m$.*

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

$$\overline{C}_t^{gh,ab} = \sum_{m=1}^{d^W} \widetilde{\sigma}_t^{gh,m} \widetilde{\sigma}_t^{ab,m}, \quad \overline{C}_t^{\prime g,ab} = \sum_{m=1}^{d^W} \sigma_t^{gm} \widetilde{\sigma}_t^{ab,m}. \tag{27}$$

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

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

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

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 Wang and Mykland (2014), Vetter (2015), and Kalnina and Xiu (2017).

## 4.2 Asymptotic Distribution

We have seen in Section 3 that all quantities of interest in (21) are functions of multiple objects of the form $[H(C), G(C)]_T^c$. Therefore, if we can obtain a multivariate asymptotic distribution for a vector with elements of the form $[H(C), G(C)]_T^c$, the asymptotic distributions for all our estimators follow by the Delta method. The current section presents this asymptotic distribution.

Let $H_1, G_1, \ldots, H_\kappa, G_\kappa$ be given smooth real-valued functions. We are interested in the asymptotic behavior of vectors

$$\begin{aligned} &\left([H_1(\widehat{C), G_1}(C)]_T^{c\,AN}, \ldots, [H_\kappa(\widehat{C), G_\kappa}(C)]_T^{c\,AN}\right)^\top \text{ and} \\ &\left([H_1(\widehat{C), G_1}(C)]_T^{c\,LIN}, \ldots, [H_\kappa(\widehat{C), G_\kappa}(C)]_T^{c\,LIN}\right)^\top. \end{aligned} \tag{28}$$

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

**Theorem 1.** *Let $[H_r(\widehat{C), G_r}(C)]_T^c$ denote either $[H_r(\widehat{C), G_r}(C)]_T^{c\,AN}$ or $[H_r(\widehat{C), G_r}(C)]_T^{c\,LIN}$ defined in equations (16) and (18), where $H_r$ and $G_r$ are three times differentiable real-valued functions, for $r = 1, \ldots, \kappa$. Suppose Assumptions 1 and 2 hold. Fix $k_n = \theta \Delta_n^{-1/2}$ for some $\theta \in (0, \infty)$. Set $u_n \asymp \Delta_n^{\varpi}$ with $\frac{2\varpi' + 9}{4(5-r)} < \varpi < \frac{1}{2}$, and $u_n' \asymp \Delta_n^{\varpi'}$ with $0 < \varpi' < \min\left(\frac{1}{2} - r, \frac{1}{8}\right)$. Then, as $\Delta_n \to 0$,*

$$\Delta_n^{-1/4} \begin{pmatrix} [H_1(\widehat{C), G_1}(C)]_T^c - [H_1(C), G_1(C)]_T^c \\ \cdots \\ [H_\kappa(\widehat{C), G_\kappa}(C)]_T^c - [H_\kappa(C), G_\kappa(C)]_T^c \end{pmatrix} \xrightarrow{L\text{-}s} MN(0, \Sigma_T). \tag{29}$$

*Let $\Sigma_T^{r,s}$ be the $(\Sigma_T)_{r,s}$ element of the $\kappa \times \kappa$ matrix $\Sigma_T$. We have*

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

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

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

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

*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 (in the absence of 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:

$$\widehat{\Omega}_T^{r,s,(1)} = \Delta_n \sum_{g,h,a,b=1}^{d} \sum_{j,k,l,m=1}^{d} \sum_{i=k_n+1}^{[T/\Delta_n]-5k_n+1} \left(\partial_{gh} H_r \partial_{ab} G_r \partial_{jk} H_s \partial_{lm} G_s(\widehat{C}_{i\Delta_n})\right)$$

$$\times\left[\widetilde{C}_{i\Delta_n}(gh,jk)\widetilde{C}_{i\Delta_n}(ab,lm)+\widetilde{C}_{i\Delta_n}(ab,jk)\widetilde{C}_{i\Delta_n}(gh,lm)\right]1_{\cap_{j=0}^{3}A_{i+jk_n}},$$

$$\widehat{\Omega}_T^{r,s,(2)}=\sum_{g,h,a,b=1}^{d}\sum_{j,k,l,m=1}^{d}\sum_{i=k_n+1}^{[T/\Delta_n]-5k_n+1}\left(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s(\widehat{C}_{i\Delta_n})\right)\left[\frac{1}{2}\widehat{\lambda}_i^{n,gh}\widehat{\lambda}_i^{n,jk}\widehat{\lambda}_{i+2k_n}^{n,ab}\widehat{\lambda}_{i+2k_n}^{n,lm}\right.$$

$$\left.+\frac{1}{2}\widehat{\lambda}_i^{n,ab}\widehat{\lambda}_i^{n,lm}\widehat{\lambda}_{i+2k_n}^{n,gh}\widehat{\lambda}_{i+2k_n}^{n,jk}+\frac{1}{2}\widehat{\lambda}_i^{n,ab}\widehat{\lambda}_i^{n,jk}\widehat{\lambda}_{i+2k_n}^{n,gh}\widehat{\lambda}_{i+2k_n}^{n,lm}+\frac{1}{2}\widehat{\lambda}_i^{n,gh}\widehat{\lambda}_i^{n,lm}\widehat{\lambda}_{i+2k_n}^{n,ab}\widehat{\lambda}_{i+2k_n}^{n,jk}\right]1_{\cap_{j=0}^{3}A_{i+jk_n}},$$

$$\widehat{\Omega}_T^{r,s,(3)}=\frac{3}{2k_n}\sum_{g,h,a,b=1}^{d}\sum_{j,k,l,m=1}^{d}\sum_{i=k_n+1}^{[T/\Delta_n]-5k_n+1}\left(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s(\widehat{C}_{i\Delta_n})\right)$$

$$\left[\widetilde{C}_{i\Delta_n}(gh,jk)\widehat{\lambda}_i^{n,ab}\widehat{\lambda}_i^{n,lm}+\widetilde{C}_{i\Delta_n}(ab,lm)\widehat{\lambda}_i^{n,gh}\widehat{\lambda}_i^{n,jk}\right.$$

$$\left.+\widetilde{C}_{i\Delta_n}(gh,lm)\widehat{\lambda}_i^{n,ab}\widehat{\lambda}_i^{n,jk}+(\widetilde{C}_{i\Delta_n}(ab,jk)\widehat{\lambda}_i^{n,gh}\widehat{\lambda}_i^{n,lm}\right]1_{\cap_{j=0}^{3}A_{i+jk_n}}.$$

with $\widehat{\lambda}_i^{n,jk}=\widehat{C}_{i+k_n}^{n,jk}-\widehat{C}_i^{n,jk}$, $\widetilde{C}_{i\Delta_n}(gh,jk)=\widehat{C}_{gj,i\Delta_n}\widehat{C}_{hk,i\Delta_n}+\widehat{C}_{gk,i\Delta_n}\widehat{C}_{hj,i\Delta_n}$, and $A_i=\{||\widehat{C}_{(i+k_n)\Delta_n}-\widehat{C}_{(i-k_n)\Delta_n}||<u_n'\}$.

The following result holds,

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

$$\frac{6}{\theta^3}\widehat{\Omega}_T^{r,s,(1)}\xrightarrow{\mathbb{P}}\Sigma_T^{r,s,(1)},\tag{30}$$

$$\frac{3}{2\theta}[\widehat{\Omega}_T^{r,s,(3)}-\frac{6}{\theta}\widehat{\Omega}_T^{r,s,(1)}]\xrightarrow{\mathbb{P}}\Sigma_T^{r,s,(3)},\ \textit{and}\tag{31}$$

$$\frac{151\theta}{140}\frac{9}{4\theta^2}[\widehat{\Omega}_T^{r,s,(2)}+\frac{4}{\theta^2}\widehat{\Omega}_T^{r,s,(1)}-\frac{4}{3}\widehat{\Omega}_T^{r,s,(3)}]\xrightarrow{\mathbb{P}}\Sigma_T^{r,s,(2)}.\tag{32}$$

The estimated matrix $\widehat{\Sigma}_T$ is symmetric but is not guaranteed to be positive semi-definite. By Theorem 1, $\widehat{\Sigma}_T$ is positive semi-definite in large samples. An interesting question is the estimation of the asymptotic variance using subsampling or bootstrap methods, see Kalnina (2011, 2023), and we leave it for future research.

**Remark 1:** The rate of convergence in equation (30) can be shown to be $\Delta_n^{-1/2}$, and the rate of convergence in (31) and (32) can be shown to be $\Delta_n^{-1/4}$.

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

**Corollary 3.** *Let* $[\widehat{H_r(C), G_r}(C)]_T^c$ *denote either* $[\widehat{H_r(C), G_r}(C)]_T^{c\;AN}$ *or* $[\widehat{H_r(C), G_r}(C)]_T^{c\;LIN}$ *defined in equations (16) and (18). Suppose the assumptions of Theorem 1 hold. Then, as* $\Delta_n \to 0$,

$$\Delta_n^{-1/4}\ \widehat{\Sigma}_T^{-1/2} \begin{pmatrix} [\widehat{H_1(C), G_1}(C)]_T^c - [H_1(C), G_1(C)]_T^c \\ \vdots \\ [\widehat{H_\kappa(C), G_\kappa}(C)]_T^c - [H_\kappa(C), G_\kappa(C)]_T^c \end{pmatrix} \xrightarrow{L} N(0, I_\kappa). \tag{33}$$

In the above, we use $L$ 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 the stable-in-law convergence. Similarly, by the Delta method, standardized asymptotic distribution can also be derived for the estimators of the quantities in (21). These standardized distributions allow the construction of confidence intervals for all the latent quantities of the form $[H_r(C), G_r(C)]_T^c$ 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 continuous components of the IdioVols of the returns on assets $j$ and $s$,

$$H_0^1 : [C_{Zj}, C_{Zs}]_T^c = 0. \tag{34}$$

Under $H_0^1$, $\Delta_n^{-1/4}[\widehat{C_{Zj}, C_{Zs}}]_T^c \widehat{V}^{-1/2} \xrightarrow{L} N(0,1)$, so we can use a t-test.

Second, we can test the hypothesis that none of the IdioVol factors $\Pi$ explaining the dynamics of IdioVol shocks of stock $j$,

$$H_0^2 : [C_{Zj}, \Pi]_T^c = 0 \tag{35}$$

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

$$\Delta_n^{-1/4} \left(\widehat{[C_{Zj}, \Pi]}_T^c\right)^\top \left(\widehat{V}\right)^{-1} \widehat{[C_{Zj}, \Pi]}_T^c \xrightarrow{L} \mathcal{X}^2_{d_\Pi, 1-\alpha}, \tag{36}$$

so we can use a Wald test. 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 of stock $j$ and $s$,

$$H_0^3 : [C_{Zj}^{resid}, C_{Zs}^{resid}]_T^c = 0. \tag{37}$$

Under $H_0^1$, $\Delta_n^{-1/4} \widehat{[C_{Zj}^{resid}, C_{Zs}^{resid}]}_T^c \widehat{V}^{-1/2} \xrightarrow{L} N(0,1)$, so we can use a t-test.

Each of the above estimators

$$\widehat{[C_{Zj}, C_{Zs}]}_T^c, \widehat{[C_{Zj}, \Pi]}_T^c, \text{ and } \widehat{[C_{Zj}^{resid}, C_{Zs}^{resid}]}_T^c$$

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

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

For the first two tests, the expression for the true asymptotic variance, $V$, is obtained using Theorem 1 and its estimation follows from Theorem 2. The asymptotic variance in the third test is obtained by applying the Delta method to the joint convergence result in Theorem 1. The expression for the estimator of the asymptotic variance, $\widehat{V}$, 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 $H_0^{1\prime} : [C_{Zj}, C_{Zs}]_t^c = 0 \text{ for all } 0 \le t \le T$, which is, in theory, stronger than our $H_0^{1\prime}$. In particular, Theorem 1 can be used to set up Kolmogorov-Smirnov type of tests for $H_0^{\prime 1}$ in the same spirit as Vetter (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 $[C_{Zj}, C_{Zs}]_t^c \geq 0 \; \forall t$, under which $H_0^1$ and $H_0^{1\prime}$ are equivalent.

## 5 Empirical Analysis

We apply our methods to study the cross-sectional dependence in IdioVols using high frequency data. One of our main findings is that stocks' IdioVols 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 IdioVols.

We use transaction prices from NYSE TAQ database for S&P 100 index constituents from 2003 to 2012. Starting with the union of constituents over this period, we select only those stocks for which complete data is available; this results in a full sample of 104 stocks. After excluding the non-trading days, our sample contains 2517 days. 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 (2020). To aid visualization, we report additional results for a subset of 30 stocks. We obtain the subset of 30 stocks by selecting at least two stocks from each of the nine GICS sectors, together with the most liquid stocks; see Table 1 for details. 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 return jump truncation threshold is the same as in simulations, see Section 6. The number of observations in the local window is taken as in Theorem 1 to be $k_n = \theta \Delta_n^{-1/2}$. We take $\theta = 2.5$ and $\Delta_n = 1/252/(6.5 \times 12)$, i.e., $\Delta_n$ is 5 minutes (with one year being a unit of time), which corresponds to the local window of approximately one week. The threshold for volatility jumps is based on the individual asset volatility changing by more than 10 percentage points. The optimal selection of this tuning parameter is a complex issue that falls outside the scope of this paper. We find that both types of estimators, AN and LIN, produce very similar results and report only the AN estimator for brevity.

To obtain the Idiosyncratic Volatilities, the preliminary step is to estimate the Return Factor Model (R-FM) for each stock. Figures G.1 and G.2 contain plots of the time series of the estimated $R^2_{Yj}$ of the R-FM for the subset of 30 stocks.[12] Each plot contains monthly $R^2_{Yj}$ from two Return Factor Models, CAPM and the Fama-French regression with market, size, and value factors. Figures G.1 and G.2 show that these time series of all stocks follow approximately the same trend with a considerable increase in the contribution around the crisis year 2008. Higher $R^2_{Yj}$ indicates that the systematic risk is relatively more important, which is typical during crises. $R^2_{Yj}$ is consistently higher in the Fama-French regression model compared to the CAPM regression model, albeit not by much. We proceed to investigate the dynamic properties of the panel of Idiosyncratic Volatilities.

We first investigate the dependence in the (total) Idiosyncratic Volatilities. Our panel has 5356 pairs of stocks. For each pair of stocks, we compute the correlation between the IdioVols, $Corr\left(C_{Zi}, C_{Zj}\right)$, see Section 3.2 for the implementation details. All pairwise correlations are positive in our sample, and their average is 0.35. Figure 1 contains a heatmap of this dependency measure in the IdioVols. We simultaneously test 5356 hypotheses of no correlation, and Figure 1 assigns non-zero correlations only for those pairs of assets, for which the null is rejected; the diagonal contains zeros, too. 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. To aid visualization, Figure 2 maps the network of dependencies in the IdioVols for the subset of 30 stocks. Similarly to Figure 1, in Figure 2, we simultaneously test 435 hypotheses of no correlation, and Figure 2 connects only the assets, for which the null is rejected. Unsurprisingly, the cross-sectional dependence between the IdioVols is also very strong among this subset of stocks.

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

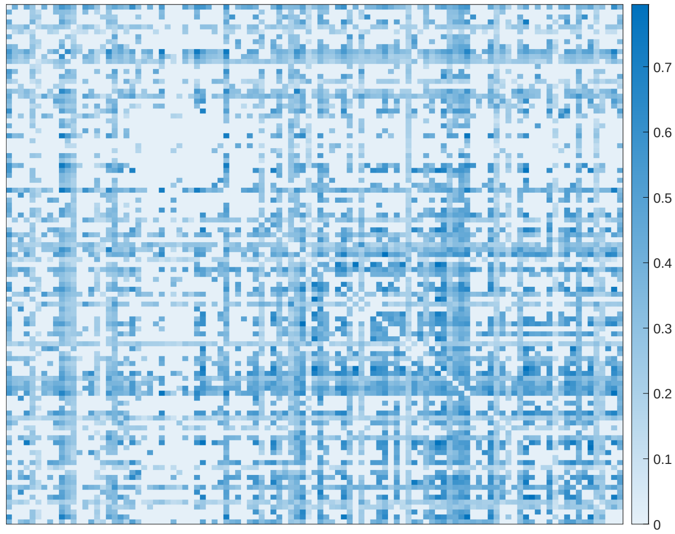


*Figure 1:* The heatmap of dependencies in total IdioVols. 104 stocks. For every pair, we test the null hypothesis of no dependence in IdioVols. If the null is rejected, the heatmap color is proportional to the estimated value of $Corr\,(C_{Zi}, C_{Zj})$, the quadratic-covariation based correlation between the IdioVols, defined in equation (8). Zero value is assigned to pairs where the null is not rejected as well as the diagonal elements.

Could missing factors in the R-FM provide an explanation? Omitted return factors in the R-FM are captured by the idiosyncratic returns, and can therefore induce correlation between the estimated IdioVols, provided these missing return factors have non-negligible volatility of volatility. To investigate this possibility, we consider the correlations between idiosyncratic returns, $Corr(Z_i, Z_j)$.[13] Table 2 presents a summary of how estimates of $Corr(Z_i, Z_j)$ are related to the estimates of correlation in IdioVols, $Corr(C_{Zi}, C_{Zj})$. In particular, different rows in Table 2 display average values of $\widehat{Corr}\,(C_{Zi}, C_{Zj})$ among those pairs, for which $|\widehat{Corr}(Z_i, Z_j)|$ is below some threshold. We observe that even among pairs with virtually uncorrelated idiosyncratic returns, the correlations among IdioVols are still high. This conclusion holds both for the idiosyncratic returns and volatilities defined with respect to CAPM, as well as the R-FM with three Fama-French factors. Moreover, we observe that IdioVol correlations, $\widehat{Corr}\,(C_{Zi}, C_{Zj})$, are similar compared among pairs

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

$$Corr(Z_i, Z_j) = \frac{\int_0^T C_{ZiZj,t}dt}{\sqrt{\int_0^T C_{Zi,t}dt}\sqrt{\int_0^T C_{Zj,t}dt}}, \quad i, j = 1, \ldots, d_S, \tag{38}$$

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

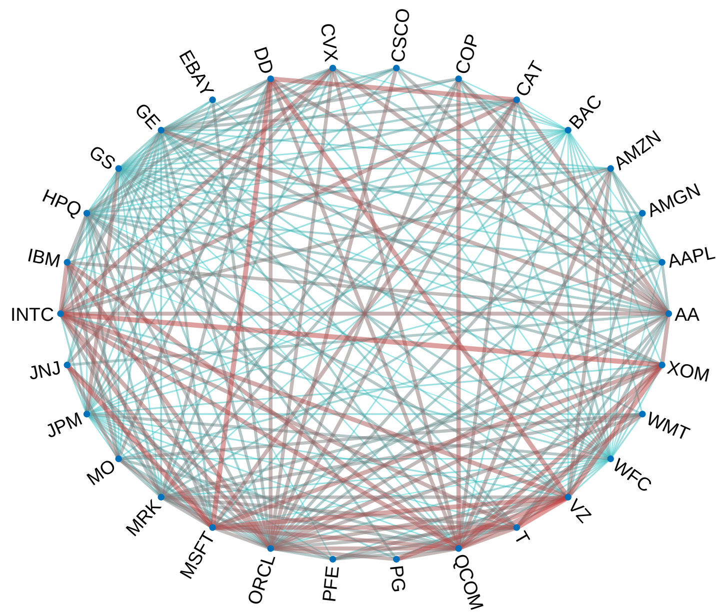


*Figure 2:* The network of dependencies in total IdioVols. 30 stocks. The color and thickness of each line is proportional to the estimated value of $Corr\,(C_{Zi}, C_{Zj})$, the quadratic-covariation based correlation between the IdioVols, defined in equation (8) (red and thick lines indicate high correlation). We simultaneously test 435 null hypotheses of no correlation, and the lines are only plotted when the null is rejected.

that have high or low idiosyncratic return correlations, $\widehat{Corr}\,(C_{Zi}, C_{Zj})$. 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 ($d_\Pi = 1$).[14] Panel (a) of Table 3 reports the estimates of the IdioVol loading ($\widehat{\gamma}_{Zi}$) and the $R^2$ of the IdioVol-FM ($R^{2,IdioVol\text{-}FM}_{Zi}$, see equation (10)). Panel (a) 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 virtually every stock, the estimated IdioVol factor loading is positive, suggesting that the Idiosyncratic Volatility co-moves with the market volatility. We have also calculated the relevant t-statistics, showing that for virtually

[14] 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.

every stock, IdioVol loading $\widehat{\gamma}_{Zi}$ is highly statistically significant. Next, Figures 3 and 5 show dependencies among residual IdioVols after accounting for the market volatility as the sole IdioVol factor. The average pairwise correlations between the residual IdioVols, $\widehat{Corr}(C_{Zi}^{resid}, C_{Zj}^{resid})$, across all pairs of stocks, decrease to 0.21. However, the market volatility cannot explain all cross-sectional dependence in residual IdioVols, as evidenced by the remaining links in both Figure 3 and 5.

Finally, we consider an IdioVol-FM with ten IdioVol factors, $d_{\Pi} = 10$, market volatility and the volatilities of nine industry ETFs. We use CAPM IdioVols. Panel (b) of Table 3 reports the corresponding $R_{Zi}^{2,Idio\,Vol\text{-}FM}$, which is considerably higher than in the one-factor case, $d_{\Pi} = 1$. Figures 4 and 6 show the implications for the cross-section of this ten-factor IdioVol-FM, for 104 and 30 stocks, respectively. The average pairwise correlations between the residual IdioVols, $\widehat{Corr}(C_{Zi}^{resid}, C_{Zj}^{resid})$, decrease further to 0.17. However, significant dependence between the residual IdioVols remains, as evidenced by the remaining links in both Figures 4 and 6. 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).

For comparison, we also calculate the naive estimators, see equation (15). Of course, since the naive estimators are inconsistent, we do not have valid confidence intervals to accompany them. We focus on the one-factor IdioVol-FM. In our data set, the absolute values of the differences between the naive and the bias-corrected estimators range, across all pairs of stocks, between 0 and 0.045 for $Corr\left(C_{Zi}, C_{Zj}\right)$, between 0 and 0.051 for $Corr(C_{Zi}^{resid}, C_{Zj}^{resid})$, and between 0.06 and 0.13 for $R_{Zj}^{2,Idio\,Vol\text{-}FM}$. However, the relative errors can be large, for example, for $R_{Zj}^{2,Idio\,Vol\text{-}FM}$, it is 42% on average. We find that in the instances where the differences are small, the multiplicative bias, i.e., the factor 2/3, dominates the additive bias both in the numerator and the denominator, so that the multiplicative bias approximately cancels out for these estimands.

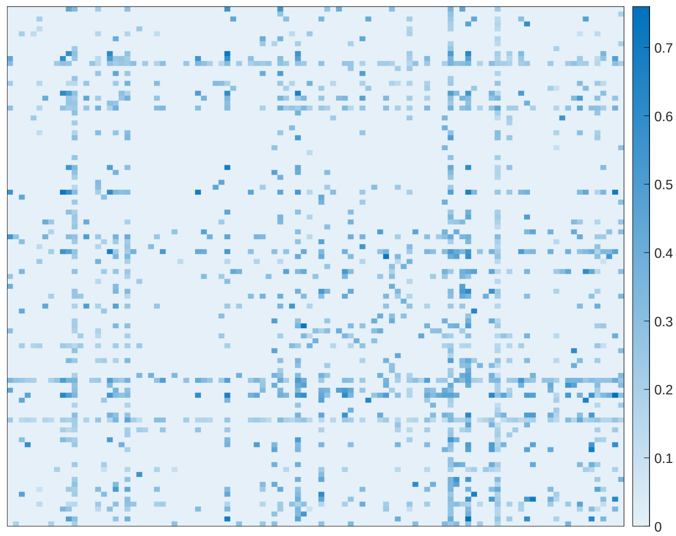


*Figure 3:* The heatmap of dependencies in residual IdioVols after accounting for a single IdioVol factor: the market variance. 104 stocks. For every pair, we test the null hypothesis of no dependence in residual IdioVols. If the null is rejected, the heatmap color is proportional to the estimated value of $Corr\left(C_{Zi}^{resid}, C_{Zj}^{resid}\right)$, the quadratic-covariation based correlation between the residual IdioVols, defined in equation (9). Zero value is assigned to pairs where the null is not rejected as well as the diagonal elements.

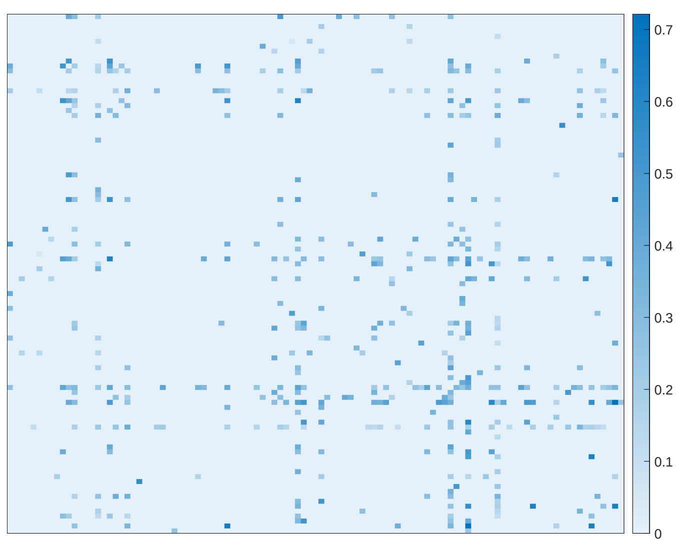


*Figure 4:* The heatmap of dependencies in residual IdioVols after accounting for ten IdioVol factors: the market variance and the variances of nine industry ETFs. 104 stocks. For every pair, we test the null hypothesis of no dependence in residual IdioVols. If the null is rejected, the heatmap color is proportional to the estimated value of $Corr\left(C_{Zi}^{resid}, C_{Zj}^{resid}\right)$, the quadratic-covariation based correlation between the residual IdioVols, defined in equation (9). Zero value is assigned to pairs where the null is not rejected as well as the diagonal elements.

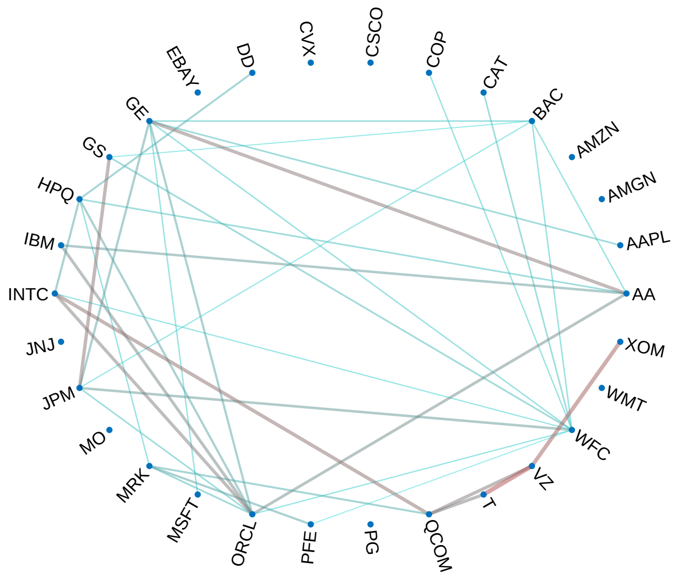


*Figure 5:* The network of dependencies in residual IdioVols after accounting for a single IdioVol factor: the market variance. 30 stocks. The color and thickness of each line is proportional to the estimated value of $Corr\left(C_{Zi}^{resid}, C_{Zj}^{resid}\right)$, the quadratic-covariation based correlation between the residual IdioVols, defined in equation (9) (red and thick lines indicate high correlation). We simultaneously test 435 null hypotheses of no correlation, and the lines are only plotted when the null is rejected.

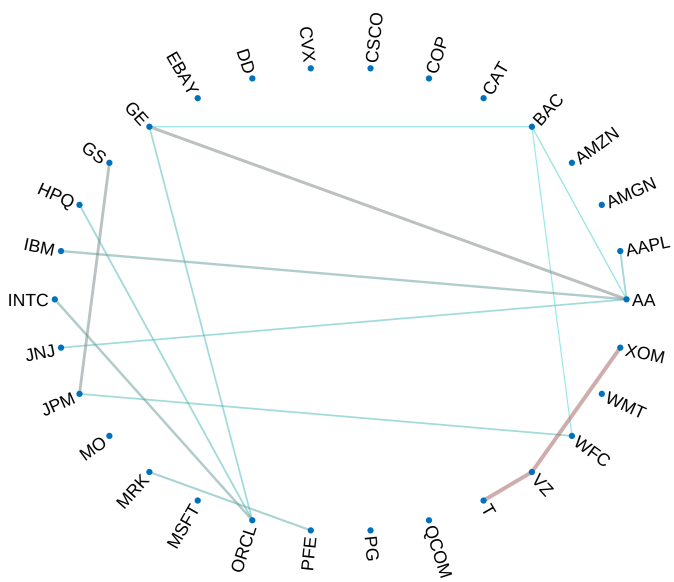


*Figure 6:* The network of dependencies in residual IdioVols after accounting for ten IdioVol factors: the market variance and the variances of nine industry ETFs. 30 stocks. The color and thickness of each line is proportional to the estimated value of $Corr\left(C_{Zi}^{resid}, C_{Zj}^{resid}\right)$, the quadratic-covariation based correlation between the residual IdioVols, defined in equation (9) (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.

| Sector | Stock | Ticker |
|---|---|---|
| Financials | Bank of America Corp | BAC |
| | Goldman Sachs Group Inc | GS |
| | JPMorgan Chase & Co | JPM |
| | Wells Fargo & Co | WFC |
| Energy | ConocoPhillips | COP |
| | Chevron Corp | CVX |
| | Exxon Mobil Corp | XOM |
| Consumer Staples | Altria Group Inc | MO |
| | Procter & Gamble Co | PG |
| | Walmart Inc | WMT |
| Industrials | Caterpillar Inc | CAT |
| | GE Aerospace | GE |
| Information Technology | Apple Inc | AAPL |
| | Cisco Systems Inc | CSCO |
| | HP Inc | HPQ |
| | Intl Business Machines Corp | IBM |
| | Intel Corp | INTC |
| | Microsoft Corp | MSFT |
| | Oracle Corp | ORCL |
| | Qualcomm Inc | QCOM |
| Health Care | Amgen Inc | AMGN |
| | Johnson & Johnson | JNJ |
| | Merck & Co | MRK |
| | Pfizer Inc | PFE |
| Consumer Discretionary | Amazon.com Inc | AMZN |
| | Ebay Inc | EBAY |
| Materials | Alcoa Corp | AA |
| | Dupont de Nemours Inc | DD |
| Communication Services | AT&T Inc | T |
| | Verizon Communications Inc | VZ |

*Table 1:* The set of 30 stocks used in the maps of network of dependencies in (residual) IdioVols in Figures 2, 5, and 6.

| | CAPM | | | FF3 Model | | |
|---|---|---|---|---|---|---|
| $\lvert\widehat{\text{Corr}}(Z_i, Z_j)\rvert$ | Pairs | Avg $\lvert\widehat{\text{Corr}}(Z_i, Z_j)\rvert$ | Avg $\widehat{\text{Corr}}\,(C_{Zi}, C_{Zj})$ | Pairs | Avg $\lvert\widehat{\text{Corr}}(Z_i, Z_j)\rvert$ | Avg $\widehat{\text{Corr}}\,(C_{Zi}, C_{Zj})$ |
| $< 0.6$ | 5356 | 0.045 | 0.347 | 5356 | 0.045 | 0.347 |
| $< 0.5$ | 5354 | 0.045 | 0.347 | 5353 | 0.045 | 0.347 |
| $< 0.4$ | 5334 | 0.044 | 0.346 | 5335 | 0.044 | 0.346 |
| $< 0.3$ | 5300 | 0.042 | 0.344 | 5300 | 0.042 | 0.345 |
| $< 0.2$ | 5236 | 0.039 | 0.343 | 5236 | 0.039 | 0.344 |
| $< 0.1$ | 4925 | 0.033 | 0.338 | 4928 | 0.033 | 0.339 |
| $< 0.075$ | 4642 | 0.030 | 0.333 | 4647 | 0.030 | 0.333 |
| $< 0.050$ | 3873 | 0.024 | 0.320 | 3895 | 0.024 | 0.320 |
| $< 0.025$ | 2049 | 0.013 | 0.296 | 2044 | 0.013 | 0.296 |
| $< 0.010$ | 757 | 0.005 | 0.293 | 748 | 0.005 | 0.293 |
| $< 0.005$ | 374 | 0.003 | 0.297 | 373 | 0.002 | 0.296 |

*Table 2:* Each row in this table describes the subset of pairs of stocks with $\lvert\widehat{\text{Corr}(Z_i, Z_j)}\rvert$ 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.

| | (a) | | | | (b) | | (a) | | | | (b) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | CAPM $d_\Pi = 1$ | | FF3 Model $d_\Pi = 1$ | | CAPM $d_\Pi = 10$ | | CAPM $d_\Pi = 1$ | | FF3 Model $d_\Pi = 1$ | | CAPM $d_\Pi = 10$ |
| Stock | $\widehat{\gamma}_Z$ | $R_Z^{2,IVFM}$ | $\widehat{\gamma}_Z$ | $R_Z^{2,IVFM}$ | $R_Z^{2,IVFM}$ | Stock | $\widehat{\gamma}_Z$ | $R_Z^{2,IVFM}$ | $\widehat{\gamma}_Z$ | $R_Z^{2,IVFM}$ | $R_Z^{2,IVFM}$ |
| AA | 0.57 | 0.12 | 0.56 | 0.12 | 0.25 | HNZ | 0.36 | 0.52 | 0.36 | 0.52 | 0.72 |
| AAPL | 0.30 | 0.07 | 0.30 | 0.07 | 0.19 | HON | 0.32 | 0.21 | 0.31 | 0.21 | 0.41 |
| ABT | 0.23 | 0.20 | 0.23 | 0.20 | 0.35 | HPQ | 0.44 | 0.15 | 0.44 | 0.15 | 0.29 |
| AEP | 0.37 | 0.29 | 0.36 | 0.29 | 0.44 | HSH | 0.19 | 0.11 | 0.19 | 0.12 | 0.18 |
| AES | 0.49 | 0.07 | 0.49 | 0.07 | 0.19 | IBM | 0.35 | 0.44 | 0.35 | 0.45 | 0.55 |
| AIG | 0.37 | 0.02 | 0.36 | 0.02 | 0.11 | INTC | 0.37 | 0.25 | 0.37 | 0.25 | 0.42 |
| ALL | 0.29 | 0.07 | 0.29 | 0.07 | 0.20 | IP | 0.34 | 0.07 | 0.33 | 0.07 | 0.22 |
| AMGN | 0.29 | 0.19 | 0.29 | 0.18 | 0.28 | JNJ | 0.37 | 0.62 | 0.37 | 0.62 | 0.69 |
| AMZN | 0.56 | 0.19 | 0.55 | 0.19 | 0.31 | JPM | 0.34 | 0.06 | 0.34 | 0.06 | 0.24 |
| APA | 0.33 | 0.14 | 0.32 | 0.14 | 0.36 | KO | 0.32 | 0.55 | 0.31 | 0.54 | 0.62 |
| APC | 0.27 | 0.05 | 0.26 | 0.04 | 0.17 | LLY | 0.40 | 0.46 | 0.39 | 0.46 | 0.55 |
| ATI | 0.35 | 0.03 | 0.35 | 0.03 | 0.14 | LMT | 0.40 | 0.28 | 0.39 | 0.28 | 0.39 |
| AVP | 0.25 | 0.04 | 0.24 | 0.03 | 0.18 | LOW | 0.47 | 0.28 | 0.45 | 0.27 | 0.41 |
| AXP | 0.40 | 0.09 | 0.39 | 0.09 | 0.29 | MCD | 0.30 | 0.20 | 0.30 | 0.20 | 0.28 |
| BA | 0.37 | 0.30 | 0.36 | 0.29 | 0.37 | MDLZ | 0.28 | 0.26 | 0.27 | 0.26 | 0.31 |
| BAC | 0.42 | 0.03 | 0.42 | 0.04 | 0.10 | MDT | 0.50 | 0.55 | 0.50 | 0.56 | 0.60 |
| BAX | 0.22 | 0.05 | 0.22 | 0.05 | 0.17 | MET | 0.25 | 0.05 | 0.25 | 0.05 | 0.14 |
| BHI | 0.25 | 0.06 | 0.25 | 0.06 | 0.21 | MMM | 0.25 | 0.30 | 0.24 | 0.29 | 0.43 |
| BK | 0.55 | 0.09 | 0.54 | 0.09 | 0.30 | MO | 0.43 | 0.36 | 0.43 | 0.36 | 0.39 |
| BMY | 0.30 | 0.25 | 0.30 | 0.25 | 0.30 | MON | 0.34 | 0.05 | 0.33 | 0.05 | 0.15 |
| C | 0.26 | 0.02 | 0.26 | 0.02 | 0.18 | MRK | 0.40 | 0.20 | 0.39 | 0.20 | 0.30 |
| CAT | 0.53 | 0.32 | 0.53 | 0.33 | 0.40 | MSFT | 0.51 | 0.59 | 0.50 | 0.60 | 0.68 |
| CI | 0.46 | 0.08 | 0.45 | 0.08 | 0.18 | NKE | 0.53 | 0.45 | 0.53 | 0.45 | 0.52 |
| CL | 0.19 | 0.29 | 0.19 | 0.30 | 0.37 | NOV | 0.30 | 0.04 | 0.29 | 0.04 | 0.18 |
| CMCSA | 0.37 | 0.20 | 0.37 | 0.20 | 0.27 | NSC | 0.41 | 0.17 | 0.41 | 0.16 | 0.36 |
| COF | 0.56 | 0.08 | 0.56 | 0.08 | 0.21 | ORCL | 0.36 | 0.25 | 0.36 | 0.25 | 0.40 |
| COP | 0.35 | 0.18 | 0.35 | 0.18 | 0.40 | OXY | 0.35 | 0.11 | 0.34 | 0.10 | 0.31 |
| COST | 0.32 | 0.30 | 0.32 | 0.30 | 0.35 | PEP | 0.27 | 0.44 | 0.27 | 0.45 | 0.57 |
| CPB | 0.17 | 0.09 | 0.17 | 0.09 | 0.26 | PFE | 0.30 | 0.20 | 0.30 | 0.21 | 0.24 |
| CSC | 0.32 | 0.08 | 0.32 | 0.08 | 0.10 | PG | 0.27 | 0.58 | 0.27 | 0.58 | 0.65 |
| CSCO | 0.39 | 0.27 | 0.39 | 0.27 | 0.44 | QCOM | 0.45 | 0.24 | 0.45 | 0.24 | 0.36 |
| CVS | 0.33 | 0.15 | 0.33 | 0.15 | 0.24 | RF | 0.50 | 0.03 | 0.50 | 0.03 | 0.14 |
| CVX | 0.29 | 0.23 | 0.28 | 0.23 | 0.47 | ROK | 0.54 | 0.22 | 0.53 | 0.22 | 0.27 |
| DD | 0.46 | 0.46 | 0.45 | 0.46 | 0.54 | S | 0.39 | 0.02 | 0.38 | 0.02 | 0.11 |
| DELL | 0.32 | 0.15 | 0.32 | 0.15 | 0.26 | SBUX | 0.49 | 0.24 | 0.48 | 0.24 | 0.32 |
| DIS | 0.42 | 0.36 | 0.42 | 0.36 | 0.50 | SO | 0.36 | 0.66 | 0.35 | 0.66 | 0.72 |
| DOW | 0.47 | 0.16 | 0.47 | 0.16 | 0.21 | T | 0.53 | 0.30 | 0.53 | 0.30 | 0.47 |
| DVN | 0.31 | 0.09 | 0.31 | 0.09 | 0.30 | TGT | 0.62 | 0.31 | 0.62 | 0.31 | 0.41 |
| EBAY | 0.50 | 0.26 | 0.50 | 0.26 | 0.45 | TWX | 0.53 | 0.41 | 0.52 | 0.41 | 0.47 |
| EMC | 0.47 | 0.20 | 0.47 | 0.20 | 0.37 | TXN | 0.42 | 0.30 | 0.42 | 0.30 | 0.46 |
| EMR | 0.26 | 0.11 | 0.26 | 0.11 | 0.19 | UIS | 0.30 | 0.01 | 0.29 | 0.01 | 0.04 |
| ETR | 0.34 | 0.32 | 0.34 | 0.32 | 0.47 | UNH | 0.63 | 0.23 | 0.64 | 0.24 | 0.28 |
| EXC | 0.44 | 0.26 | 0.42 | 0.25 | 0.40 | UNP | 0.56 | 0.29 | 0.55 | 0.29 | 0.43 |
| F | 0.52 | 0.05 | 0.51 | 0.05 | 0.15 | UPS | 0.33 | 0.49 | 0.33 | 0.49 | 0.56 |
| FDX | 0.34 | 0.30 | 0.34 | 0.30 | 0.41 | USB | 0.60 | 0.18 | 0.60 | 0.18 | 0.37 |
| GD | 0.45 | 0.43 | 0.44 | 0.44 | 0.55 | UTX | 0.38 | 0.33 | 0.38 | 0.33 | 0.49 |
| GE | 0.38 | 0.09 | 0.38 | 0.09 | 0.31 | VZ | 0.41 | 0.38 | 0.40 | 0.38 | 0.51 |
| GILD | 0.37 | 0.15 | 0.38 | 0.16 | 0.26 | WFC | 0.33 | 0.05 | 0.32 | 0.05 | 0.21 |
| GS | 0.43 | 0.12 | 0.43 | 0.12 | 0.31 | WMB | 0.35 | 0.03 | 0.35 | 0.03 | 0.10 |
| HAL | 0.29 | 0.05 | 0.29 | 0.05 | 0.21 | WMT | 0.28 | 0.47 | 0.28 | 0.48 | 0.56 |
| HD | 0.36 | 0.22 | 0.36 | 0.22 | 0.47 | XOM | 0.35 | 0.24 | 0.35 | 0.24 | 0.35 |
| HIG | 0.27 | 0.03 | 0.26 | 0.03 | 0.09 | XRX | 0.52 | 0.18 | 0.52 | 0.18 | 0.24 |

*Table 3:* Panel (a) presents estimates of the IdioVol factor loading ($\widehat{\gamma}_Z$, see eq. (6)) and the contribution of the market volatility to the variation in the IdioVols ($R_Z^{2,IVFM} = \widehat{R}_Z^{2,IdioVol\text{-}FM}$, see eq. (10)) in the one-factor IdioVol-FM, $d_\Pi = 1$. Panel (a) considers two R-FMs, CAPM or the three-factor Fama-French model (FF3). Panel (b) presents $R_Z^{2,IVFM}$ in the ten-factor IdioVol-FM, $d_\Pi = 10$, with CAPM IdioVols.

# 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 $Y_2$ the log-prices of two individual stocks, and by $X$ the log-price of the market portfolio. Recall that the superscript $c$ indicates the continuous part of a process. We assume

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

and, for $j = 1, 2$,

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

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

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

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

where $B_1, \ldots, B_4$ are independent standard Brownian Motions, which are also independent from the Brownian Motions of the return Factor Model.[15] 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 two different specifications for the IdioVol processes $C_{Z1,t}$ and $C_{Z2,t}$, 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 is $-0.8$, which

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

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).[16]

| | $C_{Z1,t}$ | $C_{Z2,t}$ |
|---|---|---|
| Model 1 | $0.1 + 1.5f_{2,t}$ | $0.1 + 1.5f_{3,t}$ |
| Model 2 | $0.1 + 0.45C_{X,t} + f_{2,t} + 0.4f_{4,t}$ | $0.1 + 0.35C_{X,t} + 0.3f_{3,t} + 0.6f_{4,t}$ |

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

We set the time span $T$ equal to 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 jump truncation threshold $u_n$ in day $t$ at $3\widehat{\sigma}_t\Delta_n^{0.49}$, where $\widehat{\sigma}_t$ is the squared root of the annualized bipower variation of Barndorff-Nielsen and Shephard (2004). We 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 $k_n = \theta/\sqrt{\Delta_n}$). We use 10,000 Monte Carlo replications in all the experiments.

We first investigate the finite sample properties of the estimators (using Model 3). We consider the following estimands:

- 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 $R_{Z1}^{2,\mathit{IdioVol\text{-}FM}}$,
- the correlation between the Idiosyncratic Volatilities of stocks 1 and 2, $Corr\,(C_{Z1}, C_{Z2})$,
- the correlation between the residual Idiosyncratic Volatilities, $Corr\left(C_{Z1}^{resid}, C_{Z2}^{resid}\right)$.

In Table 5, we report the median bias, the interquartile range (IQR), and the RMSE of the two type of the bias-corrected estimators as well as the naive estimator for each estimand using 5

[16] Notice that by Itô Lemma, each of these three models can be expressed in terms of equation (1) for the vector $(X_t, Y_{1,t}, Y_{2,t})'$ and equation (2) for the volatility matrix of this vector.

minutes data over 10 years. In Tables 5-7, in order to simplify the interpretation of the results, we fix the volatility paths $C_{X,t}$ and $(f_{j,t})_{0\leq j\leq 4}$ across simulations.

Consider first the comparison of the AN and LIN estimators. One does not consistently overperform the other in terms of the bias or the IQR. Interestingly, in terms of the RMSE, the LIN estimator outperforms the AN estimator in every scenario considered. The naive estimators are substantially biased. The comparison of the bias-corrected estimators and the naive estimators reveals the usual bias-variance trade-off, as the bias-corrected estimators have smaller bias but larger IQR than the naive estimator. In terms of RMSE, the bias-corrected estimators generally outperform the naive estimator: RMSE is significantly lower when estimating $\gamma_{Z1}$, $R^{2,IdioVol\text{-}FM}_{Z1}$, or $Corr\left(C_{Z1}, C_{Z2}\right)$, while the results for $Corr\left(C^{resid}_{Z1}, C^{resid}_{Z2}\right)$ are mixed.

It is also informative to see 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. The qualitative conclusions of Table 5 remain true in Table 6. Compared to Table 5, the bias and IQR are smaller. However, the magnitude of the decrease of the IQR is small.

Finally, Table 7 contains results from same experiment using data sampled at one minute over 5 years. Despite using more than twice as many observations than in the first experiment, the precision is not as good. In other words, increasing the time span is more effective for precision gain than increasing the sampling frequency. The qualitative conclusions generally remain the same as in Table 5.

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, $H^1_0$ : $[C_{Z1}, C_{Z2}]_T = 0$. The second null hypothesis we test is the absence of dependence between the IdioVol of the first stock and the market volatility, $H^2_0 : [C_{Z1}, C_X]_T = 0$. The third null hypothesis is the absence of dependence in the two residual IdioVols, $H^3_0 : [C^{resid}_{Z1}, C^{resid}_{Z2}]_T = 0$.

Table 8 presents the empirical rejection probabilities of the t-tests corresponding to the null hypotheses $H^1_0$, $H^2_0$, and $H^3_0$ in the above, in Model 1. In Model 1, these null hypotheses are true, so numbers in Table 8 represent empirical size. 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

| | LIN | | | | AN | | | | Naive | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\widehat{\theta}$ | 1.5 | 2 | 2.5 | 3 | 1.5 | 2 | 2.5 | 3 | 1.5 | 2 | 2.5 | 3 |
| | | | | | | **Median Bias** | | | | | | |
| $\widehat{\gamma}_{Z1}$ | -0.007 | -0.004 | -0.005 | -0.011 | -0.032 | -0.027 | -0.025 | -0.028 | -0.257 | -0.230 | -0.209 | -0.177 |
| $\widehat{R}^{2,Idio\,Vol\text{-}FM}_{Z1}$ | -0.153 | -0.138 | -0.127 | -0.115 | -0.146 | -0.132 | -0.121 | -0.110 | -0.484 | -0.465 | -0.448 | -0.417 |
| $\widehat{Corr}\left(C_{Z1}, C_{Z2}\right)$ | -0.129 | -0.104 | -0.086 | -0.059 | -0.147 | -0.118 | -0.100 | -0.070 | -0.342 | -0.334 | -0.325 | -0.307 |
| $\widehat{Corr}\left(C^{resid}_{Z1}, C^{resid}_{Z2}\right)$ | -0.089 | -0.064 | -0.045 | -0.018 | -0.109 | -0.082 | -0.061 | -0.029 | -0.245 | -0.239 | -0.232 | -0.218 |
| | | | | | | **IQR** | | | | | | |
| $\widehat{\gamma}_{Z1}$ | 0.173 | 0.157 | 0.141 | 0.118 | 0.173 | 0.154 | 0.140 | 0.118 | 0.079 | 0.078 | 0.078 | 0.076 |
| $\widehat{R}^{2,Idio\,Vol\text{-}FM}_{Z1}$ | 0.180 | 0.166 | 0.154 | 0.133 | 0.201 | 0.185 | 0.170 | 0.141 | 0.040 | 0.042 | 0.044 | 0.046 |
| $\widehat{Corr}\left(C_{Z1}, C_{Z2}\right)$ | 0.279 | 0.257 | 0.238 | 0.211 | 0.321 | 0.289 | 0.266 | 0.229 | 0.039 | 0.041 | 0.043 | 0.048 |
| $\widehat{Corr}\left(C^{resid}_{Z1}, C^{resid}_{Z2}\right)$ | 0.330 | 0.304 | 0.280 | 0.249 | 0.381 | 0.344 | 0.311 | 0.273 | 0.040 | 0.042 | 0.044 | 0.049 |
| | | | | | | **RMSE** | | | | | | |
| $\widehat{\gamma}_{Z1}$ | 0.130 | 0.116 | 0.105 | 0.090 | 0.132 | 0.118 | 0.108 | 0.093 | 0.263 | 0.238 | 0.217 | 0.185 |
| $\widehat{R}^{2,Idio\,Vol\text{-}FM}_{Z1}$ | 0.206 | 0.185 | 0.170 | 0.150 | 0.242 | 0.192 | 0.174 | 0.152 | 0.484 | 0.466 | 0.449 | 0.418 |
| $\widehat{Corr}\left(C_{Z1}, C_{Z2}\right)$ | 0.257 | 0.226 | 0.203 | 0.169 | 0.309 | 0.260 | 0.229 | 0.187 | 0.343 | 0.335 | 0.327 | 0.309 |
| $\widehat{Corr}\left(C^{resid}_{Z1}, C^{resid}_{Z2}\right)$ | 0.300 | 0.261 | 0.235 | 0.199 | 0.394 | 0.309 | 0.266 | 0.213 | 0.247 | 0.241 | 0.234 | 0.221 |

*Table 5:* Finite sample properties of our estimators using 10 years of data sampled at 5 minutes. The true values are $\gamma_{Z1} = 0.450$, $R^{2,Idio\,Vol\text{-}FM}_{Z1} = 0.342$, $Corr\left(C_{Z1}, C_{Z2}\right) = 0.523$, $Corr\left(C^{resid}_{Z1}, C^{resid}_{Z2}\right) = 0.424$. Model 2.

| | LIN | | | | AN | | | | Naive | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\widehat{\theta}$ | 1.5 | 2 | 2.5 | 3 | 1.5 | 2 | 2.5 | 3 | 1.5 | 2 | 2.5 | 3 |
| | | | | | | **Median Bias** | | | | | | |
| $\widehat{\gamma}_{Z1}$ | -0.034 | -0.029 | -0.022 | -0.013 | -0.052 | -0.044 | -0.036 | -0.025 | -0.304 | -0.295 | -0.275 | -0.267 |
| $\widehat{R}^{2,Idio\,Vol\text{-}FM}_{Z1}$ | -0.140 | -0.123 | -0.109 | -0.086 | -0.135 | -0.117 | -0.103 | -0.080 | -0.496 | -0.492 | -0.477 | -0.473 |
| $\widehat{Corr}\left(C_{Z1}, C_{Z2}\right)$ | -0.146 | -0.128 | -0.114 | -0.091 | -0.163 | -0.143 | -0.129 | -0.104 | -0.327 | -0.323 | -0.321 | -0.317 |
| $\widehat{Corr}\left(C^{resid}_{Z1}, C^{resid}_{Z2}\right)$ | -0.118 | -0.105 | -0.095 | -0.076 | -0.138 | -0.123 | -0.111 | -0.092 | -0.220 | -0.216 | -0.216 | -0.212 |
| | | | | | | **IQR** | | | | | | |
| $\widehat{\gamma}_{Z1}$ | 0.147 | 0.132 | 0.118 | 0.100 | 0.146 | 0.131 | 0.117 | 0.099 | 0.062 | 0.062 | 0.063 | 0.063 |
| $\widehat{R}^{2,Idio\,Vol\text{-}FM}_{Z1}$ | 0.165 | 0.148 | 0.137 | 0.119 | 0.176 | 0.158 | 0.145 | 0.125 | 0.032 | 0.032 | 0.034 | 0.034 |
| $\widehat{Corr}\left(C_{Z1}, C_{Z2}\right)$ | 0.260 | 0.232 | 0.209 | 0.175 | 0.287 | 0.249 | 0.224 | 0.188 | 0.032 | 0.032 | 0.033 | 0.033 |
| $\widehat{Corr}\left(C^{resid}_{Z1}, C^{resid}_{Z2}\right)$ | 0.312 | 0.280 | 0.254 | 0.211 | 0.341 | 0.303 | 0.273 | 0.225 | 0.032 | 0.032 | 0.033 | 0.033 |
| | | | | | | **RMSE** | | | | | | |
| $\widehat{\gamma}_{Z1}$ | 0.115 | 0.102 | 0.091 | 0.076 | 0.121 | 0.106 | 0.095 | 0.078 | 0.307 | 0.299 | 0.279 | 0.271 |
| $\widehat{R}^{2,Idio\,Vol\text{-}FM}_{Z1}$ | 0.192 | 0.165 | 0.147 | 0.121 | 0.198 | 0.168 | 0.148 | 0.121 | 0.496 | 0.493 | 0.478 | 0.474 |
| $\widehat{Corr}\left(C_{Z1}, C_{Z2}\right)$ | 0.251 | 0.220 | 0.196 | 0.162 | 0.283 | 0.243 | 0.215 | 0.177 | 0.328 | 0.324 | 0.322 | 0.318 |
| $\widehat{Corr}\left(C^{resid}_{Z1}, C^{resid}_{Z2}\right)$ | 0.291 | 0.249 | 0.221 | 0.182 | 0.760 | 0.279 | 0.245 | 0.199 | 0.221 | 0.218 | 0.218 | 0.214 |

*Table 6:* Finite sample properties of our estimators using 10 years of data sampled at 1 minute. The true values are $\gamma_{Z1} = 0.450$, $R^{2,Idio\,Vol\text{-}FM}_{Z1} = 0.336$, $Corr\left(C_{Z1}, C_{Z2}\right) = 0.514$, $Corr\left(C^{resid}_{Z1}, C^{resid}_{Z2}\right) = 0.408$. Model 2.

| | LIN | | | | AN | | | | Naive | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\widehat{\theta}$ | 1.5 | 2 | 2.5 | 3 | 1.5 | 2 | 2.5 | 3 | 1.5 | 2 | 2.5 | 3 |
| | | | | | | **Median Bias** | | | | | | |
| $\widehat{\gamma}_{Z1}$ | -0.075 | -0.072 | -0.068 | -0.061 | -0.096 | -0.089 | -0.083 | -0.075 | -0.323 | -0.315 | -0.299 | -0.291 |
| $\widehat{R}_{Z1}^{2,Idio\,Vol\text{-}FM}$ | -0.183 | -0.169 | -0.155 | -0.139 | -0.183 | -0.169 | -0.156 | -0.137 | -0.500 | -0.496 | -0.484 | -0.480 |
| $\widehat{Corr}\left(C_{Z1}, C_{Z2}\right)$ | -0.187 | -0.169 | -0.161 | -0.145 | -0.214 | -0.194 | -0.185 | -0.166 | -0.321 | -0.316 | -0.317 | -0.313 |
| $\widehat{Corr}\left(C_{Z1}^{resid}, C_{Z2}^{resid}\right)$ | -0.144 | -0.128 | -0.125 | -0.116 | -0.167 | -0.155 | -0.146 | -0.139 | -0.209 | -0.205 | -0.207 | -0.202 |
| | | | | | | **IQR** | | | | | | |
| $\widehat{\gamma}_{Z1}$ | 0.229 | 0.205 | 0.184 | 0.154 | 0.225 | 0.202 | 0.184 | 0.154 | 0.092 | 0.092 | 0.093 | 0.093 |
| $\widehat{R}_{Z1}^{2,Idio\,Vol\text{-}FM}$ | 0.246 | 0.223 | 0.206 | 0.177 | 0.265 | 0.238 | 0.218 | 0.187 | 0.047 | 0.047 | 0.049 | 0.049 |
| $\widehat{Corr}\left(C_{Z1}, C_{Z2}\right)$ | 0.407 | 0.357 | 0.325 | 0.281 | 0.453 | 0.394 | 0.354 | 0.299 | 0.047 | 0.046 | 0.049 | 0.048 |
| $\widehat{Corr}\left(C_{Z1}^{resid}, C_{Z2}^{resid}\right)$ | 0.475 | 0.419 | 0.387 | 0.324 | 0.529 | 0.462 | 0.420 | 0.352 | 0.047 | 0.047 | 0.049 | 0.049 |
| | | | | | | **RMSE** | | | | | | |
| $\widehat{\gamma}_{Z1}$ | 0.184 | 0.165 | 0.150 | 0.127 | 0.192 | 0.172 | 0.156 | 0.134 | 0.330 | 0.321 | 0.307 | 0.298 |
| $\widehat{R}_{Z1}^{2,Idio\,Vol\text{-}FM}$ | 0.330 | 0.240 | 0.218 | 0.188 | 0.420 | 0.246 | 0.225 | 0.192 | 0.501 | 0.497 | 0.486 | 0.482 |
| $\widehat{Corr}\left(C_{Z1}, C_{Z2}\right)$ | 0.409 | 0.342 | 0.307 | 0.260 | 0.500 | 0.388 | 0.345 | 0.285 | 0.322 | 0.318 | 0.319 | 0.314 |
| $\widehat{Corr}\left(C_{Z1}^{resid}, C_{Z2}^{resid}\right)$ | 0.510 | 0.399 | 0.355 | 0.287 | 0.813 | 0.481 | 0.417 | 0.323 | 0.212 | 0.207 | 0.209 | 0.205 |

*Table 7:* Finite sample properties of our estimators using 5 years of data sampled at 1 minute. The true values are $\gamma_{Z1} = 0.450$, $R_{Z1}^{2,Idio\,Vol\text{-}FM} = 0.35$, $Corr\left(C_{Z1}, C_{Z2}\right) = 0.517$, $Corr\left(C_{Z1}^{resid}, C_{Z2}^{resid}\right) = 0.417$. Model 2.

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 9 presents the empirical rejection probabilities of the t-tests for the same null hypotheses in Model 2. In this model, all three null hypotheses are false, so the numbers in the table represent power. The magnitude of dependence between the residual IdioVols, $[C_{Z1}^{resid}, C_{Z2}^{resid}]_T$, is of course smaller than the magnitude of the dependence between total IdioVols, $[C_{Z1}, C_{Z2}]_T$, so the power in Panel C is lower than in Panel A. However, in most of the cases the power is still nontrivial, especially for larger block sizes $\theta$, and clearly increasing with higher frequency.

| | $\Delta_n = 5$ minutes | | | | | | $\Delta_n = 1$ minute | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\theta = 1.5$ | | $\theta = 2.0$ | | $\theta = 2.5$ | | $\theta = 1.5$ | | $\theta = 2.0$ | | $\theta = 2.5$ | |
| | AN | LIN | AN | LIN | AN | LIN | AN | LIN | AN | LIN | AN | LIN |
| **Panel A :** $H_0^1 : [C_{Z1}, C_{Z2}]_T = 0$ | | | | | | | | | | | | |
| $\alpha = 10\%$ | 9.8 | 12.1 | 10.8 | 12.6 | 11.1 | 12.6 | 10.8 | 11.4 | 11.3 | 11.1 | 10.7 | 11.2 |
| $\alpha = 5\%$ | 5.5 | 5.2 | 5.4 | 6.1 | 6.0 | 6.9 | 6.6 | 6.3 | 5.7 | 5.3 | 5.3 | 5.1 |
| $\alpha = 1\%$ | 1.0 | 1.5 | 0.9 | 1.7 | 0.5 | 1.1 | 1.6 | 1.3 | 1.2 | 1.1 | 0.9 | 0.6 |
| **Panel B :** $H_0^2 : [C_{Z1}, C_X]_T = 0$ | | | | | | | | | | | | |
| $\alpha = 10\%$ | 10.2 | 10.3 | 10.4 | 10.9 | 9.9 | 10.0 | 9.7 | 8.9 | 9.2 | 9.0 | 10.4 | 10.4 |
| $\alpha = 5\%$ | 4.6 | 4.5 | 4.5 | 4.6 | 4.8 | 5.1 | 5.1 | 4.5 | 4.8 | 5.4 | 5.4 | 5.3 |
| $\alpha = 1\%$ | 0.8 | 0.5 | 1.1 | 0.9 | 1.1 | 1.3 | 1.1 | 1.3 | 1.1 | 1.2 | 0.9 | 1.1 |
| **Panel C :** $H_0^3 : [C_{Z1}^{resid}, C_{Z2}^{resid}]_T = 0$ | | | | | | | | | | | | |
| $\alpha = 10\%$ | 10.0 | 11.7 | 10.8 | 12.7 | 11.5 | 12.6 | 11.0 | 11.2 | 11.2 | 10.7 | 10.7 | 11.7 |
| $\alpha = 5\%$ | 5.9 | 4.9 | 5.7 | 6.1 | 5.9 | 7.3 | 6.4 | 6.4 | 5.4 | 5.2 | 4.9 | 4.9 |
| $\alpha = 1\%$ | 1.0 | 1.5 | 0.9 | 1.5 | 0.6 | 1.0 | 1.8 | 1.4 | 1.3 | 1.2 | 0.9 | 0.6 |

*Table 8:* The size of the t-tests. Model 1. $T = 10$ years. $\alpha$ denotes the nominal size of the test.

| | $\Delta_n = 5$ minutes | | | | | | $\Delta_n = 1$ minute | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\theta = 1.5$ | | $\theta = 2.0$ | | $\theta = 2.5$ | | $\theta = 1.5$ | | $\theta = 2.0$ | | $\theta = 2.5$ | |
| | AN | LIN | AN | LIN | AN | LIN | AN | LIN | AN | LIN | AN | LIN |
| | **Panel A :** $H_0^1 : [C_{Z1}, C_{Z2}]_T = 0$ | | | | | | | | | | | |
| $\alpha = 10\%$ | 20.3 | 31.5 | 36.8 | 47.0 | 54.2 | 64.8 | 32.5 | 39.8 | 64.6 | 69.6 | 88.0 | 91.0 |
| $\alpha = 5\%$ | 11.9 | 21.4 | 25.4 | 36.5 | 41.0 | 54.0 | 21.8 | 28.1 | 49.5 | 57.2 | 79.2 | 84.4 |
| $\alpha = 1\%$ | 3.0 | 7.0 | 8.2 | 16.9 | 20.1 | 28.6 | 9.9 | 13.2 | 27.6 | 32.2 | 54.8 | 62.2 |
| | **Panel B :** $H_0^2 : [C_{Z1}, C_X]_T = 0$ | | | | | | | | | | | |
| $\alpha = 10\%$ | 60.2 | 67.6 | 83.0 | 87.9 | 93.9 | 96.3 | 91.8 | 93.6 | 99.6 | 99.6 | 100.0 | 100.0 |
| $\alpha = 5\%$ | 45.8 | 57.2 | 72.9 | 79.0 | 88.5 | 91.9 | 86.6 | 89.5 | 98.4 | 98.8 | 100.0 | 100.0 |
| $\alpha = 1\%$ | 23.4 | 31.6 | 50.8 | 58.6 | 70.6 | 76.4 | 68.5 | 72.5 | 94.0 | 95.2 | 99.2 | 99.3 |
| | **Panel C :** $H_0^3 : [C_{Z1}^{resid}, C_{Z2}^{resid}]_T = 0$ | | | | | | | | | | | |
| $\alpha = 10\%$ | 14.2 | 19.9 | 22.6 | 29.5 | 30.9 | 38.6 | 19.6 | 22.3 | 33.5 | 36.5 | 52.9 | 58.4 |
| $\alpha = 5\%$ | 7.4 | 12.6 | 14.1 | 20.5 | 21.6 | 29.2 | 12.1 | 14.8 | 22.4 | 26.6 | 39.8 | 44.6 |
| $\alpha = 1\%$ | 1.5 | 3.3 | 4.8 | 6.9 | 8.4 | 12.1 | 3.2 | 5.2 | 10.0 | 12.1 | 19.5 | 22.9 |

*Table 9:* The power of the t-tests. Model 2. $T = 10$ 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 S&P100 index 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. For each model, we map out the network of dependencies in residual (non-systematic) Idiosyncratic Volatilities across all stocks.

## 8 Acknowledgements

We are grateful to co-editors Torben Andersen and Serena Ng, two associate editors, and four anonymous referees for numerous helpful suggestions. We benefited from discussions with Marine Carrasco, Yoosoon Chang, Valentina Corradi, Russell Davidson, Jean-Marie Dufour, Prosper Dovonon, Kirill Evdokimov, Sílvia 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, for helpful comments and suggestions. Ilze Kalnina 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.

## References


Aït-Sahalia, Y., J. Fan, R. Laeven, D. C. Wang, and X. Yang (2017): "The Estimation of Continuous and Discontinuous Leverage Effects," *Journal of the American Statistical Association*, 112, 1744–1758.

Aït-Sahalia, Y., J. Fan, and Y. Li (2013): "The Leverage Effect Puzzle: Disentangling Sources of Bias at High Frequency," *Journal of Financial Economics*, 109, 224–249.

Aït-Sahalia, Y., and J. Jacod (2014): *High-Frequency Financial Econometrics*. Princeton University Press.

Aït-Sahalia, Y., I. Kalnina, and D. Xiu (2020): "High-Frequency Factor Models and Regressions," *Journal of Econometrics*, 216, 86–105.

Aït-Sahalia, Y., and D. Xiu (2017): "Using Principal Component Analysis to Estimate a High Dimensional Factor Model with High-Frequency Data," *Journal of Econometrics*, 201, 384–399.

——— (2019): "Principal Component Analysis of High Frequency Data," *Journal of the American Statistical Association*, 114, 287–303.

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

Andersen, T. G., T. Bollerslev, F. X. Diebold, and J. Wu (2006): "Realized Beta: Persistence and Predictability," in *Econometric Analysis of Economic and Financial Time Series*, ed. by T. Fomby, and D. Terrell, vol. 20 of *Advances in Econometrics*, pp. 1–40. Elsevier.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006): "The cross-section of volatility and expected returns," *Journal of Finance*, 61(1), 259–299.

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

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

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

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

Bollerslev, T., and H. Zhou (2002): "Estimating stochastic volatility diffusion using conditional moments of integrated volatility," *Journal of Econometrics*, 109, 33–65.

Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu (2001): "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance*, 56(1), 1–43.

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

CHONG, C. H., AND V. TODOROV (2024): "Volatility of volatility and leverage effect from options," *Journal of Econometrics*, 240(1).

CHRISTOFFERSEN, P., M. FOURNIER, AND K. JACOBS (2018): "The Factor Structure in Equity Options," *Review of Financial Studies*, 31, 595–637.

CONNOR, G., R. A. KORAJCZYK, AND O. LINTON (2006): "The common and specific components of dynamic volatility," *Journal of Econometrics*, 132(1), 231–255.

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

DIEBOLD, F. X., AND K. YILMAZ (2014): "On the network topology of variance decompositions: Measuring the connectedness of financial firms," *Journal of Econometrics*, 182, 119–132.

DUARTE, J., A. KAMARA, S. SIEGEL, AND C. SUN (2014): "The Systematic Risk of Idiosyncratic Volatility," Discussion paper.

FAN, J., A. FURGER, AND D. XIU (2016): "Incorporating Global Industrial Classification Standard into Portfolio Allocation: A Simple Factor-Based Large Covariance Matrix Estimator with High Frequency Data," *Journal of Business and Economics*, 34, 489–503.

GOURIER, E. (2016): "Pricing of Idiosyncratic Equity and Variance Risks," Discussion paper, Queen Mary University.

HANSEN, P. R., A. LUNDE, AND V. VOEV (2014): "Realized Beta GARCH: a Multivariate GARCH Model with Realized Measures of Volatility," *Journal of Applied Econometrics*, 29, 774–799.

HERSKOVIC, B., B. KELLY, H. LUSTIG, AND S. V. NIEUWERBURGH (2016): "The Common Factor in Idiosyncratic Volatility: Quantitative Asset Pricing Implications," *Journal of Financial Economics*, 119, 249–283.

JACOD, J., AND P. PROTTER (2012): *Discretization of Processes*. Springer.

JACOD, J., AND M. ROSENBAUM (2013): "Quarticity and Other Functionals of Volatility: Efficient Estimation," *Annals of Statistics*, 41, 1462–1484.

JACOD, J., AND M. ROSENBAUM (2015): "Estimation of Volatility Functionals: the Case of a $\sqrt{n}$ Window," in *Large Deviations and Asymptotic Methods in Finance*, ed. by P. Friz, J. Gatheral, A. Gulisashvili, A. Jacquier, and J. Teichmann, vol. 110, pp. 559–590. Springer.

KALNINA, I. (2011): "Subsampling high frequency data," *Journal of Econometrics*, 161, 262–283.

——— (2023): "Inference for Nonparametric High-Frequency Estimators with an Application to Time Variation in Betas," *Journal of Business & Economic Statistics*, 41(2), 538–549.

KALNINA, I., AND D. XIU (2017): "Nonparametric Estimation of the Leverage Effect: A Trade-off between Robustness and Efficiency," *Journal of the American Statistical Association*, 112, 384–396.

Kanaya, S., and D. Kristensen (2016): "Estimation of Stochastic Volatility Models by Nonparametric Filtering," *Econometric Theory*, 32, 861–916.

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

Li, J., Y. Liu, and D. Xiu (2019): "Efficient Estimation of Integrated Volatility Functionals via Multiscale Jackknife," *Annals of Statistics*, 47, 156–176.

Li, J., and A. J. Patton (2018): "Asymptotic Inference about Predictive Accuracy Using High Frequency Data," *Journal of Econometrics*, 232, 223–240.

Li, J., V. Todorov, and G. Tauchen (2013): "Inference Theory on Volatility Functional Dependencies," Discussion paper, Duke University and Northwestern University.

——— (2016): "Inference Theory on Volatility Functional Dependencies," *Journal of Econometrics*, 193, 17–34.

——— (2017a): "Adaptive Estimation of Continuous-Time Regression Models using High-Frequency Data," *Journal of Econometrics*, 200, 36–47.

——— (2017b): "Jump Regressions," *Econometrica*, 85, 173–195.

Li, Y., G. Liu, and Z. Zhang (2022): "Volatility of volatility: Estimation and tests based on noisy high frequency data with jumps," *Journal of Econometrics*, 229, 422–451.

Liu, L., A. Patton, and K. Sheppard (2015): "Does Anything Beat 5-Minute RV? A Comparison of Realized Measures Across Multiple Asset Classes," *Journal of Econometrics*, 187, 293–311.

Mykland, P. A., and L. Zhang (2006): "ANOVA for diffusions and Itô processes," *Annals of Statistics*, 34(4), 1931–1963.

Pelger, M. (2019): "Large-Dimensional Factor Modeling Based on High-Frequency Observations," *Journal of Econometics*, 208, 23–42.

——— (2020): "Understanding Systematic Risk: A High-Frequency Approach," *Journal of Finance*, 75(4), 2179–2220.

Todorov, V. (2009): "Estimation of continuous-time stochastic volatility models with jumps using high-frequency data," *Journal of Econometrics*, 148, 131–148.

Vetter, M. (2015): "Estimation of Integrated Volatility of Volatility with Applications to Goodness-of-fit Testing," *Bernoulli*, 21, 2393–2418.

Wang, D. C., and P. A. Mykland (2014): "The Estimation of Leverage Effect with High Frequency Data," *Journal of the American Statistical Association*, 109, 197–215.

# Appendix

Sections A-E contain all proofs. Section F contains some numerical implementation details. Section G contains additional figures for the empirical application.

The proofs are organised as follows. Section A introduces additional notation. Section B presents auxiliary theorems and lemmas used to prove Theorems 1 and 2 in the main paper. Section C proves Theorem 1. Section D proves Theorem 2. Section E collects the proofs of the auxiliary results of Section B.

## A Notation for Proofs

Our notation is similar to that of the proofs of Jacod and Rosenbaum (2015) whenever possible. Throughout, we denote by $K$ a generic constant, which may change from line to line. We let by convention $\sum_{i=a}^{a'} = 0$ when $a > a'$. For simplicity, we omit the subscript $r$ for results involving only one object with this subscript. By the usual localization argument, there exists a $\pi$-integrable function $J$ on $E$ and a constant such that the stochastic processes in equations (26) and (27) satisfy

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

We set

$$\mathcal{F}_i^n = \mathcal{F}_{i\Delta_n},\ C_i^n = C_{i\Delta_n},\ \overline{C}_i^n = \overline{C}_{i\Delta_n},\ \text{and}\ \widehat{C}_i^n = \widehat{C}_{i\Delta_n}.$$

For any càdlàg bounded process $Z$, we set

$$\eta_{t,s}(Z) = \sqrt{\mathbb{E}\Big(\sup_{0<u\leq s} \|Z_{t+u} - Z_t\|^2 | \mathcal{F}_t\Big)}, \quad \text{and}$$

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

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 $\widehat{C}_i'^n$ be the local estimator of the spot variance of the unobservable process $Y'$, i.e.,

$$\widehat{C}_i'^n = \frac{1}{k_n\Delta_n} \sum_{u=0}^{k_n-1} (\Delta_{i+u}^n Y')(\Delta_{i+u}^n Y)'^\top = (\widehat{C}_i'^{n,gh})_{1\leq g,h\leq d}. \tag{A.2}$$

There is no price jump truncation applied in the definition of $\widehat{C}_i'^n$ since the process $Y'$ is continuous. Hence, it is more convenient to work with $\widehat{C}_i'^n$ rather than $\widehat{C}_i^n$ $(= \widehat{C}_{i\Delta_n}$, defined in equation (14)).
We also define

$$\alpha_i^n = (\Delta_i^n Y')(\Delta_i^n Y')^\top - C_{(i-1)\Delta_n}\Delta_n,\ \ \nu_i^n = \widehat{C}_i'^n - C_{(i-1)\Delta_n},\ \ \text{and}\ \ \lambda_i^n = \widehat{C}_{i+k_n}'^n - \widehat{C}_i'^n, \tag{A.3}$$

which satisfy

$$\nu_i^n = \frac{1}{k_n\Delta_n} \sum_{j=0}^{k_n-1} (\alpha_{i+j}^n + (C_{(i+j-1)\Delta_n} - C_{(i-1)\Delta_n})\Delta_n)\ \text{ and }\ \lambda_i^n = \nu_{i+k_n} - \nu_i^n + C_{(i+k_n-1)\Delta_n} - C_{(i-1)\Delta_n}. \tag{A.4}$$

The following multidimensional quantities will be used in the sequel

$$\begin{aligned}
&\zeta(1)_i^n = \tfrac{1}{\Delta_n}\Delta_i^n Y'(\Delta_i^n Y')^\top - C_{i-1}^n, && \zeta(2)_i^n = \Delta_i^n c,\\
&\zeta'(u)_i^n = \mathbb{E}(\zeta(u)_i^n|\mathcal{F}_{i-1}^n), && \zeta''(u)_i^n = \zeta(u)_i^n - \zeta'(u)_i^n,\\
&\zeta^r(u)_i^n = \left(\zeta^r(u)_i^{n,gh}\right)_{1\le g,h\le d} && \text{with } r =' \text{ or }''.
\end{aligned}$$

For $1 \le g,h \le d$ and $u,v = 1,2$, define

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

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

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

For any $u,v,m,u',v'$, we set

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

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

$$M(u,v;u',v')_n = z_{u,v}^n z_{u',v'}^n \sum_{m=1}^{2k_n-1} \lambda(u,v)_m^n \lambda(u',v')_m^n.$$

We also need some notation for volatility jumps. Denote by $N_s$ the number of jumps in $C$ from time 0 to $s$. Let

$$\begin{aligned}
L(n) &= \left\{i = k_n+1, k_n+2, ... : N_{(i+3)k_n\Delta_n} - N_{(i-1)k_n\Delta_n} = 0\right\},\\
L(n,T) &= \{i = 1,2,...,[T/\Delta_n] - 3k_n + 1\} \cap L(n),\\
L'(n,T) &= \{i = 1,2,...,[T/\Delta_n] : i - 2k_n \in L(n,T)\},\\
\overline{L}(n,T) &= \{i = 1,2,...,[T/\Delta_n] - 3k_n + 1\} \backslash L(n).
\end{aligned} \tag{A.5}$$

Additionally, set

$$\begin{aligned}
\overline{A11}(H,gh,u;G,ab,v)_T^n &= \frac{3}{2k_n^3} \sum_{i\in L'(n,T)} \Big(\sum_{j=0}^{2k_n-1} \varepsilon(u)_j^n \varepsilon(v)_j^n\Big)(\partial_{gh}H\partial_{ab}G)(C_{(i-2k_n-1)\Delta_n})\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}\\
&= \lambda(u,v)_0^n \sum_{i\in L'(n,T)} (\partial_{gh}H\partial_{ab}G)(C_{(i-2k_n-1)\Delta_n})\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab},
\end{aligned} \tag{A.6}$$

and

$$\begin{aligned}
\overline{A12}(H,gh,u;G,ab,v)_T^n &= \frac{3}{2k_n^3} \sum_{i\in L'(n,T)} (\partial_{gh}H\partial_{ab}G)(C_{(i-2k_n-1)\Delta_n}) \sum_{m=1}^{(i-1)\wedge(2k_n-1)} \sum_{j=0}^{(2k_n-m-1)} \varepsilon(u)_j^n \varepsilon(v)_{j+m}^n\\
&\quad \times \zeta_{gh}(u)_{i-m}^n \zeta_{ab}(v)_i^n.
\end{aligned} \tag{A.7}$$

Denote by $\vartheta_i^{AN}$ and $\vartheta_i^{LIN}$ the $i^{th}$ summand of $[H(\widehat{C),G}(C)]_T^{c\ AN}$ and $[H(\widehat{C),G}(C)]_T^{c\ LIN}$, without the volatility jump truncation, so they satisfy

$$[H(\widehat{C), G(C)}]_T^{c\ AN} = \sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1} \vartheta_i^{AN} 1_{\{A_i \cap A_{i+k_n}\}}, \text{ and} \tag{A.8}$$

$$[H(\widehat{C), G(C)}]_T^{c\ LIN} = \sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1} \vartheta_i^{LIN} 1_{\{A_i \cap A_{i+k_n}\}} \tag{A.9}$$

Let $\vartheta_i$ be either $\vartheta_i^{LIN}$ or $\vartheta_i^{AN}$.

# B Auxiliary Lemmas and Theorems

This section presents useful auxiliary results, which are used in the proofs of Theorems 1 and 2. The results of this section are proved in Section E below.

First, we explain why we can assume, without loss of generality, that the derivatives of functions $H_r$ and $G_r$ are bounded, for $r = 1, \ldots, \kappa$. Assumptions of Theorem 1 imply Lemma 2 of Li, Todorov, and Tauchen (2017a). Therefore, we can assume that the variables $\widehat{C}_{i\Delta_n}$ are bounded, uniformly over $i \in \{0, ..., [T/\Delta_n] - k_n + 1\}$, with probability approaching one. Using the spatial localization argument of Li, Todorov, and Tauchen (2016), which in turn uses the spatial localization argument of Li, Todorov, and Tauchen (2017a), we can assume that $H_r$ and $G_r$ are compactly supported without loss of generality. Hence, the derivatives of functions $H_r$ and $G_r$ are bounded, for $r = 1, \ldots, \kappa$.

We start with two auxiliary theorems for volatility jump truncation.

**Theorem B1.** *Under the assumptions of Theorem 1, we have*

$$\sum_{i \in L(n,T)} \vartheta_i 1_{\{A_i \cap A_{i+k_n}\}} - \sum_{i \in L(n,T)} \vartheta_i = o_p\left(\Delta_n^{1/4}\right).$$

**Theorem B2.** *Under the assumptions of Theorem 1, we have*

$$\sum_{i \in \overline{L}(n,T)} \vartheta_i 1_{\{A_i \cap A_{i+k_n}\}} = o_p\left(\Delta_n^{1/4}\right).$$

Theorems B1 and B2 allow us to focus on the simpler leading term $\sum_{i \in L(n,T)} \vartheta_i$ instead of the original estimator(s) $\sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1} \vartheta_i 1_{\{A_i \cap A_{i+k_n}\}}$ for the remaining proofs. Our next theorem shows negligibility of price jump truncation.

**Theorem B3.** *Let $\vartheta_i'^{LIN}$ and $\vartheta_i'^{AN}$ be the modifications of $\vartheta_i^{LIN}$ and $\vartheta_i^{AN}$ obtained by replacing $\widehat{C}_i^n$ by $\widehat{C}_i'^n$ in the definition of $\vartheta_i^{LIN}$ and $\vartheta_i^{AN}$ in equations (A.9) and (A.8). Under the assumptions of Theorem 1, we have*

$$\Delta_n^{-1/4}\Big( \sum_{i \in L(n,T)} \vartheta_i^{LIN} - \sum_{i \in L(n,T)} \vartheta_i'^{LIN} \Big) \xrightarrow{\mathbb{P}} 0$$

$$\text{and} \quad \Delta_n^{-1/4}\Big( \sum_{i \in L(n,T)} \vartheta_i^{AN} - \sum_{i \in L(n,T)} \vartheta_i'^{AN} \Big) \xrightarrow{\mathbb{P}} 0. \tag{B.10}$$

Theorem B3 allows, in particular, to focus on the derivation of the asymptotic distributions of $\sum_{i \in L(n,T)} \vartheta_i'^{LIN}$ and $\sum_{i \in L(n,T)} \vartheta_i'^{AN}$. The next theorem connects the *LIN* and *AN* versions of these quantities. To state the theorem, define

$$\vartheta_i^{(A)} = \frac{3}{2k_n} \sum_{g,h,a,b=1}^{d} \Bigg( \Big(\partial_{gh} H \partial_{ab} G\Big)(C_{(i-1)\Delta_n}) \Big[ (\widehat{C}_{i+k_n}'^{n,gh} - \widehat{C}_i'^{n,gh})(\widehat{C}_{i+k_n}'^{n,ab} - \widehat{C}_i'^{n,ab}) \tag{B.11}$$

$$- \frac{2}{k_n}(\widehat{C}_i'^{n,ga}\widehat{C}_i'^{n,hb} + \widehat{C}_i'^{n,gb}\widehat{C}_i'^{n,ha})\Big]\Bigg).$$

where superscript $(A)$ stands for "approximated". For simplicity, we do not index the above quantity by a prime although it depends on $\widehat{C}_i'^n$ instead of $\widehat{C}_i^n$.

**Theorem B4.** *Under the assumptions of Theorem 1, we have*

$$\Delta_n^{-1/4}\Big(\sum_{i\in L(n,T)}\vartheta_i'^{LIN} - \sum_{i\in L(n,T)}\vartheta_i^{(A)}\Big) \xrightarrow{\mathbb{P}} 0 \quad and$$
$$\Delta_n^{-1/4}\Big(\sum_{i\in L(n,T)}\vartheta_i'^{AN} - \sum_{i\in L(n,T)}\vartheta_i^{(A)}\Big) \xrightarrow{\mathbb{P}} 0, \tag{B.12}$$

*where $\vartheta_i^{(A)}$ is defined in equation (B.11).*

Theorem B4 shows that the leading terms of the the two estimators of $[H(\widehat{C),G}(C)]_T^c$, $\sum_{i\in L(n,T)}\vartheta_i'^{LIN}$ and $\sum_{i\in L(n,T)}\vartheta_i'^{AN}$ 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

$$\vartheta_i^{(A)} = \vartheta_i^{(A1)} - \vartheta_i^{(A2)}, \tag{B.13}$$

with

$$\vartheta_i^{(A1)} = \frac{3}{2k_n}\sum_{g,h,a,b=1}^{d}\big(\partial_{gh}H\partial_{ab}G\big)(C_{i-1}^n)(\widehat{C}_{i+k_n}'^{n,gh} - \widehat{C}_i'^{n,gh})(\widehat{C}_{i+k_n}'^{n,ab} - \widehat{C}_i'^{n,ab}),$$

and

$$\vartheta_i^{(A2)} = \frac{3}{k_n^2}\sum_{g,h,a,b=1}^{d}\big(\partial_{gh}H\partial_{ab}G\big)(C_{i-1}^n)(\widehat{C}_i'^{n,ga}\widehat{C}_i'^{n,hb} + \widehat{C}_i'^{n,gb}\widehat{C}_i'^{n,ha}).$$

The following theorem holds:

**Theorem B5.** *Under the assumptions of Theorem 1, we have*

$$\frac{1}{\Delta_n^{1/4}}\Bigg(\sum_{i\in L(n,T)}\vartheta_i^{(A1)} - \sum_{g,h,a,b=1}^{d}\sum_{u,v=1}^{2}\overline{A11}(H,gh,u;G,ab,v)_T^n + \overline{A12}(H,gh,u;G,ab,v)_T^n + \overline{A12}(G,ab,v;H,gh,u)_T^n\Bigg) \overset{\mathbb{P}}{\Longrightarrow} 0.$$

**Lemma B1.** *For any càdlàg bounded process $Z$, for all $t,s>0$, $j,k\geq 0$, set $\eta_{t,s}=\eta_{t,s}(Z)$. Then,*

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

**Lemma B2.** *Let $Z$ be a continuous Itô process with drift $b_t^Z$ and spot variance process $C_t^Z$, and set $\eta_{t,s}=\eta_{t,s}(b^Z,c^Z)$. Then, the following bounds hold:*

$$\Big|\mathbb{E}(Z_t\Big|\mathcal{F}_0) - tb_0^Z\Big| \leq Kt\eta_{0,t}$$

$$
\begin{aligned}
&\left|\mathbb{E}(Z_t^j Z_t^k - tC_0^{Z,jk}\Big|\mathcal{F}_0)\right| \leq Kt^{3/2}(\sqrt{\Delta_n}+\eta_{0,t}) \\
&\left|\mathbb{E}\big((Z_t^j Z_t^k - tC_0^{Z,jk})(C_t^{Z,lm}-C_0^{Z,lm})\Big|\mathcal{F}_0\big)\right| \leq Kt^2 \\
&\left|\mathbb{E}(Z_t^j Z_t^k Z_t^l Z_t^m\Big|\mathcal{F}_0) - \Delta_n^2(C_0^{Z,jk}C_0^{Z,lm}+C_0^{Z,jl}C_0^{Z,km}+C_0^{Z,jm}C_0^{Z,kl})\right| \leq Kt^{5/2} \\
&\left|\mathbb{E}(Z_t^j Z_t^k Z_t^l\Big|\mathcal{F}_0)\right| \leq Kt^2 \\
&\left|\mathbb{E}(\prod_{l=1}^{6} Z_t^{j_l}\Big|\mathcal{F}_0) - \frac{\Delta_n^3}{6}\sum_{l<l'}\sum_{k<k'}\sum_{m<m'} C_0^{Z,j_l j_{l'}}C_0^{Z,j_k j_{k'}}C_0^{Z,j_m j_{m'}}\right| \leq Kt^{7/2} \\
&\mathbb{E}\Big(\sup_{w\in[0,s]}\Big\|Z_{t+w}-Z_t\Big\|^q\Big|\mathcal{F}_t\Big) \leq K_q s^{q/2}, \text{ and } \Big\|\mathbb{E}\Big(Z_{t+s}-Z_t\Big)\Big|\mathcal{F}_t\Big\| \leq Ks. \qquad \text{(B.14)}
\end{aligned}
$$

(B.15)

**Lemma B3.** *Let $\zeta_i^n$ be a $r$-dimensional $\mathcal{F}_i^n$-measurable process satisfying $\|\mathbb{E}(\zeta_i^n|\mathcal{F}_{i-1}^n)\| \leq L'$ and $\mathbb{E}\Big(\|\zeta_i^n\|^q\Big|\mathcal{F}_{i-1}^n\Big) \leq L_q$. Also, let $\varphi_i^n$ be a real-valued $\mathcal{F}_i^n$-measurable process with $\mathbb{E}\Big(\|\varphi_{i+j-1}^n\|^q\Big|\mathcal{F}_{i-1}^n\Big) \leq L^q$ for $q \geq 2$ and $1 \leq j \leq 2k_n - 1$. Then,*

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

**Lemma B4.** *Under the assumptions of Theorem 1, we have, for $i \in L(n,T)$:*

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

**Lemma B5.** *Under the assumptions of Theorem 1, we have, for $i \in L(n,T)$:*

$$
\left|\mathbb{E}\Big(\nu_i^{n,jk}\nu_i^{n,lm}\nu_i^{n,gh}\Big|\mathcal{F}_{i-1}^n\Big)\right| \leq K\Delta_n^{3/4}\Big(\Delta_n^{1/4}+\eta_{i,k_n}^n\Big), \qquad \text{(B.16)}
$$

$$
\left|\mathbb{E}\Big(\nu_i^{n,jk}\nu_i^{n,lm}\Big(C_{i+k_n-1}^{n,gh}-C_{i-1}^{n,gh}\Big)\Big|\mathcal{F}_{i-1}^n\Big)\right| \leq K\Delta_n^{3/4}\Big(\Delta_n^{1/4}+\eta_{i,k_n}^n\Big), \qquad \text{(B.17)}
$$

$$
\left|\mathbb{E}\Big(\nu_i^{n,jk}\Big(C_{i+k_n-1}^{n,lm}-C_{i-1}^{n,lm}\Big)\Big(C_{i+k_n-1}^{n,gh}-C_{i-1}^{n,gh}\Big)\Big|\mathcal{F}_{i-1}^n\Big)\right| \leq K\Delta_n^{3/4}\Big(\Delta_n^{1/4}+\eta_{i,k_n}^n\Big), \qquad \text{(B.18)}
$$

$$
\left|\mathbb{E}\Big(\nu_i^{n,jk}\lambda_i^{n,lm}\lambda_i^{n,gh}\Big|\mathcal{F}_{i-1}^n\Big)\right| \leq K\Delta_n^{3/4}\Big(\Delta_n^{1/4}+\eta_{i,2k_n}^n\Big), \qquad \text{(B.19)}
$$

$$
\left|\mathbb{E}\Big(\lambda_i^{n,jk}\lambda_i^{n,lm}\lambda_i^{n,gh}\Big|\mathcal{F}_{i-1}^n\Big)\right| \leq K\Delta_n^{3/4}\Big(\Delta_n^{1/4}+\eta_{i,2k_n}^n\Big). \qquad \text{(B.20)}
$$

**Lemma B6.** *Under the assumptions of Theorem 1, we have:*

$$
\frac{1}{\Delta_n^{1/4}}\sum_{i\in L(n,T)}(\partial_{gh}H\partial_{ab}G)(C_{(i-2k_n-1)\Delta_n})\rho_{gh}(u,v)_i^n\zeta_{ab}'(v)_i^n \stackrel{\mathbb{P}}{\Longrightarrow} 0, \quad \forall \quad (u,v) \qquad \text{(B.21)}
$$

$$
\frac{1}{\Delta_n^{1/4}}\Big(\overline{A11}(H,gh,u;G,ab,v) - \int_0^T(\partial_{gh}H\partial_{ab}G)(C_t)\overline{C}_t^{gh,ab}dt\Big) \stackrel{\mathbb{P}}{\Longrightarrow} 0 \ \textit{ when } \ (u,v)=(2,2) \qquad \text{(B.22)}
$$

$$
\frac{1}{\Delta_n^{1/4}}\Big(\overline{A11}(H,gh,u;G,ab,v) - \frac{3}{\theta^2}\int_0^T(\partial_{gh}H\partial_{ab}G)(C_t)(C_t^{ga}C_t^{hb}+C_t^{gb}C_t^{ha})dt\Big) \stackrel{\mathbb{P}}{\Longrightarrow} 0 \qquad \text{(B.23)}
$$

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

$$\frac{1}{\Delta_n^{1/4}}\overline{A11}(H,gh,u;G,ab,v) \stackrel{\mathbb{P}}{\Longrightarrow} 0 \quad when \quad (u,v)=(1,2),(2,1) \tag{B.24}$$

# C Proof of Theorem 1

We now prove Theorem 1. By Theorem B5, we have

$$\frac{1}{\Delta_n^{1/4}}\Bigg(\sum_{i\in L(n,T)}\vartheta_i^{(A1)} - \sum_{g,h,a,b=1}^{d}\sum_{u,v=1}^{2}\overline{A11}(H,gh,u;G,ab,v)_T^n + \overline{A12}(H,gh,u;G,ab,v)_T^n + \overline{A12}(G,ab,v;H,gh,u)_T^n\Bigg) \stackrel{\mathbb{P}}{\Longrightarrow} 0.$$

Recalling the definition of $\overline{A12}(H,gh,u;G,ab,v)_T^n$ from equation (A.7), Lemma B6 implies that

$$\frac{1}{\Delta_n^{1/4}}\Bigg(\sum_{i\in L(n,T)}\vartheta_i^{(A)} - [H(C),G(C)]_T - \frac{3}{2k_n^3}\sum_{g,h,a,b}^{d}\sum_{u,v=1}^{2}\sum_{i\in L'(n,T)} \Big[(\partial_{gh}H\partial_{ab}G)(C_{(i-2k_n-1)\Delta_n})\rho_{gh}(u,v)_i^n\zeta''_{ab}(v)_i^n + (\partial_{ab}H\partial_{gh}G)(C_{(i-2k_n-1)\Delta_n})\rho_{ab}(v,u)_i^n\zeta''_{gh}(v)_i^n\Big]\Bigg) \stackrel{\mathbb{P}}{\Longrightarrow} 0. \tag{C.25}$$

Next, define

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

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

Notice that (C.25) implies

$$\frac{1}{\Delta_n^{1/4}}\Big(\sum_{i\in L(n,T)}\vartheta_i^{(A)} - [H(C),G(C)]_T\Big) \stackrel{\mathcal{L}}{=} \sum_{g,h,a,b=1}^{d}\sum_{u,v=1}^{2}\frac{1}{\Delta_n^{1/4}}\Big(Z(H,gh,u;G,ab,v)_T^n + Z(H,ab,v;G,gh,u)_T^n\Big). \tag{C.26}$$

The term $\vartheta_i^{(A)}$ depends on functions $H$ and $G$, where we have so far suppressed the subscripts $r$, $r=1,..,\kappa$, in the statement of Theorem 1 for simplicity. Denote by $\vartheta_{i,r}^{(A)}$ the term $\vartheta_i^{(A)}$ that depends on functions $H_r$ and $G_r$. Observe that to derive the asymptotic distribution of $\left(\sum_{i\in L(n,T)}\vartheta_{i,1}^{(A)},...,\sum_{i\in L(n,T)}\vartheta_{i,\kappa}^{(A)}\right)$, 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 $\frac{1}{\Delta_n^{1/4}}Z(H,gh,u;G,ab,v)_T^n$, it is enough to prove the following three properties:

$$A\Big((H,gh,u;G,ab,v),(H',g'h',u';G',a'b',v')\Big)_t^n$$

$$= \sum_{i\in L'(n,T)}\mathbb{E}(\xi(H,gh,u;G,ab,v)_i^n\xi(H',g'h',u';G',a'b',v')_i^n|\mathcal{F}_{i-1}^n) \tag{C.27}$$

$$\stackrel{\mathbb{P}}{\Longrightarrow} A\Big((H,gh,u;G,ab,v),(H',g'h',u';G',a'b',v')\Big)_t, \tag{C.28}$$

$$\sum_{i\in L'(n,T)} \mathbb{E}\Big(\Big|\xi(H,gh,u;G,ab,v)_i^n\Big|^4\Big|\mathcal{F}_{i-1}^n\big) \stackrel{\mathbb{P}}{\Longrightarrow} 0, \text{ and} \tag{C.29}$$

$$B(N;H,gh,u;G,ab,v)_t^n := \sum_{i\in L'(n,T)} \mathbb{E}\Big(\xi(H,gh,u;G,ab,v)_i^n \Delta_i^n N|\mathcal{F}_{i-1}^n\Big) \stackrel{\mathbb{P}}{\Longrightarrow} 0, \tag{C.30}$$

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 $W^j$.
Since the derivatives of $H_r$ and $G_r$ are bounded, equations (C.29) and (C.30) can be proved by an extension of (B.105) and (B.106) in Aït-Sahalia and Jacod (2014) to multivariate processes.
Next, define

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

Using again the boundedness of the derivatives of $H_r$ and $G_r$, we can show that

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

with

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

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

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

Using equation (C.26), we deduce that the asymptotic covariance between $\sum_{i\in L(n,T)}\vartheta_{i,r}^{(A)}$ and $\sum_{i\in L(n,T)}\vartheta_{i,s}^{(A)}$ is given by

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

The above expression can be rewritten as

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

which completes the proof.

# D Proof of Theorem 2

Recall that $N_s$ is the number of jumps in $C$ from time 0 to $s$. Let

$$\begin{aligned} L''(n) &= \left\{i=k_n+1,k_n+2,\ldots: N_{(i+5)k_n\Delta_n}-N_{(i-1)k_n\Delta_n}=0\right\},\\ L''(n,T) &= \{i=1,2,\ldots,[T/\Delta_n]-5k_n+1\}\cap L''(n),\\ \overline{L}''(n,T) &= \{i=1,2,\ldots,[T/\Delta_n]-5k_n+1\}\backslash L''(n). \end{aligned}$$

Denote by $\widehat{\omega}_T^{r,s,(1)}$, $\widehat{\omega}_T^{r,s,(2)}$, and $\widehat{\omega}_T^{r,s,(3)}$ the $i^{th}$ summand of $\widehat{\Omega}_T^{r,s,(1)}$, $\widehat{\Omega}_T^{r,s,(2)}$, and $\widehat{\Omega}_T^{r,s,(3)}$, without the volatility jump truncation, so they satisfy

$$\widehat{\Omega}_T^{r,s,(m)}=\sum_{i=k_n+1}^{[T/\Delta_n]-5k_n+1}\widehat{\omega}_T^{r,s,(m)}1_{\{A_i\cap A_{i+k_n}\cap A_{i+2k_n}\cap A_{i+3k_n}\}}\text{ for } m=1,2,\text{ and }3.$$

The same methods as in Theorems B1 and B2 can be used to show

$$\begin{aligned} \sum_{i\in L''(n,T)}\widehat{\omega}_T^{r,s,(m)}1_{\{A_i\cap A_{i+k_n}\cap A_{i+2k_n}\cap A_{i+3k_n}\}}-\sum_{i\in L''(n,T)}\widehat{\omega}_T^{r,s,(m)} &= o_p(1)\text{ and}\\ \sum_{i\in\overline{L}''(n,T)}\widehat{\omega}_T^{r,s,(m)}1_{\{A_i\cap A_{i+k_n}\cap A_{i+2k_n}\cap A_{i+3k_n}\}} &= o_p(1). \end{aligned}$$

We conclude that the probability limit of $\widehat{\Omega}_T^{r,s,(m)}$ is the same as $\sum_{i\in L''(n,T)}\widehat{\omega}_T^{r,s,(m)}$ for $m=1,2,3$.

Using boundedness of the derivatives of $H_r, G_r, H_s$ and $G_s$ and Theorem 2.2 in Jacod and Rosenbaum (2015), one can show that

$$\frac{6}{\theta^3}\sum_{i\in L''(n,T)}\widehat{\omega}_T^{r,s,(1)}\xrightarrow{\mathbb{P}}\Sigma_T^{r,s,(1)}.$$

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

$$\frac{3}{2\theta}\left(\sum_{i\in L''(n,T)}\widehat{\omega}_T^{r,s,(3)}-\frac{6}{\theta}\sum_{i\in L''(n,T)}\widehat{\omega}_T^{r,s,(1)}\right)\xrightarrow{\mathbb{P}}\Sigma_T^{r,s,(3)}.$$

Finally, to show that

$$\frac{151\theta}{140}\frac{9}{4\theta^2}\left(\sum_{i\in L''(n,T)}\widehat{\omega}_T^{r,s,(2)}+\frac{4}{\theta^2}\sum_{i\in L''(n,T)}\widehat{\omega}_T^{r,s,(1)}-\frac{4}{3}\sum_{i\in L''(n,T)}\widehat{\omega}_T^{r,s,(3)}\right)\xrightarrow{\mathbb{P}}\Sigma_T^{r,s,(2)},$$

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

$$\begin{aligned}
\widehat{W}_T^n &= \sum_{i\in L''(n,T)}(\partial_{gh}H_r\partial_{ab}G_r\partial_{gh}H_s\partial_{lm}G_s)(\widehat{C}_i^n)\lambda_i^{n,gh}\lambda_i^{n,jk}\lambda_{i+2k_n}^{n,ab}\lambda_{i+2k_n}^{n,lm},\\
\widehat{w}(1)_i^n &= (\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(C_{i-1}^n)\mathbb{E}(\lambda_i^{n,gh}\lambda_i^{n,jk}\lambda_{i+2k_n}^{n,ab}\lambda_{i+2k_n}^{n,lm}|\mathcal{F}_i^n),\\
\widehat{w}(2)_i^n &= (\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(C_{i-1}^n)(\lambda_i^{n,gh}\lambda_i^{n,jk}\lambda_{i+2k_n}^{n,ab}\lambda_{i+2k_n}^{n,lm}-\mathbb{E}(\lambda_i^{n,gh}\lambda_i^{n,jk}\lambda_{i+2k_n}^{n,ab}\lambda_{i+2k_n}^{n,lm}|\mathcal{F}_i^n)),\\
\widehat{w}(3)_i^n &= \Big((\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(\widehat{C}_i^n)-(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(C_{i-1}^n)\Big)\lambda_i^{n,gh}\lambda_i^{n,jk}\lambda_{i+2k_n}^{n,ab}\lambda_{i+2k_n}^{n,lm},\\
\widehat{W}(u)_t^n &= \sum_{i\in L''(n,T)}\widehat{w}_i(u),\ u=1,2,3.
\end{aligned}$$

Now, note that we also have $\widehat{W}_t^n=\widehat{W}(1)_t^n+\widehat{W}(2)_t^n+\widehat{W}(3)_t^n$. By Taylor expansion and using repeatedly the boundedness of $C_t$, we obtain, for $i\in L''(n,T)$

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

which implies $\mathbb{E}(|\widehat{w}(3)_i^n|)\le K\Delta_n^{5/4}$ and hence $\widehat{W}(3)_t^n\xrightarrow{\mathbb{P}}0$. Using Cauchy-Schwartz inequality and the bound $\mathbb{E}(\|\lambda_i^n\|^q|\mathcal{F}_i^n)\le K\Delta_n^{q/4}$, we have $\mathbb{E}(|\widehat{w}(2)_i^n|^2)\le K\Delta_n^2$ for $i\in L''(n,T)$. Observing furthermore that $\widehat{w}(2)_i^n$ is $\mathcal{F}_{i+4k_n}-$measurable, Lemma B.8 in Aït-Sahalia and Jacod (2014) implies $\widehat{W}(2)_t^n\xrightarrow{\mathbb{P}}0$.
Next, define

$$\begin{aligned}
w_i^n &= (\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(C_{i-1}^n)\Big[\frac{4}{k_n^2\Delta_n}(C_{i-1}^{n,ga}C_{i-1}^{n,hb}+C_{i-1}^{n,gb}C_{i-1}^{n,ha})(C_{i-1}^{n,jl}C_{i-1}^{n,km}+C_{i-1}^{n,jm}C_{i-1}^{n,kl})\\
&+\frac{4}{3}(C_{i-1}^{n,jl}C_{i-1}^{n,km}+C_{i-1}^{n,jm}C_{i-1}^{n,kl})\overline{C}_{i-1}^{n,gh,ab}+\frac{4}{3}(C_{i-1}^{n,ga}C_{i-1}^{n,hb}+C_{i-1}^{n,gb}C_{i-1}^{n,ha})\overline{C}_{i-1}^{n,jk,lm}\\
&+\frac{4(k_n^2\Delta_n)}{9}\overline{C}_{i-1}^{n,gh,ab}\overline{C}_{i-1}^{n,jk,lm}\Big],\\
W_T^n &= \Delta_n\sum_{i\in L''(n,T)}w_i^n.
\end{aligned}$$

Using the cadlag property of $c$ and $\overline{C}$, $k_n\sqrt{\Delta_n}\to\theta$, and the Riemann integral convergence, we conclude that $W_T^n\xrightarrow{\mathbb{P}}W_T$ where

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

In addition, by Lemma B4, it holds that

$$\mathbb{E}(|\widehat{W}(1)_T^n-W_T^n|)\le\Delta_n\mathbb{E}\left(\sum_{i\in L''(n,T)}(\Delta_n^{1/8}+\eta_{i,4k_n})\right).$$

Hence, by the third result of Lemma B1 we have $\widehat{W}_T^n \xrightarrow{\mathbb{P}} W_t$, from which it follows that

$$\frac{9}{4\theta^2}\Big[\widehat{W}(1)_T^n + \frac{4}{k_n^2}\sum_{i\in L''(n,T)}(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(\widehat{C}_i^n)[C_i^n(jk,lm)C_i^n(gh,ab)]$$
$$-\frac{2}{k_n}\sum_{i\in L''(n,T)}(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(\widehat{C}_i^n)C_i^n(gh,ab)\lambda_i^{n,jk}\lambda_i^{n,lm}$$
$$-\frac{2}{k_n}\sum_{i\in L''(n,T)}(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(\widehat{C}_i^n)C_i^n(jk,lm)\lambda_i^{n,gh}\lambda_i^{n,ab}\Big]$$
$$\xrightarrow{\mathbb{P}} \int_0^T(\partial_{gh}H_r\partial_{ab}G_r\partial_{jk}H_s\partial_{lm}G_s)(C_t)\overline{C}_t^{gh,ab}\overline{C}_t^{jk,lm}dt.$$

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

# E Proofs of Auxiliary Lemmas and Theorems

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

## E.1 Proof of Theorem B1

The proof proceeds in three steps. In Step 1, we prove, for $i \in L(n,T)$,

$$P\left(\overline{A}_i\right) \leq K a_n \Delta_n^{(2-r)\varpi-\varpi'}, \tag{E.31}$$

where $a_n$ is a sequence converging to zero, and $\overline{A}_i$ is the complement of $A_i$. In Step 2, we prove, for $p \geq 1$ and $i \in L(n,T)$,

$$E\left[|\vartheta_i|^p\right] \leq K\Delta_n^p + K a_n \Delta_n^{(4p-r)\varpi+1-\frac{3}{2}p}. \tag{E.32}$$

Step 3 completes the proof of Theorem B1.

*Step 1.* We now prove equation (E.31). Recall $\widehat{C}_i'^n$ notation in (A.2). For $i \in L(n,T)$,

$$\begin{aligned} P\left(\overline{A}_i\right) &= P\left(\left\|\widehat{C}_{i+k_n}^n - \widehat{C}_{i-k_n}^n\right\| \geq u_n'\right) \\ &\leq P\left(\left\|\widehat{C}_{i+k_n}'^n - \widehat{C}_{i-k_n}'^n\right\| + \left\|\widehat{C}_{i+k_n}^n - \widehat{C}_{i+k_n}'^n\right\| + \left\|\widehat{C}_{i-k_n}^n - \widehat{C}_{i-k_n}'^n\right\| \geq u_n'\right) \\ &\leq P\left(\left\|\widehat{C}_{i+k_n}'^n - \widehat{C}_{i-k_n}'^n\right\| \geq \frac{u_n'}{2}\right) + P\left(\left\|\widehat{C}_{i+k_n}^n - \widehat{C}_{i+k_n}'^n\right\| + \left\|\widehat{C}_{i-k_n}^n - \widehat{C}_{i-k_n}'^n\right\| \geq \frac{u_n'}{2}\right). \end{aligned} \tag{E.33}$$

Using standard results in the literature, we have for $q \geq 2$ and $i \in L(n,T)$,

$$E\left(\left\|\widehat{C}_{i+k_n}'^n - \widehat{C}_{i-k_n}'^n\right\|^q\right) \leq K\Delta_n^{q/4}, \tag{E.34}$$

see, for example, equation (3.26) in Jacod and Rosenbaum (2015). Therefore, the first term in (E.33) satisfies, by Markov's inequality, for $p \geq 2$,

$$P\left(\left\|\widehat{C}_{i+k_n}'^n - \widehat{C}_{i-k_n}'^n\right\| \geq \frac{u_n'}{2}\right) \leq K\Delta_n^{p/4-\varpi' p}. \tag{E.35}$$

By (4.8) in Jacod and Rosenbaum (2013), there exists a sequence of real numbers $a_n$ converging to zero such that

$$\mathbb{E}(\|\widehat{C}_i^n - \widehat{C}_i'^n\|^q) \leq K_q a_n \Delta_n^{(2q-r)\varpi+1-q}, \text{ for any } q \geq 1, \tag{E.36}$$

where for later use, we note that this result also holds in the presence of volatility jumps. Therefore, the second term in (E.33) satisfies, by Markov's inequality,

$$\begin{aligned} & P\left(\left\|\widehat{C}^n_{i+k_n}-\widehat{C}'^n_{i+k_n}\right\|+\left\|\widehat{C}^n_{i-k_n}-\widehat{C}'^n_{i-k_n}\right\|\geq\frac{u'_n}{2}\right) \\ \leq & \frac{1}{u'_n/2}E\left(\left\|\widehat{C}^n_{i+k_n}-\widehat{C}'^n_{i+k_n}\right\|+\left\|\widehat{C}^n_{i-k_n}-\widehat{C}'^n_{i-k_n}\right\|\right)\leq Ka_n\Delta_n^{(2-r)\varpi-\varpi'}. \end{aligned} \tag{E.37}$$

Since $\varpi' < \frac{1}{8}$ and by choosing sufficiently large $p$ in (E.35), equations (E.35) and (E.37) give (E.31).

*Step 2.* We now prove equation (E.32). First, note that for $q \geq 1$, by (E.36),

$$\mathrm{E}\left(\left\|\widehat{C}^n_{i\Delta_n}\right\|^q\right)\leq K\mathrm{E}\left[\left|\widehat{C}^n_{i\Delta_n}-\widehat{C}'^n_{i\Delta_n}\right|^q\right]+K\mathrm{E}\left[\left|\widehat{C}'^n_{i\Delta_n}\right|^q\right]\leq Ka_n\Delta_n^{(2q-r)\varpi+1-q}+K. \tag{E.38}$$

By Taylor expansion and $H$ and $G$ having bounded derivatives, for $i \in L(n,T)$ and $p \geq 1$,

$$\begin{aligned} & E\left[|\vartheta_i|^p\right] \\ \leq & K\frac{1}{k_n^p}\mathrm{E}\left(\left\|\widehat{C}^n_{(i+k_n)\Delta_n}-\widehat{C}^n_{i\Delta_n}\right\|^{2p}\right)+K\frac{1}{k_n^{2p}}\mathrm{E}\left(\left\|\widehat{C}^n_{i\Delta_n}\right\|^{2p}\right) \\ \leq & K\Delta_n^{p/2}\mathrm{E}\left(\left\|\widehat{C}'^n_{(i+k_n)\Delta_n}-\widehat{C}'^n_{i\Delta_n}\right\|^{2p}+\left\|\widehat{C}^n_{(i+k_n)\Delta_n}-\widehat{C}'^n_{(i+k_n)\Delta_n}\right\|^{2p}+\left\|\widehat{C}^n_{i\Delta_n}-\widehat{C}^{n\prime}_{i\Delta_n}\right\|^{2p}\right)+K\Delta_n^p\mathrm{E}\left(\left\|\widehat{C}^n_{i\Delta_n}\right\|^{2p}\right) \\ \leq & K\Delta_n^{p/2}\left(\Delta_n^{p/2}+a_n\Delta_n^{(4p-r)\varpi+1-2p}\right)+K\Delta_n^p\left[Ka_n\Delta_n^{(4p-r)\varpi+1-2p}+K\right] \\ = & K\Delta_n^p+Ka_n\Delta_n^{(4p-r)\varpi+1-\frac{3}{2}p}, \end{aligned} \tag{E.39}$$

where the third inequality uses (E.38), (E.34) and (E.36).

*Step 3.* We now complete the proof of Theorem B1. By the triangle and Cauchy-Schwarz inequalities,

$$\begin{aligned} & E\left|\sum_{i\in L(n,T)}\vartheta_i 1_{\{A_i\cap A_{i+k_n}\}}-\sum_{i\in L(n,T)}\vartheta_i\right| \\ \leq & \sum_{i\in L(n,T)}E\left|\vartheta_i\left(1_{\{A_i\cap A_{i+k_n}\}}-1\right)\right| \\ \leq & \sum_{i\in L(n,T)}\sqrt{E\left|\vartheta_i^2\right|}\sqrt{P\left(\overline{A_i}\cup\overline{A_{i+k_n}}\right)} \\ \leq & \sum_{i\in L(n,T)}\sqrt{E\left|\vartheta_i^2\right|}\sqrt{P\left(\overline{A_i}\right)+P\left(\overline{A_{i+k_n}}\right)} \\ \leq & K\Delta_n^{-1}\left(\Delta_n^{(8-r)\varpi-2}\right)^{1/2}\left(a_n\Delta_n^{(2-r)\varpi-\varpi'}\right)^{1/2} \\ = & \Delta_n^{l(\varpi,\varpi')}, \end{aligned}$$

where 4th inequality follows by (E.32) with $p = 2$, and (E.31). In the above,

$$l(\varpi,\varpi')=-1+\frac{1}{2}\left[(8-r)\varpi-2\right]+\frac{1}{2}\left[(2-r)\varpi-\varpi'\right].$$

A straightforward calculation shows that $\varpi > \frac{2\varpi'+9}{4(5-r)}$ implies $l(\varpi,\varpi') > \frac{1}{4}$, which completes the proof of Theorem B1.

## E.2 Proof of Theorem B2

Without loss of generality, we can assume that there is at most one volatility jump in $((i-k_n)\Delta_n, (i+3k_n)\Delta_n]$ for any $i \in \overline{L}(n,T)$. To study the behavior of $\vartheta_i 1_{\{A_i \cap A_{i+k_n}\}}$ on $i \in \overline{L}(n,T)$, we will distinguish between two cases, depending on whether or not there is a volatility jump in $(i\Delta_n, (i+2k_n)\Delta_n]$. So define $B_i$ as the event that there is a volatility jump in $(i\Delta_n, (i+2k_n)\Delta_n]$ (we omit indexing $B_i$ by $n$ for brevity). Denote by $\overline{B}_i$ the complement of $B_i$. Intuitively, for $i \in \overline{L}(n,T)$, $\vartheta_i 1_{\{A_i \cap A_{i+k_n}\}}$ is small because, on the one hand, $P(A_i \cap A_{i+k_n})$ is small on $B_i$, on the other hand, $\vartheta_i$ is small on $\overline{B}_i$.

We have

$$\begin{aligned} \mathrm{E}\left|\sum_{i\in\overline{L}(n,T)} \vartheta_i 1_{\{A_i\cap A_{i+k_n}\}}\right| &= \mathrm{E}\left|\sum_{i\in\overline{L}(n,T):B_i} \vartheta_i 1_{\{A_i\cap A_{i+k_n}\}} + \sum_{i\in\overline{L}(n,T):\overline{B}_i} \vartheta_i 1_{\{A_i\cap A_{i+k_n}\}}\right| \\ \leq \quad & \sum_{i\in\overline{L}(n,T):B_i} \mathrm{E}\left|\vartheta_i 1_{\{A_i\cap A_{i+k_n}\}}\right| + \sum_{i\in\overline{L}(n,T):\overline{B}_i} \mathrm{E}\left|\vartheta_i 1_{\{A_i\cap A_{i+k_n}\}}\right|, \end{aligned} \tag{E.40}$$

where "$i \in \overline{L}(n,T) : B_i$" denotes those terms in $\overline{L}(n,T)$, for which $B_i$ is true.

First, we show that the second term in (E.40) is $o_p\left(\Delta_n^{1/4}\right)$. For $i \in \overline{L}(n,T)$ such that $B_i$ if false, we can use the bound on $\mathrm{E}\left[|\vartheta_i|^p\right]$ in (E.39) for $p \geq 1$. The second term in (E.40) satisfies

$$\begin{aligned} & \sum_{i\in\overline{L}(n,T):\overline{B}_i} \mathrm{E}\left|\vartheta_i 1_{\{A_i\cap A_{i+k_n}\}}\right| \\ \leq \quad & \sum_{i\in\overline{L}(n,T):\overline{B}_i} \mathrm{E}\left|\vartheta_i\right| \\ \leq \quad & K\Delta_n^{-1/2}\left(\Delta_n + \Delta_n^{(4-r)\varpi-\frac{1}{2}}\right) \\ = \quad & K\Delta_n^{1/2} + K\Delta_n^{(4-r)\varpi-1}. \end{aligned}$$

Theorem 1 assumptions imply $(4-r)\varpi - 1 > \frac{1}{4}$, so the second term in (E.40) is $o_p\left(\Delta_n^{1/4}\right)$.

The rest of the proof is devoted to showing that the first term in (E.40) is $o_p\left(\Delta_n^{1/4}\right)$. This will complete the proof of Theorem B2.

The first term in (E.40) involves those $i \in \overline{L}(n,T)$, for which $B_i$ is true. We will show below that

$$P(A_i \cap A_{i+k_n}) \leq K\Delta_n^{1/2} \text{ for } i \in \overline{L}(n,T) \text{ such that } B_i \text{ holds.} \tag{E.41}$$

We use the following bound in the presence of the volatility jump,

$$\mathrm{E}\left(|\vartheta_i|^p\right) \leq K\frac{1}{k_n^p}\left[\mathrm{E}\left[\left|\widehat{C}_i^n\right|^p\right] + \mathrm{E}\left[\left|\widehat{C}_{i+k_n}^n\right|^p\right]\right] \leq K\frac{1}{k_n^p}\left(a_n\Delta_n^{(2p-r)\varpi+1-p} + K\right), \tag{E.42}$$

where the first inequality uses Taylor expansion and bounded derivatives of $H$ and $G$, and the last transition uses (E.38).

The first term in (E.40) satisfies, for $p \geq 1$, by Holder inequality, (E.41) and (E.42),

$$\begin{aligned} & \sum_{i\in\overline{L}(n,T):B_i} \mathrm{E}\left|\vartheta_i 1_{\{A_i\cap A_{i+k_n}\}}\right| \\ \leq \quad & \sum_{i\in\overline{L}(n,T):B_i} \left(\mathrm{E}\left[|\vartheta_i|^p\right]\right)^{1/p}\left(\mathrm{P}\left(A_i\cap A_{i+k_n}\right)\right)^{(p-1)/p} \end{aligned}$$

$$
\begin{aligned}
&\le \sum_{i\in\overline{L}(n,T):B_i} K\left[\Delta_n \times \left(\Delta_n^{(2p-r)\varpi+1-p}+1\right)\right]^{1/p}\left[\Delta_n^{1/2}\right]^{(p-1)/p} \\
&= \sum_{i\in\overline{L}(n,T):B_i} K\Delta_n^{l(r,\varpi)}.
\end{aligned}
$$

Since the number of terms in $\overline{L}(n,T)$ is bounded by $Kk_n$ ($k_n$ arises due to overlapping blocks defining $\vartheta_i$), the first term in (E.40) is $o_p\left(\Delta_n^{1/4}\right)$ if $l(r,\varpi) > \frac{3}{4}$. To study $l(r,\varpi)$, we distinguish two cases, depending on whether $(2p-r)\varpi + 1 - p \ge 0$ holds.

Case 1. When $(2p-r)\varpi + 1 - p \ge 0$, $l(r,\varpi) = \frac{1}{2p}(p+1)$, so $l(r,\varpi) > \frac{3}{4}$ if $p < 2$.

Case 2. When $(2p-r)\varpi + 1 - p < 0$, $l(r,\varpi) = \frac{1}{p}((2p-r)\varpi + 1 - p) + \frac{p-1}{2p}$. We have $l(r,\varpi) > \frac{3}{4}$ if $\varpi > \frac{5p-6}{4(2p-r)}$. This is satisfied if we choose, for example, $p = 1.5$.

The last step in the proof of Theorem B2 is to show that (E.41) is true. In order to do that, we first prove that if there is a volatility jump on $(i\Delta_n, (i+k_n)\Delta_n]$, then

$$
P\left(\left\|\widehat{C}^n_{i+k_n} - \widehat{C}^n_{i-k_n}\right\| < u_n'\right) = o_p\left(\Delta_n^{1/4}\right). \tag{E.43}
$$

Denote by $S$ the time of the volatility jump on $(i\Delta_n, (i+k_n)\Delta_n]$, so the jump is $\Delta C_S$. Denote $\xi_n \equiv \widehat{C}^n_{i+k_n} - \widehat{C}^n_{i-k_n} - \Delta C_S$, so $\widehat{C}^n_{i+k_n} - \widehat{C}^n_{i-k_n} = \Delta C_S + \xi_n$. We know $\xi_n = o_p(1)$. We know that there exists $\epsilon$, independent of $i$ or $S$, such that $\|\Delta C\| > \epsilon$.

We will first show that if there is a volatility jump on $(i\Delta_n, (i+k_n)\Delta_n]$, for $s \ge 0$, it follows that

$$
P\left(\left\|\widehat{C}^n_{i+k_n} - \widehat{C}^n_{i-k_n}\right\| < u_n'\right) \le \frac{E\left(\left\|\widehat{C}^n_{i+k_n} - \widehat{C}^n_{i-k_n} - \Delta C_S\right\|^s\right)}{(\epsilon/2)^s}. \tag{E.44}
$$

To prove (E.44), note that the reverse triangle inequality gives $\left\|\widehat{C}^n_{i+k_n} - \widehat{C}^n_{i-k_n}\right\| = \|\Delta C + \xi_n\| \ge |\|\Delta C\| - \|\xi_n\||$. Thus,

$$
\begin{aligned}
& P\left(\left\|\widehat{C}^n_{i+k_n} - \widehat{C}^n_{i-k_n}\right\| < u_n'\right) \\
\le\ & P\left(|\|\Delta C\| - \|\xi_n\|| < u_n'\right) \\
\le\ & P\left(\|\xi_n\| > \frac{\epsilon}{2}\right),
\end{aligned}
$$

where the second inequality follows by distinguishing two cases, depending on whether $\|\Delta C\| \ge \|\xi_n\|$. Case 1: if $\|\Delta C\| \ge \|\xi_n\|$, $\{|\|\Delta C\| - \|\xi_n\|| < u_n'\} = \{\|\Delta C\| - \|\xi_n\| < u_n'\} = \{\|\Delta C\| - u_n' < \|\xi_n\|\}$, so we deduce $\{\epsilon - u_n' < \|\xi_n\|\}$. For $n$ large enough, this implies $\left\{\|\xi_n\| > \frac{\epsilon}{2}\right\}$ since $u_n' \to 0$. Case 2: if $\|\Delta C\| < \|\xi_n\|$, we have $P\left(\{|\|\Delta C\| - \|\xi_n\|| < u_n'\} \cap \{\|\Delta C\| < \|\xi_n\|\}\right) \le P\left(\|\xi_n\| > \|\Delta C\|\right) \le P\left(\|\xi_n\| > \epsilon\right) \le P\left(\|\xi_n\| > \frac{\epsilon}{2}\right)$. Finally, (E.44) follows by Markov's inequality.

By (E.44), we obtain, for $s \ge 2$,

$$
\begin{aligned}
P\left(\left\|\widehat{C}^n_{i+k_n} - \widehat{C}^n_{i-k_n}\right\| < u_n'\right) &\le \frac{E\left(\left\|\widehat{C}^n_{i+k_n} - \widehat{C}^n_{i-k_n} - \Delta C\right\|^s\right)}{(\epsilon/2)^s} \\
&\le KE\left(\left\|\widehat{C}^n_{i-k_n} - C_{S-}\right\|^s\right) + KE\left(\left\|\widehat{C}^n_{i+k_n} - C_S\right\|^s\right).
\end{aligned} \tag{E.45}
$$

The first term in (E.45) satisfies, for $s \ge 2$, by (E.36) and (E.49)

$$
\begin{aligned}
E\left(\left\|\widehat{C}^n_{i-k_n} - C_{S-}\right\|^s\right) &\le KE\left(\left\|\widehat{C}^n_{i-k_n} - \widehat{C}^{n\prime}_{i-k_n}\right\|^s\right) + KE\left(\left\|\widehat{C}^{n\prime}_{i-k_n} - C_{S-}\right\|^s\right) \\
&\le K_q a_n \Delta_n^{(2s-r)\varpi+1-s} + K\Delta_n^{s/4}.
\end{aligned}
$$

The second term in (E.45) has the same bound by the same arguments as the first term. Choosing $s = 2$ in the above, and taking into account that $(2 - r)\,\varpi \geq \frac{3}{4}$ and $\varpi \geq \frac{3}{8}$, we obtain (E.43).

Given (E.43), it is simple to obtain (E.41) as follows. By (E.43), if there is a jump on $(i\Delta_n, (i + k_n)\,\Delta_n]$, we know $P(A_i) = o_p\left(\Delta_n^{1/4}\right)$, thus $(A_i \cap A_{i+k_n}) \leq P(A_i) = o_p\left(\Delta_n^{1/4}\right)$. Applying (E.43) with $i + k_n$ instead of $i$, if there is a jump on $((i + k_n)\,\Delta_n, (i + 2k_n)\,\Delta_n]$, $P(A_{i+k_n}) = o_p\left(\Delta_n^{1/4}\right)$. Thus, $P(A_i \cap A_{i+k_n}) \leq P(A_{i+k_n}) = o_p\left(\Delta_n^{1/4}\right)$. We conclude that if there is a jump on $(i\Delta_n, (i + 2k_n)\,\Delta_n]$, i.e., event $B_i$ is true, then $(A_i \cap A_{i+k_n}) \leq P(A_{i+k_n}) = o_p\left(\Delta_n^{1/4}\right)$. This concludes the proof of (E.41) and hence Theorem B2.

## E.3 Proof of Theorem B3

To show this result, let us define the functions

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

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

$$\begin{aligned}
&\sum_{i\in L(n,T)} \vartheta_i^{AN} - \sum_{i\in L(n,T)} \vartheta_i'^{AN}\\
&= \frac{3}{2k_n}\sum_{i\in L(n,T)} \Big[\big(S(\widehat{C}_i^n, \widehat{C}_{i+k_n}^n) - S(\widehat{C}_i'^n, \widehat{C}_{i+k_n}'^n)\big) - \frac{2}{k_n}\big(U(\widehat{C}_i^n) - U(\widehat{C}_i'^n)\big)\Big],\\
&\sum_{i\in L(n,T)} \vartheta_i^{LIN} - \sum_{i\in L(n,T)} \vartheta_i'^{LIN}\\
&= \frac{3}{2k_n}\sum_{i\in L(n,T)} \Big[\big(R(\widehat{C}_i^n, \widehat{C}_{i+k_n}^n) - R(\widehat{C}_i'^n, \widehat{C}_{i+k_n}'^n)\big) - \frac{2}{k_n}\big(U(\widehat{C}_i^n) - U(\widehat{C}_i'^n)\big)\Big].
\end{aligned}$$

Since $H$ and $G$ are three times continuously differentiable with bounded derivatives, the functions $R$ and $S$ are continuously differentiable and satisfy

$$\|\partial J(x,y)\| \leq K \text{ for } J \in \{S,R\}, \tag{E.46}$$
$$\|\partial U(x)\| \leq K, \tag{E.47}$$

where $\partial J$ (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 $(x,y)$ (respectively, $x$). Using the Taylor expansion, (E.46) and (E.47), it holds that, for $J \in \{S,R\}$,

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

By equation (E.36), the following condition is sufficient for Theorem B3 to hold:

$$(2-r)\varpi - \frac{3}{4} \geq 0.$$

The above condition follows from our assumptions of Theorem 1. Using the fact that $0 < \varpi < \frac{1}{2}$, we can see that Theorem B3 holds when $3/4(2-r) \leq \varpi < \frac{1}{2}$, which completes the proof.

### E.4 Proof of Theorem B4

Note that we have

$$\sum_{i\in L(n,T)} \vartheta_i'^{LIN} - \sum_{i\in L(n,T)} \vartheta_i^{(A)} = \frac{3}{2k_n} \sum_{g,h,a,b=1}^{d} \sum_{i\in L(n,T)} \psi_i^n(g,h,a,b),$$

$$\sum_{i\in L(n,T)} \vartheta_i'^{AN} - \sum_{i\in L(n,T)} \vartheta_i^{(A)} = \frac{3}{2k_n} \sum_{i\in L(n,T)} \Big(\chi_i^n - \sum_{g,h,a,b=1}^{d} (\partial_{gh}H\partial_{ab}G)(C_i^n)\lambda_i^{n,gh}\lambda_i^{n,ab}\Big),$$

with

$$\psi_i^n(g,h,a,b) = \Big((\partial_{gh}H\partial_{ab}G)(\widehat{C}_i'^n) - (\partial_{gh}H\partial_{ab}G)(C_i^n)\Big)\lambda_i^{n,gh}\lambda_i^{n,ab},$$

$$\chi_i^n = \Big(H(\widehat{C}_{i+k_n}'^n) - H(\widehat{C}_i'^n)\Big)\Big(G(\widehat{C}_{i+k_n}'^n) - G(\widehat{C}_i'^n)\Big).$$

By Taylor expansion, we have

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

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

and

$$S(\widehat{C}_{i+k_n}'^n) - S(\widehat{C}_i'^n) = \sum_{gh} \partial_{gh}S(C_i^n)\lambda_i^{n,gh} + \sum_{j,k,g,h} \partial^2_{jk,gh}S(C_i^n)\lambda_i^{n,gh}\nu_i^{n,jk}$$

$$+\frac{1}{2}\sum_{x,y,g,h} \partial^2_{xy,gh}S(C_i^n)\lambda_i^{n,gh}\lambda_i^{n,xy} + \frac{1}{2}\sum_{x,y,j,k,g,h} \partial^3_{xy,jk,gh}S(CC_i^{n,S})\lambda_i^{n,gh}\nu_i^{n,xy}\nu_i^{n,jk}$$

$$+\frac{1}{6}\sum_{j,k,x,y,g,h} \partial^3_{jk,xy,gh}S(C_i^{n,S})\lambda_i^{n,jk}\lambda_i^{n,gh}\lambda_i^{n,xy},$$

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

By (4.10) in Jacod and Rosenbaum (2013) we have

$$\mathbb{E}\Big(\Big\|\alpha_i^n\Big\|^q\Big|\mathcal{F}_{(i-1)\Delta_n}\Big) \leq K_q\Delta_n^q \text{ for all } q\geq 0 \text{ and } \mathbb{E}\Big(\Big|\sum_{j=0}^{k_n-1}\alpha_{i+j}^n\Big|^q\Big|\mathcal{F}_{(i-1)\Delta_n}\Big) \leq K_q\Delta_n^q k_n^{q/2} \text{ for } q\geq 2. \tag{E.48}$$

Combining (E.48), (A.4), (B.14) with $Z = C$ and the Hölder inequality yields for $q \geq 2$, for $i \in L(n,T)$

$$\mathbb{E}\Big(\Big\|\nu_i^n\Big\|^q\Big|\mathcal{F}_{(i-1)\Delta_n}\Big) \leq K_q\Delta^{q/4}, \text{ and } \mathbb{E}\Big(\Big\|\lambda_i^n\Big\|^q\Big|\mathcal{F}_{(i-1)\Delta_n}\Big) \leq K_q\Delta^{q/4}. \tag{E.49}$$

The bound in the first equation of (E.49) is tighter than that in (4.11) of Jacod 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 derivatives of $H$ and $G$,

$$\left| \left(\partial^3_{jk,xy,ab} G \partial_{gh} H + \partial^2_{xy,gh} H \partial^2_{jk,ab} G\right)(\widetilde{c}^n_i) \nu^{n,xy}_i \nu^{n,jk}_i \lambda^{n,gh}_i \lambda^{n,ab}_i \right| \leq K \|\nu^n_i\|^2 \|\lambda^n_i\|^2. \tag{E.50}$$

Using the Taylor expansion, we have

$$\chi^n_i - \sum_{g,h,a,b} (\partial_{gh} H \partial_{ab} G)(C^n_i) \lambda^{n,gh}_i \lambda^{n,ab}_i =$$

$$\sum_{g,h,a,b,j,k} (\partial_{gh} H \partial^2_{jk,xy} G + \partial_{gh} G \partial^2_{jk,xy} H)(C^n_i)(\lambda^{n,gh}_i + \frac{1}{2}\nu^{n,gh}_i)\lambda^{n,ab}_i \lambda^{n,jk}_i + \varphi^n_i, \quad \text{and}$$

$$\sum_{g,h,a,b} \left(\partial_{gh} H \partial_{ab} G\right)(\widehat{C}'^n_i) - \left(\partial_{gh} H \partial_{ab} G\right)(C^n_i) =$$

$$\sum_{g,h,a,b,x,y} (\partial_{gh} H \partial^2_{ab,xy} G + \partial_{ab} G \partial^2_{gh,xy} G)(C^n_i)(\nu^{n,xy}_i)\lambda^{n,gh}_i \lambda^{n,ab}_i + \delta^n_i$$

with $\mathbb{E}(|\varphi^n_i| \big| \mathcal{F}^n_i) \leq K\Delta_n$ and $\mathbb{E}(|\delta^n_i| \big| \mathcal{F}^n_i) \leq K\Delta_n$ which follow by the Cauchy-Schwartz inequality together with equation (E.49). Given that $k_n = \theta(\Delta_n)^{-1/2}$, the previous inequalities imply

$$\frac{3\Delta_n^{-1/4}}{2k_n} \sum_{i \in L(n,T)} \varphi^n_i \overset{\mathbb{P}}{\Longrightarrow} 0 \quad \text{and} \quad \frac{3\Delta_n^{-1/4}}{2k_n} \sum_{i \in L(n,T)} \delta^n_i \overset{\mathbb{P}}{\Longrightarrow} 0.$$

Therefore, it suffices to show that

$$\frac{3\Delta_n^{-1/4}}{2k_n} \sum_{i \in L(n,T)} \sum_{g,h,a,b,j,k} (\partial_{gh} H \partial^2_{jk,ab} G + \partial_{gh} H \partial^2_{jk,ab} G)(C^n_i) \lambda^{n,gh}_i \lambda^{n,ab}_i \lambda^{n,jk}_i \overset{\mathbb{P}}{\longrightarrow} 0, \tag{E.51}$$

$$\frac{3\Delta_n^{-1/4}}{2k_n} \sum_{i \in L(n,T)} \sum_{g,h,a,b,j,k} (\partial_{gh} H \partial^2_{jk,ab} G + \partial_{gh} H \partial^2_{jk,ab} G)(C^n_i) \nu^{n,gh}_i \lambda^{n,ab}_i \lambda^{n,jk}_i \overset{\mathbb{P}}{\longrightarrow} 0. \tag{E.52}$$

These results hold by the bounds in Lemma B5.

## E.5 Proof of Theorem B5

In Section E.5, to simplify the notational burden, we adopt the following strategy. Instead of studying $\sum_{i \in L(n,T)} \vartheta^{(A)}_i$, we work with all indices $i$, i.e., $\sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1} \vartheta^{(A)}_i$, together with the assumption that there are no volatility jumps. The difference between the two quantities is $o_p\left(\Delta_n^{1/4}\right)$ because in the absence of volatility jumps, $\vartheta^{(A)}_i$ satisfies the bound in equation (E.39). Recall the decomposition from from B.13,

$$\vartheta^{(A)}_i = \vartheta^{(A1)}_i - \vartheta^{(A2)}_i. \tag{E.53}$$

Given the boundedness of the derivatives of $H$ and $G$ and the fact that $k_n = \theta(\Delta_n)^{-1/2}$, by Theorem 2.2 in Jacod and Rosenbaum (2015) we have

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

which yields

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

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

$$\widehat{C}_i'^n=C_{i-1}^n+\frac{1}{k_n}\sum_{j=0}^{k_n-1}\sum_{u=1}^{2}\overline{\varepsilon}(u)_j^n\zeta(u)_{i+j}^n,\qquad \widehat{C}_{i+k_n}'^n-\widehat{C}_i'^n=\frac{1}{k_n}\sum_{j=0}^{2k_n-1}\sum_{u=1}^{2}\varepsilon(u)_j^n\zeta(u)_{i+j}^n,$$

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

Changing the order of the summation in the last term yields

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

Therefore, we can further rewrite $\sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1}\vartheta_i^{(A1)}$ as

$$\begin{aligned}\sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1}\vartheta_i^{(A1)}&=\sum_{i\in L(n,T)}\vartheta_i^{(A11)}+\sum_{i\in L(n,T)}\vartheta_i^{(A12)}+\sum_{i\in L(n,T)}\vartheta_i^{(A13)},\text{with}\\\sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1}\vartheta_i^{(A1w)}&=\sum_{g,h,a,b=1}^{d}\sum_{u,v=1}^{2}\widehat{A1w}(H,gh,u;G,ab,v)_T^n,\quad w=1,2,3,\end{aligned}$$

where

$$\begin{aligned}\widehat{A11}(H,gh,u;G,ab,v)_T^n&=\frac{3}{2k_n^3}\sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1}\sum_{j=0}^{2k_n-1}(\partial_{gh}H\partial_{ab}G)(C_{i-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n\zeta(u)_{i+j}^{n,gh}\zeta(v)_{i+j}^{n,ab},\\\widehat{A12}(H,gh,u;G,ab,v)_T^n&=\frac{3}{2k_n^3}\sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1}\sum_{j=0}^{2k_n-2}\sum_{q=j+1}^{2k_n-1}(\partial_{gh}H\partial_{ab}G)(C_{i-1}^n)\varepsilon(u)_j^n\varepsilon(v)_q^n\zeta(u)_{i+j}^{n,gh}\zeta(v)_{i+q}^{n,ab},\\\widehat{A13}(H,gh,u;G,ab,v)_T^n&=\frac{3}{2k_n^3}\sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1}\sum_{j=0}^{2k_n-2}\sum_{q=j+1}^{2k_n-1}(\partial_{gh}H\partial_{ab}G)(C_{i-1}^n)\varepsilon(v)_j^n\varepsilon(u)_q^n\zeta(v)_{i+j}^{n,ab}\zeta(u)_{i+q}^{n,gh},\end{aligned}$$

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

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

$$
\begin{aligned}
&\times (C^n_{i-j-1})\varepsilon(u)^n_j\varepsilon(v)^n_j\zeta(u)^{n,gh}_i\zeta(v)^{n,ab}_i,\\
\widehat{A12}(H,gh,u;G,ab,v)^n_T &= \frac{3}{2k_n^3}\sum_{i=2}^{[T/\Delta_n]}\sum_{m=1}^{(i-1)\wedge(2k_n-1)}\sum_{j=0\vee(i+2k_n-1-m-[T/\Delta_n])}^{(2k_n-m-1)\wedge(i-m-1)}(\partial_{gh}H\partial_{ab}G)(C^n_{i-1-j-m})\\
&\times \varepsilon(u)^n_j\varepsilon(v)^n_{j+m}\zeta_{gh}(u)^n_{i-m}\zeta_{ab}(v)^n_i.
\end{aligned}
$$

Now, set

$$
\begin{aligned}
\widetilde{A11}(H,gh,u;G,ab,v)^n_T &= \frac{3}{2k_n^3}\sum_{i=3k_n}^{[T/\Delta_n]-k_n}\sum_{j=0}^{2k_n-1}(\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1})\varepsilon(u)^n_j\varepsilon(v)^n_j\zeta(u)^{n,gh}_i\zeta(v)^{n,ab}_i,\\
\widetilde{A12}(H,gh,u;G,ab,v)^n_T &= \frac{3}{2k_n^3}\sum_{i=3k_n}^{[T/\Delta_n]-k_n}\sum_{m=1}^{(i-1)\wedge(2k_n-1)}\sum_{j=0}^{(2k_n-m-1)}(\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1-m})\varepsilon(u)^n_j\varepsilon(v)^n_{j+m}\\
&\times \zeta_{gh}(u)^n_{i-m}\zeta_{ab}(v)^n_i.
\end{aligned}
$$

We show below that the following results hold:

$$
\frac{1}{\Delta_n^{1/4}}\left(\widehat{A1w}(H,gh,u;G,ab,v)^n_T-\widetilde{A1w}(H,gh,u;G,ab,v)^n_T\right)\xrightarrow{\mathbb{P}}0 \tag{E.54}
$$

$$
\frac{1}{\Delta_n^{1/4}}\left(\widetilde{A1w}(H,gh,u;G,ab,v)^n_T-\overline{A1w}(H,gh,u;G,ab,v)^n_T\right)\xrightarrow{\mathbb{P}}0 \tag{E.55}
$$

for all $(H,gh,u,G,ab,v)$ and $w=1,2$.

### E.5.1 Proof of Equation (E.54) for $w=1$

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

$$
\begin{aligned}
&\widehat{A11}-\widetilde{A11}=\widetilde{\widehat{A11}}(1)+\widetilde{\widehat{A11}}(2)+\widetilde{\widehat{A11}}(3)\quad\text{with}\\
&\widetilde{\widehat{A11}}(1)=\sum_{i=1}^{(2k_n-1)\wedge[T/\Delta_n]}\left(\frac{3}{2k_n^3}\sum_{j=0\vee(i+2k_n-1-[T/\Delta_n])}^{(2k_n-1)\wedge(i-1)}(\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1})\varepsilon(u)^n_j\varepsilon(v)^n_j\right)\zeta(u)^{n,gh}_i\zeta(v)^{n,ab}_i,\\
&\widetilde{\widehat{A11}}(2)=\sum_{i=[T/\Delta_n]-2k_n+2}^{[T/\Delta_n]}\frac{3}{2k_n^3}\left(\sum_{j=0\vee(i+2k_n-1-[T/\Delta_n])}^{(2k_n-1)\wedge(i-1)}(\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1})\varepsilon(u)^n_j\varepsilon(v)^n_j\right.\\
&\qquad\left.-\sum_{j=0}^{(2k_n-1)}(\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1})\varepsilon(u)^n_j\varepsilon(v)^n_j\right)\zeta(u)^{n,gh}_i\zeta(v)^{n,ab}_i,\\
&\widetilde{\widehat{A11}}(3)=\sum_{i=2k_n}^{[T/\Delta_n]-2k_n+1}\frac{3}{2k_n^3}\left(\sum_{j=0\vee(i+2k_n-1-[T/\Delta_n])}^{(2k_n-1)\wedge(i-1)}(\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1})\varepsilon(u)^n_j\varepsilon(v)^n_j\right.\\
&\qquad\left.-\sum_{j=0}^{(2k_n-1)}(\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1})\varepsilon(u)^n_j\varepsilon(v)^n_j\right)\zeta(u)^{n,gh}_i\zeta(v)^{n,ab}_i.
\end{aligned}
$$

It is easy to see that $\widetilde{\widehat{A12}}(3)=0$. Using equation (B.14) with $Z=c$ and equation (E.48), we obtain

$$
\mathbb{E}(\|\zeta(1)^n_i\|^q|\mathcal{F}^n_{i-1})\le K_q,\ \ \mathbb{E}(\|\zeta(2)^n_i\|^q|\mathcal{F}^n_{i-1})\le K_q\Delta_n^{q/2}. \tag{E.56}
$$

By the boundedness of the derivatives of $H$ and $G$, the random quantities

$\left(\frac{3}{2k_n^3}\sum_{j=0\vee(i+2k_n-1-[T/\Delta_n])}^{(2k_n-1)\wedge(i-1)}(\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n\right)$ and $\frac{3}{2k_n^3}\sum_{j=0}^{(2k_n-1)}(\partial_{gh}H\partial_{ab}G)(C_{i-j-1}^n)\varepsilon(u)_j^n\varepsilon(v)_j^n$ are $\mathcal{F}_{i-1}^n-$ measurable and are bounded by $\widetilde{\lambda}_{u,v}^n$ defined as

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

Similarly, the quantity

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

is $\mathcal{F}_{i-1}^n-$ measurable and bounded by $2\widetilde{\lambda}_{u,v}^n$. Note also that, by equation (E.56) and the Cauchy Schwartz inequality, we have

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

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

### E.5.2 Proof of Equation (E.54) for $w=2$

First, observe that $\widehat{A12}-\widetilde{A12}=\widetilde{\widetilde{A12}}(1)+\widetilde{\widetilde{A12}}(2)$, with

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

Notice that the quantity

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

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

$$\kappa^n_i = \sum_{m=1}^{(i-1)} \frac{3}{2k_n^3}\Big(\sum_{j=0\vee(i+2k_n-1-m-[T/\Delta_n])}^{(2k_n-m-1)\wedge(i-m-1)} (\partial_{gh}H\partial_{ab}G)(C^n_{i-1-j-m})\varepsilon(u)^n_j\varepsilon(v)^n_{j+m}\Big)\zeta_{gh}(u)^n_{i-m}.$$

It follows that $\kappa^n_i$ is $\mathcal{F}^n_{i-1}$-measurable and we have

$$\begin{aligned}
\mathbb{E}(|\kappa^{m,n}_i|^z|\mathcal{F}_0) &\leq (\widetilde{\lambda}^n_{u,v})^z, \\
|\mathbb{E}(\zeta(u)^n_{i-m}|\mathcal{F}_{i-m-1})| &\leq \begin{cases} K\sqrt{\Delta_n} & \text{if } u=1 \\ K\Delta_n & \text{if } u=2 \end{cases}, \\
\mathbb{E}(\|\zeta(u)^n_{i-m}\|^z|\mathcal{F}_{i-m-1}) &\leq \begin{cases} K_z & \text{if } u=1 \\ K_z\Delta_n^{z/2} & \text{if } u=2 \end{cases}.
\end{aligned}$$

Using Lemma B3, we deduce that for $z \geq 2$,

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

Using the above result, we obtain $\frac{1}{\Delta_n^{1/4}}\widetilde{\widehat{A12}}(1) \overset{\mathbb{P}}{\Rightarrow} 0$. A similar argument yields $\frac{1}{\Delta_n^{1/4}}\widetilde{\widehat{A12}}(2) \overset{\mathbb{P}}{\Rightarrow} 0$, which completes the proof of the equation (E.54) for $w=2$.

### E.5.3 Proof of Equation (E.55) for $w = 1$

Define

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

By Taylor expansion, boundedness of the derivatives of $H$ and $G$, and using (B.14) with $Z=c$, we have

$$\begin{aligned}
&\Big|\mathbb{E}\Big((\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1}) - (\partial_{gh}H\partial_{ab}G)(C^n_{i-2k_n})\big|\mathcal{F}^n_{i-2k_n}\Big)\Big| \leq K(k_n\Delta_n) \leq K\sqrt{\Delta_n} \\
&\mathbb{E}(|(\partial_{gh}H\partial_{ab}G)(C^n_{i-j-1}) - (\partial_{gh}H\partial_{ab}G)(C^n_{i-2k_n})|^q|\mathcal{F}^n_{i-2k_n})| \leq K(k_n\Delta_n)^{q/2} \leq K\Delta_n^{q/4},
\end{aligned}$$

for $q \geq 2$ and for $j = 0,\ldots,2k_n-1$. Next, observe that $\Theta(u,v)^{(C),i,n}_0$ is $\mathcal{F}^n_{i-1}$ -measurable and satisfies $|\Theta(u,v)^{(C),i,n}_0| \leq \widetilde{\lambda}^n_{u,v}$, $|\mathbb{E}\Big(\Theta(u,v)^{(C),i,n}_0|\mathcal{F}^n_{i-2k_n}\Big)| \leq K\Delta_n^{1/2}\widetilde{\lambda}^n_{u,v}$ and $\mathbb{E}\Big(|\Theta(u,v)^{(C),i,n}_0|^q|\mathcal{F}^n_{i-2k_n}\Big) \leq K_q\Delta_n^{q/4}(\widetilde{\lambda}^n_{u,v})^q$ where the latter follows from the Hölder inequality. We aim to prove that

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

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

$$\widehat{E}(1) = \frac{1}{\Delta_n^{1/4}}\left[\sum_{i=3k_n}^{[T/\Delta_n]-k_n}\Theta(u,v)^{(C),i,n}_0\mathbb{E}(\zeta(u)^{n,gh}_i\zeta(v)^{n,ab}_i|\mathcal{F}^n_{i-1})\right]$$

$$\widehat{E}(2) = \frac{1}{\Delta_n^{1/4}} \left[ \sum_{i=3k_n}^{[T/\Delta_n]-k_n} \Theta(u,v)_0^{(C),i,n} \big( \zeta(u)_i^{n,gh} \zeta(v)_i^{n,ab} - \mathbb{E}(\zeta(u)_i^{n,gh} \zeta(v)_i^{n,ab} | \mathcal{F}_{i-1}^n) \big) \right],$$

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

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

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

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

The latter inequality implies $\widehat{E}(2) \overset{\mathbb{P}}{\Rightarrow} 0$ for all $(u,v)$. It remains to show that $\widehat{E}(1) \overset{\mathbb{P}}{\Rightarrow} 0$. Here, we recall some bounds under Assumption 2,

$$|\mathbb{E}(\zeta(1)_i^{n,gh} \zeta(2)_i^{n,ab} | \mathcal{F}_{i-1}^n)| \leq K \Delta_n, \tag{E.57}$$

$$|\mathbb{E}(\zeta(1)_i^{n,gh} \zeta(1)_i^{n,ab} | \mathcal{F}_{i-1}^n) - \big( C_{i-1}^{n,ga} C_{i-1}^{n,hb} + C_{i-1}^{n,gb} C_{i-1}^{n,ha} \big)| \leq K \Delta_n^{1/2}, \tag{E.58}$$

$$|\mathbb{E}(\zeta(2)_i^{n,gh} \zeta(2)_i^{n,ab} | \mathcal{F}_{i-1}^n - \overline{C}_{i-1}^{n,gh,ab} \Delta_n)| \leq K \Delta_n^{3/2} (\sqrt{\Delta_n} + \eta_i^n). \tag{E.59}$$

**Case** $(u,v) \in \{(1,2),(2,1)\}$. By equation (E.57) we have

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

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

$$\widehat{E}'(1) = \frac{1}{\Delta_n^{1/4}} \left[ \sum_{i=3k_n}^{[T/\Delta_n]-k_n} \Theta(u,v)_0^{(C),i,n} V_{i-2k_n}^n \right]$$

$$\widehat{E}''(1) = \frac{1}{\Delta_n^{1/4}} \left[ \sum_{i=3k_n}^{[T/\Delta_n]-k_n} \Theta(u,v)_0^{(C),i,n} \big( V_{i-1}^n - V_{i-2k_n}^n \big) \right]$$

$$\widehat{E}'''(1) = \frac{1}{\Delta_n^{1/4}} \left[ \sum_{i=3k_n}^{[T/\Delta_n]-k_n} \Theta(u,v)_0^{(C),i,n} \Big( \mathbb{E}(\zeta(u)_i^{n,gh} \zeta(v)_i^{n,ab} | \mathcal{F}_{i-1}^n) - V_{i-1}^n \Big) \right]$$

where

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

Note that we have $\widehat{E}(1) = \widehat{E}'(1) + \widehat{E}''(1) + \widehat{E}'''(1)$. Using equations (E.58) and (E.59), it can be shown that

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

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

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

Using the $\mathcal{F}_{i+2k_n-2}^n$-measurability of the last sum, we are able to show

$$\frac{1}{\Delta_n^{1/4}} \left[ \sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1} |\mathbb{E}(\Theta(u,v)_0^{(C),i-1+2k_n,n} V_{i-1}^n | \mathcal{F}_{i-1}^n)| \right] \overset{\mathbb{P}}{\Rightarrow} 0 \quad \text{and}$$
$$\frac{2k_n-2}{\Delta_n^{1/2}} \left[ \sum_{i=k_n+1}^{[T/\Delta_n]-3k_n+1} \mathbb{E}\Big( |\Theta(u,v)_0^{(C),i-1+2k_n,n} V_{i-1}^n)|^2 \Big) \right] \Rightarrow 0.$$

The first result readily follows from the inequality

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

while the second is a direct consequence of

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

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

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

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

### E.5.4 Proof of Equation (E.55) for $w = 2$

Our aim here is to show that

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

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

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

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

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

Lemma B3 implies that for $q \geq 2$,

$$\mathbb{E}(|\rho(u,v)^{(C),i,n,gh}|^q) \leq \begin{cases} K_q(\Delta_n^{1/4}\widetilde{\lambda}_{u,v}^n)^q k_n^{q/2} & \text{if } u=1 \\ K_q(\Delta_n^{1/4}\widetilde{\lambda}_{u,v}^n)^q / k_n^{q/2} & \text{if } u=2 \end{cases} \leq \begin{cases} K_q/k_n^{2q} & \text{if } v=1 \\ K_q k_n^q & \text{if } v=2 \end{cases}. \tag{E.60}$$

Set

$$\widehat{E}'(2) = \frac{1}{\Delta_n^{1/4}} \sum_{i=3k_n}^{[T/\Delta_n]-k_n} \rho(u,v)^{(C),i,n,gh} \mathbb{E}(\zeta(v)_i^{n,ab} | \mathcal{F}_{i-1}^n),$$
$$\widehat{E}''(2) = \frac{1}{\Delta_n^{1/4}} \sum_{i=3k_n}^{[T/\Delta_n]-k_n} \rho(u,v)^{(C),i,n,gh} (\zeta(v)_i^{n,ab} - \mathbb{E}(\zeta(v)_i^{n,ab} | \mathcal{F}_{i-1}^n)).$$

The martingale increments property implies $\mathbb{E}(|\widehat{E}''(2)|^2) \leq K\Delta_n^{1/2}$ in all the cases, which in turn implies $\widehat{E}''(2) \stackrel{\mathbb{P}}{\Longrightarrow} 0$. Next, using the bounds on $\rho(u,v)^{(C),i,n,gh}$, we obtain that $\widehat{E}'(2) \stackrel{\mathbb{P}}{\Longrightarrow} 0$.

We refer to Jacod and Rosenbaum (2015) for the proofs of Lemma B1 and Lemma B2.

## E.6 Proof of Lemma B3

Set

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

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

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

Therefore, on the one hand we have

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

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

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

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

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

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

## E.7 Proof of Lemma B4

We use $i \in L(n,T)$ throughout the proof of Lemma B4. We use the terminology "successive conditioning" to refer to either of the following two equalities,

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

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

To prove Lemma B4, we first note that $\lambda_i^{n,jk}\lambda_i^{n,lm}$ is $\mathcal{F}_{i+2k_n}^n$-measurable. Therefore, by the law of iterated expectations, we have

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

By equation (3.27) in Jacod and Rosenbaum (2015), we have

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

From the above, it follows that

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

where the last inequality follows from Lemma B1.

Now, using equation (B.14) successively with $Z = C$ and $Z = \overline{C}$ (recall that the latter holds under Assumption 2), together with the successive conditioning, we also have

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

The result derives from the last inequality.

## E.8 Proof of Lemma B5

### E.8.1 Proof of Equation (B.16) in Lemma B5

We start by obtaining some useful bounds for some important quantities. First, using the second statement in Lemma B2 applied to $Z = Y'$, we have

$$|\mathbb{E}(\alpha_i^{n,jk}|\mathcal{F}_{i-1}^n)| \leq K\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_{i,1}^n). \tag{E.61}$$

Second, by repeated application of the Cauchy-Schwartz inequality and making use of the third and last statements in Lemma B2 as well as equation (B.14) with $Z = C$, it can be shown that

$$\left|\mathbb{E}(\alpha_i^{n,jk}\alpha_i^{n,lm}|\mathcal{F}_{i-1}^n) - \Delta_n^2\Big(C_i^{n,jl}C_i^{n,km} + C_i^{n,jm}C_i^{n,kl}\Big)\right| \leq K\Delta_n^{5/2}. \tag{E.62}$$

Next, by successive conditioning and using the bound in equation (B.14) for $Z = C$ as well as equations (E.61) and (E.62), we have for $0 \leq u \leq k_n - 1$,

$$\left|\mathbb{E}(\alpha_{i+u}^{n,jk}|\mathcal{F}_{i-1}^n)\right| \leq K\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_{i,u}^n), \tag{E.63}$$

$$\left|\mathbb{E}(\alpha_{i+u}^{n,jk}\alpha_{i+u}^{n,lm}|\mathcal{F}_{i-1}^n) - \Delta_n^2\Big(C_i^{n,jl}C_i^{n,km} + C_i^{n,jm}C_i^{n,kl}\Big)\right| \leq K\Delta_n^{5/2}. \tag{E.64}$$

To prove equation (B.16), we first observe that $\nu_i^{n,jk}\nu_i^{n,lm}\nu_i^{n,gh}$ can be decomposed as

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

with $\zeta_{i,u}^n = \alpha_{i+u}^n + (C_{i+u}^n - C_i^n)\Delta_n$, which satisfies $\mathbb{E}(\|\zeta_{i,u}^n\|^q|\mathcal{F}_{i-1}^n) \leq K\Delta_n^q$ for $q \geq 2$.
Set

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

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

The following bounds complete the proof of equation (B.16),

$$|\mathbb{E}(\xi_i^n(1)|\mathcal{F}_{i-1}^n)| \leq K\Delta_n \tag{E.65}$$
$$|\mathbb{E}(\xi_i^n(2)|\mathcal{F}_{i-1}^n)| \leq K\Delta_n \tag{E.66}$$
$$|\mathbb{E}(\xi_i^n(3)|\mathcal{F}_{i-1}^n)| \leq K\Delta_n \tag{E.67}$$

$$|\mathbb{E}(\xi_i^n(4)|\mathcal{F}_{i-1}^n)| \leq K\Delta_n^{3/4}(\Delta_n^{1/4} + \eta_{i,k_n}). \tag{E.68}$$

These bounds are proved below.

### Proof of Equation (E.65)

The result readily follows from an application of the Cauchy Schwartz inequality coupled with the bound $\mathbb{E}(\|\zeta_{i+u}^n\|^q|\mathcal{F}_{i-1}^n) \leq K_q\Delta_n^q$ for $q \geq 2$.

### Proof of Equation (E.66)

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

$$\mathbb{E}(\zeta_{i+u}^{n,jk}\zeta_{i+v}^{n,lm}\zeta_{i+v}^{n,gh}|\mathcal{F}_{i-1}^n) = \mathbb{E}(\zeta_{i+u}^{n,jk}\mathbb{E}(\zeta_{i+v}^{n,lm}\zeta_{i+v}^{n,gh}|\mathcal{F}_{i+u+1}^n)\big|\mathcal{F}_{i-1}^n). \tag{E.69}$$

By successive conditioning, equation (E.62), and the Cauchy-Schwartz inequality, we also have

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

Given that $\mathbb{E}(|\zeta_{i+u}^{n,jk}|^q\big|\mathcal{F}_{i-1}^n) \leq \Delta_n^q$, the approximation error involved in replacing $\mathbb{E}(\zeta_{i+v}^{n,lm}\zeta_{i+v}^{n,gh}|\mathcal{F}_{i+u+1}^n)$ by $\Delta_n^2(C_{i+u+1}^{n,lg}C_{i+u+1}^{n,mh} + C_{i+u+1}^{n,lh}C_{i+u+1}^{n,mg}) + \Delta_n^2(C_{i+u+1}^{n,gh} - C_i^{n,gh})(C_{i+u+1}^{n,lm} - C_i^{n,lm})$ in equation (E.69) is smaller than $\Delta_n^{7/2}$.
We can also easily show that

$$|\mathbb{E}(\alpha_{i+u}^{n,jk}(C_{i+u+1}^{n,lm} - C_{i+u}^{n,lm})|\mathcal{F}_{i-1}^n)| \leq K\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_{i,k_n}^n). \tag{E.70}$$

Since $(C_{i+u}^n - C_i^n)$ is $\mathcal{F}_{i+u}^n$-measurable, we use the successive conditioning, the Cauchy-Schwartz inequality, equation (E.61), equation (E.62), and the fifth statement in Lemma B2 applied to $Z = c$ to obtain

$$\begin{aligned}
|\mathbb{E}(\alpha_{i+u}^{n,gh}(C_{i+u}^{n,lm} - C_i^{n,lm})(C_{i+u}^{n,jk} - C_i^{n,jk})|\mathcal{F}_{i-1}^n)| &\leq K\Delta_n^{5/2} \\
|\mathbb{E}(\alpha_{i+u}^{n,jk}\alpha_{i+u}^{n,lm}(C_{i+u}^{n,gh} - C_i^{n,gh})|\mathcal{F}_{i-1}^n)| &\leq K\Delta_n^{5/2} \\
|\mathbb{E}\big((C_{i+u}^{n,lm} - C_i^{n,lm})(C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+u}^{n,gh} - C_i^{n,gh})\big)|\mathcal{F}_{i-1}^n)| &\leq K\Delta_n.
\end{aligned} \tag{E.71}$$

The following inequalities can be established using equation (E.61), the successive conditioning together with equation (B.14) for $Z = C$,

$$\begin{aligned}
\Big|\mathbb{E}(\alpha_{i+u}^{n,jk}(C_{i+u+1}^{n,lg}C_{i+u+1}^{n,mh} + C_{i+u+1}^{n,lh}C_{i+u+1}^{n,mg})|\mathcal{F}_{i-1}^n)\Big| &\leq K\Delta_n^{3/2} \\
\Big|\mathbb{E}\Big((C_{i+u}^{n,jk} - C_i^{n,jk})\big(C_{i+u+1}^{n,lg}C_{i+u+1}^{n,mh} + C_{i+u+1}^{n,lh}C_{i+u+1}^{n,mg})|\mathcal{F}_{i-1}^n\Big)\Big| &\leq K\Delta_n^{1/2} \\
\Big|\mathbb{E}(\alpha_{i+u}^{n,jk}(C_{i+u+1}^{n,gh} - C_i^{n,gh})(C_{i+u+1}^{n,lm} - C_i^{n,lm})|\mathcal{F}_{i-1}^n)\Big| &\leq K\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_{i,k_n}^n).
\end{aligned}$$

The last three inequalities together yield $|\mathbb{E}(\xi_i^n(2)|\mathcal{F}_{i-1}^n)| \leq K\Delta_n$.

### Proof of Equation (E.67)

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

$$\mathbb{E}(\zeta_{i+u}^{n,jk}\zeta_{i+u}^{n,lm}\zeta_{i+v}^{n,gh}|\mathcal{F}_{i-1}^n) = \mathbb{E}(\zeta_{i+u}^{n,jk}\zeta_{i+u}^{n,lm}\mathbb{E}(\zeta_{i+v}^{n,gh}|\mathcal{F}_{i+u}^n)\big|\mathcal{F}_{i-1}^n). \tag{E.72}$$

By successive conditioning and equation (E.61), we have

$$|\mathbb{E}(\alpha_{i+w}^{n,gh}|\mathcal{F}_{i+v}^n)| \le K\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_{i+v+1,w-v}). \tag{E.73}$$

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

$$\begin{aligned} & |\mathbb{E}\big((C_{i+w}^{n,gh} - C_{i+v+1}^{n,gh}))|\mathcal{F}_{i-1}^n\big) - \Delta_n(w-v-1)\widetilde{b}_{i+v+1}^{n,gh}| \\ \le \quad & K(w-v-1)\Delta_n\eta_{i+v+1,w-v} \le K\Delta_n^{1/2}\eta_{i+v+1,w-v}. \end{aligned}$$

The last two inequalities together imply

$$\left|\mathbb{E}\Big(\zeta_{i+w}^{n,gh}|\mathcal{F}_{i+v}^n\Big) - (C_{i+v+1}^{n,gh} - C_i^{n,gh})\Delta_n - \Delta_n^2(w-v-1)\widetilde{b}_{i+v+1}^{n,gh}\right| \le K\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_{i+v+1,w-v}). \tag{E.74}$$

Since $\mathbb{E}(|\zeta_{i,u}^{n,jk}|^q|\mathcal{F}_{i-1}^n) \le \Delta_n^q$, the error induced by replacing $\mathbb{E}(\zeta_{i+v}^{n,gh}|\mathcal{F}_{i+u}^n)$ by $(C_{i+v+1}^{n,gh} - C_i^{n,gh})\Delta_n + \Delta_n^2(w-v-1)\widetilde{b}_{i+v+1}^{n,gh}$ in equation (E.72) is smaller that $\Delta_n^{7/2}$.
Using Cauchy Schwartz inequality, successive conditioning, equation (E.71), equation (B.14) for $Z = C$ and the boundedness of $\widetilde{b}_t$ and $C_t$ we obtain

$$\begin{aligned} \left|\mathbb{E}\Big(\alpha_{i+u}^{n,jk}\alpha_{i+u}^{n,lm}(C_{i+u+1}^{n,jk} - C_i^{n,gh})|\mathcal{F}_{i+u-1}^n\Big)\right| &\le K\Delta_n^{5/2} \\ \left|\mathbb{E}\Big(\alpha_{i+u}^{n,jk}\alpha_{i+u}^{n,lm}\widetilde{b}_{i+u+1}^{n,gh}|\mathcal{F}_{i+u-1}^n\Big)\right| &\le K\Delta_n^2 \\ \left|\mathbb{E}\Big(\alpha_{i+u}^{n,jk}(C_{i+u}^{n,lm} - C_i^{n,lm})(C_{i+u+1}^{n,gh} - C_i^{n,gh})|\mathcal{F}_{i-1}^n\Big)\right| &\le K\Delta_n^{1/4}\Delta_n^{3/2}(\sqrt{\Delta_n} + \eta_{i,k_n}^n) \\ \left|\mathbb{E}\Big(\alpha_{i+u}^{n,jk}(C_{i+u}^{n,lm} - C_i^{n,lm})\widetilde{b}_{i+u+1}^{n,gh}|\mathcal{F}_{i-1}^n\Big)\right| &\le \Delta_n^{5/4} \\ \left|\mathbb{E}\Big((C_{i+u}^{n,jk} - C_i^{n,gh})(C_{i+u}^{n,lm} - C_i^{n,lm})\widetilde{b}_{i+u+1}^{n,gh}|\mathcal{F}_{i-1}^n\Big)\right| &\le K\Delta_n^{1/2} \\ \left|\mathbb{E}\Big((C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+u}^{n,lm} - C_i^{n,lm})(C_{i+u+1}^{n,gh} - C_i^{n,gh})|\mathcal{F}_{i-1}^n\Big)\right| &\le K\Delta_n. \end{aligned}$$

The above inequalities together yield $|\mathbb{E}(\xi_i^n(3)|\mathcal{F}_{i-1}^n)| \le K\Delta_n$.

**Proof of Equation (E.68)**

We first observe that $\xi_i^n(4)$ can be rewritten as

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

where

$$\begin{aligned} \zeta_{i+u}^{n,jk}\zeta_{i+v}^{n,lm}\zeta_{i+w}^{n,gh} = \Big[ & \alpha_{i+u}^{n,jk}\alpha_{i+v}^{n,lm}\alpha_{i+w}^{n,gh} + \alpha_{i+u}^{n,jk}\Delta_n\alpha_{i+v}^{n,lm}(C_{i+w}^{n,gh} - C_i^{n,gh}) + \alpha_{i+u}^{n,jk}\Delta_n(C_{i+v}^{n,lm} - C_i^{n,lm})\alpha_{i+w}^{n,gh} \\ & + \Delta_n^2\alpha_{i+u}^{n,jk}(C_{i+v}^{n,lm} - C_i^{n,lm})(C_{i+w}^{n,gh} - C_i^{n,gh}) + \Delta_n(C_{i+u}^{n,jk} - C_i^{n,jk})\alpha_{i+v}^{n,lm}\alpha_{i+w}^{n,gh} \end{aligned}$$

$$+ \Delta_n^2 (C_{i+u}^{n,jk} - C_i^{n,jk}) \alpha_{i+v}^{n,lm} (C_{i+w}^{n,gh} - C_i^{n,gh}) + \Delta_n^2 (C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+v}^{n,lm} - C_i^{n,lm}) \alpha_{i+w}^{n,gh}$$
$$+ \Delta_n^3 (C_{i+u}^{n,jk} - C_i^{n,jk})(C_{i+v}^{n,lm} - C_i^{n,lm})(C_{i+w}^{n,gh} - C_i^{n,gh}) \Bigg].$$

Based on the above decomposition, we set

$$\xi_i^n(4) = \sum_{j=1}^{8} \chi(j),$$

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

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

Upon changing the order of the summation, we have

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

Define also

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

Note that $\mathbb{E}(\chi(1) | \mathcal{F}_{i-1}^n) = \mathbb{E}(\chi'(1) | \mathcal{F}_{i-1}^n)$.
By Lemma B3, we have for $q \geq 2$,

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

The Cauchy-Schwartz inequality yields

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

where the last iteration is obtained using equation (E.73) 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 B1. It follows that

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

Next, we introduce

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

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

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

Similar steps to $\chi(1)$ lead to

$$|\mathbb{E}(\chi(2)\big|\mathcal{F}_{i-1}^n)| \leq K\Delta_n^{3/4}(\sqrt{\Delta_n}+\eta_{i,k_n}^n) \quad \text{and} \quad |\mathbb{E}(\chi(j)\big|\mathcal{F}_{i-1}^n)| \leq K\Delta_n(\sqrt{\Delta_n}+\eta_{i,k_n}^n) \quad \text{for} \quad j=3,4.$$

Define

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

where we have $\mathbb{E}(\chi(5)|\mathcal{F}_{i-1}^n) = \mathbb{E}(\chi'(5)|\mathcal{F}_{i-1}^n)$. Recalling equation (E.74), we further decompose $\chi'(5)$ as,

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

with

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

$$\chi\prime(5)[2] =\frac{1}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\Delta_n(C^{n,gh}_{i+v}-C^{n,gh}_{i})\Big(\sum_{u=0}^{v-1}\alpha^{n,jk}_{i+u}\Big)\alpha^{n,lm}_{i+v}$$
$$\chi\prime(5)[3] =\frac{1}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\Big(\sum_{u=0}^{v-1}\alpha^{n,jk}_{i+u}\Big)\Delta_n(C^{n,gh}_{i+v+1}-C^{n,gh}_{i+v})\alpha^{n,lm}_{i+v}$$
$$\chi\prime(5)[4] =\frac{1}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\Big(\sum_{u=0}^{v-1}\alpha^{n,jk}_{i+u}\Big)\Delta_n^2(w-v-1)(\widetilde{b}^{n,gh}_{i+v+1}-\widetilde{b}^{n,gh}_{i+v})\alpha^{n,lm}_{i+v}$$
$$\chi\prime(5)[5] =\frac{1}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\sum_{v=0}^{w-1}\Delta_n^2(w-v-1)\widetilde{b}^{n,gh}_{i+v}\Big(\sum_{u=0}^{v-1}\alpha^{n,jk}_{i+u}\Big)\alpha^{n,lm}_{i+v}.$$

Using equations (E.74), (E.73), and (E.70) and following the same strategy proof as for $\chi(1)$, it can be shown that

$$|\mathbb{E}\Big(\chi\prime(5)[j]\big|\mathcal{F}^n_{i-1}\Big)| \le K\Delta_n^{3/4}(\sqrt{\Delta_n}+\eta^n_{i,k_n}), \quad \text{for} \quad j=1,\dots,5,$$

which in turn implies

$$|\mathbb{E}\Big(\chi(5)\big|\mathcal{F}^n_{i-1}\Big)| \le K\Delta_n^{3/4}(\sqrt{\Delta_n}+\eta^n_{i,k_n}), \quad \text{for} \quad j=1,\dots,5.$$

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

$$|\mathbb{E}\Big(\chi(6)\big|\mathcal{F}^n_{i-1}\Big)| \le K\Delta_n^{3/4}(\sqrt{\Delta_n}+\eta^n_{i,k_n}).$$

Next, we set

$$\chi(7) = \frac{1}{(k_n\Delta_n)^3}\sum_{w=2}^{k_n-1}\left(\sum_{v=0}^{w-1}\Big(\sum_{u=0}^{v-1}\alpha^{n,jk}_{i+u}\Big)\Delta_n(C^{n,lm}_{i+v}-C^{n,lm}_{i})\Delta_n(C^{n,gh}_{i+w}-C^{n,gh}_{i})\right).$$

Define

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

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

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

To handle the remaining term $\chi(7)[4]$, we decompose it $\chi(7)[4] = \sum_{j=1}^{9} \chi(7)[4][j]$, where

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

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

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

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

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

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

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

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

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

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

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

Using arguments similar to those involved for the treatment of $\chi(1)$, it can be shown that

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

which yields

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

Next, define

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

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

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

These bounds can be used to deduce

$$|\mathbb{E}(\chi(8)\big|\mathcal{F}_{i-1}^n)| \leq K\Delta_n.$$

This completes the proof.

### E.8.2 Proof of Equations (B.17) and (B.18) in Lemma B5

Observe that

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

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

Hence, equations (B.17) and (B.18) can be proved using the same strategy as for (B.16).

#### E.8.3 Proof of Equations (B.19) and (B.20) in Lemma B5

Note that we have

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

and

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

From (A.4), notice that $\nu_i^n$ is $\mathcal{F}_{i+k_n-1}^n$-measurable and satisfies $\|\mathbb{E}(\nu_i^n|\mathcal{F}_{i-1}^n)\| \le K\Delta_n^{1/2}$. The law of iterated expectations and existing bounds imply

$$\begin{aligned}|\mathbb{E}(\nu_i^{n,lm}\nu_{i+k_n}^{n,jk}|\mathcal{F}_{i-1}^n)| &\le K\Delta_n^{3/4},\\
|\mathbb{E}(\nu_i^{n,lm}\nu_i^{n,gh}\nu_{i+k_n}^{n,jk}|\mathcal{F}_{i-1}^n)| &\le K\Delta_n,\\
|\mathbb{E}(\nu_i^{n,lm}(C_{i+k_n}^{n,gh} - C_i^{n,gh})\nu_{i+k_n}^{n,jk}|\mathcal{F}_{i-1}^n)| &\le K\Delta_n,\\
|\mathbb{E}(\nu_{i+k_n}^{n,lm}(C_{i+k_n}^{n,jk} - C_i^{n,jk})|\mathcal{F}_{i-1}^n)| &\le K\Delta_n^{3/4},\\
|\mathbb{E}((C_{i+k_n}^{n,jk} - C_i^{n,jk})(C_{i+k_n}^{n,lm} - C_i^{n,lm})(C_{i+k_n}^{n,gh} - C_i^{n,gh})|\mathcal{F}_{i-1}^n)| &\le K\Delta_n. \qquad \text{(E.75)}\end{aligned}$$

It can also be readily verified that

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

Hence, for $\varphi_i^{n,gh} \in \{\nu_i^{n,gh}, C_{i+k_n}^{n,gh} - C_i^{n,gh}\}$, which satisfies $\mathbb{E}(|\varphi_i^{n,gh}|^q\Big|\mathcal{F}_{i-1}^n) \le K\Delta_n^{q/4}$ and $\mathbb{E}(\varphi_i^{n,gh}|\mathcal{F}_{i-1}^n) \le K\Delta_n^{1/2}$. One can show that

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

$\leq K\Delta_n^{3/4}(\Delta_n^{1/4} + \eta_{i,2k_n}^n)$.

Next, by combining the successive conditioning together with existing bounds, we have

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

which together imply

$$|\mathbb{E}(\varphi_i^{n,gh}\nu_{i+k_n}^{n,jk}\nu_{i+k_n}^{n,lm}|\mathcal{F}_{i-1}^n)| \leq K\Delta_n^{3/4}(\Delta_n^{1/4} + \eta_{i,2k_n}^n). \tag{E.76}$$

It is easy to see that equations (B.16), (E.75) and (E.76) and the inequality $\eta_{i,k_n}^n \leq \eta_{i,2k_n}^n$ together yield equations (B.19) and (B.20).

### E.9 Proof of Lemma B6

Equation (B.21) can be proved easily using the bounds of $\rho(u,v)_i^{n,gh}$ in equation (E.60). To show equations (B.22), (B.23) and (B.24), we set

$$\overline{\overline{A11}}(H,gh,u;G,ab,v) = \lambda(u,v)_0^n \sum_{i\in L'(n,T)} (\partial_{gh}H\partial_{ab}G)(C_{i-1})\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}.$$

Then,

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

The above result is proved following similar steps as for equation (E.54) in case $w=1$ by replacing $\Theta(u,v)_0^{(C),i,n}$ by $\lambda(u,v)_0^n((\partial_{gh}H\partial_{ab}G)(C_{i-1}) - (\partial_{gh}H\partial_{ab}G)(C_{i-2k_n}))$, which has the same bounds as the former. Next, decompose $\overline{\overline{A11}}$ as follows,

$$\begin{aligned} \overline{\overline{A11}}(H,gh,u;G,ab,v) = \lambda(u,v)_0^n\Bigg[ &\sum_{i\in L'(n,T)} (\partial_{gh}H\partial_{ab}G)(C_{i-1})V_{i-1}^n \\ &+ \sum_{i\in L'(n,T)} (\partial_{gh}H\partial_{ab}G)(C_{i-1})\Big(\mathbb{E}(\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}|\mathcal{F}_{i-1}^n) - V_{i-1}^n\Big) \\ &+ \sum_{i\in L'(n,T)} (\partial_{gh}H\partial_{ab}G)(C_{i-1})\Big(\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab} - \mathbb{E}(\zeta(u)_i^{n,gh}\zeta(v)_i^{n,ab}|\mathcal{F}_{i-1}^n)\Big)\Bigg]. \end{aligned}$$

We follow the proof of equation (E.55) for $w=1$, and we replace $\Theta(u,v)_0^{(C),i,n}$ by $\lambda(u,v)_0^n(\partial_{gh}H\partial_{ab}G)(C_{i-1})$, which satisfies only the condition $|\lambda(u,v)_0^n(\partial_{gh}H\partial_{ab}G)(C_{i-1})| \leq \widetilde{\lambda}_{u,v}^n$. This calculation shows that the last two terms in the above decomposition vanish at a rate faster than $\Delta_n^{1/4}$. Therefore,

$$\frac{1}{\Delta_n^{1/4}}\left(\overline{\overline{A11}}(H,gh,u;G,ab,v) - \lambda(u,v)_0^n\Big(\sum_{i\in L'(n,T)} (\partial_{gh}H\partial_{ab}G)(C_{i-1})V_{i-1}^n\Big)\right) \Rightarrow 0.$$

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

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

The results follow from the following observation,

$$\frac{1}{\Delta_n^{1/4}} \Bigg( \lambda(u,v)_0^n \Big( \sum_{g,h,a,b=1}^{d} \sum_{i \in L'(n,T)} (\partial_{gh} H \partial_{ab} G)(C_{i-1}) V_{i-1}^n(u,v) \Big) - \frac{3}{\theta^2} \int_0^T (\partial_{gh} H \partial_{ab} G)(C_t)(C_t^{ga} C_t^{hb} + C_t^{gb} C_t^{ha}) dt \Bigg) \Rightarrow 0, \quad \text{for } (u,v) = (2,2),$$

$$\frac{1}{\Delta_n^{1/4}} \Bigg( \sum_{g,h,a,b=1}^{d} \lambda(u,v)_0^n \Big( \sum_{i \in L'(n,T)} (\partial_{gh} H \partial_{ab} G)(C_{i-1}) V_{i-1}^n(u,v) \Big) - [H(C), G(C)]_T \Bigg) \Rightarrow 0,$$

for $(u, v) = (1, 1)$.

## F Numerical Implementation

We now discuss some details for the numerical implementation of our estimators. Section 3.2 explains how the main quantities of interest can be expressed in terms of $[H(C), G(C)]_T$, where $C$ is the spot variance matrix of all $d$ assets. However, in practice many quantities of interest involve only a much smaller subset of assets, which greatly reduces the computational burden.

For example, suppose we want to calculate the variance of the IdioVol for a single stock, where R-FM is the CAPM, and IdioVol-FM has one volatility factor – the market volatility. Then, we only need to consider two assets, the stock and the market, e.g., SPY, so $d_S = d_F = 1$ and $d = 2$. Denote the relevant spot variance-covariance matrix by

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix},$$

where $C_{22} = C_F$ is the spot variance of the market, and $C_{11}$ is the spot variance of the individual stock. The quantity of interest is

$$[H(C), H(C)]_T = [C_{Z1}, C_{Z1}]_T,$$

where $C_{Z1,t} = C_{11} - C_{12} C_{22}^{-1} C_{21}$. The estimators in equations (16) and (18) involve the first derivatives $\partial_{ab} H(C)$ for $a, b = 1, ..., d$, which are

$$\partial_{ab} H(C) \equiv \frac{\partial H(C)}{\partial C_{ab}} = \frac{\partial C_{Zj}}{\partial C_{ab}} = \frac{\partial \left( C_{11} - C_{12} C_{22}^{-1} C_{21} \right)}{\partial C_{ab}} = \begin{cases} C_{12} C_{22}^{-2} C_{21} & \text{if } (a,b) = (2,2) \\ -C_{22}^{-1} C_{21} & \text{if } (a,b) = (1,2) \\ -C_{12} C_{22}^{-1} & \text{if } (a,b) = (2,1) \\ 1 & \text{if } (a,b) = (1,1) \end{cases}$$

If we are interested in the stock's IdioVol $\gamma_Z$, by equation (23) we also need the volatility factor $\Pi_t = G(C_t) = C_{22,t}$, and $[\Pi, C_{Z1}]_T^c$. The derivatives are $\partial_{ab} G(C) \equiv \partial G(C) / \partial C_{ab} = 1\{a = b = 2\}$.

# G Additional Figures

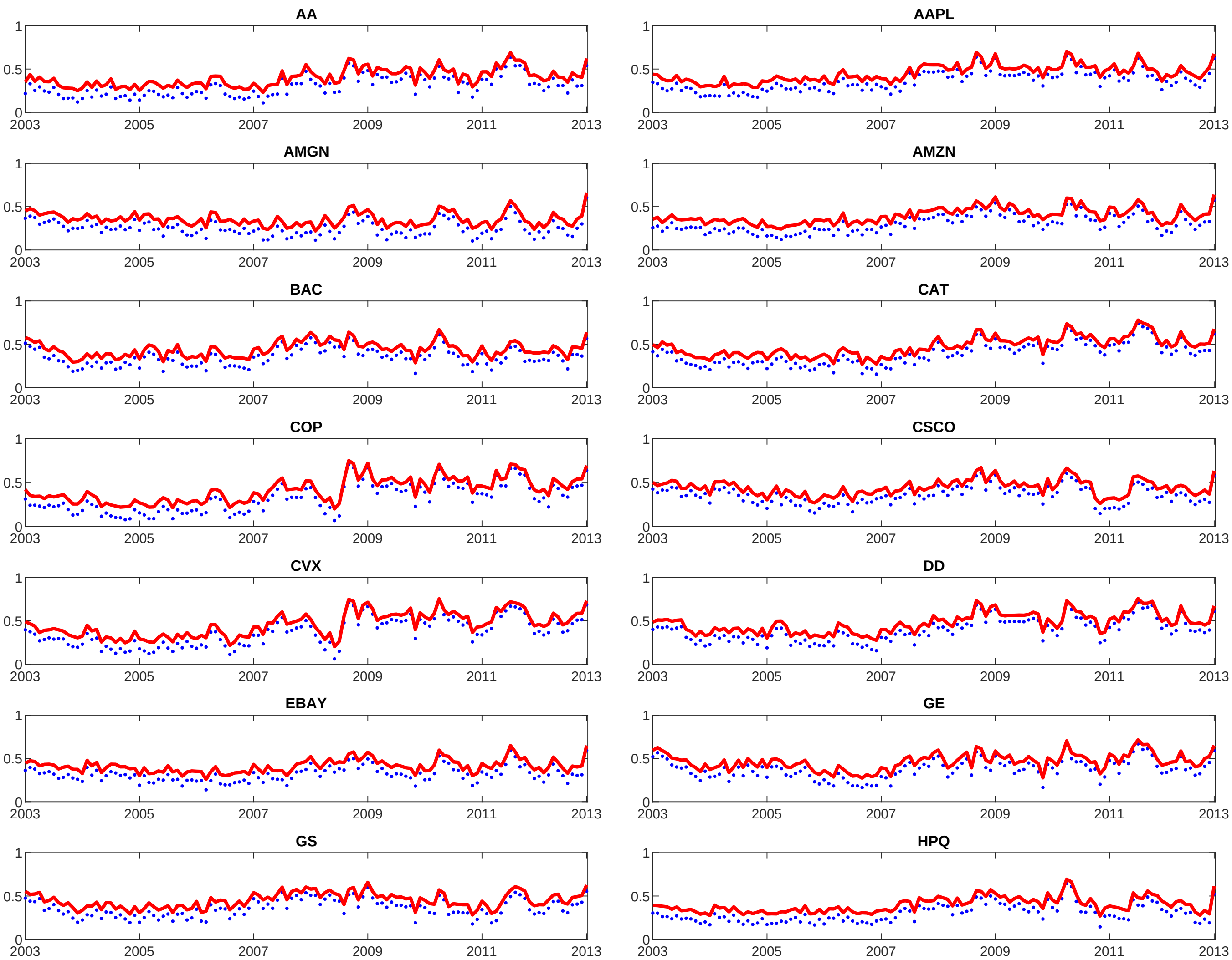


*Figure G.1:* Monthly $R^2$ of two Return Factor Models ($\widehat{R}^2_{Yj}$): the CAPM (the blue dotted line) and the Fama-French three factor model (the red solid line). Stocks are represented by tickers (see Table 1 for full stock names).

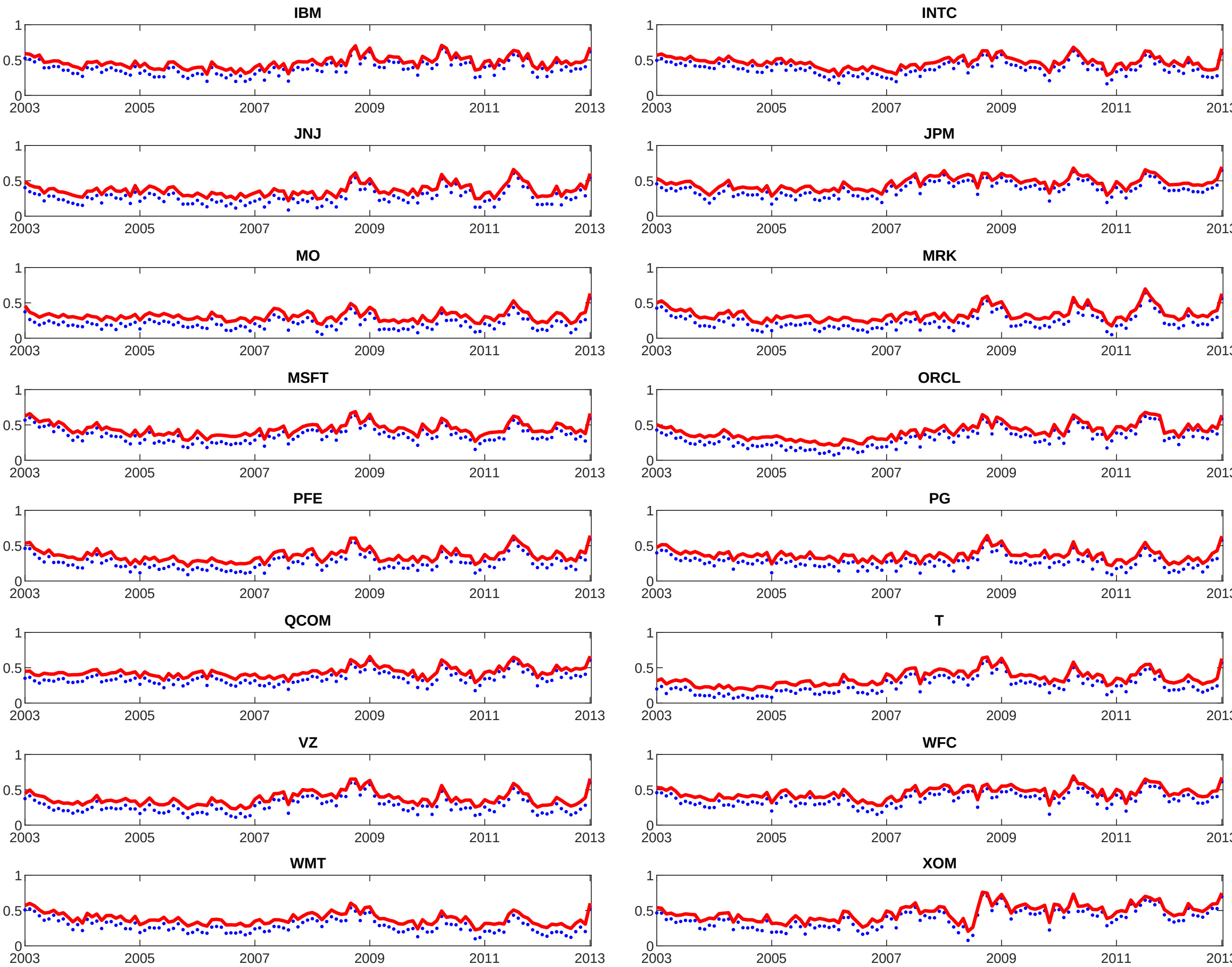


*Figure G.2:* Monthly $R^2$ of two Return Factor Models ($\widehat{R}^2_{Y j}$): the CAPM (the blue dotted line) and the Fama-French three factor model (the red solid line). Stocks are represented by tickers (see Table 1 for full stock names).